文章编号：1004-4353(2014)02-0138-04

基于 MCMC 的分位回归 AR-ARCH 模型的贝叶斯分析

曾惠芳[1]， 熊培银[2]
（1. 湖南科技大学 商学院；2. 湖南科技大学 信息与电气工程学院：湖南 湘潭 411201）

摘要： 针对时间序列分布特征的高峰厚尾性，提出了一类分位回归 ARCH 模型. 在贝叶斯理论框架下，通过选择适当的先验分布，并基于非对称 Laplace 分布构建模型的似然函数，实现了模型的贝叶斯推断. 仿真试验和分析表明，该分位回归 ARCH 模型可全面刻画时间序列的非对称性和高峰厚尾性.

关键词： 贝叶斯；分位数；AR-ARCH 模型；仿真
中图分类号： F224.9；O212　　**文献标识码：** A

Bayesian analysis of the quantile AR-ARCH models based on MCMC algorithm

ZENG Huifang[1]， XIONG Peiyin[2]
（1. *College of Business*，*Hunan University of Science and Technology*；2. *School of Information and Electrical Engineering*，*Hunan University of Science and Technology*：*Xiangtan* 411201，*China*）

Abstract: Since many time series with asymmetric and heavier tails, we adapt the quantile regression ideas to the ARCH models. In the framework of Bayesian theory, we employ the proper prior, the likelihood function based on the asymmetric Laplace distribution was employed irrespective of the original distribution of the data, and derive the posterior distribution of the model parameters. The simulation result shows that the quantile ARCH models are effective to capture the diversity of time series distribution.

Key words: Bayesian; quantile; AR-ARCH models; simulation

金融市场波动性是近年来金融理论研究中较为活跃的一个课题，其中波动率是衡量金融资产风险的重要指标. 传统的金融时间序列模型假设收益率序列服从正态分布，但是，近年来的大量实证研究表明，金融市场上绝大多数资产收益率序列的特征往往无法用一些标准的金融时间序列模型来刻画，比如尖峰厚尾性、非对称性等，尤其是当条件方差不存在时，传统的 GARCH 模型很难实现对金融市场波动特征的刻画. 为了更全面和准确地描述金融时间序列的波动特征，研究者们提出了一些更加灵活和稳健的分位回归方法. 如：Koenker 等讨论了异方差模型的分位估计[1]，且提出了分位自回归条件异方差模型[2]，并推导了其估计量的渐近分布；Xiao 等[3]提出了分位数 GARCH 模型，并给出了该模型的两步估计方法以及估计量的渐近性质；Chen 等[4]提出了 ARCH 模型的一步估计，并提出了分位回归模型的格兰杰因果检验；王新宇等[5]利用 MCMC 方法对间接 TARCH-CAViaR 模型进行了贝叶斯分析，并分析了中国股市的风险价值. 虽然对 ARCH 模型的分位回归分析取得了一些成果，但它仍处于起步阶段，有

收稿日期： 2014－03－11　　**作者简介：** 曾惠芳(1981—)，女，博士，讲师，研究方向为贝叶斯计量经济、极值理论.
基金项目： 国家自然科学基金资助项目(41301421)；教育部人文社会科学研究项目(13YJC790203)

待于进一步研究. 本文讨论了分位回归 ARCH 模型的结构特征及其相应的估计方法,提出了 AR-ARCH 模型的两步估计方法,并对其进行了仿真验证分析.

1　模型结构分析

假设随机变量 $\{y_t\}$ 服从如下的 AR-ARCH 过程:

$$\begin{cases} y_t = a_0 + a_1 y_{t-1} + \cdots + a_p y_{t-p} + \varepsilon_t, \\ \varepsilon_t = \sigma_t e_t, \\ \sigma_t = \alpha_0 + \alpha_1 |\varepsilon_{t-1}| + \cdots + \alpha_q |\varepsilon_{t-q}|, \end{cases} \tag{1}$$

