Multivariate Volatility Modelling

Leopoldo Catania
Aarhus University and CREATES
leopoldo.catania@econ.au.dk

Why modelling multivariate returns?

Economics and financial economics present problems whose solutions need the specification and estimation of a multivariate distribution.

- the standard portfolio allocation problem
- the risk management of a portfolio of assets
- pricing of derivative contracts based on a more than one underlying asset (e.g., Quanto options)
- Financial contagion (shocks transmission volatility and returns)

What and How

Stylized facts:

- Volatility clustering
- Dynamic covariances and dynamic correlations

Financial variables have time-dependent second order moments.
Parametric models:

- Multivariate GARCH models
- Multivariate Stochastic volatility models
- Multivariate realized volatility models

General multivariate volatility model

Vector of returns:

$$\mathbf{y}_t = (y_{1t}, \dots, y_{Nt})' \quad (N \times 1)$$

$$\mathbf{y}_t - \mu_t = \epsilon_t = H_t^{1/2} \mathbf{z}_t$$

Let $\{\mathbf{z}_t\}$ be a sequence of $(N \times 1)$ i.i.d. random vector with the following characteristics:

$$E[\mathbf{z}_t] = \mathbf{0}$$

$$E\left[\mathbf{z}_t \mathbf{z}_t'\right] = \mathbf{I}_N$$

So

$$\mathbf{z}_t \overset{iid}{\sim} G(\mathbf{0}, \mathbf{I}_N),$$

for some distribution G

General multivariate volatility model

$$E_{t-1}\left(\epsilon_t\right) = 0$$

$$\begin{aligned} E_{t-1}\left(\epsilon_t\epsilon_t'\right) &= \mathbf{H}_t \\ E\left(\epsilon_t\epsilon_t'\right) &= \Sigma \end{aligned}$$

$$E_{t-1}[\cdot] = E[\cdot|\mathcal{F}_{t-1}]$$

where $\mathbf{H}_t$ is a matrix $(N \times N)$ positive definite and measurable with respect to the information set $\mathcal{F}_{t-1}$, that is the σ-field generated by the past observations: $\{\epsilon_{t-1}, \epsilon_{t-2}, \dots\}$.
The correlation matrix:

$$\mathrm{Corr}_{t-1}(\epsilon_t) = \mathbf{R}_t = \mathbf{D}_t^{-1/2}\mathbf{H}_t\mathbf{D}_t^{-1/2}$$

$$\mathbf{D}_t = \mathrm{diag}(h_{11,t}, \dots, h_{NN,t})$$

Conditions on $\mathbf{H}_t$

MVMs provide a parametric structure for the dynamic evolution of $\mathbf{H}_t$. MVMs must satisfy:

1. Diagonal elements of $\mathbf{H}_t$ must be strictly positive;
2. Positive definiteness of $\mathbf{H}_t$;
3. Stationarity: $E[\mathbf{H}_t]$ exists, finite and constant w.r.t. t.

Ideal Multivariate volatility model

Ideal characteristics of a MVM:

1. Estimation should be flexible for increasing N
2. It should allow for covariance spillovers and feedbacks;
3. Coefficients should have an economic or financial interpretation

Different approaches

Three approaches for constructing multivariate GARCH models:

1. direct generalizations of the univariate GARCH model of Bollerslev (1986); (VEC, BEKK)
2. linear combinations of univariate GARCH models; ((generalized) orthogonal models and latent factor models.)
3. nonlinear combinations of univariate GARCH models; (constant and dynamic conditional correlation models, copula-GARCH models)

Targeting: a way to simplify estimation

GARCH(1,1):

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

Long-run variance (if $(\alpha + \beta) < 1$):

$$\sigma^2 = E[\sigma_t^2] = \omega(1 - \alpha - \beta)^{-1}$$

Variance targeting:

$$\sigma_t^2 = \widehat{\sigma}^2(1 - \alpha - \beta) + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

$$\widehat{\sigma}^2 = T^{-1}\sum_t \widehat{\epsilon}_t^2$$

Introduction of targeting transforms the model estimation into a two-step estimation approach:

1. $\widehat{\sigma}^2$
2. $\widehat{\alpha}, \widehat{\beta}$

Theory for this two step estimator is presented by Francq et al. (2011).

