Amath 546/Econ 589
Multivariate GARCH Models

Eric Zivot

May 8, 2012

Lecture Outline

- Exponentially weighted covariance estimation
- Multivariate GARCH models
- Prediction from multivariate GARCH models

Reading

- FRF chapter 3.
- QRM chapter 4, sections 5 and 6;
- FMUND, chapter 6

Exponentially Weighted Covariance Estimate

let $\mathbf{y}_t$ be a $k \times 1$ vector of multivariate time series:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{c} + \boldsymbol{\epsilon}_t,\ t = 1, 2, \cdots, T \\ & \boldsymbol{\epsilon}_t \sim WN(0, \boldsymbol{\Sigma}) \end{aligned}$$

The sample covariance matrix is given by:

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T-1} \sum_{t=1}^{T} \hat{\boldsymbol{\epsilon}}_t \hat{\boldsymbol{\epsilon}}_t' = \frac{1}{T-1} \sum_{t=1}^{T} (\mathbf{y}_t - \bar{\mathbf{y}})(\mathbf{y}_t - \bar{\mathbf{y}})',\ \bar{\mathbf{y}} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{y}_t$$

To allow for time varying covariance matrix, an *ad hoc* approach uses exponentially decreasing weights as follows:

$$\begin{aligned} \hat{\boldsymbol{\Sigma}}_t &= \lambda \hat{\boldsymbol{\epsilon}}_{t-1} \hat{\boldsymbol{\epsilon}}_{t-1}' + \lambda^2 \hat{\boldsymbol{\epsilon}}_{t-2} \hat{\boldsymbol{\epsilon}}_{t-2}' + \cdots \\ &= \sum_{i=1}^{\infty} \lambda^i \hat{\boldsymbol{\epsilon}}_{t-i} \hat{\boldsymbol{\epsilon}}_{t-i}',\ 0 < \lambda < 1 \end{aligned}$$

Since

$$\lambda + \lambda^2 + \cdots = \frac{\lambda}{1-\lambda}$$

the weights are usually scaled so that they sum up to one:

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

The above equation can be easily rewritten to obtain the following recursive form for exponentially weighted covariance matrix:

$$\hat{\mathbf{\Sigma}}_t = (1-\lambda)\hat{\boldsymbol{\epsilon}}_{t-1}\hat{\boldsymbol{\epsilon}}'_{t-1} + \lambda \hat{\mathbf{\Sigma}}_{t-1}$$

From the above equation, given λ and an initial estimate $\hat{\mathbf{\Sigma}}_1$, the time varying exponentially weighted covariance matrices can be computed easily.

Example: Bivariate Model

$$\begin{pmatrix} \hat{\sigma}_t^{11} & \hat{\sigma}_t^{12} \\ \hat{\sigma}_t^{12} & \hat{\sigma}_t^{22} \end{pmatrix} = (1-\lambda) \begin{bmatrix} \hat{\epsilon}_{t-1}^1 \hat{\epsilon}_{t-1}^1 & \hat{\epsilon}_{t-1}^1 \hat{\epsilon}_{t-1}^2 \\ \hat{\epsilon}_{t-1}^1 \hat{\epsilon}_{t-1}^2 & \hat{\epsilon}_{t-1}^2 \hat{\epsilon}_{t-1}^2 \end{bmatrix} + \lambda \begin{bmatrix} \hat{\sigma}_{t-1}^{11} & \hat{\sigma}_{t-1}^{12} \\ \hat{\sigma}_{t-1}^{12} & \hat{\sigma}_{t-1}^{22} \end{bmatrix}$$

Note

$$\begin{aligned} \hat{\sigma}_t^{11} &= (1-\lambda)\left(\hat{\epsilon}_{t-1}^1\right)^2 + \lambda \hat{\sigma}_{t-1}^{11} \\ \hat{\sigma}_t^{12} &= (1-\lambda)\hat{\epsilon}_{t-1}^1 \hat{\epsilon}_{t-1}^2 + \lambda \hat{\sigma}_{t-1}^{12} \end{aligned}$$

Estimating λ

- In practice, the value of λ is usually chosen in an *ad hoc* way as typified by the RiskMetrics proposal.

