Chapter 11

Financial Data Mining Using Flexible ICA–GARCH Models

Philip I.H. Yu
The University of Hong Kong, Hong Kong

Edmond H.C. Wu
The Hong Kong Polytechnic University, Hong Kong

W.K. Li
The University of Hong Kong, Hong Kong

ABSTRACT

As a data mining technique, independent component analysis (ICA) is used to separate mixed data signals into statistically independent sources. In this chapter, we apply ICA for modeling multivariate volatility of financial asset returns which is a useful tool in portfolio selection and risk management. In the finance literature, the generalized autoregressive conditional heteroskedasticity (GARCH) model and its variants such as EGARCH and GJR-GARCH models have become popular standard tools to model the volatility processes of financial time series. Although univariate GARCH models are successful in modeling volatilities of financial time series, the problem of modeling multivariate time series has always been challenging. Recently, Wu, Yu, & Li (2006) suggested using independent component analysis (ICA) to decompose multivariate time series into statistically independent time series components and then separately modeled the independent components by univariate GARCH models. In this chapter, we extend this class of ICA-GARCH models to allow more flexible univariate GARCH-type models. We also apply the proposed models to compute the value-at-risk (VaR) for risk management applications. Backtesting and out-of-sample tests suggest that the ICA-GARCH models have a clear cut advantage over some other approaches in value-at-risk estimation.

DOI: 10.4018/978-1-60566-908-3.ch011
INTRODUCTION

In econometrics, volatility modeling of financial time series has received a lot of attention due to its wide applications in finance such as option pricing and risk management. Among the existing volatility models, one of the most important models is the autoregressive conditional heteroscedasticity (ARCH) model proposed by (Engle, 1982) which was further extended to generalized ARCH (GARCH) model by (Bollerslev, 1986). After the success of ARCH and GARCH models, researchers further proposed different types of GARCH models such as EGARCH (Nelson, 1991) and GJR-GARCH (Glosten, Jagannathan, & Runkle, 1993), etc. These univariate GARCH models are capable in capturing the dynamics of volatilities from the characteristics of financial time series.

Although GARCH models are successful in modeling volatilities of univariate financial time series, the problem of modeling multivariate time series still raises challenges in this research area. It is mainly because in existing multivariate GARCH models, the number of unknown parameters grows very fast with the number of time series in the model. For example, in (Engle, 2002), Engle et al. compared the complexity of several multivariate GARCH models and most of them have the complexity of \( O(N^3) \) or even \( O(N^2) \), where \( N \) is the number of time series.

However, in practice, we often need to extend the volatility modeling to high dimensional cases. For instance in portfolio optimization, a portfolio could contain several hundred of stocks. Therefore, new approaches are needed to deal with such situations.

Recently, (Wu, Yu & Li, 2006) suggested using independent component analysis (ICA) to decompose multivariate time series into statistically independent time series components and then separately modeled the independent components (ICs) by univariate GARCH models. Their experiment results showed that the ICA-GARCH models are more effective in capturing the time-varying features of volatilities and provide better value-at-risk estimate than existing methods including DCC (Engle, 2002), PCA-GARCH (Alexander, 2001) and RiskMetrics.

In this chapter, we extend this class of ICA-GARCH models to allow more flexible univariate GARCH-type models. In addition to the popular ICs extraction methods—FastICA algorithms, we also consider other ICA algorithms which can extract ICs from non-stationary data (Hyvärinen, 2001) as most financial time series are non-stationary.

The rest of this chapter is organized as follows: In Section 2, we introduce the univariate GARCH process and several of its extended models. Then, we propose flexible ICA-GARCH models for multivariate volatility modeling in Section 3. In Section 4, we consider the estimation of value at risk of a single asset or a portfolio based on the flexible ICA-GARCH models. Experimental results are given in Section 5. Finally, we conclude in Section 6.

VOLATILITY MODELS

In the following, we introduce several prevailing volatility models for financial time series.

GARCH \((p,q)\) Model

In GARCH models, financial time series \( \{y_t\} \) are assumed to be generated by a stochastic process with (conditional) time-varying volatility \( \{\sigma_t\} \). The general GARCH \((p,q)\) model \((p > 0 \text{ and } q \geq 0 \text{ are integers})\) is defined as