

VaR Method Based on HMM-GARCH Model and Its Application in Agricultural Stock Market

Lanlan Rong¹, Shuang Chen², Ru Feng¹

¹School of Science, Hebei University of Technology, Tianjin

²Department of Mathematics and Physics, Beijing Institute of Petrochemical Technology, Beijing

Email: hebutchenshuang@163.com

Received: Sep. 5th, 2017; accepted: Sep. 20th, 2017; published: Sep. 26th, 2017

Abstract

In this paper, the HMM and GARCH models are used to establish the HMM-GARCH model to measure the financial asset risk value (VaR) by using the advantage of the hidden Markov model in state division. First, the hidden Markov model of the financial asset return sequence is set up. The Baum-Welch algorithm is used to estimate the parameters of the model. Then, the Viterbi algorithm is used to estimate the corresponding hidden state sequence of the return sequence. According to the hidden state sequence, the return sequence is classified in two categories. And the HMM-GARCH model is established to estimate VaR for each state sequence respectively. Finally, the Beidahuang stock (600598) data are analyzed by using the GARCH model, MRS-GARCH model and HMM-GARCH model respectively. The accuracy of the VaR is tested by Kupiec failure frequency method. The results show that the HMM-GARCH model can better describe and predict the risk of the stock.

Keywords

HMM-GARCH Model, MRS-GARCH Model, VaR Method, Kupiec Failure Frequency Test

基于HMM-GARCH模型的VaR方法及其在农业股市的应用

容兰兰¹, 陈爽², 冯茹¹

¹河北工业大学理学院, 天津

²北京石油化工学院数理系, 北京

Email: hebutchenshuang@163.com

收稿日期：2017年9月5日；录用日期：2017年9月20日；发布日期：2017年9月26日

摘要

文章利用隐马尔科夫模型在状态划分上的优势，将隐马尔科夫模型(HMM)与GARCH模型结合建立HMM-GARCH模型来度量金融资产风险价值(VaR)。首先对收益率序列建立隐马尔科夫模型，将股票市场分为正常状态与异常状态，用Baum-Welch算法估算模型的参数，再采用Viterbi算法估算收益率序列所对应的隐状态序列，根据隐状态序列将收益率序列分为两类，分别建立HMM-GARCH模型估算VaR。最后利用GARCH模型、MRS-GARCH模型、HMM-GARCH模型对北大荒股票(600598)数据进行了实证分析，采用Kupiec-失败频率检验法对估计的VaR值进行检验。实证结果表明，基于HMM-GARCH模型的VaR方法能更好的描绘和预测该股票的风险。

关键词

HMM-GARCH模型，MRS-GARCH模型，VaR方法，Kupiec-失败频率检验法

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

自从2008年金融危机爆发以来，广大投资者和研究者都希望金融市场能具有提醒投资者和管理者抵抗非预期风险的能力。随着我国资本市场的进一步开放，我国的金融机构面临的风险将更加复杂，金融市场管理、金融监管原则和技术都要和国际接轨[1]。VaR方法作为国际上衡量金融风险的领先技术，必将给我国的风险管理工作带来一定的借鉴，在风险管理的实践中产生很大的现实意义。

1994年JP Morgan公布了Riskmetrics体系，对VaR模型的原理和计算方法进行了系统的总结，由于VaR计算简单，易于理解及其合理性，很快受到推崇和研究。Riskmetrics在第一版技术文件中对VaR的原理做了详细的阐述，包括历史模拟法、参数法、蒙特卡罗模拟法等具体模型原理的说明[2]；在第二版技术文件中，对非线性期权头寸的风险计量方法与如何处理实际分布的非正态性进行了补充[3]；后又分别对理论界与实务界提出的问题与建议进行了研究和改进工作[4]。2001年，巴塞罗那委员会正式指定VaR方法作为银行风险度量的标准。新Basle协议依据VaR方法对跨国银行的市场风险提出了具体的监管措施，使VaR方法成为金融风险管理的国际标准和理论基础。Zhang等在研究投资组合风险时，引入连接函数(Copula)方法，并与历史模拟法、均值-方差法、DCC-GARCH方法进行了比较[5]。Feroni在混合频率VaR中引入状态转换方法和马尔科夫模型，提出了MSMF-VaR，通过实证表明在国民生产总值和经济预测方面效果较好[6]。Marcucci通过比较不同的GARCH模型在持有期为一天到一个月内对股市波动率的预测能力，发现在短于一星期的持有期内，具有马尔科夫结构转换的GARCH模型(MRS-GARCH模型)比单一结构的GARCH模型对股市的VaR的估计效果更准确[7]。

国内对VaR方法的研究起步较晚，最早研究是郑文通的《金融风险管理的VaR方法及其应用》。我国许多学者主要应用波动模型来描述金融市场的波动性，估计VaR值，对我国的金融市场风险进行深入分析。主要采用的波动模型包括GARCH模型、EGARCH模型、MRS-GARCH模型等[8]。范英通过对VaR方法的深入研究，并对股票市场在不同置信水平下的风险值得实例计算，表明将VaR方法用于我国

股票市场的投资风险分析是可行的[9]。魏捷等以我国上证指数为例探讨了用 GARCH 模型计算 VaR 的新方法[10]。郭名媛和张世英提出了基于 MRS-GARCH 模型的 VaR 方法,并对上海股票市场的风险进行了实证分析,结论证明 MRS-GARCH 模型计算 VaR 值更加有效[11]。罗健英等针对中国钢铁期货市场波动率具有的结构突变特征,使用 MRS-GARCH 模型对其波动率建模,实证结果表明基于 MRS-GARCH 模型下的 VaR 方法能够有效地测度出钢铁期货市场的风险变化[12]。

