

# 动态混合 HGARCH 模型的估计和预测<sup>①</sup>

李木易<sup>1, 2, 3</sup>, 方 颖<sup>1, 2, 3\*</sup>

- (1. 教育部计量经济学重点实验室(厦门大学), 厦门 361005;
2. 厦门大学王亚南经济研究院与经济学院, 厦门 361005;
3. 福建省统计科学重点实验室(厦门大学), 厦门 361005)

**摘要:** 在 GARCH 模型框架下, 提出过新的双曲 GARCH 形式(记为 HGARCH), 不仅与 HY-GARCH 模型一样可以同时刻画波动的强烈振幅和长记忆衰减两个性质, 并且较之 HY-GARCH 模型, 有更简单的条件方差非负约束条件. 然而, 当时间序列较长时, 用单一参数结构不能充分捕捉可能发生的结构变化. 为此, 提出新的动态混合 HGARCH 模型(DM-HGARCH), 使之可以同时拥有协方差平稳、长记忆和结构变化 3 个特性. 讨论了新模型的弱平稳解存在条件, 利用 EM 算法进行参数估计, 并且用蒙特卡罗模拟给出估计在有限样本下的表现. 最后将该模型分别用于 1995 年~2014 年中国上证指数和美国标普 500 指数的日波动率建模. 结果表明, 在给定样本期间内, 动态混合 HGARCH 模型(DM-HGARCH) 对标普 500 指数有更好的样本内拟合和样本外预测表现.

**关键词:** 动态混合; 长记忆波动率; 双曲 GARCH 模型; EM 算法

**中图分类号:** O212      **文献标识码:** A      **文章编号:** 1007-9807(2020)05-0001-12

## 0 引 言

自从 Ding 等<sup>[1]</sup>首次在美国标普 500 指数收益率的幂律转换序列  $\{|r_t|^d\}$  上发现长记忆现象以来, 金融资产波动率的长记忆性已作为资产收益率的 10 大经验特征之一而被广泛接受<sup>[2]②</sup>. 忽略波动率的长记忆性将会提高模型误设的风险, 不能精确量化波动率的产生机制和对未来波动率的预测, 从而导致对整个市场风险的错判. 为刻画波动率的长记忆性质, Baillie<sup>[3]</sup>借用 ARFIMA 模型的思想, 建立了 FIGARCH 模型, 用来刻画二阶矩序列的长记忆性质. 但 FIGARCH 模型的无

条件方差不存在, 因而无法对 FIGARCH 建立相应的统计推断. Davidson<sup>[4]</sup>通过将经典 GARCH 模型<sup>[5]</sup>和 FIGARCH 模型进行加权组合, 得到双曲 GARCH 模型(hyperbolic GARCH, HY-GARCH), 并讨论了此模型宽平稳解的存在性和参数估计的统计推断. 基于脉冲响应系数的衰减速度, 可以将经典 GARCH 模型称为短记忆 GARCH(脉冲响应系数呈快速的几何衰减), 而将 FIGARCH 和 HY-GARCH 称为长记忆 GARCH(脉冲响应系数呈缓慢的双曲衰减).

波动率有两个本质特征, 一个是波动的振幅大小, 一个是波动的衰减速度. HY-GARCH 模型

① 收稿日期: 2019-04-28; 修订日期: 2020-03-09.

基金项目: 国家自然科学基金重点资助项目(71631004); 国家自然科学基金资助项目(71671150; 11771361); 国家杰出青年科学基金资助项目(71625001); 国家基础科学中心资助项目(71988101).

通讯作者: 方 颖(1973—), 男, 上海人, 博士, 教授. Email: yifst1@xmu.edu.cn

② 目前对于长记忆波动率产生的内在原因学术界并没有统一的解释. Mikosch 和 Stărică<sup>[6]</sup>对金融时间序列的非平稳性、长记忆性以及 IGARCH 产生的内在原因进行了系统的研究. Baillie 和 Morana<sup>[7]</sup>指出条件方差出现结构上的变化是导致波动率产生长记忆性的潜在原因之一.

中既包含 GARCH 模型中没有的记忆参数,也包含 FIGARCH 中没有的振幅参数,因而可以从两个维度刻画波动率的经验特征.该模型也是目前最为广泛使用的长记忆波动率模型之一.但 HY-GARCH 模型的表达形式不够直观,且因为振幅参数对脉冲响应系数的非线性影响,使得非负约束条件比在 FIGARCH 模型下更加复杂,继而加大了优化难度.基于这些原因, Li 等<sup>[8]</sup>建立了新的双曲 GARCH 模型(记为 HGARCH),该模型具有简单的参数化形式,且与 HY-GARCH 一样可以同时捕捉波动率的振幅大小和衰减速度两个维度.与 HY-GARCH 模型相比, HGARCH 所容许的参数空间与 FIGARCH 相同,简化了参数约束条件的表达形式.实证分析也表明,该模型在拟合和预测上与 HY-GARCH 模型有相同的优良表现.

在时间序列涵盖较长时间范围时,用单一的模型结构进行拟合是不合理的,使用含有结构变化的模型更能捕捉数据的动态特征. Davidson<sup>[4]</sup>也发现,用单一 HY-GARCH 模型对 1997 年~1998 年亚洲金融危机爆发期间中国台湾、韩国和印度尼西亚的美元汇率进行拟合时,3 条序列的加权系数均大于 1 或者接近于 1,这意味着整个过程仍然是爆炸的或非平稳的;该现象在 Li 等<sup>[8]</sup>中也有类似发现.

基于上述研究的启发,本文考虑在长记忆的基础上,采用时变模型来解释波动率的结构变化.混合模型作为特殊的时变模型,已有很多文献对其进行研究,例如 Wong 和 Li<sup>[9]</sup>通过混合 ARCH 来刻画波动率的变化情况,并捕捉极端厚尾情形; Cheng 等<sup>[10]</sup>则将混合 ARCH 推广到混合 GARCH,指出该模型在 VaR 的失败率检验上超过混合 ARCH 以及单一 GARCH 类模型.但以上这些模型都属于短记忆波动率模型,不足以捕捉波动率的慢衰减性质.为此, Li 等<sup>[11]</sup>将 HY-GARCH 模型中的加权系数看成区制变量,把 HY-GARCH 模型中的条件方差看成是短记忆平稳 GARCH 模型和长记忆非平稳 FIGARCH 模型的混合,构造了混合记忆(mixture memory) GARCH 模型.该模型同时具备协方差平稳、平方项双曲衰减(波动率长记忆性)以及结构时变的特点. Klein 和

Walther<sup>[12]</sup>将其应用到原油价格波动率的预测时发现,相对其他离散型的 GARCH 模型,该模型具有更好的预测能力. Li 等<sup>[13]</sup>进一步提出混合双自回归模型来同时刻画均值和方差序列可能存在的不同成分结构.近年来混合记忆模型得到进一步扩展,如 Catania<sup>[14]</sup>提出的动态自适应(dynamic adaptive)混合模型和 Basatini 和 Rezakhanlou<sup>[15]</sup>提出的马尔可夫平滑转移 HY-GARCH 模型.

在 Li 等<sup>[11]</sup>的混合记忆 GARCH 模型中,由于限定了混合的两个成分必须是 GARCH 和 FIGARCH,因而对混合模型的灵活性造成一定损失.为了克服这一缺点,并且考虑到 GARCH、IGARCH 以及 FIGARCH 本身就是 HGARCH 模型的特殊形式,在本文中直接建立新的动态混合 HGARCH 模型,记为 DM-HGARCH. DM-HGARCH 模型的各个成分更具有一般性,可以由数据驱动决定每一个成分的实际结构,并且在一定条件下,可以同时具备协方差平稳、波动率双曲衰减以及结构变化 3 个重要性质.

