

# 基于隐马尔可夫模型的收益率波动性研究 ——以沪深 300 股指期货为例

杨海霞

(北京物资学院 北京 101149)

**摘要:** 我国 2010 年推出沪深 300 股指期货, 股指期货市场还在发展过程中。本文以沪深 300 股指期货的收益率波动性为研究对象, 运用隐马尔可夫模型对收益率的波动性进行研究, 并对沪深 300 股指期货的指数进行预测。

**关键词:** 收益率; 波动性; 隐马尔可夫模型; TGARCH 模型

## 一、隐马尔可夫收益率波动模型介绍

隐马尔可夫模型的一般形式如下: 通过一个五元组  $(S, O, A, B, \pi)$  来表示隐马尔可夫模型, 其中 (1)  $S = \{s_1, s_2, \dots, s_N\}$  为隐含状态集合, 状态数为  $N$ 。 (2)  $O = \{o_1, o_2, \dots, o_N\}$  为观测值集合, 状态数为  $N$ 。 (3)  $A = \{a_{ij}\}$  为状态转移概率矩阵,  $a_{ij} = p\{q_{t+1} = s_j | q_t = s_i\}$ ,  $1 \leq i, j \leq N$ ,  $a_{ij}$  表示  $t$  时刻从状态  $s_i$  转移到  $s_j$  的概率。 (4)  $B = \{b_{jk}\}$  为输出概率分布,  $1 \leq j \leq N$ ,  $1 \leq k \leq M$  其中  $b_{jk}$  为状态  $s_j$  输出观测符号。 (5)  $\pi = \{\pi_1, \pi_2, \pi_3, \dots, \pi_n\}$  记为初始状态的概率分布,  $\pi_i = p\{q_1 = s_i\}$  表示初始时刻某一状态下的概率。

## 二、实证

本文选择 2015 年 1 月 5 日至 2019 年 3 月 22 日的沪深 300 股指期货数据, 样本总量为 1028, 第一部分为 2015 年 1 月 5 日至 2019 年 2 月 1 日共有 994 个数据, 使用这些数据进行建模, 并对之后 30 日的状态进行预测; 第二部分为 2019 年 2 月 11 日至 2019 年 3 月 22 日共 30 个数据, 使用这些数据进行验证模型的预测效果。

### (一) 数据基本分析

对沪深 300 股指期货对数收益率进行波动性分析得:

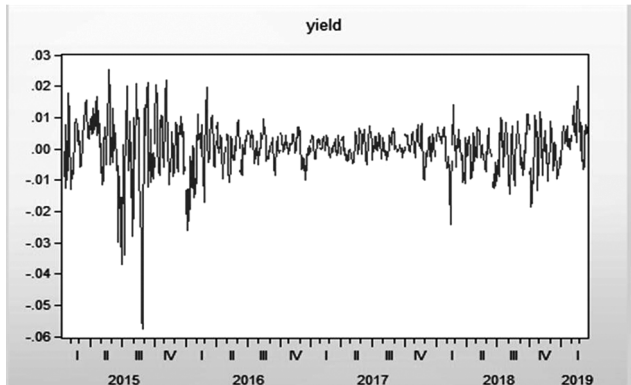


图 1 沪深 300 股指期货对数收益率的金融时间序列

通过图 1 可以发现沪深 300 股指期货收益率存在波动聚集现象, 2016 年到 2018 年波动较小, 2016 年之前波动较大, 2018 年开始波动在逐渐变大。

### (二) 基于改进的隐马尔可夫模型实证

利用 Python 对沪深 300 股指期货进行 HMM 模型实证得出隐含层的初始状态、转移矩阵和最终的稳态。

隐含层初始状态:  $[5.56291552e-01 \ 4.43296222e-01 \ 4.12225679e-04]$

隐含层转移矩阵:

$[0.45797525 \ 0.53454649 \ 0.00747825]$

$[0.45563766 \ 0.53853528 \ 0.00582706]$

$[0.06722476 \ 0.06121235 \ 0.87156289]$

稳态:  $[6.19665310e-05 \ 5.43889727e-05 \ 4.03463645e-04]$

利用 HMM 模型对后 30 天数据进行预测。首先对后 30 天真实数据绘制折线图, 如下:

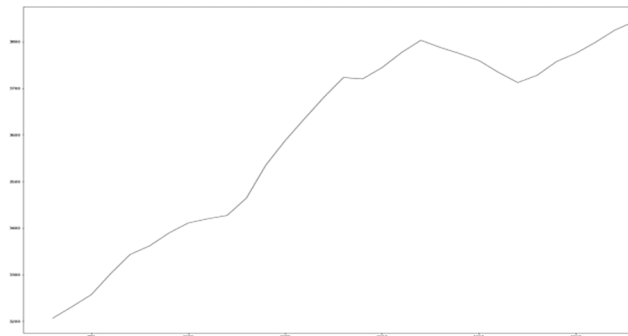


图 2 沪深 300 股指期货后 30 天真实数据折线图

图中可以看出后 30 天的大体趋势。然后利用 HMM 模型对后 30 天进行预测, 对三种不同隐含状态下后 30 天的预测图进行对比:

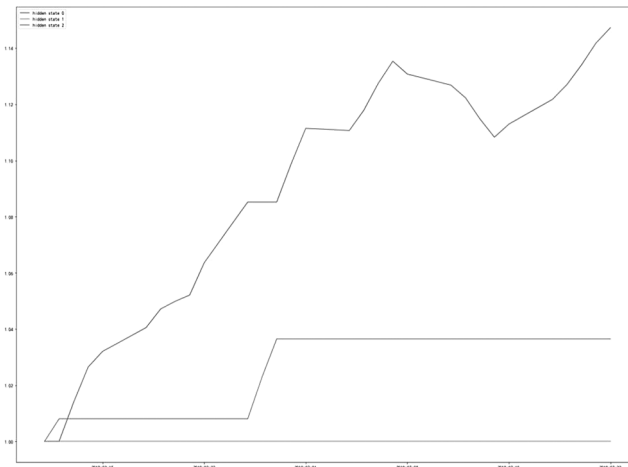


图 3 沪深 300 股指期货后 30 天预测数据折线图

如图分析发现, 状态 0 与实际数值的走势最为接近, 进而说明状态 0 对数据有更好的预测。由 python 得出预测值的对数似然值为: 101.0195888591508, AIC 为 -9.2306289。由上面两图的对比可以看出隐马尔可夫模型的预测精确度较高。