- If $\boldsymbol{\epsilon}_t \sim N(0, \boldsymbol{\Sigma}_t)$ where $\boldsymbol{\Sigma}_t = \text{Cov}_{t-1}(\boldsymbol{\epsilon}_t)$ then the log-likelihood function of the observed time series can be written as:

$$\log L = -\frac{kT}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}|\boldsymbol{\Sigma}_t| - \frac{1}{2}\sum_{t=1}^{T}(\mathbf{y}_t - \mathbf{c})'\boldsymbol{\Sigma}_t^{-1}(\mathbf{y}_t - \mathbf{c}).$$

- The mean vector $\mathbf{c}$ and λ can be treated as unknown model parameters and estimated using quasi-maximum likelihood estimation (MLE), given the initial value $\boldsymbol{\Sigma}_1$.

Multivariate GARCH Models

The general multivariate GARCH model has the form

$$\begin{aligned}
\underset{k\times 1}{\mathbf{y}_t} &= \mathbf{c} + \sum_{i=1}^{r} \underset{k\times k}{\boldsymbol{\Phi}_i}\,\mathbf{y}_{t-i} + \sum_{l=1}^{L} \underset{k\times m}{\boldsymbol{\beta}_l}\,\underset{m\times 1}{\mathbf{x}_{t-l}} + \sum_{s=1}^{S} \underset{k\times k}{\boldsymbol{\Theta}_s}\,\boldsymbol{\epsilon}_{t-s} \\
\underset{k\times 1}{\boldsymbol{\epsilon}_t} &= \underset{k\times k}{\boldsymbol{\Sigma}_t^{1/2}}\mathbf{z}_t,\ \mathbf{z}_t \sim iid\ (0, \mathbf{I}_k) \\
\boldsymbol{\Sigma}_t &= \boldsymbol{\Sigma}_t^{1/2}\boldsymbol{\Sigma}_t^{1/2'},\ \boldsymbol{\Sigma}_t^{1/2} = \text{ Cholesky factor}
\end{aligned}$$

Here

$$\text{var}_{t-1}(\mathbf{y}_t) = \boldsymbol{\Sigma}_t^{1/2}\text{var}_{t-1}(\mathbf{z}_t)\boldsymbol{\Sigma}_t^{1/2'} = \boldsymbol{\Sigma}_t$$

Practical Issues in Modeling $\boldsymbol{\Sigma}_t$

- Models should be easy to understand and estimate, and allow for flexible dynamics in conditional variances and correlations
- Models should have a limited number of parameters
- $\boldsymbol{\Sigma}_t$ should be positive semi-definite
  - diagonal elements of $\boldsymbol{\Sigma}_t^{1/2}$ should be greater than or equal to zero

Diagonal VEC Models

Bollerslev, Engle and Wooldridge (1988) extended the univariate GARCH specification to a multivariate setting with the diagonal VEC model

$$\begin{aligned} \underset{k\times k}{\boldsymbol{\Sigma}_t} &= \underset{k\times k}{\mathbf{A}_0} + \sum_{i=1}^{p} \underset{k\times k}{\mathbf{A}_i} \odot (\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}_{t-i}') + \sum_{j=1}^{q} \underset{k\times k}{\mathbf{B}_j} \odot \boldsymbol{\Sigma}_{t-j} \\ \odot &= \text{Hadamard product (element by element)} \end{aligned}$$

where $\mathbf{A}_0$, $\mathbf{A}_i$ and $\mathbf{B}_j$ are symmetric matrices. For example, for $k = 2$ and $p = q = 1$

