A Teaching Summary on the Estimation of DCC-type Models

Qunxing Pan*

School of Finance, Nanjing University of Finance and Economics
Corresponding author. Email: pqxjs2018@nufe.edu.cn

ABSTRACT

The three-stage estimation technique that was designed to estimate the dynamic conditional correlation (DCC) model and its variants is computationally convenient but it may be less efficient than the quasi-maximum likelihood method for simultaneous estimation of all parameters, due to the inconsistent moment estimators used in the second stage. In this article, we have made an in-depth survey on the advantages and disadvantages of this technique, to facilitate readers to carry out relevant empirical research. We suggest that the consistent DCC (cDCC) model and its variants should be widely applied to typical financial series since this three-stage estimation is more feasible for them.

Keywords: *Teaching summary, DCC-type models, Three-stage estimation, Consistence of estimators.*

1. INTRODUCTION

In the past 30 years, modeling the time-varying covariance of asset returns has become a part of financial econometrics. In risk management, it has been proved to be a useful tool to estimate the time-varying volatility and co-movement between financial series. The correlation between returns is related to the time-varying beta coefficient in the CAPM model, the hedge ratio and the Value-at-Risk of a portfolio.

Alternative multivariate GARCH models, which aim to provide simple solutions to the problems described above, have become a widely discussed topic. Among them, the most famous one is the dynamic conditional correlation (DCC) model (Engle, 2002), which is one of the most cited works in parametric modeling of time-varying correlations of multivariate portfolios.

The DCC model is a generalization of the constant conditional correlation (CCC) model (Bollerslev, 1990) and its extended (ECCC) model (Jeantheau, 1998), where volatilities are time-varying, but conditional correlations are assumed to be constant. Nevertheless, it is generally believed that DCC model does not perform well for the high-dimension case. The reason is the assumption that the same parameters drive all the correlations (that is, the correlation evolve according to a process that has the same innovation impact and smoothing parameters for all pairs of variables) is very strong. To better capture the heterogeneity in conditional correlations, Cappiello et al. (2006) proposed the asymmetric generalized DCC (AG-DCC) model by introducing two modifications: one is the asset-specific correlation evolution parameters, the other is the asymmetric parameters in conditional correlations. The asymmetric DCC (A-DCC), as a special case of the AG-DCC, has more flexibility and practicability. Li et al. (2002), Bauwens et al. (2006), and Silvennoinen and Teräsvirta (2008) reviewed a large number of univariate and multivariate GARCH-type models from different perspectives.

Let $a_t = (a_{1,t}, a_{2,t}, \cdots, a_{k,t})'$ be the vector of innovations or shocks from k financial asset returns at time t for $t = 1, 2, \cdots, T$ with $E(a_t|\mathcal{F}_{t-1}) = 0$ and $Cov(a_t|\mathcal{F}_{t-1}) = \Sigma_t$, where $\mathcal{F}_{t-1}$ is the innovation set up to time $t-1$. The DCCs cited above, including the CCC and ECCC, use the fact that the conditional covariance matrix Σ_t is decomposed as follows:

$$\Sigma_t = D_t R_t D_t \tag{1}$$

where $D_t = D_t(\boldsymbol{\theta})$ is the diagonal matrix of time-varying volatilities $\sigma_{ii,t}$, $i = 1, 2, \cdots, k$, $\boldsymbol{\theta}$ is a parameter vector, and R_t is the $k \times k$ (possibly) time-varying conditional correlation matrix. Since Engle's DCC model was designed to consider a three-stage estimation of Σ_t, these variants of DCC have also adopted this estimation method. This three-stage process is roughly as follows. In the first stage, univariate volatility models that assume Gaussian distribution errors (regardless of the real error distribution) are used to model the variances $\sigma^2_{ii,t}$ for each innovation, and estimates of $\sigma^2_{ii,t}$ are obtained. In the second stage, residuals, transformed by their estimated standard deviations from the first stage, are used to estimate the intercept parameters of the conditional correlation. The third stage is to estimate the coefficients controlling the correlation dynamics after the intercept parameters have been estimated.

