Nonlinear Parameter Estimation for State-Space ARCH Models with Missing Observations

SEBASTIÁN OSSANDÓN
Pontificia Universidad Católica de Valparaíso
Instituto de Matemáticas
Blanco Viel 596, Cerro Barón, Valparaíso
CHILE
sebastian.ossandon@ucv.cl

NATALIA BAHAMONDE
Pontificia Universidad Católica de Valparaíso
Instituto de Estadística
Blanco Viel 596, Cerro Barón, Valparaíso
CHILE
natalia.bahamonde@ucv.cl

Abstract: A new mathematical representation, based on a discrete time nonlinear state space formulation, is presented to characterize AutoRegresive Conditional Heteroskedasticity (ARCH) models. A novel parameter estimation procedure for state-space ARCH models with missing observations, based on an Extended Kalman Filter (EKF) approach, is described and successfully evaluated herein. Finally, through a comparison analysis between our proposed estimation method and a Quasi Maximum Likelihood Estimation (QMLE) technique based on different methods of imputation, some numerical results with simulated data, which make evident the effectiveness and relevance of the proposed nonlinear estimation technique are given.

Key–Words: ARCH models, Missing observations, Nonlinear state space model, Nonlinear estimation, Extended Kalman Filter.

1 Introduction

Autoregressive conditionally heteroscedastic (ARCH) type modeling, introduced by [4], are often used in finance because their properties are close to the observed properties of empirical financial data such as heavy tails, volatility clustering, white noise behavior or autocorrelation of the squared series. Financial time series often exhibit that the conditional variance can change over time, namely heteroskedasticity. The ARCH family of model is a class of nonlinear time series models introduced by [1]. The ARCH(p) model with normal error is

$$\begin{cases} r_k = \sigma_k \varepsilon_k, \\ \sigma_k^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{k-i}^2, \end{cases} \tag{1}$$

where r_k and σ_k ($> 0, \forall k$) are, respectively the return and the volatility, in the discrete time $k \in \mathbb{Z}$, associated to a financial process, and $\{\varepsilon_k\}_{k\in\mathbb{Z}}$ is a i.i.d. Gaussian sequence, with $\mathbb{E}(\varepsilon_k) = 0, \mathbb{E}(\varepsilon_k \cdot \varepsilon_j) = Q\delta_{k-j}$, and parameters $\alpha_0 > 0$, $\alpha_i \geq 0, p \geq i \geq 1$ and $\sum_{i=1}^{p} \alpha_i < 1$. Moreover, let us consider r_0 independent of sequence $\{\varepsilon_k\}_{k>0}$.

The goal of this research is to develop a novel nonlinear parameter estimation procedure, based on an EKF approach, for ARCH models considering missing observations. The EKF technique proposed is derived from a nonlinear state space formulation of the discrete time ARCH equation (see equation (1)). This method is adequate to obtain initial conditions for a maximum-likelihood iteration, or to provide the final estimation of the parameters and the states when maximum-

ISBN: 978-1-61804-141-8

likelihood is considered inadequate or costly.

2 Nonlinear state space formulation for ARCH models

Let us consider the following general discrete time nonlinear state space mathematical model:

$$\begin{cases} \boldsymbol{X}_k = \boldsymbol{f}(\boldsymbol{X}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{\theta}) + \boldsymbol{\sigma}(\boldsymbol{u}_{k-1}, \boldsymbol{\theta}) \cdot \boldsymbol{w}_k, \\ \boldsymbol{Y}_k = \boldsymbol{h}(\boldsymbol{X}_k, \boldsymbol{u}_k, \boldsymbol{\theta}) + \boldsymbol{\nu}_k, \end{cases} \tag{2}$$

where $\boldsymbol{X}_k \in \mathbb{R}^n$ is the state unknown vector, $\boldsymbol{u}_k \in \mathbb{R}^r$ is the input known vector, $\boldsymbol{Y}_k \in \mathbb{R}^m$ is the noisy observation vector or output vector of the stochastic process, $\boldsymbol{w}_k \in \mathbb{R}^n$ and $\boldsymbol{\nu}_k \in \mathbb{R}^m$ are, respectively, the process noise (due, mainly, to disturbances and modelling inaccuracies of the process) and the measurement noise (due, mainly, to sensor inaccuracy). Moreover $\boldsymbol{\theta} \in \mathbb{R}^l$ is the parameter vector that is generally unknown, $\boldsymbol{f}(\cdot) \in \mathbb{R}^n$, $\boldsymbol{\sigma}(\cdot) \in \mathbb{R}^{n\times n}$ and $\boldsymbol{h}(\cdot) \in \mathbb{R}^m$ are nonlinear functions that characterize the stochastic system.

