Comparing Multivariate GARCH-DCC Models using Hamiltonian Monte Carlo and Stan

Rafael Soares Paixão,
Ricardo Sandes Ehlers

Instituto de Ciências Matemáticas e Computação - USP
Departamento de Estatística - UFSCar
rafaelspx@usp.br, ehlers@icmc.usp.br

Introduction

The demand for practical statistical tool to modeling volatility in Financial and Econometric time-series has been the object of much attention ever since the introduction of ARCH model (Autoregressive conditional Heteroscedasticity) (Engle [1982]). Subsequently, numerous variants and extensions ARCH models have been proposed. A large body this has been literature devoted to univariate models.

While modeling the volatility of returns has been the main center of attention, understanding the financial return comovements is of great practical importance. Therefore, it is important to extend the considerations multivariate GARCH (MGARCH). The complexity of MGARCH models has been the obstacle to their major use in applied works, because as the number of parameters in an MGARCH model often increases rapidly with the dimension of the model, for example for a number of k-dimensional returns there are k(k + 1)/2 elements in the covariance matrix. Therefore, the specification should be parsimonious enough to allow for relatively easy estimation of the model and also allow for easy interpretation of the model parameters.

In this context, the objective of this study is to compare and apply stochastic simulation techniques to the Bayesian estimates for the parameters of the GARCH-EDCC model, which is an extension of the GARCH-DCC model, considering the multivariate distributions. It will be necessary to distributions specification a prior for the parameters of the proposed model. A posterior to distributions will be obtained through the HMC method.

Extended Dynamic Conditional Correlation Garch Model

Consider a multivariate time series a $\mathbf{y_t} = (y_{1t}, y_{2t}, ..., y_{kt})^T$. The model is given by

$$\begin{aligned} \mathbf{y_t} &= \mathbf{H_t}^{1/2}\mathbf{e_t}, \\ \mathbf{H_t} &= \mathbf{D_t R_t D_t} \end{aligned} \tag{1}$$

where

$\mathbf{R_t} = diag(\mathbf{Q_t})^{-1/2}\mathbf{Q_t}diag(\mathbf{Q_t})^{-1/2}$,
$\mathbf{Q_t} = (1-\theta_1-\theta_2)\mathbf{R} + \theta_1\mathbf{u_t}\mathbf{u_t}^T + \theta_2\mathbf{Q_{t-1}}$,
$\mathbf{D_t} = diag(h_{1t}^{1/2}, ..., h_{kt}^{1/2})$,
$h_t = \omega + \sum_{j=1}^{p}\mathbf{A_j}y_{t-j}^2 + \sum_1^q \mathbf{B_j}h_{t-j}$,

$\mathbf{u_t} = \mathbf{D_t}^{-1}\mathbf{y_t}$ are the standardized returns, R is the unconditional covariance matrix of $\mathbf{u_t}$, $\theta_1 > 0$, $\theta_2 > 0$, $\theta_1 + \theta_2 < 1$, ω is k vector, $y_t^2 = y_t \odot y_t$ and A = $[\alpha_{ij}]$ and B = $[\beta_{ij}]$ are $k \times k$ matrix, if A and B is a diagonal matrix then model 1 is GARCH - DCC (Engle [2002]). The error vectors ($\mathbf{e_t}$) assumed independence and identically distributed with $\mathbb{E}(\mathbf{e_t}) = 0$ and $Var(\mathbf{e_t}) = I_k$. In this work, we will consider $\mathbf{e_t}$ follow a Normal/Independent Distribution with $p = q = 1$.

Normal/Independent Distribution

A d-dimensional random vector Y follow a normal/independent (NI) distribution with location parameter $\eta \in \mathbb{R}^k$ and positive definite scale matrix Σ (see Lange and Sinsheimer [1993]), if its probability density function has the form

$$f_Y(y) = \int_{\mathbb{R}^+} \phi_k(y \mid \eta, u^{-1}\Sigma)\partial H(u \mid \nu), \tag{2}$$

where $H(u \mid \eta)$ is a cumulative distribution function of a unidimensional positive random variable U indexed by the parameter vector ν and ϕ_k denotes the d-dimensional multivariate normal density with mean η and covariance matrix Σ. For a random vector with a probability density function as in 2, we denote it as $Y \sim NI_k(\eta, \Sigma, H)$.

Examples of NI distributions: Normal, Student's-t (t_ν), Slash (Sl_ν), Slash-T ($Slt_{\nu,\varpi}$) , and Contaminated-Normal ($CN_{\nu,\varpi}$) distributions.

Bayesian Inference

Bayesian inference is based on Bayes theorem:

$$\pi(\theta \mid y) \propto l(\theta)\pi(\theta)$$

where $\pi(\theta \mid y)$ is denoted of posterior distribution, $l(\theta)$ is likelihood function and $\pi(\theta)$ is denoted of prior distribution.

The posterior distribution $\pi(\theta \mid \{y_t\})$ is analytically intractable. Therefore, we adopt Hamiltonian Monte Carlo (HMC) sampling strategies for obtaining samples from the joint posterior distributions.