这些模型大都没有考虑波动结构变化, MRS-GARCH 模型考虑了结构波动变化,在一定程度上可以对波动率结构发生突变前后的不同状态进行建模,从而得到不同状态下的定量模型,但是 MRS-GARCH 模型在进行参数估计时需要先用 GARCH 模型对时间序列的数据进行模拟估计波动率,再根据波动率大致的范围主观确定一个基准对波动率状态进行划分,最后用 MRS-GARCH 模型对样本数据再进行一次模拟。此种划分方法难免存在主观性差异,影响模型对样本数据模拟的精确度。

因此本文采用隐马尔科夫结构转换 GARCH 模型,即 HMM-GARCH 模型(Hidden Markov regime switching GARCH model)来度量我国金融市场的波动,并估计 VaR 值,希望为相关领域的研究提供一个参考方案。

2. 波动率模型及其估计方法

2.1. GARCH 模型

GARCH 模型是由 Bollerslev [13]基于 ARCH 模型提出的一个推广形式,广义 ARCH 模型。由于 GARCH 模型能够捕捉金融时间序列中一些重要的特征,例如,随时间变化的异方差性和波动率聚集性,使其成为波动率研究中一个标准方法。

一般的 GARCH 模型:

$$\begin{aligned} r_t &= \phi_0 + \sum_{i=1}^{p_1} \phi_i r_{t-i} + a_t - \sum_{j=1}^{q_1} \phi_j a_{t-j} \\ a_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \omega + \sum_{i=1}^{p_2} \alpha_i a_{t-i}^2 + \sum_{j=1}^{q_2} \beta_j \sigma_{t-j}^2 \end{aligned} \quad (1)$$

其中 r_t 表示金融时间序列, $\{\varepsilon_t\}$ 是均值为 0、方差为 1 的独立同分布随机变量序列, σ_t^2 是条件方差,即条件波动率。 $\omega > 0, \alpha_i \geq 0, \beta_j \geq 0, \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$ (这里对 $i > p, \alpha_i = 0$, 对 $j > q, \beta_j = 0$)。对 $\alpha_i + \beta_i$ 的限制条件是保证 a_t 的无条件方差是有限的,同时条件方差 σ_t^2 是随时间变化的。通常假定 ε_t 服从标准正态分布或标准化的学生-t 分布或广义误差分布(GED)。

2.2. MRS-GARCH 模型

由于 GARCH 模型在对波动率进行预测时,假设条件波动率在整个时期只有一个机制,没有考虑到突发事件和预期的改变等冲击可能会使影响金融资产价格的结构发生显著变化,前后往往表现出不同的特征,为了反映这种结构性变化,将马尔科夫状态转换方法引入波动率模型中,刻画了金融资产在不同状态下的特征及其转换行为。MRS-GARCH 模型的形式为:

$$\begin{aligned} r_t &= \phi_0 + \sum_{i=1}^{p_1} \phi_i r_{t-i} + a_t - \sum_{j=1}^{q_1} \phi_j a_{t-j} \\ a_t &= \varepsilon_t \sigma_t^2 \\ \sigma_t^2 &= \omega_{s_t} + \sum_{i=1}^{p_2} \alpha_{s_t,i} a_{t-i}^2 + \sum_{j=1}^{q_2} \beta_{s_t,j} \sigma_{t-j}^2 \end{aligned} \quad (2)$$

s_t 表示 r_t 在 t 时刻的状态, $\omega_{s_t}, \alpha_{s_t, i}, \beta_{s_t, j}$ 都是依赖于 t 时刻状态为 s_t 的参数。

状态之间的转移服从一阶马尔科夫过程:

$$p_{mn} = P(s_{t+1} = n | s_t = m, s_{t-1} = k, \dots, r_t, r_{t-1}, \dots) = P(s_{t+1} = n | s_t = m) \quad (3)$$

p_{mn} 表示状态 m 到状态 n 的一步转移概率。 k 表示状态的个数。本文只考虑两个状态。则转移概率矩阵为:

$$P = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix}$$

且有 $\sum_{n=1}^2 p_{mn} = 1, 0 \leq p_{mn} \leq 1, (m=1, 2)$ 。

对于 MRS-GARCH 模型的估计一般使用极大似然估计, 待估参数为 $\theta = (\mu, \omega_1, \omega_2, \alpha_1, \alpha_2, \beta_1, \beta_2, p_{11}, p_{22})$ 。为得到样本的对数似然函数, 需要进行滤波过程。对数似然函数为 $L(\theta | a) = \sum_{t=2}^T \ln f(a_t | \theta, \Omega_{t-1})$, 其中 $a = (a_1, a_2, \dots, a_T)'$, $\Omega_{t-1} = (a_{t-1}, a_{t-2}, \dots, a_1)$ 表示 $t-1$ 时刻的信息集。通过计算可以得到极大似然估计 $\hat{\theta}$ 。

2.3. HMM-GARCH 模型

2.3.1. 隐马尔科夫模型的理论知识

隐马尔科夫模型是由两个随机过程组成的一个双随机过程, 用来描述一个含有未知参数的马尔科夫过程。主要从可观察到的变量中确定该过程的隐含参数, 然后再用这些参数进行下一步分析。

隐马尔科夫模型可以由一个五元组 $\lambda = (M, N, A, B, \Pi)$ 表示, 其中:

N 表示隐含状态数量。隐含状态之间满足马尔科夫性质, 是马尔科夫模型中实际隐含的状态。

M 表示可观察状态数目。可观察状态在模型中与隐含状态相关联, 可通过直接观察得到, 可观察状态数目 M 不一定要和隐含状态数目 N 一致。

A 表示隐含状态转移概率矩阵。可以表示为 $a_{ij} \in A$, a_{ij} 表示在 t 时刻从状态 i 转移到 j 的概率, 其中 $1 \leq i, j \leq N$ 。

B 表示观察状态概率矩阵。 $b_j(k) \in B$, $b_j(k)$ 表示在 t 时刻从状态 j 中选择观察特征为 x_k 的概率, 其中, $1 \leq j \leq N; 1 \leq k \leq M$ 。

Π : 表示初始状态的概率分布, $\Pi = \{\pi_1, \pi_2, \dots\}$, $\pi_i = P\{s_t = i\}$, π_i 表示在 t 时刻选择状态 i 的概率, 其中, s_t 表示 t 时刻的状态, $1 \leq i \leq N$ 。

2.3.2. HMM-GARCH 模型

由于 HMM-GARCH 模型与 MRS-GARCH 模型都是基于 GARCH 模型对市场结构突变发生前后不同状态的波动率进行估计, 只是两者对状态的分类方法不同。因此两者的波动率模型相同, 具有相同的表达式:

$$\begin{aligned} r_t &= \phi_0 + \sum_{i=1}^{p_1} \phi_i r_{t-i} + a_t - \sum_{j=1}^{q_1} \phi_j a_{t-j} \\ a_t &= \varepsilon_t \sigma_t^2 \\ \sigma_t^2 &= \omega_{s_t} + \sum_{i=1}^{p_2} \alpha_{s_t, i} a_{t-i}^2 + \sum_{j=1}^{q_2} \beta_{s_t, j} \sigma_{t-j}^2 \end{aligned} \quad (4)$$

s_t 表示 r_t 在 t 时刻的状态, $\omega_{s_t}, \alpha_{s_t, i}, \beta_{s_t, j}$ 都是依赖于 t 时刻状态为 s_t 的参数。并且状态之间的转移服从

马尔科夫状态转移概率矩阵 P 。

2.3.3. HMM-GARCH 模型

HMM 模型的参数估计通常使用 Baum-Welch 算法。具体思想是：首先给定一个观察序列，调整模型参数 λ ，使其产生观察序列的概率 $P(O|\lambda)$ 最大，再利用迭代过程确定模型的参数。实际中，在给定的观察序列有限时，不存在一个最佳的方法来估计 λ ，只可能找到某些方法，让参数 λ 在特定的几个性能上具有较强的优势。在这种情况下，Baum-Welch 算法利用递推的思想，使得条件概率局部最大，最后得到模型参数的最终估计值。具体步骤如下：

首先在给出模型参数和观察序列的前提下，定义 $\xi_t(i, j)$ 表示 t 时刻的状态为 s_i ， $t+1$ 时刻的状态为 s_j 的概率， q_t 表示 t 时刻的状态值，即

$$\xi_t(i, j) = P(q_t = s_i, q_{t+1} = s_j | O, \lambda) \quad (5)$$

根据向前和向后变量的定义，可以将 $\xi_t(i, j)$ 写成下面的形式：

$$\xi_t(i, j) = \frac{P(q_t = s_i, q_{t+1} = s_j, O | \lambda)}{P(O | \lambda)} = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)} \quad (6)$$

其中，分子为 $P(q_t = s_i, q_{t+1} = s_j, O | \lambda)$ 。定义 $\gamma_t(i)$ 为 t 时刻状态为 s_i 的概率，则

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j) \quad (7)$$

将 $\gamma_t(i)$ 关于 t 求和，得到 $\sum_{t=1}^{T-1} \gamma_t(i)$ ，为从状态 s_i 转移出去的期望次数，将 $\xi_t(i, j)$ 关于 t 求和，得到

$\sum_{t=1}^{T-1} \xi_t(i, j)$ ，为从 s_i 转移到 s_j 的期望次数。

模型的参数 Π, A, B 也可表示为：

$$\tilde{\pi}_i = \gamma_1(i), \tilde{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \tilde{b}_j(k) = \frac{\sum_{t=1}^{T-1} \gamma_t(j)}{\sum_{t=1}^{T-1} \gamma_t(j)} \quad (8)$$

其中， $\tilde{\pi}_i$ 表示 $t=1$ 时刻状态 s_i 出现的概率。即把原来的参数 $\lambda = (A, B, \Pi)$ 更新为 $\tilde{\lambda} = (\tilde{A}, \tilde{B}, \tilde{\Pi})$ ，Baum [14]

证明了一般情况下 $P(O|\tilde{\lambda}) > P(O|\lambda)$ ，即 $\tilde{\lambda}$ 更接近于优化目标，因此经反复迭代就可以得到 λ 的极大似

然值，而且满足 $\sum_{i=1}^N \tilde{\pi}_i = 1, \sum_{j=1}^N \tilde{a}_{ij} = 1, \sum_{k=1}^N \tilde{b}_j(k) = 1$ 。

在实际应用中，常常假设观察序列服从连续分布，建立连续 HMM 模型，参数 B 为观测概率密度函数，可近似的表示为多个混合正态分布的组合。连续概率密度函数为：

$$b_j(O_t) = \sum_{k=1}^M w_{jk} b_{jk}(O_t), j=1, \dots, N \quad (9)$$

M 是混合分布的个数， w_{jk} 是在隐状态 j 时第 k 个混合分布的权重，且 $\sum_{k=1}^M w_{jk} = 1, j=1, \dots, N, k=1, \dots, M$ 。

