Two-stage least absolute power deviation estimation for a general class of conditionally heteroskedastic models

Abdelhakim AKNOUCHE [1]

[1] *Faculté de Mathématiques, Université des sciences et de Technologie Houari Boumediene, Alger, Algérie, aknouche_ ab@yahoo.com.*

Résumé. Dans ce travail, nous proposons une méthode, dite des moindres déviations fonctionnelles absolues en puissances en deux étapes ($2S$-$LAPD$), pour l'estimation d'une classe générale de modèles conditionnellement hétéroscédastiques, comprenant notamment le modèle $GARCH$, le modèle $GARCH$ asymétrique en puissance et le modèle $ARCH$ infini. L'estimateur proposé est indexé par une fonction instrumentale dont le choix permet de contrôler et alléger les hypothèses sur les moments du processus d'innovation, hypothèses sur la base desquelles nous montrons consistance et normalité asymptotique (CAN) de l'estimateur $2S$-$LAPD$. Dans le cas d'une puissance du carré, l'estimateur $2S$-$LAPD$ possède la même variance asymptotique que le quasi-maximum de vraisemblance généralisé et ce pour certaines classes de fonctions instrumentales et même pour des innovations à queues lourdes et/ou asymétriques. De plus, pour une puissance unité, l'estimateur $2S$-$LAPD$ se réduit à des variantes en deux-étapes de l'estimateur des moindres déviations absolues ($2S$-LAD).

Mots-clés. Modèles conditionnellement hétéroscédastiques, moindres puissances de déviations absolues, moindres carrés pondérés en deux étapes, quasi-maximum de vraisemblance généralisé, moindres déviations absolues, consistance et normalité asymptotique.

Abstract. This work proposes a two-stage least absolute power functional deviation ($2S$-$LAPD$) method for estimating a general class of conditionally heteroskedastic models that includes, in particular, the $GARCH$ model, the asymmetric power $GARCH$ model and the infinite $ARCH$ model. The proposed estimate is indexed by an instrumental function which allows to control and weaken the innovation moment assumptions under which we establish consistency and asymptotic normality (CAN) of the $2S$-$LAPD$ estimate. In the case of power two, the $2S$-$LAPD$ estimate has the same asymptotic variance as the generalized quasi-maximum likelihood estimate for certain instrumental functions and even with heavy tailed and skewed innovations. Moreover, for a unit power, the $2S$-$LAPD$ estimate reduces to some two-stage variants of the least absolute deviation ($2S$-LAD) estimate.

Keywords. Conditionally heteroskedastic models, least absolute power deviation estimate, two-stage weighted least squares, generalized quasi-maximum likelihood estimate, least absolute deviation estimate, consistency and asymptotic normality.

1. Introduction

Consider an observable process $\{\epsilon_t, t \in \mathbb{Z}\}$ which is solution of the general *conditionally heteroskedastic* (CH) model

$$\begin{aligned} \epsilon_t &= \sigma_t \eta_t, & t \in \mathbb{Z}, \\ \sigma_t &= \sigma\left(\epsilon_{t-1}, \epsilon_{t-2}, ...; \theta_0\right), \end{aligned} \tag{1.1}$$

where $\{\eta_t, t \in \mathbb{Z}\}$ is a sequence of independent and identically distributed (iid) random variables with η_t independent of $\{\epsilon_i, i < t\}$, $\theta_0 \in \mathbb{R}^m$ is an unknown parameter belonging to a parameter space Θ and $\sigma : \mathbb{R}^\infty \times \Theta \to (0, \infty)$. Model (1.1) is a fairly general one since most of conditional volatility models known in practice may be cast in (1.1): stable $GARCH\,(p,q)$, stable asymmetric power $GARCH$ $(APGARCH\,(p,q))$, Linear $ARCH$, infinite $ARCH(\infty)$ (see e.g. Bardet and Wintenberger, 2009; Francq and Zakoïan, 2013 and the references therein).