$$\begin{aligned} \begin{pmatrix} \sigma_t^{11} & \sigma_t^{12} \\ \sigma_t^{12} & \sigma_t^{22} \end{pmatrix} &= \begin{bmatrix} a_0^{11} & a_0^{21} \\ a_0^{21} & a_0^{22} \end{bmatrix} + \begin{bmatrix} a_1^{11} & a_1^{21} \\ a_1^{21} & a_1^{22} \end{bmatrix} \odot \begin{bmatrix} \epsilon_{t-1}^1\epsilon_{t-1}^1 & \epsilon_{t-1}^1\epsilon_{t-1}^2 \\ \epsilon_{t-1}^1\epsilon_{t-1}^2 & \epsilon_{t-1}^2\epsilon_{t-1}^2 \end{bmatrix} \\ &\quad + \begin{bmatrix} b_1^{11} & b_1^{21} \\ b_1^{21} & b_1^{22} \end{bmatrix} \odot \begin{bmatrix} \sigma_{t-1}^{11} & \sigma_{t-1}^{12} \\ \sigma_{t-1}^{12} & \sigma_{t-1}^{22} \end{bmatrix} \end{aligned}$$

Element-by-element we have GARCH(1,1)-type models for variances and covariances:

$$\begin{aligned}
\sigma_t^{11} &= a_0^{11} + a_1^{11}\left(\epsilon_{t-1}^1\right)^2 + b_1^{11}\sigma_{t-1}^{11} \\
\sigma_t^{12} &= a_0^{21} + a_1^{21}\epsilon_{t-1}^1\epsilon_{t-1}^2 + b_1^{21}\sigma_{t-1}^{12} \\
\sigma_t^{22} &= a_0^{22} + a_1^{22}\left(\epsilon_{t-1}^2\right)^2 + b_1^{22}\sigma_{t-1}^{22}
\end{aligned}$$

Remark: There are no cross-volatility or cross-covariance feedback effects. For example, σ_t^{11} does not depend on $\sigma_t^{12}, \epsilon_{t-2}^2$, or σ_t^{22}.

To isolate the unique elements of $\mathbf{\Sigma}_t$ the lower triangular elements are extracted using the vech$(\cdot)$ operator

$$\mathbf{\Sigma}_t = \begin{pmatrix} \sigma_t^{11} & \sigma_t^{12} \\ \sigma_t^{12} & \sigma_t^{22} \end{pmatrix}, \ \text{vech}(\mathbf{\Sigma}_t) = \begin{pmatrix} \sigma_t^{11} \\ \sigma_t^{12} \\ \sigma_t^{22} \end{pmatrix}$$

Example: Unique components of bivariate DVEC(1,1) Model

$$\begin{aligned}
\begin{bmatrix} \sigma_t^{11} \\ \sigma_t^{12} \\ \sigma_t^{22} \end{bmatrix} &= \begin{bmatrix} a_0^{11} \\ a_0^{12} \\ a_0^{22} \end{bmatrix} + \begin{bmatrix} a_1^{11} & & \\ & a_1^{21} & \\ & & a_1^{22} \end{bmatrix} \begin{bmatrix} \epsilon_{t-1}^1\epsilon_{t-1}^1 \\ \epsilon_{t-1}^1\epsilon_{t-1}^2 \\ \epsilon_{t-1}^2\epsilon_{t-1}^2 \end{bmatrix} \\
&\quad + \begin{bmatrix} b_1^{11} & & \\ & b_1^{21} & \\ & & b_1^{22} \end{bmatrix} \begin{bmatrix} \sigma_{t-1}^{11} \\ \sigma_{t-1}^{12} \\ \sigma_{t-1}^{22} \end{bmatrix}
\end{aligned}$$

Hence, only need to specify the lower triangular elements of $\mathbf{A}_0$, $\mathbf{A}_1$ and $\mathbf{B}_1$.

In the bivariate DVEC(1,1),

$$\begin{bmatrix} a_0^{11} \\ a_0^{12} \\ a_0^{22} \end{bmatrix} = \text{vech}(\mathbf{A}_0)$$

$$\begin{bmatrix} a_1^{11} & & \\ & a_1^{21} & \\ & & a_1^{22} \end{bmatrix} = \text{diag}(\text{vech}(\mathbf{A}_1))$$

$$\begin{bmatrix} b_1^{11} & & \\ & b_1^{21} & \\ & & b_1^{22} \end{bmatrix} = \text{diag}(\text{vech}(\mathbf{B}_1))$$