其中使用最多的是高斯分布，且均值是 μ_{jk} ，协方差是 Σ_{jk} ，那么 $b_{jk}(O_t) = N(O_t, \mu_{jk}, \Sigma_{jk})$ 。

当隐状态为 j 时，第 k 个正态分布产生的显状态的均值和协方差为

$$\begin{aligned}\mu_{jk} &= E[O(t)|Q(t)=j, M(t)=k] \\ \Sigma_{jk} &= \text{cov}[O(t)|Q(t)=j, M(t)=k]\end{aligned}\quad (10)$$

多个混合正态分布的函数可表示为

$$b_{jk}(O_t) = N(O_t, \mu_{jk}, \Sigma_{jk}) = \frac{1}{(2\pi)^{U/2} |\Sigma_{jk}|^{1/2}} \exp\left\{-\frac{1}{2}(O_t - \mu_{jk})^T \Sigma_{jk}^{-1} (O_t - \mu_{jk})\right\}$$

$k=1$ 时,

$$b_{jk}(O_t) = N(O_t, \mu_{jk}, \Sigma_{jk}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\quad (11)$$

首先随机产生几组数值, 包括初始状态的概率分布 Π 、状态转移矩阵 A 、可近似表示为混合正态分布的观测概率密度函数 B , 运用 Baum-Welch 算法对数据集进行训练, 经过多次迭代后, 当达到设定的阈值时, 停止计算, 参数逐渐趋于最优值。

通过这些计算可以得到波动率状态的划分结果, 为建立结构突变前后不同状态的波动率模型提供基础。

3. VaR 方法及其检验

3.1. VaR 方法

VaR 表示处于风险中的价值, 是指在市场正常波动下, 某一金融资产或证券组合面临的未来最大可能损失。根据 Jorion [15] 定义, VaR 是在一定置信水平下, 某一金融资产或证券组合在未来特定的一段时间内的最大预期损失, 可用公式表示为:

$$P(\Delta P \leq \text{VaR}) = \alpha \quad (12)$$

上式表示资产价值损失小于等于可能损失上限的概率, 其中, ΔP 表示某一金融资产在一定持有期 Δt 的价值损失额, $1-\alpha$ 为置信水平。

根据 VaR 的定义, 首先假设金融资产的初始价值为 K_0 , 收益率 r_t 服从正态分布 $N(\mu, \sigma^2)$, z_α 为正态分布下置信水平为 $1-\alpha$ 的标准正态分布的分位数, 则在 $1-\alpha$ 概率水平下, 有

$$\text{VaR} = K_0 [\mu - (\mu - z_\alpha \sigma)] = K_0 z_\alpha \sigma \quad (13)$$

若持有期为 T , 可得均值和标准差为 μT 和 $\sigma\sqrt{T}$, 公式可以写为 $\text{VaR} = K_0 z_\alpha \sigma \sqrt{T}$ 。因此在假设下, 只要估计出金融资产的波动率 σ_t , 则可求出其 VaR 值, 在本文的 VaR 估计中, 波动率 σ_t 将通过 GARCH 模型、MRS-GARCH 模型和 HMM-GARCH 模型估计, 在实证中假设只考虑 1 天的持有期, 令 $K_0=1$ 。

同理可知道[16], 当金融资产收益率服从学生 t 分布或 GED 分布时, 可得到 VaR 的计算公式:

学生 t 分布:

$$\text{VaR}_t = K_{t-1} t_{\alpha, \nu} \sigma_t \quad (14)$$

$t_{\alpha, \nu}$ 表示自由度为 ν 的学生 t 分布的 α 分位数, K_{t-1} 表示滞后一期的股票收盘价格。

GED 分布:

$$\text{VaR}_{ged} = K_{t-1} g_\alpha \sigma_t \quad (15)$$

K_{t-1} 表示滞后一期的股票收盘价格, g_α 表示在置信水平 α 下的分位数。

3.2. VaR 检验

本文采用 Kupiec 提出的失败频率检验法对 VaR 的估计结果进行检验, 即是考察实际损失超过 VaR 的概率。实际损失计算公式为 $W_t = r_t \times P_t$, 即当期收益率与当期价格乘积[17]。

用 T 表示样本总数, F 表示样本中实际损失大于 VaR 值的个数, 即失败的次数, 且有 $F \sim B(T, \alpha)$, 则 F/T 表示失败率。Kupiec 提出的似然比检验的原假设是基于样本期计算得到的失败率与估计的 VaR 的左尾概率无显著性差异, 原假设 $H_0: \frac{F}{T} = \alpha$ 。检验统计量为:

$$LR = 2 \ln \left(\left(1 - \frac{F}{T} \right)^{T-F} \left(\frac{F}{T} \right)^F \right) - 2 \ln \left((1-\alpha)^{T-F} (\alpha)^F \right) \quad (16)$$

在原假设成立的情况下, 统计量 LR 服从自由度为 1 的 χ^2 分布, 其非拒绝域为:

$$\chi_{(1-\alpha/2)}^2(1) < LR < \chi_{(1-\alpha)}^2(1) \quad (17)$$

4. 实证研究

农业属于第一产业, 在我国的经济建设中具有战略性的地位, 农业的健康发展是推动工业文明不断前进和建设社会主义和谐社会的重要基础和物质保障, 其发展状况直接关系到我国改革开放和现代化建设全局。随着农业产业化经营的发展以及中国股票的建立, 农业上市公司开始顺应农业现代化的要求而不断出现; 2004 年至 2014 年连续 11 年中央“一号文件”中农业新政策的提出, 为农业经济发展在宏观方面创造了条件, 对于农业上市公司来说也提供了新的发展机遇。因此本文基于农业上市公司的股票, 建立基于状态转换的 GARCH 类模型来度量我国金融市场的波动, 并估计 VaR 值, 希望为农业方面的研究与发展提供一个参考方案。