Different approaches

N assets: N variances $+\ \frac{1}{2}N(N-1)$ covariances $= \frac{N}{2}(N+1)$.
Two alternative approaches:

- Models of $\mathbf{H}_t$
- Models of $\mathbf{D}_t$ and $\mathbf{R}_t$

The parametrization of $\mathbf{H}_t$ as a multivariate GARCH, which means as a function of the information set $\mathcal{F}_{t-1}$, allows each element of $\mathbf{H}_t$ to depend on q lagged of the squares and cross-products of ϵ_t, as well as p lagged values of the elements of $\mathbf{H}_t$. So the elements of the covariance matrix follow a vector of ARMA process in squares and cross-products of the disturbances.

Notation

Let **vech** denote the vector-half operator, which stacks the lower triangular elements of an $N \times N$ matrix as an $[N(N+1)/2] \times 1$ vector.
Let $\mathbf{A}$ be (2×2), then vech$(\mathbf{A})$

$$\text{vech}(\mathbf{A}) = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{22} \end{bmatrix}$$

Since the conditional covariance matrix $\mathbf{H}_t$ is symmetric, vech$(\mathbf{H}_t)$, which is of dimension $N(N+1)/2$, contains all the unique elements in $\mathbf{H}_t$.

The VEC GARCH

A natural **multivariate extension** of the univariate GARCH(p,q) model is

$$\begin{aligned} \text{vech}\left(\mathbf{H}_t\right) &= \mathbf{W} + \sum_{i=1}^{q} \mathbf{A}_i^* \text{vech}\left(\epsilon_{t-i}\epsilon_{t-i}'\right) + \sum_{j=1}^{p} \mathbf{B}_j^* \text{vech}\left(\mathbf{H}_{t-j}\right) \\ &= \mathbf{W} + \mathbf{A}^*\left(L\right)\text{vech}\left(\epsilon_t\epsilon_t'\right) + \mathbf{B}^*\left(L\right)\text{vech}\left(\mathbf{H}_t\right) \end{aligned}$$

$$\begin{aligned} \mathbf{A}^*\left(L\right) &= \mathbf{A}_1^* L + \ldots + \mathbf{A}_q^* L^q \\ \mathbf{B}^*\left(L\right) &= \mathbf{B}_1^* L + \ldots + \mathbf{B}_q^* L^p \end{aligned}$$

The VEC GARCH: $N = 2$

$$N^* \equiv \frac{N(N+1)}{2}$$

$$\begin{aligned} \mathbf{W} \quad &: \quad [N(N+1)/2] \times 1 \\ \mathbf{A}_i^*, \ \mathbf{B}_j^* \quad &: \quad [N^* \times N^*] \end{aligned}$$

$N = 2$, Vech-GARCH(1,1):

$$\begin{bmatrix} h_{11,t} \\ h_{21,t} \\ h_{22,t} \end{bmatrix} = \begin{bmatrix} w_1^* \\ w_2^* \\ w_3^* \end{bmatrix} + \begin{bmatrix} a_{11}^* & a_{12}^* & a_{13}^* \\ a_{21}^* & a_{22}^* & a_{23}^* \\ a_{31}^* & a_{32}^* & a_{33}^* \end{bmatrix} \begin{bmatrix} \epsilon_{1,t-1}^2 \\ \epsilon_{1,t-1}\epsilon_{2,t-1} \\ \epsilon_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} b_{11}^* & b_{12}^* & b_{13}^* \\ b_{21}^* & b_{22}^* & b_{23}^* \\ b_{31}^* & b_{32}^* & b_{33}^* \end{bmatrix} \begin{bmatrix} h_{11,t-1} \\ h_{21,t-1} \\ h_{2,t-1} \end{bmatrix}$$

The VEC GARCH

- This general formulation is termed *vec representation* by Engle and Kroner (1995).
- Earlier references can be found in Kraft and Engle (1982) and Bollerslev et al. (1992).
- The number of parameters is $\frac{N(N+1)}{2}\left[1+(p+q)\frac{N(N+1)}{2}\right]$.
- Even for low dimensions of N and small values of p and q the number of parameters is very large; for $N=5$ and $p=q=1$ the unrestricted version of (1) contains 465 parameters.
- The number of parameters is of order $O(N^4)$: the *curse of dimensionality*.

