

基于上证指数高频数据的中国资本市场微观特性研究

唐振鹏 陈尾虹 冉梦

Microscopic characteristics of Chinese capital market based on the high frequency data of Shanghai composite index

Tang Zhen-Peng Chen Wei-Hong Ran Meng

引用信息 Citation: *Acta Physica Sinica*, 66, 120203 (2017) DOI: 10.7498/aps.66.120203

在线阅读 View online: <http://dx.doi.org/10.7498/aps.66.120203>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2017/V66/I12>

---

您可能感兴趣的其他文章

Articles you may be interested in

中国西南地区干旱 Copula 函数模型对样本量的敏感性分析

Sensitivity analysis of sample number on the drought descriptive model built by Copula function in south-west China

物理学报.2015, 64(10): 100203 <http://dx.doi.org/10.7498/aps.64.100203>

具有质量及频率涨落的欠阻尼线性谐振子的随机共振

Stochastic resonance of an underdamped linear harmonic oscillator with fluctuating mass and fluctuating frequency

物理学报.2015, 64(2): 020202 <http://dx.doi.org/10.7498/aps.64.020202>

多载波微放电中二次电子横向扩散的概率分析

Probabilistic analysis of the lateral diffusion of secondary electrons in multicarrier multipactor

物理学报.2014, 63(22): 220205 <http://dx.doi.org/10.7498/aps.63.220205>

无线传感器网络中无标度拓扑的动态容错性分析

Dynamic fault-tolerance analysis of scale-free topology in wireless sensor networks

物理学报.2014, 63(11): 110205 <http://dx.doi.org/10.7498/aps.63.110205>

超网络中标度律的涌现

Emergence of scaling in hypernetworks

物理学报.2014, 63(9): 090207 <http://dx.doi.org/10.7498/aps.63.090207>

# 基于上证指数高频数据的中国资本市场 微观特性研究\*

唐振鹏<sup>1)2)3)</sup> 陈尾虹<sup>1)†</sup> 冉梦<sup>1)</sup>

1) (福州大学经济与管理学院, 福州 350116)

2) (福建省金融科技创新重点实验室, 福州 350116)

3) (福建省企业发展研究中心, 福州 350116)

(2017年1月6日收到; 2017年4月6日收到修改稿)

以上证指数高频数据为研究对象, 基于上涨、平缓和下跌三个市场状态分析我国金融市场的微观特性. 通过分析上证指数在不同时间间隔下的概率分布、自相关性和多分形三个特性, 发现上证指数对数增量序列存在厚尾、列维非高斯分布特征, 且随着时间间隔的增大, 收益序列愈收敛于正态分布, 其中, 下降趋势收敛于正态分布的速度更快, 拟合于列维分布的效果更好. 最为突出的是, 在自相关函数分析中, 上证指数的收益率无长期记忆性, 而波动率则具有较强的记忆性. 同时, 波动率的自相关性存在明显的周期性特征, 即  $T = 240 \text{ min}$ , 且在下降趋势时其相关性最高. 在以时间增量刻画的多重分形结构中, 对于不同的时间序列、时间间隔, 由于受投资期限和流动性的影响, 三种股市状态的收益率波动存在着短期和长期性的差异. 上证指数的总体宏观行为与国际成熟股市较为一致, 但在微观特性上仍存在显著差异, 其所特有的周期性是投资者的惯性反冲所致, 而自相关性函数较之成熟股市衰减较慢, 则表明投资者的投资行为更多地受历史信息的影响.

**关键词:** 资本市场, 微观特性, 上证指数, 高频数据

**PACS:** 02.50.-r, 02.50.Cw, 02.50.Ng, 02.60.Ed

**DOI:** 10.7498/aps.66.120203

## 1 引言

1995年, Mantegna和Stanley<sup>[1]</sup>在《Nature》发表了关于经济指数标度行为的研究成果, 这一开创性的工作拉开了金融物理学研究的帷幕, 吸引了众多物理学家加入到相关研究中. 近年来, 经济学家同样将统计物理学的概念和方法引入金融、经济、交通物流和社会财富等系列问题的研究中<sup>[2-5]</sup>, 金融物理学作为一门新的交叉学科应运而生, 其中最为系统和深入的研究集中于股票市场. 这一方面是由于各个股票市场保存了大量的交易数据, 特别是高频和超高频数据的发展为研究提供了足够

的数据深度; 另一方面是不同于传统金融学的研究视角, 金融物理学旨在揭示股票市场的普适性、自组织结构等. 经济物理学家对西方发达国家主要股市的数据做了大量系统的分析和研究, 发现了一系列重要的共性结果, 即在股市错综复杂的涨跌趋势中, 存在一些普遍适用的统计规律<sup>[1,6]</sup>. 而在中国股票市场的研究中, Wei和Wang<sup>[7]</sup>将多分形分析运用到金融市场的波动率测度和预测领域, 并提出基于多分形的波动率测度MFV(multifractal volatility). 吴斌哲和马红孺<sup>[8]</sup>对上证指数的概率密度分布和自相关特性进行分析, 发现上证高频指数的负相关性效应和投机性特征. Gao等<sup>[9]</sup>将网络理论应用于中国股市, 研究揭示了中国股票市场

\* 国家自然科学基金(批准号: 71573042, 71171056)和福建省社科基金青年博士项目(批准号: FJ2016C200)资助的课题.

† 通信作者. E-mail: tingling69@163.com

的拓扑属性以及上市公司间的交互影响关系. 在国内外股票市场的对比研究中, Qiu等<sup>[10]</sup>的研究指出我国股市基本与国际成熟股市的统计行为趋同, 但在某些关联行为上表现出一些差异. Gao等<sup>[11]</sup>的研究则表明我国股票市场与发达金融市场均存在与多重分形标度相关的层级结构特征, 但层级结构各异.

但上述研究忽视了股市普适性背后隐匿的差异性及其原理, 同时未能区分上涨、平缓和下跌等不同股市状态并对其微观特性进行研究. 而值得注意的问题是: 当股票市场处于不同的市场状态下, 其概率分布、自相关性和多分形性等统计规律有何差异? 经济意义如何? 当股市分别处于上涨、平缓和下跌趋势时, 投资者的决策有何行为动机和依据? 识别不同股市状态的微观特性对投资决策有何指导意义? 本文以上证指数1 min高频数据为研究对象, 类比文献<sup>[8]</sup>的国际股市研究, 对中国金融市场微观特性进行分析, 并尝试解释其中的差异, 以期对投资决策提出有效的指导意见.

## 2 物理金融方法的理论基础

### 2.1 概率分布函数

#### 2.1.1 厚尾分布

早期的金融理论一直认为金融资产价格或者收益率满足正态分布, 而Mandelbrot<sup>[12]</sup>提出了平稳列维分布, 并创造性地应用平稳帕累托(stable Paretian)定律描述证券市场价格收益率的尾部分布, 研究发现在刻画棉花价格波动的概率分布上, 平稳帕累托分布比正态分布的拟合效果更好, 这一发现彻底颠覆了传统的布朗运动模型. 具有胖尾特征的金融资产收益的概率分布被越来越多的学者所发现, 学界一般以互补概率或对数分段来捕捉这一特性. 本文引入互补概率的类似方法, 即分阶排序法来刻画金融资产收益的胖尾分布, 该法具有易于实现、不损失信息量和噪声小的优点<sup>[13,14]</sup>.