4.1. 样本选取与数据预处理

为了研究农业上市公司股票价格的波动特征, 本文数据选择北大荒股票(600598)每日的收盘价为研究对象, 考虑到中共中央在 2004 年 1 月至 2014 年月连续 11 年发布农业、农村和农民为主题的中央一号文件, 同时考虑到我国从 2001 年开始实施股票交易制度的新规则, 方便以后研究国家农业经济政策对农业上市公司股票价格波动的影响, 选取时间范围是从 2005 年 1 月 4 日至 2017 年 4 月 11 日, 样本数量为 2881, 数据来源于同花顺股票软件。本文将数据分为样本内和样本外两部分, 其中样本内数据: 2005 年 1 月 4 日到 2016 年 11 月 11 日共 2781 个, 用于模型的建模; 样本外数据: 2016 年 11 月 12 日到 2017 年 4 月 11 日共 100 个数据, 用于检验模型的预测效果。选取的这段时间是全球受金融危机影响, 在这段时间我国股市也受其影响, 波动现象严重, 因此适合于考察中国股票市场收益率的波动特征。本文的数据处理均采用 R3.3.2 软件进行处理。

为了研究北大荒股票的股票价格波动和风险价值, 先对其数据的基本统计特征进行分析。通过统计分析, 我们可以了解北大荒公司的收益率是否具有“尖峰后尾”分布特征, 观察收益率序列是否服从正态分布的假设, 是否可以利用 GARCH 模型对股票数据价格波动进行估计。

在对日收益率的处理上采用对数收益率 $R_t = \ln(P_t/P_{t-1})$, 其中, P_t 表示 t 时刻北大荒的日收盘价, 通过计算可以得到 2780 个收益率。

4.2. ARCH 效应检验

4.2.1. 农业上市公司股票价格波动的统计特征分析

图 1 是北大荒股票 2780 个收益率的时序图。从图 1 可知, 收益率在 0 处上下波动, 而且较大的波动

后面会跟着一个较大的波动，同时较小的波动后面会跟着一个较小的波动，这说明北大荒日收益率具有集群性特征。这也验证了收益率的方差与时间有关，具有异方差性，即过去的收益率波动会影响未来的收益率波动。

图 2 给出了收益率序列的直方图。由计算结果可知，北大荒对数收益率序列均值为 0.000224，表明整体收益是大于 0 的；标准差为 0.032538，偏度为 -0.452，小于 0，说明序列分布有较长的左拖尾。峰度

