Barcelona Graduate School of Economics

Efficient Estimation of Dynamic Conditional Correlation Models

Euan Dowers

June 24, 2017

Contents

1 Introduction

The aim of this paper is to efficiently estimate the Dynamic Conditional Correlation model in large dimensions (both in time and the number of assets included in the model), using two innovations. The first, due to Engle, Ledoit and Wolf [?], is to use the linear shrinkage method of Ledoit and Wolf [?] to estimate the correlation targeting matrix of the DCC model, and the second, which is original to this paper, is to calculate the log-likelihood of the DCC model using a series of rank-one updates to the cholesky decomposition of the quasi-correlation matrix at each time period, thus avoiding having to refactor this matrix at every time period.

The structure of this report is as follows: Section 2 will describe the Dynamic Conditional Correlation model, including the estimation problem; Section 3 will describe the linear shrinkage estimation for covariance matrices of Ledoit and Wolf.

2 Dynamic Conditional Correlation Model

In order to describe the Dynamic Conditional Correlation Model, it is necessary to build up some notation.

1. Let $r_{i,t}$ denote the return of asset i at time t, and r_t be the N-dimensional vector of all returns at time t.
2. Let $d_{i,t}^2 = V_{t-1}(r_{i,t}|\mathcal{F}_{t-1})$ be the conditional variance of asset i at time t, given information up to time $t-1$.
3. Let D_t be a diagonal matrix such that $D_{i,i,t} = d_{it}$.
4. Let H_t denote the conditional covariance matrix such that $H_{i,i,t} = \text{cov}_{t-1}(r_{i,t}, r_{j,t}|\mathcal{F}_{t-1})$.
5. Let $\epsilon_{i,t} = \frac{r_{i,t}}{d_{i,t}}$ be the standardised residuals of asset i at time t.
6. Let R_t be the conditional correlation matrix, whose i,j-th entry is given by the conditional correlation between asset i and asset j at time t.
7. Let σ_i^2 be the unconditional variance of the series r_i, t
8. Let R be the unconditional correlation matrix of the system.

2.1 Dynamic Conditional Correlation Model

Assume we have N financial assets, and we observe for each of these assets a series of returns over a time period $1, \ldots, T$.

Since the conditional correlation and conditional covariance are related by the equation

$$\rho_{i,j,t} = \frac{E_{t-1}((r_{i,t} - E_{t-1}(r_{i,t}))(r_{j,t} - E_{t-1}(r_{j,t})))}{\sqrt{V_{t-1}(r_{i,t})V_{t-1}(y_{j,t})}}, \tag{2.1}$$

these entries are therefore also given by

$$\frac{H_{i,j,t}}{\sqrt{H_{i,i,t}H_{j,j,t}}}. \tag{2.2}$$

Therefore, the conditional correlation matrix and conditional variance matrix are given by

$$R_t = D_t^{-1} H_t D_t^{-1} \quad D_t^2 = diag[H_t] \tag{2.3}$$

In the mean-reverting DCC model, the dynamics of the quasi-correlation matrix Q_t are governed by the process

$$Q_t = \Omega + \alpha \epsilon_{t-1}\epsilon'_{t-1} + \beta Q_{t-1}. \tag{2.4}$$

In the correlation targeting version of this model, the intercept matrix Ω is given

$$\Omega = (1 - \alpha - \beta) R \tag{2.5}$$

The matrix Q_t does not define a correlation matrix, as the diagonal elements are not necessarily 1. To convert the matrix Q_t into a correlation matrix we just divide the i, j-th entry by $\sqrt{Q_{i,i,t}Q_{j,j,t}}$, in other words

$$R_t = \text{diag}\{Q_t\}^{-\frac{1}{2}} Q_t \text{diag}\{Q_t\}^{-\frac{1}{2}} \tag{2.6}$$

We will discuss in Section 4 the difficulty of estimating the parameters of this model when the number of assets considered, N, is large.

3 Shrinkage Estimation for Large Unconditional Correlation Matrices

As seen in Equation refeq: correlation targeting, a key part of fitting the DCC model is estimating the unconditional correlation matrix R. We know from the cross-sectional literature that using the sample covariance matrix (Equation 3.1) works poorly in large dimensions, since for large N, with $N \approx T$, the sample covariance marix overfits to the data, and provides a good in-sample fit, but not out-of-sample.

$$\hat{\Sigma} = \frac{1}{T} \sum \epsilon_t \epsilon'_t \tag{3.1}$$

As the concentration ratio N/T becomes nearer to 1 the sample covariance matrix contains a lot of estimation error, and is ill-conditioned, and in the case of financial data this is the case: we want to analyse the covariance structure of as many stocks as possible.

In fact, Ledoit and Wolf [?] show that the eigenvalues of the sample covariance matrix are further from their mean than the true eigenvalues, so they suggest an estimator of the form of Equation 3.2, where λ_i is the i-th largest eigenvalue, and u_i is its corresponding eigenvector.

$$\bar{C} = \sum_{i=1}^{N} \left(\rho \bar{\lambda} + (1 - \rho)\lambda_i \right) u_i u'_i \tag{3.2}$$

Ledoit and Wolf also show [?] that using a non-linear version of shrinkage estimation works better, but employing this technique is beyond the scope of this paper, and is an opportunity for further research. The main difference being that in nonlinear shrinkage the shrinkage parameter ρ is determined for each eigenvector individually rather than being forced to be the same for each eigenvector.

