AMSI VACATION RESEARCH SCHOLARSHIPS 2018-2019

Predicting Behaviour of Financial Systems based on a Multivariate GARCH Perspective and Dynamical Network

Shidan Liu

Supervised by Dr Laleh Tafakori

RMIT University

Vacation Research Scholarships are funded jointly by the Department of Education and Training and the Australian Mathematical Sciences Institute.

AMSI Australian Mathematical Sciences Institute

Australian Government Department of Education and Training

Abstract

The development of models for the dependency structure of risk among financial entities is of great interest to risk management. This project aims to develop a new tool based on existing tools for the identification of dependence relations among various parties in a financial network that will enable understanding systemic risk in our current financial systems. We also endeavour to improve the model performance in large dimensions.

1 Introduction

Owing to several economic shocks such as Asian crisis in 1997, European currency in 1999, bankruptcy of WolrdCom in 2002 and global financial crisis (GFC) in 2008, there has been a strong interest in the development of accurate models for the dependency structure of default risk among financial firms. As evidence in this concern, risk management measures need to be at hand in order to prepare for reducing the effect of extreme events. The aim of this project is to develop fundamental theories and tools for the identification of dependence (influence) relations among various parties in a financial network that will enable understanding systemic risk in our current financial systems. Thereby, by identifying the changes and differences of links and dependency structure in the complex financial networks, portfolio and financial managers can set up a strategy to avoid or reduce the impact of the extreme events (economic shocks such as financial crisis) in the financial market. In fact, it gives a dynamic vision to managers to identify and evaluate the consistent and fragile entities. The project will be roughly divided into three parts.

The first will be introducing of a Multivariate GARCH model derived from the GARCH family started by Engle (1982). Indeed, these models are popular tools for risk management and portfolio selection, but they suffer from the curse of dimensionality. In other words, when the number of assets in the investment universe is large, many multivariate GARCH models exhibit unsatisfactory performance or cannot even be estimated in the first place due to computational problems. Therefore in this project we will use the dynamic conditional correlation (DCC) model originally proposed by Engle (2002) against large dimensions to improve a spatial approach for modeling risk spillovers using financial time-varying proximity matrices based on observable networks.

The second part of the project will be using the methodology to combine two tools. The first tool is the composite likelihood method of Pakel et al. (2014), which makes the estimation of a DCC model in large dimensions computationally feasible and the nonlinear shrinkage method which results in improved estimation of the correlation targeting matrix of a DCC model. Further, using this approach can lead to an improvement to the BEKK model presented in Engle and Kroner (1995).

Finally we would look at the spatial approach for modeling risk spillovers using financial time-varying proximity matrices based on observable networks and the approach which is mentioned in the second part. We show how these methods could be useful in investigating the role of portfolio composition in risk transfer, and its ability to reduce the forecasted system variance and thus the risk of the system.

2 The DCC Model

2.1 Notation

In this report, the subscript i indexes variables/assets and ranges from 1 to N, where N is the dimension of the covariance matrix. The subscript t indexes the dates and ranges from 1 to T, where T is the sample size. Diag(·) sets all off-diagonal elements of a matrix to zero.

- $r_{i,t}$: observed data series for variable/asset i
- $d_{i,t}^2 := var(r_{i,t}|\mathcal{F}_{t-1})$: conditional variance of the ith variable/asset at date t
- D_t is the N-dimensional matrix whose ith diagonal element is $d_{i,t}$
- $H_t := cov(\boldsymbol{r}_t|\mathcal{F}_{t-1})$: conditional covariance matrix at date t
- $s_{i,t} := r_{i,t}/d_{i,t}$: devolatilized series
- $R_t := Corr(\boldsymbol{r}_t|\mathcal{F}_{t-1}) = cov(\boldsymbol{s}_t|\mathcal{F}_{t-1})$: conditional correlation matrix at date t
- $C := Corr(\boldsymbol{r}_t) = cov(\boldsymbol{s}_t)$

2.2 Model Definition

One of the challenges for modelling financial time series is heteroskedasticity. ARCH (auto-regressive conditional heteroskedasticity) models were introduced by Robert Engle in 1982 to account for this behaviour. Tim Bollerslev (1986) extended the ARCH model to GARCH(p,q) (generalized ARCH) model. The standard GARCH(1,1) specification will be used for simplicity as it produces good enough results, and is given by:

$$d_{i,t}^2 = \omega_i + a_i r_{i,t-1}^2 + b_i d_{i,t-1}^2 \,,$$

The DCC (Dynamic Conditional Covariance) model with correlation targeting introduced by Engle et al. (2017) is expressed as:

$$Q_t = (1-\alpha-\beta)C + \alpha \boldsymbol{s}_{t-1}\boldsymbol{s'}_{t-1} + \beta Q_{t-1}\,, \tag{1}$$

$$R_t := Diag(Q_t)^{-1/2} Q_t Diag(Q_t)^{-1/2}\,,$$

$$H_t := D_t R_t D_t\,.$$

where (α, β) are the DCC parameters, Q_t can be interpreted as a conditional pseudo-correlation matrix.