With respect to the noises of the process, we assume the following assumptions: i) the vector $\boldsymbol{w}_k$ is assumed to be Gaussian, zero-mean $\mathbb{E}(\boldsymbol{w}_k) = 0$ and white noise with covariance matrix $\mathbb{E}(\boldsymbol{w}_k \cdot \boldsymbol{w}_j^T) = \mathbf{Q} \cdot \boldsymbol{\delta}_{k-j}$; ii) the vector $\boldsymbol{\nu}_k$ is assumed to be Gaussian, zero-mean $\mathbb{E}(\boldsymbol{\nu}_k) = 0$ and white noise with covariance matrix $\mathbb{E}(\boldsymbol{\nu}_k \cdot \boldsymbol{\nu}_j^T) = \mathbf{R} \cdot \boldsymbol{\delta}_{k-j}$. Where $\boldsymbol{\delta}_{k-j}$ = identity matrix when $k = j$, otherwise, $\boldsymbol{\delta}_{k-j}$ = zero matrix.

2.1 Case 1: State space formulation for ARCH(1)

In this case a possible state space representation of equation (1) when $p = 1$ is,

$$\begin{cases} \boldsymbol{X}_k := \begin{bmatrix} X_k^{(1)} \\ X_k^{(2)} \end{bmatrix} = \boldsymbol{f}(\boldsymbol{X}_{k-1}, \boldsymbol{\theta}) + \mathcal{B} \cdot \boldsymbol{w}_k, \\ \boldsymbol{Y}_k := r_k = X_k^{(2)} \sqrt{X_k^{(1)}}, \end{cases} \tag{3}$$

where $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) \in \mathbb{R}^3$, $\mathcal{B} = (0, 1)^\top \in \mathbb{R}^2$, $X_k^{(1)} = \sigma_k^2 \in \mathbb{R}$ $(> 0, \forall k)$, $X_k^{(2)} = r_k/\sigma_k \in \mathbb{R}$, $\boldsymbol{w}_k = \varepsilon_k$ and

$$\boldsymbol{f}(\boldsymbol{X}_{k-1}, \boldsymbol{\theta}) := \begin{bmatrix} \alpha_0 \\ 0 \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ 0 \end{bmatrix} X_{k-1}^{(1)} (X_{k-1}^{(2)})^2. \tag{4}$$

For details see [6].

The construction leading to the model of equations (3) and (4) is easily extended to ARCH models with higher orders yielding a different state space formulation that will be studied in the next subsection.

2.2 Case 2: State space formulation for ARCH(p), $p > 1$

The following formulation is a possible state space representation of equation (1) when $p > 1$:

$$\begin{cases} \boldsymbol{X}_k := \begin{bmatrix} \boldsymbol{X}_k^{(1)} \\ X_k^{(2)} \\ X_k^{(3)} \end{bmatrix} = \mathcal{A}_{\boldsymbol{\theta}} \cdot \boldsymbol{X}_{k-1} + \boldsymbol{f}(\boldsymbol{X}_{k-1}, \boldsymbol{\theta}) + \\ \qquad + \mathcal{B} \cdot \boldsymbol{w}_k, \\ \boldsymbol{Y}_k := r_k = X_k^{(3)} \sqrt{X_k^{(2)}}, \end{cases} \tag{5}$$

where $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \cdots, \alpha_p, \beta_1)^\top \in \mathbb{R}^{p+2}$, $\mathcal{B} = (0, 0, \cdots, 0, 1)^\top \in \mathbb{R}^{p+2}$, $\boldsymbol{w}_k = \varepsilon_k$

$$\boldsymbol{X}_k^{(1)} := \begin{bmatrix} r_{k-p}^2 \\ r_{k-p+1}^2 \\ \vdots \\ r_{k-1}^2 \end{bmatrix} \in \mathbb{R}^p, \tag{6}$$