Due to its many desirable properties, the Gaussian quasi-maximum likelihood estimate ($QMLE$ in short) has been the most common used method for estimating the parameters of most particular classes of (1.1). Two important objectives have motivated interest in studying the $QMLE$ for many classes of (1.1). The first one, of older interest, was to establish the consistency and asymptotic normality (CAN) property with minimal assumptions on the moments of the observed process $\{\epsilon_t, t \in \mathbb{Z}\}$, in particular without any moment requirement and even outside the strict stationarity domain. The second one, which has been the subject of more recent research activity, was to get the CAN property with, in addition, minimal moment assumptions on the innovation $\{\eta_t, t \in \mathbb{Z}\}$, in particular for η_t with heavy-tailed distribution.

The Gaussian $QMLE$ has been extensively studied to achieve the first objective and we know at present that the $QMLE$ for models (1.1) has the CAN property without any moment conditions on $\{\epsilon_t, t \in \mathbb{Z}\}$, but under the fourth moment requirement on $\{\eta_t, t \in \mathbb{Z}\}$. When the condition $E\left(\eta_1^4\right) < \infty$ is dropped, the Gaussian $QMLE$ would not have the CAN and even is inconsistent when $E\left(\eta_1^2\right) = \infty$.

For the second objective, robust estimates such as LAD-type estimate (Peng and Yao, 2003, Francq and Zakoïan, 2013) and M-estimate (Mukherjee, 2008) have been considered. In addition, the generalized $QMLE$ $(GQMLE)$ (Berkes and Horvàth, 2004; Francq and Zakoïan, 2013; Fan, Qi and Xiu, 2014; Zhu and Li, 2014) calculated on the basis of some "instrumental density" has been introduced as an alternative to the $QMLE$, especially when the assumption $E\left(\eta_1^4\right) < \infty$ is not necessarily satisfied. Indeed, under general and mild conditions on the moments of $\{\epsilon_t, t \in \mathbb{Z}\}$ and $\{\eta_t, t \in \mathbb{Z}\}$, it has been proved (Berkes and Horvàth, 2004; Francq and Zakoïan, 2013) that the $GQMLE$ has the CAN property. As an application, a one-step procedure based on the $GQMLE$ has been proposed by Francq and Zakoïan (2013) to get prediction of powers for the class of CH models given by (1.1).

In this work, we explore an alternative two-stage (functional) least absolute power deviation method for the general model (1.1). The proposed method depending on an instrumental function h, has the same properties as the $GQMLE$ for some classes of omnibus functions, under some quite mild conditions on $\{\epsilon_t, t \in \mathbb{Z}\}$ and $\{\eta_t, t \in \mathbb{Z}\}$ as well. Some advantages of the proposed method are:

- In some cases where the conditional volatility of the model is linear in the parameters ($ARCH$, Asymmetric Power $ARCH$), the proposed method has a closed-form contrary to the generalized $QMLE$.

- The general formulation of the method easily allows for some other generalizations and gives the link between many existing LAD, LS, WLS and M-type methods.

Thus, we establish consistency and asymptotic normality of the proposed method for the general conditional volatility model (1.1) under some mild assumptions. Application to some specific models such as: the $GARCH$ model, the asymmetric power $GARCH$ ($APGARCH$) model and $ARCH(\infty)$ model is considered.

2. Two-stage functional least squares estimate for conditionally heteroskedastic models

Let $\epsilon_1, \epsilon_2, ..., \epsilon_n$ be a series generated from model (1.1) which is subject to the following assumption.

A0: $\{\epsilon_t, t \in \mathbb{Z}\}$ *is a strictly stationary and ergodic solution of* (1.1).

