

# R/S 分析方法在股票市场分形特征中的应用研究

## 摘 要

分形和混沌理论是当今科学最前沿的学科，本文通过使用 R/S 分析方法，研究了中国股市的分形特征。首先，通过使用 ARMA 模型来作回归消去线性依赖，再对消去线性依赖的序列进行 R/S 分析，R/S 分析方法是一个很稳健的方法，能够发现系统中的非线性依赖。然后使用 Matlab 和 Excel 软件作为工具，对上证综合指数的收盘数据进行了 R/S 分析，发现中国股市不服从正态分布，而是个具有长期记忆的有偏随机游走。通过对数据进行分段回归，发现股市存在一个非周期循环，并通过 V 统计计算出了该循环的长度。股市中存在许多种类型的噪声，将 R/S 分析应用于有噪声的序列以检验 R/S 分析方法是否依然有效，发现 R/S 分析依然能发现过程中的非线性并能得出存在的非周期循环，可见，R/S 分析对于噪声来讲也是非常稳健的。最后，通过使用 BDS 统计，进一步证实了中国股市并非独立过程，而是二阶相关的，中国股票市场并不是个有效的市场。本论文得出的结论基本上能够如实地反映中国股票市场的基本分形特征，能给股票投资者提供了一个好的分析方法和建议。

**关键词** 分形，R/S 分析，赫斯特指数，随机游走，非周期循环，BDS 统计，

# Research and Application of Rescaled Range (R/S) analysis in Fractal Characteristic of Stock Market

## Abstract

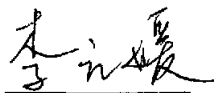
Fractal and Chaos theories are among the most advanced sciences in the world. The application of Rescaled Range (R/S) analysis in fractal characteristic of security market is researched in this paper. Firstly, ARMA model is used to regress a series to recede linear depending. Then R/S analysis is used to analyze the series has been receded linear depending. It proves that R/S analysis is a stable method to find the nonlinear depending in system. Secondly, by using Matlab and Excel softwares, closing quotation datas of aggregative index number of Shanghai Security Market is analyzed by R/S analysis. It proves that Stock Market isn't obey normal distribution and is a random walk with deviation. A non-periodic cycle is found in stock market by regressing the date. Using V-Statistic can calculate the length of this non-periodic cycle. It exist many kinds of noise in stock market, R/S analysis is used to analysed series with noise in order to analyze if it is still efficient to the series with noise. R/S Analysis can find the nonlinear depending and the non-periodic cycle in process with noise, it proves that R/S Analysis is stable for noise. At last, BDS statistic is used to prove that stock market in China is not an independent process but a process with two dimension correlation in further. Stock market in china is not an efficient market. The conclusion in this paper can faithfully evaluated the fractal characteristic of stock market in China. So this method can provide appropriate advices for investors in stock market.

**Key words** fractal, Rescaled Range (R/S) analysis, Hurst exponent, random walk, non-periodic cycle, BDS statistic

## 学位论文独创性声明

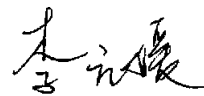
本人郑重声明：

- 1、坚持以“求实、创新”的科学精神从事研究工作。
- 2、本论文是我个人在导师指导下进行的研究工作和取得的研究成果。
- 3、本论文中除引文外，所有实验、数据和有关材料均是真实的。
- 4、本论文中除引文和致谢的内容外，不包含其他人或其它机构已经发表或撰写过的研究成果。
- 5、其他同志对本研究所做的贡献均已在论文中作了声明并表示了谢意。

作者签名：   
日期： 2006. 5. 19

## 学位论文使用授权声明

本人完全了解南京信息工程大学有关保留、使用学位论文的规定，学校有权保留学位论文并向国家主管部门或其指定机构送交论文的电子版和纸质版；有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅；有权将学位论文的内容编入有关数据库进行检索；有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

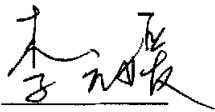
作者签名：   
日期： 2006. 5. 19

## 关于学位论文使用授权的说明

本人完全了解南京信息工程大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

**(保密的论文在解密后应遵守此规定)**

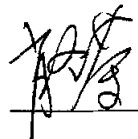
作者签名：



日期：

2006.5.19

导师签名：



日期：

2006.5.19

# 第 1 章 绪论

## 1.1 引言

几乎所有的交易者都希望能够找到适用于股市的理论和分析方法，但是往往结果却并不如预期的那样好。传统的股市运行规律都是基于有效市场假说（Efficient Market Hypothesis, 以下简称为 EMH）做出的研究，当今股市运行规律的理论都是基于以下几个关键性的概念：理性投资者；有效市场；随机游走。因此，得出概率分布近似服从正态分布或对数正态分布的结论。随着研究的深入，虽然大多数实证研究支持了 EMH，但是人们发现了正态收益率假设与实际之间的偏差和越来越多的异常现象，所以 EMH 面临着两个不可避免的挑战：一是正态收益率假设与实际不符，二是存在某些获取超常收益的现象。许多实证研究表明：证券和股市的日、周或月收益率并非正态分布，它们显著地偏离了正态，呈现偏态、宽尾、扁平或尖峰的特征。

由 Fama (1965)<sup>[1]</sup> 最先发现，之后由 Sharpe (1970)<sup>[2]</sup> 注意到，到最近 Turner 和 Weigel (1990)<sup>[3]</sup> 实证研究都证实了的股票市场客观存在着陡峰胖尾现象，这一点动摇了股市运行的正态假定。彼得斯 (Peters, 1991)<sup>[4,5,6,7]</sup> 从非线性的观点出发，提出了看起来不那么完美但却更符合实际的资本市场基本假设——分形市场假说 (Fractal Market Hypothesis, 以下简称为 FMH)，并以大量的实证研究证据支持这一新的假说，从而为人们准确刻画股票市场结构特征开辟了新的视角，为解释 EMH 等现有资本市场理论所不能解释的异常现象提供了新的理论框架。

为了探求股票市场的内在特征，研究人员将近年来研究自然科学发展起来的新方法（非线性分析、动力系统、混沌、分形等）应用于股票市场。其中赫斯特 (H·E·Hurst)<sup>[8]</sup> 创立的 R/S 分析 (Rescaled Range Analysis, 重标极差) 就是这些新方法中较为成功的一个。R/S 分析法是 Hurst 在大量实证研究的基础上提出的一种时间序列分析方法，后经过多人努力逐渐完善。Hurst 发现许多自然现象和经济行为并不呈现纯粹的随机游，而是遵循一个有偏的随机游走过程或分形布朗运动。Mandelbort<sup>[9,10]</sup> 首次将 R/S 分析法应用于美国的证券市场分析股票收益行为。随后，许多学者用 R/S 分析研究了其他国家的证券市场和经济等领域的问题，取得了很多有价值的学术成果。

R/S 分析已经存在了 40 多年了，除了它的稳健性和通用性之外，R/S 分析在很大程度上还不为人知，伴随着其他在传统上已发展的工具以及混沌分析，它在任何分析者的工具包里都会占有一席之地。分形市场假说也会在时间上被修改和完善，并将由投资机构详尽的研究而独领风骚。本文就是使用了 R/S 分析这个显著的工具并联合其他技术来检验了 FMH。真实世界的美景开始浮现，R/S 分析与分形市场假说在分形市场的引导之下走到我们面前。

## 1.2 有效市场假说和分形市场假说<sup>[14~20]</sup>

### 1.2.1 有效市场假说

市场有效性理论是现代金融市场学的基础，美国学者法玛(Fama,1965)提出“如果所有股票价格反映了所有相关信息，市场即有效”。有效市场假说(Efficient Market Hypothesis,EMH)是一个被广泛接受的定义。

为检验这一假设，统计学家从统计学上找出所谓“价格已经反映了所有相关的信息”是什么意思，最终得出的往往是随机游动模型。而随机游动模型是一个布朗运动，它要求连续价格波动之间是独立同分布的。当被度量的事件是“独立的和同一分布的”IID的，中心极限定理表明，只要有足够的样本容量，随机系数的极限分布是正态分布，从而标准高斯统计在这种情况下能正确的运用，即它依赖于正态分布。因为理性的投资者以当时可获得的信息为基础进行交易，其交易价格反映了已有的信息，市场未来的价格变化与现在的信息无关，只反映未来的新信息。而未来的信息又是随机出现的，所以未来的价格变化也是随机的、不可预测的。正是由于资本市场价格变化的独立性，同时理性的投资者又被假设为能够根据期望收益率及其方差进行投资决策。因此，当观测的收益率足够多的时候，EMH 隐含着收益率正态性的假设。

事实上，人们有充分的证据表明股价收益率不服从正态分布。任何依赖于正态分布的统计方法均可能得出错误的结论。为了使检验切合数据，需要一种没有先验假设的概率分布方法。现在提供关于市场结构的另一种理论—分形市场假说(Fractal Market Hypothesis,FMH)，它是有效市场假说的替换物。

### 1.2.2 分形市场假说

FMH 认为：资本市场是由大量的不同投资期望的投资者组成的，信息对各种不同投资者的交易时间有着不同的影响，在每日、周或月时段内的交易未必是均匀，而且投资者的理性是有限的，未必按照理性预期的方式行事。他们甚至不知道如何解读已知的信息，也许仅仅根据经验做出反应，即把过去的信息作为决策的依据。因此，资产价格的变化不是随机游动，而是具有增强趋势的持久性时间序列，今天或未来的资产价格变动与初始状态之间并非相互独立，而是持续相关的。

FMH 强调市场的流动性(Liquidity)和投资者的投资期望水平(Investment horizons)。分形市场假说认为所有的投资者在市场上具有不同的投资期望水平，信息对于具有不同投资期望水平的投资者价值是不同的。流动性的产生是因为投资者具有不同的投资期望水平，不同的信息集随之产生不同的“公平价格”。具有不同投资期望水平的投资者面临的风险水平是相同的，这可通过比较日收益率和周收益率的频率分布图看出。当突发事件爆发时，如经济危机、政治危机、石油危机改变了投资者对长期基本面的看法，从

而降低了投资期望。这与熊市不同，熊市是因为长期投资价值的降低，市场剧变的是因为出现大量的短线交易者，结果是剧烈的下跌或上升。FMH 着眼于“少量的大变化”，EMH 着眼于“大量的小变化”。与 EMH 认为信息无价值不同，FMH 认为信息依据不同的投资期望水平有不同的价值，故信息的扩散是不平衡的。在任何时间股价也许不是反映了所有公开的信息，而是仅反映了对于特定投资期望水平的信息。

### 1.2.3 EMH 与 FMH 的区别

不同于有效市场假说 EMH, 分形市场假说 FMH 认为，信息依其投资者的投资起点而被评价。因为不同的投资起点对信息的估价是不同的，信息的传播也是参差不齐的。价格不可能每一次都反映所有可适合的信息，而只反映对投资起点重要的信息。

Vaga(1991)<sup>[11]</sup>的协同市场理论 (Coherent market Hypothesis) 和 Larrain(1991)<sup>[12]</sup>的 K-Z 模型(K-Z Model)支持了 FMH。FMH 基于这样一个前提，即市场假定不同的状态并能在稳定与不稳定的状态下变换。当投资者对长期基础信息失去信心时，发生了混沌方式。在许多方面，FMH 通过使用投资起点连结了这两个模型。当体制发生变化，它就特征化了，并且，当基础信息失去它的价值时，市场就变得不稳定。当它没有特征性时间规模变化或投资起点时，FMH 就可以称市场是稳定的。当市场失去它的分形结构并且假定有唯一恰好的投资起点时，不稳定性发生。