The three-stage estimation technique is much more convenient in computation, but it may be less efficient than the quasi-maximum likelihood method for all parameters estimated simultaneously. In this article, we shall make a comprehensive summary of the three-stage estimation technique and procedure, to facilitate readers to carry out relevant empirical research. The remainder of this article is organized as follows: Section 2 discusses the specification of univariate volatility models, Section 3 discusses positive definiteness and estimation of the DCC-type models, Section 4 focuses on the two-dimensional case regarding element versions, and Section 5 gives some comments.

2. UNIVARIATE VOLATILITY SPECIFICATIONS

Engle and Sheppard (2008) have asymptotically studied the efficiency of the three-stage estimation process, and Engle and Sheppard (2001) have simulated it. By comparing the estimated univariate model with the exact knowledge of univariate parameters, they found that the univariate estimations had almost no results. A similar test of conditional correlation process shows that, compared with the exact knowledge of the unconditional correlation matrix, the separate estimation of intercept and dynamic parameters leads to small finite sample bias. However, the existing empirical studies almost assume that the univariate models are correctly specified. If the univariate models are not well specified, the estimation of correlations will no more be consistent. The choice of univariate models does not affect the sign of standardized residuals, and many of them generate relatively similar volatility patterns, so this correlation may not be very sensitive to univariate models, at least in a reasonable range. Nevertheless, people should not just rely on this intuition, but should carry out a broader selection process for univariate models to minimize the risk of inconsistent correlation estimation caused by the univariate models. Cappiello et al. (2006) suggested using the BIC to select univariate volatility specifications from a class of models that can capture the common characteristics of asset return variances. They claimed that it was appropriate to use BIC, since it can make, as long as it is a member of the group, an asymptotical selection of the correct specification.

We describe several widely used univariate GARCH specifications with 1-lag innovation and 1-lag volatility as empirical studies have shown that the (1, 1)-order GARCH specifications are sufficient for most financial asset return series, thereby effectively reducing the lag length in the ARCH model, which may induce cumbersome computation.

GARCH (1, 1) model:

$$a_{i,t} = \sigma_{ii,t} \times \epsilon_{i,t},\ \sigma_{ii,t}^2 = \omega_{i0} + \alpha_{i1} a_{i,t-1}^2 + \beta_{i1}\sigma_{ii,t-1}^2,\ \epsilon_{i,t} \sim i.i.d.(0,1) \quad (2)$$

TGARCH (1, 1) model:

$$a_{i,t} = \sigma_{ii,t} \times \epsilon_{i,t},\ \sigma_{ii,t} = \omega_{i0} + (\alpha_{i1}^+ a_{i,t-1}^+ \alpha_{i1}^- a_{i,t-1}^-) + \beta_{i1}\sigma_{ii,t-1},\ \epsilon_{i,t} \sim i.i.d.(0,1) \quad (3)$$

where $a_{i,t-1}^+ = max(a_{i,t-1}, 0)$, $a_{i,t-1}^- = min(a_{i,t-1}, 0)$.

GJR-GARCH (1, 1) model:

$$a_{i,t} = \sigma_{ii,t} \times \epsilon_{i,t},\ \sigma_{ii,t}^2 = \omega_{i0} + (\alpha_{i1} + \gamma_{i1} I_{a_{i,t-1}<0}) a_{i,t-1}^2 + \beta_{i1}\sigma_{ii,t-1}^2,\ \epsilon_{i,t} \sim i.i.d.(0,1) \quad (4)$$

where the indicator variable $I_{a_{i,t-1}<0}$ takes on value1 if $a_{i,t-1} < 0$, and 0 otherwise.

EGARCH (1, 1) model:

$$a_{i,t} = \sigma_{ii,t} \times \epsilon_{i,t},\ log(\sigma_{ii,t}^2) = \omega_{i0} + \alpha_{i1} g(\epsilon_{i,t-1}) + \beta_{i1} log(\sigma_{ii,t-1}^2),\ \epsilon_{i,t} \sim i.i.d.(0,1) \quad (5)$$

where $g(\epsilon_{i,t-1}) = \theta_{i1}\epsilon_{i,t-1} + \gamma_{i1}(|\epsilon_{i,t-1}| - E|\epsilon_{i,t-1}|)$.