For any parametrization to be sensible, we require that $\mathbf{H}_t$ be positive definite for all values of ϵ_t in the sample space. Unfortunately, in the *vech* representation this restriction can be difficult to check, let alone impose during estimation.

Reduce the number of parameters: The Diagonal VEC

A natural restriction is the *diagonal representation*, in which each element of the covariance matrix depends only on past values of itself and past values of $\varepsilon_{jt}\varepsilon_{kt}$. In the diagonal model the $\mathbf{A}_i^*$ and $\mathbf{B}_j^*$ matrices are all taken to be diagonal.
For $N = 2$ and $p = q = 1$, the diagonal model is written as:

$$\begin{bmatrix} h_{11,t} \\ h_{21,t} \\ h_{22,t} \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} + \begin{bmatrix} a_{11}^* & 0 & 0 \\ 0 & a_{22}^* & 0 \\ 0 & 0 & a_{33}^* \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} b_{11}^* & 0 & 0 \\ 0 & b_{22}^* & 0 \\ 0 & 0 & b_{33}^* \end{bmatrix} \begin{bmatrix} h_{11,t-1} \\ h_{21,t-1} \\ h_{22,t-1} \end{bmatrix}$$

$$h_{ij,t} = w_i^* + a_{ii}^* \epsilon_{i,t-1}\epsilon_{j,t-1} + b_{ii}^* h_{ij,t-1}$$

This specification has been used by Bollerslev et al. (1988). Note that your book writes this as

$$\boldsymbol{H_t} = \boldsymbol{\Omega} + \sum_{i=1}^{p} A_i \odot (\varepsilon_{t-i}\varepsilon_{t-i}') + \sum_{j=1}^{q} B_j \odot \boldsymbol{H}_{t-j}$$

BEKK model

Engle and Kroner (1995) propose a parametrization that imposes positive definiteness restrictions.

Consider the following model

$$\mathbf{H}_t = \mathbf{C}\mathbf{C}' + \sum_{i=1}^{q} \mathbf{A}_i \epsilon_{t-i} \epsilon_{t-i}' \mathbf{A}_i' + \sum_{j=1}^{p} \mathbf{B}_j \mathbf{H}_{t-j} \mathbf{B}_j' \tag{1}$$

where $\mathbf{C}$, $\mathbf{A}_i$ and $\mathbf{B}_j$ are $(N \times N)$.

- The intercept matrix is decomposed into $\mathbf{C}\mathbf{C}'$, where $\mathbf{C}$ is a lower triangular matrix.
- Without any further assumption $\mathbf{C}\mathbf{C}'$ is positive semidefinite.
- This representation is general, it includes all positive definite diagonal representations and nearly all positive definite *vech* representations.

(The acronym BEKK stands for Baba, Engle, Kraft, and Kroner.)

BEKK model

Consider the simple BEKK(1,1) model:

$$\mathbf{H}_t = \mathbf{C}\mathbf{C}' + \mathbf{A}_1\epsilon_{t-1}\epsilon_{t-1}'\mathbf{A}_1' + \mathbf{B}_1\mathbf{H}_{t-1}\mathbf{B}_1' \tag{2}$$

BEKK identification (Engle and Kroner (1995))

Suppose that the diagonal elements in $\mathbf{C}$ are restricted to be positive and that a_{11} and b_{11} are also restricted to be positive. Then there exists no other $\mathbf{C}$, $\mathbf{A}_1$, $\mathbf{B}_1$ in the model (2) that will give an equivalent representation.