2.3 Model Estimation

The estimation will be carried out in three stages:

1. For each variable/asset, fit a univariate GARCH(1,1) model and use the fitted model to devolatilize the return series.

2. Estimate the unconditional correlation matrix and use it for correlation targeting.

3. Maximize the composite likelihood to estimate correlation dynamics.

Engle et al.(2017) proposed to use the maximum composite likelihood method to overcome the computational hurdle in large dimensions, and the composite log-likelihood is computed by summing up the log-likelihood of contiguous pairs of variables/assets. They also introduced a nonlinear shrinkage method for the estimation of the unconditional correlation matrix C, constructed as:

$$\widetilde{C} := \sum_{i=1}^{N} \widetilde{\lambda}_i(\widetilde{\boldsymbol{\tau}}) \mu_i \mu_i'\,.$$

where $\widetilde{\boldsymbol{\lambda}}(\widetilde{\boldsymbol{\tau}})$ is the shrunk eigenvalues, computed using Theorem 4 of Ledoit and Pch (2011) given the population eigenvalues $\widetilde{\boldsymbol{\tau}}$ that are estimated from the sample eigenvalues by numerically inverting the QuEST function of Ledoit and Wolf (2015).

3 The Spatial BEKK Model

3.1 Model Definition

The BEKK model is another popular model for the estimation of conditional covariance matrices, and the simplest BEKK model is given by:

$$\Sigma_t = CC' + A\mu_{t-1}\mu_{t-1}'A' + B\Sigma_{t-1}B'\,.$$

The full BEKK model is computationally expensive as it has $(2n^2+0.5n(n+1))$ parameters to estimate. For this reason, A and B are usually restricted to be scalar or diagonal. However, in this case the risk spillover and variance feedback effect in the model are excluded or limited. To overcome this limitation, Caporin and Paruolo (2015) introduced the spatial BEKK GARCH model where A and B are replaced by proximity matrices:

$$\Sigma_t = CC' + A_M(W)\mu_{t-1}\mu_{t-1}'A_M(W)' + B_M(W)\Sigma_t B_M(W)'\,.$$

where $M = L, R$, indicating left or right multiplication, and W is the weight matrix, representing the financial relationship between variables/assets.

If we look at the left multiplication case and focus on the ARCH part (matrix A) of the model, and note that:

$$\begin{aligned} \upsilon_{L,t} &= A_L(W)\mu_t = (A_{0,L} + A_{1,L}W)\mu_t = \upsilon_{L,t}^0 + \upsilon_{L,t}^1\,, \\ \upsilon_{L,t}^0 &= [A_{0,L}\mu_t]_i = a_{0,L,i}\mu_{i,t}\,, \\ \upsilon_{L,t}^1 &= [A_{1,L}W\mu_t]_i = a_{1,L,i}\sum_{j=1}^{n}\omega_{i,j}\mu_j\,. \end{aligned} \tag{2}$$

As discussed in Monica et al. (2015), the ith element of $\upsilon_{L,t}$ is dependant on its own past shock, weighted by $a_{0,L,j}$ (direct effect), and on the past shocks of its neighbours, weighted by the distance, all loaded with the same coefficient, $a_{1,L,j}$ (indirect effect).

As shown in equation (2), the left multiplication specification allows us to investigate which are the risk receivers in a network, and the model parameters stacked into the vector $a_{1,L}$ monitors the reaction of risk recipients to shock originated from the neighbors.

In the case of right multiplication and focusing again on the ARCH part (A matrix) of the model, we have:

$$\begin{aligned} \upsilon_{R,t} &= A_R(W)\mu_t = (A_{0,R} + A_{1,R}W)\mu_t = \upsilon_{R,t}^0 + \upsilon_{R,t}^1\,, \\ \upsilon_{R,t}^0 &= [A_{0,R}\mu_t]_i = a_{0,R,i}\mu_{i,t}\,, \\ \upsilon_{R,t}^1 &= [A_{1,R}W\mu_t]_i = \sum_{j=1}^{n}\omega_{i,j}a_{1,R,j}\mu_j\,. \end{aligned} \tag{3}$$

From equation (3), we see that the parameters magnify the effect of the source of risk, allowing us to investigate risk spreaders.

