

doi:10.16055/j.issn.1672-058X.2020.0001.016

# 区间型金融时间序列的长记忆性研究\*

丁勤祥<sup>a</sup>, 王哲<sup>a</sup>, 王艺宁<sup>b</sup>, 李铭源<sup>a</sup>

(安徽大学 a. 经济学院; b. 数学科学学院, 合肥 230601)

**摘要:**研究金融时序的长记忆性能够帮助人们更加准确地刻画金融市场的特征,而在现有研究中,有关区间型金融时序长记忆性的研究很少。因此,考虑了区间型金融时序蕴含的长记忆性特征及其基于现有实值金融时序长记忆性建模的区间值时序预测模型,首先,将区间数表示成区间中心和区间半径的形式;然后分别对中心和半径序列进行长记忆性检验,并对具有长记忆性的序列进行组合预测;最后,以上证综指和深证综指的区间股指为实证对象进行验证。实证结果表明:上证综指的区间股指具有明显的长记忆性,且组合预测能够显著提高区间型金融时序的预测精度。

**关键词:**区间金融时序;长记忆性;Hurst 指数

**中图分类号:**O221.1

**文献标志码:**A

**文章编号:**1672-058X(2019)06-0104-08

## 0 引言

在金融市场日益复杂的趋势下,“有效市场假说”以及“随机游走理论”正遭受着人们的质疑。金融时序是否具有长记忆性以及如何对具有长记忆性的金融时序进行预测成为近些年的研究热点。有关金融时间序列长记忆性的研究主要包括两方面:一方面,检验金融时间序列是否具有长记忆的特性;另一方面,构建能够刻画金融时序长记忆性的预测模型。在现有研究中,常用 Hurst 指数来衡量一个金融时序是否具有长记忆特性,而计算 Hurst 指数的方法主要有 R/S 法、DFA 法、DMA 法、GHE 法等<sup>[1-4]</sup>。同时,也可以通过假设检验的方法检验长记忆性,如修正的 R/S 法、KPSS 法、V/S 法等<sup>[5-6]</sup>。首次明确提出长记忆性的是水文学家 Hurst,他在对水文数据进行分析时发现其时间序列

中存在着长记忆的特性,并给出了检验时间序列长记忆性的方法——重标极差分析法,即 R/S 分析法<sup>[1]</sup>。但 R/S 分析法也不是完全没有缺点。1991 年 Lo 指出<sup>[5]</sup>,当特定时间序列中包含短记忆性、存在异质性、具有非平稳性时,R/S 分析的结果将会大打折扣。因此在 R/S 分析法的基础上,Lo 对经典的 R/S 分析法进行了修正,提出了修正的重标极差分析法,即修正的 R/S 法(MRS 法)。除此之外,Peng<sup>[2]</sup>在做心跳和 DNA 研究时最先采用了去趋势法(DFA)来检验时序的长记忆性。Matteo<sup>[4]</sup>采用广义 Hurst 指数作为工具,研究了 4 个国家的股票市场指数的标度特点,发现市场标度特性均偏离了“随机游走假设”下的情形,其偏离程度和市场发展程度有明显的关系。有关检验时间序列长记忆性的方法还有很多,如 KPSS 检验法、拉格朗日乘子检验法(LM)等,但这些方法各有各的优势,具体的表现还需要在实证中进行进一步验证。刻画长记忆

收稿日期:2019-04-09;修回日期:2019-05-15.

\* 基金项目:安徽大学大学生创新创业训练计划(201710357455;201810357203;201810357206;201810357518;201810357519).

作者简介:丁勤祥(1994—),男,安徽阜阳人,硕士研究生,从事组合预测与经济预测研究.

性的模型可以分为长记忆均值模型和长记忆波动率模型。较为人所熟知的 ARMA 模型体系是一种短记忆均值模型,当序列具有长记忆性时,这种模型将不再适用,因此需要对现有的短记忆模型进行改进。首先,Granger 等<sup>[7]</sup>提出了分数维差分噪声模型,即 FDN 模型。FDN 模型可以很好地拟合时间序列中的长记忆特性,但其仅考虑了时序中的长记忆性忽略了短记忆性。在此基础上,Granger 和 Joyeux<sup>[8]</sup>提出了分整自回归移动平均模型,即 ARFIMA( $p, d, q$ )模型。ARFIMA 模型用参数  $p$  和  $q$  来描述时间序列的短记忆性,用参数  $d$  来描述时间序列的长记忆性。ARFIMA 模型弥补了 FDN 模型的不足,既刻画了时序中的长记忆性,又考虑了短记忆性。和 ARFIMA 模型类似,Baillie 等<sup>[9]</sup>提出了 FIGARCH( $p, d, q$ )模型,用来刻画波动率序列中的长记忆特性。FIGARCH 模型实际上就是关于扰动项的 ARFIMA 模型,原理和收益率的长记忆性十分类似,其中参数  $d$  反映了其作用于远距离观测值之间的效果以及双曲率缓慢下降的长记忆性。