ISBN: 978-1-61804-141-8

$X_k^{(2)} = \sigma_k^2 \in \mathbb{R}\ (>0, \forall k)$, $X_k^{(3)} = r_k/\sigma_k \in \mathbb{R}$, and

$$\boldsymbol{f}(\boldsymbol{X}_{k-1}, \boldsymbol{\theta}) := \begin{bmatrix} \mathbf{0}_{p\times 1} \\ \alpha_0 \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{(p-1)\times 1} \\ X_{k-1}^{(2)}(X_{k-1}^{(3)})^2 \\ \alpha_1 X_{k-1}^{(2)}(X_{k-1}^{(3)})^2 \\ 0 \end{bmatrix}. \tag{7}$$

Moreover,

$$\mathcal{A}_{\boldsymbol{\theta}} = \begin{pmatrix} \mathbf{A}_{p\times p} & \mathbf{0}_{p\times 2} \\ \boldsymbol{\alpha}_{2\times p} & \mathbf{0}_{2\times 2} \end{pmatrix} \in \mathbb{R}^{(p+2)\times(p+2)}, \tag{8}$$

with

$$\mathbf{A}_{p\times p} = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{p\times p},$$
$$\boldsymbol{\alpha}_{2\times p} = \begin{pmatrix} 0 & \alpha_p & \cdots & \alpha_2 \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{2\times p}, \tag{9}$$

and where $\mathbf{0}_{m\times n}$, $1 \leq \forall m, n \leq p$, are matrices of zeros whose size is $m \times n$.

Let us notice the obvious nonlinearity of all state space representations, presented in this section, due clearly to nonlinearity of the process and observation equations.

3 The Extended Kalman Filter

Let $\mathcal{Y}_N = [\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_k, ..., \boldsymbol{y}_N]$ be a known sequence of measurement or observations. The functions $\boldsymbol{f}$ and $\boldsymbol{h}$ (see equation (2)) are used to compute the predicted state and the predicted measurement from the previous estimate state. The following equation shows the computation of the predicted state from the previous estimate:

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{\theta}). \tag{10}$$

To compute the predicted estimate covariance a matrix $\mathbf{A}$ of partial derivatives (the Jacobian matrix) is previously computed. This matrix is evaluated, with the predicted states, at each discrete timestep and used in the KF equations. In other words, $\mathbf{A}$ is a linearized version of the nonlinear function $\boldsymbol{f}$ around the current estimate.

$$\mathbf{P}_{k|k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{A}_{k-1}^{\top} + \mathbf{Q}. \tag{11}$$

After making the prediction stage, we need to update the equations. So we have the residual measure innovation:

$$\tilde{\boldsymbol{y}}_k = \boldsymbol{y}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{u}_k, \boldsymbol{\theta}) \tag{12}$$

and the conditional covariance innovation

$$\mathbf{S}_{k|k-1} = \mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^{\top} + \mathbf{R}, \tag{13}$$

where $\mathbf{C}$ is a linearized version of the nonlinear function $\boldsymbol{h}$ around the current estimate.

The Kalman gain is given by

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{C}_k^{\top}\mathbf{S}_{k|k-1}^{-1}, \tag{14}$$

and the corresponding updates by

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \mathbf{K}_k\tilde{\boldsymbol{y}}_k \tag{15}$$

and

$$\mathbf{P}_{k|k} = (I - \mathbf{K}_k\mathbf{C}_k)\mathbf{P}_{k|k-1}. \tag{16}$$

The state transition and observation matrices (the linearized versions of $\boldsymbol{f}$ and $\boldsymbol{h}$) are defined, respectively, by

$$\mathbf{A}_{k-1} = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}}\right|_{\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}} \tag{17}$$

ISBN: 978-1-61804-141-8

and

$$\mathbf{C}_k = \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right|_{\hat{\boldsymbol{x}}_{k|k-1}} \tag{18}$$

4 Nonlinear estimation with missing observations

Given a sequence of measurement or observations $\mathcal{Y}_N$, the likelihood function is given by the following joint probability density function:

$$L(\boldsymbol{\theta}; \mathcal{Y}_N) = p(\boldsymbol{y}_0|\boldsymbol{\theta}) \prod_{k=1}^{N} p(\boldsymbol{y}_k|\mathcal{Y}_{k-1}, \boldsymbol{\theta}). \tag{19}$$