BEKK model

The purpose of the restrictions is to eliminate all other observationally equivalent structures.

For example, as relates to the term $\mathbf{A}_1 \epsilon_{t-1} \epsilon'_{t-1} \mathbf{A}'_1$ the only other observationally equivalent structure is obtained by replacing $\mathbf{A}_1$ by $-\mathbf{A}_1$. The restriction that a_{11} (b_{11}) be positive could be replaced with the condition that a_{ij} (b_{ij}) be positive for a given i and j, as this condition is also sufficient to eliminate $-\mathbf{A}_1$ from the set of admissible structures.

BEKK model, $N = 2$

MGARCH(1,1)-BEKK, $N = 2$:

$$
\begin{aligned}
\mathbf{H}_t &= \mathbf{C}\mathbf{C}' + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-1}^2 & \varepsilon_{1t-1}\varepsilon_{2t-1} \\ \varepsilon_{2t-1}\varepsilon_{1t-1} & \varepsilon_{2t-1}^2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}' \\
&+ \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} h_{11t-1} & h_{12t-1} \\ h_{21t-1} & h_{22t-1} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}'
\end{aligned}
$$

Positive definiteness of $\boldsymbol{H}_t$

BEKK-GARCH(p,q) model (Engle and Kroner (1995)):

Sufficient condition for positive definiteness of $\mathbf{H}_t$

If $\mathbf{H}_0, \mathbf{H}_{-1}, \dots, \mathbf{H}_{-p+1}$ are all positive definite, then the BEKK parametrization yields a positive definite $\mathbf{H}_t$ for all possible values of ε_t if $\mathbf{C}$ is a full rank matrix or if any $\mathbf{B}_j$ $j = 1, \dots, p$ is a full rank matrix.

BEKK: Three cases

1. BEKK
$$\mathbf{H}_t = \Sigma + \mathbf{A}_1 \left(\epsilon_{t-1}\epsilon'_{t-1}\right) \mathbf{A}'_1 + \mathbf{B}_1 \left(\mathbf{H}_{t-1}\right) \mathbf{B}'_1$$
Number of parameters is $\frac{N(N+1)}{2} + 2N^2 = O(N^2) + O(N^2) = O(N^2)$.

2. Diagonal BEKK
$$\mathbf{H}_t = \Sigma + \mathbf{A}_1 \left(\epsilon_{t-1}\epsilon'_{t-1}\right) \mathbf{A}'_1 + \mathbf{B}_1 \left(\mathbf{H}_{t-1}\right) \mathbf{B}'_1$$
where $\mathbf{A}_1$ and $\mathbf{B}_1$ are diagonal. Number of parameters is $\frac{N(N+1)}{2} + 2N = O(N^2) + O(N) = O(N^2)$.

3. Scalar BEKK
$$\mathbf{H}_t = \Sigma + \alpha \left(\epsilon_{t-1}\epsilon'_{t-1}\right) + \beta \left(\mathbf{H}_{t-1}\right)$$
for $\alpha + \beta < 1$. Number of parameters is $\frac{N(N+1)}{2} + 2 = O(N^2) + O(1) = O(N^2)$.

BEKK: from $O(N^2)$ to $O(N)$ and $O(1)$

The same targeting approach used for univariate GARCH can be used with the BEKK model (assuming that a stationary solution exists).
For example, we can set

$$\Sigma = \left(I - \sum_{i=1}^{p} \boldsymbol{A}_i - \sum_{j=1}^{q} \boldsymbol{B}_j \right) \hat{\boldsymbol{S}},$$

where $\hat{\boldsymbol{S}} = \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \epsilon_t'$. This is discussed by Engle and Mezrich (1996).
The estimation procedure then reduces to:

1) Estimate the empirical covariance of ϵ_t, $\hat{\boldsymbol{S}}$.
2) Estimate the other parameters of the model fixing $\boldsymbol{S} = \hat{\boldsymbol{S}}$

In this way, the second estimation step (which is a numerical optimization) is of order

1) $O(N^2)$ for the unrestricted BEKK.
2) $O(N)$ for the diagonal BEKK.
3) $O(1)$ for the scalar BEKK.

