doi:10.16055/j.issn.1672-058X.2015.0011.003

基于傅立叶级数的半参数 CAViaR 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,湖南 湘潭 411201; 2.湖南科技大学 信息与电气工程学院,湖南 湘潭 411201)

摘　要:针对金融市场复杂性及不确定性,为了更灵活地测度金融市场的波动风险,提出了一类半参数 CAViaR 模型,利用傅立叶级数拟合前一期信息对当前 VaR 风险值的非线性影响,选择合适的先验分布,基于非对称 Laplace 分布构建了相应的似然函数,推导了参数的后验分布,从而实现了对 CAViaR 模型的贝叶斯推断;最后利用贝叶斯 CAViaR 模型研究了上海综合指数的风险波动特征,结果发现上证综指的风险波动存在自回归性.

关键词:贝叶斯分析;傅立叶级数;半参数方法;CAViaR 模型

中图分类号:F224.9;O212　　**文献标志码**:A　　**文章编号**:1672-058X(2015)11-0000-05

随着经济全球化及投资自由化,全球范围内,汇率、利率、股票价格和商品价格高度波动,呈现不断加剧的趋势.这些基础市场价格因子的高度波动,直接反映为金融市场风险的不断增大.自 20 世纪 90 年代以来,国际金融市场经历了一些重大的结构性变化,对金融市场风险产生了更为深远且深刻的影响.它们大大增加了金融市场和工具的关联度、复杂性、不确定性和波动性,使得金融市场风险上升,结构成分复杂,难以测量分析.

由于金融资产价格行为的复杂性,度量金融风险的方法越来越多,其中利用分位回归方法度量金融风险是目前研究的热点.分位回归方法不对金融时间序列的分布做任何假设,为有效地度量金融风险提供了一种灵活的方法.Taylor(1999)提出利用分位回归模型估计多期风险值的情形,并研究德国马克、英镑与日元汇率,发现分位回归模型用于估计多期风险值有良好的表现;Engle 和 Manganelli(2004)提出了非线性动态 VaR 模型(CAViaR),CAViaR 模型只描述条件分布尾部的行为,不受异常点的影响;刘新华,黄大山(2005)采用 Hansen 检验方法讨论了中国股市风险 CAViaR 建模的稳定性问题;张颖和孙和风(2012)提出了一类包含信息不对称性的 GJR-CAViaR 模型,并把它应用于刻画中美股票市场风险的差异;闫昌荣(2012)提出了一类流动性调整 CAViaR 模型,并计算了资产未来经过流动性调整的风险 VaR;王新宇等(2010)考虑到金融资产收益与正负收益对分位数冲击的不对称性,建立含有不对称绝对值和斜率设定的 AAVS-CAViaR 模型,对 1996-2008 年期间沪深港股票指数进行了实证研究;王新宇等(2013)提出了带有结构变点的条件分位数自回归模型,利用非对称拉普拉斯分布实现了模型的贝叶斯推断.此处将基于傅立叶级数构建一类半参数 CAViaR 模型,并实现对模型的贝叶斯推断,最后把贝叶斯 CAViaR 模型应用于中国股票市场的风险测度,发现中国股票市场的 VaR 风险存在自相关性.

收稿日期:2015-05-20;**修回日期**:2015-07-20.

* **基金项目**:国家自然科学基金项目(41301421).

作者简介:曾惠芳(1981-),女,湖南省邵阳人,讲师,博士,从事贝叶斯空间统计研究.

1 模型结构分析

Engle 和 Manganelli(2004)提出了 4 种 CAViaR 模型,分别是自适应 CAViaR,对称 CAViaR,非对称 CAViaR,GARCH(1,1)-CAViaR 模型.实际上,金融市场收益率的波动特征更丰富,而且 Engle 和 Ng(1993)曾经也提出了利用非参数方法,即线性样条方法,估计滞后收益率对波动率的影响.此处将提出一类新的更灵活的模型来度量 VaR 风险,即半参数 CAViaR 模型:

$$y_t = q_t(\tau) + \varepsilon_t = \beta_0 + \beta_1 q_{t-1}(\tau) + g(y_{t-1}) + \varepsilon_t \tag{1}$$