考虑变量 $x$ 的 $n$ 个观察值, 真实概率密度为 $f(x)$ , 则其互补概率分布为 $C(x) = \int_x^\infty f(y)dy$ . 将这 $n$ 个观察值从大到小排序, 得到

$$x_1 \geq x_2 \geq \dots \geq x_R \dots \geq x_n, \quad (1)$$

其中 $R$ 为排名, 即阶. 于是,  $nC(x_R)$ 是不小于 $x_R$

的观察值的数量, 即

$$nC(x_R) = R. \quad (2)$$

如若变量 $x$ 具有幂律尾分布, 即 $f(x) \sim x^{-(1+\omega)}$ , 那么 $C(x) \sim x^{-\omega}$ , 可得到 $x_R$ 与 $R$ 之间的近似关系为

$$x_R \sim R^{-1/\omega}. \quad (3)$$

(3)式的精确表达式可从(2)式推导出, 即

$$x_R \sim \left( \frac{\omega n + 1}{\omega R + 1} \right)^{1/\omega}. \quad (4)$$

#### 2.1.2 列维分布

大量的统计研究表明, 股票指数对数增量的概率密度函数尾部胖于正态分布, 同时又瘦于负幂律分布, 因此, 满足截尾的列维(truncated Levy flight, TLF)分布<sup>[15-18]</sup>, 截尾从理论上保证了分布具有有限的二阶矩, 在靠近中心部分, TLF分布和列维分布并没有明显差别. 为了简化模型, 进一步用列维分布对上证指数的对数增量分布进行处理. 均值为零的列维分布的概率分布函数由下式给出

$$P_L(x) \equiv \frac{1}{\pi} \int_0^\infty e^{-\gamma|q|^\alpha} \cos(qx) dq, \quad (5)$$

其中,  $x$ 为随机变量;  $\alpha$ 为一个标志分布特性的参数, 且 $0 < \alpha \leq 2$ ;  $\gamma$ 为一个正的标度因子. (5)式表示的列维平稳分布只在特殊情况下存在解析式, 其中, 当 $\alpha = 1/2, \beta = 1$ 时, 对应的是列维-斯米尔洛夫分布; 当 $\alpha = 1, \beta = 0$ 时, 对应的是洛伦兹分布,  $P_L(x) = \frac{\gamma}{\pi} \frac{1}{x^2 + \gamma^2}$ ; 当 $\alpha = 2$ 时, 对应的是高斯分布,  $P_L(x) = \frac{1}{2\sqrt{\pi\gamma}} e^{-\frac{x^2}{4\gamma}}$ . 文中仅考虑具有零均值( $\mu = 0$ )的对称稳定分布( $\beta = 0$ ), 即具有指数 $\alpha$ 和尺度因子 $\gamma$ 的对称稳定分布. 则对于任意股指对数增量 $R_k(t) = \ln P(t+k) - \ln P(t)$ , 其中,  $k$ 为时间间隔,  $P(t)$ 表示某一资产在时刻 $t$ 的价格, 并且 $R_k(t)$ 满足对称的列维平稳分布

$$P(R_k) = \frac{1}{\pi} \int_0^\infty e^{-k\gamma|q|^\alpha} \cos(qR_k) dq. \quad (6)$$

由于0对数增量处于概率分布的峰值且数据集中, 同时考虑到计算的便利性, 文中用列维分布拟合数据时, 选择 $P(R=0)$ (即对数增量的概率密度为0)的数据来求参数. 如果分布具有自相似结构, 则不同的 $k$ 对应的 $\alpha$ 相同. 列维平稳分布的重要特征之一是自相似性, 则当 $R_k = 0$ 时, (6)式可以转化为

$$P(R_k = 0) = \frac{1}{\pi} \int_0^\infty e^{-k\gamma|q|^\alpha} dq = \frac{\Gamma(1/\alpha)}{\pi\alpha(\gamma k)^{1/\alpha}}. \quad (7)$$

对等式两边取对数, 得

$$\log P(R_k = 0) = \log \left[ \frac{\Gamma(1/\alpha)}{\pi\alpha\gamma^{1/\alpha}} \right] - \frac{1}{\alpha} \log k. \quad (8)$$

由(8)式拟合可得列维分布的指数分布参数 $\alpha$ .

### 2.2 自相关函数

研究金融资产对数价格变化的两两独立性, 通常是分析对数价格时间变化的自相关函数或讨论对数价格时间序列的频谱密度. 对于平稳随机过程, 研究这二类统计特性是等值的. 对数价格变化的自相关函数是一个快速衰减的函数, 刻画它的相关时间通常小于一个交易日. 因此, 只有通过分析高频数据才可以准确地估计序列的相关时间. 自相关系数 $C(\tau)$ 定义为

$$C_R(t) = \frac{\langle R_k(x)R_k(x+t) \rangle - \langle R_k(x) \rangle \langle R_k(x+t) \rangle}{[\sigma(R_k)]^2}, \quad (9)$$

其中, 在给定时间间隔 $k$ 下, 当 $t = 0$ 时,  $C_R(t) = 1$ . 如果同一 $k$ 水平下间隔为 $t$ 的对数增量之间没有关联, 即 $\langle R_k(x)R_k(x+t) \rangle = \langle R_k(x) \rangle \langle R_k(x+t) \rangle$ , 则 $C_R(t) = 0$ .

### 2.3 多分形

多分形是一种随着时间标度 $t$ 的增加, 时间序列的动力学特征随之变化的性质, 即收益率的绝对矩随着取样间隔的不同呈现出幂律变化<sup>[19,20]</sup>. Mandelbrot<sup>[21]</sup>对多分形过程的定义如下:  $X(t) = \ln[P(t)]$ , 令 $X(t)$ 为具有平稳增量 $\delta_k X(t)$ 的随机过程, 即 $\delta_k X(t) = X(t) - X(t-k)$ . 假设 $E[|\delta_k X(t)|^q]$ 对于所有的 $q \in Q$ 存在, 其中 $Q \rightarrow \mathbb{R}$ 是包含间隔 $[0, 1]$ 的合适集, 令 $c_q = E[|\delta_1 X(t)|^q]$ . 进一步地,  $\zeta: Q \rightarrow \mathbb{R}$ 定义一个可微的函数,  $T$ 是正实数值.