国内学术界关于金融市场波动率的长记忆性和混合分布的理论研究较少,实证研究较多.在股票指数方面,林宇<sup>[16]</sup>利用 HY-GARCH 模型及有偏学生  $t$  分布来捕捉中国沪深股市资产收益率的分布形态;曹广喜等<sup>[17]</sup>利用双长记忆模型对中国沪深股市收益率和波动率序列同时进行拟合,并比较了不同分布下的拟合与预测精度;李云红等<sup>[18]</sup>对上证指数、深证指数、日经 225 指数、标普 500 指数在不同时间区间内的序列进行分析,指出收益序列的长记忆性具有时变特征;王安兴和谭鲜明<sup>[19]</sup>利用混合分布和 GARCH 模型对中国上市公司股票波动率进行分组,可将异常股票找出并归为同一组;近几年对 GARCH 类模型波动率的长记忆性和时变性的研究,还可以参考文献 [20-26].

## 1 动态混合 HGARCH 模型

### 1.1 单一双曲 GARCH 模型

Davidson<sup>[4]</sup>通过加权组合的思路,将平稳短记忆 GARCH 与非平稳长记忆 FIGARCH 相组合,得到新的 HY-GARCH( $p, d, q$ )模型

$$\begin{cases} y_i = \sigma_i \varepsilon_i \\ \sigma_i^2 = \frac{\gamma}{\beta(1)} + \theta(L) y_i^2 \theta(L) \\ \quad = 1 - \frac{\delta(L)}{\beta(L)} [1 - \alpha + \alpha(1-L)^d] \\ \beta(L) = 1 - \beta_1 L - \beta_2 L^2 - \dots - \beta_p L^p \\ \delta(L) = 1 - \delta_1 L - \delta_2 L^2 - \dots - \delta_q L^q \\ (1-L)^d = 1 - \sum_{j=1}^{\infty} \pi_j L^j \\ \pi_j = \frac{d\Gamma(j-d)}{\Gamma(1-d)\Gamma(1+j)} = O(j^{-1-d}) \end{cases} \quad (1)$$

式中  $\varepsilon_i$  是均值为 0 且方差为 1 的独立同分布序列;  $L$  为滞后算子;  $\alpha \geq 0$ ;  $0 \leq d \leq 1$ ;  $\Gamma(\cdot)$  代表伽马函数。

相比经典 GARCH 模型, 此模型中有两个额外参数  $\alpha$  和  $d$ . 参数  $\alpha$  控制了波动的振幅大小(称为振幅参数), 继而决定过程的矩存在条件. 参数  $d$  控制了波动带来的冲击消散的速度(称为记忆参数). 由  $\pi_j$  的表达式可知, 记忆的衰减速度与  $d$  成反比,  $d$  值越大, 脉冲系数衰减速度越快; 相反,  $d$  值越小, 脉冲系数衰减速度越慢. 从 HY-GARCH 条件方差的形式可以进一步发现, 此模型包含了经典 GRACH、IGARCH 以及 FIGARCH 等模型. 例如, 当  $\alpha=1, 0 < d < 1$  时, 该模型退化为 FIGARCH. 此时只有长记忆性, 不具备方差平稳性. 当  $\alpha=1, d=1$  时, 该模型退化为 IGARCH. 当  $0 < \alpha < 1, d=1$  时, 该模型则退化为一般平稳 GARCH. 当  $0 < \alpha < 1$  且  $0 < d < 1$  时, 模型同时具备协方差平稳和长记忆波动两个性质.

在上述 HY-GARCH 模型中, 由于  $\alpha$  和  $d$  这两个参数的引入, 使得条件方差  $\sigma_i^2$  的非负约束条件变得复杂很多<sup>[27]</sup>, 进而增加了优化难度. 重新考虑条件方差的结构后, Li 等<sup>[8]</sup> 提出了新的双曲 GARCH 模型, 为区别起见, 这里记为 HGARCH 模型. 该模型参数化形式为

$$\begin{cases} y_i = \sigma_i \varepsilon_i, \sigma_i^2 = \frac{\gamma}{\beta(1)} + \theta(L) y_i^2 \\ \theta(L) = \alpha \left[ 1 - \frac{\delta(L)}{\beta(L)} (1-L)^d \right] \end{cases} \quad (2)$$

式中  $\alpha \geq 0, 0 \leq d \leq 1$ .

关于该模型的可识别性、(弱) 平稳解存在条件、高阶矩条件、拟似然估计(QMLE) 的渐近性质

以及双曲记忆检验, 详见文献[8], 这里只做简单概括. 该模型仍然由  $\alpha$  和  $d$  两个参数来控制波动的振幅和消散速度. 相比 HY-GARCH, 该模型中  $\theta(L)$  只是简单的将 FIGARCH 的脉冲响应重新尺度化(rescale). 当  $\alpha=1$  时, 该模型退化为 FIGARCH. 同时, 由于  $\alpha \geq 0$ , 使得该模型中条件方差的非负约束和 FIGARCH 模型中相同<sup>[28]</sup>. 因此, HGARCH 模型同时拥有协方差平稳(当  $0 < \alpha < 1$ ) 以及波动率双曲衰减两个性质, 同时参数化形式比 HY-GARCH 模型更为直观简单. Li 等<sup>[8]</sup> 实证分析表明, 该模型与 HY-GARCH 模型在有限样本下, 有同样的拟合和预测表现.

### 1.2 混合 HGARCH 模型

受混合模型的启发, 考虑如下含有  $K$  个成分的混合 HGARCH 模型

$$\begin{cases} y_i = \sigma_i \varepsilon_i, \sigma_i^2 = \sum_{i=1}^K Z_{it} \sigma_{it}^2 \\ \sigma_{it}^2 = \frac{\gamma_i}{\beta_i(1)} + \alpha_i \left[ 1 - \frac{\delta_i(L)}{\beta_i(L)} (1-L)^{d_i} \right] y_i^2 \\ i = 1, \dots, K \end{cases} \quad (3)$$

记  $F(y_i | \mathcal{F}_{t-1})$  是给定过去信息后  $y_i$  的累积分布函数(CDF),  $\mathcal{F}_t$  是由  $\{y_t, y_{t-1}, \dots\}$  产生的  $\sigma$ -代数;  $G(\cdot)$  是  $\varepsilon_i$  的分布函数, 且满足以下条件  $\int x dG(x) = 0, \int x^2 dG(x) = 1$ .  $Z_{it}$  可看成是混合成分的指示变量, 当  $Z_{it} = 1$  时, 表示  $t$  时刻的波动率来自第  $i$  个成分. 令  $P(Z_{it}=1) = p_i$ , 则  $\sum_{i=1}^K p_i = 1$ . 基于这些, 可以得到  $\{y_i\}$  的条件分布函数如下

$$F(y_i | \mathcal{F}_{t-1}) = \sum_{i=1}^k p_i G\left(\frac{y_i}{\sigma_{it}}\right) \quad (4)$$

值得注意的是, Wong 和 Li<sup>[9]</sup> 的混合 ARCH 模型, Cheng 等<sup>[10]</sup> 的混合 GARCH 模型, 以及 Li 等<sup>[11]</sup> 的混合记忆 GARCH 模型都是模型(3)的特殊形式, 因而模型(3)更具有一般性, 可以降低模型误设的风险. 但对于模型(3), 需要重新给出弱平稳解和高阶矩的存在条件. 为此, 将模型(3)中的条件方差重新表达成如下式的 ARCH( $\infty$ ) 模型

$$\begin{cases} \sigma_{i,t}^2 = \omega_0^{(i)} + \alpha_i \sum_{j=1}^{\infty} \omega_j^{(i)} \varepsilon_{i-t}^2 \\ i = 1, 2, \dots, K \end{cases} \quad (5)$$

式中对每一个  $i$  都有

$$\sum_{j=1}^{\infty} \omega_j^{(i)} = 1$$

**假设 1**

$$\omega_0^{(i)} > 0, 0 < \alpha_i < 1, \omega_j^{(i)} > 0$$

$$j = 1, 2, \dots, i = 1, \dots, K$$

1) 当  $0 < d_i < 1$  时, 该假设下,  $\omega_j^{(i)}$  的非负性等同于 FIGARCH 模型的非负约束, 具体约束条件可参考文献 [28].