其中,y_t 可以表示金融时间序列的收益率,$q_t(\tau)=F_{y_t}^{-1}(\tau)$ 表示 y_t 的 τ 分位数函数,参数 $\beta_0>0$,β_1 为非负的自回归系数,$g(y_{t-1})$ 表示 y_{t-1} 未知的非线性函数,ε_t 为扰动项.用傅立叶级数方法来估计半参数 CAViaR 模型,并与线性样条方法进行比较.根据 Lenk(1999)的研究,可以利用傅立叶级数来近似式(1)的非参数部分,即利用余弦函数构成连续函数 g 在区间$[a,b]$上的正交基.

$$g(y_{t-1}) = \sum_{k=1}^{\infty} \theta_k \phi_k(y_{t-1}), \quad a \leqslant y_{t-1} \leqslant b \tag{2}$$

其中,

$$\phi_k(y_{t-1}) = \left(\frac{2}{b-a}\right)^{1/2} \cos\left\{\pi k\left(\frac{y_{t-1}-a}{b-a}\right)\right\},\ k=1,2,\cdots$$

$$\theta_k = \int_a^b g(y_{t-1})\phi_k(y_{t-1})\mathrm{d}y_{t-1}$$

为简便起见,可利用式(2)中的前 k 项近似 $g(y_{t-1})$,并记 $\boldsymbol{\Theta}=(\theta_{11},\cdots,\theta_k)$,这样可以通过解如下最优化问题来实现模型的估计

$$(\hat{\beta}_0,\hat{\beta}_1,\hat{\boldsymbol{\Theta}}) = \underset{\beta_0,\beta,\theta}{\arg\min}\, \rho \sum_{t=1}^{n} \Big(y_t - \beta_0 - \beta_1 q_{t-1}(\tau) - \sum_{k=1}^{K} \theta_k \phi_k(y_{t-1})\Big) \tag{3}$$

其中 $\rho_\tau(u)=u(\tau-I(u<0))$.因为分位回归估计最小化问题可以等价于以非对称 *Laplace* 分布为似然函数的极大似然估计问题,因此似然函数可以表示为

$$L(\boldsymbol{Y}|\beta_0,\beta_1,\theta) \propto \tau^n(1-\tau)^n \exp\Big\{-\sum_{t=1}^{n}\rho_\tau\Big(y_t - \beta_0 - \beta_1 q_{t-1}(\tau) - \sum_{k=1}^{K}\theta_k\phi_k(y_{t-1})\Big)\Big\} \tag{4}$$

其估计量可以看作是分位回归模型的伪极大似然估计,其中 $\boldsymbol{Y}=(y_1,\cdots,y_n)$.

为了实现模型的贝叶斯估计,首先给出参数的先验分布.因为对函数 g 进行平滑时会依赖于分位数 τ,一般在极值分位水平下,数据比较少,需要较大的光滑.因此可以设计参数 g 的先验分布为

$$\theta_k \sim N\{0,\delta^2\exp(-\gamma c_k)\}$$

其中,

$$c_k = \begin{cases} k, & \text{几何平滑} \\ \log(k), & \text{代数平滑} \end{cases} \tag{5}$$

$$\gamma = \begin{cases} \tau, & \text{几何平滑} \\ 1/\tau, & \text{代数平滑} \end{cases} \tag{6}$$

式(6)可以保证傅立叶级数以概率收敛于被拟合的函数.从式(5)来看,傅立叶级数系数的先验分布的方差有两种形式.代数平滑器平滑程度集中于 γ 的函数,而几何平滑器先验平滑程度由解析函数控制.参数 δ^2 控制着未知函数 g 的全局不确定性,即用于确定先验分布和似然函数之间的平衡关系.参数 γ 确定了傅立叶系数的衰减速度,因此能起到对函数 g 平滑的作用.显然,其他基函数以及先验分布可以用于模拟函数 g,并且假设 δ^2 独立于$\{\theta_k\}$,其先验分布为服从参数为 $r_0/2$,$s_0/2$ 的逆伽玛分布,可以表示为

$$\delta^2 \sim IG(r_0/2, s_0/2)$$

2　模型的贝叶斯推断

以非对称 Laplace 分布为基础的非参数分位回归模型的贝叶斯推断很难得到傅立叶系数的解析后验分布,即无法得到解析的后验均值和后验方差,以及无法得到参数解析的完全条件分布,不能用 Gibbs 抽样来实现对模型的模拟,引入标准指数分布和标准正态分布的混合分布,从而可以得到模型参数解析的完全条件分布,实现模型的 Gibbs 抽样.