综合现有研究发现,有关金融时序长记忆性的研究存在以下几方面问题:首先,研究对象多是收益率序列或波动率序列等点值序列,极少有针对区间型的金融时间序列而展开的研究,而区间型金融时序在金融市场中广泛存在;其次,检验长记忆性时使用方法过少,不足以刻画长记忆性特征,检验结果不具有说服力;最后,预测模型多为单一模型,不能很好地反映金融时序的长记忆特性,预测精度不高,有必要对现有模型进行整合,发挥各个模型的优势。针对存在的问题,以国内上证综指和深证综指的区间型股指序列为研究对象,首先把区间型数据表示成区间中心和区间半径的形式,其中,区间中心表示序列的期望水平,而区间半径表示序列的波动水平;然后,分别使用 R/S 法、修正的 R/S 法、DFA 法、DMA 法和 GHE 法检验区间中心和区间半径是否具有长记忆性,并对计算结果进行显著性检验;之后,以具有长记忆性的金融时序为基础,针对区间中心序列,分别建立 ARFIMA、GM-ARIMA、GM-ARFIMA 3 种单项预测模型,针对区间半径序列,分别建立 FIGARCH、GM-GARCH、GM-FIGARCH 3

种单项预测模型;最后,以 3 种单项预测结果为基础,建立区间中心和区间半径的组合预测模型,并根据 MSEP、MSEL、MSEI、MRIE 4 种评价指标对模型的精度和有效性进行评价比较。经过实例分析对比,结果表明:相比较于单一的预测模型,组合预测模型更能发挥各模型的优势,更好地刻画区间型金融时序可能蕴含的长记忆特性。

## 1 基本概念

**定义 1**<sup>[10]</sup> 令  $a = [a^-, a^+] = (c; r)$ , 其中  $\{x | a^- \leq x \leq a^+, a^-, a^+ \in \mathbf{R}\}$ , 则称  $a$  为一个区间数。其中  $c = (a^- + a^+)/2$  为区间数的中心,  $r = (a^+ - a^-)/2$  为区间数的半径。当  $a^- = a^+$  时, 区间数退化为普通的实数。若  $a^- \geq 0$ , 则称  $a$  为非负区间数。设  $\Omega$  为所有非负区间数的集合, 同时,  $a = [a^-, a^+]$ ,  $b = [b^-, b^+] \in \Omega$  为任意两个非负区间数,  $\lambda \geq 0$ , 则区间数的运算法则如下:

$$a+b = [a^-+b^-, a^++b^+]$$

$$a-b = [a^- - b^+, a^+ - b^-]$$

$$\lambda a = [\lambda a^-, \lambda a^+]$$

设某一区间序列观测值  $\{x_t = [x_t^-, x_t^+], t = 1, 2, \dots, N\}$  为金融时间序列观测值, 则称此序列为区间型金融时间序列。

**定义 2**<sup>[11]</sup> 由于平稳且可逆的 ARMA 模型的自相关函数  $\rho_k$  是有界的, 即  $|\rho_j| \leq cm^{-k}, 0 < m < 1$ , 对于给定的一个离散时间序列  $\{y_t, t = 1, 2, \dots, N\}$ , 其滞后  $j$  阶自相关函数为  $\rho_j$ , 那么如果  $\lim_{n \rightarrow \infty} \sum_{j=-n}^n |\rho_j|$  是非有限的, 则序列  $\{y_t, t = 1, 2, \dots, N\}$  具有长记忆性。

相较于时间序列的短记忆性, 长记忆性又被称之为分数单积过程。当一个序列的自相关函数类似于非平稳过程, 而差分后的序列表现为过度差分, 则表明这个序列很可能具有长记忆性。一般来说, 常用 Hurst 指数来判断一个序列是否具有长记忆性。当 Hurst 指数明显大于 0.5 时, 说明序列具有长记忆性, 且 Hurst 指数越大, 长记忆性越明显。