Note that in Monica et al. (2015), the weight matrix W is time-varying that can be varying on a lower time scale than that of the evolution of the entities in the system, indicating a dynamical network. In this report, we will assume the weight matrix W to be constant, and will investigate in its time-varying property in future study.

3.2 Model Estimation

We will endeavour to adopt the nonlinear shrinkage method in the estimation of the unconditional covariance matrix target C, and the composite likelihood method in the estimation of the model parameters. If we denote by $\theta \equiv (a_{0,M}, a_{1,M}, b_{0,M}, b_{1,M})$ the vector of model parameters, the composite log-likelihood is:

$$\ell(\theta) = \sum_{i=1}^{N-1} -\frac{1}{2}(2log(2\pi) + log(det(\Sigma_t)) + \mu_t(\Sigma_t)^{-1}\mu_t')/(N-1)\,.$$

According to Caporin and Paruolo (2015), a simple identification restriction consists of constraining the first element of $a_{0,M}$, $a_{1,M}$, $b_{0,M}$, $b_{1,M}$, to be positive, which allows the coefficients to be either positive or negative.

To ensure the ergodicity and stationarity of the process, the maximum spectral radius is constrained in the following way:

$$\max_{t\in[1,T]} \rho(A_M(W) \otimes A_M(W) + B_M(W) \otimes B_M(W)) < 1\,.$$

3.3 Inferred Network

Monica et al. (2015) introduced methods to build inference-based networks that revise our knowledge of the underlying network. Depending on the information used to filter the network, we have four different cases:

- The choice of ARCH or GARCH parameter matrices;
- The choice of left or right multiplication in the proximity matrices.

Take the GARCH parameter matrix and right multiplication for example, the filtered network is constructed as:

$$[W_t^{A,L}]_{i,j} = a_{1,L,i}w_{t,i,j} * (1 - p\text{-}value(a_{1,L,i}w_{t,i,j} = 0)) = a_{1,L,i}w_{t,i,j} * (1 - p\text{-}value(a_i^L = 0))\,.$$

where the p-values are computed under the asymptotic normality of the maximum composite likelihood estimator (Varin and Vidoni, 2005). Then the test statistics is computed as:

$$t\text{-}stat(a_{1,R,i}\omega_{t,i,j} = 0) = \frac{\hat{a}_{1,R,i}\omega_{t,i,j}}{\sqrt{\mathbb{V}ar(\hat{a}_{1,R,i}\omega_{t,i,j})}}$$
$$= \frac{\hat{a}_{1,R,i}\omega_{t,i,j}}{\omega_{t,i,j}\sqrt{\mathbb{V}ar(\hat{a}_{1,R,i}}} = t\text{-}stat(a_{1,R,i} = 0)\ for\ each\ j.$$

4 Data

We will use the test data set provided by the Engle et al. (2017) paper to test spatial BEKK model with the composite likelihood method and the nonlinear shrinkage method against the restricted diagonal BEKK model without network dependence. The test data set contains n = 1000 variables/assets, and T = 1250 observations. For the purpose of demonstration, we will only consider 6 variables/assets in this report. For the weight matrix, we will use a simple grouping criteria at the moment and obtain the following:

$$W = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 \end{bmatrix}.$$

5 Result

We wrote Matlab codes for the new model, adapted from the Matlab codes used in Engle et al. (2017) and the MFE Toolbox by Kevin Sheppard, and estimated three models: (i) a restricted diagonal model that has no network dependence, (ii) a left multiplication model magnifies the risk receivers of the network, and (iii) a right multiplication model focuses on the risk spreader of the network. The maximum composite likelihood estimation results for the relevant parameters of these models are reported in Table 1.

As we can see in Table 1, both the left multiplication (risk receivers) and right multiplication (risk spreaders) spatial models outperform the restricted diagonal BEKK model, demonstrated by the likelihood ratio test statistics.

It is also notable the risk receiving propensity of entity 3 and entity 4 in both short term (as shown in the ARCH parameters of the model) and long term (as shown in the GARCH parameters of the model). In both the left and right multiplication models, the different signs for entity 2 could be evidence that it plays a diversifying role in the covariance contribution with respect to other entities in the network in both short and long term, though in the left multiplication case, the coefficients are not statistically different than zero.