Due to the central limit theorem when the number of experiment replications becomes large the distribution of the mean becomes approximately normal and centered at the mean of each original observation. So it seems reasonable to assume that, for a large number of experiment replications, the probability density functions can be approximated by functions of Gaussian probability densities. Therefore we can rewrite equation (19) as follows:

$$L(\boldsymbol{\theta}; \mathcal{Y}_N) = \frac{p(\boldsymbol{y}_0|\boldsymbol{\theta})}{(2\pi)^{m/2}} \prod_{k=1}^{N} \frac{g(k)}{\det(\mathbf{S}_{k|k-1})^{1/2}}, \tag{20}$$

where $g(k) = \exp\{-0.5\tilde{\boldsymbol{y}}_k^\top \cdot \mathbf{S}_{k|k-1}^{-1} \cdot \tilde{\boldsymbol{y}}_k\}$, $\tilde{\boldsymbol{y}}_k$ is the residual measure innovation defined in equation (12), $\widehat{\boldsymbol{y}}_{k|k-1} = \mathbb{E}(\boldsymbol{y}_k|\mathcal{Y}_{k-1}, \boldsymbol{\theta})$ is the conditional mean of $\boldsymbol{y}_k$ given $\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_{k-1}$ and $\boldsymbol{\theta}$, and finally $\mathbf{S}_{k|k-1}$ is the conditional covariance innovation, defined in equation (13), given $\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_{k-1}$ and $\boldsymbol{\theta}$. Conditioning on $\boldsymbol{y}_0$, and considering the function:

$$\ell(\boldsymbol{\theta}) = -\ln(L(\boldsymbol{\theta}; \mathcal{Y}_k|\boldsymbol{y}_0)), \tag{21}$$

the maximum likelihood estimator of $\boldsymbol{\theta}$ can be obtained solving the following nonlinear optimization problem:

$$\widehat{\boldsymbol{\theta}} = arg \min_{\boldsymbol{\theta}} (\ell(\boldsymbol{\theta})). \tag{22}$$

State space formulation and EKF provide a powerful tool for the analysis of data in the context of maximum likelihood estimation. Let us remark that for a fixed $\boldsymbol{\theta}$, the values of $\tilde{\boldsymbol{y}}_k$ and $\mathbf{S}_{k|k-1}$, at each discrete timestep, can be obtained from the Kalman filter equations, described in section 3, and subsequently used in the construction of the log-likelihood function. In this context, the success of the optimization of log-likelihood function depends strictly on the behavior of the EKF designed.
In a first time, we give the evaluation of the Gaussian log-likelihood function bases on $\mathcal{Y}_r = [\boldsymbol{y}_{i_0}, \boldsymbol{y}_{i_1}, \ldots, \boldsymbol{y}_{i_r}]$, where $\mathcal{I}_r = [i_0, i_1, \ldots, i_r]$ are positive integers such that $0 \le i_0 < i_1 < \cdots < i_r \le N$. This allows for observation of the process at irregular intervals, or equivalently for the possibility that $(N-r)$ observations are missing from the sequence $\mathcal{Y}_N$.

To deal with possibly irregularly spaced observations or data with missing values (see [5]), we introduce a new series $\{\boldsymbol{Y}_k^*\}_k$, related to the state $\{\boldsymbol{X}_k\}_k$ by the modified observation equation (see equation (2)):

$$\boldsymbol{Y}_k^* = \boldsymbol{h}^*(\boldsymbol{X}_k, \boldsymbol{u}_k, \boldsymbol{\theta}) + \boldsymbol{\nu}_k^*, \qquad k = 0, 1, \ldots, \tag{23}$$

where

$$\boldsymbol{h}^*(\boldsymbol{X}_k, \boldsymbol{u}_k, \boldsymbol{\theta}) = \begin{cases} \boldsymbol{h}(\boldsymbol{X}_k, \boldsymbol{u}_k, \boldsymbol{\theta}) & \text{if } k \in \mathcal{I}_r, \\ \mathbf{0} & \text{otherwise,} \end{cases}$$

$$\boldsymbol{\nu}_k^* = \begin{cases} \boldsymbol{\nu}_k & \text{if } k \in \mathcal{I}_r, \\ \boldsymbol{\eta}_k & \text{otherwise,} \end{cases}$$