**定义 3**<sup>[11]</sup> 将长度为  $M$  的时间序列  $\{X_i\}$  划分为  $A$  个长度为  $N$  的等长子区间, 重标极差可以表示

为  $(R/S)_N = \frac{1}{A} \sum_{a=1}^A R_a/S_a$ , 其中,  $R_a$  为子区间  $a$  内数据的极差,  $S_a$  为子区间  $a$  内数据的标准差。对于不同的区间划分  $N$ , 存在以下关系:

$$(R/S)_N = kN^H$$

其中:  $H$  为 Hurst 指数,  $k$  为常数。则称上述方法为重标极差分析法, 即 R/S 分析法。

**定义 4**<sup>[8]</sup> 若平稳时间序列  $\{x_t\}$  满足分数阶差分方程, 即

$$\Phi(B)(1-B)^d(x_t - \mu) = \Theta(B)\varepsilon_t$$

则称时序  $\{x_t\}$  服从分整自回归移动平均模型, 即 ARFIMA( $p, d, q$ ) 模型。其中:  $B$  为滞后算子,  $\mu$  为  $x_t$  的均值,  $\{\varepsilon_t\}$  为白噪声序列,  $|d| < 0.5$ ,  $(1-B)^d$  为分数差分算子,  $\Phi(B)$  和  $\Theta(B)$  根的模均在单位圆外。ARFIMA( $p, d, q$ ) 模型用  $p+q$  个参数来描述短记忆特征, 用  $d$  个参数来描述长记忆特征, 综合考虑了长短期记忆的特性。

相较于单一的预测模型, 组合预测模型更能发挥各个模型的优势, 能够显著提高模型的预测精度。但传统的组合预测模型中, 单项预测模型在每个时刻的权系数是固定的, 不能使组合预测模型的精度达到最高。为了克服单项预测模型在每个时刻取固定权系数的缺陷, 可在组合预测中引入诱导有序加权几何平均算子。

**定义 5**<sup>[12]</sup> 设  $\langle u_1, a_1 \rangle, \langle u_2, a_2 \rangle, \dots, \langle u_n, a_n \rangle$  为  $n$  个二维数组, 令

$$IOWGA_w(\langle u_1, a_1 \rangle, \langle u_2, a_2 \rangle, \dots, \langle u_n, a_n \rangle) =$$

$$\prod_{i=1}^n a_{u-index(i)}^{w_i}$$

则称函数  $IOWGA_w$  是由  $u_1, u_2, \dots, u_n$  所产生的  $n$  维诱导有序加权几何平均算子, 简记为 IOWGA 算子, 称  $u_i$  为  $a_i$  的诱导值, 其中  $u-index(i)$  是  $u_1, u_2, \dots, u_n$  中按从大到小的顺序排列的第  $i$  个大的数的下标,  $W = (w_1, w_2, \dots, w_n)^T$  是 IOWGA 的加权向量。

**定义 6** 设偏好系数  $\alpha, \alpha \in [0, 1]$ , 有如下模型:

$$\min F(W) = \alpha \left| \sum_{t=1}^N \left| \sum_{i=1}^m w_i e_{p-index(it)} \right| \right| + (1 - \alpha) \left| \sum_{t=1}^N \left| \sum_{i=1}^m w_i e_{q-index(it)} \right| \right|$$

$$\text{s. t. } \begin{cases} \sum_{i=1}^m w_i = 1 \\ w_i \geq 0, i = 1, 2, \dots, m \end{cases}$$

则称该模型为基于  $L_1$  范数的 IOWGA 算子的区间组合预测模型。其中,  $e_{p-index(it)}$  为基于 IOWGA 算子的区间中心预测误差,  $e_{q-index(it)}$  为基于 IOWGA 算子的区间半径预测误差。

## 2 理论研究过程

目前, 有关区间型时间序列的主要处理方式是把区间型数据转化为点值数据, 然后分别对点值数据进行研究。把区间型数据转化为点值数据的方式有两种, 一种是把区间数据的左端和右端分别看成两个独立的点值序列; 另一种是把区间型数据转化成区间中心和区间半径的形式。设某区间型金融时间序列观测值为  $\{x_t = [x_t^-, x_t^+], t = 1, 2, \dots, N\}$ , 采用第二种方式对此区间型序列进行处理。下面介绍区间型金融时间序列长记忆性的研究过程。