Hence, the DVEC(1,1) can be expresses as

$$\begin{aligned} \text{vech}(\boldsymbol{\Sigma}_t) &= \text{vech}(\mathbf{A}_0) + \text{diag}(\text{vech}(\mathbf{A}_1))\text{vech}(\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}') \\ &\quad + \text{diag}(\text{vech}(\mathbf{B}_1))\text{vech}(\boldsymbol{\Sigma}_{t-1}) \end{aligned}$$

or, more simply, as

$$\mathbf{h}_t = \mathbf{a}_0 + \text{diag}(\mathbf{a}_1)\mathbf{v}_{t-1} + \text{diag}(\mathbf{b}_1)\mathbf{h}_{t-1}$$

where

$$\begin{aligned} \mathbf{h}_t &= \text{vech}(\boldsymbol{\Sigma}_t),\ \mathbf{v}_t = \text{vech}(\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}') \\ \mathbf{a}_0 &= \text{vech}(\mathbf{A}_0),\ \mathbf{a}_1 = \text{vech}(\mathbf{A}_1),\ \mathbf{b}_1 = \text{vech}(\mathbf{B}_1) \end{aligned}$$

General Diagonal VEC Model

Let $\mathbf{\Sigma}_t$ be $k \times k$ and define the $k(k+1)/2 \times 1$ vectors $\mathbf{h}_t = \text{vech}(\mathbf{\Sigma}_t)$, $\mathbf{v}_t = \text{vech}(\varepsilon_t \varepsilon_t')$, $\mathbf{a}_j = \text{vech}(\mathbf{A}_j)$ and $\mathbf{b}_j = \text{vech}(\mathbf{B}_j)$. Then the DVEC(p,q) model has the form

$$\mathbf{h}_t = \mathbf{a}_0 + \sum_{j=1}^{p} \text{diag}(\mathbf{a}_j)\mathbf{v}_{t-j} + \sum_{j=1}^{q} \text{diag}(\mathbf{b}_j)\mathbf{h}_{t-j}$$

where $\text{diag}(\mathbf{a}_j)$ and $\text{diag}(\mathbf{b}_j)$ denote diagonal matrices with the elements of $\mathbf{a}_j$ and $\mathbf{b}_j$ along the diagonals, respectively.

Problems

- Large number of parameters: $(p+q+1)k(k+1)/2$.
  - $p = q = 1, k = 2 \Rightarrow$ 9 parameters
  - $p = q = 1, k = 10 \Rightarrow$ 135 parameters
- $\mathbf{\Sigma}_t$ is not guaranteed to be psd
  - could have negative variances or absolute correlations bigger than one

Unconditional Covariance in DVEC model

Consider the DVEC(1,1)

$$\mathbf{h}_t = \mathbf{a}_0 + \text{diag}(\mathbf{a}_1)\mathbf{v}_{t-1} + \text{diag}(\mathbf{b}_1)\mathbf{h}_{t-1}$$

Then

$$\begin{aligned} E[\mathbf{h}_t] &= \mathbf{a}_0 + \text{diag}(\mathbf{a}_1)E[\mathbf{v}_{t-1}] + \text{diag}(\mathbf{b}_1)E[\mathbf{h}_{t-1}] \\ &= \mathbf{a}_0 + \text{diag}(\mathbf{a}_1)E[\mathbf{h}_t] + \text{diag}(\mathbf{b}_1)E[\mathbf{h}_t] \text{ by stationarity} \\ \Rightarrow\ & E[\mathbf{h}_t] = (\mathbf{I}_p - \text{diag}(\mathbf{a}_1) - \text{diag}(\mathbf{b}_1))^{-1}\mathbf{a}_0 \end{aligned}$$

where $p = k(k+1)/2$. Here, stationarity requires the eigenvalues of $\text{diag}(\mathbf{a}_1) + \text{diag}(\mathbf{b}_1)$ to have modulus less than unity.