其中:$\alpha_0 > 0$; $(\alpha_1, \cdots, \alpha_q)' \in \mathbf{R}_+^q$; $\{e_t\}$ 是独立同分布的序列,其均值为 0 且方差 $\sigma^2 < \infty$,并且可假设残差 e_t 的分布未知. 因为方差 $\sigma_t > 0$,对于任意 t, y_t 的条件分位数函数满足单调性,所以 y_t 的条件分位数函数为 $Q_{y_t}(\tau \mid I_{t-1}) = a_0 + \sum_{i=1}^{p} a_i y_{t-i} + (\alpha_0 + \sum_{j=1}^{q} \alpha_j |\varepsilon_{t-j}|) F^{-1}(\tau)$,其中 $F^{-1}(\tau)$ 是 e_t 的 τ 分位数. 为了简便起见,本文考虑均值等于零的 ARCH 模型

$$y_t = \sigma_t e_t = (\alpha_0 + \alpha_1 |y_{t-1}| + \cdots + \alpha_q |y_{t-q}|) e_t, \tag{2}$$

因此,分位回归 ARCH 模型可用如下的随机系数分位回归 AR 来表示:

$$y_t = \sigma_t e_t = (\alpha_0 + \alpha_1 |y_{t-1}| + \cdots + \alpha_q |y_{t-q}|) e_t = \alpha_0(e_t) + \alpha_1(e_t) |y_{t-1}| + \cdots + \alpha_q(e_t) |y_{t-q}|, \tag{3}$$

因为 $|y_{t-j}|$ 以及系数都大于 0,所以它是关于 e_t 的单调递增函数. 相应地, y_t 的条件分位数函数为

$$Q_{y_t}(\tau \mid I_{t-1}) = \alpha_0(\tau) + \sum_{j=1}^{q} \alpha_j(\tau) |y_{t-j}|. \tag{4}$$

在经济系统中,经济变量是互相联系的,因此,在对资产收益率时间序列建模时,需要在分位回归 ARCH 模型中引入其他经济解释变量,即

$$y_t = a_0 + a_1 y_{t-1} + \cdots + a_p y_{t-p} + a_{p+1} x_{1t} + \cdots + a_{p+v} x_{vt} + \varepsilon_t, \tag{5}$$

并假设残差项为

$$\varepsilon_t = (\alpha_0 + \alpha_1 |\varepsilon_{t-1}| + \cdots + \alpha_q |\varepsilon_{t-q}| + \alpha_{q+1} |x_{1t}| + \cdots + \alpha_{q+v} |x_{vt}|) e_t, \tag{6}$$

其中 $(x_1, \cdots, x_v)'$ 表示其他经济解释变量, $\alpha_0 > 0$, $(\alpha_1, \cdots, \alpha_q, \alpha_{q+1}, \cdots, \alpha_{q+v})' \in \mathbf{R}_+^{q+v}$,并且假设残差 e_t 的分布 $F(\cdot)$ 为未知. y_t 的条件分位数函数为 $Q_{y_t}(\tau \mid I_{t-1}) = a_0 + \sum_{i=1}^{p} a_i y_{t-i} + \sum_{j=1}^{v} a_{p+j} x_{jt} + Z_t' \alpha(\tau)$,其中 $F^{-1}(\tau)$ 是 e_t 的 τ 分位数, $Z_t = (1, |\varepsilon_{t-1}|, \cdots, |\varepsilon_{t-q}|, |x_{1t}|, \cdots, |x_{vt}|)'$, $\alpha(\tau) = (\alpha_0 F^{-1}(\tau), \alpha_1 F^{-1}(\tau), \cdots, \alpha_{q+v} F^{-1}(\tau))'$. 首先通过普通最小一乘法(Least absolute deviations, LAD) 估计 $\hat{\varepsilon}_t = y_t - \hat{a}_0 - \sum_{i=1}^{p} \hat{a}_i y_{t-i} - \sum_{j=1}^{v} \hat{a}_{p+j} x_{jt}$,然后用 $\hat{\varepsilon}_t$ 代替(6) 式中的 ε_t,则有

$$\hat{\varepsilon}_t = (\alpha_0 + \alpha_1 |\hat{\varepsilon}_{t-1}| + \cdots + \alpha_q |\hat{\varepsilon}_{t-q}| + \alpha_{q+1} |x_{1t}| + \cdots + \alpha_{q+v} |x_{vt}|) e_t. \tag{7}$$

用分位回归方法估计参数 $\alpha(\tau)$,解决优化问题 $\hat{\alpha}(\tau) = \arg\min_\alpha \sum_{t=q+1}^{n} \rho_\tau(\hat{\varepsilon}_t - Z' \alpha_t)$ 可以实现对模型参数 $\alpha = (\alpha_0 F^{-1}(\tau), \alpha_1 F^{-1}(\tau), \cdots, \alpha_{q+v} F^{-1}(\tau))'$ 的分位回归估计.