下面的表格列出了 EMH 和 FMH 的区别<sup>[20]</sup>。

EMH 和 FMH 的区别

比较项目	EMH	FMH
市场特性	线性、孤立系统	非线性、开放、耗散系统
均衡状态	均衡	允许非均衡
系统复杂性	简单系统	具有分形、混沌等特性的复杂系统
反馈机制	无反馈	正反馈
对信息的反应	线性因果关系	非线性因果关系
收益序列	白噪声、不相关	分数噪声、长记忆、对于初始值敏感
价格序列	布朗运动 ( $H=0.5$ )	分数布朗运动 ( $0.5 < H < 1$ )
风险度量指标	有限方差	Hurst 指数
可预测性	不可预测	提供了一个预测的新方法
波动有序性	无序	有序
二者之间的关系	有效市场是分形市场的一个特例, 分形市场拓展了有效市场的含义, FMH 更广泛的, 准确的刻画市场	

FMH 勾勒了一个市场结构的新概念。大部分标准的市场分析都假定市场过程基本上是随机的, 因为要检验 EMH, 这一假设几乎没有任何问题。对于 FMH, 虽然许多标准检验失效, 这并不是说这些检验无用。在第二章的篇幅中, 将检测这样一个方法论: R/S 分析。它是一个时间序列分析的稳健形式。

### 1.3 论文内容概述

本文将在参考国内外有关文献<sup>[1~47]</sup>的基础上, 运用 R/S 分析方法来研究中国股票市场的分形特征问题, 以进一步揭示中国股票市场的内在特性。

在导师的精心指导下, 本人在论文中主要对以下几个方面进行了研究。

1、通过对上证综指收益率序列的 R/S 分析, 分析了其 Hurst 指数, 发现了中国股市存在的分形特征。中国股市不是独立的过程, 而是具有长期记忆的有偏随机游走。

2、通过对上证综指的日收益率、周收益率进行分段回归, 发现了股市中存在的非周期循环, 并通过 R/S 分析和 V 统计计算出了非周期循环的长度。同时发现中国股市存在自相似性, 股市中存在的非周期循环在不一样的分段回归中同样的出现, 并且周期基本上一致。

3、通过对相同的数据进行 BDS 统计, 发现中国股市是二阶相关的, 更进一步的证明了中国股市不是独立过程, 中国股票市场不是有效的市场, 而是分形的市场。

### 1.4 论文的创新点或者说是改进之处

(1) 数据采样区间的不同。很多论文采用 R/S 分析对股市进行分析, 但是他们大多是选取了中国股市 1990 年左右期间的数据, 本论文选取了上证综指的 1999 年 1 月 1 日—2005 年 12 月 31 日的数据, 也可以说是选取了中国股市比较近期的数据。因此, 得出的结论相信能够更紧随经济发展的趋势, 表现中国股市的现状。

(2) 数据处理、分析的角度不同。很多论文只是对沪市或深市的单日或周收益率进行了研究以概括沪深股市的规律, 比较片面。而本文从纵向角度分析了上证综指的 1 日, 5 日, 20 日, 30 日, 90 日和 180 日收益率的分形结构, 找出了沪市不同时间增量收益率的不同分形结构, 并根据它们的变化趋势, 找出了中国股市分形结构的时间敏感性, 或者说是时间稳定性规律。



## 第 2 章 R/S 分析方法与 Hurst 指数

标准的统计分析源于一个假定，即系统基本是在随机情况下研究的，也就是产生时间序列的因果过程有许多组成部分或自由度，而且那些组成部分的相互关系是如此复杂以至于确定性的说明是不可能的。标准的高斯统计在非常有限的假定下才能很好的运行。中心极限定理（或大数法则）表明，只要我们做越来越多的试验，随机系统的极限分布将是正态分布，或是钟形曲线。被度量的事件必须是“独立的和同一分布的”（IID）。但是，倘若研究下的系统不是 IID，或者接近，那么我们将需要一个无参数的方法。英国水文学家 H. E. Hurst 提供了一种方法，叫做重标极差方法，或是 R/S 分析<sup>[15, 20~24]</sup>，用于从分形式间序列里区分随机序列。下面将讨论这一方法。

### 2.1 R/S 分析的发展

布朗运动是随机游走的基本模型，涵盖时间平方根增长的随机分子被用来测度它的距离，或：

$$R = T^{0.50} \quad (2.1)$$

这里，R = 被涵盖的距离；T = 时间指数。

等式 (2.1) 被叫做二分之一法则，在金融经济学上应用它来做易变性或标准差的年度变化。

以一时间序列来开始， $x = x_1, \dots, x_n$ ，代表 n 个连续值。时间指标对于市场来说，它可能是股票指数的价格每日变化。时间系列 x 的均值  $x_m$  可定义如下：

$$x_m = (x_1 + \dots + x_n) / n \quad (2.2)$$

标准差  $s_n$ ，是如下估计：

$$s_n = n^{-1/2} * \sqrt{(x_r - x_m)^2} \quad (2.3)$$

重标极差由一阶重标或“标准化”的减去样本均值的数据计算出来。

$$Z_r = (x_r - x_m); \quad r = 1, \dots, n \quad (2.4)$$

接下来的序列 Z，现在有零均值。下一步产生一个累计时间序列 Y：

$$Y_1 = (Z_1 + Z_r); \quad r = 2, \dots, n \quad (2.5)$$

由定义知道，最后一个  $Y(Y_n)$  将永远是零，因为  $Z$  有一个零均值。调整过的极差， $R_n$  是  $Y_t$  的最大值减去最小值。  $R_n = \max(Y_1, \dots, Y_n) - \min(Y_1, \dots, Y_n)$  (2.6)

对于  $R_n$ ，下标  $n$  现在表明对于  $x_1, \dots, x_n$  是一个调整过的极差。因为  $Y$  已被调整为零均值， $Y$  的最大值总是大于或等于零，而且最小值总是小于或等于零。这样调整的极差  $R_n$ ，将永远是非负的。这个调整的极差， $R_n$  是系统对于时间指数遍历的距离。等式 (2.1) 仅适用于布朗运动的时间序列：具有零均值，且方差为 1。对于不是布朗运动的时间序列运用这一概念，我们需要一般化等式 (2.1)，并且解释非独立的系统。下列是等式 (2.1) 更一般的形式：

$$(R/S)_n = c * n^H \quad (2.7)$$

对于  $(R/S)_n$ ，下标  $n$  被考虑为对  $x_1, \dots, x_n$  的  $R/S$  值； $c = a$  常数。因为具有零的均值和以局部标准差形式出现，等式 (3.7) 被称作重标极差。在一般情况下， $R/S$  值以我们增加的时间增量而规模变化， $n$ ，基于幂法则，值等于  $H$ ，通常称作赫斯特指数。赫斯特指数能够由绘制的  $\text{Log}(R/S_n)$  与  $\text{Log}(n)$  的标绘图逼近，而且能通过最小二乘法回归解出斜率。

$$\text{Log}(R/S_n) = \text{Log}(c) + H * \text{Log}(n) \quad (2.8)$$

假如一个系统是独立分布，那么  $H = 0.50$ ，Hurst 第一个调查了尼罗河，他发现  $H = 0.91$ 。Hurst 决定检查其它河流，然后将自然现象检验数据变为更丰富多彩的一降雨、太阳黑子、无价值情绪、年轮等类型，或任何具有长时间序列的东西。会不会存在一种共有现象发生的可能性？结果如图 2.1<sup>[15]</sup>。

所有的这些情况，Hurst 发现  $H$  都大于 0.5。他对  $H$  通常取得逼近 0.70 的这一值的现象发生了兴趣。一般现象的发生会具有某种品质？

Hurst 感觉有偏随机游走将导致他的结果，他需要方法来模拟。他设计了一个精美的过程来充当赫斯特过程的模拟。接下来这部分将帮助你了解赫斯特现象

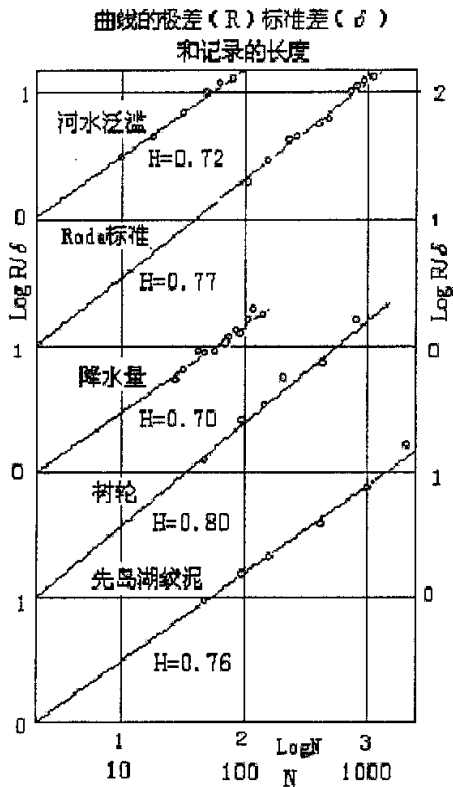


图 2.1

## 2.2 对赫斯特过程的模拟——王牌现象

Hurst 用一副牌来模拟随机过程，那就是说，用标有  $\pm 1$ 、 $\pm 3$ 、 $\pm 5$ 、 $\pm 7$ 、 $\pm 9$  的 52 张牌来逼近正态分布。洗乱这副牌，记住重复卡牌的数，产生一个随机时间序列。在已经产生逼近 0.5 的赫斯特指数的结果序列上实施 R/S 分析。它足够的接近以满足活动时期标准。Hurst 实施了 1000 次试验而没有发现变化的斜率。

接下来模拟一个有偏的随机行走。他首先洗牌，卡牌并记录号码。例如，我们用 +3 作为最初的卡牌，再重新放回并洗牌。他将处理各自 26 张的两手牌，我们称其为 A 手牌和 B 手牌。因为最初的号码是 +3，我们将从 A 手牌中拿出 3 张最大的并放到 B 手牌。然后从 B 手牌中去掉 3 张最小的，这样，B 手牌将有偏 3 个水平。最后，将王牌放在 B 手牌并洗牌。他用一个现已有偏的 B 手牌作为他的时间序列发生器，直到他卡到王牌。然后 Hurst 可再产生一副新的有偏的牌。

Hurst 用 100 手牌作了 1000 次试验。他计算  $H = 0.72$ ，远大于他实际上所作的。考虑过程所包括的：首先，每副牌是有偏的，它们决定于随机卡牌的次数；再者，时间序列自身的产生，它是随机卡牌数的另一序列；最后，王牌的出现，也是在随机情况下出现。不管所有这些随机事件的运用， $H = 0.72$  却总是出现。局部的随机性和整体的结构性极像混沌现象。在这种情况下，它是一个整体统计结构而不是几何结构。

假如市场是赫斯特过程，它们展现一个一直到王牌出现的从大格局上，或方向上，或两者兼之的改变有偏的经济等价物的趋势。

### 2.3 解读赫斯特指数

根据原始理论， $H = 0.50$  将暗示一个独立过程。R/S 分析不要求划分基础过程为高斯，仅要求为独立。这当然包括正态分布，但它也包括非高斯独立过程。R/S 分析是非参数的，因此，对划分分布没有任何基础形态上的要求。

$0.5 < H \leq 1.00$  暗示了一个持续的时间序列，而这一序列被长期记忆效用特征化了。从理论上讲，今天所发生的一切都将永远影响将来。用混沌的话来说就是，存在对初始条件敏感的依赖情形。这个长期记忆发生并不考虑时间的比例变化。所有逐日变化是相关所有未来的逐日变化；所有周的变化是相关所有将来周的变化。不存在特征性时间规模变化，即分形时间序列的关键特征。

$0 < H \leq 0.50$  表明一个反持续性的时间序列。一个反持续性系统比随机系统覆盖较短的距离。因为一个系统覆盖较短的距离，它必定比随机过程以较高的频率翻转自身。具有正统背景的理论家会将这一行为等同为均值回复过程。因此，反持续假定研究下的系统具有稳定的均值。我们却不能做出这一假定。