Hamiltonian Monte Carlo - HMC

The HMC is a method alternately combine Gibbs updates with Metropolis updates and avoids the random walk behavior. This method process a new state by computing a trajectory obeying Hamiltonian dynamics Neal et al. [2011].

Consider a random vector $\theta \in \mathbb{R}^k$ as position variables (parameters) and $\omega \in \mathbb{R}^k$ an independence auxiliary random vector, where $\omega \sim N_k(0, M)$. In physical analogy, the negative logarithm of the joint probability density for the parameters of interest, $-L(\theta) = -\ln p(\theta|y)$, denote a potential energy function, the ω is analogous to a momentum variable and the covariance matrix M denotes a mass matrix.

The Hamiltonian dynamics system is described by a function of the two variables known as the Hamiltonian function, $H(\theta, \omega)$, which is a sum the potential energy $U(\theta)$ and kinetic energy $K(\omega)$, Neal et al. [2011],

$$H(\theta, \omega) = U(\theta) + K(\omega),$$

where $U(\theta) = -L(\theta)$ and $K(\omega) = \frac{1}{2}\omega^T M^{-1}\omega$. These method propose a candidate (θ^*, ω^*) by solving the Hamiltonian equations, given by

$$\frac{\partial\theta}{\partial\tau} = \frac{\partial H(\theta,\omega)}{\partial\omega} = M^{-1}\omega \quad \frac{\partial\omega}{\partial\tau} = -\frac{\partial H(\theta,\omega)}{\partial\theta} = \nabla L(\theta)$$

over a fictitious time τ. In practice, the differential Hamiltonian equations are simulated by leapfrog method.

The state (θ^*, ω^*) is then accepted as the next state of the Markov chain with probability

$$\mathbb{P}(\theta, \omega; \theta^*, \omega^*) = \min\{1, \exp\{-H(\theta^*, \omega^*) + H(\theta^*, \omega^*)\}\}$$

Application

We analyze the daily observations of the hundredfold log-returns of daily indices of stock markets in Frankfurt (DAX), Paris (CAC40) and Tokyo (NIKKEI), from 10 October 1991 until 30 December 1997, , which leads to 1628 observations. The stock market data is freely available at `http://robjhyndman.com/tsdldata/data/FVD1.dat`.

To simulate the posterior distribution we use the HMC algorithm of **Rstan** package of **R** program. We have then run a total of 15,000 iterations discarding the first 5,000 realizations as burn-in. Posterior results are then based on 10,000 realizations of the Markov chain with the prior distributions as $\omega_i \sim N(0,1)\mathbb{I}(\omega_i > 0)$; $(\alpha_{ii}, \beta_{ii}) \sim$ Dirichlet$(1,1,1)$; $\alpha_{i,j} \sim N(0,1)\mathbb{I}(\alpha_{i,j} > 0)$; $\beta_{i,j} \sim N(0,1)\mathbb{I}(\beta_{i,j} > 0)$; $(\theta_1, \theta_i) \sim$ Dirichlet$(1,1,1)$; $\nu \sim N(0,10)\mathbb{I}(\nu > 2)$, if $e_t \sim t_\nu$; $\nu \sim N(0,10)\mathbb{I}(\nu > 1)$, if $e_t \sim Sl_\nu$; $\nu \sim N(0,10)\mathbb{I}(\nu > 2)$ and $\varpi > 1$ if $e_t \sim SlT_{\nu,\varpi}$; and $\nu \sim U(0,1)$ and $\varpi \sim U(0,1)$ if $e_t \sim CN_{\nu,\varpi}$.

After we obtained our own estimations, we used the Deviance Information Criterion (DIC), Watanabe-Akaike information criterion (WAIC) and Leave-one-out cross-validation information criterion (LOOIC) for comparison.

Table 2: DIC, WAIC and LOOIC values for the ten multivariate GARCH models compared.

	DIC	WAIC	LOOIC
DCC-NORMAL	13996.73	14009.62	14010.46
DCC-T	13803.68	**13806.44**	**13806.46**
DCC-SLASH	13811.93	13814.84	13814.85
DCC-SLASH-T	**13803.54**	13808.94	13808.91
DCC-CONT.NORMAL	13823.56	13830.01	13830.42
EDCC-NORMAL	14001.34	14017.13	14017.71
EDCC-T	13816.24	13820.96	13820.98
EDCC-SLASH	13824.03	13829.01	13829.02
EDCC-SLASH-T	13816.88	13823.82	13823.79
EDCC-CONT.NORMAL	13830.74	13845.78	13846.04

The Dynamic Conditional Correlation model with t innovations ($DCC - T$) has the lowest looic and waic.

