Signal Processing 89 (2009) 921–932

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Estimating multivariate ARCH parameters by two-stage least-squares method

Saman Mousazadeh [a,b,*], Mahmood Karimi [a]

[a] Department of Electrical Engineering, Shiraz University, Shiraz, Iran
[b] Iran Telecommunication Research Center (ITRC), Iran

ARTICLE INFO

Article history:
Received 16 March 2008
Received in revised form
5 October 2008
Accepted 17 November 2008
Available online 28 November 2008

Keywords:
Multivariate ARCH
Parameter estimation
QMLE
TSLS

ABSTRACT

This paper discusses the asymptotic properties of the two-stage least-squares (TSLS) estimator of the parameters of multivariate autoregressive conditional heteroscedasticity (ARCH) model. The estimator is easy to obtain since it involves solving sets of linear equations. It will be shown that, under some conditions, this TSLS estimator is asymptotically consistent and its rate of convergence is the same as that of the quasi maximum likelihood estimator (QMLE). At the same time, the computational load of the TSLS estimator is extremely lower than that of the QMLE. The performance of the TSLS estimator will be evaluated and compared with QMLE using simulations. Simulation results show that the performances of the two estimators are comparable, even for small data records.

1. Introduction

The autoregressive conditional heteroscedasticity (ARCH) model was first introduced by Engle in 1982 [1]. This model represents a powerful tool for the analysis and forecasting of volatility in financial markets. ARCH is a statistical model which explicitly parameterizes a time-varying conditional variance using squared absolute values, while considering volatility clustering and excess kurtosis (i.e. heavy-tailed distribution). This feature of ARCH models enables them to be applied to numerous economic and financial data to model unpredictability; the strong dependence of the instantaneous variability of a time series on its own past. Another feature of the ARCH process is that it is a white process. So, in some applications it may be useful for modeling heavy-tailed white processes.

Since 1982, the ARCH model has been expanded by several authors such as Bollerslev [2,3] and others. One of these expanded models is the multivariate generalized ARCH (multivariate GARCH) model introduced by Bollerslev [3]. The multivariate ARCH model with constant correlation is a special case of this model.

The most obvious application of multivariate ARCH models is the study of the relations between the volatilities and co-volatilities of several markets [3]. As an example, it was used to model coherence in short-run nominal exchange rates [3]. As econometric variables usually depend on each other, the multivariate GARCH model has recently been used for modeling the econometrics (see [4–8]).

Recently, GARCH modeling has been used in many signal processing applications such as speech denoising [9], blind speech source separation [10], voice activity detection [11], speech recognition in isolated digits [12], and ambient sea noise modeling [13]. In [12], a GARCH model has been utilized in the time domain for speech recognition applications. The model parameters, characterizing the speech phonemes, are assumed speaker independent. However, in most speech processing

* Corresponding author at: Department of Electrical Engineering, Shiraz University, Shiraz, Iran. Tel.: +98 935 426 7897.
E-mail addresses: smousazadeh@gmail.com (S. Mousazadeh), karimi@shirazu.ac.ir (M. Karimi).

0165-1684/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2008.11.012

applications like speech denoising and blind speech source separation, the speech is modeled in the short time Fourier transforms (STFT) domain as a GARCH process. For the simplicity and lowering the computational load, most authors assume that signals in different frequency bins are independent and can be modeled separately [9–11,13]. However, it is mentioned that if one can model the signals as a multivariate GARCH process better results can be achieved [10]. Since the GARCH process can be modeled as an ARCH process with high order [14], it seems that one can use the ARCH model of high order instead of the GARCH model in speech processing applications [14].

Parameter estimation is the first step in model based signal processing. The most popular method for estimating the parameters of the multivariate ARCH model is quasi maximum likelihood (QML) estimation which needs numerical maximization because it does not admit a closed form expression [15]. In this method the process noise, $\boldsymbol{\varepsilon}(t)$, (see Section 2) is assumed to be Gaussian (this is because the true distribution might not be known), and by using this assumption one can maximize the likelihood function over the parameters. Note that usually the true distribution of $\boldsymbol{\varepsilon}(t)$ is not known and the maximum likelihood (ML) estimator cannot be used. So, it is assumed that $\boldsymbol{\varepsilon}(t)$ is Gaussian and the estimator is denoted as QML. It is shown in [16] that the QML estimator is asymptotically consistent. The drawback of this method is its computational load, which is very high [15]. Therefore this method cannot be used in real time signal processing applications such as speech enhancement, which needs to be real time in some cases. In this paper we propose an asymptotically consistent estimator for the multivariate ARCH parameters. In the scalar case, this multivariate estimator reduces to the estimator proposed in [17]. Our estimator is very simple because it just involves the solution of linear equations, and therefore its computational load is extremely lower than the QML estimator. The main achievement of this paper is to show that the rate of convergence for our estimator (as the number of data increases) is the same as the QML estimator. Simulation results show that the performance of our estimator is approximately the same as the performance of the QML estimator.