APARCH (1, 1) model:

$$a_{i,t} = \sigma_{ii,t} \times \epsilon_{i,t},\ \sigma_{ii,t}^\nu = \omega_{i0} + \alpha_{i1}(|a_{i,t-1}| - \gamma_{i1} a_{i,t-1})^\nu + \beta_{i1}\sigma_{ii,t-1}^\nu,\ \epsilon_{i,t} \sim i.i.d.(0,1) \quad (6)$$

where $\nu \geq 0, -1 < \gamma_{i1} < 1$.

FIGARCH (1, 1) model:

$$a_{i,t} = \sigma_{ii,t} \times \epsilon_{i,t},\ \sigma_{ii,t}^2 = \frac{\omega_{i0}}{1-\beta_{i1}} + [1 - (1 - L)^d \frac{1-\alpha_{i1}L}{1-\beta_{i1}L}] a_{i,t}^2,\ \epsilon_{i,t} \sim i.i.d.(0,1) \quad (7)$$

where

$(1 - L)^d = 1 - \frac{dL}{1!} - \frac{d(1-d)}{2!} L^2 - \cdots - \frac{d(1-d)\cdots(n-1-d)}{n!} L^n - \cdots$ is the fractional difference, L is the lag operator, $0 < d < 1$.

HYGARCH (1, 1) model:

$$a_{i,t} = \sigma_{ii,t} \times \epsilon_{i,t},\ \sigma_{ii,t}^2 = \frac{\omega_{i0}}{1-\beta_{i1}} + \{1 - [1 - \tau + \tau(1 - L)^d \frac{1-\alpha_{i1}L}{1-\beta_{i1}L}]\} a_{i,t}^2,\ \epsilon_{i,t} \sim i.i.d.(0,1) \quad (8)$$

where $\tau \geq 0$.

The simplest of the volatility models is GARCH, followed by TGARCH, GJR-GARCH and EGARCH (all of which considering threshold/leverage effects), and APARCH (encompassing both threshold/leverage effects and estimated power of variance evolution). FIGARCH and HYGARCH models are often used to describe volatility long memory, but we should be cautious due to the existence of covariance stationarity.

If $\epsilon_t | \mathcal{F}_{t-1} \sim N(0, I_k)$, then the quasi-log-likelihood function of this vector estimator $\boldsymbol{\theta}$, composed of these parameters in the volatility models, can be expressed as

$$QL_1(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{t=1}^T [klog(2\pi) + 2log|D_t| + a'_t D_t^{-2} a_t] \quad (9)$$

where $\epsilon_t = D_t^{-1}(\boldsymbol{\theta}) a_t$.

3. ESTIMATION OF DCC-TYPE MODELS: MATRIX VERSION

Once the univariate volatility models are estimated, the standardized residuals, $\hat{\epsilon}_{i,t} = a_{i,t}/\hat{\sigma}_{ii,t}$, may be used to estimate the correlation parameters of DCC-type models.

3.1. Specification of DCCs

The evolution of the correlation in Engle's (2002) DCC model is given by

$$R_t = diag(Q_t)^{-1/2} Q_t diag(Q_t)^{-1/2} \quad (10)$$

$$Q_t = (1-a-b)S + a\epsilon_{t-1}\epsilon'_{t-1} + bQ_{t-1} \quad (11)$$

where R_t is the time-varying conditional correlation matrix with all diagonal elements 1, a and b are two nonnegative parameter such that $a+b<1$, S is a positive definite matrix (see below), and $diag(Q_t)$ is a diagonal matrix composed of the diagonal elements of the matrix Q_t.

It is noted that the model described by Eqs. (10) and (11) does not allow for asset-specific innovations and smoothing parameters or asymmetries. Hence Cappiello et al. (2006) modified the correlation evolution to be an asymmetric generalized DCC

$$Q_t = (S - A'SA - B'SB) + A'\epsilon_{t-1}\epsilon'_{t-1}A + B'Q_{t-1}B + \Gamma'[\eta_{t-1}\eta'_{t-1} - E(\eta_t\eta'_t)]\Gamma \quad (12)$$