Table 4: Summary of the HMC simulations for the model with multivariate skew t errors

	mean	se_mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
ω_1	0.02	0.00	0.01	0.01	0.01	0.02	0.02	0.04	1530	1.00
ω_2	0.03	0.00	0.01	0.01	0.02	0.02	0.03	0.06	2187	1.00
ω_3	0.02	0.00	0.01	0.01	0.02	0.02	0.03	0.05	546	1.01
α_1	0.06	0.00	0.01	0.04	0.05	0.06	0.07	0.09	1611	1.00
α_2	0.03	0.00	0.01	0.02	0.03	0.03	0.04	0.06	1760	1.00
α_3	0.07	0.00	0.01	0.05	0.06	0.07	0.08	0.10	404	1.01
β_1	0.92	0.00	0.02	0.88	0.91	0.92	0.94	0.95	1323	1.00
β_2	0.95	0.00	0.02	0.90	0.94	0.95	0.96	0.97	2099	1.00
β_3	0.92	0.00	0.01	0.89	0.91	0.92	0.93	0.95	346	1.01
θ_1	0.04	0.00	0.01	0.02	0.03	0.04	0.04	0.06	10000	1.00
θ_2	0.70	0.00	0.13	0.39	0.64	0.73	0.79	0.89	10000	1.00
ν	8.09	0.01	0.86	6.60	7.49	8.03	8.62	9.94	10000	1.00

In Figure 3 shows the estimated marginal density function of the posterior and function of autocorrelation of HMC sampler. We observed that the dependence between samples quickly decays to increase as the distance of the sample.

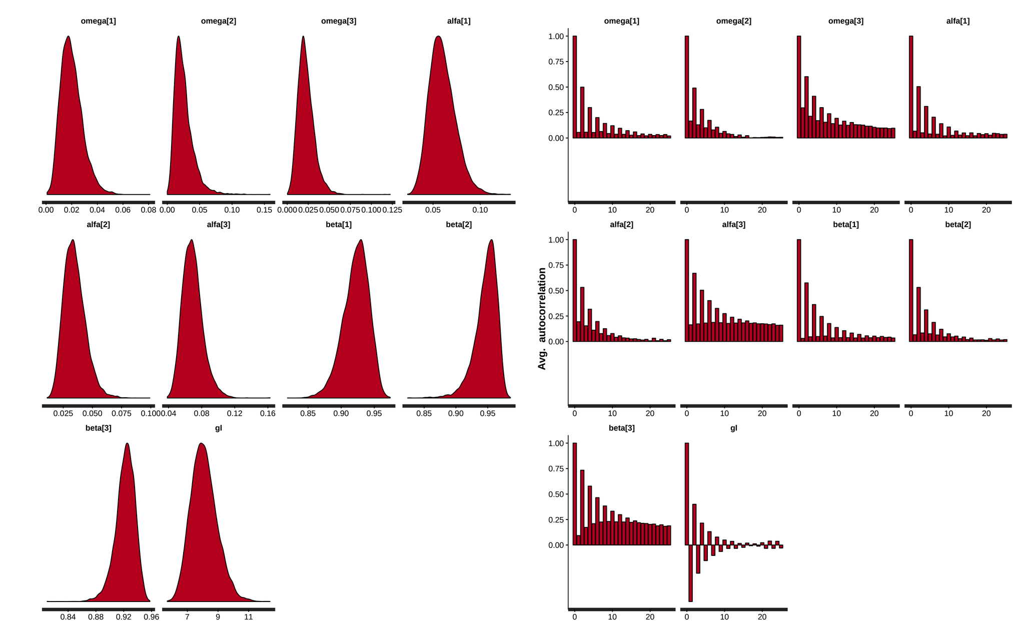

Figure 2: Estimates of the marginal probability density function of posterior and Function of autocorrelation of HMC sampler

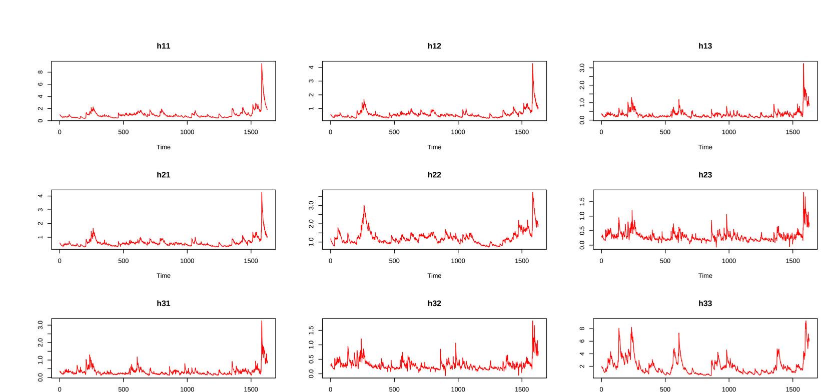

Figure 4: Volatility estimates for the DAX, CAC40, and Nikkei log-returns of model DCC-T

Conclusions

In the present work, we studied the Extended Dynamic Conditional Correlation model with different innovations following Normal/independent distribution. The HMC algorithm was used to simulate the posterior distribution.

To compare of the proposed models using the criteria DIC, WAIC and LOOIC, the better performance is attained by the Dynamic Conditional Correlation model with t innovations.

References

Robert Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

Kenneth Lange and Janet S Sinsheimer. Normal/independent distributions and their applications in robust regression. *Journal of Computational and Graphical Statistics*, 2(2):175–198, 1993.

Radford M Neal et al. Mcmc using hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*, 2:113–162, 2011.