如果对于所有的 $\tau \in [0, T]$ 以及所有的矩 $q \in Q$ , 满足

$$M(q, k) = E[|\delta_k X(t)|^q] = c_q k^{\zeta(q)}. \quad (10)$$

则标度不变性过程增量的绝对矩作为取样间隔 $k$ 的函数呈现幂律分布, 那么我们可以称过程 $X(t)$

是标度不变性. 标度不变性过程分为两种情况: 当 $\zeta(q)$ 是线性时比如 $\zeta(q) = qH$ ( $H$ 为Hurst指数), 该过程是单分形过程, 这种情况下尺度可以扩展到一个无限制的尺度范围( $T = +\infty$ ). 单分形过程就是所谓的自相似过程比如分形布朗运动或 $\alpha$ -stable运动; 当 $\zeta(q)$ 为非线性时, 该过程是多分形过程. 非线性指数 $\zeta(q)$ 是严格的凹函数. 对于一个凹函数, 其多分形性质不会对于任意大的标度都成立. 通常我们定义 $T$ 为整体尺度(integral scale), 超过时间 $T$ 该过程即停止呈现多分形特征. 上述等式在 $k \leq T$ 时成立.

### 3 数据和时间的选取

在处理股市数据的时间序列对价格的随机动态进行分析与建模的过程中, 时间尺度和价格尺度的选择至关重要. 关于时间尺度主要有3种不同的选择: 1) 物理时间, 该选择无法对休市期间的信息与价格随机动态特性建模; 2) 交易时间, 即在开市期间流逝的时间, 但这一选择无法刻画休市期间价格动态特性信息的时间效应; 3) 交易次数, 即以每笔交易为时间单位, 这种选择下金融市场上观测到的一类随机性就消失了, 也即每笔交易间的时间间隔. 目前, 学者大量使用的是交易时间<sup>[6,8]</sup>, 为了研究结果的可比性, 本文同样选择交易时间, 即选用上证指数的1 min数据. 通过周数据所得的涨跌趋势图(如图1所示)将市场分为上涨、下跌和平缓振荡三个状态(如表1所列), 本文将相对于前期低点涨幅超过30%的走势定义为一个上涨趋势, 将相对于前期高点跌幅超过30%的走势定义为一个下跌趋势<sup>[22]</sup>, 并将变化幅度位于5%—10%的趋势定义为一个平缓趋势. 同时, 为剔除样本量对研究结果的影响, 将数据区间控制在同一时间长度, 遇非交易日则日期顺延. 其中, 上涨区间为2015年2月12日至2015年6月12日, 平缓区间为2014年3月10日至2014年7月10日, 下跌区间为2015年6月12日至2015年10月12日, 采用的数据均来自于同花顺数据库(研究过程中曾试验过多组不同的样本数据, 研究结论不变. 考虑到篇幅问题以及时间的同步性, 故选取2012年—2015年的样本数据).

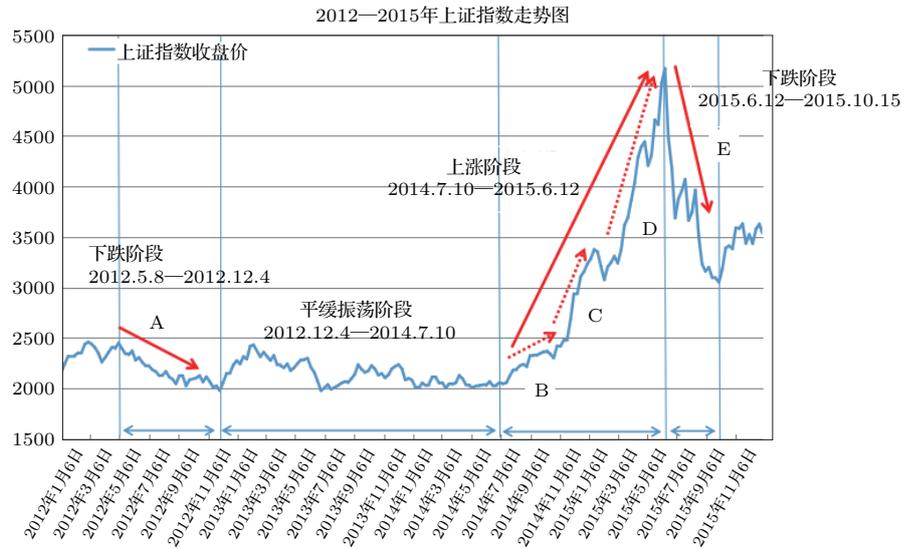


图1 上证指数走势

Fig. 1. Trend of Shanghai Stock Exchange composite index.

表1 上证指数市场状态区间分布

Table 1. Interval distribution of Shanghai Stock Exchange composite index in market state.

	下跌阶段		平缓振荡阶段	上涨阶段	
	阶段 A	阶段 E		阶段 B	阶段 D
时间	2012.5.8—2012.12.4	2015.6.12—2015.10.15	2012.12.4—2014.7.10	2014.7.10—2014.11.21	2015.2.9—2015.6.12
指数分布区间	2448.88—1975.14	5178.19—3052.78	1974.14—2038.34	2038.34—2486.79	3095.12—5178.19

股票数据的价格尺度选择不惟一，根据研究目的的不同，可以区分为股指  $P(t)$ ，股指涨跌  $P(t+k) - P(t)$  等。本文考虑股指涨跌的统计特性，因此，选用股指的相对变化

$$R(t) = (P(t+k) - P(t))/P(t) = Z(t)/P(t), \quad (11)$$

其中， $k$  为时间间隔。作为统计量，股票价格是基于货币，指数的变化中也包含货币价值的变化，当  $k$  较小时，假定货币本身的波动在一段时间内变化不大，那么选择  $R(t)$  作为统计量也就自然修正了货币波动的影响。但当  $k$  较大时， $t+k$  和  $t$  则不对称，从而导致对货币价值变化修正的偏离。注意到当  $k \rightarrow 0$  时， $Z(t)$  实际上代表股指对数的变化。因此，更为合理的选择是股指的对数增量，即

$$R_k(t) = \ln P(t+k) - \ln P(t). \quad (12)$$

将  $Y$  定义为一个时间序列，则

$$R_Y = \frac{dY}{dt}, \quad (13)$$

$$R_Y^k(t_i) = Y(t_i+k) - Y(t_i), \quad (14)$$

其中， $Y = \ln(P)$ 。作为统计变量，在下面的统计分析中，最小时间间隔  $k$  取为 1 min，这一时间间隔既可较好地保持数据的连续性，亦可较好地处理指数变化时间序列间的短程相关性对数据统计特性分析的影响<sup>[23]</sup>。则

$$|R_Y^k| = |Y(t_i+k) - Y(t_i)|. \quad (15)$$

由 (15) 式可得标准化绝对收益：

$$|\hat{R}_Y^k| = \frac{|R_Y^k|}{\left(\sum_{t_i}^{N_T-k} |R_Y^k|\right) / (N_T - k)}. \quad (16)$$