Some theory for this estimation procedure is discussed in Francq et al. (2016).

Estimation by ML

Let θ be the vector that contains all the parameters of the model. The log-likelihood function for θ given a sequence of T observations $\{\epsilon_1, \ldots, \epsilon_T\}$ obtained under the assumption of conditional multivariate normality is:

$$\log L_T(\theta; \epsilon_1, \ldots, \epsilon_T) = -\frac{1}{2}\left[TN \log(2\pi) + \sum_{t=1}^{T}\left(\log|\mathbf{H}_t| + \epsilon_t' \mathbf{H}_t^{-1} \epsilon_t\right)\right]$$

- The assumption of conditional normality can be quite restrictive.
- The symmetry imposed under normality is difficult to justify, and the tails of even conditional distributions often seem fatter than that of normal distribution.

If the interest in only on the variances, the estimator can be regarded as QML.

Direct Modelling of Correlations

- These models are based on a decomposition of the $\boldsymbol{H}_t$.
- The conditional var-cov matrix is expressed as

$$\mathbf{H}_t = \mathbf{D}_t^{1/2}\mathbf{R}_t\mathbf{D}_t^{1/2}$$

  where $\mathbf{R}_t$ is possibly time-varying and $\mathbf{D}_t$ is a diagonal matrix containing the conditional variances.
- Conditional correlations and variances are separately modeled.

Constant Conditional Correlations (CCC)

Bollerslev et al. (1990)'s Constant Conditional Correlations model:
The time-varying conditional covariances are parameterized to be proportional to the product of the corresponding conditional standard deviations. The model assumptions are:

$$E_{t-1}\left[\epsilon_t \epsilon_t'\right] = \mathbf{H}_t$$

$$\{\mathbf{H}_t\}_{ii} = h_{it} \quad i = 1, \dots, N$$

$$\{\mathbf{H}_t\}_{ij} = h_{ijt} = \rho_{ij} h_{it}^{1/2} h_{jt}^{1/2} \qquad i \neq j \quad i,j = 1, \dots, N$$

$$\mathbf{D}_t = \mathrm{diag}\left\{h_{1t}, \dots, h_{Nt}\right\}$$

Constant Conditional Correlations (CCC)

The conditional covariance matrix can be written as:

$$\mathbf{H}_t = \mathbf{D}_t^{1/2}\mathbf{R}\mathbf{D}_t^{1/2}$$

$$\mathbf{H}_t = \begin{bmatrix} h_{1t}^{1/2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_{Nt}^{1/2} \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} & \ldots & \rho_{1N} \\ \rho_{21} & 1 & \ldots & \vdots \\ \vdots & \vdots & \ldots & \rho_{N-1N} \\ \rho_{N1} & \ldots & \rho_{NN-1} & 1 \end{bmatrix} \begin{bmatrix} h_{1t}^{1/2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_{Nt}^{1/2} \end{bmatrix}$$

When $N = 2$

$$\begin{aligned} \mathbf{H}_t &= \begin{bmatrix} h_{1t}^{1/2} & 0 \\ 0 & h_{2t}^{1/2} \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} \\ \rho_{21} & 1 \end{bmatrix} \begin{bmatrix} h_{1t}^{1/2} & 0 \\ 0 & h_{2t}^{1/2} \end{bmatrix} \\ &= \begin{bmatrix} h_{1t} & \rho_{12} h_{1t}^{1/2} h_{2t}^{1/2} \\ \rho_{12} h_{1t}^{1/2} h_{2t}^{1/2} & h_{2t} \end{bmatrix}. \end{aligned}$$

Constant Conditional Correlations (CCC)