4 DCC Model Estimation

In Section

The problem of estimating the DCC model can be formulated as a maximum likelihood problem. In the DCC model where the conditional distribution of r_t

$$l = -\frac{1}{2}\sum_{t=1}^{T} n\log(2\pi) + 2\log|D_t| + y_t' D_t^2 y_t - \epsilon_t'\epsilon_t + \log|R_t| + \epsilon_t' R_t^{-1}\epsilon_t \tag{4.1}$$

This log-likelihood can be split into two parts, the first part contains only the variane parameters from the GARCH specification in Equation 4.2, and the second contains only the DCC parameters, α and β. This log-likelihood can therefore be maximised in two stages, the first maximising with respect to the variance parameters $\omega_i, \alpha_i, \beta_i$, and the second stage maximising with respect to α, β. The first stage is done by fitting a univariate GARCH(1,1) model to each asset independently.

To model the quantities $d_{i,t}$, and from this extract the standardised residuals ϵ_t we apply the univariate GARCH(1,1) model to each asset serparately to obtain the dynamics of its conditional variance.

$$H_{i,i,t} = \omega_i + \alpha_i r_{t-1}^2 + \beta_i H_{i,i,t-1}. \tag{4.2}$$

Then, the standardised residuals $\epsilon_{i,t}$ are given by

$$\epsilon_{i,t} = \frac{r_{i,t}}{d_{i,t}} \tag{4.3}$$

For the second stage, after fitting our univariate, independent GARCH(1,1) models, we extract the standardised residuals $\epsilon_{i,t}$ and use these to calculate the second part of the log-likelihood function

$$l_2 = -\frac{1}{2}\sum_{t=1}^{T} \log|R_t| + \epsilon_t' R_t^{-1}\epsilon_t. \tag{4.4}$$

To clarify how this relates to the specification of the mean-reverting DCC model in terms of α and β, let us remind ourselves of the dynamics of the quasi-correlation matrix Q_t, and how Q_t relates to the conditional correlation matrix R_t.

$$Q_t = \Omega + \alpha\epsilon_{t-1}\epsilon_{t-1}' + \beta Q_{t-1}, \tag{4.5}$$

$$\text{and } R_t = \text{diag}\{Q_t\}^{-\frac{1}{2}} Q_t \text{diag}\{Q_t\}^{-\frac{1}{2}}. \tag{4.6}$$

Therefore, in order to calculate the log-likelihood of our observed data, we need to, at every time interval t, invert the matrix R_t. Therefore, if T or N is large, directly estimating the log-likelihood in a naive way (by calculating the inverse of the matrix R_t for each $t \in \{1, \ldots, T\}$) becomes computationally infeasible.

4.1 Cholesky Decomposition

The first thing to note is that we do not ever calculat the inverse of the matrix R_t. Rather we use the *Cholesky decomposition*, where, for a positive definite matrix A, $A = LL'$, where L is a lower triangular matrix.

$$L = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{bmatrix}. \tag{4.7}$$

This allows us to solve the system of linear equations $Ax = b$ (and thus $x = A^{-1}b$) using forward and back-substitution, without having to calculate A^{-1}. In particular, we want to, at each time interval t, calculate the quantity $R_t^{-1}\epsilon_t$. So, given the Cholesky decomposition of the matrix R_t, $\tilde{L}_t$ we can compute $x_t = R_t^{-1}\epsilon$ and therefore $= \epsilon_t' x_t = \epsilon_t' R_t^{-1}\epsilon_t$, the second part of l_2. Therefore, if we can come up with an efficient way of updating the cholesky decomposition of R_t at each time interval t, then we can improve on the

4.2 Updating the Cholesky Decomposition

Fortunately, there are algorithms for updating the cholesky decomposition of a matrix A when A is modified by

$$\tilde{A} = A + x'x \tag{4.8}$$

where x is a vector of length N. Such an update is called a rank one update, and the algorithm implemented for finding the Cholesky decomposition of $\tilde{A}$ is given by Algorithm 1

Algorithm 1 Rank One Update to Cholesky decomposition

procedure RankOneUp
  for $i \in 1 : N$ **do**
    $r \leftarrow \sqrt{L_{ii}^2 + x_i^2}$
    $c \leftarrow r/L_{ii}$
    $s \leftarrow x_i/L_{ii}$
    for $j \in i+1 : N$ **do**
      $L_{ji} \leftarrow \frac{L_{ji} + s x_j}{c}$
      $x_j \leftarrow c x_j - s L_{ji}$

Noting that if we take x to be $\sqrt{\frac{\alpha}{\beta}}\epsilon_t$, this gives us an algorithm for finding the Cholesky decomposition of $Q_{t-1} + \frac{\alpha}{\beta}\epsilon_t\epsilon_t'$. Now note that if we multiply the resulting decomposition $\tilde{L}$ by $\sqrt{\beta}$ we get the Cholesky decomposition of $\alpha\epsilon_t'\epsilon_t + \beta Q_{t-1}$. Therefore, this algorithm has given us an efficient way to update the Cholesky decomposition of the quasi-correlation matrix Q_t. However, the problem occurs when we add the full rank correlation target matrix Ω. There is no way of updating the Cholesky decomposition upon adding a full-rank matrix that is any more efficient than refactoring the matrix Q_t at every time interval t.

The innovation of this paper is to approximate each update by

5 Monte Carlo Simulation

6 Conclusion

References

[1] Engle, R. *Dynamic Conditional Correlation - a simple class of multivariate models.* Journal of Business and Economic Statistics, 2002.

[2] Bollerslev, D *Dynamic Conditional Correlation - a simple class of multivariate models.* Journal of Business and Economic Statistics, 1990.

[3] Engle, R. *Anticipating Correlations: a new paradigm for risk management*