## 4 实证分析

### 4.1 $|\hat{R}_Y^k|$ 的概率分布函数

#### 4.1.1 厚尾分布

为使不同时间间隔的序列具有可比性，以标准化绝对收益序列为研究样本，分别取  $k = 1 \text{ min}$  和

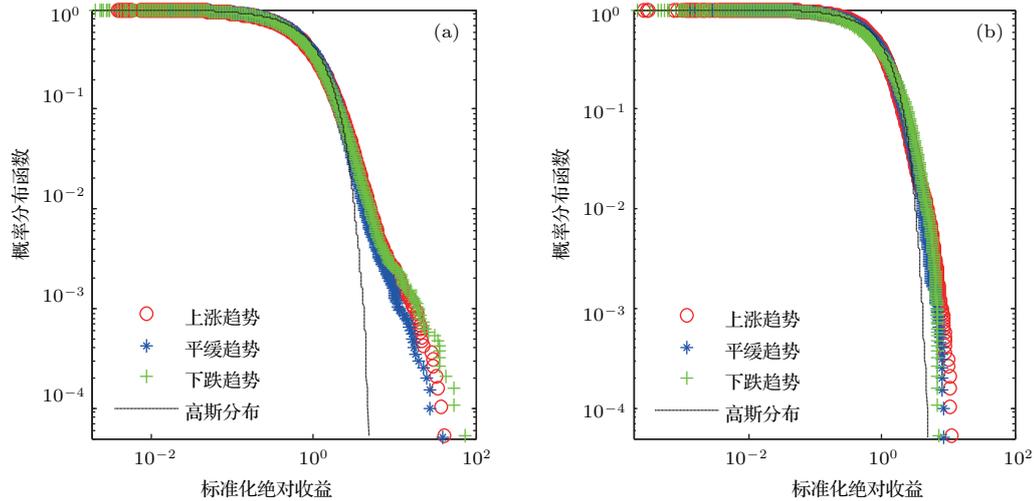


图2 互补概率分布图 (a)  $k = 1$  min; (b)  $k = 100$  min

Fig. 2. Distribution of complementary probability with (a)  $k = 1$  min and (b)  $k = 100$  min.

$k = 100$  min. 基于高频数据对上证指数的上涨、平缓、下跌状态进行微观特性分析, 互补概率分布如图2所示.

从图2可得, 上证指数标准化绝对收益序列的分布函数呈现出厚尾特征, 同时, Kolmogorov-Smirnov (KS) 估计结果也表明三个阶段的厚尾分布存在显著性差异. 但随着时间间隔  $k$  由 1 min 增至 100 min, 收益序列的分布聚敛于正态分布. 其主要原因是股票每分钟的收盘价受到噪声 (短时间大量交易造成不连续) 的影响, 而随着  $k$  的增加, 可以滤掉噪声, 因此,  $k = 100$  min 时的收益序列比  $k = 1$  min 时更加收敛于正态分布. 对比2组图形可以看出, 当  $k = 1$  min 时, 股市下跌趋势的尾部分布比平缓、上涨趋势更加扩散, 而这一分布特征也符合股市信息冲击的非对称性, 其主要是受政府对股市干预的宏观政策、投资者掌握信息的不对称性、股市缺乏做空机制和散户投资者普遍具有从众行为等因素的影响. 投资者的行为决策深受投资者情绪的传染, 这种传染效应在下跌期间尤甚. 如 2015 年股灾期间股票指数不断下跌, 投资者恐慌情绪迅速传染, 恐慌性抛售股票攀升, 以致 2016 年初启动熔断机制后, 股市多次触碰熔断而中断. 而当  $k = 100$  min 时, 下降趋势的尾部分布较之平缓、上涨趋势, 则更拟合于正态分布, 这主要是由于人本质上具有社会性, 个人的投资决策易受社会因素的影响, 因此, 具有“恐慌”心理的投资者会加剧市场“恐慌”情绪的蔓延, 我国股市在下降趋势时存在较

高的投资同质性, 而较长时间间隔的收益序列则熨平了短期时间间隔内噪声因素所带来的发散特性.

而较之文献[6]的研究, 成熟股市的收益率在  $k = 1$  min 时尾部呈现截尾现象, 比我国股市的尾部分布更厚, 我们认为这主要是由于我国股市有涨跌停限制, 股票在一个交易日中的成交价格不能高于或低于限定的涨跌幅度, 当超过该范围时报价将视为无效, 所以我国股市的振幅没有国外股市那么剧烈.

#### 4.1.2 列维分布

采用 Mantegna-Stanley 方法研究分布的最中心峰值即零收益的概率  $P(R = 0)$  作为  $k$  的函数, 这一方法可使我们只研究每一概率分布中受噪声影响最少的点. 首先分析  $R_k(t)$  的概率分布函数. 分别取  $k = 1, 3, 10, 32, 100, 316, 1000$  min, 求出不同市场状态下对应的  $R_k(t)$  并统计其概率分布, 即得到  $R_k(t)$  的概率分布函数. 其次根据(8)式拟合得出列维分布的指数分布参数  $\alpha$ , 如图3所示.

图3是上证指数0对数增量的概率密度  $P_k(0)$  与  $k$  的双对数图, 由图可见数据点基本在一条直线上, 表明列维分布假定是合理的. 由最小二乘法进行数值拟合, 可得上涨、平缓和下跌状态下拟合直线的斜率  $(-1/\alpha)$  分别为  $-0.6899, -0.5681, -0.5667$ , 这一结果与标准普尔(S&P)指数的  $-0.71$ 、欧元美元汇率的  $-0.71$  均存在差异, 说明我国股市与成熟股市尚存在较大差距. 同时, 研究结果也与文献[6]所得的  $0.65$  不同, 这主要是由

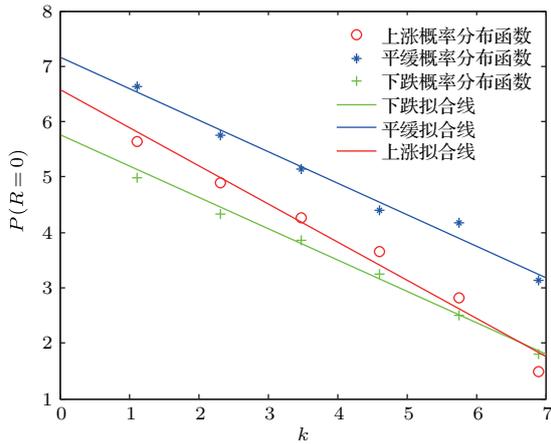


图3 在上涨、平缓和下跌趋势时  $P(R = 0)$  与  $k$  的双对数图  
 Fig. 3. Situation in rising trend, steady trend and falling trend log-log plot of the probability  $P(R = 0)$  versus  $k$ .

于本文考虑了股市的不同市场状态. 列维分布的指数分布参数  $\alpha$  分别为 1.4494, 1.7604, 1.7646, 说明截断列维飞行能够描述尖峰胖尾特征, 且具有缓慢渐近收敛到高斯域的特性. 可知我国股市在下跌趋势时 ( $\alpha = 1.7646$ ) 更趋向于高斯分布 ( $\alpha = 2$ ). 而拟合优度  $R^2$  分别为 0.9820, 0.9866, 0.9929, 则进一步表明股市在下跌时列维分布的拟合效果更好, 对于投资决策更具预测价值.

#### 4.2 收益率 $R_k(t)$ 与波动率 $|R_Y^k|$ 的自相关性分析

以  $R_k(t)$  为上证指数的收益率,  $|R_Y^k|$  为上证指数的波动率, 则上证指数的收益率和波动率自相关性分析如图 4—图 6 所示.