**Step1** 把区间数转化为区间中心和区间半径的形式。

根据定义 1 把区间型金融时间序列表示成区间中心和区间半径的形式。转化后的区间序列观测值为  $\{x_t = [x_t^-, x_t^+] = (c_t; r_t), t = 1, 2, \dots, N\}$ 。其中,  $\{c_t, t = 1, 2, \dots, N\}$  为区间中心序列, 表示区间序列观测值  $\{x_t\}$  的期望水平,  $\{r_t, t = 1, 2, \dots, N\}$  为区间半径序列, 表示区间序列观测值  $\{x_t\}$  的波动水平。

**Step2** 分别对区间中心序列和区间半径序列进行长记忆性检验。

分别采用 R/S 分析法、修正的 R/S 分析法、DFA 法、DMA 法以及 GHE 法对区间中心序列  $\{c_t, t = 1, 2, \dots, N\}$  和区间半径序列  $\{r_t, t = 1, 2, \dots, N\}$  进行长记忆性检验。其中, R/S 分析法、DFA 法、DMA 法以及 GHE 法通过计算 Hurst 指数来检验序列是否具有长记忆性, 而修正的 R/S 分析法则通过假设检验的方式来进行。

在用 R/S 分析法、DFA 法、DMA 法以及 GHE 法计算序列的 Hurst 指数时, 由于样本序列的有限性, 可能会出现抽样误差, 导致测量结果并不一定

准确,因此要对检验结果进行显著性检验。这里将采用 Bootstrapping 算法(自举法)对以上4种方法的 Hurst 指数计算结果进行进一步的显著性检验。

基于以上5种方法以及显著性检验的结果,综合判断区间中心序列  $\{c_t, t=1, 2, \dots, N\}$  和区间半径序列  $\{r_t, t=1, 2, \dots, N\}$  是否具有长记忆性。

**Step3** 针对具有长记忆性的序列建立单项预测模型。

由于区间中心序列  $\{c_t\}$  和区间半径序列  $\{r_t\}$  分别表示区间序列观测值  $\{x_t\}$  的期望水平和波动水平,因此需要分别对具有长记忆性的区间中心序列和区间半径序列进行建模。具体来说,针对具有长记忆性的区间中心序列,分别建立 ARFIMA、GM-ARIMA、GM-ARFIMA 模型来反映区间序列  $\{x_t\}$  的期望水平,针对具有长记忆性的区间半径序列,分别建立 FIGARCH、GM-GARCH、GM-FIGARCH 模型来反映区间序列  $\{x_t\}$  的波动水平。

其中,GM-ARIMA、GM-ARFIMA、GM-GARCH、GM-FIGARCH 模型为灰色模型 GM 分别和 ARIMA、ARFIMA、GARCH、FIGARCH 模型的组合模型。这4个组合模型先分别用主体模型 ARMA 模型、ARFIMA 模型、GARCH 模型、FIGARCH 模型对序列进行拟合,然后由 GM 模型对模拟值和真实值之间的偏差进行估计,最后将两个预测结果求和。这种组合模型可以发挥两个模型的特点,提高最终预测精度。

**Step4** 建立区间中心和区间半径的组合预测模型。

根据定义6,以区间中心序列和区间半径序列的单项预测结果为基础,建立基于  $L_1$  范数的 IOWGA 算子的长记忆区间组合预测模型。这里,偏好系数  $\alpha$  的确定没有统一的准则,由决策者进行选择。如果决策者更加看重区间中心,则  $\alpha$  相对较大;反之,则  $\alpha$  相对较小。一般来说,会对  $\alpha$  取不同值时的情况进行对比分析,选择模型预测精度较大时的  $\alpha$  取值。

**Step5** 将区间中心预测值和区间半径预测值转化为区间数的形式,并计算模型的预测精度。

根据定义1,将区间中心预测值和区间半径预测值相结合,转化为区间数的形式。具体来说,把

模型 ARFIMA 和 FIGARCH 的预测值相结合,把模型 GM-ARIMA 和 GM-GARCH 的预测值相结合,把 GM-ARFIMA 和 GM-FIGARCH 模型的预测值相结合,最后把组合预测模型的预测值相结合。为了反映模型的预测性能,这里用以下4个误差指标来度量模型的精度<sup>[13]</sup>。