2) 当  $d_i = 0$  或者  $d_i = 1$  时, 此时退化成 GARCH 模型或者 IGARCH 模型, 可参考经典 GARCH 模型的非负条件 [29].

**假设 2**

$$(E\varepsilon_t^2) \sum_{j=1}^{\infty} \sum_{i=1}^K p_i \alpha_i \omega_j^{(i)} < 1$$

3) 当  $\sum_{j=1}^{\infty} \omega_j^{(i)} = 1$  时, 假设 2 可以进一步简化成  $(E\varepsilon_t^2) \sum_{i=1}^K p_i \alpha_i < 1$ .

**定理 1** 在假设 1 和假设 2 下, 模型 (3) 存在唯一的不可预期 (non-anticipative) 且方差有限的严平稳解.

**假设 3**

$$[E(\varepsilon_t^{2M})]^{1/M} \sum_{j=1}^{\infty} \left[ \sum_{i=1}^K p_i (\alpha_i \omega_j^{(i)})^M \right]^{1/M} < 1$$

$M$  是正整数.

**定理 2** 在假设 1 和假设 3 下, 模型 (3) 存在唯一的不可预期 (non-anticipative) 的严平稳解  $E(y_t^{2M}) < \infty$ .

4) 定理 2 给出了模型 (3) 高阶矩存在的充分条件, 是定理 1 的推广. 特别地, 当  $M = 1$  时, 定理 2 退化成定理 1.

考虑更一般的混合概率是动态时变函数的情形, 此时  $p_i$  变成  $p_{it}$ , 表示  $t$  时刻波动率来自第  $i$  个成分的概率. 由于金融市场的动态性绝大多数时候都可以用震荡或平静两个状态来刻画, 因此只考虑有两个混合成分的简单情形. 这里采用逻辑斯特函数作为连接函数 (link function). 连接函数的具体形式则依赖于实际经济意义和研究目的, 可以是解释变量的线性或者非线性函数, 这里不做具体讨论. 值得注意的是, 如果在逻辑斯特函数中引入外生变量  $X_t$ , 此时平稳性的讨论将更为复杂. 由于本文的重点是考虑动态混合时间序列自回归模型, 因此这里假设连接函数只依赖自身的

历史信息, 是个自激励过程. 本文使用的连接函数具体形式如下

$$\ln\left(\frac{p_{it}}{1-p_{it}}\right) = \lambda_0 + \lambda_1 y_{t-1} + \dots + \lambda_p y_{t-p} \quad (6)$$

即

$$p_{it} = \frac{e^{\lambda_0 + \sum_{i=1}^p \lambda_i y_{t-i}}}{1 + e^{\lambda_0 + \sum_{i=1}^p \lambda_i y_{t-i}}}$$

进一步, 可以利用极大似然比检验, 针对原假设  $H_0: \lambda_1 = \dots = \lambda_p = 0$ , 检验混合概率是否为常数. 如果接受原假设, 则表示混合概率不随时间变化而变化.

## 2 模型估计方法

### 2.1 EM 算法

由于混合模型中存在不可观测的哑变量  $Z_{it}$ , 在这种情况下文献中通常采用 EM 算法进行参数估计. 暂且假设  $\varepsilon_t$  服从正态分布, 且假设均值为 0, 则伪对数似然函数 (pseudo-log-likelihood) 为

$$\begin{aligned} L^*(\theta) &= \sum_{i=1}^n \ln \left\{ \frac{dF(y_t | \mathcal{F}_{t-1}, \Omega_t)}{dy} \Big|_{y=y_t} \right\} \\ &= \sum_{i=1}^n \ln \left\{ \sum_{i=1}^K \frac{p_{it}(\theta)}{\sigma_{it}(\theta)} \exp\left(-\frac{y_t^2}{2\sigma_{it}^2(\theta)}\right) \right\} - \\ &\quad \frac{n}{2} \ln(2\pi) \end{aligned} \quad (7)$$

式中  $\theta$  为未知参数向量,  $\theta = (\theta_{1h}, \dots, \theta_{kh}, \theta_p)'$ ,  $\theta_{ih} = (\gamma_i, \beta_{i1}, \dots, \beta_{ip_i}, \delta_{i1}, \dots, \delta_{i q_i}, \alpha_i, d_i)'$ ,  $i = 1, \dots, K$  是第  $i$  个条件方差包含的未知参数,  $\theta_p = (\lambda_0, \lambda_1, \dots, \lambda_p)'$  是连接函数包含的未知参数.

在模型满足一定条件下, 利用 EM 算法可以求得  $L^*(\theta)$  的最大值点, 即  $\theta$  的伪最大似然估计量. 但为了更容易处理似然函数  $L^*(\theta)$  的优化问题, 将哑变量  $Z_{it}$  加入似然函数, 并且将  $\{y_t, z_{it}, t = 1, \dots, n, i = 1, \dots, K\}$  看作样本. 现考虑如下对数似然函数形式

$$\begin{aligned} L(\theta) &= \sum_{i=1}^n l_i(\theta) \quad (8) \\ l_i(\theta) &= \sum_{i=1}^K Z_{it} \ln [p_{it}(\theta)] - \\ &\quad \frac{1}{2} \sum_{i=1}^K Z_{it} \left\{ \ln [\sigma_{it}^2(\theta)] + \frac{y_t^2}{\sigma_{it}^2(\theta)} \right\} \end{aligned}$$

将函数  $L(\boldsymbol{\theta})$  称为完整对数似然函数, 相应地,  $L^*(\boldsymbol{\theta})$  称为不完整对数似然函数.

由于  $Z_{it}$  不可观测, 按照混合模型的一般惯例, 采用迭代 EM 算法对  $L(\boldsymbol{\theta})$  求解伪极大似然估计<sup>[30]</sup>. 迭代 EM 算法的两个步骤如下:

①期望 (E 步) 假定  $\boldsymbol{\theta}$  的初始值  $\boldsymbol{\theta}_0$ , 对缺失数据  $\{z_{it}\}$  用其条件期望 (用  $\tau_{it}$  表示) 来替代. 依据 Wong 和 Li<sup>[9]</sup>, 该步可表述为

$$\begin{cases} l_t(\boldsymbol{\theta}) = \sum_{i=1}^K Z_{it} \ln [p_{it}(\boldsymbol{\theta})] - \\ \frac{1}{2} \sum_{i=1}^K Z_{it} \left\{ \ln [\sigma_{it}^2(\boldsymbol{\theta})] + \frac{y_{it}^2}{\sigma_{it}^2(\boldsymbol{\theta})} \right\} \\ \tau_{it} = \frac{p_{it} g\left(\frac{y_{it}}{\sigma_{it}}\right)}{\sum_{j=1}^K \frac{p_{jt} g\left(\frac{y_{jt}}{\sigma_{jt}}\right)}{\sigma_{jt}}}, i = 1, \dots, K \end{cases} \quad (9)$$

式中  $g(\cdot)$  为标准正态分布的密度函数. 显然, 当混合成分只有两个时, 有  $\tau_{1t} = 1 - \tau_{2t}$ .