### 2.4 Hurst 指数特性与 EMH

时序的分形维  $D$  是时间轨迹的分形维，度量时间序列的参差不齐性，它与 Hurst 指数的关系为：

$$D = 2 - H \quad (2.9)$$

为了度量现在对未来的影响，可以引入相关性度量  $C$ ，

$$C = 2^{(2H-1)} - 1 \quad (2.10)$$

现在把 Hurst 指数和分形维  $D$ 、相关性度量  $C$  结合起来分析 Hurst 指数的特性

(1)  $H = 0.5, D = 1.5, C = 0$ ，变量之间是相互独立的，相应的相关系数是零，现在不会影响未来，时序是随机的。若该时序是股指或者股价时，这表明市场是有效的。

(2)  $0 \leq H < 0.5, D > 1.5, -0.5 \leq C < 0$ , 变量之间是负相关, 称之为反持续性的, 遍历性的或均值回复的时间序列。如果某一时刻序列向上(下), 那么下一时刻它很可能反转向下(上)。由于频繁地出现逆转, 反持续性的时序具有比随机噪声更剧烈的波动性。若该时序是股指或者股价时, 这表明市场不是有效的。

(3)  $0.5 < H \leq 1, D < 1.5, 0 < C \leq 1$ , 持续性或分形时序, 波动比较平缓。其变量之间不是相互独立, 而是正相关, 时序具有持久性或趋势增强的特性。如果某一时刻序列向上(下), 那么下一时刻它很可能继续向上(下)。H 越接近 1, 这种持续性越强。若该时序是股指或者股价时, 这也表明市场不是有效的。

经验研究发现, 经济系统中的大多数时间序列  $H > 0.5$ , 表明其中具有长期记忆效应, 并不遵循随机游走。

## 2.5 R/S 分析的实际操作指南

R/S 分析是一个高度数据集中的简单过程。这部分通过等式 (2.8) 打破了等式 (2.2) 进入一系列执行步骤。这些是后续步骤:

1. 以长度  $M$  为开始, 并把长度  $M$  转换成对数比的长度  $N = M - 1$  的时间序列:

$$N_i = \text{Log}(M_{(i+1)} / M_i) \quad i = 1, 2, 3, \dots, (M - 1) \quad (2.11)$$

2. 均分这个时间区间长度为  $N$  的相邻的子区间  $A$ , 因而  $A * n = N$ 。标记每个子区间为  $I_a$ ,  $a = 1, 2, 3, \dots, A$ 。在子区间  $I_a$  中, 每一个元素标记为  $N(k, m)$ ,  $k = 1, 2, 3, \dots, n$ 。长度为  $N$  的子区间的平均值被定义为:

$$e_a = (1/n) * \sum_{k=1}^n N_{k,a} \quad (2.12)$$

这里,  $e_a$  = 长度为  $n$ , 包括在子区间  $I_a$  中的  $N_i$  的均值。

3. 作为每一个子区间  $I_a$  对于均值的累积横距 (XKA) 的时间序列定义如下:

$$X_{k,a} = \sum_{i=1}^k (N_{i,a} - e_a) \quad k = 1, 2, 3, \dots, n \quad (2.13)$$

4. 极差被定义为在每一个子区间  $I_a$ ,  $X_{k,a}$  的最大减去  $X_{k,a}$  的最小:

$$R_{I_a} = \max(X_{k,a}) - \min(X_{k,a}) \quad (2.14)$$

这里:  $1 \leq k \leq n$

5. 对于每一个子区间  $I_a$  的样本标准差:

$$S_{I_a} = (1/n) * \sum_{k=1}^n (X_{k,a} - \bar{e}_a^2)^{0.50} \quad (2.15)$$

6. 每一个极差  $R_{I_a}$ , 是由对应于它的标准差  $S_{I_a}$  分割而正式化的。这样一来, 每一个  $I_a$  子区间的重标极差即等于  $R_{I_a} / S_{I_a}$ 。从第二步往后的步骤, 我们有  $A$  个长度为  $n$  的相邻子区间, 这样一来, 作为长度为  $n$  的 averages 的  $R/S$  值定义如下:

$$(R/S)_n = (1/A) * \sum_{a=1}^A (R_{I_a} / S_{I_a}) \quad (2.16)$$

7. 对于下一个较高的值, 长度  $N$  是增加的, 而且  $(M-1)/n$  是一个整数值。我们使用那个包括时间序列的起止点  $N$  的值, 步骤从 1 到 6 将重复到  $n = (M-1)/2$ 。

我们现在对  $\text{Log}(n)$  为独立变量,  $\text{Log}(R/S)_n$  为因变量, 应用等式 (2.7) 和 (2.8) 实施普通最小平方回归。截距是对于常数  $\text{Log}(c)$  的估计。斜率是赫斯特指数  $H$  的估计。

现在加进笔者在进行分析时候得出的一个经验法则: 一般而言, 对于  $n \geq 10$  运用回归。当样本容量小,  $n$  的小值产生不稳定估计。

## 第3章 检验 R/S 分析<sup>[15, 25~31]</sup>

### 3.1 随机零假设

假如我们理解不了诸如股票市场这样特殊程序背后的机制,那么,独立的,同分布(IID)的,而且是被随机行走特征化了的这样一个统计结构,是我们最好的猜想。结构是高斯的,而且它的概率密度函数是正态分布,或者钟形曲线。我们称这一最初的猜想为零假设。我们选择高斯情形作为零假设的条件,只因为它简单。从数学上讲,检验一个过程是否是随机游走、或能否说它不是随机游走,那就是,将要证明分形布朗运动的存在(或者某种其他的长期记忆过程)。因为高斯情形提供最优解,而且容易被模拟。除此之外,有效市场假说依据高斯条件,通过违约,把它设为零假设[15]。

Hurst(1951)在他的二项分布上的零假设,结果对于随机游走而言,等式(3.1)是一特殊情形:

$$(R/S)_n = (n * \pi / 2)^{0.50} \quad (3.1)$$

这里,  $n$  = 观察值的个数。

Feller(1951)发现了相似的结果,他严谨地研究了调整过的极差  $R'$ 。Hurst 认为等式(2.17)为重标极差,但他没有正式地证明它。Feller 研究了调整过的极差(具有删掉样本均值的累计偏差),而且发展了  $R'$  的期望值和它的方差。因为样本标准差的行为,重标极差  $R/S$  被认为是难驾驭的,特别是  $N$  取小值的时候,人们认为,调整过的极差可以解出并渐近(即在无限的情况下)为重标极差的等价物,这个结果充分逼近重标极差。

Feller(1951)发现了下列公式,在调整过的极差的期望值上,而且他也能算出它的方差,他们基本上等同于赫斯特等式(3.1):

$$E(R'(n)) = (n * \pi / 2)^{0.50} \quad (3.2)$$

$$Var(E(R'(n))) = (\pi^2 / 6 - \pi / 2) * n \quad (3.3)$$

方差公式,等式(3.3),提供了  $R'(n)$  的一个值的方差,因为我们能够预期随机数的  $R/S$  值将是正态分布的,  $R'(n)$  的方差将下降。等式(3.1)(3.2)是在布朗运动的零假设条件下的标准假定。极差随时间的平方根增长。Feller 讲过,这个极差的方差随时间呈线性增长。

### 3.2 蒙特卡洛 (Monte Carlo) 模拟

对于由随机数发生器产生的随机数，可以通过等式 (2.7) (2.8) 来模拟 R/S 分析，并能根据经验计算平均数和方差，观察它们是否与等式 (3.1) (3.2) 和 (3.3) 一致。这个过程即为著名的“蒙特卡洛”模拟方法，它特别适用于检验高斯假设。

开始做这件事之前，必须处理一下“随机数值”的神话。任何一个随机数发生器都不产生一个真正的随机数。相反的，根据大部分高斯检验，算法产生的伪随机数是统计独立分布的。它们开始重复之后，这些伪随机数实际上有很长的循环或记忆。通常，这个循环长到足以保证重复而不被察觉的地步。当大量数据被应用到蒙特卡洛模拟中时，伪随机数能够销蚀结果。对于作随机数的发生器的大多数算法来说，它们是混沌系统的翻版。R/S 分析特别擅长揭示确定性混沌和长期记忆过程。

开始于 5000 个（具有零均值和标准差为 1 的正态分布）拼凑了两次伪随机数序列，对所有能够整除 5000 的 n 计算 R/S 值，那就是，每个  $R/S_n$  值总是包括全部时间序列的开始和终止值。我们重复这个过程 300 次，结果就得到了每个 n 的 300 个  $R/S_n$  的值。对于高斯随机数的系统，这些  $R/S_n$  的均值是一期望值  $E(R/S_n)$ 。方差可以计算出来，终值可以和通过等式 (3.1) (3.2) (3.3) 计算得到的数值相比较。结果列在表 3.1。

当 n 大于 20 时，被模拟的  $R/S_n$  的值收敛于 (3.1) (3.2) 的值。因此，对于 n 的小组织，存在一个一致性偏差。由模拟所产生的  $R/S_n$  值从整体上低于从 Feller 和 Hurst 等式得到的值。  $R/S_n$  的方差也总是系统的低于费勒等式 (3.3)。

多年以后，Anis 和 Liloyd (1976) 围绕小 n 的 R/S 统计的系统偏差发展了下列方程：

$$E(R/S_n) = [\Gamma\{0.5 * (n-1)\} / \sqrt{\pi} * \Gamma(0.5 * n)] * \sum_{r=1}^{n-1} \sqrt{(n-r)/r} \quad (3.4)$$

等式可以化简为：

$$E(R/S_n) = (n * \pi / 2)^{-0.50} * \sum_{r=1}^{n-1} \sqrt{(n-r)/r} \quad (3.5)$$

通过用校正因子乘以等式 (3.4) (3.5) 可以对 Anis 和 Liloyd 公式校正：

$$E(R/S_n) = ((n-0.5)/n * (n * \pi / 2))^{-0.50} * \sum_{r=1}^{n-1} \sqrt{(n-r)/r} \quad (3.6)$$



假如独立过程不是高斯过程而是一个胖尾陡峰的独立分布，均值如同等式 (3.6) 所预期，但是由于没有被正态分布，方差就不相同。置信区间仅仅对于 IID 随机变量有效，这样，我们就需要一个能够过滤掉短期依赖性的方法。下面就将介绍用于滤去短期依赖性的几个经济模型。

观测数	Monte Carlo	Hurst	Anis 和 Lloyd
10	0.4577	0.5981	0.4805
20	0.6530	0.7486	0.6638
25	0.7123	0.7970	0.7208
40	0.8332	0.8991	0.8382
50	0.8891	0.9475	0.8928
100	1.0577	1.0981	1.0589
125	1.1097	1.1465	1.1114
200	1.2190	1.2486	1.2207
250	1.2710	1.2970	1.2720
500	1.4292	1.4475	1.4291
625	1.4801	1.4960	1.4795
1000	1.5869	1.5981	1.5851
1250	1.6351	1.6465	1.6349
2500	1.7839	1.7970	1.7889

表 3.1

### 3.3 金融时间序列的几个随机模型

假设，短期记忆过程的五个类型为金融时间序列：

1. 自回归 (AR);
2. 移动平均 (MA);
3. 自回归移动平均 (ARMA);
4. 自回归集中移动平均 (ARIMA);
5. 自回归条件异方差性 (ARCH);

以上每一个时间序列都有许多变量，他们都是基本模型的提纯。这些提纯了的变量使得时间序列的特征更接近实际数据，下面依次考察这些过程。

### 3.3.1 自回归过程

自回归过程是一个变量在时间的某一点的变化是线性地相关于前期的变化的过程。一般来说, 相关随时间呈指数性的下降而且在相对比较短的周期消失。一般过程如下:

$$C_n = e_n + a * C_{n-1} + b * C_{n-2}$$

(3.7)

这里,  $C_n = C$  在时间  $n$  的变化,  $0 \leq C \leq 1$ ;  $a, b$  = 具有  $|a| \leq 1, |b| \leq 1$  的常数;  $e$  = 具有零均值和方差的白噪声序列  $\delta_e^2$ 。