The remainder of this paper is organized as follows. In Section 2 we derive our estimator named the two-stage least-squares (TSLS) estimator. In Section 3 we derive the statistical properties of the TSLS estimator in the large data case. In Section 4 we use simulations to evaluate the performance of our estimator and compare it with that of the QML estimator. Finally, the proof of some equations is given in Appendix A.

2. TSLS estimator

Multivariate ARCH model with constant correlation is defined as

$$\mathbf{x}(t) = \boldsymbol{\Sigma}(t)\boldsymbol{\varepsilon}(t) \tag{1}$$

where $\mathbf{x}(t)$ is a $d \times 1$ vector, $\boldsymbol{\varepsilon}(t)$ is a $d \times 1$ independent identically distributed (i.i.d.) zero mean normal vector with unknown covariance matrix

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & \gamma_{12} & \cdots & \gamma_{1d} \\ \gamma_{21} & 1 & \cdots & \gamma_{2d} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & & \cdot \\ \gamma_{d1} & \gamma_{d2} & \cdots & 1 \end{bmatrix} \tag{2}$$

and $\boldsymbol{\Sigma}(t)$ is a $d \times d$ matrix defined as

$$\boldsymbol{\Sigma}(t) = diag(\sigma_1(t), \sigma_2(t), \ldots, \sigma_d(t)) \tag{3}$$

$$[\sigma_1^2(t), \sigma_2^2(t), \ldots, \sigma_d^2(t)]^{\mathrm{T}} = \mathbf{w} + \sum_{i=1}^{q} \mathbf{A_i}(\mathbf{x}(t-i) \odot \mathbf{x}(t-i)) \tag{4}$$

where $d \times 1$ vector $\mathbf{w}$ and the $d \times d$ matrices $\mathbf{A_i}$ are model parameters having all positive elements and

$$\mathbf{B} = [\mathbf{w} \vdots \mathbf{A}_1 \vdots \mathbf{A}_2 \vdots \cdots \vdots \mathbf{A}_q] \tag{5}$$

is the matrix of these parameters. q is the known order of the model and d is the dimension of the model. $[.]^{\mathrm{T}}$ denotes transpose of a matrix or vector and $\odot$ is the Hadamard product (term by term multiplication) operator. It can be shown that if the polynomial $p(\lambda) = \det(\mathbf{I} - \sum_{i=q}^{q} \mathbf{A_i}\lambda^i)$, where $\mathbf{I}$ is the identity matrix, has all of its roots outside the unit circle, then this model is strictly stationary and ergodic [16].

The definition of multivariate ARCH model given in (1) shows that this model is a multivariate nonlinear autoregressive process. The term $\boldsymbol{\Sigma}(t)$ on the right hand side of (1) shows the dependence of $\mathbf{x}(t)$ on its past values, and the process noise $\boldsymbol{\varepsilon}(t)$ is the input noise of the model.

If we square the elements of $\mathbf{x}(t)$ and call the produced vector $\mathbf{y}(t)$, i.e.

$$\mathbf{y}(t) = [y_1(t), y_2(t), \ldots, y_d(t)]^{\mathrm{T}} = \mathbf{x}(t) \odot \mathbf{x}(t) \tag{6}$$

then

$$\begin{cases} x_1(t) = \sigma_1(t)\varepsilon_1(t) \\ x_2(t) = \sigma_2(t)\varepsilon_2(t) \\ \qquad \cdot \\ \qquad \cdot \\ x_d(t) = \sigma_d(t)\varepsilon_d(t) \end{cases}$$

$$\Rightarrow \begin{cases} y_1(t) = x_1^2(t) = \sigma_1^2(t) + \sigma_1^2(t)(\varepsilon_1^2(t) - 1) \\ y_2(t) = x_2^2(t) = \sigma_2^2(t) + \sigma_2^2(t)(\varepsilon_2^2(t) - 1) \\ \qquad \cdot \\ \qquad \cdot \\ y_d(t) = x_d^2(t) = \sigma_d^2(t) + \sigma_d^2(t)(\varepsilon_d^2(t) - 1) \end{cases} \tag{7}$$

where x_i and ε_i are the ith elements of $\mathbf{x}(t)$ and $\boldsymbol{\varepsilon}(t)$, respectively. If we define

$$\eta_i(t) = \varepsilon_i^2(t) - 1 \tag{8}$$