2　模型的贝叶斯估计

为了实现分位回归 ARCH 模型的贝叶斯推断,可以假设模型为

$$y_t = Z_t' \alpha(\tau) + \xi_t, \tag{8}$$

其中 ξ_t 是服从非对称 Laplace 分布的随机变量, ξ_t 的 τ 分位数等于 0. 这样,分位回归 ARCH 模型的似然

函数可以表示为 $L(Y\mid\alpha)\propto\exp\left\{-\sum\limits_{t=q+1}^{n}\rho_\tau(y_t-Z_t'\alpha(\tau))\right\}$，其中 $Y=(y_1,y_2,\cdots,y_n)'$. 因为非对称 Laplace 分布可以表示为标准指数分布和标准正态分布的混合分布，即 $\xi_t=\varphi w_t+\lambda\sqrt{w_t}u_t$，所以分位回归 ARCH 模型可以表示为 $y_t=Z_t'\alpha(\tau)+\varphi w_t+\lambda\sqrt{w_t}u_t$，其中 w_t 服从标准指数分布，u_t 服从标准正态分布. y_t 可以认为是服从均值为 $Z_t'\alpha(\tau)+\varphi w_t$、方差等于 $\lambda^2 w_t$ 的正态分布，于是模型的似然函数可以表示为

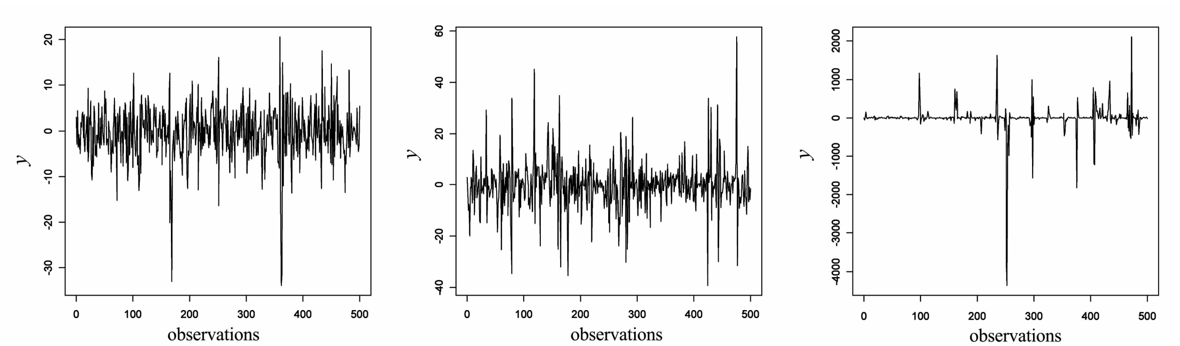
$$L(Y\mid\alpha)\propto\Big(\prod_{t=q+1}^{n}w_t^{-1/2}\Big)\exp\left\{-\sum_{t=q+1}^{n}\frac{(y_t-Z_t'\alpha(\tau)-\varphi w_t)^2}{2\lambda^2 w_t}\right\},\tag{9}$$

其中 $W=(w_{p+1},\cdots,w_n)'$. 假设参数 α 的先验分布为扩散先验分布，即 $\pi(\alpha)\propto 1$，根据贝叶斯公式，可以实现对参数 α 的贝叶斯分位回归估计.