Covariance Targeting

Using the result $E[\mathbf{h}_t] = (\mathbf{I}_p - \text{diag}(\mathbf{a}_1) - \text{diag}(\mathbf{b}_1))^{-1}\mathbf{a}_0$, the vector $\mathbf{a}_0$ can be expressed as

$$\mathbf{a}_0 = (\mathbf{I}_p - \text{diag}(\mathbf{a}_1) - \text{diag}(\mathbf{b}_1))E[\mathbf{h}_t]$$

The parameters in $\mathbf{a}_0$ can be eliminated by specifying a vector for $E[\mathbf{h}_t]$. For example, set $E[\mathbf{h}_t] = \text{vech}(\hat{\mathbf{\Sigma}})$,where $\hat{\mathbf{\Sigma}}$ is the sample covariance matrix. This is called *covariance targeting*.

Example: DVEC(1,1) with covariance targeting

$$\mathbf{h}_t = (\mathbf{I}_p - \text{diag}(\mathbf{a}_1) - \text{diag}(\mathbf{B}_1))\text{vech}(\hat{\mathbf{\Sigma}}) + \text{diag}(\mathbf{a}_1)\mathbf{v}'_{t-1} + \text{diag}(\mathbf{b}_1)\mathbf{h}_{t-1}$$

This eliminates $p = k(k+1)/2$ parameters from the model.

Simplification of DVEC model (Scalar DVEC)

- Restrict $\mathbf{A}_i$ to have common element a_i and $\mathbf{B}_j$ to have common element b_j.

- Total parameters $= k(k+1) + (p+q)$.
  - $k = 2, p = q = 1 \Rightarrow 5$ parameters;
  - $k = 10, p = q = 1 \Rightarrow 112$ parameters

Example: Bivariate DVEC(1,1) Model

$$\begin{bmatrix} \sigma_t^{11} \\ \sigma_t^{12} \\ \sigma_t^{22} \end{bmatrix} = \begin{bmatrix} a_0^{11} \\ a_0^{12} \\ a_0^{22} \end{bmatrix} + \begin{bmatrix} a_1 & & \\ & a_1 & \\ & & a_1 \end{bmatrix} \begin{bmatrix} \epsilon_{t-1}^1 \epsilon_{t-1}^1 \\ \epsilon_{t-1}^1 \epsilon_{t-1}^2 \\ \epsilon_{t-1}^2 \epsilon_{t-1}^2 \end{bmatrix} + \begin{bmatrix} b_1 & & \\ & b_1 & \\ & & b_1 \end{bmatrix} \begin{bmatrix} \sigma_{t-1}^{11} \\ \sigma_{t-1}^{12} \\ \sigma_{t-1}^{22} \end{bmatrix}$$

The scalar bivariate DVEC(1,1) can then be re-expressed as

$$\mathbf{h}_t = \mathbf{a}_0 + a_1 \mathbf{I}_3 \mathbf{v}_{t-1} + b_1 \mathbf{I}_3 \mathbf{h}_{t-1}$$

Notice that

$$E[\mathbf{h}_t] = (1 - a_1 - b_1)^{-1} \mathbf{a}_0$$

and that stationarity requires $a_1 + b_1 < 1$

BEKK Models

- DEV and model restricts conditional variances and covariances to only depend on their own lagged values and the corresponding cross-product of the error terms.

- The BEKK model (Baba, Engle, Kraft and Kroner) gives a richer dynamics and is given by

  $$\mathbf{\Sigma}_t = \mathbf{A}_0 \mathbf{A}_0' + \sum_{i=1}^{p} \mathbf{A}_i (\boldsymbol{\epsilon}_{t-i} \boldsymbol{\epsilon}_{t-i}') \mathbf{A}_i' + \sum_{j=1}^{q} \mathbf{B}_j \mathbf{\Sigma}_{t-j} \mathbf{B}_j'$$

  where $\mathbf{A}_0$ is a lower triangular matrix, but $\mathbf{A}_i$ $(i = 1, \cdots, p)$ and $\mathbf{B}_j$ $(j = 1, \cdots, q)$ are unrestricted square matrices.