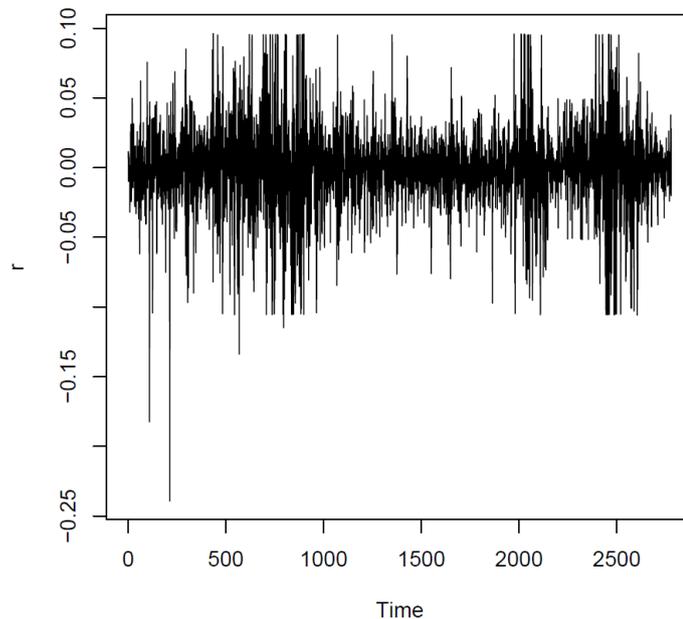


Figure 1. Day logarithmic yield sequence

图 1. 日对数收益率序列

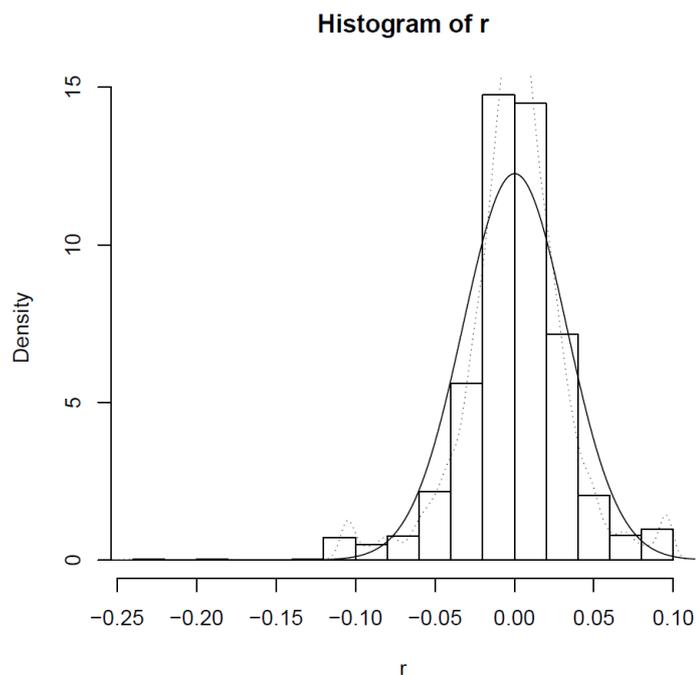


Figure 2. The histogram of return sequence

图 2. 收益率序列直方图

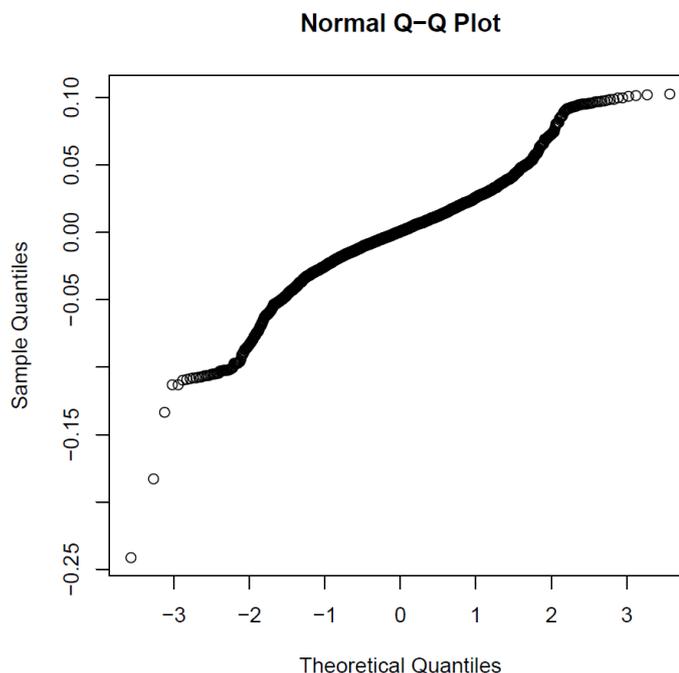


Figure 3. Normal Q-Q plot
图 3. 正态分布 Q-Q 图

为 3.197, 高于正态分布的峰度值 3, 说明收益率序列具有“尖峰后尾”的特征。Jarque-Bera 统计量为 1281.1599, 远大于在 5% 的显著性水平下的临界值 5.99, P 值为 0.000000, 拒绝该对数收益率序列服从正态分布的假设。从其正态 Q-Q 图也可进一步得出收益率序列的非正态性, 图 3 中可以看出收益率序列在正态直线外散布有大量的点, 数据点组成的线呈曲线状, 左右尾均有明显的摆动, 说明实际分布两侧具有后尾现象。所以, 可以判断该收益率不服从正态分布。

4.2.2. ARCH 效应检验

(1) 对收益率 r_t 序列进行平稳性检验

通过计算可得, ADF 统计量为 -37.7574, 均显著小于 1%、5% 和 10% 下的临界值, 因此, 北大荒收益率序列是平稳的, 可用于建模。

(2) 对样本收益率的自回归进行检验

为了确定其均值方程, 对收益率序列采用最小二乘估计, 得到最优模型是 ARMA(1,1), 具体均值方程表达式为: $r_t = \phi_0 + \phi_1 r_{t-1} + a_t + \theta_1 a_{t-1}$ 。

经检验 ARMA 模型中估计参数所对应的 P 值显示均显著, 表明该模型是可靠的。并且上述模型的每个统计检验结果都通过检验。对模型的残差以及残差的平方的自相关系数进行分析, 经检验得出残差序列不存在序列相关性。

(3) ARCH 效应检验

对上述均值方程的残差序列进行 ARCH-LM 检验, 验证结果如图 4。

由图 4 可以看出在高阶时 P 值仍为 0, 说明该序列具有高阶的 ARCH 效应, 因此可以考虑建立 GARCH(p,q)模型。

4.3. GARCH 类模型的建立

由于对数收益率序列拒绝了正态分布的假设, 本文则建立了基于广义误差分布(GED)下的 GARCH

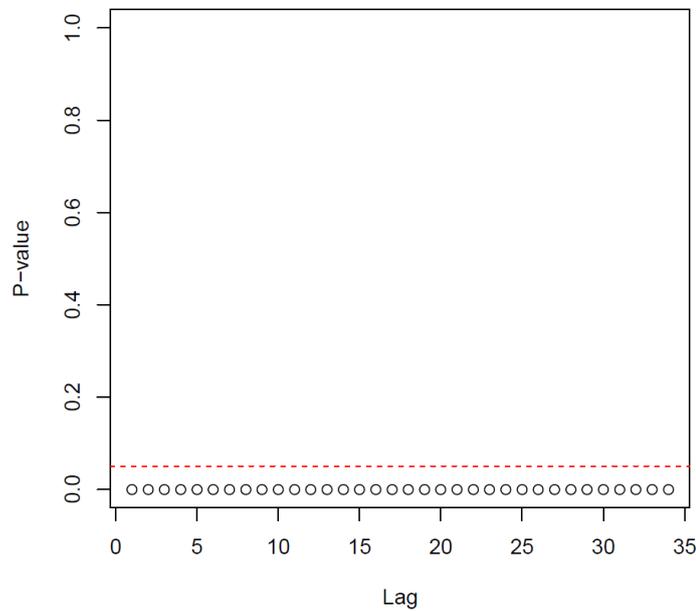


Figure 4. The test result of ARCH effect
图 4. ARCH 效应检验结果

模型，暂不考虑学生 t 分布。

4.3.1. GARCH 模型

对收益率序列建立 ARMA-GARCH，使用 R 软件进行参数估计，可得出 ARMA-GARCH 模型的均值方程和方差方程表达式：

$$\begin{aligned} r_t &= 0.000622 - 0.528r_{t-1} + a_t - 0.55a_{t-1} \\ \sigma_t^2 &= 0.0000137 + 0.075a_{t-1}^2 + 0.913\sigma_{t-1}^2 \end{aligned}$$