For arbitrary initial values $\widetilde{\epsilon}_0, \widetilde{\epsilon}_{-1}, ...$, define

$$\widetilde{\sigma}_t(\theta) = \sigma\left(\epsilon_{t-1}, \epsilon_{t-2}, ..., \epsilon_1, \widetilde{\epsilon}_0, \widetilde{\epsilon}_{-1}, ...; \theta\right),$$

as a proxy for

$$\sigma_t(\theta) = \sigma\left(\epsilon_{t-1}, \epsilon_{t-2}, ..., \epsilon_1, \epsilon_0, \epsilon_{-1}, ...; \theta\right).$$

For any fixed known $\theta_1 \in \Theta$ and any function $h : \mathbb{R} \to \mathbb{R}$, the two-stage functional least squares estimate is given by

$$\begin{aligned}
\widehat{\theta}_{1,h} &= \arg\min_{\theta\in\Theta} \sum_{t=1}^{n} \left(h\left(\tfrac{\epsilon_t}{\widetilde{\sigma}_t(\theta_1)}\right) - h\left(\tfrac{\widetilde{\sigma}_t(\theta)}{\widetilde{\sigma}_t(\theta_1)}\right)\right)^2, \\
\widehat{\theta}_{2,h} &= \arg\min_{\theta\in\Theta} \sum_{t=1}^{n} \left(h\left(\tfrac{\epsilon_t}{\widetilde{\sigma}_t(\widehat{\theta}_{1,h})}\right) - h\left(\tfrac{\widetilde{\sigma}_t(\theta)}{\widetilde{\sigma}_t(\widehat{\theta}_{1,h})}\right)\right)^2,
\end{aligned}$$

for some compact space Θ. By analogy to the Generalized $QMLE$, h is called the *instrumental* function. To study the asymptotic properties of $(\widehat{\theta}_{1,h}, \widehat{\theta}_{2,h})$ let

$$g_{\theta_1}\left(\epsilon_t, \widetilde{\sigma}_t(\theta)\right) = \left(h\left(\tfrac{\epsilon_t}{\widetilde{\sigma}_t(\theta_1)}\right) - h\left(\tfrac{\widetilde{\sigma}_t(\theta)}{\widetilde{\sigma}_t(\theta_1)}\right)\right)^2,$$

and assume that:

A1: *For all* θ_1, $\left\{h\left(\frac{\epsilon_t}{\sigma_t(\theta_1)}\right)-h\left(\frac{\sigma_t(\theta_0)}{\sigma_t(\theta_1)}\right), t\in\mathbb{Z}\right\}$ *is a square-integrable martingale difference with respect to* $\{\mathcal{F}_t, t\in\mathbb{Z}\}$, *where* $\mathcal{F}_t=\sigma\{\epsilon_t,\epsilon_{t-1},...\}$.

For example, when $h(x)=|x|^r$, $r\neq 0$, assumption **A1** reduces to $E(|\eta_1|^r)=1$ and when $h(x)=\log|x|$, it becomes $E(\log(|\eta_1|))=0$. Define

$$\begin{aligned}
A_{1,h} &= 2E\left(\frac{1}{\sigma_t^2(\theta_1)}h_1^2\left(\frac{\sigma_t(\theta_0)}{\sigma_t(\theta_1)}\right)\frac{\partial\sigma_t(\theta_0)}{\partial\theta}\frac{\partial\sigma_t(\theta_0)}{\partial\theta'}\right)\\
B_{1,h} &= 4E\left(\frac{1}{\sigma_t^2(\theta_1)}\frac{\partial\sigma_t(\theta_0)}{\partial\theta}\frac{\partial\sigma_t(\theta_0)}{\partial\theta'}h_1^2\left(\frac{\sigma_t(\theta_0)}{\sigma_t(\theta_1)}\right)\left(h\left(\frac{\epsilon_t}{\sigma_t(\theta_1)}\right)-h\left(\frac{\sigma_t(\theta_0)}{\sigma_t(\theta_1)}\right)\right)\right)^2\\
J_{1,h} &= A_{1,h}^{-1}B_{1,h}A_{1,h}^{-1}
\end{aligned}$$