ISBN: 978-1-61804-141-8

and $\boldsymbol{\eta}_k$ is a i.i.d. Gaussian, zero-mean and white noise sequence with covariance matrix $\mathbb{E}(\boldsymbol{\eta}_k \cdot \boldsymbol{\eta}_s^T) = \mathbf{I}_{w\times w} \cdot \boldsymbol{\delta}_{k-s}$ (w is the dimension of noise $\boldsymbol{\nu}_k$ and $\boldsymbol{\eta}_k$ at each discrete timestep k). Moreover $\boldsymbol{\eta}_s \perp \boldsymbol{X}_0$, $\boldsymbol{\eta}_s \perp \boldsymbol{w}_k$, and $\boldsymbol{\eta}_s \perp \boldsymbol{\nu}_k$ $(s, k = 0, \pm 1, \dots)$.

Let us consider the joint Gaussian probability density function $L_1(\boldsymbol{\theta}; \mathcal{Y}_r)$ and the related Gaussian log-likelihood function $\ell_1(\boldsymbol{\theta})$ based on the measured values $\mathcal{Y}_r$. From these measured values, let us define the sequence $\mathcal{Y}_N^* = [\boldsymbol{y}_0^*, \boldsymbol{y}_1^*, \dots, \boldsymbol{y}_N^*]$ as follows:

$$\boldsymbol{y}_k^* = \begin{cases} \boldsymbol{y}_k & \text{if } k \in \mathcal{I}_r, \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{24}$$

Let $L_2(\boldsymbol{\theta}; \mathcal{Y}_N^*)$ be the joint Gaussian probability density function based on the values $\mathcal{Y}_N^*$ and let $\ell_2(\boldsymbol{\theta})$ be the related log-likelihood function. So it is clear that $L_1(\boldsymbol{\theta}; \mathcal{Y}_r)$ and $L_2(\boldsymbol{\theta}; \mathcal{Y}_N^*)$ are related by:

$$L_1(\boldsymbol{\theta}; \mathcal{Y}_r) = (2\pi)^{(N-r)w/2} L_2(\boldsymbol{\theta}; \mathcal{Y}_N^*), \tag{25}$$

and consequently $\ell_1(\boldsymbol{\theta})$ and $\ell_2(\boldsymbol{\theta})$ are related by:

$$\ell_1(\boldsymbol{\theta}) = \frac{\ln(2\pi)(N-r)w}{2} + \ell_2(\boldsymbol{\theta}). \tag{26}$$

Then we can now compute the required $L_1(\boldsymbol{\theta}; \mathcal{Y}_r)$ and the corresponding $\ell_1(\boldsymbol{\theta})$ of the realized sequence $\mathcal{Y}_r$, using the Kalman approach described above and applying equations (25) and (26).

5 Numerical results

Let us notice that for the treatment of missing data using QMLE technique is necessary to replace the unobserved data by other values (which is know as imputation methods). Two probabilities $1 - q$ (5% and 20%) of missing observation are considered. All the experiments are based on 50000 replications of an ARCH(1) model, with parameters $\boldsymbol{\theta} = (\alpha_0, \alpha_1)^\top = (0.3, 0.5)^\top$. Also let us assume a noise process covariance $\mathbf{Q} = \mathbf{I}_{1\times 1}$. The sample mean and standard error of the EKF estimator $\widehat{\boldsymbol{\theta}} = (\widehat{\alpha}_0, \widehat{\alpha}_1)^\top$ (new proposed estimator) are compared to those of estimators based on a complete data set obtained after filling the missing observations by some imputation procedures. Different imputation methods can be applied and we consider the following estimators obtained with the old methods (see [2] and [3]):

- $\widetilde{\boldsymbol{\theta}}_{\mathrm{b}} = (\widetilde{\alpha}_{0\,\mathrm{b}}, \widetilde{\alpha}_{1\,\mathrm{b}})^\top$, $\widetilde{\boldsymbol{\theta}}_{\mathrm{m}} = (\widetilde{\alpha}_{0\,\mathrm{m}}, \widetilde{\alpha}_{1\,\mathrm{m}})^\top$ and $\widetilde{\boldsymbol{\theta}}_{\mathrm{p}} = (\widetilde{\alpha}_{0\,\mathrm{p}}, \widetilde{\alpha}_{1\,\mathrm{p}})^\top$ are the QMLE estimators of $\boldsymbol{\theta}$ using different imputation methods, obtained respectively when each missing values is (i) deleted and the observed data are bound together, (ii) replaced by the sample mean of the observed values, and (iii) replaced by the first non missing value prior to it.