Table 1: Estimated parameters of Diagonal BEKK (top panel), spatial risk receivers (left multiplication) BEKK (central panel), and spatial risk spreaders (right multiplication) BEKK (bottom panel) models on the test data set

	a_0	b_0		
entity1	0.0000(0.0001)	0.0066(0.2524)		
entity2	0.3691***(0.0436)	0.0055(0.1230)		
entity3	0.0000(0.0707)	0.0607(0.4658)		
entity4	0.4220***(0.0660)	0.0000(0.1106)		
entity5	0.3897***(0.0605)	0.2424***(0.0803)		
entity6	0.2719***(0.0883)	0.2090(0.1394)		
log-likelihood:	6099			
	$a_{0,L}$	$a_{1,L}$	$b_{0,L}$	$b_{1,L}$
entity1	0.0000(0.0438)	0.0000(0.5197)	0.1161(0.2737)	0.0008(0.1023)
entity2	0.4341***(0.1323)	0.1245(0.2271)	-0.0438(0.1936)	-0.4135(0.5047)
entity3	0.1537(0.1068)	0.1732***(0.0001)	-0.1773(0.1687)	0.2000***(0.0001)
entity4	-0.3407***(0.1167)	0.7015***(0.2132)	0.2188***(0.0809)	0.6483***(0.1887)
entity5	-0.1817*(0.1034)	0.4590***(0.1319)	-0.2358***(0.0806)	-0.2296(0.1626)
entity6	-0.3724*(0.1949)	0.1151(0.1319)	0.0833(0.1357)	-0.2144(0.1578)
log-likelihood:	6141			
likelihood ratio:	84***			
	$a_{0,R}$	$a_{1,R}$	$b_{0,R}$	$b_{1,R}$
entity1	0.0000(0.1362)	0.1319(0.1227)	0.0000(0.1735)	0.3742***(0.1208)
entity2	0.5419***(0.0676)	-0.2849***(0.0707)	0.0647(0.1034)	0.0181(0.0873)
entity3	0.1567**(0.0792)	0.1732***(0.0001)	0.1750(0.1668)	0.2000***(0.0001)
entity4	-0.3091***(0.0582)	0.4137***(0.0618)	-0.0463(0.1279)	0.0310(0.0923)
entity5	0.1336***(0.0421)	0.3783***(0.0682)	-0.1365(0.1157)	0.0560(0.5582)
entity6	0.3783***(0.0905)	-0.2087**(0.0913)	-0.0516(0.7359)	-0.1381(0.9021)
log-likelihood:	6150			
likelihood ratio:	102***			

* parameters significant at the 10% level, ** parameters significant at the 5% level, *** parameters significant at the 1% level..

We can also use the inferred network method proposed by Monica et al. (2015) to produce a simple graphical representation of the networks that allows us monitor the edges of the network, or the level of exposure between two specific entities. In our case, we assume that the weight matrix to be time-invariant as we have discussed in previous sections, and that all the suitable regularity conditions are met so that the maximum composite likelihood estimators are asymptotically normally distributed. The preliminary results are shown in Figure 1 and Figure 2. The second column in Figure 1 represents the short-term response to shocks and shows that the inference mainly magnifies the role of entity 6, which is the most fragile entity in the network in the short term. Similarly, Figure 2 magnifies the risk spreading role of entity 4 in the short term and entity 1 in the long term.