(1) 平均区间位置误差平方和:

$$F_{\text{MSEP}} = \frac{1}{N} \sum_{t=1}^N (c_t - \hat{c}_t)^2 \quad (1)$$

(2) 平均区间长度误差平方和:

$$F_{\text{MSEP}} = \frac{1}{N} \sum_{t=1}^N (r_t - \hat{r}_t)^2 \quad (2)$$

(3) 平均区间误差平方和:

$$F_{\text{MSEP}} = F_{\text{MSEP}} + F_{\text{MSEP}} \quad (3)$$

(4) 平均区间相对误差和:

$$F_{\text{MSEP}} = \frac{1}{N} \sum_{t=1}^N \frac{|c_t - \hat{c}_t|}{r_t + \hat{r}_t} \quad (4)$$

### 3 实证研究

首先,分别选取上证综指和深证综指的区间股指作为研究对象,运用5种检验长记忆性的方法对其进行长记忆性检验。其中,区间股指由相应股指每日的最高价和最低价构成。然后,针对具有长记忆性的区间股指,建立基于  $L_1$  范数的 IOWGA 算子区间组合预测模型。上证综指和深证综指的股票指数采样区间均为 2000-01-04—2018-12-28,数据均为日数据(数据来源:同花顺网)。

#### 3.1 区间中心和区间半径的长记忆性检验

为了进一步研究区间型数据,首先采用均值、标准差、偏度、峰度以及 normtest. p 等指标对区间中心和区间半径序列进行正态性检验。检验结果如表1所示。

从表1可以看出,上证综指的区间中心和区间半径序列均呈现为右偏、尖峰分布,且没有通过正态性检验。深证综指的区间中心序列呈现右偏,但峰度小于0,不具有尖峰特征,但区间半径序列的右偏以及尖峰特征较为明显,区间中心和区间半径序列也没有通过正态性检验。综合来看,上证综指和深证综指的区间中心和区间半径序列具有“尖峰厚

尾”特征,不服从正态分布,下面采用长记忆性检验方法对序列进行进一步的研究。

表 1 区间型金融时序正态性检验结果

Table 1 Normality test results of interval financial time series

金融时序	序列	均值	标准差	偏度	峰度	normtest. p
上证综指	区间中心	2434.01	901.53	0.92	1.13	0.00
	区间半径	23.11	21.93	2.95	12.25	0.00
深证综指	区间中心	999.32	571.85	0.71	-0.29	0.00
	区间半径	10.06	10.67	3.75	22.98	0.00

针对上证综指和深证综指的区间中心和区间半径序列,分别采用经典的 R/S 分析法、修正的 R/S 分析法、DFA 法、DMA 法以及 GHE 法对其进行长记忆性检验。由于修正的 R/S 分析法需要首先确定参数  $q$ ,而目前对于参数  $q$  的选取没有统一的标准,因此,分别针对  $q=1,3,5,7,9$  时的情况进行检验。检验结果见表 2 和表 3。

从表 2 可以看出,上证综指的区间中心和区间半径序列在 4 种方法下的 Hurst 指数计算结果明显高于 0.5,说明上证综指的区间中心和区间半径序列具有明显的长记忆性。相比较于上证综指,深证综指区间中心序列在 4 种方法下的 Hurst 指数计算结果并不具

有一致性,在 R/S 分析法下,区间中心序列具有长记忆性,但其他 3 种方法的计算结果并不支持这个结论。而深证综指区间半径序列在 4 种方法下的 Hurst 指数计算结果均表明区间半径序列具有长记忆性。由表 3 看出,在修正的 R/S 分析法下,在参数  $q$  取不同值时,上证综指的检验结果表明,其区间中心和区间半径序列均具有长记忆性。深证综指只在  $q$  取 1 和 3 时,在 10% 的显著性水平下是显著的。在用经典的 R/S 分析法、DFA 法、DMA 法以及 GHE 法对序列进行长记忆性检验之后,为了保证检验结果的有效性,采用 Bootstrapping 算法(自举法)对计算结果进行显著性检验,显著性检验结果如表 4 所示。

表 2 上证综指和深证综指区间中心和区间半径序列的 Hurst 指数计算结果

Table 2 Hurst index calculation results for the central and interval radius sequences of the Shanghai Composite Index and the Shenzhen Component Index