②最大化 (M 步) 假定缺失数据  $\{z_{it}\}$  由 E 步中的  $\{\tau_{it}\}$  代替. 式 (8) 中只剩下  $\boldsymbol{\theta}$  未知. 接下来通过最大化伪似然函数  $L(\boldsymbol{\theta}) = \sum_{t=1}^n l_t(\boldsymbol{\theta})$  求得  $\boldsymbol{\theta}$  的伪极大似然估计值, 即

$$\hat{\boldsymbol{\theta}}_n = \arg \max_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta})$$

通过不断迭代上述两步, 直至达到给定的收敛精度, 将最后一步的估计值作为  $\boldsymbol{\theta}$  的估计值, 记为  $\hat{\boldsymbol{\theta}}_{EM}$ . 值得注意的是, 在没有均值结构的情况下, M 步的整体优化问题可以简化为如下独立的优化问题

$$\hat{\theta}_{it} = \arg \max_{\theta_{it}} \sum_{t=1}^n Z_{it} \left\{ \ln [\sigma_{it}^2(\boldsymbol{\theta})] + \frac{y_{it}^2}{\sigma_{it}^2(\boldsymbol{\theta})} \right\}$$

$$i = 1, \dots, K$$

以及

$$\hat{\theta}_p = \arg \max_{\theta_p} \sum_{t=1}^n \left\{ \sum_{i=1}^K Z_{it} \ln [p_{it}(\boldsymbol{\theta})] \right\}$$

### 2.2 估计的渐近方差

根据 Louis<sup>[31]</sup> 的缺失数据原则, 观测信息矩阵可以通过以下方式计算

$$\mathbf{I} = \mathbf{I}_c - \mathbf{I}_m = \mathbf{E} \left( - \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \middle|_{\boldsymbol{\theta}, \mathcal{F}_{t-1}} \right)_{\hat{\boldsymbol{\theta}}_{EM}} - \text{Var} \left( \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \middle|_{\boldsymbol{\theta}, \mathcal{F}_{t-1}} \right)_{\hat{\boldsymbol{\theta}}_{EM}} \quad (10)$$

式中  $\mathbf{I}_c$  为完整信息矩阵,  $\mathbf{I}_m$  为缺失信息矩阵,  $\hat{\boldsymbol{\theta}}_{EM}$  为 EM 算法求解的参数估计值. 如果分布是对称的, 信息矩阵可以看成近似分块对角矩阵. 此时每个部分的参数估计值的标准误可以单独计算. 这里需要注意的是, 哑变量  $\{Z_{it}\}$  条件独立, 所以观测信息矩阵可以用另一种方式求得, 即

$$\mathbf{I} = \sum_{t=1}^n \mathbf{J}_t(\hat{\boldsymbol{\theta}}_{EM}) \mathbf{J}_t'(\hat{\boldsymbol{\theta}}_{EM}) \quad (11)$$

其中

$$\begin{aligned} \mathbf{J}_t(\boldsymbol{\theta}) &= \mathbf{E} \left( \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \middle|_{\boldsymbol{\theta}, \mathcal{F}_{t-1}} \right) \\ &= \sum_{i=1}^K \tau_{it} \left\{ \frac{1}{p_{it}} \frac{\partial p_{it}}{\partial \boldsymbol{\theta}} + \frac{1}{2\sigma_{it}^2} \left[ \frac{y_{it}^2}{\sigma_{it}^2} - 1 \right] \frac{\partial \sigma_{it}^2}{\partial \boldsymbol{\theta}} \right\} \end{aligned}$$

利用上述两种方法计算的信息矩阵理论上等价, 详细证明可参考文献 [11]. 但在数值运算中, 因为后者保证信息矩阵在任何情况下都是非负定的, 因此在实际应用中本文采用第二种方法. 当  $\{y_{it}\}$  满足严平稳遍历和一定的矩条件, 同时误差项满足  $\mathbf{E}(\boldsymbol{\varepsilon}_t^4) < \infty$  时, 可以得到  $\hat{\boldsymbol{\theta}}_{EM}$  的一致性和渐近正态性, 具体而言, 当样本量  $n$  足够大时

- 1)  $\hat{\boldsymbol{\theta}}_{EM} \xrightarrow{p} \boldsymbol{\theta}_0$ ,
- 2)  $\sqrt{n}(\hat{\boldsymbol{\theta}}_{EM} - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \mathbf{I}^{-1})$

详细证明可以参考文献 [9, 11]. 值得注意的是, 当误差项  $\boldsymbol{\varepsilon}_t$  不满足正态分布时, 参数估计的渐近方差将有更复杂的表达式. 这将留作以后研究.

## 3 模拟结果

利用蒙特卡罗模拟试验, 研究 EM 算法在 DM-HGARCH 模型中参数估计的有限样本表现. 由于本文主要研究二阶矩性质, 所以在整个试验过程中将均值项  $\mu_i$  都设定为 0. 为考察序列本身的自反馈行为, 逻辑斯特连接函数设定为只依赖于  $y_t$  的历史值, 不包含其他外生变量, 且考虑只有两个混合成分的情况. 数据产生过程如下

$$\begin{cases} y_t = \sigma_t \boldsymbol{\varepsilon}_t, \sigma_t = \sqrt{Z_{1t} \sigma_{1t}^2 + Z_{2t} \sigma_{2t}^2} \\ \sigma_{1t}^2 = 0.15 + 1.2 [(1 - 0.6B) - (1 - B)^{0.7}] y_t^2 + 0.6 \sigma_{1t-1}^2 \\ \sigma_{2t}^2 = 0.4 + 0.8 [(1 - 0.4B) - (1 - B)^{0.6}] y_t^2 + 0.4 \sigma_{2t-1}^2 \end{cases} \quad (12)$$

式中  $\varepsilon_t \sim i. i. d. N(0, 1)$  .

$$P(Z_{1t} = 1) = p_{1t} = \frac{1}{1 + \lambda_0 + \lambda_1 y_{t-1}}$$

$$P(Z_{2t} = 1) = p_{2t} = 1 - p_{1t}$$

参数向量  $\theta$  表示为  $\theta = (\theta'_{1h}, \theta'_{2h}, \theta'_p)'$ . 波动率成分 1 中的参数  $\theta_{1h} = (\gamma_1, \beta_1, \alpha_1, d_1)'$  = (0.15, 0.6, 1.2, 0.7)'. 波动率成分 2 中的参数  $\theta_{2h} = (\gamma_2, \beta_2, \alpha_2, d_2)'$  = (0.4, 0.4, 0.8, 0.6)', 连接函数中参数  $\theta_p = (1.8, -0.7)'$ .

以上参数设置满足条件方差的非负约束条件<sup>[28]</sup>, 并且在两个条件方差中,  $d$  的取值均大于 0.5, 这是为了保证  $d$  的估计的渐近正态性<sup>[32]</sup>. 两个波动率成分均为 HGARCH(1,  $d$ , 0) 结构, 都没有包含自回归多项式  $\delta(\cdot)$  部分, 原因在于即使包含  $\delta(\cdot)$  部分, 其对应的脉冲响应系数呈几何衰减, 当滞后阶数足够大时, 衰减速度则完全由更慢的双曲衰减所控制, 因而该部分不改变波动率记忆的根本属性. 不失一般性, 本文只考虑上述简化的模型.