假设 z 服从标准指数分布,u 服从标准正态分布,如果变量 ε 服从非对称 Laplace 分布,其密度函数表示为

$$f_\tau(\varepsilon)=\tau(1-\tau)\exp\{-\rho_\tau(\varepsilon)\}$$

那么 ε 可以表示为

$$\varepsilon=\varphi z+\omega\sqrt{z}u$$

其中 $\varphi=(1-2\tau)/\tau(1-\tau)$,$\omega^2=2/\tau(1-\tau)$.这样,半参数 CAViaR 模型可以表示为

$$y_t=\beta_0+\beta_1 q_{t-1}+\sum_{k=1}^{K}\theta_k\phi_k(y_{t-1})+\varphi z_t+\omega\sqrt{z_t}u_t \tag{7}$$

为了简便起见,式(7)可以表示为

$$\boldsymbol{Y}=\beta_0\boldsymbol{I}+\beta_1\boldsymbol{q}+\boldsymbol{\Phi\Theta}+\varphi Z+\sqrt{\sum}\boldsymbol{U} \tag{8}$$

其中 $\boldsymbol{Y}=(y_1,\cdots,y_n)'$,$\boldsymbol{I}=(1,\cdots,1)'_{n\times1}$,$\boldsymbol{q}=(q_0,\cdots,q_{n-1})'$,$\boldsymbol{\Theta}=(\theta_1,\cdots,\theta_K)'$,$\boldsymbol{Z}=(z_1,\cdots,z_n)'$,$\sum=\mathrm{diag}(\omega^2z_1,\cdots,\omega^2z_n)$,$\boldsymbol{U}=(u_1,\cdots,u_n)'$,以及 $\mathrm{Var}(\Theta)=\delta^2\Psi$,$\Psi$ 是以 $\exp(-\gamma c_k)$ 为第(k,k)对角元素的对角矩阵.

这样,在一定的条件下 Y 条件分布可以表示为以 $\beta_0\boldsymbol{I}+\beta_1\boldsymbol{q}+\boldsymbol{\Phi\Theta}+\varphi Z$ 为均值,以 $\sum$ 为方差的正态分布,因此 Y 的联合分布可以表示为

$$L(\boldsymbol{Y}|\beta_0,\beta,\theta)=\frac{1}{(\sqrt{2\pi})^n\left|\sum\right|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{Y}-\beta_0\boldsymbol{I}-\beta_1\boldsymbol{q}-\boldsymbol{\Phi\Theta}-\varphi Z)'\sum{}^{-1}(\boldsymbol{Y}-\beta_0\boldsymbol{I}-\beta_1\boldsymbol{q}-\boldsymbol{\Phi}\theta-\varphi Z)\right\} \tag{9}$$

3　实证分析

金融风险的度量方法有很多,要建立一个金融风险的测度机制,研究金融市场的动态特征非常重要.迄今为止,描述金融市场动态特征的模型当首推广义自回归条件异方差模型(GARCH).为了研究上证综指周数据的风险值,样本选择的时间段为 2006 年 1 月到 2011 年 2 月.图 1 给出了周收益率的波动轨迹.

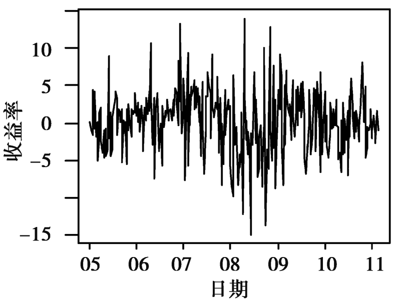

图 1　上证综指周收益率的波动轨迹

2008 年的全球金融危机对中国经济也产生了巨大的影响.在度量中国股票市场的风险时,金融危机对中国股市的影响不容忽视.为了更灵活地度量中国股票市场的风险,此处将采用半参数 CAViaR 模型来度量中国股票市场的风险,利用三阶傅立叶级数拟合模型中的非参数部分,用一条马尔可夫链模拟模型参数的后验分布,将最初的 4 000 次迭代得到的样本舍去,用第 4 001 次到第 10 000 次迭代得到 6 000 个模拟样本去估计模型参数.图 2 给出了模型参数的后验分布核密度函数.图 2 密度曲线表现比较平滑且呈钟型,说明 MCMC (Markov Chain Monte Carlo)算法有效地模拟了模型中各参数的边缘后验分布.表 1 给出了半参数 CAViaR 模型参数的后验均值、标准方差、蒙特卡罗误差和贝叶斯置信区间.