金融时序	序列	R/S 分析法	DFA 法	DMA 法	GHE 法
上证综指	区间中心	0.672 4	0.636 3	0.605 2	0.615 8
	区间半径	0.651 3	0.601 8	0.597 9	0.625 4
深证综指	区间中心	0.563 2	0.519 8	0.472 4	0.485 7
	区间半径	0.602 1	0.573 2	0.594 8	0.568 1

表 3 上证综指和深证综指区间中心和区间半径序列的修正 R/S 分析结果

Table 3 Corrected R/S analysis results of the central and interval radius sequences of the Shanghai Composite Index and the Shenzhen Component Index

金融时序	序列	$q=1$	$q=3$	$q=5$	$q=7$	$q=9$
上证综指	区间中心	3.56 ***	3.71 ***	3.12 **	3.35 ***	2.98 **
	区间半径	2.94 **	2.87 **	2.96 ***	3.01 ***	2.81 **
深证综指	区间中心	2.19 *	2.14 *	2.01	1.98	1.86
	区间半径	2.41 *	2.03	1.91	1.67	1.56

(注:\*, \*\*, \*\*\*分别表示在 10%、5% 以及 1% 的水平下统计显著)

表 4 长记忆性计算结果的显著性检验

Table 4 Significance test of long memory calculation results

金融时序	序列	R/S 法	DFA 法	DMA 法	GHE 法	置信区间
上证综指	区间中心	0.672 4	0.636 3	0.605 2	0.615 8	[0.442 8,0.561 2]
	区间半径	0.651 3	0.601 8	0.597 9	0.625 4	[0.436 9,0.575 1]
深证综指	区间中心	0.563 2	0.519 8	0.472 4	0.485 7	[0.407 4,0.614 6]
	区间半径	0.602 1	0.573 2	0.594 8	0.568 1	[0.421 3,0.572 7]

由显著性检验结果可以看出,在 95% 的置信水平下,上证综指的区间中心和区间半径列在 4 种方法下计算的 Hurst 指数均落在置信区间以外,拒绝不具有长记忆性的原假设,说明上证综指的区间股指具有长记忆性。而深证综指的中心序列在 4 种方法下的 Hurst 指数没有通过显著性检验,但半径序列只在前 3 种方法下通过显著性检验,在 GHE 法下没有通过显著性检验。因此,显著性检验结果进一步说明深证综指区间股指序列的长记忆性并不明显。

综上所述,仅有上证综指的区间序列股指具有明显的长记忆性,而深证综指的区间股指并没有发现明显的长记忆性。接下来,对具有长记忆性的上证综指的区间序列股指进行组合预测建模。

### 3.2 基于 IOWGA 算子的长记忆性区间组合预测

采用 ARFIMA、GM-ARIMA、GM-ARFIMA 3 种模型对上证综指的区间中心序列进行单项预测建模;采用 FIGARCH、GM-GARCH、GM-FIGARCH 3 种模型对上证综指区间半径序列进行单项预测建模。上证综指的采样区间为 2000-01-04—2018-12-28,其中,2018 年 11 月 30 日之前的样本构成训练集,用于模型参数的估计,11 月 30 日之后的 20 个样本构成测试集,用于检验模型的预测精度。

经过计算,上证综指区间中心序列和区间半径序列的长记忆参数  $d$  分别为 0.213 4、0.207 1。而参数  $p$ 、 $q$  滞后阶数的确定则根据 AIC 和 SIC 准则来进行,得到最终的模型见表 5。

表 5 上证综指区间中心和区间半径模型的参数

Table 5 Parameters of the central and interval radius models of the Shanghai Composite Index

序列	单项方法 I	单项方法 II	单项方法 III
区间中心	ARFIMA(2,0.213 4,1)	GM(1,1)-ARIMA(2,1,0)	GM(1,1)-ARFIMA(2,0.213 4,1)
区间半径	FIGARCH(1,0.207 1,1)	GM(1,1)-GARCH(1,1)	GM(1,1)-FIGARCH(1,0.207 1,1)

运用定阶之后的模型对采样区间的样本进行拟合,得到各时刻的预测值。然后以各单项预测方法的预测结果为基础,对上证综指的区间股指进行区间组合预测建模。

在确定偏好系数时,不同的决策者有不同的准

则,为了探究不同偏好系数对预测结果的影响,分别计算了偏好系数取不同值时权系数的计算结果。由于区间中心和区间半径的预测方法不同,所以区间中心和区间半径的权系数也不相同。表 6 给出了上证综指权系数的计算结果。