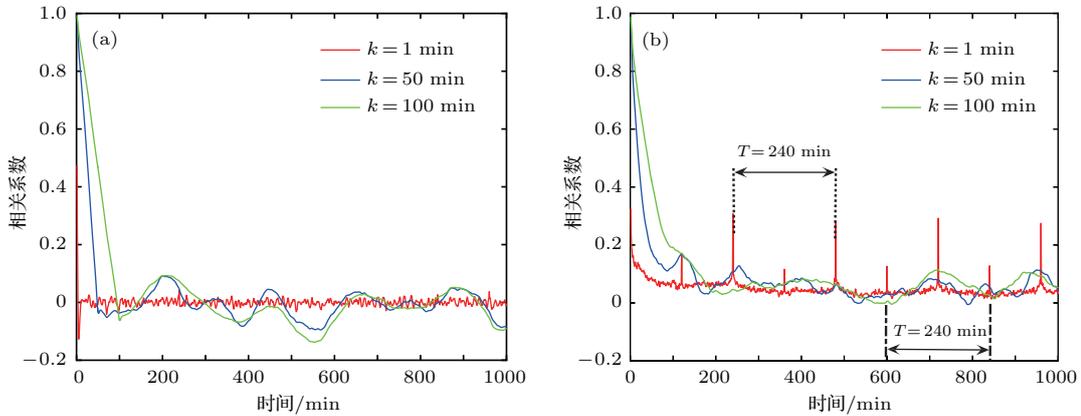


图4 上涨趋势情况 (a) 收益率自相关函数; (b) 波动率自相关函数

Fig. 4. Situation in rising trend: (a) Autocorrelation functions of the return; (b) autocorrelation functions of volatility.

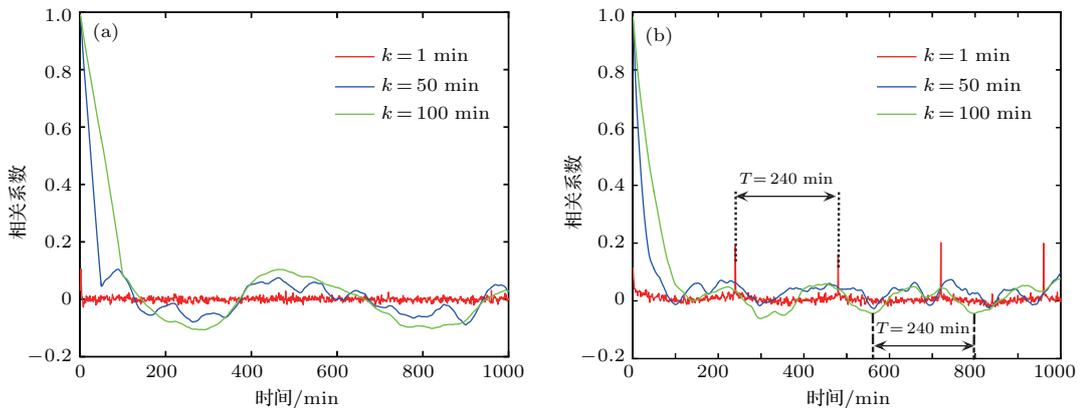


图5 平缓趋势情况 (a) 收益率自相关函数; (b) 波动率自相关函数

Fig. 5. Situation in steady trend: (a) Autocorrelation functions of the return; (b) autocorrelation functions of volatility.

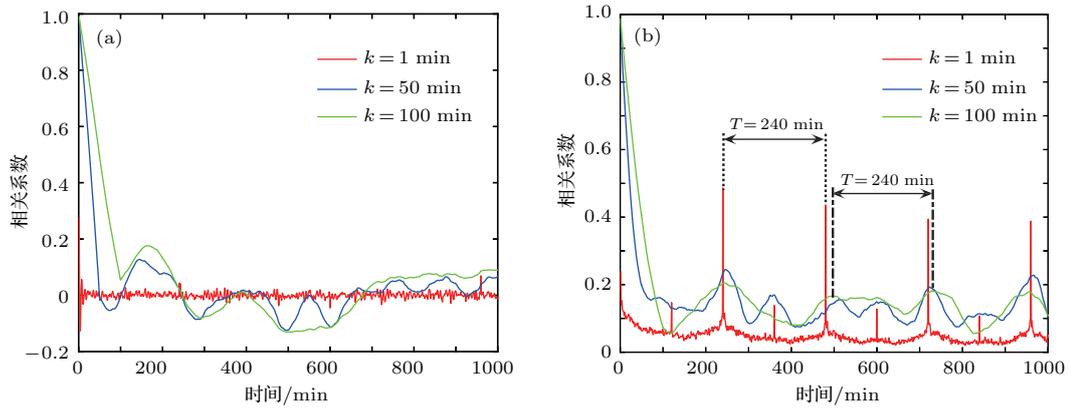


图6 下跌趋势情况 (a) 收益率自相关函数; (b) 波动率自相关函数

Fig. 6. Situation in falling trend: (a) Autocorrelation functions of the return; (b) autocorrelation functions of volatility.

从图4—图6可得, 上涨、平缓和下跌趋势波动率的自相关性均高于收益率的自相关性. 当  $k = 1$  时, 收益率的自相关系数趋于0, 这与传统金融理论中的价格变动的独立性假设一致. 在非线性的时域关联中, 通常以收益率的绝对值, 即波动率来考虑自相关性, 并以它度量收益率涨落幅度的时间关联. 图4(b)、图5(b)、图6(b)也证实了在较长的时间范围内,  $k = 1$  min的自相关系数都不为0. 在  $k = 1, 50, 100$  min时, 收益率自相关性曲线的振幅均大于波动率自相关性曲线,  $k = 1$  min的曲线呈微弱式下降,  $k = 50$  min与  $k = 100$  min的曲线正负走势则一致.

从图4(b)上涨趋势波动率自相关函数图可得, 在自相关函数的正负值特性中, 当  $k = 1$  min时, 自相关函数均为正值; 当  $k = 50$  min时, 自相关函数在  $t = 530—537$  min,  $t = 786—796$  min呈负相关; 当  $k = 100$  min时, 自相关函数在  $t = 571—595$  min,  $t = 617$  min呈负相关. 在周期性特性中, 当  $k = 1$  min时, 在滞后期  $t = 120$  min, 相关性系数趋于0.17, 并以  $T = 240$  min的周期出现; 在滞后期  $t = 240$  min, 相关性系数趋于0.31, 并以  $T = 240$  min的周期出现, 即波动率曲线以  $T = 120$  min的周期呈现出低相关性和高相关性的交替变换. 而这一周期性规律在  $k = 100$  min时依然不变.

从图5(b)可得, 在平缓趋势时,  $k = 1, 50, 100$  min的波动率自相关系数均出现负值. 在周期性特性中, 当  $k = 1$  min时, 在滞后期  $t = 240$  min, 相关性曲线向上跳跃, 以趋于0.2的值并以  $T =$

240 min的周期规律性地出现, 且在  $k = 100$  min时这一周期性规律依然存在. 在滞后期  $t = 240, 480, 720, 960$  min时, 三条曲线呈现正相关.