where A, B and Γ are $k\times k$ parameter matrices, $\eta_t = I_{\epsilon_t<0} \bowtie \epsilon_t$, $I_{\epsilon_t<0}$ is a $k\times 1$ indicator function as the case of one variable in Eq. (4), while "$\bowtie$" indicates the Hadamard product. They referred to the model in Eqs. (10) and (12) as the AG-DCC. The asymmetric diagonal DCC (AD-DCC) is obtained if A, B and Γ are replaced by diagonal matrices, the asymmetric DCC (A-DCC) is further obtained if A, B and Γ are replaced by scalar matrices. The symmetric DCC is a special case of A-DCC when $\Gamma = 0$. The AG-DCC is at the cost of added parameters and complexity, so it is not very applicable. The AD-DCC is appropriate for applications to many financial assets as its parameters are easier to interpret. Of course, the scalar specifications (A-DCC and DCC) are preferred if the asset number is very large.

It can be clearly seen from Eq. (12) that a sufficient condition for Q_t to be positive definite is the initial matrix Q_0 is positive definite and the intercept matrix $S - A'SA - B'SB - \Gamma'E(\eta_t\eta'_t)\Gamma$ is positive semi-definite (see Bauwens et al. (2016) for further details). In the DCC, the condition is only $a \geq 0, b \geq 0, a+b<1$. In the A-DCC,

$$Q_t = (1-a-b)S + a\epsilon_{t-1}\epsilon'_{t-1} + bQ_{t-1} + \gamma[\eta_{t-1}\eta'_{t-1} - E(\eta_t\eta'_t)] \quad (13)$$

The sufficient condition for Q_t to be positive definite is $a+b+\lambda\gamma<1$, $a\geq 0$, $b\geq 0$, $\gamma\geq 0$, where λ is the maximum eigenvalue of matrix $S^{-1/2}E(\eta_t\eta'_t)S^{-1/2}$.

It is not very difficult to further extend the models to allow for exogenous variables in dynamics. For example, the ADCCE and ADCCX models proposed by Li (2011, 2015). The AG-DCCX generalization is as follows:

$$Q_t = (S - A'SA - B'SB) + A'\epsilon_{t-1}\epsilon'_{t-1}A + B'Q_{t-1}B + \Gamma'[\eta_{t-1}\eta'_{t-1} - E(\eta_t\eta'_t)]\Gamma + \delta I_{k\times k}x_{t-1} \quad (14)$$

where δ measures the impact of an exogenous variable x_t on the DCC, $I_{k\times k}$ is the matrix that all elements are 1.

3.2. Estimation of DCCs

If $\Sigma_t = D_tR_tD_t$, $a_t|\mathcal{F}_{t-1} \sim N(0,\Sigma_t)$, then the joint quasi-log-likelihood function can be written as

$$QL = -\frac{1}{2}\sum_{t=1}^{T}[klog(2\pi) + log|\Sigma_t| + a'_t\Sigma_t^{-1}a_t]$$

$$= -\frac{1}{2}\sum_{t=1}^{T}[klog(2\pi) + log|D_tR_tD_t| + a'_tD_t^{-1}R_t^{-1}D_t^{-1}a_t]$$

$$= -\frac{1}{2}\sum_{t=1}^{T}[klog(2\pi) + 2log|D_t| + log|R_t| + \epsilon'_tR_t^{-1}\epsilon_t]$$

$$= -\frac{1}{2}\sum_{t=1}^{T}[klog(2\pi) + 2log|D_t| + (a'_tD_t^{-2}a_t - \epsilon'_t\epsilon_t) + log|R_t| + \epsilon'_tR_t^{-1}\epsilon_t]$$

$$= -\frac{1}{2}\sum_{t=1}^{T}[klog(2\pi) + 2log|D_t| + a'_tD_t^{-2}a_t] - \frac{1}{2}\sum_{t=1}^{T}[-\epsilon'_t\epsilon_t + log|R_t| + \epsilon'_tR_t^{-1}\epsilon_t]$$

$$\triangleq QL_1(\boldsymbol{\theta}) + QL_2(\boldsymbol{\theta}, \boldsymbol{\phi}, S, N) \quad (15)$$

where $\boldsymbol{\phi} = (A, B, \Gamma, \delta)$, $N = E(\eta_t\eta'_t)$.