The Table 1 show the results for 2000 simulated observations. We can see the estimates for $\boldsymbol{\theta} = (\alpha_0, \alpha_1)^\top$ and the corresponding standard error in parentheses. We can see that the proposed nonlinear estimation method presents better performances than the classical imputation techniques considered here. Also our methodology is readily extended to other nonlinear time series models and nonstationary and asymmetric cases.

6 Conclusion

This article introduces a new nonlinear state space representation to characterize ARCH(p) ($p \geqslant 1$) time series models. Also an efficient numerical method, based on an EKF approach, for nonlinear parameter estimation of ARCH processes, has been presented. The

ISBN: 978-1-61804-141-8

$1-q$	5%	20%
$\widehat{\boldsymbol{\theta}}$	0.3001 (0.0030) 0.4991 (0.0135)	0.3001 (0.0033) 0.4994 (0.0160)
$\widetilde{\boldsymbol{\theta}}_{\mathrm{b}}$	0.308 (1.591e-2) 0.485 (4.625e-2)	0.334 (1.991e-2) 0.442 (5.297e-2)
$\widetilde{\boldsymbol{\theta}}_{\mathrm{m}}$	0.316 (2.041e-2) 0.432 (4.527e-2)	0.323 (2.620e-2) 0.303 (4.412e-2)
$\widetilde{\boldsymbol{\theta}}_{\mathrm{p}}$	0.293 (1.512e-2) 0.511 (4.397e-2)	0.267 (1.616e-2) 0.554 (4.329e-2)

Table 1: Sample mean (standard error in parentheses) of the different estimators when $n = 2000$.

framework that we propose it is valuable if the processes have unobserved values. The strategy of estimation associated with this representation allows computationally efficient parameter estimation.

Let us notice that quasi maximum likelihood methods are very accurate by estimating ARCH parameters in a complete time series. In these cases, the EKF involve a more complex procedure for the parameter estimation and the gain is not realy significant. On the other hand, it is well known that time series with missing values presents a serious problem to conventional parameter estimation methodologies such as QMLE. Studies show that, for linear time series models, the standard KF approach can be easily modified in order to obtain an efficient method to deal with missing observations. A natural extension, for nonlinear time series models, has been proposed in this work to treat problems with missing values. The numerical results presented herein demonstrate the effectiveness of this methodology, and show that it is more appropriate than other QMLE-imputation methods used in practice (see Table 1).

In conclusion, the methodology proposed in this article is an innovative and effective way to solve the problem of estimation of parameters in ARCH processes with missing observations.

Acknowledgements: The research was supported by the projects DII 037.397/2012 and DII 037.395/2012, Dirección de Investigación e Innovación, Pontificia Universidad Católica de Valparaíso.

References:

[1] Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity, J. Econometrics 31(3), 307-327.

[2] Bondon, P. and Bahamonde, N. (2012). Least squares estimation of ARCH models with missing observations. *J. Time Ser. Anal.* In Press.

[3] Bose, A. and Mukherjee, K. (2003). Estimating the ARCH parameters by solving linear equations. *J. Time Ser. Anal.*, 24(2):127-136.

[4] Engle, R. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50(4), 987-1007.

[5] Little, R. J. A. and Rubin, D. B. (2002). *Statistical analysis with missing data*. Wiley Series in Probability and Statistics. Wiley-Interscience John Wiley & Sons, Hoboken, NJ, second edition.

[6] Ossandón, S. and Bahamonde, N. (2011), On the Nonlinear Estimation of GARCH Models Using an Extended Kalman Filter, Lectures Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering 2011, WCE 2011, 6-8 July, 2011, London, U.K., pp 148-151.