等式 (3.7) 是一个二阶自回归过程, 或  $AR(2)$ , 因为在时间  $n$  的变化涉及到最后两个周期的变化。也可能有  $AR(q)$  过程, 即  $C$  在时间  $n$  的变化依赖前  $q$  个周期。检验  $AR$  过程的可能性, 当在时间  $n$  的变化依赖变量, 而且在前  $q$  个周期 (滞后) 的变化是作为独立变量使用, 可视作一个回归的运行。每一个滞后的  $t$ -统计都已经被评价了。假如在 5% 的水平所有  $t$ -统计是显著的, 我们就可以形成  $AR$  过程在运行的假设。就系数的值的极差的限制确保过程是稳定的, 这意味着, 在均值或方差方面, 都没有或往上、或往下的长期趋势。

高频金融时间序列 (日变化或更频繁) 通常展示明显的自回归倾向。因为高频率数据是基本的交易数据, 而交易者的确互相影响。

### 3.3.2 移动平均过程

在移动平均过程 (MA) 里, 时间序列是一个未观测到的时间序列的移动平均的结果:

$$C_n = c * e_{n-1} + e_n \quad (3.8)$$

这里,  $e$  = 一个 IID 随机变量;  $c$  = 常数, 且  $|c| < 1$ 。

在移动平均参数  $c$  上的限制确保了过秤时可转换的。 $c > 1$  暗示: (1) 未来事件影响现在的时间, 可他有点不现实; (2) 过程是稳定的。对于  $e$  的限制, 就如同  $AR$  过程中的随机震颤, 它是一个具有零均值和方差为  $r$  的 IID 变量。

已观测到的时间序列  $C$ , 是未观测的随机时间序列的移动平均的结果  $e$ 。由于移动平均过程, 在过去和短期记忆结果上存在一个线性依赖。可是, 与  $AR$  过程不同, 随机震颤只有一个周期的记忆。

### 3.3.3 ARMA 模型

在这类模型中，既有自回归项，又有移动平均项。移动平均项也是一个未被观测到的随机序列：
$$C_t = a * C_{t-1} + e_t - b * e_{t-1} \quad (3.9)$$

这种类型的模型被称为混合模型，而且被冠之  $ARMA(p, q)$  模型。 $p$  是自回归项的个数，而  $q$  是移动平均项的个数；即，一个  $ARMA(2, 0)$  过程和  $AR(2)$  过程一样，因为它们没有移动平均项。一个  $ARMA(0, 2)$  过程和  $MA(2)$  过程相同，因为它没有自回归项。

### 3.3.4 ARIMA 模型

$AR$  和  $ARMA$  两模型可合并成为一个更一般类型的过程。 $ARIMA$  模型是专门用于不稳定时间序列的。这些过程在它们的均值和方差里，有一个基础的倾向。可是，由于采取了数据的累次差分，结果是稳定的。

例如，只因为有了长期增长因素，价格序列就是不稳定的了。它能够无边界的增长，以至于价格自身将不再倾向平均值。一般来讲，有效市场假说  $EMH$  是能够接受价格（或收入）的变化是稳定的。通常，价格变化是专门用百分比来表示的，或是对数差分，这仅仅是一阶差分。在一些序列里，高阶差分可能用来使数据稳定。例如，不同差分的区别是一个二阶  $ARIMA$  过程，它可能走到更高阶的差分。

$C_t$  是一个  $d$  阶齐次非稳定过程，假如：

$$w_t = \Delta^d C_t \quad (3.10)$$

是稳定的。 $\Delta$  代表了差分化，而  $d$  代表了什么程度的差分是必须的。例如：

$$\Delta C_t = C_t - C_{t-1}$$

$$\Delta^2 C_t = \Delta C_t - \Delta C_{t-1}$$

等等。

假如， $W_t$  是一个  $ARMA(p, q)$  过程，那么  $C_t$  被认作是  $(p, d, q)$  阶的整和自回归移动平均过程，或是  $ARIMA(p, d, q)$  过程。而且， $p$  是自回归项的个数， $q$  是移动平均项的个

数。参数  $d$  被认为是所需差分运算的个数。过程不必混合。假如  $C_t$  是一个  $ARIMA(p, d, 0)$  过程, 那么  $W_t$  是一个  $AR(p)$  过程。同样, 假如  $C_t$  是一个  $ARIMA(0, d, q)$  过程, 那么  $W_t$  是一个  $MA(0, q)$  过程。

### 3.4.5 ARCH 模型

展示自回归条件异方差性的模型 (ARCH) 主要基于以下几点理由:

1. 它们, 针对于  $AR$  或  $MA$  线性依赖过程, 是一组非线性随机过程;
2. 它们的频率分布是陡峰胖尾分布;
3. 经验研究已表明, 金融时间序列从统计上显示了显著的 ARCH 特征。

基本的 ARCH 模型是由 Engle(1982)发展起来的。Engle 认为时间序列是由正态概率分布所定义的, 而非时间依赖方差所定义; 一个过程的期望方差, 以它的前期为其先决条件。尽管方差对于个体分布是稳定的, 但也会呈现“时间变化”, 条件异方差性过程因而得名。这个过程也是自回归的, 它具有时间依赖特征。一个样本频率分布将是这些扩大和收缩的正态分布的平均值。正因为如此, 在时间的任何一点上, 它可能呈现出陡峰胖尾分布。基本的 ARCH 模型如下定义:

$$C_n = s_n * e_n$$

$$S_n^2 = f_0 + f * e_{n-1}^2$$

(3.11)

这里,  $e$  是一个标准的正态随机变量;  $f$  是一个常数。

为了方便的缘故,  $f_0 = 1$ , 而且  $f = 0.50$ , 是典型值。我们再次看见 ARCH 模型同以前讨论的 AR 模型有相似性: 被观测到的值中,  $C$  再次是未观测到的序列的结果,  $e$  是依赖过去自身的实现。然而, ARCH 模型是非线性的。小的变化似乎跟随着小的变化, 而大的变化跟着大的变化, 可是, 征兆就不会被预测。因为 ARCH 是一个非线性的模型, 大的变化将放大, 而小的变化将收缩。这导致了胖尾陡峰分布。

ARCH 模型被改为也依赖于过去的  $s$  变量。Bollerslev(1986)以下列方程正规化了一般的 ARCH (或 GARCH) 模型:

$$C_n = s_n * e_n$$

$$s_n^2 = f_0 + f * e_{n-1}^2 + g * s_{n-1}^2 \quad (3.12)$$

对于 GARCH 一般是典型地设  $f_0 = 1, f = 0.1$  和  $g = 0.8$ , 尽管所有这三个变量能从 0 到 1 变化。GARCH 也产生了胖尾陡峰频率分布。等式 (3.11) 和 (3.12) 是基本的 ARCH 和 GARCH 模型; 存在许多变化。扩大了 ARCH 和 GARCH 模型成为和谐特征的典范, 模型能较好地与经验观测一致。然而, 在 ARCH 和 GARCH 过程的规模变化的性质里, 没有变化, 尽管变化改变了模型的理论模型。

因为基本的 ARCH 和 GARCH 模型有许多符合经验数据的性质, 模拟, 被模拟的 ARCH 和 GARCH 的值, 对 R/S 分析是非常优秀的检验。

### 3.4 随机模型中存在的问题

以上简要汇总的几个模型就市场而言, 对于赫斯特过程来说, 是最流行的替换模型。每一个模型都似乎抓住了某种市场的经验发现物, 但是, 至今完全令人满意的却一个也没有。问题好像是, 每个模型都提出一个局部的市场性质。许多这样的局部性质似乎都与某些、而不是与所有的投资起点联系在一起。例如, AR 过程具有极高频率的数据特性, 诸如月收益; GARCH 则是一个非相似性的胖尾陡峰分布, GARCH 的参数表现出周期依赖性, 它不是常数, 且一再被作为规模调整。总之, 这些模型确与分形市场假说不服, 而当调查周期性特殊数据时他们肯定会被认为如此。有一个例外是 ARIMA 族模型的分形的说法。另一个例外是 IGARCH, 它具有无限无条件方差而非有限的条件方差。

## 第 4 章 周期循环和非周期循环

对于一些技术性分析，发现循环是市场分析的同义词。市场像许多自然现象一样，有规律性的落潮和流动。技术人员相信，隐藏在噪声和不规则扰动背后，存在一个规律的市场循环，这个循环驾驭和驱使着市场的基础，像钟表一样工作的机制。这一“循环”对粗心大意的投资者已证明是无常和易变的。有时它们循环，有时它们不循环。诸如谱分析这样的统计检验，仅仅发现相关噪声。所有与之相涉的，在市场和经济里对循环的研究都已证明无效。

在时间序列分析中，它的焦点一直放在规则、周期循环上。在傅立叶分析中，假定形成时间序列不规则形状的是许多周期 sine 波集合，它们各自都有其不同的频率和振幅。谱分析试图用不明显的循环来分裂已观察到的不规则时间序列为这些正弦波。功率谱里的陡峰被认作循环行为的证据，谱分析强加给已观察的无周期时间序列上一个未观察到的周期结构。它不是一个循环，而是一个正弦波或余弦波。

Granger (1964) 第一个提出，谱分析可以供市场时间序列使用。人们做了各种关于这些数据的转换，以便找出直觉感到循环存在的证据，可就是找不到。随后，这个领域里的大多数人放弃了这种寻找，并且得出结论，这是一个错觉。

遗憾的是，不存在任何直观的理由使人相信，市场或经济的循环基础与正弦波或任何别的周期循环有关系。谱分析将是一个不恰当的市场循环分析的工具。在混沌理论里，非周期循环存在。这些循环有一个平均持续，但一个未来循环的确定的持续是未知数。我们需要一个稳健的循环分析工具，一个既能探测周期又能探测非周期的工具<sup>[18~32]</sup>，R/S 分析就具有这一功能。

### 4.1 周期循环

Hurst (1951) 第一个意识到，基础周期能被 R/S 分析探索到。周期系统依赖于有限循环或一个与吸引子同类的循环。在正弦波的情形下，时间序列可能被波的放大所界定。因为极差不可能在放大之外部增长，在一个循环之后，R/S 值可能达到最大。Mandelbrot 和 Wallis (1969a-1969d) 做了一个计算机模拟的扩大序列，特别是考虑到时间上的技术实用性。我们将重复和扩大那些实验丧钟的某些实验，以便当者众多的周期因素，表明 R/S 分析的行为。

我们先使用简单的正弦波： $Y_t = \sin(t)$

这里， $t$  = 一个时间指数。

图 4.1 显示了正弦波为 100 个重复的循环长度的 Log/Log 标绘图。在  $t=100$  处的断点明显的体现出来。R/S 正是用了这个方法抓住这个重要的过程。实际上，当正弦波被涵盖了一个完全循环时，它的极差停止增长。因为它达到了它的极大振幅。它的最大极差，