从图6(b)可得, 在下跌趋势时,  $k = 1, 50, 100$  min的波动率自相关系数均为正值. 在周期性特性中, 当  $k = 1$  min时, 在滞后期  $t = 120$  min, 相关性系数趋于0.15, 并以  $T = 240$  min的周期微下落式出现; 在滞后期  $t = 240$  min, 相关性系数趋于0.5, 并以  $T = 240$  min的周期微下落式出现. 即波动率曲线以  $T = 120$  min的周期呈现出低相关性和高相关性的交替变换, 且这一周期性规律在  $k = 100$  min时依然存在. 在不同  $k$  值的相关性特性中, 当滞后期  $t = 240, 480, 720, 960$  min时,  $k = 1, 50, 100$  min的波动率相关性系数呈正相关, 均出现一个大波峰. 在滞后期  $t = 120, 360, 600, 840$  min时,  $k = 1$  与  $k = 50$  min的相关性函数呈正相关, 均为一个小波峰, 而对应  $k = 100$  min的波动率自相关曲线则呈现出一个小波谷.

上涨、平缓、下跌趋势的波动率自相关性均高于收益率的自相关性, 这与经济学家的主流趋向一致, 即股票价格收益率无长期记忆性, 而波动率存在较强的记忆性. 而从上述三种趋势的自相关周期性分析图, 我们发现上证指数在  $k = 1$  min与  $k = 100$  min时波动率自相关曲线存在明显的周期性, 即我国股市每日开盘价存在较高的相关性, 其主要原因可能是由于存在日内模式, 这与Zhou等<sup>[24]</sup>的研究结论一致. 但本文尝试改变样本的起始值, 分别以时间9:27, 10:34, 11:39, 13:51, 14:47, 15:26的收盘价作为样本数据的起点, 结果发现这

一周期性规律仍然不变. 这说明了上证指数一分钟的股价波动深受前期同一时刻股价波动的影响, 而不仅仅是日内模式的作用. 因此, 投资者的研究视角不应仅局限于股市收盘价、开盘价, 还应挖掘高频数据的内在属性, 利用自相关性的周期性规律进行投资决策. 同时, 研究也发现: 下降趋势的收益率、波动率自相关曲线高于平缓 and 上涨趋势的曲线, 这也进一步验证我国股市信息冲击的非对称特性, 即对利坏信息冲击的反应大于利好信息的特征一直持续存在.

### 4.3 多分形 $\langle |R^k|^q \rangle / \langle |R^k| \rangle^q$

2008年, Calvet 和 Fisher<sup>[25]</sup>对多重分形的标度不变性的形式进行改进, 多重分形可由时间间隔  $k$  的收益比率  $\langle |R^k|^q \rangle / \langle |R^k| \rangle^q$  来表示, 比率若为常数则表示单分形, 反之则为多分形. 在下面的统计分析中, 每组的样本数据为 500 个, 时间间隔  $k$  的取值为 5 min,  $q$  值分别取 1.5, 2.0, 2.5, 3.0, 并且在  $q = 3.0$  时, 曲线的变化幅度最为明显, 上涨、平缓 and 下跌三种趋势下的收益比率 ( $\langle |R^k|^q \rangle / \langle |R^k| \rangle^q$ ) 曲线的线性回归斜率分别为  $-0.3429$ ,  $-0.1564$  以及  $-0.1515$ , 即三者从小到依次为大依次为上涨趋势线性斜

率、平缓趋势线性斜率、下跌趋势线性斜率. 具体如图 7 所示.

对不同幅度的波动性进行幂次方, 相当于对波动的幅度进行放大和缩小, 即大波动更大, 小波动更小. 不同  $q$  值对应的尺度函数刻画不同时点的分形特征, 随着时间标度的变化, 可以观察到不同时间标度下的价格波动信息. 从图 7 可以看出, 不论收益率数据趋于上涨、平缓还是下跌, 图形均为非线性, 即股票市场具有多重分形结构. 其中, 在不同时点上收益的波动幅度虽然各异, 但不同市场状态均呈现出趋势相似的衰减式长尾图形, 即股价的波动具有持久性或长期依赖关系. 由上述分析可知, 上证指数在不同的波动趋势下都具有较强的长记忆性特征. 此外, 从图中可以看出, 当  $q$  越大时, 价格增量较大的那些部分的标度行为变化越明显, 这说明价格在不同的变化范围内, 有不同的标度关系, 即股市中存在多标度行为. 而一般地, 对正时间标度方向来说, 行情较好时市场的交易量也相对较高, 高交易量意味着市场具有高流动性, 而高流动性能引发市场的大的价格波动. 因此, 上涨趋势较平缓 and 下跌趋势最先到达谷底.

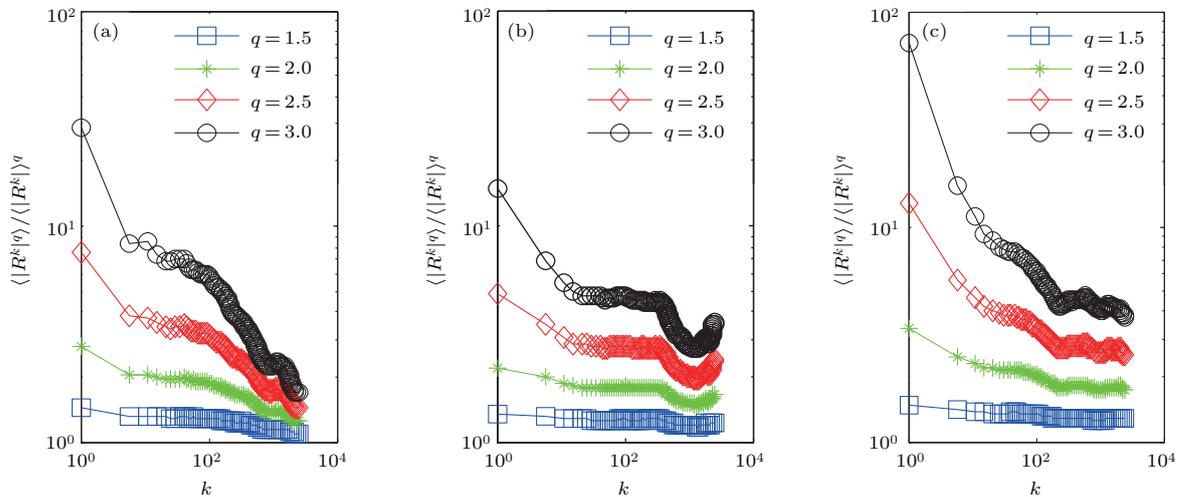


图 7  $k = 5 \text{ min}$  时不同趋势下的收益比率 ( $\langle |R^k|^q \rangle / \langle |R^k| \rangle^q$ ) 曲线 (a) 上涨趋势; (b) 平缓趋势; (c) 下跌趋势  
 Fig. 7. Curves of ratio  $\langle |R^k|^q \rangle / \langle |R^k| \rangle^q$  versus number  $k = 5 \text{ min}$  in different trends: (a) Rising trend; (b) steady trend; (c) falling trend.