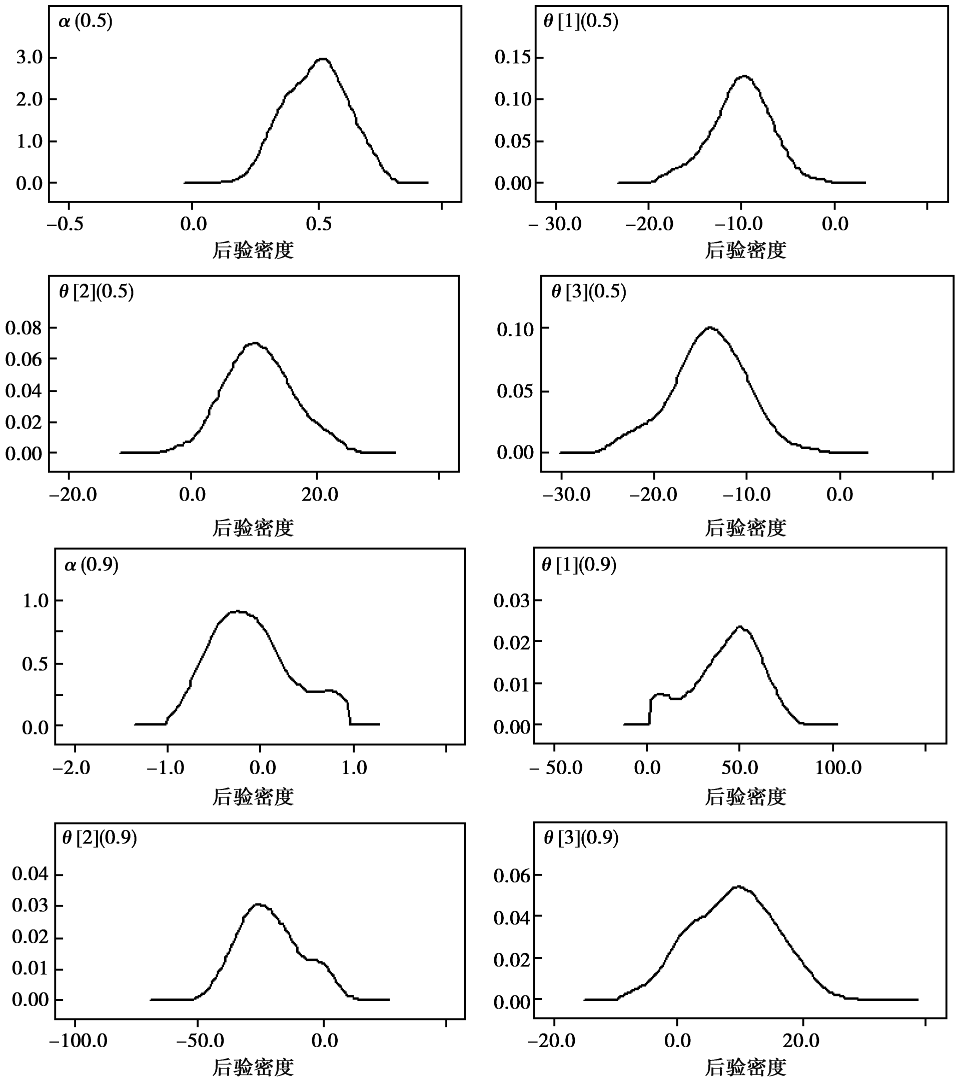

图 2 参数的后验核密度函数

表 1 参数的贝叶斯分位估计

参　数	均　值	标准差	MC 误差	0.025分位数	0.975分位数
$\alpha(0.5)$	0.465 3	0.117 4	0.014 0	0.270	0.720
$\theta_1(0.5)$	−9.929 0	3.230	0.296 0	−17.06	−3.755
$\theta_2(0.5)$	11.060 0	5.555	0.503 0	0.259	22.570
$\theta_3(0.5)$	−13.920 0	4.140	0.384 5	−22.920	−5.842
$\alpha(0.9)$	−0.014 3	0.496 9	0.046 8	−0.796	0.872
$\theta_1(0.9)$	42.840	18.330	1.529	4.994	72.620
$\theta_2(0.9)$	−20.620	12.580	1.008	−42.090	4.679
$\theta_3(0.9)$	9.140	6.939	0.507 5	−4.494	22.29