从峰顶到谷底来说，对 500 个观测值和对 100 个观测值差异不会很大。平均的 R/S 值在 100 个观测值之后停止增长。

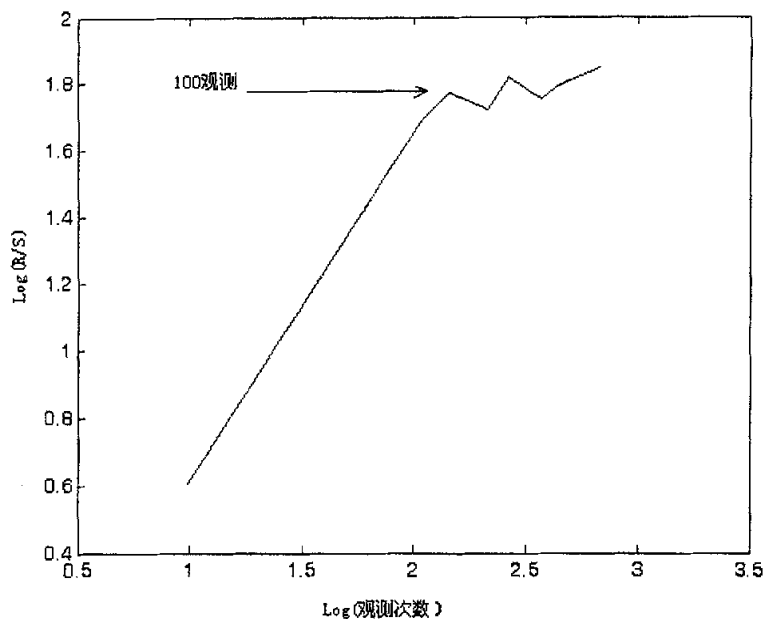


图 4.1

## 4.2 非周期循环

一个非周期循环没有绝对的频率。但它有一个平均频率。非周期循环可能有两种来源：  
1. 它们可能是统计循环，以持续（长期相关）和在方向上突然变化的赫斯特现象为例；  
2. 它们可能是一个非线性动态系统的结果，或决定性的混沌。

下面将概括的讨论一下这两个系统的不同之处。

### 4.2.1 统计循环

赫斯特过程是一个可以被描述为有偏的随机游走过程。不过，这个有偏在方向上和宏观上可以突然变化。这些有偏的突然变化由 Hurst 以他的概率扑克的王牌所模型化，给予了循环的显示。除了统计结构的稳定性，王牌的显露是一个随机事件。因为王牌的概率基于重置（Replacement），没有什么方法可预测王牌何时将现。假如经济循环是一个赫斯特过程，循环的持续性没有意义并且不是时间序列单独的产物。相反，王牌的出现归因于某

些外生的可能或不可能预测的事件。从这个眼光看，赫斯特“循环”没有平均长度，而且 Log/Log 的标绘图无限的规模变化。

## 4.2.2 混沌循环

非线性动态系统是一个能展示飘忽不定行为的确定性系统。当讨论混沌时，它通常考虑混沌映射，映射通常是一个迭代的微分方程系统，比如著名的逻辑斯蒂克方程 (Logistic Equation)：

$$X_t = a * X_{t-1} * (1 - X_{t-1}), \quad 0 < X < 1$$

这种类型的方程从统计上决定性的产生随机数。因此，作为市场和经济分析的工具，方程并不适用。迭代映射，像逻辑斯蒂克方程，展示了每一次迭代的混沌；那就是，它们的记忆长度是极短的。它们不能展示我们在经济学和投资中看到的循环类型。

相反，我们将研究混沌流，相依微分方程的连续系统。这种系统被用来建立大的生态系统（像气候）和热动力系统。这种类型的最有名的系统是洛伦茨吸引子 (Lorenz Attractor)。

一个简单的系统就是麦凯-格拉斯 (Mackey-Glass) 等式。它的基础假定是，当前的生产是根据过去的生产以及当前的测量而进行的。在生产与当前水平的测量之间的耽搁产生了一个涉及到延滞的“循环”。因为系统是非线性的，过量或不足量产生趋向被放大，导致非周期循环。非周期循环的平均长度却是非常靠近延滞的时间。麦凯-格拉斯等式的另一个性质是它是一个延滞微分方程：它具有有限数目的自由度，更像市场。延滞微分方程可以转变成如下的微分方程：

$$X_t = 0.9 * X_{t-1} + 0.2 * X_{t-n} \quad (4.1)$$

不规则的程度以及依赖时间滞后 n 的基础分形维，方程给变化的滞后及使用过的循环提供了方便。我们能够使用麦凯-格拉斯等式检验我们的假设：R/S 分析能够估计非周期循环的平均长度。

麦凯-格拉斯等式在等式 (4.1) 的描述是初始的延滞微分方程变为微分方程。在这个形式里，它能以容易的电子表格来模拟。起始具有 n=50 的滞后，步骤是：

1. 嵌入 0.10 在 A1 方格。在 A 栏向下复制 0.10 前 50 个方格。
2. 在方格 A51，敲入：0.9\*A50+0.2\*a1。
3. 向下复制方格 A51，共 8000 个方格。

图 4.2 表明了这个检验中所用的 8000 个观测值中的前 500 个。它们标记了不规则循环的长度是一个，非线性动态系统的典型。图 4.3 表明，当 n<50 时，H=0.93，当 >50 时，斜率几乎为零，表明极差已经达到最大，所以循环长度在逼近 50 的时候显示。



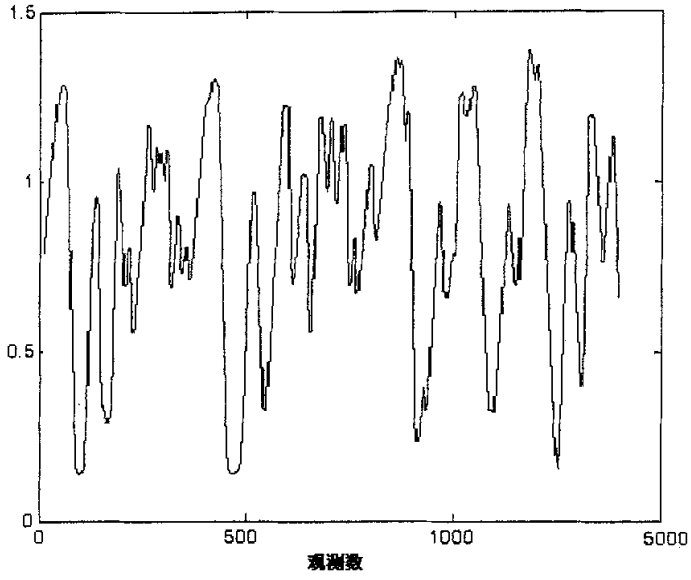


图 4.2 麦凯-格拉斯等式：观测值滞后=50

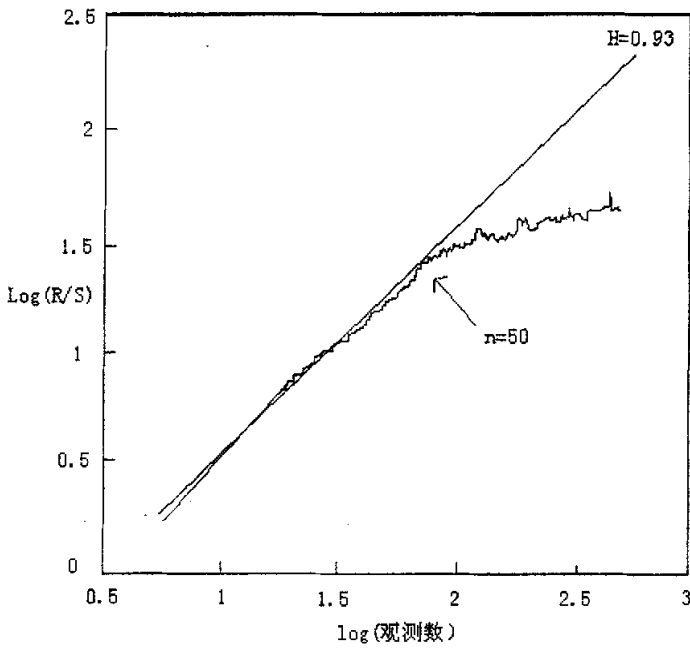


图 4.3 R/S 分析，麦凯-格拉斯等式：观测值滞后=50

### 4.2.3 添加噪声

R/S 能确定一个大  $H$  值的非周期循环的平均长度。在缺乏噪声的情况下，许多检验工作做得很好，可是，一旦一个小数量的噪声被添加上去，检验过程就会失效。因为 R/S 分析被用作测度系统里的噪声数量，我们可以预期 R/S 分析对于噪声的作用将会更稳健。

在动态系统里有两种类型的噪声。第一个叫做可观测的或可加性噪声。系统不被这些噪声所影响；相反，噪声是一个测度问题。观测者有一个精确测量系统输出的麻烦，以至于记录的值有一个添加的噪声增量。例如，假定你在研究龙头水滴之间的时间。你在桌子上放置一个测量仪器并在滴水现场放一个麦克风，记录精确的水滴击打底部的声音。遗憾的事，你是在一个许多人也在做实验的忙忙碌碌的实验室里。每次有人从身边走过时，你的测量仪器有点晃动，当水滴击打麦克风时它改变了时间。添加噪声对于过程是外生的。它是观察者的问题，不是系统的问题。

大多数人们在考虑噪声时，他们考虑添加的噪声。然而，第二种噪声的类型，叫做动态噪声，甚至可能更普通而且有更多的问题。当系统解释噪声输出为输入时，我们有动态噪声，因为噪声侵袭了系统。

现在，我们将处理噪声。图 4.4 显示了与图 4.2 所显示相同的点，具有一个标准差的添加噪声。

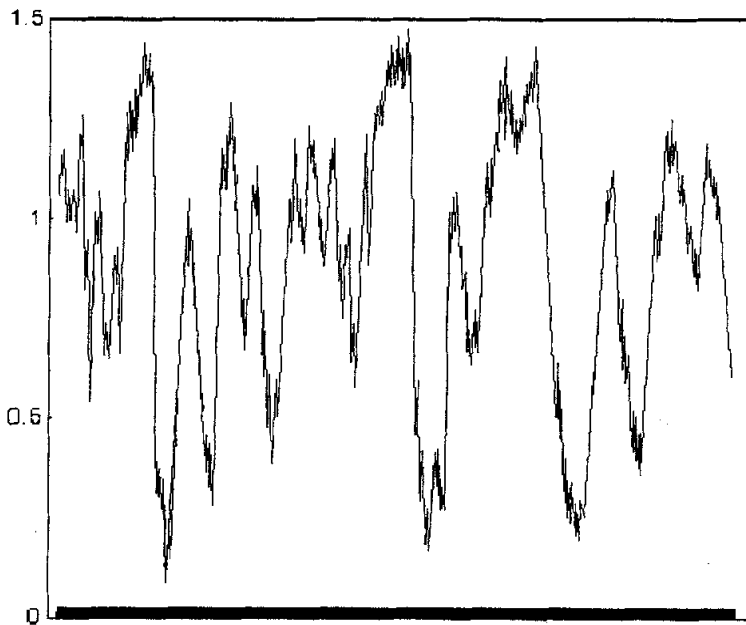


图 4.4 麦凯—格拉斯等式，添加的观测噪声

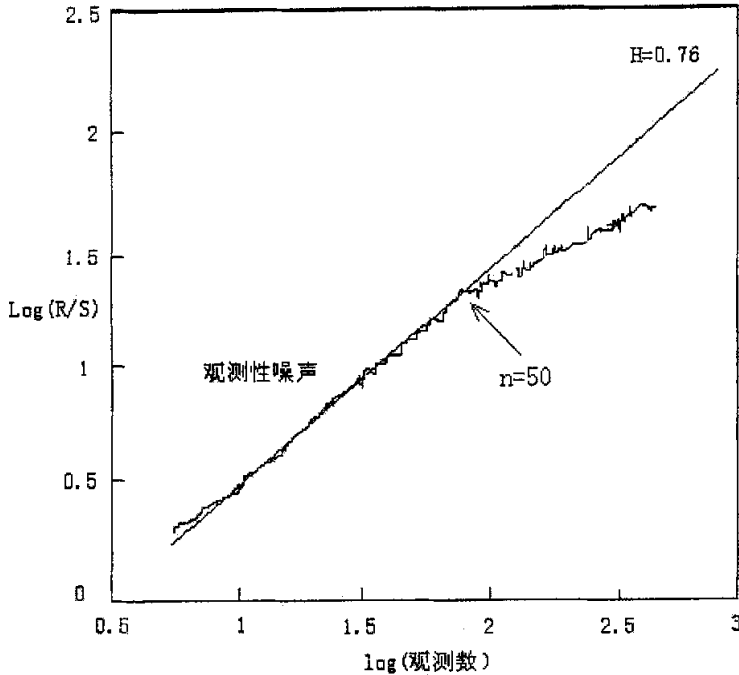


图 4.5 R/S 分析，具有观测性噪声的麦凯—格拉斯等式

由图 4.4 可以看出，时间序列看上去更像一个自然的时间序列。图 4.5 显示了 R/S 标绘图，这里， $H=0.76$ 。添加一个标准差的噪声已经降低了赫斯特指数，时间序列现在更是参差不齐，但是未受大量噪声的影响，循环长度在  $n=50$  的时候，仍然可以被估计。R/S 分析就噪声来说特别稳健。