- If the conditional variances along the diagonal in the $\mathbf{D}_t$ matrices are all positive, and the conditional correlation matrix $\mathbf{R}$ is positive definite, then the sequence of conditional covariance matrices $\{\mathbf{H}_t\}$ is guaranteed to be positive definite a.s. for all t.
- Furthermore the inverse of $\mathbf{H}_t$ is given by

$$\mathbf{H}_t^{-1} = \mathbf{D}_t^{-1/2}\mathbf{R}^{-1}\mathbf{D}_t^{-1/2}.$$

  When calculating the log-likelihood function only one matrix inversion is required for each evaluation.
- CCC is generally estimated in two steps:
  1. conditional variances are estimated employing the marginal likelihoods
  2. $\mathbf{R}$ is estimated using the sample estimator of standardized residuals $\widehat{\boldsymbol{\eta}}_t = \widehat{\mathbf{D}}_t^{-1/2}\mathbf{y}_t$ (assuming $\mu_t = \mathbf{0}$).

Constant Conditional Correlations (CCC)

- The CCC solves the *curse of dimensionality* problem of MGARCH models
- The number of parameters is $O(N^2)$ but these are not jointly estimated. The two-step estimation procedure impacts on the computational issues.
- Asymptotic properties of QMLE estimators studied in Ling and McAleer (2003).

Dynamic Conditional Correlations (DCC)

The CCC has two main limitations:

1. No spillover neither feedback effects across conditional variances
2. Correlations are static

The evolution of CCC is the Dynamic Conditional Correlation (DCC) Model of Engle (2002). The DCC is an extension of the Bollerslev's CCC Model.

Dynamic Conditional Correlations (DCC)

The conditional correlation between two random variables, X_t and Y_t is defined as:

$$\rho_{YX,t} = \frac{Cov_{t-1}(X_t Y_t)}{\sqrt{E_{t-1}(X_t - \mu_{x,t})^2 E_{t-1}(Y_t - \mu_{Y,t})^2}}$$

Assets returns conditional distribution:

$$\mathbf{y}_t | \mathcal{F}_{t-1} \sim N(\mathbf{0}, \mathbf{H}_t)$$

$$\mathbf{H}_t = \mathbf{D}_t^{1/2} \mathbf{R}_t \mathbf{D}_t^{1/2}.$$

$$\mathbf{D}_t = \mathrm{diag}(\mathrm{Var}_{t-1}(y_{1t}), \dots, \mathrm{Var}_{t-1}(y_{Nt}))$$

where the $\mathrm{Var}_{t-1}(y_{it})$, $i = 1, \dots, N$ are modeled as univariate GARCH processes.

Dynamic Conditional Correlations (DCC)

The standardized returns are:

$$\eta_t = \mathbf{D}_t^{-1/2}\mathbf{y}_t$$

$$E_{t-1}(\eta_t\eta_t') = \mathbf{D}_t^{-1/2}\mathbf{H}_t\mathbf{D}_t^{-1/2} = \mathbf{R}_t = \{\rho_{ij,t}\}$$

To ensure that $\mathbf{R}_t$ is positive defined we can employ the following tranformation

$$\boldsymbol{R}_t = \widetilde{\boldsymbol{Q}}_t^{-1/2}\boldsymbol{Q}_t\widetilde{\boldsymbol{Q}}_t^{-1/2},$$

where $\widetilde{\boldsymbol{Q}}_t$ is a diagonal matrix with typical elements $\widetilde{q}_{ii,t} = q_{ii,t}$. This implies that The conditional correlation are

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{\widetilde{q}_{ii,t}\widetilde{q}_{jj,t}}}.$$

Where $q_{ij,t}$ are assumed to follow a GARCH(1,1) model

$$q_{ij,t} = \overline{q}_{ij}(1-\alpha-\beta) + \alpha\eta_{i,t-1}\eta_{j,t-1} + \beta q_{ij,t-1} \tag{3}$$

The term $\overline{q}_{ij}$ is not the unconditional correlation ($\overline{q}_{ij} \neq \overline{\rho}_{ij}$) between η_{it} and η_{jt}; the unconditional correlation between η_{it} and η_{jt} has no closed form.

Wrong targeting!