Hence, the joint quasi-log-likelihood function is split into two parts and maximized sequentially. First, the parameters of the univariate volatility process are estimated by Eq. (9), and the estimate of $\boldsymbol{\theta}$ is given by

$$\widehat{\boldsymbol{\theta}} = arg\max_{\boldsymbol{\theta}} QL_1(\boldsymbol{\theta}) \quad (16)$$

Once $\widehat{\boldsymbol{\theta}}$ is obtained, S and N are replaced with sample analogues, $\hat{S} = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_t\hat{\epsilon}'_t$ and $\widehat{N} = \frac{1}{T}\sum_{t=1}^{T}\hat{\eta}_t\hat{\eta}'_t$, respectively, where $\hat{\epsilon}_t = D_t^{-1}(\widehat{\boldsymbol{\theta}})a_t$. Finally, the parameters of the correlation process are estimated by

$$QL_2(\widehat{\boldsymbol{\theta}}, \boldsymbol{\phi}, \hat{S}, \widehat{N}) = -\frac{1}{2}\sum_{t=1}^{T}[-\epsilon'_t\epsilon_t + log|R_t| + \epsilon'_tR_t^{-1}\epsilon_t] \quad (17)$$

and the estimate of $\boldsymbol{\phi}$ is given by

$$\widehat{\boldsymbol{\phi}} = arg\max_{\boldsymbol{\phi}} QL_2(\widehat{\boldsymbol{\theta}}, \boldsymbol{\phi}, \hat{S}, \widehat{N}) \quad (18)$$

The three-stage estimation technique was called ''correlation targeting'' in Engle et al. (2009) rather than its revised version (Pakel et al., 2020), since neither S (which is assumed to be the second moment of ϵ_t, i.e., $S = E(\epsilon_t\epsilon'_t)$) nor N is estimated by the quasi-maximum-likelihood method but are replaced respectively by the moment estimators $\hat{S}$ and $\widehat{N}$, which aim to match the sample covariance matrix of ϵ_t with the unconditional expectation of Q_t. In the DCC, however, Aielli (2013) showed that $\hat{S}$ is both biased and inconsistent for S, due to the fact that $E(\epsilon_t\epsilon'_t) \neq E(Q_t)$. The DCC driving process, Q_t, is usually regarded as a linear process (e.g., Eqs. (11) and (13)), and the roles of the parameters a, b and S are then explained accordingly. In fact, the conditional covariance matrix of ϵ_t is not Q_t but R_t, so Q_t is not a linear process. A serious consequence is that the traditional interpretation of a, b and S may lead to misleading conclusions.

4. ESTIMATION OF DCC-TYPE MODELS: ELEMENT VERSION

In order to understand the dynamic process at a glance, we expand the matrices in these DCCs into the elements. At the same time, for concision, we only investigate the two-dimensional case.

Let $a_t = (a_{1,t}, a_{2,t})'$ for $t = 1, 2, \cdots, T$ with $E(a_t|\mathcal{F}_{t-1}) = 0$ and $Cov(a_t|\mathcal{F}_{t-1}) = \begin{bmatrix} \sigma_{11,t}^2 & \sigma_{12,t}^2 \\ \sigma_{21,t}^2 & \sigma_{22,t}^2 \end{bmatrix}$, and let $\epsilon_{i,t} = a_{i,t}/\sigma_{ii,t}$ for $i = 1, 2$ with the conditional correlation

$$\rho_{12,t} = \sigma_{12,t}^2/\sigma_{11,t}\sigma_{22,t} \quad (19)$$

between $a_{1,t}$ and $a_{2,t}$. Clearly, $\rho_{12,t} = Cov(\epsilon_{1,t}, \epsilon_{2,t}|\mathcal{F}_{t-1}) = \rho_{21,t}$.

4.1. Specification of DCCs

Let's first describe the element versions of these DCC-type models, including scale and diagonal cases.

DCC model:

$$\rho_{12,t} = q_{12,t}/\sqrt{q_{11,t}q_{22,t}}$$

$$q_{ij,t} = (1 - a - b)S_{ij} + a\epsilon_{i,t-1}\epsilon_{j,t-1} + bq_{ij,t-1} \quad (20)$$

where S_{ij} is often assumed to be the unconditional correlation between $\epsilon_{i,t}$ and $\epsilon_{j,t}$ ($S_{11} = S_{22} = 1$).