### 4.3 非周期循环的估计方法

R/S 分析能探察到非周期循环，可以通过两种方法估计平均循环长度：

一种是通过绘制  $\text{Log}((R/S)_n)/\text{Log}(n)$  图来估计。对于 R/S 分析的 Log/Log 图中，每一个循环的结尾，也就是下一个循环的开始，通过图中的转折点，可以估计平均循环长度。

另一种是 V 统计方法。它给出了一个更精确的循环长度的度量。它可以表示为：

$$V_n = (R/S)_n / \sqrt{n} \quad (4.2)$$

如果过程是独立的，随机的，( $H=0.50$ ) 联系公式 (3.7)，针对  $\text{Log}(n)$  绘制的 V 统计量

应该是平的。如果过程是反持续的，R/S 以慢于时间方根 ( $H < 0.50$ ) 的比率规模变化，那么图形应该是向下倾斜的。如果过程是持续的，R/S 以快于时间方根 ( $H > 0.50$ ) 的比率规模变化，那么图形应该是向上倾斜的。通过 V 统计图，可以发现断点，检查在每一区间的 V 的最大值，从而估计每一频率的循环长度。

#### 4.4 消除序列线性依赖的方法论

我们将就资本市场对数收益的 AR(1) 提出以下的方法。AR(1) 残差被用来消除，或至少最小化线性依赖。线性依赖可能偏离赫斯特指数（当没有长期记忆过程存在时，可能使这一点看起来显著），或是一个 1 型错误。由于使用 AR(1) 残差，我们最小化有偏，而且我们希望减少那些无关紧要的结果。过程通常叫做“预先白化”或“从属倾向”。从属倾向不适用于所有的统计检验，作为一些检验，从属倾向可能掩盖显著性信息。而且，在 R/S 分析的情况下，从属倾向将消除序列相关或短期记忆，以及通货膨胀增长。前者是一个具有非常高的频率数据的问题，诸如 5 分钟受益。后者是一个低频率数据问题，诸如 60 年的月收益。这样，对于 R/S 分析，短期记忆过程比通货膨胀问题有更多的问题。

我们开始于一个对数收益序列：

$$S_t = \text{Log}(P_t / P_{t-1}) \quad (4.3)$$

这里， $S_t$  = 在时间 t 的对数收益； $P_t$  = 在时间 t 的价格。

那么，针对于独立变量  $S_{t-1}$ ，我们回归依赖变量  $S_t$ ，得到截距 a，斜率 b。  $S_t$  的 AR(1)

残差减掉了  $S_t$  对  $S_{t-1}$  的依赖性：

$$X_t = S_t - (a + b * S_{t-1}) \quad (4.4)$$

AR(1) 的残差方法不能消除所有的线性依赖。然而，它消除了足够的依赖，降低了非显著性水平的影响。而不论 AR 过程是在水平 2 还是水平 3 上。

#### 4.5 数据抽样中的问题

大量的数据不是 100% 所要求的。我们将计算不同时间起点的收益，以观察 R/S 行为是否依赖所用的时间增量而变化。对于在不同区间抽样时间序列样本的这一数量。对于长的序列，我们能够调查“过量抽样”是否导致有偏。

我们能够预期大量当我们改变抽样区间时所发生的事：

1. 当我们增加抽样区间时，赫斯特指数能够被预期增加。在较短的区间或较高的频率，

在数据里对于较多的噪声有一边界。低频率抽样将最小化噪声的影响，并且消除了可能存在于较高频率的任何分数式的噪声。如在魏尔斯特拉斯函数所见，高频率循环的相加使得时间序列更参差不齐，以至于减少赫斯特指数（或增加分形维）。较低频率的抽样“略过”了较高的频率。

2. 存在于较长区间的任何“循环”应该保留。假如一个循环显露在 1 日的 1000 个区间，那么，在 10 日的 100 个区间也应该显露。

3. 假如过程是一个高斯随机行走，前两点将不成立。白噪声显示在相同的所有频率中（就像我们在录音机中听到的“嘶嘶”声，它好像在所有速度都是一样的）。而且，没有循环。假如 R/S 图在逐日区间显示一个断裂点可在 10 日区间没有出现，那么，逐日区间图的断裂点是人工制品，而非真正的循环。

## 第 5 章 使用 R/S 分析进行实例分析

### 5.1 Matlab 简介<sup>[33]</sup>

1984 年美国的 Mathworks 公司推出 Matlab<sup>[33]</sup>，到目前为止，它已经发展成为国际上最优秀的科技应用软件之一。其强大的科学计算和可视化功能、简单易用的开放式可扩展环境以及多达 30 多个面向不同领域而扩展的工具箱 (Toolbox) 支持，使得 Matlab 在许多科学领域中成为计算机辅助设计与分析、算法研究和应用开发的基本工具和首选平台。现在 Matlab 已经发展到了 Matlab7.1 版本。

Matlab 最初用于自动控制系统的辅助设计，而后采用了开放性开发的思想，不断吸收各学科领域权威人士所编写的实用程序，形成了一系列规模庞大、覆盖面极广的工具箱 (Toolbox)。所谓工具箱，其实是一组一组的函数。主要包括信号处理工具箱 (Signal Processing Toolbox)，图像处理工具箱 (Image Processing Toolbox)，通信工具箱 (Communication Toolbox)，定点运算工具箱 (Fixed-Point Blockset Toolbox)，小波分析工具箱 (Wavelet Toolbox)，高阶谱分析工具箱 (High-Order Spectral Analysis Toolbox)，地理信息处理工具箱 (Mapping Toolbox)，金融工具箱 (Financial Toolbox)，金融时间序列工具箱 (Financial Time Series Toolbox)，金融衍生工具工具箱 (Financial Derivatives Toolbox)，统计工具箱 (Statistic Toolbox)，GARCH 工具箱 (GARCH Toolbox)，偏微分方程工具箱 (Partial Differential Equation Toolbox)，线性矩阵不等式控制工具箱 (LMI Control Toolbox)，符号数学工具箱 (Symbolic Math Toolbox)，样条工具箱 (Spline Toolbox)，优化工具箱 (Optimization Toolbox)，系统识别工具箱 (System Identification Toolbox)，控制系统工具箱 (Control System Toolbox)，鲁棒控制工具箱 (Robust Control Toolbox)，模型预测控制工具箱 (Model Predictive Toolbox)，模糊逻辑工具箱 (Fuzzy Logic Toolbox)， $\mu$  分析与综合工具箱 ( $\mu$ -Analysis and Synthesis Toolbox) 等等。

目前，我国金融风险防范手段、金融工具和信息技术还相对落后，究其原因主要有三：首先，一般金融软件公司无法获得财政部、人民银行、证监会、银监会、保监会和国家计委等国家单位的信息资源，因而只能片面追求金融软件的指标数量、表现形式和功能；其次，国内金融理论界长期以来重文轻理，与国际金融实务脱轨，往往“知其然，不知其所以然”，尤其是缺乏对现代金融理论、数学模型、算法程序和信息技术方面的系统研究和把握；第三，由于国内信息技术人员的流动性和软件工程管理的差异性，金融软件开发商很难具备构建国家金融风险防范体系的具体设计和开发能力。工欲善其事，必先利其器。现时国外许多大型银行如大通曼哈顿银行、美洲银行等和非银行金融机构如高盛、美林、摩根斯坦利等投资银行，都拥有相当数量的、高水平的金融工程师，他们利用大型计算机和先进的通讯设备，开发各种新型金融产品和风险管理技术，对于客户的特殊需求诸如金融产品的流动性和收益、风险特性等，已经发展到“量体裁衣”的程度。与国外相比，目前国内金融行业的资金实力、知识素质、管理水平和科技水准尚不足与国际对手抗衡。金融软件系统的核心技术的数学模型与复杂算法不易被竞争对手模仿，所以保密性强。

Matlab 数学软件，以精确的运算能力，高品质的可视化功能，高度的开放性与延展性

著称，提供使用者一个整合的互动式金融工程运算环境。Matlab 在金融领域的应用是 MathWorks 公司业务增长最快的模块，国外的应用已经相当普及。在国内，却只有少数几个单位在用。因此，Matlab 的金融数学模块在中国具有非常巨大的发展潜力和应用前景。

## 5.2 对上证综指的收益率进行实际的分析<sup>[34~47]</sup>

对 Hurst 指数的计算没有固定的统计软件能够完成，本文通过使用 Matlab 和 Excel 软件计算了 Hurst 指数。

本研究所使用的数据是上证综合指数，时间为 1999 年 1 月 1 日至 2005 年 12 月 31 日，选取其中的 1637 个日收盘价和 313 个周收盘价。数据来源于证券网上的历史数据。

在估计和回归过程中，我们使用对数收益率。因为对于 R/S 分析而言，对数收益率比价格百分比变化更为适用，R/S 分析中的极差是对于平均值的累积横距，对数收益率加起来等于累积收益率，而百分比变化却不是。

具体估计步骤如下：

1. 计算对数收益率

$$S_t = \ln(P_t / P_{t-1})$$

其中， $P_t$  = 股票 t 时的价格， $S_t$  = 股票 t 时的对数收益率。

2. 对  $S_t$  进行 AR(1) 回归，消除序列  $S_t$  序列的线性相关得残差序列：

$$x_t = S_t - (a + bS_t)$$

其中，a 和 b 是 AR(1) 的系数， $x_t$  为残差。

对数收益的 AR(1) 残差能消除或降低线性依赖，而线性依赖可能偏离 Hurst 指数。通过取 AR(1) 的残差，可以降低偏离程度并降低结果的不显著性程度。同时，对于 R/S 分析，消去趋势法可以消去序列相关或短期记忆，但该方法并非适用所有统计检验，因为对于某些统计检验，消去趋势法可能会消去了某些重要的信息。

3. 将长度为 M 的残差序列  $\{x_t\}$  分成 A 个长度为 N 的相邻子区间， $AN=M$ ，利用等式 (3.14) 和 (3.15) 分别计算每个子区间  $I_a$  的极差  $R_{I_a}$  和标准差  $S_{I_a}$  ( $a=1,2,\dots,A$ )。

4. 利用等式 (3.16) 来计算  $(R/S)_n$  序列，并根据计算结果利用等式 (5.2) 计算 V 统计量。计算出的 V 统计量用于测量序列“长期记忆”的长度，也就是非周期循环的长度。按照混沌理论，在任何非线性系统中都有一点，到这一点，初始条件的“记忆”就会消失。

### 5.2.1 检验 Hurst 指数的有效性

检验 Hurst 指数的有效性的一种最简单的方法是随机的打乱数据。由于实际观测还在那里，其频数分布保持不变。对打乱后的数据计算 Hurst 指数，如果序列是一个真正独立的序列，观测之间没有长期记忆效应或相关性，那么，Hurst 指数应该保持不变，打乱数据应该对于数据的定性没有影响。反之，如果有长期记忆效应，那么，观测数据的次序就是重要的，打乱数据将破坏系统的结构。计算的 Hurst 指数应有很大差别或接近于 0.50。

取上证综指 1999 年 1 月 1 日到 2005 年 12 月 31 日的日收盘价 1637 个数据样本，再取 1999 年 1 月 1 日到 2005 年 12 月 31 日的周收盘价 313 数据样本，把两组数据通过随机打乱以后重新进行 R/S 分析，计算 Hurst 指数。结果显示于表 5.1。