3 仿真试验和分析

由随机数生成器生成服从 ARCH 过程的 500 个数据，并把它看作是一个随机系数 AR 过程，则 ARCH 的仿真过程为 $y_t=(3+0.5|y_{t-1}|)\varepsilon_t$，其中假定 ε_t 分别服从标准正态分布、自由度为 3 的 t 分布和标准柯西分布. 取 y_t 的条件分位数函数为 $Q_{y_t}(\tau\mid y_{t-1})=\alpha_0(\tau)+\alpha_1(\tau)|y_{t-1}|$，其中 $\alpha_0(\tau)=3F_\varepsilon^{-1}(\tau)$，$\alpha_1(\tau)=0.5F_\varepsilon^{-1}(\tau)$，$F_\varepsilon^{-1}(\tau)$ 表示 ε_t 的分位数.

图 1 分别给出了残差项服从标准正态分布、自由度为 3 的 t 分布和标准柯西分布的时间序列样本轨迹图. 从图 1 可以看出：扰动项服从柯西分布的时间序列，其异常点最多；其次是扰动项服从 t 分布的时间序列；正态分布的异常点相对较少.



(a) 正态分布 (b) t 分布 (c) 柯西分布

图 1 时间序列样本轨迹图

选择先验分布为均匀分布，可以由 M-H 算法模拟得到各个参数的边缘后验分布. 在模型运行的过程中，一共迭代了 11 000 次. 为确保参数估计的一致性，丢弃开始时的 1 000 次迭代，用 1 001 次到 11 000 次迭代得到的样本来估计参数. 图 2 给出了残差项服从标准正态分布时，模型参数在不同概率水平($\tau=0.05,0.25,0.5,0.75,0.95$)下的后验密度直方图. 由于篇幅所限，本文略去残差项服从自由度为 3 的 t 分布和标准柯西分布情况下的模型参数估计的后验密度直方图. 从图 2 可以看出，不同情况下参数的后验估计直方图都呈倒钟型，说明 MCMC 模拟过程是平稳的，即该 MCMC 抽样算法能有效地模拟分位回归 ARCH 过程中各参数的边缘后验分布.

表 1 给出了当分位数 $\tau=0.05,0.25,0.5,0.75,0.95$ 时参数的后验估计. 从表 1 可知，扰动项服从正态分布或 t 分布时，因为其本身是对称分布，所以参数的估计也是对称的. 当扰动项服从正态分布时，$\alpha_0(0.05)=-9.4150$，$\alpha_0(0.95)=9.8960$；当扰动项服从 t 分布时，$\alpha_0(0.05)=-14.8810$，$\alpha_0(0.95)=14.0202$. 由于 t 分布具有厚尾性，当分位数相等时，t 分布的尾部概率比正态分布更大. 因为柯西分布比 t 分布具有更厚的尾部特征，所以其分位数在尾部的绝对值更大. 模型参数的估计结果与真实值比较接近，这进一步说明了该方法的有效性.