表 6 上证综指不同偏好系数所对应的权系数

Table 6 The weight coefficient corresponding to different preference coefficients of the Shanghai Composite Index

偏好系数	$l_{c,1}$	$l_{c,2}$	$l_{c,3}$	$l_{r,1}$	$l_{r,2}$	$l_{r,3}$
$\alpha=0.1/0.2$	0.923 7	0.076 3	0	0.942 8	0.057 2	0
$\alpha=0.3/0.4/0.5$	0.851 7	0.128 4	0.019 9	0.897 2	0.094 1	0.008 7
$\alpha=0.6/0.7/0.8$	0.907 3	0.061 5	0.031 2	0.962 9	0.037 1	0
$\alpha=0.9$	0.942 0	0.058 0	0	0.971 6	0.028 4	0

上证综指权系数计算结果表明,偏好系数取不同值时,有 4 组不同的权系数。当偏好系数取

0.1/0.2 时,权系数相同;当偏好系数取 0.3/0.4/0.5 时,权系数相同;当偏好系数取 0.6/0.7/0.8

时,权系数相同。而在一组权系数中,3 个权系数是依次递减的,这也符合诱导有序加权几何平均算子的思想,预测精度较高的单项预测方法应该优先赋予较大的权重。

根据不同偏好系数所对应的权系数,可以依次计算训练集以及测试集中不同时间点的组合预测拟

合值。为了评估模型的预测效果,根据式(1)~(4)提供的 4 个误差指标计算不同方法的区间预测误差。表 7 给出了上证综指测试集中不同模型的区间预测误差。其中,组合预测 I—IV 依次代表偏好系数取不同值时所得到的 4 组不同权系数所对应的模型。

表 7 上证综指测试集区间预测误差

Table 7 Interval prediction error of the Shanghai Composite Index testing set

预测模型	$F_{MSEP}$	$F_{MSEP}$	$F_{MSEP}$	$F_{MSEP}$
ARFIMA&FIGARCH	254.707 3	1.754 6	256.461 9	0.506 3
GM-ARIMA&GM-GARCH	875.453 8	4.958 2	880.4120	0.989 3
GM-ARFIMA&GM-FIGARCH	42.201 3	0.756 4	42.957 7	0.214 0
组合预测 I	24.324 9	0.105 3	24.430 2	0.061 5
组合预测 II	18.714 6	0.084 3	18.798 9	0.024 1
组合预测 III	15.664 8	0.063 5	15.728 3	0.013 4
组合预测 IV	21.374 3	0.094 0	21.468 3	0.047 9

从上证综指测试集的 4 个区间预测误差指标可以看出,在 3 种单项预测方法中,混合长记忆性模型 GM-ARFIMA 和 GM-FIGARCH 表现得较好,4 种预测误差均较低。而相比较于单项预测模型,组合预测模型的表现要普遍好于单项预测模型,预测精度明显提高,说明组合预测能够提高区间型金融时间序列的预测精度。在 5 种组合预测模型中,当偏好系数取 0.6~0.8 时,组合预测模型的预测精度相对较高,因此可以采用组合预测 III 对未来的区间股指进行预测。

## 4 结束语

首先总结了现有研究中有关金融时间序列长记忆性的研究状况,针对发现的问题,提出了区间型金融时间序列的长记忆性检验方法及其预测模型,并以上证综指和深证综指的区间型股票指数为基础,验证了区间型长记忆性检验方法以及预测模型的有效性。根据预测连贯性的原则,可以根据提出的模型对未来进行区间预测。在单项预测方法已取得预测结果的基础上,在预测区间  $[n+1, n+2, \dots, n+k, k=1, 2, \dots]$  上预测精度的确定原则是依据各个单项预测方法在样本区间内近几期拟合平均精度来计算的。即在进行  $k$  步区间预测时,用第  $i$

种预测方法的最近  $k$  期的拟合平均精度  $\frac{1}{k} \sum_{t=n-k+1}^n P_{it}$  和  $\frac{1}{k} \sum_{t=n-k+1}^n q_{it}$  来反映预测区间上  $n+k$  期的区间中心预测精度和半径预测精度,然后再根据预测精度进行组合预测。