指数	重排前 Hurst 指数	重排后 Hurst 指数
日收益	0.640534	0.51178
周收益	0.653974	0.52159

表 5.1

结论：

(1) 通过对两组数据进行 R/S 分析，发现上证综指的日收益率和周收益率的赫斯特指数分别为 0.64 和 0.70，均 $>0.50$ ，这说明上证综指具有长期记忆性。打乱数据后，两组数据的赫斯特指数都发生了明显的减小，趋向于 0.50，这说明这两组数据不是真正的独立序列，数据之间的次序很重要，打乱了数据的次序就破坏了系统的结构。

(2) 上证综合指数日收益率和周收益率的赫斯特指数都明显偏离 0.50，上证综指服从分形布朗运动或是有偏的随机游走过程，变量之间不是相互独立，而是正相关。市场表现出趋势增强行为，而不是均值回复行为。其原因是信息以非线性的方式呈现，同时人们对信息的消化与确认也是不均等的。他们以非线性的方式对信息作出反应，所有这些最终都将通过市场交易活动反映在股票价格上，使得股票价格呈现为有偏的随机游走。

这一结论与有效市场假说的描述不符，EMH 假定所有信息均能及时充分的反映到股票价格中，投资者接受信息，并通过价格的变化来反映新的信息。但实际的信息分布是陡峰厚尾的，人们也是以非线性的方式对信息作出反应，即只有当信息水平达到某个临界值时，人们才对所有过去他们曾忽视的信息作出反应，这就意味着当前的价格受过去历史信息的影响。



## 5.2.2 中国股市的自相似性研究

取上证综指 1999 年 1 月 1 日到 2005 年 12 月 31 日的日收盘价 1637 个数据样本，再取 1999 年 1 月 1 日到 2005 年 12 月 31 日的周收盘价 313 个数据样本，进行 R/S 分析。对日收益收益数据的分析结果见表 5.2，图 5.1，图 5.2。

N	Ln(N)	R/S	Ln(R/S)	V(R/S)
4	1.386294	1.678527	0.517917	0.839264
8	2.079441	2.526730	0.926783	0.893333
10	2.302585	2.968591	1.088087	0.938751
20	2.995732	4.596847	1.525371	1.027886
40	3.688879	7.315284	1.989966	1.156648
80	4.382027	11.18276	2.414373	1.250271
120	4.787492	14.36269	2.664703	1.311128
150	5.010635	16.39257	2.796839	1.338448
200	5.298317	20.12926	3.002174	1.423354
250	5.521461	23.86547	3.172433	1.509384
270	5.598422	25.98969	3.257700	1.581682
302	5.703782	29.55363	3.386207	1.706280
303	5.736572	29.41897	3.381640	1.670875
400	5.991465	26.79397	3.288177	1.339690
800	6.684612	40.41393	3.699175	1.428848
1000	6.907755	41.56813	3.723333	1.314499
1300	7.170119	45.91576	3.826808	1.273474
1500	7.313220	51.00235	3.931875	1.316875

表 5.2

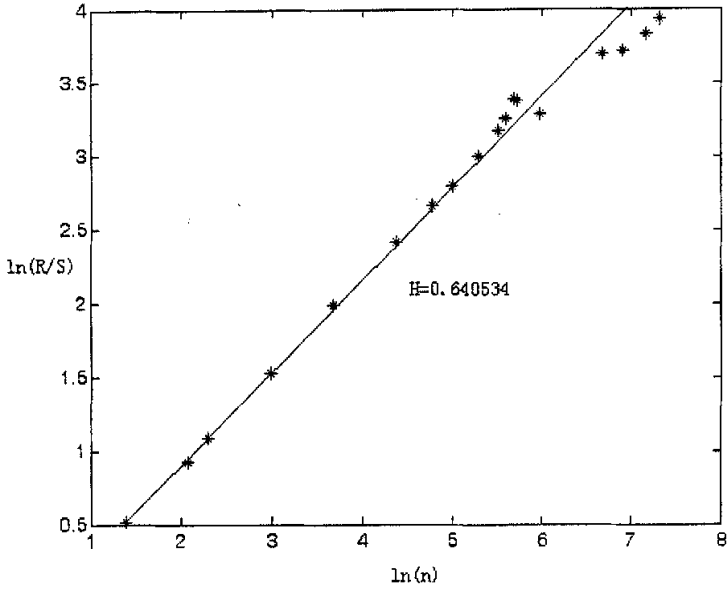


图 5.1 日收益的 R/S 分析

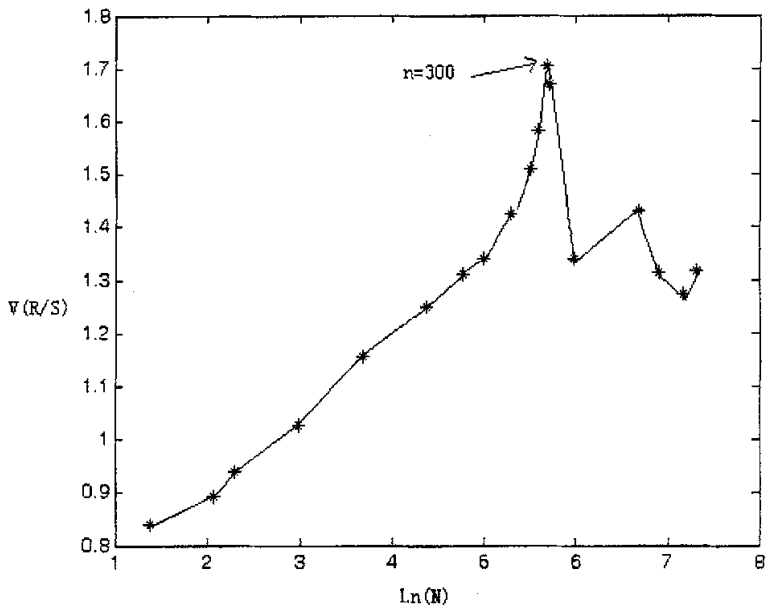


图 5.2 日收益的 V 统计

对周收益数据的分析结果显示于表 5.3，图 5.3，图 5.4。

N	Ln(N)	R/S	Ln(R/S)	V(R/S)
2	0.693147	2.986593	1.094133	2.112159
4	1.386294	4.568474	1.519182	2.284237
8	2.079442	7.318419	1.990394	2.587452
10	2.302585	9.218631	2.470327	2.915188
24	3.178054	14.62953	2.683042	2.986241
30	3.401197	16.92786	2.828961	3.090590
40	3.688879	21.29643	3.058539	3.367261
50	3.912023	24.68571	3.206224	3.491086
54	3.988984	26.99692	3.283424	3.673816
60	4.090345	30.63138	3.422025	3.954495
61	4.110874	30.18750	3.407428	3.865114
80	4.383027	27.79974	3.325027	3.108105
160	5.075174	41.13938	3.716966	3.252346
200	5.298317	42.81353	3.756854	3.027374
260	5.560682	46.15768	3.832063	2.862578
300	5.703782	53.53846	3.980400	3.091044

表 5.3

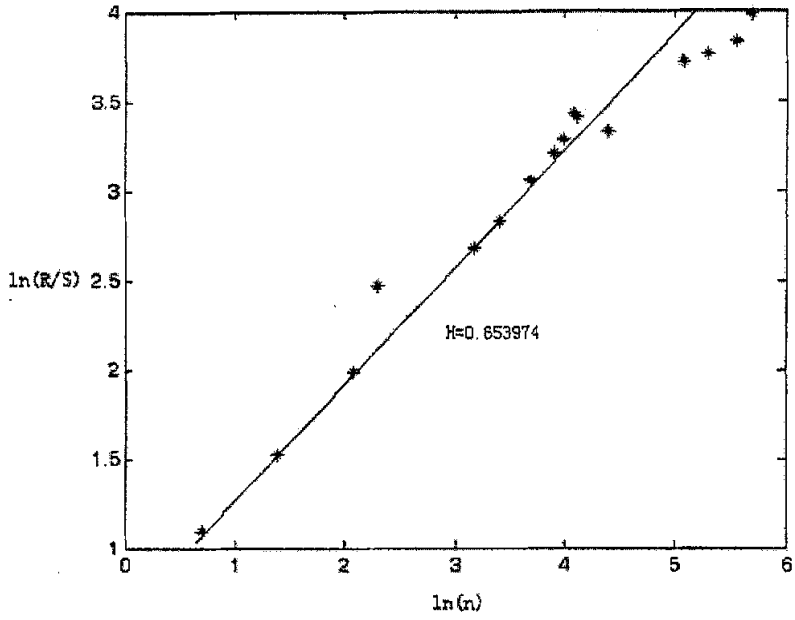


图 5.3 周收益的 R/S 分析

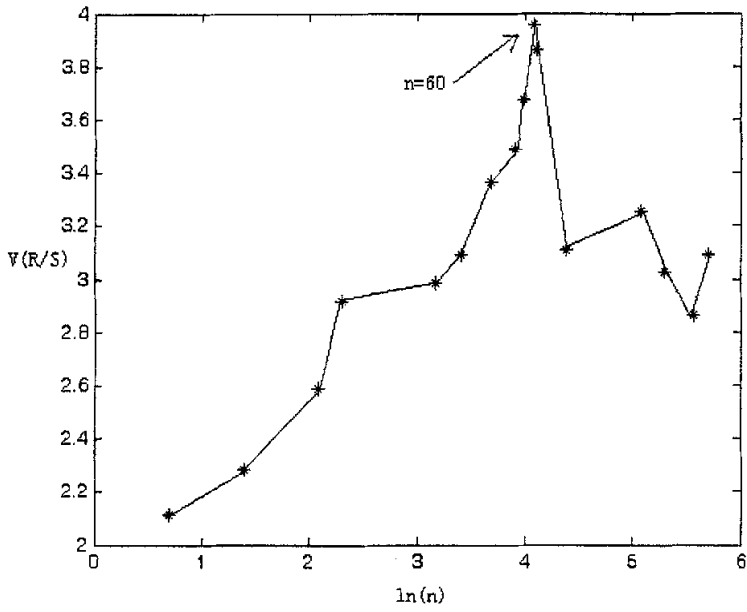


图 5.4 周收益的 V 统计

结论：中国股票市场存在明显的自相似性，具体表现在两个方面：

(1) 通过日收益率和周收益率计算得出的赫斯特指数值非常接近，这说明我国股票市场的确具有标度不变性。

(2) 中国股票市场的自相似性还表现在通过不同时间标度计算出的周期非常接近。在股票市场，每周有 5 日开盘，为了避免周末的影响，我们以 5 天为一周。日收益的循环断点发生在第 302 天，根据周收益，发现周收益的循环断点发生在第 60 周，也就是  $60 \times 5 = 300$  天左右，两者产生了很好的吻合。

### 5.3.3 分析数据的非周期循环

现在采用中国股市比较有代表性的上证综指 1999 年 1 月 1 日到 2005 年 12 月 31 日的 1637 个数据，运用 R/S 分析来对中国股市进行分析研究。对数据进行分段回归，发现出现了循环断点，表 (5.4) 显示了分段回归的结果：

表 5.4 数据分段回归结果

数据类别	区间 1	赫斯特指数 1	区间 2	赫斯特指数 2
上证综指单日收益率	(20, 300)	0.640534	(300, 1550)	0.493607
上证综指 5 日收益率	(20, 300)	0.653974	(300, 1500)	0.531743
上证综指 20 日收益率	(20, 300)	0.820325	(300, 1550)	0.573642
上证综指 30 日收益率	(20, 300)	0.889732	(300, 1500)	0.602814
上证综指 90 日收益率	(20, 1200)	0.853673	(20, 1550)	0.820327
上证综指 180 日收益率	(20, 1200)	0.964158	(20, 1500)	0.943845