- $k(k-1)(p+q)/2$ more parameters than DVEC(p,q)

Example: bivariate BEKK(1,1)

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

Consider the $(2,2)$ element of $\mathbf{\Sigma}_t$ in the BEKK$(1,1)$ model:

$$\sigma_t^{22} = a_0^{22}a_0^{22} + [a_1^{21}\epsilon_{t-1}^1 + a_1^{22}\epsilon_{t-1}^2]^2 + [b_1^{21}b_1^{21}\sigma_{t-1}^{11} + 2b_1^{21}b_1^{22}\sigma_{t-1}^{21} + b_1^{22}b_1^{22}\sigma_{t-1}^{22}]$$

Notice that both ϵ_{t-1}^1 and ϵ_{t-1}^2 enter the equation. In addition, σ_{t-1}^{11}, the volatility of the first series, also has direct impacts on σ_t^{22}, the volatility of the second series. However, for the bivariate BEKK$(1,1)$ model, flexibility is achieved at the cost of two extra parameters, i.e., a_1^{12} and b_1^{12}, which are not needed for the DVEC$(1,1)$ model.

Multivariate GARCH Prediction

- Predictions from multivariate GARCH models can be generated in a similar fashion to predictions from univariate GARCH models.

- For multivariate GARCH models, predictions can be generated for both the levels of the original multivariate time series and its conditional covariance matrix. Predictions of the levels are obtained just as for vector autoregressive (VAR) models. Compared with VAR models, the predictions of the conditional covariance matrix from multivariate GARCH models can be used to construct more reliable confidence intervals for predictions of the levels.

Forecasting from DVEC(1,1)

Consider the conditional variance equation for the DVEC$(1,1)$ model:

$$\mathbf{h}_t = \mathbf{a}_0 + \text{diag}(\mathbf{a}_1)\mathbf{v}_{t-1} + \text{diag}(\mathbf{b}_1)\mathbf{h}_{t-1}$$

which is estimated over the time period $t = 1, 2, \cdots, T$.

- To obtain $E_T(\mathbf{h}_{T+k})$, use the forecasts of conditional covariance matrix at time $T+k$ for $k > 0$, given information at time T.

- For one-step-ahead prediction:

$$\begin{aligned} E_T(\mathbf{h}_{T+1}) &= \mathbf{a}_0 + \text{diag}(\mathbf{a}_1)E_T[\mathbf{v}_T] + \text{diag}(\mathbf{b}_1)E_T[\mathbf{h}_T] \\ &= \mathbf{a}_0 + \text{diag}(\mathbf{a}_1)\mathbf{v}_T + \text{diag}(\mathbf{b}_1)\mathbf{h}_T \end{aligned}$$

  since an estimate of $\boldsymbol{v}_T$ and $\mathbf{h}_T$ already exists after estimating the DVEC model.

- When $k = 2$,

$$\begin{aligned} E_T(\mathbf{h}_{T+2}) &= \mathbf{a}_0 + \text{diag}(\mathbf{a}_1)E_T[\mathbf{v}_{T+1}] + \text{diag}(\mathbf{b}_1)E_T[\mathbf{h}_{T+1}] \\ &= \mathbf{a}_0 + (\text{diag}(\mathbf{a}_1) + \text{diag}(\mathbf{b}_1))E_T[\mathbf{h}_{T+1}]. \end{aligned}$$

  where $E_T(\mathbf{h}_{T+1})$ is obtained in the previous step.

- This procedure can be iterated to obtain $E_T(\mathbf{h}_{T+k})$ for $k > 2$.

- Forecasts converge to the unconditional covariance matrix defined by

$$\text{vech}(\bar{\boldsymbol{\Sigma}}) = (\mathbf{I}_k - \text{diag}(\mathbf{a}_1) - \text{diag}(\mathbf{b}_1)^{-1}\mathbf{a}_0$$

Univariate GARCH-Based Models

- For BEKK, DVEC and matrix diagonal models, the conditional covariance matrix is modeled directly.
  - This approach can result in a large number of parameters since the covariance terms need to be modeled separately.
- Another approach in multivariate GARCH modeling is to first model individual series using univariate GARCH and then model the conditional correlations between the series. The main types of models are
  - Constant conditional correlation (CCC) model, Dynamic conditional correlation (DCC) model, and orthogonal principal component (OGARCH) model.