GED 的自由度为 1.242，小于 2，说明 GED 分布相对于标准正态分布有厚尾。

4.3.2. MRS-GARCH 模型

通过对北大荒的股票收益率建模，得到的 MRS-GARCH 模型可表示为：

当 $s_t = 1$ 时，即正常波动状态下模型为：

$$\begin{aligned} r_t &= 1.50186 + a_t \\ \sigma_t^2 &= 0.00031 + 0.08012a_{t-1}^2 + 0.91978\sigma_{t-1}^2 \end{aligned}$$

当 $s_t = 2$ 时，即异常波动状态下建立的模型为：

$$\begin{aligned} r_t &= 1.10136 + a_t \\ \sigma_t^2 &= 0.01634 + 0.00181a_{t-1}^2 + 0.99745\sigma_{t-1}^2 \end{aligned}$$

得到状态的一步转移概率矩阵为

$$P = \begin{pmatrix} 0.949 & 0.308 \\ 0.051 & 0.692 \end{pmatrix}$$

$p_{11} > p_{22}$ 表明正常状态的稳定性高于异常状态，这一点与实际相符。基于马尔科夫的 GARCH 模型， $\alpha_1^1 + \beta_1^1 = 0.9999$ ， $\alpha_1^2 + \beta_1^2 = 0.9996$ 均小于 1，说明基于马尔科夫建立条件异方差模型有助于降低波动的持续性，提高模型的可预测性。

4.3.3. HMM-GARCH 模型

首先对时间序列建立隐马尔科夫模型，将股票市场分为正常状态与异常状态。假设初始时刻为正常状态的概率为 1，异常状态为 2。运用 R3.3.2 软件，用 Baum-Welch 算法估算模型的参数，再采用 Viterbi 算法给出隐状态序列的最佳估计。大幅度的波动存在重要的信息，在处理时我们选择保留。小幅度的波动看作噪声，具有干扰的作用，不具有分析和预测价值，因此，在处理时只选择将小幅波动的噪声去除，保留有用信息，防止信号过度失真。

通过分析正常状态与异常状态下的数据，分别建立 GARCH 模型。

当 $s_t = 1$ 时，即正常波动状态下的模型为：

$$\begin{aligned} r_t &= 0.000492 + a_t \\ \sigma_t^2 &= 0.00000321 + 0.0085a_{t-1}^2 + 0.983\sigma_{t-1}^2 \end{aligned}$$

当 $s_t = 2$ 时，即异常波动状态下的模型为：

$$\begin{aligned} r_t &= 0.000315 - 0.188r_{t-1} - 0.928r_{t-2} + a_t + 0.259a_{t-1} + 0.943a_{t-2} + 0.082a_{t-3} \\ \sigma_t^2 &= 0.000536 + 0.125a_{t-1}^2 + 0.646\sigma_{t-1}^2 \end{aligned}$$

得到一步状态转移概率矩阵为：

$$P = \begin{pmatrix} 0.974 & 0.026 \\ 0.054 & 0.946 \end{pmatrix}$$

$p_{11} > p_{22}$ 表明正常状态的稳定性高于异常状态，与实际相符。在基于隐马尔科夫模型的 GARCH 模型中， $\alpha_1^1 + \beta_1^1 = 0.9915$ ， $\alpha_1^2 + \beta_1^2 = 0.9$ ，均小于 1，说明基于隐马尔科夫模型建立条件异方差模型有助于降低波动的持续性，提高模型的可预测性。

4.4. VaR 的计算与检验

根据建模得到波动率方程，可以计算出波动率 σ_t 的估计值，从而利用 VaR 的计算公式可以算出北大荒股票收盘价的日 VaR 值，由于经检验北大荒股票数据拒绝了正态分布，本文研究的是基于 GED 分布下的情况，因此 VaR 的计算公式为 $\text{VaR}_{ged} = K_{t-1} g_\alpha \sigma_t$ 。

本文采用估计的 GARCH-GED 模型、MRS-GARCH-GED 模型、HMM-GARCH-GED 模型对北大荒股票的日 VaR 值进行了计算，分别计算了左尾概率为 1% 和左尾概率为 5% 时的 VaR 值。其失败频率检验结果见表 1。

根据表 1 的检验结果，可以得到以下结论：

1) 对于不同的波动率模型计算 VaR 值，基于 GARCH 模型与 MRS-GARCH 模型估计的 VaR 值全部通过了检验，而 HMM-GARCH 模型估计的 VaR 值在置信水平为 95% 时未通过检验，说明对于北大荒股票数据，基于 GARCH 模型与 MRS-GARCH 模型的 VaR 估计方法优于 HMM-GARCH 模型的 VaR 估计方法。

2) 在 95% 和 99% 的置信水平下，HMM-GARCH 模型的失败率均小于 GARCH 模型与 MRS-GARCH 模型的失败率，并且在 99% 的置信水平下，HMM-GARCH 模型的失败率为 0，估计 VaR 值比另外两个模型估计的 VaR 值更接近实际损失，其估计最理想。因此，基于 HMM-GARCH 模型的 VaR 方法优于 GARCH 模型与 MRS-GARCH 模型的 VaR 方法。

3) 置信水平越高，各种方法计算的 VaR 值越大，并且失败率越小。

总结以上结论，对于北大荒股票数据，虽然 HMM-GARCH-VaR 方法在置信水平为 95% 时未通过检验，但是其失败天数明显少于其他两个模型，并且在 99% 的置信水平下，失败率为 0，估计效果很好，