上文是基于不同的  $q$  值, 以时间增量矩的尺度特性刻画股市的多重分形结构. 下文将选取  $q = 3$  的曲线测度股市的多重分形特征, 并分别分析处于上涨、平缓、下跌市场状态下收益率的长期、短期

波动趋势. 其中, 样本时间序列为 3790 个,  $k$  的取值为 5 min. 图 8 中, 红色曲线为上涨趋势的收益比率, 蓝色曲线为平缓趋势的收益比率, 绿色曲线为下跌趋势的收益比率.

当  $q = 3$  时, 按不同尺度的时间间隔对同一价格时间序列求取收益序列, 并对所得收益序列进行适当的尺度变换, 图中各市场状态下的股市都具有向下倾斜的结构特点, 反映了序列整体的稳定性, 但上涨趋势和下跌趋势的收益率序列衰减速度更快, 在短期内具有更明显的变化趋势, 说明收益率的变化趋势在一定程度上影响收益率序列的分布. 总体而言, 不同时间标度(分钟, 日, 周, 月, 年等)下股市收益序列在统计意义上具有相似性. 文中分别对不同长度的时间数据和时间间隔取值进行研究, 结果表明  $k$  越大, 所刻画的图形越粗糙, 但图形的收敛速度也越快, 即可缩短时间尺度并近似估计图形的变化趋势. 由分形结构的标度不变性可得, 不同的时间长度和时间间隔变化并不会改变股市的分形结构.

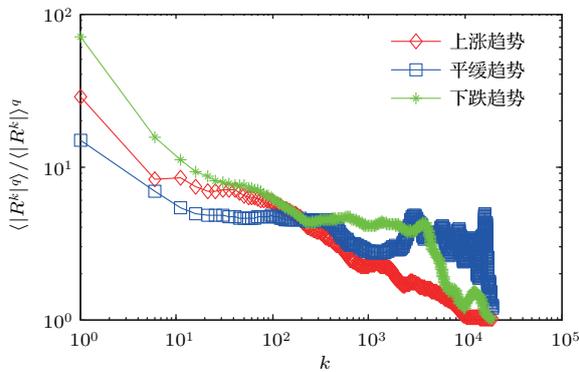


图8  $k = 5 \text{ min}$ ,  $q = 3$  的收益比率  $\langle |R^k|^q \rangle / \langle |R^k| \rangle^q$  曲线

Fig. 8. Curves of ratio  $\langle |R^k|^q \rangle / \langle |R^k| \rangle^q$  versus number  $k = 5 \text{ min}$  with  $q = 3$ .

股市分形结构的存在是由于市场中存在大量的具有不同投资期限的投资者和良好的流动性. 由图8可得, 上涨趋势收益率的变化会引起较大的价格波动. 这主要是因为投资者在股市处于上涨趋势时对未来预期乐观, 产生了“乐队车效应”, 从而出现资产流动性过剩, 助推资产价格泡沫的形成, 即高流动性易引发市场价格的大波动. 反之, 在短期内, 当股市处于下跌趋势时, 市场出现恐慌情绪, 由于股权质押、定增或购并中杠杆的广泛使用, 为求自保, 有影响力的股东和投资者以各种借口对股票实行停牌处理, 这显著挤压了场内的流动性; 受强制平仓和主动降杠杆的要求, 仍然在交易的股票大量跌停, 场内流动性基本枯竭; 面对潜在的赎回压力和现金要求, 投资者被迫抛售在境外市场、商品

市场和债券市场的头寸, 流动性枯竭迅速冲击到其他领域和地区; 同时, 流动性枯竭加剧了市场的恐慌情绪, 导致整个市场的流动性进一步下降. 投资者对未来预期悲观, 容易引发金融资产的大量恐慌抛售, 致使市场流动性严重不足, 最终出现金融市场的资产流动性好像瞬间被掏空的现象, 即“流动性黑洞”. 因此, 此时投资者在股市上谨慎寻求投资机会, 但由于市场流动性较低, 不活跃的市场往往会导致小的价格波动, 即波动趋势衰减速度较为平缓.

而当股市长期处于低位价格时, 对于长期投资者来说, 可以低价进入股市, 寻求时机以在未来更长的时间内获得更高收益, 因此, 当市场从流动性不足恢复至原先的流动性水平, 价格波动幅度比上涨趋势的波动更加显著. 同时, 受卖空限制的影响, 繁荣情绪更容易在投资者之间传递并反映到股价上, 即出现“集中式极化”. 而卖空限制会导致投资者的悲观预期无法及时反映到市场上, 即减慢资产价格对悲观信息的调整速度, 致使恐慌信息逐渐扩散且传递速度较慢, 即出现“分散式极化”. 对于平缓趋势的收益率序列, 股市的长期记忆性特征最弱, 且随着时间尺度的变化, 价格变化出现无规律的振荡式波动. 这一现象也与现有的研究相符, 当股市处于平缓时, 股票价格的趋势最不易捕捉.

## 5 结 论

本文分析了上证指数不同市场状态下的对数增量在不同时间间隔下的概率分布密度、自相关性和多分形特性. 首先, 同成熟股市一样, 当时间间隔由 1 min 变为 100 min 时, 标准化收益率序列的分布向高斯趋近, 且在市场为下跌状态时, 尾部更拟合高斯分布. 在时间间隔  $k = 1-1000 \text{ min}$  的范围内, 三种趋势对数增量序列的分布均能较好地用列维分布描述, 上涨、平缓、下跌趋势的分布指数、拟合优度不同, 而其中在股市下跌时列维分布的拟合效果最好, 对于投资决策更具预测价值. 其次, 在自相关性分析中, 上证指数的收益率无长期记忆性, 而波动率则存在较强的记忆性, 同时, 其波动率的自相关性存在明显的周期性, 且在下降趋势时相关性最高. 这是上证指数所特有的周期性, 这一效应表明投资者对于市场变化的过度反应导致投资

的惯性反冲. 而上证指数的自相关性函数较之成熟股市衰减更慢, 则表明投资者的投资热情更多地受到历史信息的影响. 因此, 投资者的研究视角不应只限于历史信息, 而应触及实时的高频数据, 并挖掘其内在属性, 利用自相关性的周期性等规律进行投资决策. 最后, 上证指数在不同的波动趋势下都具有很强的长记忆性特征, 其主要是由于长期、短期投资决策所引致的流动性水平的变化. 上证指数的总体宏观行为与国际成熟股市较为一致, 但在微观特性存在的差异说明我国股市受宏观政策、投资者情绪和流动性水平等因素的影响仍较强烈. 而股市价格的变化除了受外生性因素的影响外, 近年的研究也表明由于内源性反馈流程<sup>[26,27]</sup>, 股市内生性的影响水平已显著增加, 即我国股市价格的变化更多地受来自市场订单、股市结构等微观因素的影响.