The positivity constraints: $a + b < 1$, $a \geq 0$ and $b \geq 0$.

A-DCC model:

$$\rho_{12,t} = q_{12,t}/\sqrt{q_{11,t}q_{22,t}}$$

$$q_{ij,t} = (1 - a - b)S_{ij} + a\epsilon_{i,t-1}\epsilon_{j,t-1} + bq_{ij,t-1} + \gamma[\eta_{i,t-1}\eta_{j,t-1} - E(\eta_{i,t}\eta_{j,t})] \quad (21)$$

where $\eta_{i,t} = I_{\epsilon_{i,t}<0} \times \epsilon_{i,t}$.

The positivity constraints: $a + b + \lambda\gamma < 1$, $a \geq 0$, $b \geq 0$, $\gamma \geq 0$, λ is the maximum eigenvalue of matrix $[S_{ij}]^{-\frac{1}{2}}[Cov(\eta_{i,t}, \eta_{j,t})][S_{ij}]^{-\frac{1}{2}}$.

In these diagonal cases below, let $A = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}$, $B = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$ and $\Gamma = \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix}$.

D-DCC model:

$$\rho_{12,t} = q_{12,t}/\sqrt{q_{11,t}q_{22,t}}$$

$$q_{ij,t} = (1 - a_i a_j - b_i b_j)S_{ij} + a_i a_j \epsilon_{i,t-1}\epsilon_{j,t-1} + b_i b_j q_{ij,t-1} \quad (22)$$

The positivity constraints:

$(1 - a_1^2 - b_1^2)(1 - a_2^2 - b_2^2) \geq (1 - a_1 a_2 - b_1 b_2)^2$, $1 - a_1^2 - b_1^2 \geq 0$ and $1 - a_2^2 - b_2^2 \geq 0$.

AD-DCC model:

$$\rho_{12,t} = q_{12,t}/\sqrt{q_{11,t}q_{22,t}}$$

$$q_{ij,t} = (1 - a_i a_j - b_i b_j)S_{ij} + a_i a_j \epsilon_{i,t-1}\epsilon_{j,t-1} + b_i b_j q_{ij,t-1} + \gamma_i\gamma_j[\eta_{i,t-1}\eta_{j,t-1} - E(\eta_{i,t}\eta_{j,t})] \quad (23)$$

The positivity constraints:

$1 - a_1^2 - b_1^2 \geq \gamma_1^2$, $1 - a_2^2 - b_2^2 \geq \gamma_2^2$, and

$(1 - a_1^2 - b_1^2 - \gamma_1^2)(1 - a_2^2 - b_2^2 - \gamma_2^2) \geq [(1 - a_1 a_2 - b_1 b_2)S_{12} - \gamma_1\gamma_2 E(\eta_{1,t}\eta_{2,t})]^2$.

Moreover, the one-lag variable x_{t-1} can be embedded into Eqs. (20)-(23) to investigate the effect of an exogenous variable (x_t) on these conditional correlations.

4.2. Estimation of DCCs

Next, we describe the element versions of the three-stage estimation for DCCs.

If $\epsilon_{i,t} = a_{i,t}/\sigma_{ii,t} \sim i.i.d.N(0,1)$ for $i = 1, 2$, then the volatility part $QL_1(\boldsymbol{\theta})$ in the joint quasi-log-likelihood function, Eq. (15), is obviously the sum of individual GARCH-likelihood functions

$$QL_1(\theta_1, \theta_2) = -\frac{1}{2}\sum_{i=1}^{2}\sum_{t=1}^{T}[log(2\pi) + log(\sigma_{ii,t}^2) + \frac{a_{i,t}^2}{\sigma_{ii,t}^2}] \quad (24)$$

which is jointly maximized by separately maximizing each term

$$QL_{1,i}(\theta_i) = -\frac{1}{2}\sum_{t=1}^{T}[log(2\pi) + log(\sigma_{ii,t}^2) + \frac{a_{i,t}^2}{\sigma_{ii,t}^2}] \quad (25)$$

where θ_i is composed of these parameters in the volatility model followed by $a_{i,t}$ for $i = 1, 2$.