针对这个模型, 分别产生样本量为  $n = 2\ 000$ ,

4 000 和 8 000 的序列, 每个样本量都进行 500 次独立试验. 为了减少初始值的影响, 每次重复试验都产生  $n + 5\ 000$  个数据点, 然后将前 5 000 个数据剔除. 在展开式  $(1 - B)^d = 1 - \sum_{j=1}^{\infty} \pi_j B^j$  中, 截取前 200 项. 实际上, 当  $j = 50$  时,  $\pi_j$  已经非常小, 因此并不会对最终结果产生任何显著影响. 估计结果的乖离率 (bias) 和均方差 (MSE) 汇报在表 1 中.

由表 1 可以看出, 随着样本量的增大, 两个波动率成分和连接函数部分的参数估计的乖离率和均方差都在逐渐减小, 与估计的渐近性质基本保持一致. 并且观察到, 在第一部分条件方差 HGARCH1 中,  $\alpha$  的真实值大于 1, 即此过程是非平稳的, 但参数估计仍然表现良好, 说明 EM 迭代估计法对于非平稳过程的参数估计也是有效的. 另外, 由于  $(1 - B)^d$  展开式中,  $\pi_j = O(j^{-1-d})$ , 表明脉冲响应系数的衰减速度与  $d$  的取值呈负相关, 即  $d$  值越大, 衰减越快,  $d$  的参数估计的渐近方差越小;  $d$  值越小, 衰减越慢,  $d$  的参数估计的渐近方差越大. 这在表 1 中也有所体现.

表 1 模型 (12) 中 EM 估计在有限样本下的表现

Table 1 Finite-sample performance of EM estimation in model (12)

模型	参数	真实值	$n$					
			2 000		4 000		8 000	
			偏差	均方差	偏差	均方差	偏差	均方差
HGARCH1	$\gamma_1$	0.15	0.141 7	0.113 8	0.063 2	0.024 1	0.033 3	0.006 6
	$\beta_1$	0.6	0.013 8	0.009 9	0.007 4	0.004 7	0.004 1	0.002 7
	$\alpha_1$	1.2	-0.017 9	0.004 5	-0.015 2	0.002 5	-0.007 7	0.001 0
	$d_1$	0.7	0.023 3	0.010 7	0.013 1	0.005 0	0.008 3	0.002 9
HGARCH2	$\gamma_2$	0.4	0.161 8	0.430 2	0.036 2	0.179 3	0.012 5	0.107 8
	$\beta_2$	0.4	0.079 5	0.043 7	0.064 2	0.032 1	0.019 9	0.016 0
	$\alpha_2$	0.8	-0.037 5	0.012 1	-0.035 1	0.006 0	-0.031 0	0.003 0
	$d_2$	0.6	0.100 0	0.047 0	0.077 6	0.034 2	0.033 7	0.014 9
连接函数	$\lambda_0$	1.8	0.055 0	0.107 4	0.070 9	0.064 3	0.078 3	0.019 2
	$\lambda_1$	-0.7	0.083 0	0.020 7	0.087 4	0.013 8	0.085 8	0.010 3

### 4 实证分析

为考察本模型在实际金融市场的适用性, 选取上证指数和美国标普 500 指数从 1995 - 01 - 01 ~ 2014 - 04 - 30 的日收盘价, 分别包含 4 687 与 4 866 个数据点. 这段时间中国经济经历高速增长、面临全球金融危机的冲击到逐步回归平稳的过程. 目标序列为对数收益率序列, 定义为

$y_t = \ln(P_t) - \ln(P_{t-1})$ , 其中  $P_t$  为第  $t$  天的收盘价. 图 1 和 2 分别描绘了上证指数和标普 500 指数的原始价格序列  $P_t$ 、对数收益率序列  $y_t$  以及  $y_t$  和  $y_t^2$  的自相关图. 从中可以看出, 上证指数在一阶矩  $\{y_t\}$  序列和二阶矩  $\{y_t^2\}$  序列都没有表现出明显的长记忆性, 而标普 500 指数则在  $\{y_t^2\}$  序列的自相关函数上展现出显著的长记忆现象. 这个经验特征与接下来的拟合结果也是一致的.

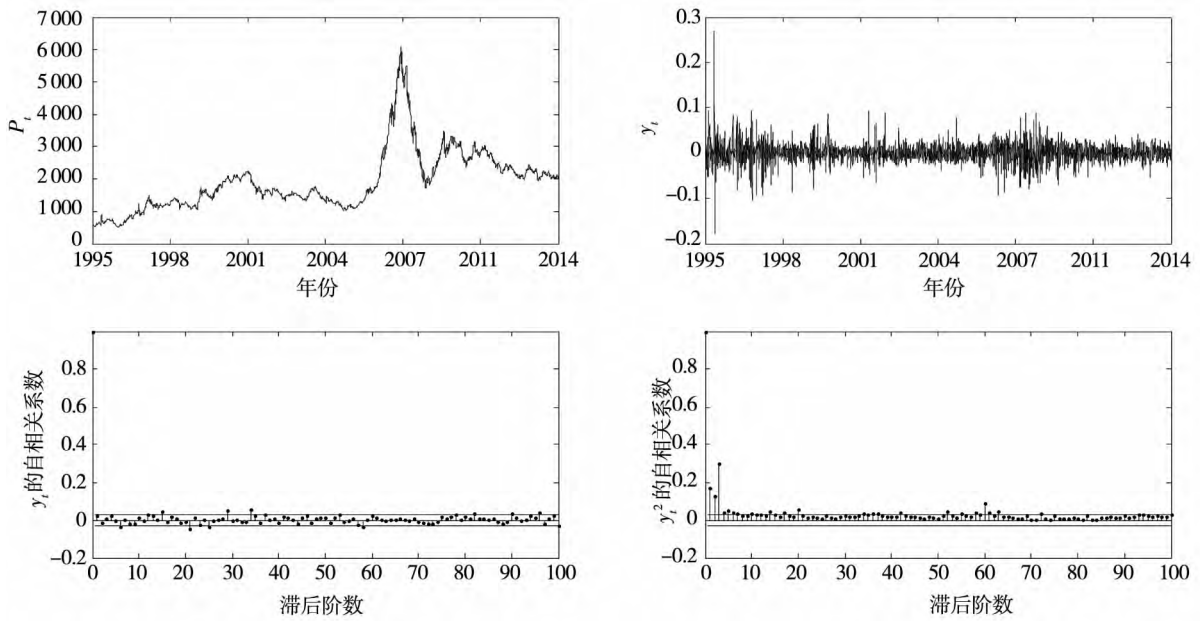


图1 上证指数的原始价格序列  $P_t$ 、对数收益率  $y_t$ 、 $y_t$  的自相关图和  $y_t^2$  的自相关图

Fig.1 Shangzheng index: Original price  $P_t$ , log-return  $y_t$ , autocorrelations of  $y_t$ , autocorrelations of  $y_t^2$