Constant Conditional Correlation (CCC) Model

- Given that $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$ a $k \times k$ covariance matrix $\mathbf{\Sigma}$ can be decomposed into:

$$\mathbf{\Sigma} = \mathbf{DRD}$$

where $\mathbf{R}$ is the correlation matrix, $\mathbf{D}$ is a diagonal matrix with the vector $(\sigma_1, \cdots, \sigma_k)'$ on the diagonal, and σ_i is the standard deviation of the i-th series.

$$\mathbf{R} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{12} & 1 & \cdots & \rho_{2k} \\ \vdots & & \ddots & \vdots \\ \rho_{1k} & \rho_{2k} & \cdots & 1 \end{bmatrix}, \ \mathbf{D} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_k \end{bmatrix}$$

- Based on the observation that the correlation matrix of foreign exchange rate returns is usually constant over time, Bollerslev (1990) suggested modelling the time varying covariance matrix as:

$$\mathbf{\Sigma}_t = \mathbf{D}_t \mathbf{R} \mathbf{D}_t$$

where $\mathbf{R}$ is the constant conditional correlation matrix, and $\mathbf{D}_t$ is a diagonal matrix:

$$\mathbf{D}_t = \begin{bmatrix} \sigma_{1t} & & \\ & \ddots & \\ & & \sigma_{kt} \end{bmatrix}$$

with σ_{it} following any univariate GARCH process, for $i = 1, \cdots, k$. Here, the conditional correlations ρ_{ij} are constant but the conditional covariances $\sigma_{ij,t} = \rho_{ij}\sigma_{it}\sigma_{jt}$ are time varying.

- $\mathbf{R}$ can be the sample correlation matrix or a matrix of specified values

Dynamic Conditional Correlation (DCC) Model

Engle (2002) extended Bollerslev's CCC model to allow the conditional correlations to be time varying.

$$\underset{k\times 1}{\mathbf{r}_t} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t, \ \boldsymbol{\varepsilon}_t | I_{t-1} \sim iid \ N(\mathbf{0}, \mathbf{\Sigma}_t)$$

$$var(\boldsymbol{\varepsilon}_t | I_{t-1}) = \underset{k\times k}{\mathbf{\Sigma}_t} = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t$$

$$\mathbf{R}_t = \begin{bmatrix} 1 & \rho_{12,t} & \cdots & \rho_{1k,t} \\ \rho_{12,t} & 1 & \cdots & \rho_{2k,t} \\ \vdots & & \ddots & \vdots \\ \rho_{1k,t} & \rho_{2k,t} & \cdots & 1 \end{bmatrix}, \ \mathbf{D}_t = \begin{bmatrix} \sigma_{1t} & 0 & \cdots & 0 \\ 0 & \sigma_{2t} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{kt} \end{bmatrix}$$

For each univariate series ($i = 1, \ldots, k$), we have

$$\begin{aligned} \varepsilon_{it} &= \sigma_{it} z_{it},\ z_{it} \sim N(0,1) \\ var(\varepsilon_{it} | I_{t-1}) &= \sigma_{it}^2 \end{aligned}$$

Define the vector of standardized errors (returns) $\mathbf{z}_t = (z_{1t}, \ldots, z_{kt})'$. Then

$$E[\mathbf{z}_t \mathbf{z}_t' | I_{t-1}] = \mathbf{R}_t \neq \mathbf{I}_k$$

Why? Consider

$$\begin{aligned} cor(\varepsilon_{it}, \varepsilon_{jt} | I_{t-1}) &= \frac{cov(\varepsilon_{it}, \varepsilon_{jt} | I_{t-1})}{\sigma_{it}\sigma_{jt}} = \frac{cov(\sigma_{it} z_{it}, \sigma_{jt} z_{it} | I_{t-1})}{\sigma_{it}\sigma_{jt}} \\ &= \frac{\sigma_{it}\sigma_{jt} cov(z_{it}, z_{it} | I_{t-1})}{\sigma_{it}\sigma_{jt}} = cov(z_{it}, z_{it} | I_{t-1}) \\ &= E[z_{it} z_{jt}] \end{aligned}$$

Idea behind DCC:

- Estimate univariate GARCH models (e.g. GARCH(1,1)) for each ε_{it} ($i = 1, \ldots, k$)

  $$\hat{\sigma}_{it}^2 = \hat{a}_0 + \hat{a}_1 \hat{\varepsilon}_{it-1}^2 + \hat{b}_1 \hat{\sigma}_{it-1}^2$$

  and form estimated standardized residuals

  $$\hat{z}_{it} = \frac{\hat{\varepsilon}_{it}}{\hat{\sigma}_{it}}$$

- Model the pairwise conditional covariances between the standardized residuals

  $$\hat{\rho}_{ij,t} = \widehat{cov}(\hat{z}_{it}, \hat{z}_{jt} | I_{t-1})$$

- Estimate the conditional covariance matrix from the univariate volatility estimates and bivariate conditional correlation estimates

$$\hat{\mathbf{\Sigma}}_t = \hat{\mathbf{D}}_t \hat{\mathbf{R}}_t \hat{\mathbf{D}}_t$$

Modeling Conditional Correlations

Engle proposed two ways to model $\hat{\rho}_{ij,t} = \widehat{cov}(\hat{z}_{it}, \hat{z}_{jt} | I_{t-1})$

1. EWMA covariance matrix for $\hat{\mathbf{z}}_t = (\hat{z}_{1t}, \ldots, \hat{z}_{kt})'$

$$\hat{\mathbf{Q}}_t^{EWMA} = (1-\lambda)\hat{\mathbf{z}}_{t-1}\hat{\mathbf{z}}'_{t-1} + \lambda \hat{\mathbf{Q}}_{t-1}^{EWMA}$$

Then re-scale EWMA covariances to get EWMA correlations

$$\hat{\rho}_{ij,t}^{EWMA} = \frac{\hat{q}_{ij,t}^{EWMA}}{\left(\hat{q}_{ii,t}^{EWMA} \times \hat{q}_{jj,t}^{EWMA}\right)^{1/2}}$$

Note: re-scaling is necessary because the elements of $\hat{\mathbf{Q}}_t^{EWMA}$ are not guaranteed to lie between -1 and 1.

Note: We use the same λ to model all conditional covariances. This greatly reduces the number of parameters to estimate!

2. Common GARCH(1,1) model for $\hat{q}_{ij,t} = cov(\hat{z}_{it}, \hat{z}_{jt}|I_{t-1})$

$$\begin{aligned} \hat{q}_{ij,t} &= \omega_{ij} + \alpha \cdot \hat{z}_{i,t-1}\hat{z}_{j,t-1} + \beta \cdot \hat{q}_{ij,t-1}, \text{ for all } i,j \\ &\Rightarrow \underset{k\times k}{\hat{\mathbf{Q}}_t^{DCC}} = \boldsymbol{\Omega} + \alpha \cdot \hat{\mathbf{z}}_{t-1}\hat{\mathbf{z}}_{t-1}' + \beta \cdot \hat{\mathbf{Q}}_{t-1}^{DCC} \end{aligned}$$

- Use the same values of α and β for all $\hat{q}_{ij,t}$! This reduces the number of estimated parameters like in EWMA.

- Use covariance targeting to eliminate ω_{ij} in each equation

$$\begin{aligned} \omega_{ij} &= \hat{E}[\hat{z}_{it}\hat{z}_{jt}](1-\alpha-\beta) \\ \hat{E}[\hat{z}_{it}\hat{z}_{jt}] &= \text{sample covariance b/w } \hat{z}_{it} \text{ and } \hat{z}_{jt} = \hat{s}_{ij} \\ &\Rightarrow \boldsymbol{\Omega} = (1-\alpha-\beta) \times \hat{\mathbf{S}} \end{aligned}$$

- Re-scale to get conditional correlation matrix

$$\begin{aligned} \hat{\mathbf{R}}_t^{DCC} &= \hat{\mathbf{D}}_t^{-1}\hat{\mathbf{Q}}_t^{DCC}\hat{\mathbf{D}}_t^{-1} \\ \hat{\mathbf{D}}_t &= \begin{bmatrix} \hat{q}_{11,t}^{DCC} & 0 & \cdots & 0 \\ 0 & \hat{q}_{22,t}^{DCC} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{q}_{kk,t}^{DCC} \end{bmatrix} \end{aligned}$$