Once $\hat{\theta}_i = arg\max_{\theta_i} QL_{1,i}(\theta_i)$ is obtained, $S_{12} = E(\epsilon_{1,t}\epsilon_{2,t})$ and $N_{ij} = E(\eta_{i,t}\eta_{j,t})$ are replaced with the moment estimators, $\hat{S}_{12} = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_{1,t}\hat{\epsilon}_{2,t}$ and $\hat{N}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\hat{\eta}_{i,t}\hat{\eta}_{j,t}$, respectively. Thus, these parameters $\boldsymbol{\phi}$ (e.g., a_i, b_i, γ_i) of the correlation process are estimated by

$$QL_2(\hat{\theta}_1, \hat{\theta}_2; \boldsymbol{\phi}; \hat{S}_{12}, \hat{N}_{ij}) = -\frac{1}{2}\sum_{t=1}^{T}[log(1 - \rho_{12,t}^2) + \frac{\rho_{12,t}^2}{1-\rho_{12,t}^2}\frac{a_{1,t}^2}{\sigma_{11,t}^2} - \frac{2\rho_{12,t}}{1-\rho_{12,t}^2}\frac{a_{1,t}a_{2,t}}{\sigma_{11,t}\sigma_{22,t}} + \frac{\rho_{12,t}^2}{1-\rho_{12,t}^2}\frac{a_{2,t}^2}{\sigma_{22,t}^2}] \quad (26)$$

All the above quasi-log-likelihood functions are usually solved by the BHHH algorithm, which uses the gradient information of the objective function to iterate and optimize. However, the gradient calculation is very complex and does not necessarily exist, due to more parameters. At the same time, there may be some oscillations in the search, leading to convergence to the local optimal solution, and the results are sensitive to the selection of the initial value. Some statistical software packages, such as OxMetrics and R, contain the programs for DCC and A-DCC. So far, these programs for D-DCC and AD-DCC have not been discovered, and thus we need to program them ourselves.

5. CONCLUSIONS AND COMMENTS

In the above DCCs, the conditional variances follow univariate GARCH models. The conditional correlation is then modeled as a special function of the past GARCH standardized residuals. In its original intention, such a modeling should be able to provide two advantages. Firstly, due to the modular structure of conditional covariance matrix Σ_t, a consistent three-stage estimator for large systems should be easily obtained. Secondly, due to the explicit parameterization of conditional correlation process, the examination of some correlation hypotheses, such as whether the correlation process is integrated or not, should be more direct than with other data-driven volatility models.

The DCCs, however, are less tractable than expected. First of all, it is generally believed that they do not perform well in high-dimensional cases (there exists the dimension curse), as the parameter estimates encounter serious negative biases, resulting in the correlation trajectories being smoother and smoother, and eventually becoming almost flat and constant. Pakel et al. (2020) suggested a composite likelihood (CL) estimator based on the summation of the quasi-likelihood functions of innovation subsets and thus to avoid the operation of high-dimensional matrix. This method, worth trying in the future, allows people to estimate the model even when the cross-section size is larger than the sample one. Secondly, the conjecture on the consistency of the second stage estimator, in which the correlation intercept are estimated, has been proved to be untenable. As a result, the third stage estimator, where the dynamic correlation parameters are estimated, may be inconsistent in turn. This issue motivated Aielli (2013) to correct the DCC, and introduce a consistent DCC (cDCC) model, in which the Q_t process is a slightly different form that in Eq. (11):

$$Q_t = (1 - a - b)K + a[diag(Q_{t-1})^{-\frac{1}{2}}\epsilon_{t-1}\epsilon'_{t-1}diag(Q_{t-1})^{-\frac{1}{2}}] + bQ_{t-1} \quad (27)$$

where K is a correlation parameter matrix. Note that the corrected standardized residual $diag(Q_t)^{-\frac{1}{2}}\epsilon_t$ does not depend on K since $diag(Q_t)$ is only associated with the diagonal elements of K, which are all equal to 1. The three-stage estimation technique, the same quasi-log-likelihood function as the DCC, is feasible for the cDCC with large systems since K can be estimated consistently by the sample covariance matrix of $diag(Q_t)^{-\frac{1}{2}}\epsilon_t$, i.e., $\widehat{K} = \frac{1}{T}\sum_{t=1}^{T} diag(Q_t)^{-\frac{1}{2}}\hat{\epsilon}_t diag(Q_t)^{-\frac{1}{2}}\hat{\epsilon}'_t$. In fact,