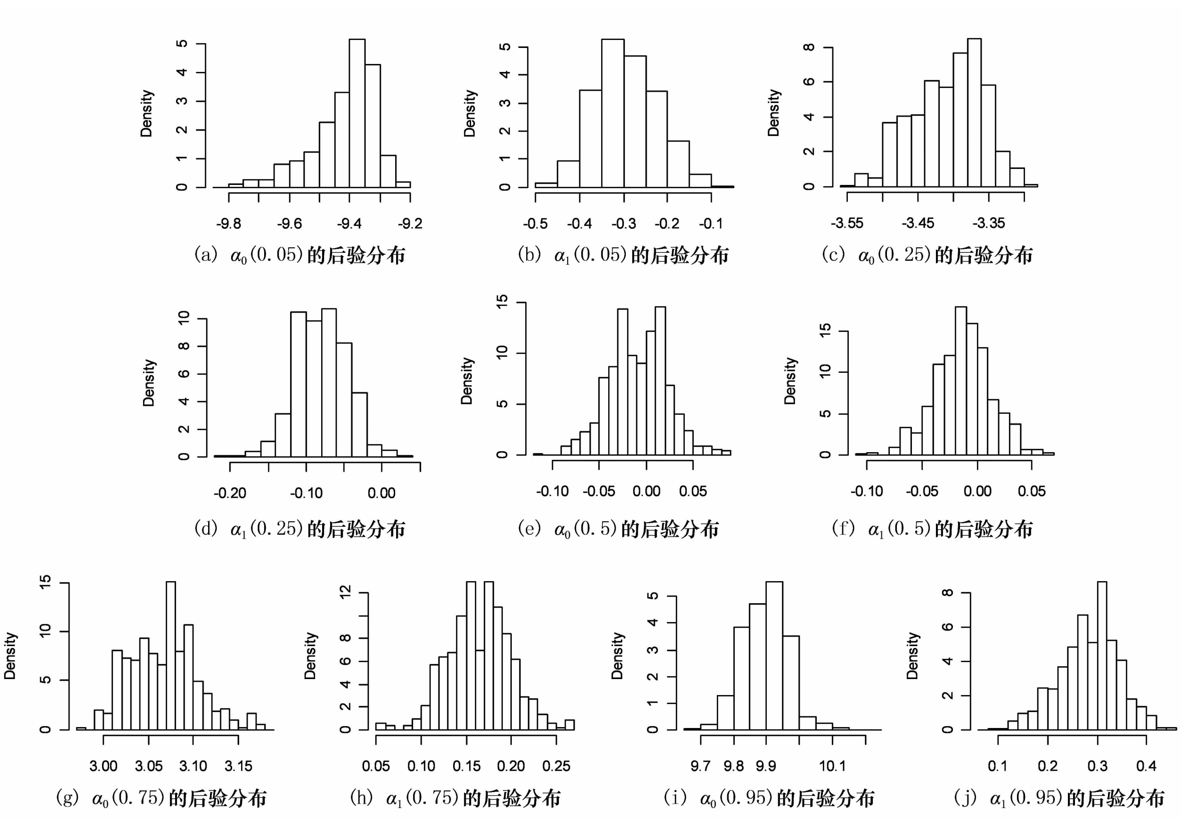


图 2　不同概率水平下参数的后验密度(残差项服从标准正态分布)

表 1　参数的贝叶斯分位回归估计

模型	参数	不同分位数时参数的后验估计				
		0.05	0.25	0.50	0.75	0.95
正态分布	$\alpha_0(\tau)$	−9.415 0	−3.406 0	−0.008 7	3.067 1	9.896 0
	$\alpha_1(\tau)$	−0.293 2	−0.080 0	−0.012 4	0.165 1	0.283 5
t 分布	$\alpha_0(\tau)$	−14.881 0	−4.529 0	0.147 0	3.858 0	14.020 2
	$\alpha_1(\tau)$	−0.498 4	−0.170 7	−0.065 5	0.145 9	0.479 3
柯西分布	$\alpha_0(\tau)$	−18.420 1	−10.840 0	−0.073 6	10.250 0	27.330 4
	$\alpha_1(\tau)$	−1.274 0	−0.361 4	−0.109 4	0.339 3	1.887 8

参考文献：

[1] Koenker R, Zhao Q. L-estimation for linear heteroscedastic models[J]. Journal of Nonparametric Statistics, 1994, 3:223-235.

[2] Koenker R, Zhao Q. Conditional quantile estimation and inference for ARCH models[J]. Econometric Theory, 1996,12:793-813.

[3] Xiao Z, Koenker R. Conditional quantile estimation for generalized autoregressive conditional heteroscedasticity models[J]. Journal of the American Statistical Association, 2009,104(488):1696-1712.

[4] Chen C W S, Gerlach R, Wei D C M. Bayesian causal effects in quantiles: accounting for heteroscedasticity[J]. Computational Statistics & Data Analysis, 2009,53:1993-2007.

[5] 王新宇，宋学锋. 间接 TARCH-CAViaR 模型及其 MCMC 参数估计与应用[J]. 系统工程理论与实践，2008(9):46-51.

[6] 曾惠芳，朱慧明，李素芳. 基于 MH 算法的贝叶斯分位自回归模型[J]. 湖南大学学报：自然科学版，2010,37(2):88-92.

[7] 朱慧明，王彦红，曾惠芳. 基于逆跳 MCMC 的贝叶斯分位自回归模型研究[J]. 统计与信息论坛，2010,25(1):9-13.