根据表 1 可知,在分位水平 $\tau=0.5$ 时,参数部分自回归系数 α 后验均值的估计为0.465 3,95%置信区间为(0.270,0.72);非参数部分傅立叶级数 θ_1 后验均值的估计为−9.929,95%置信区间为(−17.06,−3.755);θ_2 后验均值的估计为11.06,95%置信区间为(0.256,22.57),θ_3 后验均值的估计为−13.92,95%置信区间为(−22.92,−5.842).在分位水平 $\tau=0.9$ 时,参数部分自回归系数 α 后验均值的估计为−0.014 3,95%置信区间

为(-0.796,0.872);非参数部分傅立叶级数 θ_1 后验均值的估计为42.84,95%置信区间为(4.994,72.62);θ_2 后验均值的估计为-20.62,95%置信区间为(-42.09,4.679),θ_3 后验均值的估计为9.14,95%置信区间为(-4.494,22.29).从结果可以看出,VaR 过程存在自回归过程,即滞后一阶的风险对目前的风险会有影响,同时,滞后一阶的冲击对当前的风险也存在非线性的影响.

4　结　论

提出了一类基于傅立叶级数的半参数 CAViaR 模型,并给出了模型的贝叶斯推断和 MCMC 抽样估计算法,对于灵活地测度金融市场的 VaR 风险值具有重要的意义.把贝叶斯半参数 CAViaR 模型应用于中国股票市场的风险测度,结果发现中国股票市场的 VaR 风险存在自相关性和羊群效应.但是,如何把极值理论与非参数方法结合起来估计金融风险值是未来研究的方向.

参考文献:

[1] TAYLOR J W.A Quantile Regression Approach to Estimating the Distribution of Multiperiod Returns[J].Journal of Derivatives,1999(7):64-78

[2] ENGLE R, MANGANELLI S. Conditional Autoregressive Value at Risk by Regression Quantiles[J]. Journal of Business & Economic Statistics,2004(22):367-381

[3] 刘新华,黄大山.中国股市风险 CAViaR 方法的稳定性分析及其时变建模[J].系统工程理论与实践,2005(3):1-6

[4] 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,2008,28(9):46-51

[5] 叶五一,缪柏其,谭常春.基于分位点回归模型变点检测的金融传染分析[J].数量经济技术经济研究,2007(10):151-160

[6] 王新宇,赵绍娟.基于分位数回归模型的沪深股市风险测量研究[J].中国矿业大学,2008,37(3):416-421

[7] 关静,史道济.分位数回归与上证综指 VaR 研究[J].统计与信息论坛,2008,23(12):15-19

[8] 钱争鸣,郭鹏飞.上海证券交易市场量价关系的分位回归分析[J].数量经济技术经济研,2007(10):141-150

[9] KUESTER K, MITTIK S, PAOLELLA M S. Value-at-risk Prediction: A Comparison of Alternative Strategies[J]. Journal of Financial Econometrics,2005,4(1):53-89

[10] 张颖,孙和风.中美股票市场风险差异的新解释——收益对市场风险不对称效应的 CAViaR 模型与实证[J].南开经济研究,2012(5):111-120

[11] 闫昌荣.基于流动性调整 CAViaR 模型的风险度量方法[J].数量经济技术经济研究,2012(3):131-140

[12] 王新宇,宋学峰,吴瑞明.基于 AAVS-CAViaR 模型的股市风险测量研究[J].系统工程学报,2010,25(3):326-333

[13] 王新宇,吴祝武,宋学锋.变点 CAViaR 市场风险测量模型及创业板应用[J].中国矿业大学学报,2013,42(3):506-512

Bayesian Analysis of Semi-parametric CAViaR Model Based on Fourier Series

ZENG Hui-fang[1], XIONG Pei-gen[2]

(1.Business School,Hunan University of Science and Technology,Hunan Xiangtan 411201,China; 2.School of Information and Electrical Engineering,Hunan University of Science and Technology,Hunan Xiangtan 411201,China)

Abstract: According to the complexity and uncertainty in financial market,in order to more flexibly measure the volatility risk of financial market,this paper presents a class of semi-parametric CAViaR model by using the non-linear influence of former term fitting information of Fourier series on current VaR risk value to select suitable priori distribution, the related likelihood function is constructed based on asymmetric Laplace distribution, the posterior distribution of the parameters is deduced, and thus, the Bayesian deduction for CAViaR model is implemented.Finally,the feature of the risk volatility of Shanghai Composite Index is studied based on Bayesian CAViaR model and its results show the risk volatility of Shanghai Composite Index has auto-regression.

Key words: Bayesian analysis; Fourier series; semi-parametric method; CAViaR model