$$E\left(diag(Q_t)^{-\frac{1}{2}}\epsilon_t\epsilon'_t diag(Q_t)^{-\frac{1}{2}}\right) = K,$$

$$Cov(diag(Q_t)^{-\frac{1}{2}}\epsilon_t|\mathcal{F}_{t-1}) = Q_t.$$

Regarding the performance of the three-stage estimators, Aielli (2013) revealed that the cDCC correlation forecasts perform equally or significantly better than the DCC correlation forecasts for typical financial series. Accordingly, it is worth exploring to extend the correction in the cDCC to other DCC-type models.

This summary is not only of great significance to promote “teaching and learning”, but also very helpful to investors in asset allocation and risk management.

ACKNOWLEDGMENTS

The author thanks the Project of Humanities and Social Sciences, Ministry of Education in China (No. 20YJAZH080) and the Project of Philosophy and Social Sciences for the Universities in Jiangsu Province (No. 2019SJA0258) for their support.

REFERENCES

[1] G.P. Aielli, Consistent estimation of large scale dynamic conditional correlations. Journal of Business and Economic Statistics, 31, 2013, pp. 282-299.

[2] L. Bauwens, S. Laurent, J. Rombouts, Multivariate GARCH models: A survey. Journal of Applied Econometrics, 21, 2006, pp. 79-109.

[3] L. Bauwens, L. Grigoryeva, J.P. Ortega, Estimation and empirical performance of non-scalar dynamic conditional correlation models. Computational Statistics and Data Analysis, 100, 2016, pp. 17-36.

[4] T. Bollerslev, Modeling the coherence in short run nominal exchange rates: A multivariate generalized ARCH model. Review of Economics and Statistics, 72, 1990, pp. 498-505.

[5] L. Cappiello, R.F. Engle, K. Sheppard, Asymmetric dynamics in the correlations of global equity and bond returns. Journal of Financial Econometrics, 4(4), 2006, pp. 537-572.

[6] R.F. Engle, Dynamic conditional correlation: A simple class of multivariate GARCH models. Journal of Business and Economic Statistics, 20, 2002, pp. 339-350.

[7] R.F. Engle, K. Sheppard, Theoretical and empirical properties of dynamic conditional correlation MVGARCH. Working Paper, 2001, No. 2001-15, University of California, San Diego.

[8] R.F. Engle, K. Sheppard, Evaluating the specification of covariance models for large portfolios. Working Paper, 2008, University of Oxford.

[9] R.F. Engle, N. Shephard, K. Sheppard, Fitting vast dimensional time-varying covariance models. Working Paper, 2009, No. FIN-08-009, New York University.

[10] C.M., Hafner, O. Reznikova, On the estimation of dynamic conditional correlation models. Computational Statistics and Data Analysis, 56, 2012, pp. 3533-3545.

[11] T. Jeantheau, Strong consistency of estimators for multivariate ARCH models. Econometric Theory, 14, 1998, pp. 70-86.

[12] W.K. Li, S.Q. Ling, M. McAleer, Recent theoretical results for time series models with GARCH errors. Journal of Economic Surveys, 16(3), 2002, pp. 245-269.

[13] X.M. Li, How do exchange rates co-move? A study on the currencies of five inflation-targeting countries. Journal of Banking & Finance, 35, 2011, pp. 418-429.

[14] X.M. Li, B. Zhang, R.Z. Gao, Economic policy uncertainty shocks and stock-bond correlations: Evidence from the US market, Economics Letters, 132, 2015, pp. 91-96.

[15] C. Pakel, N. Shephard, K. Sheppard, R.F. Engle, Fitting vast dimensional time-varying covariance models. Journal of Business & Economic Statistics, 2020, Doi: 10.1080/07350015.2020.1713795.

[16] A. Silvennoinen, T. Teräsvirta, Multivariate GARCH Models, in Handbook of Financial Time Series. Andersen, T. G. Davis, R. A., Kreiβ, J. P., Mikosch, T., Berlin, Heidelberg: Springer-Verlag, 2008, pp. 201-229.