- Engle (2002) assumes that $\overline{\rho}_{ij} \simeq \overline{q}_{ij}$, and writes

$$q_{ij,t} = \overline{\rho}_{ij}(1 - \alpha - \beta) + \alpha\eta_{i,t-1}\eta_{j,t-1} + \beta q_{ij,t-1} \tag{4}$$

- Aielli (2013) and Engle et al. (2007) suggest to modify the standard DCC in order to correct the asymptotic bias which is due to the fact that $\frac{1}{T}\sum_t \epsilon_t\epsilon_t'$ does not converge to $\overline{\mathbf{Q}}$.
- It is known though that the impact of this is very small, see Engle and Sheppard (2001).

DCC has a number of additional issues which are largely ignored in empirical applications, see Caporin and McAleer (2013).

Positive definiteness

The matrix $\mathbf{Q}_t$ is positive definite for all t as long as it is a weighted average of positive definite matrices and positive semidefinite matrices.
To ensure p.-d-ness of $\mathbf{Q}_t$ we must impose $\alpha + \beta < 1$ In matrix from:

$$\mathbf{Q}_t = \overline{\mathbf{Q}}(1 - \alpha - \beta) + \alpha(\boldsymbol{\eta}_{t-1}\boldsymbol{\eta}'_{t-1}) + \beta\mathbf{Q}_{t-1}$$

where $\overline{\mathbf{Q}}$ is positive definite and set to the unconditional covariance matrix of $\boldsymbol{\eta}_t$.

The log-likelihood function can be written as:

$$
\begin{aligned}
\log L_T &= -\frac{1}{2}\sum_{t=1}^{T}\left(N\log(2\pi) + \log|\mathbf{H}_t| + \mathbf{y}_t'\mathbf{H}_t^{-1}\mathbf{y}_t\right) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(N\log(2\pi) + \log|\mathbf{D}_t^{1/2}\mathbf{R}_t\mathbf{D}_t^{1/2}| + \mathbf{y}_t'\mathbf{D}_t^{-1/2}\mathbf{R}_t^{-1}\mathbf{D}_t^{-1/2}\mathbf{y}_t\right) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(N\log(2\pi) + \log|\mathbf{D}_t| + \log|\mathbf{R}_t| + \eta_t'\mathbf{R}_t^{-1}\eta_t\right)
\end{aligned}
$$

Adding and subtracting $\mathbf{y}_t'\mathbf{D}_t^{-1/2}\mathbf{D}_t^{-1/2}\mathbf{y}_t = \eta_t'\eta_t$

$$
\begin{aligned}
\log L_T &= -\frac{1}{2}\sum_{t=1}^{T}\Big(N\log(2\pi) + \log|\mathbf{D}_t| + \mathbf{y}_t'\mathbf{D}_t^{-1/2}\mathbf{D}_t^{-1/2}\mathbf{y}_t \\
&\qquad -\eta_t'\eta_t + \log|\mathbf{R}_t| + \eta_t'\mathbf{R}_t^{-1}\eta_t\Big) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(N\log(2\pi) + \log|\mathbf{D}_t| + \mathbf{y}_t'\mathbf{D}_t^{-1}\mathbf{y}_t\right) \\
&\quad -\frac{1}{2}\sum_{t=1}^{T}\left(\eta_t'\mathbf{R}_t^{-1}\eta_t - \eta_t'\eta_t + \log|\mathbf{R}_t|\right)
\end{aligned}
$$

Likelihood decomposition

Volatility component:

$$\mathcal{L}_V(\theta) \equiv \log L_{V,T}(\theta) = -\frac{1}{2}\sum_{t=1}^{T}\left(N\log(2\pi) + \log|\mathbf{D}_t| + \mathbf{y}_t'\mathbf{D}_t^{-1}\mathbf{y}_t\right)$$

Correlation component:

$$\mathcal{L}_C(\theta,\phi) \equiv \log L_{C,T}(\theta,\phi) = -\frac{1}{2}\sum_{t=1}^{T}\left(\eta_t'\mathbf{R}_t^{-1}\eta_t - \eta_t'\eta_t + \log|\mathbf{R}_t|\right)$$