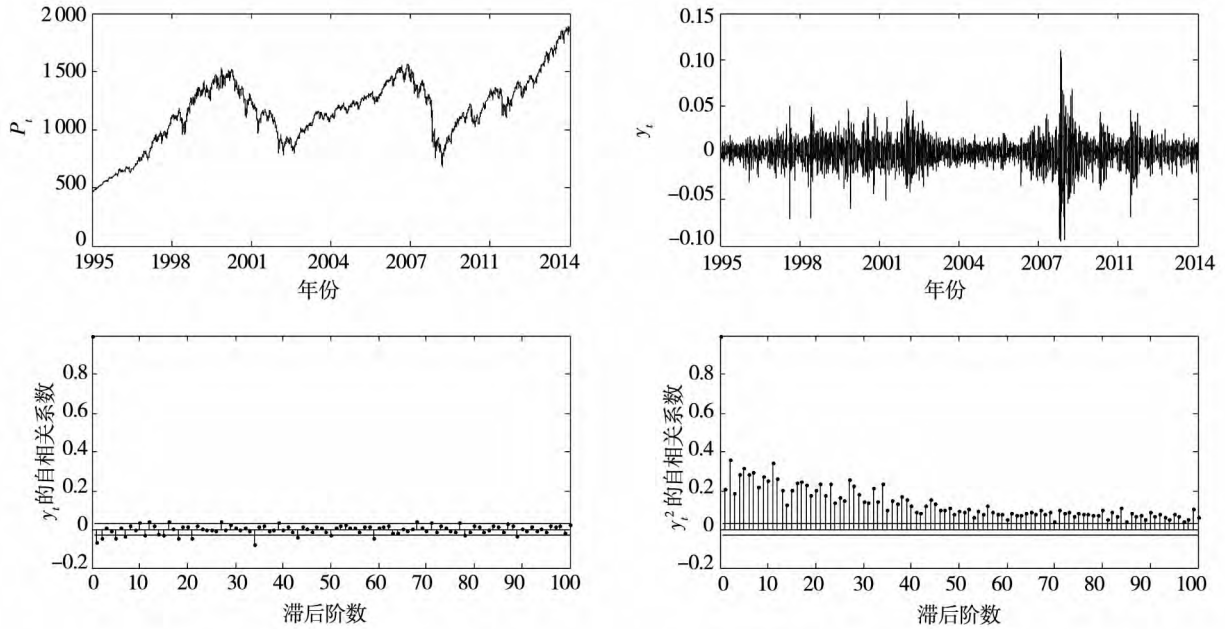


图2 标普 500 指数的原始价格序列  $P_t$ 、对数收益率  $y_t$ 、 $y_t$  的自相关图和  $y_t^2$  的自相关图

Fig.2 S&P500 index: Original price  $P_t$ , log-return  $y_t$ , autocorrelations of  $y_t$ , autocorrelations of  $y_t^2$

在拟合部分, 将每一个样本集都分为两个部分, 1 至  $n - 500$  用于样本内 (训练集) 拟合, 最后 500 个数据用于样本外 (测试集) 预测效果的评价. 表 2 报告了 DM-HGARCH 模型和单一 HY-GARCH 模型的拟合结果. 由于 HY-GARCH 和 HGARCH 有相同的表现, 为节约篇幅, 此处仅汇报单一 HY-GARCH 模型的实证

结果.

从表 2 发现, 对于标普 500 指数, DM-HGARCH 模型的拟合效果明显好于单一的 HY-GARCH 模型. 在 DM-HGARCH 下, 得到更大的对数似然函数值 ( $LLF$ ) 和更小的  $BIC$  值, 并且在两个波动率成分中, 记忆参数  $d$  的估计值都是显著区别于 1, 这表明长记忆性的确存在. 但是用 DM-HGARCH

对上证指数的拟合效果并没有超过单一-HYGARCH模型. 通过进一步分析,发现: 1) 在图 1 中,上证指数日收益率的平方序列在给定时间段上并没有展现长记忆特征,因而使用复杂模型不会带来更大的收益; 2) 用 DM-HGARCH 模型对上证指数的拟合结果中,  $\lambda_1$  的估计值为  $-0.115 2(0.063 5)$ . 这意味着在 5% 的显著水平下,该估计值并不显著,进一步表明动态混合的建模是不必要的. 为了对这两个结果做更精确的描述,实际上可以分别做以下两个假设检验: 对波动率长记忆存在性的检验和对混合概率是否是常数的检验. 关于两个检验的详细讨论可分别参考文献 [33] 和文献 [11],在此处不做展开; 3) 用 DM-HGARCH 对标普

500 指数的拟合结果中,得到两个完全不同的波动率成分. 其中一个波动率成分为振幅相对剧烈、衰减较为缓慢的过程(对应的  $\hat{\alpha} = 1.412 8, \hat{d} = 0.396 1$ ),另外一个则为振幅相对平缓、衰减较为迅速的过程(对应的  $\hat{\alpha} = 0.416 2, \hat{d} = 0.720 9$ ). 这在一定程度上刻画了金融市场始终动荡和平静两个状态交织的过程. 同时,连接函数中  $\hat{\lambda}_0 = 0.516 6(0.111 3)$  和  $\hat{\lambda}_1 = -0.595 2(0.039 7)$ ,表明在 1% 的显著水平下,该时间段内标普 500 指数的波动率序列存在不止一个成分,且混合概率是随着时间变化而变化的. 综上所述,本文认为在该时间段内,相对于上证指数,DM-HGARCH 模型对标普 500 指数的拟合效果更为理想.

表 2 参数估计结果

Table 2 Estimation results of parameters

参数	上证指数		标普 500 指数	
	模型			
	DM-HGARCH	HY-GARCH	DM-HGARCH	HY-GARCH
$\gamma_1$	0.127 0 (0.603 8)	0.001 5 (0.025 1)	0.135 1 <sup>***</sup> (0.018 6)	0.029 6 <sup>**</sup> (0.015 1)
$\beta_1$	0.346 3 <sup>***</sup> (0.092 0)	0.570 1 <sup>***</sup> (0.130 4)	0.366 5 <sup>***</sup> (0.050 6)	0.679 2 <sup>***</sup> (0.101 9)
$\delta_1$	—	0.219 8 <sup>***</sup> (0.080 6)	—	0.093 6 <sup>**</sup> (0.040 1)
$\alpha_1$	4.957 3 <sup>***</sup> (0.454 9)	1.088 0 <sup>***</sup> (0.030 4)	1.412 8 <sup>***</sup> (0.055 0)	1.030 5 <sup>***</sup> (0.013 9)
$d_1$	0.377 9 <sup>***</sup> (0.119 4)	0.409 6 <sup>***</sup> (0.080 3)	0.396 1 <sup>***</sup> (0.057 2)	0.597 9 <sup>***</sup> (0.100 8)
$\gamma_2$	0.227 1 <sup>***</sup> (0.043 6)		0.000 1 (0.003 1)	
$\beta_2$	0.282 8 <sup>***</sup> (0.055 2)		0.720 9 <sup>***</sup> (0.102 7)	
$\delta_2$	—		—	
$\alpha_2$	0.609 2 <sup>***</sup> (0.038 5)		0.416 2 <sup>***</sup> (0.009 7)	
$d_2$	0.356 9 <sup>***</sup> (0.062 2)		0.720 9 <sup>***</sup> (0.113 7)	
$\lambda_0$	-1.894 3 <sup>***</sup> (0.167 7)		0.516 6 <sup>***</sup> (0.111 3)	
$\lambda_1$	-0.115 2 <sup>*</sup> (0.063 5)		-0.595 2 <sup>***</sup> (0.039 7)	
<i>LLF</i>	-8 020.8	-8 034.2	-6 226.1	-6 343.4
<i>BIC</i>	16 124.9	16 110.2	12 536.0	12 728.7

注: \*\*\*、\*\* 和 \* 分别表示 1%、5%、10% 的显著性水平; 括号内为检验统计量相应的  $p$  值.



在预测阶段, 本文只考虑拟合效果更好的 DM-HGARCH 模型. 利用每个指数序列的最后 500 个数据, 考察 DM-HGARCH 模型对 VaR 的预测, 汇报上下尾 95% 的预测失败率以及区间覆盖率. VaR 的定义如下: 在给定显著水平  $\alpha$  下,  $P(y_t < -\text{VaR}_t^{(\alpha)}) = \alpha$  (下尾) 和  $P(y_t > \text{VaR}_t^{(\alpha)}) = \alpha$  (上尾). 在原假设  $H_0: f = \alpha$  ( $f$  为预测失败率)

下, 预测失败个数应该近似服从二项分布  $B(N, f)$ , 其中  $N = 500$ . 分别考虑 Kupiec<sup>[34]</sup> 无条件失败率检验和 Christoffersen<sup>[35]</sup> 条件失败率检验(后者考虑前后两天的预测失败率有相依关系). 前者检验统计量近似  $\chi_1^2$  分布, 后者检验量则近似  $\chi_2^2$  分布. 预测步长分别为 1 步和 5 步, 预测结果汇报在表 3 中.