$$\begin{aligned}
A_{2,h} &= 2h_1^2(1)\,E\left(\frac{1}{\sigma_t^2(\theta)}\frac{\partial\sigma_t(\theta_0)}{\partial\theta}\frac{\partial\sigma_t(\theta_0)}{\partial\theta'}\right)\\
B_{2,h} &= 4h_1^2(1)\,E\left(h(\eta_t)-h(1)\right)^2E\left(\frac{1}{\sigma_t^2(\theta_1)}\frac{\partial\sigma_t(\theta_0)}{\partial\theta}\frac{\partial\sigma_t(\theta_0)}{\partial\theta'}\right)^2,\\
J_{2,h} &= A_{2,h}^{-1}B_{2,h}A_{2,h}^{-1}=\frac{E(h(\eta_1)-h(1))^2}{4h_1^2(1)}\left(E\left(\frac{1}{\sigma_t^4(\theta)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta}\frac{\partial\sigma_t^2(\theta_0)}{\partial\theta'}\right)\right)^{-1}.
\end{aligned}$$

Under **A0**-**A1** and some other additional assumptions we have the following result which establishes the CAN for $\widehat{\theta}_{1,h}$ and $\widehat{\theta}_{2,h}$.

Theorem 2.1

i) $\widehat{\theta}_{1,h}\to\theta_0$ *a.s.* *and* $\widehat{\theta}_{2,h}\to\theta_0$ *a.s.*

ii) $\sqrt{n}\left(\widehat{\theta}_{1,h}-\theta_0\right)\xrightarrow{\mathcal{L}}N(0,J_{1,h})$ *and* $\sqrt{n}\left(\widehat{\theta}_{2,h}-\theta_0\right)\xrightarrow{\mathcal{L}}N(0,J_{2,h})$.

When $h(x)=|x|^r$ and $r\neq 0$ (resp. $h(x)=\log|x|$ and $r=0$), $\widehat{\theta}_{2,h}$ has the same asymptotic distribution as the Generalized $QMLE$ with the instrumental distribution belonging to $\mathcal{C}(r)$ (cf. Francq and Zakoïan, 2013). So if the distribution of η_t belongs to $\mathcal{C}(r)$ then $\widehat{\theta}_{2,h}$ is asymptotically efficient.

3. Two-stage functional least power deviation estimate for conditionally heteroskedastic models

For any fixed known $\theta_1\in\Theta$ and any function $h:\mathbb{R}\to\mathbb{R}$, the two-stage functional least absolute power deviation estimate is given by

$$\left.\begin{aligned}
\widehat{\theta}_{1,h,s} &= \arg\min_{\theta\in\Theta}\sum_{t=1}^{n}\left|h\left(\frac{\epsilon_t}{\widetilde{\sigma}_t(\theta_1)}\right)-h\left(\frac{\widetilde{\sigma}_t(\theta)}{\widetilde{\sigma}_t(\theta_1)}\right)\right|^s,\\
\widehat{\theta}_{2,h,s} &= \arg\min_{\theta\in\Theta}\sum_{t=1}^{n}\left|h\left(\frac{\epsilon_t}{\widetilde{\sigma}_t(\widehat{\theta}_{1,h})}\right)-h\left(\frac{\widetilde{\sigma}_t(\theta)}{\widetilde{\sigma}_t(\widehat{\theta}_{1,h})}\right)\right|^s,
\end{aligned}\right\}\quad s>0.$$

$$\left.\begin{aligned}
\widehat{\theta}_{1,h,0} &= \arg\min_{\theta\in\Theta}\sum_{t=1}^{n}\log\left|h\left(\frac{\epsilon_t}{\widetilde{\sigma}_t(\theta_1)}\right)-h\left(\frac{\widetilde{\sigma}_t(\theta)}{\widetilde{\sigma}_t(\theta_1)}\right)\right|,\\
\widehat{\theta}_{2,h,0} &= \arg\min_{\theta\in\Theta}\sum_{t=1}^{n}\log\left|h\left(\frac{\epsilon_t}{\widetilde{\sigma}_t(\widehat{\theta}_{1,h})}\right)-h\left(\frac{\widetilde{\sigma}_t(\theta)}{\widetilde{\sigma}_t(\widehat{\theta}_{1,h})}\right)\right|,
\end{aligned}\right\}\quad s=0.$$

Some remarks are in order:

- When $s=2$ we find the two-stage functional least squares estimate given above.