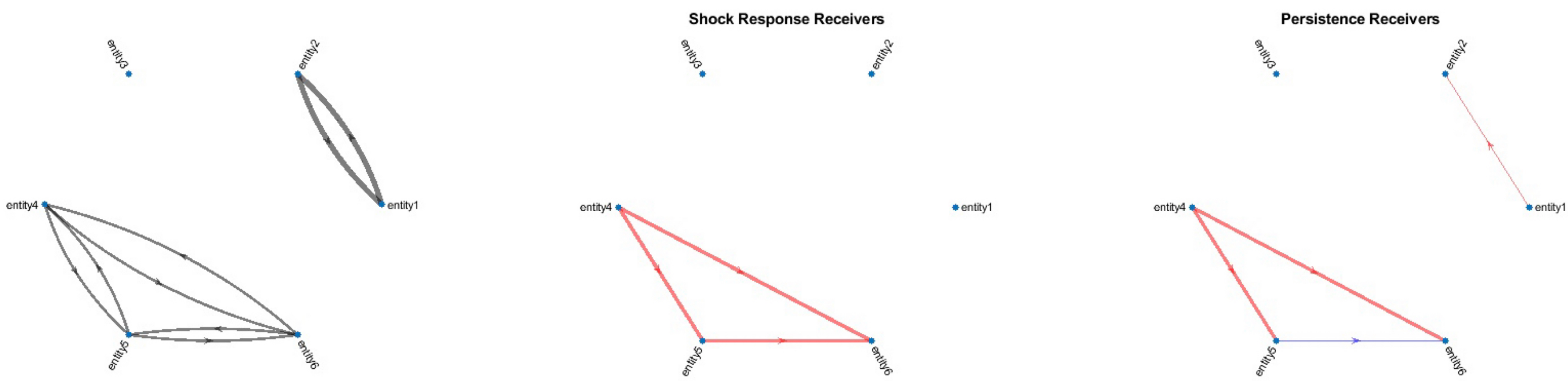

Figure 1: Normalized network ω_t, Shock response risk receivers (left multiplication) network $\mathcal{W}_t^{A,L}$, and persistence risk receivers (left multiplication) network $\mathcal{W}_t^{B,L}$. Positive weights are in red while negative are in blue.

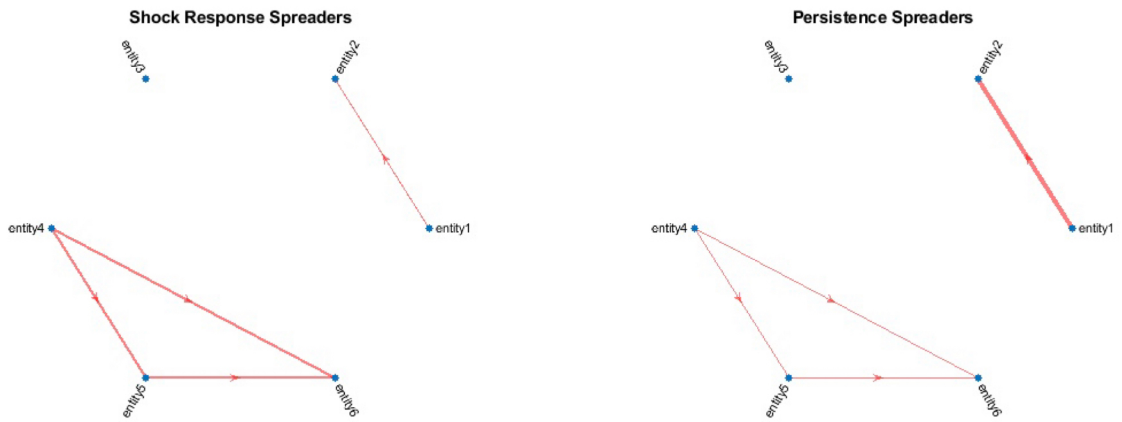

Figure 2: Shock response risk spreaders (right multiplication) network $\mathcal{W}_t^{A,R}$, and persistence risk spreaders (right multiplication) network $\mathcal{W}_t^{B,R}$. Positive weights are in red while negative are in blue.

6 Conclusion

In this project, we reviewed the DCC GARCH model and the Spatial BEKK model. We incorporated two methods, the composite likelihood method and the nonlinear shrinkage method into the Spatial BEKK model, and compared the new model with the restricted diagonal BEKK model, using a test data set. The results showed that the new model outperforms the restricted diagonal BEKK model. In our future study, we will be looking into the possibility of applying the new model to forecasting and portfolio selection process.

References

[1] Monica Billio et al. "Networks in Risk Spillovers: A Multivariate GARCH Perspective". In: 3 (2016).

[2] Tim Bollerslev. "Generalized Autoregressive Conditional Heteroskedasticity". In: *Journal of Econometrics* 31.3 (1986), pp. 307–327.

[3] Tim Bollerslev. "Modeling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model". In: *Review of Economics and Statistics* 72.3 (1990), pp. 498–505.

[4] Massimilano Caporin and Paolo Paruolo. "Proximity-Structured Multivariate Volatility Models". In: *Econometric Reviews* 36.4 (2015), pp. 493–493.

[5] Robert F. Engle. "Dynamic Conditional Correlation: A Simple Class of Multivariate GARCH Models". In: *Journal of Business Economic Statistics* 20.3 (1990), pp. 339–350.

[6] Robert F. Engle, Olivier Ledoit, and Michael Wolf. "Large Dynamic Covariance Matrices". In: *Journal of Business Economic Statistics* (2017), pp. 1–13.

[7] Olivier Ledoit and Michael Wolf. "Honey, I Shrunk the Sample Covariance Matrix". In: *The Journal of Portfolio Management* 30 (2004), pp. 110–119.

[8] Cavit Pakel et al. "Fitting Vast Dimensional Time-Varying Covariance Models". In: (2008).

[9] Cristiano Varin and Paolo Vidoni. "A Note on Composite Likelihood Inference and Model Selection". In: *Biometrika* 92.3 (2005), pp. 519–528.