Table 1. The statistical results of VaR failure frequency test
表 1. VaR 失败频率检验统计结果

波动模型	置信水平	VaR 的均值	VaR 的方差	失败次数	失败频率/%	LR 统计量	是否通过检验
GARCH	99%	0.84	0.01092	1	1	-1.78E-15	是
	95%	0.528	0.00432	3	3	0.977	是
MRS-GARCH	99%	0.464	0.00262	3	3	2.632	是
	95%	0.298	0.00023	6	6	0.198	是
HMM-GARCH	99%	0.617	0.02382	0	0	1.97	是
	95%	0.388	0.00942	1	1.02	4.783	否

则说明基于 HMM-GARCH 模型的 VaR 方法比基于 GARCH 模型与 MRS-GARCH 模型的 VaR 方法能更好的描绘和预测该股票的风险。

5. 结论

本文主要研究了基于状态转换 GARCH 模型的 VaR 实证分析。首先简单介绍了状态转换 GARCH 模型和其参数估计方法。再利用 R3.3.2 软件对北大荒股票数据进行基本统计分析, 包括正态性检验、平稳性检验、异方差检验, 选取拟合较好的 GARCH 模型, 分别在 95%、99% 置信水平下计算 VaR 值, 利用 Kupiec 提出的失败频率检验法检验模型的准确性和精度。

经过对北大荒的股票收盘价统计分析, 其收益率序列具有尖峰后尾性, 因此拒绝正态分布的假设, 则本文研究了基于广义误差分布(GED)下的波动率模型, 分别建立了 ARMA(1,1)-GARCH(1,1), MRS-GARCH(1,1), 正常状态下的 HMM-GARCH(1,1), 异常状态下的 HMM-ARMA(2,3)-GARCH(1,1)。结果表明基于隐马尔科夫模型的条件异方差模型可以有效降低 GARCH 模型高估波动持续性的问题, 较好的刻画了波动的集聚性、时变性和波动的结构性转变等特点。然后利用估计的波动率模型分别计算了置信水平为 95% 和 99% 的 VaR 值, 用 Kupiec-失败频率检验法对估计的 VaR 值进行检验。实验结果表明对于本文所选的股票, 基于 HMM-GARCH 模型的 VaR 方法比基于 GARCH 模型与 MRS-GARCH 模型的 VaR 方法能更好的描绘和预测该股票的风险。

致 谢

本工作受到国家自然科学基金和北京市教委面上项目资助。同时, 罗路和常欣欣同学在 Baum-Welch 算法推导过程中给予了宝贵的指导, 卢静会、王晓繁同学等在软件编程方面提供了协助, 在此一并致谢!

基金项目

国家自然科学基金资助项目(71401124); 北京市教委面上项目资助(KM201510017002)。

参考文献 (References)

- [1] 袁梁, 霍学喜, 吴后宽. VaR 模型在商业银行风险管理中的应用[J]. 西北农林科技大学学报(社会科学版), 2005, 5(2): 56-60.
- [2] Morgan Guaranty Trust Company (1995) RiskMetrics-Technical Document.
- [3] Morgan Guaranty Trust Company, Reuters Ltd. (1996) Risk Metrics-Technical Document. 2nd Edition.
- [4] Riskmetrics Group (2011) Return to Risk Metrics: The Evolution of a Standard.
- [5] Zhang, B., Wei, Y., Yu, J., et al. (2014) Forecasting VaR and ES of Stock Index Portfolio: A Vine Copula Method. *Physica A: Statistical Mechanics and Its Applications*, 112-124. <https://doi.org/10.1016/j.physa.2014.08.043>

- [6] Foroni, C., Guérin, P. and Marcellino, M. (2014) Markov-Switching Mixed-Frequency VaR Models. *International Journal of Forecasting*.
- [7] Marcucci, J. (2005) Forecasting Stock Market Volatility with Regime-Switching Models. *Studies in Nonlinear Dynamics & Econometrics*, **9**, Article 6. <https://doi.org/10.2202/1558-3708.1145>
- [8] 陈守东, 俞世典. 基于 GARCH 模型的 VaR 方法对中国股市的分析[J]. 吉林大学社会科学学报, 2002(4): 11-17.
- [9] 范英. VaR 方法及其在股市风险分析中的应用初探[J]. 中国管理科学, 2000, 8(3): 26-32.
- [10] 魏捷, 王冬梅, 李威. 基于 GARCH 模型的上证综合指数 VaR 计算[J]. 统计与决策, 2010(23): 135-137.
- [11] 郭名媛, 张世英. 基于 MRS-GARCH 模型的 VaR 方法及其在上海股市的应用[J]. 西北农林科技大学学报(社会科学版), 6(2): 52-55.
- [12] 罗健英, 陈宴祥, 陈粘. 基于 MRS-GARCH 的钢铁期货市场 VaR 风险测度[J]. 管理现代化, 2014, 4: 92-94.
- [13] Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307-327.
- [14] Baum, L.E. and Sell, G.R. (1986) Growth Functions for Transformations on Manifolds. *Pacific Journal of Mathematics*, **27**, 211-227.
- [15] Jorion, P. (1997) Value-at-Risk: The New Benchmark for Controlling Market Risk. The McGraw-Hill Companies.
- [16] 张竞文. 基于 GARCH-VaR 模型的沪、深指数风险度量分析[D]: [硕士学位论文]. 大庆: 东北石油大学, 2016.
- [17] 徐伟浩. 沪深 300 股指期货 VaR-GARCH 模型风险管理研究——基于恒指期货的比较视角[J]. 证券与保险, 2011, 10: 31-35.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2324-7991, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: aam@hanspub.org