## 参考文献(References):

- [1] HURST H E. Long Term Storage Capacity of Reservoirs[J]. Transactions of the American Society of Civil Engineers, 1951(116):770—799
- [2] PENG C K, BULDYREV S V, HAVLIN S, et al. Goldberger Mosaic Organization of DNA Nucleotides[J]. Phys Rev E, 1994(49):1691—1695
- [3] SANDRO T. Matplotlib for Python Developers [M]. Birmingham:Packt Publishing, 2009
- [4] MATTEO T, CRAMATICA R, ASTE T. Dynamical Generalized Hurst Exponent as a Tool to Monitor Unstable Periods in Financial Time Series [J]. Physica a-Statistical Mechanics and Its Applications, 2012, 391(11):3180—3189
- [5] LO A W. Long-term Memory in Stock Market Prices[J]. Econometrica, 1991, 59(6):1279—1313
- [6] GIRAITIS L, KOKOSZKA P, LEIPUS R. Rescaled Variance and Related Tests for Long Memory in Volatility and Levels[J]. Journal of Econometrics, 2003, 112(2):

- 265—294
- [7] GRANGER C W J. Long Memory Relationships and the Aggregation of Dynamic Models [J]. *Journal of Econometrics*, 1980, 14(2): 227—238
- [8] GRANGER C W J, JOYEUX R. An Introduction to Long Memory Time Series Models and Fractional Differencing [J]. *Journal of Time Series Analysis*, 1980(1): 15—39
- [9] BAILLIE. Fractional Integrated Generalized Autoregressive Conditional Heteroskedasticity [J]. *Journal of Econometrics*, 1996, 74(1): 3—30
- [10] 周礼刚, 陈华友. 两类区间数判断矩阵的一致性研究 [J]. *运筹与管理*, 2005(4): 47—51  
ZHOU L G, CHEN H Y. Research on Consistency of Two Interval Judgement Matrices [J]. *Operations Research and Management Science*, 2005(4): 47—51 (in Chinese)
- [11] MCLEOD A I, HIPELI K W. Preservation of the Rescales Adjusted Range: A Recessment of the Hurst Pjnenomenon [J]. *Water Resources Research*, 1978, 14, 491—508
- [12] 陈华友. 组合预测方法有效性理论及其应用 [M]. 北京: 科学出版社, 2008  
CHEN H Y. The Validity Theory of Combined Forecasting Method and Its Application [M]. Beijing: Science Press, 2008 (in Chinese)
- [13] 袁宏俊, 韦晨珺, 钟梅. 基于联系数贴近度的区间型组合预测及其有效性 [J]. *统计与信息论坛*, 2017(6): 31—37  
YUAN H J, WEI C J W, ZHONG M. Research on Interval Combination Forecasting Model and Its Effectiveness Based on Contact Number Degree [J]. *Statistics & Information Forum*, 2017(6): 31—37 (in Chinese)

## Long Memory Study of Interval Financial Time Series

DING Qin-xiang<sup>a</sup>, WANG Zhe<sup>a</sup>, WANG Yi-ning<sup>b</sup>, LI Ming-yuan<sup>a</sup>

(a. School of Economics; b. School of Mathematical Sciences, Anhui University, Hefei 230601, China)

**Abstract:** Studying the long memory of financial time series can help people characterize financial markets accurately. However, studies on the interval financial time series are fewer. Therefore, the long-memory features of interval-based financial time series and its interval time series prediction model based on existing point-valued financial time series long memory modeling are considered in this paper. Firstly, the interval number is expressed as the interval center and the interval radius. Then, the center series and radius series are tested respectively whether they have long memory, and the interval-combination forecast model is proposed to predict the series which has long memory. Finally, an example about Shanghai Composite Index and Shenzhen Component Index is given to show that the interval stock index of Shanghai Composite Index has long memory, and the interval combination forecasting model can improve the prediction accuracy of interval financial time series.

**Key words:** interval financial time series; long memory; Hurst Index

责任编辑: 李翠薇

引用本文/Cite this paper:

丁勤祥, 王哲, 王艺宁, 李铭源. 区间型金融时间序列的长记忆性研究 [J]. *重庆工商大学学报(自然科学版)*, 2020, 37(1): 104—111

DING Q X, WANG Z, WANG Y N, LI M Y. Long Memory Study of Interval Financial Time Series [J]. *Journal of Chongqing Technology and Business University (Natural Science Edition)*, 2020, 37(1): 104—111