表 3 基于 1 步向前和 5 步向前预报区间覆盖率的预报效果比较(95% 置信水平下)

Table 3 Comparison of forecasting performance based on the 1-day-ahead and 5-days-ahead predictive interval coverage rate (under 95% confidence level)

指数	步长	无条件失败率检验 统计量		条件失败率检验 统计量		样本外覆盖概率(%)	
		下尾	上尾	下尾	上尾	下尾	上尾
上证指数	1 步向前	3.850 8 (0.049 7)	3.850 8 (0.049 7)	4.911 0 (0.085 8)	4.911 0 (0.085 8)	96.8	96.8
	5 步向前	36.520 6 (0.000 0)	31.829 5 (0.000 0)	36.536 8 (0.000 0)	31.866 1 (0.000 0)	99.6	99.4
SP500	1 步向前	5.368 2 (0.020 5)	1.438 6 (0.230 4)	5.395 5 (0.067 4)	2.032 7 (0.361 9)	92.6	93.8
	5 步向前	0.431 7 (0.511 1)	0.133 3 (0.715 1)	0.706 1 (0.702 5)	0.859 7 (0.650 6)	94.35	95.36

注: 括号内为检验统计量相应的  $p$  值.

从表 3 可以看出, 当预测步长为 1 时, 上证指数和标普 500 指数基本都能通过检验( $p$  值超过 0.05 表示通过检验), 且上下尾覆盖率接近 95%. 但是当预测步长增加到 5 时, 只有标普 500 指数通过相应检验, 而对上证指数检验的所有  $p$  值均为极端小值, 因而拒绝原假设. Conrad<sup>[27]</sup> 指出, 长记忆 GARCH 模型在多步预测上会有更好的表现, 这与表 3 中给出的对标普 500 指数的 1 步预测和 5 步预测结果相吻合. 综上所述, 在对上证指数和标普 500 指数两支序列的实证分析中, 本文发现 DM-HGARCH 模型对后者在 1995 年 ~ 2014 年的日收益率序列的波动具有更好的解释和预测能力.

值得注意的是, 田存志等<sup>[36]</sup> 通过对现代经典金融理论和新金融理论进行梳理和归纳, 指出金融市场具有多元的记忆模式, 即在不同时间区段上会表现出不同的记忆类型, 有时表现为长记忆, 有时表现为短记忆, 甚至是白噪声. 因而这里的实证分析只是在基于给定的有限样本, 通过模型选择、风险测度的预测等标准进行比较, 得出 DM-HGARCH 模型在该时间段内对标普 500 指数

的波动率具有更好的解释和预测能力. 但是对不同时间段的不同样本, 可能会有不一样的实证结果. 对“金融市场是否存在长记忆”和长记忆的产生机制, 相应的理论基础在文献中尚且匮乏. 这也是作者未来的研究方向之一.

## 5 结束语

当时间序列跨度较长时, 其观测值往往来自不同的驱动机制, 混合模型是简单有效的捕捉结构变化的模型. 本文在 Li 等<sup>[8]</sup> 的基础上提出一类动态混合 HGARCH 模型, 该模型可以同时刻画波动率的长记忆性和结构变化, 并且在一定条件下, 存在协方差平稳解. 本文讨论了模型的弱平稳解存在条件, 研究了 EM 算法进行参数估计的大样本性质, 并用蒙特卡罗模拟展现该估计方法在有限样本下的表现. 最后利用该模型对上证指数和标普 500 指数进行实证分析. 从拟合结果和预测结果两方面来看, DM-HGARCH 模型对标普 500 指数的拟合和预测能力远好于单一的 HY-GARCH 模型, 但对上证指数的拟合效果一般. 值得注意的

是,本文中给出的动态混合模型可以推广到多个混合成分(即混合成分的个数大于2),但是如何确定混合成分的个数,目前仍是混合时间序列研究的难点,也是本文的局限所在。

### 参 考 文 献:

- [1] Ding Z, Granger C W J, Engle R F. A long memory property of stock market returns and a new model [J]. *Journal of Empirical Finance*, 1993, 1(1): 83 – 106.
- [2] Cont R. Empirical properties of asset returns: Stylized facts and statistical issues [J]. *Quantitative Finance*, 2001, 1(2): 223 – 236.
- [3] Baillie R T, Bollerslev T, Mikkeslen H O. Fractionally integrated generalized autoregressive conditional heteroscedasticity [J]. *Journal of Econometrics*, 1996, 74(1): 3 – 30.
- [4] Davidson J. Moment and memory properties of linear conditional heteroscedasticity models, and a new model [J]. *Journal of Business & Economic Statistics*, 2004, 22(1): 16 – 29.
- [5] Bollerslev T. Generalized autoregressive conditional heteroscedasticity [J]. *Journal of Econometrics*, 1986, 31(3): 307 – 327.
- [6] Mikosch T, Stărică C. Nonstationarities in financial time series, the long range dependence, and the IGARCH effects [J]. *Review of Economics and Statistics*, 2004, 86(1): 378 – 390.
- [7] Baillie R T, Morana C. Modelling long memory and structural breaks in conditional variances: An adaptive FIGARCH approach [J]. *Journal of Economic Dynamics and Control*, 2009, 33(8): 1577 – 1592.
- [8] Li M, Li G, Li W K. On a new hyperbolic GARCH model [J]. *Journal of Econometrics*, 2015, 189(2): 428 – 436.
- [9] Wong C, Li W K. On a mixture autoregressive conditional heteroscedastic model [J]. *Journal of the American Statistical Association*, 2001, 96(455): 982 – 995.
- [10] Cheng X X, Yu L H, Li W K. On a dynamic mixture GARCH models [J]. *Journal of Forecasting*, 2009, 28(3): 247 – 265.
- [11] Li M, Li W K, Li G. On mixture memory GARCH models [J]. *Journal of Time Series Analysis*, 2013, 34(6): 606 – 624.
- [12] Klein T, Walther T. Oil price volatility forecast with mixture memory GARCH [J]. *Energy Economics*, 2016, 58: 46 – 58.
- [13] Li G, Zhu Q, Liu Z, et al. On mixture double autoregressive time series models [J]. *Journal of Business & Economic Statistics*, 2017, 35(2): 306 – 317.
- [14] Catania L. Dynamic adaptive mixture models with an application to volatility and risk [J/OL]. *Journal of Financial Econometrics*, <https://doi.org/10.1093/jfinec/nbz018>, 2019 – 06 – 14.
- [15] Basatini F M, Rezakhah S. Markov switch smooth transition HYGARCH model: Stability an estimation [J]. *Communications in Statistics-Theory and Methods*, 2020, 49(10): 2384 – 2409.
- [16] 林 宇. 基于双曲线记忆 HY-GARCH 模型的动态风险 VaR 测度能力研究 [J]. *中国管理科学*, 2011, 19(6): 15 – 24.  
Lin Yu. Study on dynamic VaR measurement with hyperbolic memory GARCH [J]. *Chinese Journal of Management Science*, 2011, 19(6): 15 – 24. (in Chinese)
- [17] 曹广喜, 曹 杰, 徐龙炳. 双长记忆 GARCH 族模型的预测能力比较研究——基于沪深股市数据的实证分析 [J]. *中国管理科学*, 2012, 20(2): 41 – 49.  
Cao Guangxi, Cao Jie, Xu Longbing. Comparative research on forecast ability of double long memory GARCH family models: Empirical analysis of Shanghai and Shenzhen stock markets [J]. *Chinese Journal of Management Science*, 2012, 20(2): 41 – 49. (in Chinese)
- [18] 李云红, 魏 宇, 张帮正. 股票市场历史信息的长记忆性特征研究 [J]. *中国管理科学*, 2015, 23(9): 37 – 45.  
Li Yunhong, Wei Yu, Zhang Bangzheng. Research on the long memory characters of stock market's history information [J]. *Chinese Journal of Management Science*, 2015, 23(9): 37 – 45. (in Chinese)