- When $s=1$ and $h(x)=\log(x^2)$ we find a LAD estimate ($\widehat{\theta}_2$ of Peng and Yao, 2003) with only one stage

$$\widehat{\theta}_{1,h,1}=\widehat{\theta}_{2,h,1}=\arg\min_{\theta\in\Theta}\sum_{t=1}^{n}\left|\log\left(\epsilon_t^2\right)-h\left(\widetilde{\sigma}_t^2\left(\theta\right)\right)\right|.$$

- When $s=1$ and $h(x)=x^2$ the proposed estimate reduces to the following $2S$-LAD

$$\begin{array}{c}\widehat{\theta}_{1,x^2,1}=\arg\min_{\theta\in\Theta}\sum_{t=1}^{n}\left|\frac{\epsilon_t^2}{\widetilde{\sigma}_t^2(\theta_1)}-\frac{\widetilde{\sigma}_t^2(\theta)}{\widetilde{\sigma}_t^2(\theta_1)}\right|,\\ \widehat{\theta}_{2,x^2,1}=\arg\min_{\theta\in\Theta}\sum_{t=1}^{n}\left|\frac{\epsilon_t^2}{\widetilde{\sigma}_t^2\left(\widehat{\theta}_{1,h}\right)}-\frac{\widetilde{\sigma}_t^2(\theta)}{\widetilde{\sigma}_t^2\left(\widehat{\theta}_{1,h}\right)}\right|.\end{array}$$

- More generally, a two-stage functional LAD estimate is given by

$$\begin{array}{c}\widehat{\theta}_{1,h,1}=\arg\min_{\theta\in\Theta}\sum_{t=1}^{n}\left|h\left(\frac{\epsilon_t}{\widetilde{\sigma}_t(\theta_1)}\right)-h\left(\frac{\widetilde{\sigma}_t(\theta)}{\widetilde{\sigma}_t(\theta_1)}\right)\right|,\\ \widehat{\theta}_{2,h,1}=\arg\min_{\theta\in\Theta}\sum_{t=1}^{n}\left|h\left(\frac{\epsilon_t}{\widetilde{\sigma}_t\left(\widehat{\theta}_{1,h}\right)}\right)-h\left(\frac{\widetilde{\sigma}_t(\theta)}{\widetilde{\sigma}_t\left(\widehat{\theta}_{1,h}\right)}\right)\right|.\end{array}$$

Bibliography

[1] Aknouche, A. (2014). Estimation and strict stationarity testing of $ARCH$ processes based on weighted least squares. *Mathematical Methods of Statistics*, **23**, 81-102.

[2] Bardet, J-M. and Wintenberger, O. (2009). Asymptotic normality of the Quasi-maximum likelihood estimator for multidimensional causal processes. *The Annals of Statistics*, **37**, 2730-2759.

[3] Berkes, I. and Horvàth, L. (2004). The efficiency of the estimators of the parameters in $GARCH$ processes. *Annals of Statistics*, **32**, 633-655.

[4] Fan, J., Qi, L. and Xiu, D. (2014). Quasi Maximum Likelihood Estimation of $GARCH$ Models with Heavy-Tailed Likelihoods. *Journal of Business and Economic Statistics*, **32**, 178-191.

[5] Francq, C. and Zakoïan, J.M. (2013). Optimal predictions of powers of conditionally heteroskedastic processes. *Journal of the Royal Statistical Society*, **B75**, 345-367.

[6] Mukherjee, K. (2008). M-estimation in $GARCH$ models. *Econometric Theory*, **24**, 1530-1553.

[7] Peng, L. and Yao, Q. (2003). Least absolute deviations estimation for $ARCH$ and $GARCH$ models. *Biometrika*, **90**, 967-975.

[8] Preminger A. and Storti, G. (2014). Least squares estimation for $GARCH(1,1)$ model with heavy tailed errors. *$MPRA$ document*, N° 59082.

[9] Zhu, K. and Li, W.K. (2014). A new Pearson-type $QMLE$ for conditionally heteroskedastic models. *$MPRA$ document*, N° 52732.