θ denotes the parameters in $\mathbf{D}_t$ and ϕ the parameters in $\mathbf{R}_t$.

$$\mathcal{L}(\theta,\phi) = \mathcal{L}_V(\theta) + \mathcal{L}_C(\theta,\phi)$$

$$\mathcal{L}_V(\theta) = -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{N}\left(\log(2\pi) + \log(h_{i,t}) + \frac{r_{i,t}^2}{h_{i,t}}\right).$$

The likelihood is apparently the sum of individual GARCH likelihoods, which will be jointly maximized by separately maximizing each term.

Two-step procedure:

1. $$\hat{\theta} = \arg\max\{\mathcal{L}_V(\theta)\}$$

2. $$\max_{\phi}\{\mathcal{L}_C(\hat{\theta}, \phi)\}.$$

Under regularity conditions, consistency of the first step will ensure consistency of the second step. The maximum of the second step will a function of the first step parameter estimates. If the first step is consistent then the second step will be too as long as the function is continuous in a neighborhood of the true parameters.

Time-varying correlation of Tse and Tsui (2002)

Tse and Tsui (2002) proposes a specification similar to the DCC. In the time-varying correlation (TVC) model of Tse and Tsui (2002) the correlation matrix evolves as an ARMA process

$$\mathbf{R}_t = (1 - \alpha - \beta)\mathbf{R} + \alpha \boldsymbol{\Psi}_{t-1} + \beta \mathbf{R}_{t-1},$$

where $\mathbf{R}$ is a $N \times N$ constant matrix as in the DCC specification, and the driving force $\boldsymbol{\Psi}_t$ is a $N \times N$ matrix with general element:

$$\psi_{ij,t} = \frac{\sum_{h=0}^{m-1} \eta_{i,t-h}\eta_{j,t-h}}{\sqrt{\left(\sum_{h=0}^{m-2} \eta_{i,t-h}^2\right)\left(\sum_{h=0}^{m-2} \eta_{j,t-h}^2\right)}}.$$

Note that $m \geq 1$ is an integer that must be selected by the econometrician. Usually $m = 1$ is used. $\mathbf{R}_t$ is positive definite if $\alpha \geq 0$, $\beta \leq 1$ with $\alpha + \beta \leq 1$, and $\mathbf{R}$ and $\mathbf{R}_1$ are positive definite.

References I

Aielli, G. P. (2013). Dynamic conditional correlation: on properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299.

Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of econometrics*, 52(1-2):5–59.

Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of political Economy*, 96(1):116–131.

Bollerslev, T. et al. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *Review of Economics and statistics*, 72(3):498–505.

Caporin, M. and McAleer, M. (2013). Ten things you should know about the dynamic conditional correlation representation. *Econometrics*, 1(1):115–126.

References II

Engle, R. and Mezrich, J. (1996). Garch for groups: A round-up of recent developments in garch techniques for estimating correlation. *RISK-LONDON-RISK MAGAZINE LIMITED-*, 9:36–40.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric theory*, 11(1):122–150.

Engle, R. F., Shephard, N., and Sheppard, K. (2007). Fitting and testing vast dimensional time-varying covariance models.

Engle, R. F. and Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate garch. Technical report, National Bureau of Economic Research.

Francq, C., Horvath, L., and Zakoïan, J.-M. (2011). Merits and drawbacks of variance targeting in garch models. *Journal of Financial Econometrics*, 9(4):619–656.

Francq, C., Horvath, L., and Zakoian, J.-M. (2016). Variance targeting estimation of multivariate garch models. *Journal of Financial Econometrics*, 14(2):353–382.

References III

Kraft, D. F. and Engle, R. F. (1982). Autoregressive conditional heteroskedasticity in multiple time series. *UCSD manuscript*.

Ling, S. and McAleer, M. (2003). Asymptotic theory for a vector arma-garch model. *Econometric theory*, 19(2):280–310.

Tse, Y. K. and Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, 20(3):351–362.