- [19] 王安兴, 谭鲜明. 基于混合分布的中国股票波动风险因素的识别与分析 [J]. 中国管理科学, 2018, 26(2): 86-95.  
Wang Anxing, Tan Xianming. The analysis of Chinese stock volatile risk factors based on mixture distribution [J]. Chinese Journal of Management Science, 2018, 26(2): 86-95. (in Chinese)
- [20] 冯 玲, 吴江樵. 股票市场组合的市场风险度量研究——基于相关性模型 [J]. 系统科学与数学, 2016, 36(12): 2307-2324.  
Feng Ling, Wu Jiangqiao. Research on market risk measure for portfolio composed of stocks: Based on correlation models [J]. Journal of Systems Science and Mathematical Sciences, 2016, 36(12): 2307-2324. (in Chinese)
- [21] 龚玉婷, 郑 旭. 基于 Copula 模型的尾部相依性长记忆效应研究 [J]. 系统科学与数学, 2016, 36(6): 783-799.  
Gong Yuting, Zheng Xu. Study the long memory effects in tail dependence based on Copula models [J]. Journal of Systems Science and Mathematical Sciences, 2016, 36(6): 783-799. (in Chinese)
- [22] 李红权, 何敏园. 我国股市的对外溢出效应与国际影响力研究——基于 Copula-DCC-GARCH 模型 [J]. 系统科学与数学, 2017, 37(8): 1790-1806.  
Li Hongquan, He Minyuan. The spillover effect and international influence of China's stock market: Based on Copula-DCC-GARCH model [J]. Journal of Systems Science and Mathematical Sciences, 2017, 37(8): 1790-1806. (in Chinese)
- [23] 赵大萍, 王杨孜, 高杰英. 基于 GARCH 类模型的空头市场情景生成与投资组合 [J]. 系统科学与数学, 2017, 37(5): 1287-1299.  
Zhao Daping, Wang Yangzi, Gao Jieying. Scenario generation and portfolio selection in bear market with GARCH models [J]. Journal of Systems Science and Mathematical Sciences, 2017, 37(5): 1287-1299. (in Chinese)
- [24] 王天一, 刘 浩, 黄 卓. 基于混频数据抽样的已实现波动率长记忆模型 [J]. 系统工程学报, 2018, 33(6): 812-822.  
Wang Tianyi, Liu Hao, Huang Zhuo. Model for the long memory of realized volatility based on mixed data sampling [J]. Journal of Systems Engineering, 2018, 33(6): 812-822. (in Chinese)
- [25] 蒋 伟, 顾 研. 基于广义已实现测度的 Realized GARCH 模型改进及应用 [J]. 数量经济技术经济研究, 2019, (7): 156-173.  
Jiang Wei, Gu Yan. Forecasting VaR: Realized GARCH incorporating generalized realized measures [J]. The Journal of Quantitative & Technical Economics, 2019, (7): 156-173. (in Chinese)
- [26] 袁金建, 刘海龙, 刘小涛. 基于时变波动率的存款保险定价研究 [J]. 管理科学学报, 2019, 22(3): 113-126.  
Yuan Jinjian, Liu Hailong, Liu Xiaotao. Deposit insurance pricing based on time-varying volatility [J]. Journal of Management Sciences in China, 2019, 22(3): 113-126. (in Chinese)
- [27] Conrad C. Non-negativity conditions for the hyperbolic GARCH model [J]. Journal of Econometrics, 2010, 157(2): 441-457.
- [28] Conrad C, Haag B R. Inequality constraints in the fractionally integrated GARCH model [J]. Journal of Financial Econometrics, 2006, 4(3): 413-449.
- [29] Nelson D B, Cao C. Inequality constraints in the univariate GARCH model [J]. Journal of Business & Economic Statistics, 1992, 10(2): 229-235.
- [30] Dempster A P, Laird N M, Rubin D B. Maximum likelihood from incomplete data via the EM algorithm [J]. Journal of the Royal Statistical Society, Series B, 1977, 39(1): 1-38.
- [31] Louis T A. Finding the observed information matrix when using the EM algorithm [J]. Journal of the Royal Statistical Society: Series B, 1982, 44(2): 226-233.
- [32] Robinson P M, Zaffaroni P. Pseudo-maximum likelihood estimation of ARCH( $\infty$ ) models [J]. The Annals of Statistics, 2006, 34(3): 1049-1074.
- [33] Li M, Li G, Li W K. Score tests for hyperbolic GARCH models [J]. Journal of Business & Economic Statistics, 2011, 29(4): 579-586.
- [34] Kupiec P. Techniques for verifying the accuracy of risk measurement models [J]. Journal of Derivatives, 1995, 3(2): 73

– 84.

[35] Christoffersen P. Evaluating interval forecasts [J]. *International Economic Review*, 1998, 39(4): 841 – 862.

[36] 田存志, 程富强, 付 辉. 关于金融市场长记忆性研究的若干争论 [J]. *经济学动态*, 2016, (6): 141 – 149.

Tian Cunzhi, Cheng Fuqiang, Fu Hui. Some debates on the study of long memory in financial market [J]. *Economic Perspectives*, 2016, (6): 141 – 149. (in Chinese)

## Estimation and forecasting of mixture HGARCH model

LI Mu-yi<sup>1, 2, 3</sup>, FANG Ying<sup>1, 2, 3\*</sup>

1. MOE Key Laboratory of Econometrics ( Xiamen University) , Xiamen 361005 , China;
2. The Wang Yanan Institute for Studies in Economics ( WISE) , Xiamen University and Department of Statistics, School of Economics, Xiamen University, Xiamen 361005, China;
3. Fujian Key Laboratory of Statistical Science ( Xiamen University) , Xiamen 361005, China

**Abstract:** In the framework of GARCH models, Li et al. (2015) proposed a new hyperbolic GARCH model (denoted by HGARCH), which can parameterize both long memory decay and dramatic amplitude of volatilities as the HY-GARCH model (Davidson, 2004). What's more, the non-negative restrictions on the parameters in the HGARCH model are more tractable than those counterparts in HY-GARCH models. However, it is well known that when the time series covers a long time span, a constant structure is usually inadequate to capture possible structure changes. To address this issue, this paper constructs a new dynamic mixture hyperbolic GARCH (denoted by DM-HGARCH) model. The DM-HGARCH model accommodates covariance stationarity, long memory and structural changes in volatilities simultaneously. Conditions for the existence of weak stationary solutions are investigated and the EM algorithm is employed for parameter estimation. Monte-Carlo simulations are conducted to evaluate the finite-sample performance. Finally, the new model is applied to the daily log returns of Shanghai stock exchange (SSE) index in China and S&P 500 index in USA respectively. The empirical study illustrates that in the given sample period, the DM-HGARCH performs better on the latter index than the former one in terms of both in-sample fitting and out-of-sample forecasting.

**Key words:** dynamic mixture; long memory volatility; HGARCH model; EM algorithm