#### 参考文献

- [1] Mantegna R N, Stanley H E 1995 *Nature* **376** 46
- [2] Chatfield C 2016 *The Analysis of Time Series: An Introduction* (Boca Raton: CRC Press) pp92–135
- [3] Chakraborti A, Toke I M, Patriarca M, Abergel F 2011 *Quant. Financ.* **11** 991
- [4] Huang J P 2013 *Economy Physics* (Beijing: Higher Education Press) pp1–7 (in Chinese) [黄吉平 2013 经济物理学 (北京: 高等教育出版社) 第1—7页]
- [5] Zhou R W, Li J C, Dong Z W, Li Y X, Qian Z W 2017 *Acta Phys. Sin.* **66** 040501 (in Chinese) [周若微, 李江城, 董志伟, 李云仙, 钱振伟 2017 物理学报 **66** 040501]
- [6] Parisi D R, Sornette D, Helbing D 2013 *Phys. Rev. E* **87** 012804
- [7] Wei Y, Wang P 2008 *Physica A* **387** 1585
- [8] Wu P Z, Ma H R 2008 *J. Shanghai Jiaotong Univ.* **42** 147 (in Chinese) [吴斌哲, 马红孺 2008 上海交通大学学报 **42** 147]
- [9] Gao Y C, Zeng Y, Cai S M 2015 *J. Stats. Mech.* **2015** 03017
- [10] Qiu T, Zheng B, Ren F, Trimper S 2007 *Physica A* **378** 387
- [11] Gao Y C, Cai S M, Wang B H 2012 *J. Stats. Mech.* **2012** 12016
- [12] Mandelbrot B B 1963 *J. Polit. Econ.* **71** 421
- [13] Sornette D, Knopoff L, Kagan Y, Vanneste C 1996 *J. Geophys. Res.* **101** 13883
- [14] Zhou W X 2007 *Introduction of Econophysics* (Shanghai: Shanghai University of Finance and Economics Press) pp17–33 (in Chinese) [周炜星 2007 金融物理学导论 (上海: 上海财经大学出版社) 第17—33页]
- [15] Mantegna R N, Stanley H E 1994 *Phys. Rev. Lett.* **73** 2946
- [16] Wang B H, Hui P M 2001 *Eur. Phys. J. B* **20** 573
- [17] Issler M, Höller J, Imamoğlu A 2016 *Phys. Rev. B* **93** 081414
- [18] Ncheuguim E K, Appiah-Kubi S, Ofori-Dankwa J 2014 *Res. Meth. Strateg. Manag.* **10** 215
- [19] Liu S D, Fu Z T, Liu S K 2014 *Chin. J. Geophys.* **57** 2751 (in Chinese) [刘式达, 付遵涛, 刘式适 2014 地球物理学报 **57** 2751]
- [20] Ma F, Wei Y, Huang D, Chen Y 2014 *Physica A* **405** 171
- [21] Mandelbrot B 1974 *J. Appl. Probab.* **11** 437
- [22] Zhang T W, Lu W B, Li S 2013 *Economist* **9** 97 (in Chinese) [张腾文, 鲁万波, 李隋 2013 经济学家 **9** 97]
- [23] Wu M C, Huang M C, Yu H C, Chiang T C 2006 *Phys. Rev. E* **73** 019908
- [24] Zhou W C, Xu H C, Cai Z Y, Wei J R, Zhu X Y, Wang W, Zhao L, Huang J P 2009 *Physica A* **388** 891
- [25] Calvet L E, Fisher A 2008 *Multifractal Volatility: Theory, Forecasting, and Pricing* (Massachusetts: Academic Press) pp23–31
- [26] Toth B, Eisler Z, Lillo F, Bouchaud J P, Kockelkoren J, Farmer J D 2012 *Quant. Financ.* **12** 1015
- [27] Flimonov V, Sornette D 2012 *Phys. Rev. E* **85** 056108

# Microscopic characteristics of Chinese capital market based on the high frequency data of Shanghai composite index\*

Tang Zhen-Peng<sup>1)2)3)</sup> Chen Wei-Hong<sup>1)†</sup> Ran Meng<sup>1)</sup>

1) (School of Economics and Management, Fuzhou University, Fuzhou 350116, China)

2) (Fujian Province Key Laboratory for Financial Innovation of Science and Technology, Fuzhou 350116, China)

3) (Fujian Province Center for Enterprise Development and Research, Fuzhou 350116, China)

( Received 6 January 2017; revised manuscript received 6 April 2017 )

## Abstract

This paper mainly uncovers the typical microscopic characteristics of Chinese capital market in three different stock price stages of rising, steady and falling based on the high frequency data of Shanghai composite index. Firstly, by analyzing the probability distribution of the Shanghai composite index in different time intervals, we clearly find that the logarithmic change of the index presents an obvious heavy tail feature as well as non-Gaussian Levy distribution, and the return series converges to a normal distribution with the increase of the time interval, which becomes more significant especially in the falling stage of stock prices. Secondly, by calculating the autocorrelation function, we observe that unlike the return rate, the fluctuation ratio of Shanghai composite index demonstrates remarkable long memory volatility with a periodicity of about 240 min, and the autocorrelation curve in falling stage is much higher than in rising and steady stages. Thirdly, in the multi-fractal structure, the volatility of return series has significant short-term and long-term differences among three different stages of rising, steady and falling due to the effects of time limitation and liquidity of investment. Finally, the macroscopic behavior of the Shanghai composite index is relatively consistent with that of the international mature stock market, however, the corresponding microscopic characteristics demonstrate significant differences due to the fact that the Chinese capital market is strongly dependent on the macroeconomic policy, investor sentiment, and liquidity levels. It is quite remarkable that the tail distribution of mature stock market is much fatter than that of Chinese stock market because of the special control and limit mechanism of stock prices in China, which finally causes the considerably lower amplitude of price fluctuation. Moreover, it is also found that the attenuation speed of the autocorrelation function in the Chinese capital market is obviously slower than that in the mature stock market, which suggests that the behaviors of investors in Chinese stock market are more likely to be influenced by the historic exchange information. At the same time, the periodicity of autocorrelation function is actually caused by the inertia recoil of investors, which further verifies the information asymmetry of Chinese stock market. Especially, by changing the starting values of the samples, we find that the periodicity of autocorrelation function still remains the same, which indicates that the periodicity characteristic of stock price is not dominated only by the intraday pattern of trading activity. Therefore, the investors should discover the underlying rules of high-frequency data and extract more useful knowledge in order to guide their investment decisions more effectively.

**Keywords:** capital market, microscopic characteristics, Shanghai composite index, high frequency data

**PACS:** 02.50.-r, 02.50.Cw, 02.50.Ng, 02.60.Ed

**DOI:** 10.7498/aps.66.120203

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 71573042, 71171056) and the Society Science Foundation for Young Ph. D. of Fujian Province, China (Grant No. FJ2016C200).

† Corresponding author. E-mail: [tingling69@163.com](mailto:tingling69@163.com)