图 (5.5) 至 图 (5.10) 显示了不同时段收益率的赫斯特图形。

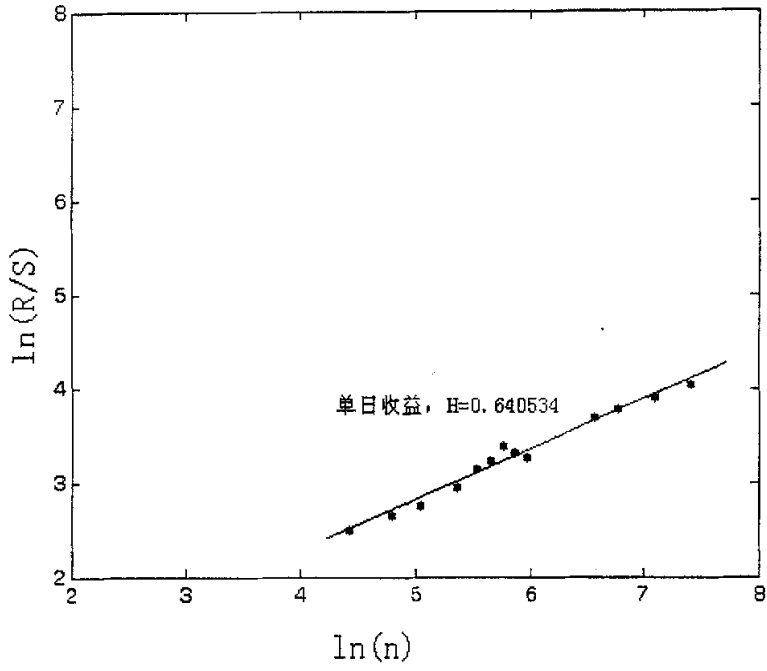


图 5.5 上证综指单日收益率

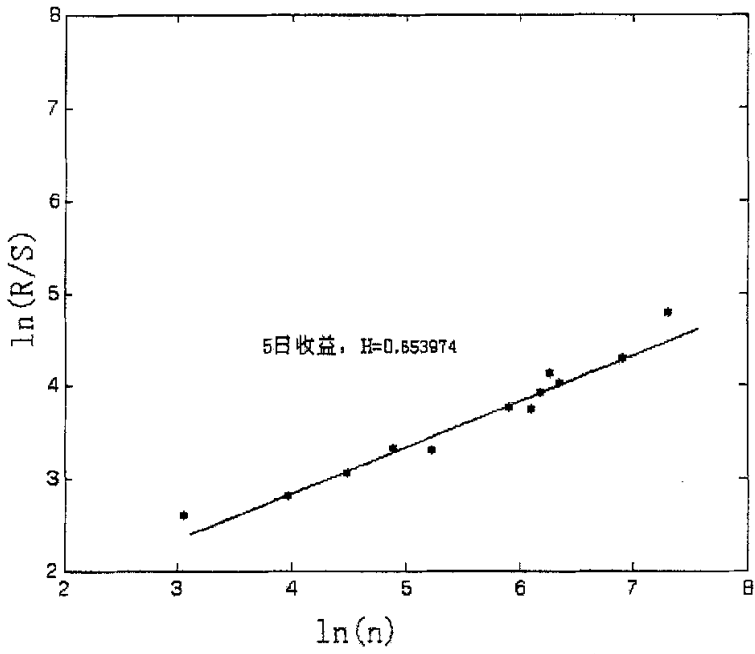


图 5.6 上证综指 5 日收益率

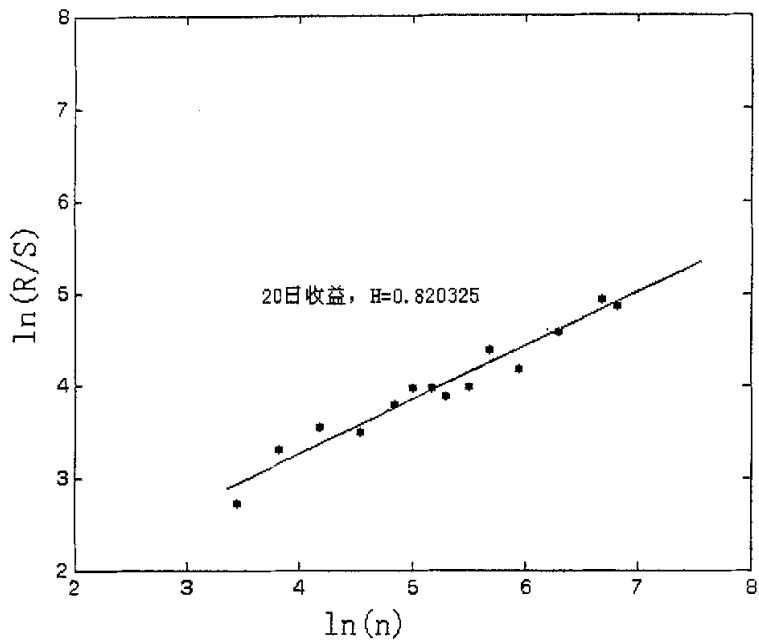


图 5.7 上证综指 20 日收益率

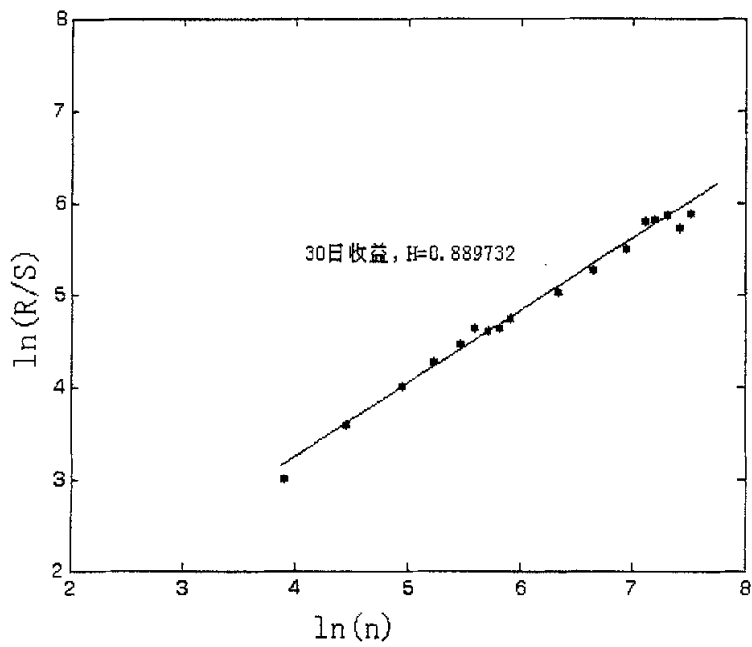


图 5.8 上证综指 30 日收益率

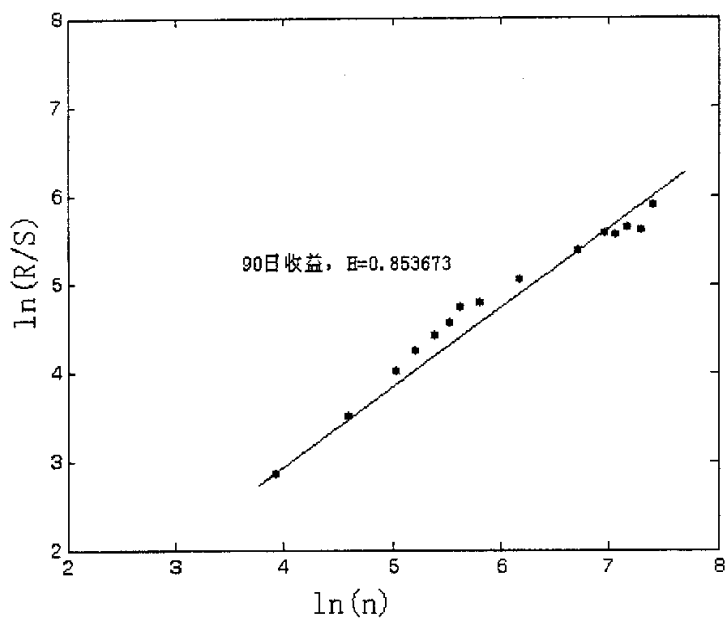


图 5.9 上证综指 90 日收益率

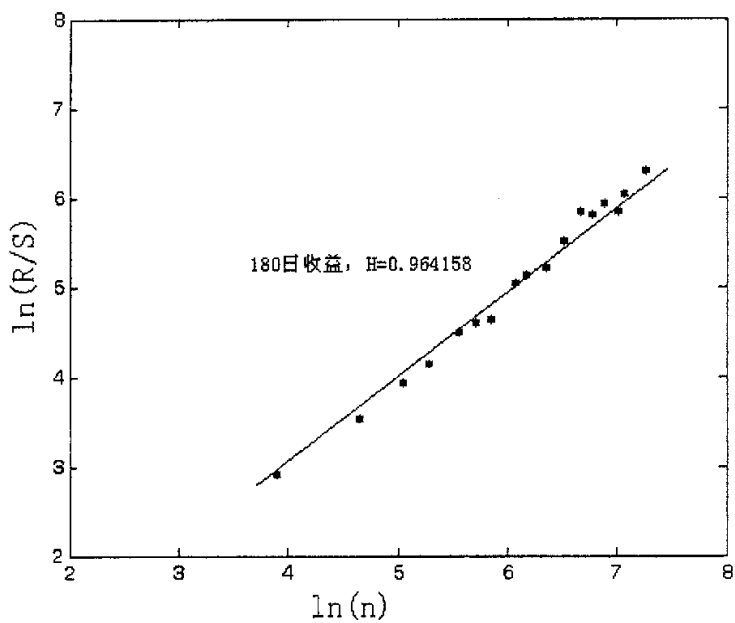


图 5.10 上证综指 180 日收益率



表 5.5 显示了对上证综指的循环周期的分析结果

数据类别	赫斯特指数	循环周期
上证综指单日收益率	0.640534	300 个交易日左右
上证综指 5 日收益率	0.653974	300 个交易日左右
上证综指 20 日收益率	0.820325	300 个交易日左右
上证综指 30 日收益率	0.889732	300 个交易日左右
上证综指 90 日收益率	0.853673	1200 个交易日左右
上证综指 180 日收益率	0.964158	1200 个交易日左右

表 5.5 上证综指指数波动及循环周期

结论:

1、中国股市存在分形结构,也存在长期记忆。上证综指在不同的时间增量收益率上  $H$  都大于 0.5,所以分形结构存在是明显的。也就是说中国股市不完全是一个随机游走的过程,而是一个有偏的随机过程,是一个趋势加上一个噪声,以前的事将影响到现在和将来。

2、中国股市在不同时间增量收益率上的稳定性是不同的,这种稳定性随着时间增量的增加而增加。从单日到 5 日、20 日、30 日、90 日、180 日,上证综指的收益率的赫斯特指数在不断增加,从 0.640534 增加到 0.964158,这显示了随着时间增量的增加,指数收益率的稳定性在增加,噪声越来越少。反映出在中国股市上,短期内投机成分占主导地位,但是在长期内,股市总要回归,经济基本面的成分将最终主导股市波动。同时应验了分形结构的短期随机、长期确定的性质,体现了随机性与确定性的统一。

3、可以看到,对于上证综指,随着区间的扩展, $H$  在逐渐减小,这显示了一个  $H$  不断下降的过程。不难推测,这个扩展的最终结果应该是收敛到 0.5,也就是直到过去对未来之存在一个随机影响,也就是不确定性影响。这说明股市对于过去的记忆是存在的,但是却是有限的,过去可以影响到未来,但时不时无限期的影响下去。过去对未来的作用在逐渐减小,直至减小到零。

## 第 6 章 R/S 分析应用于分形噪声与混沌噪声 [34~47]

### 6.1 分形噪声与 R/S 分析

#### 6.1.1 噪声的色彩

人们一想到噪声，总是认为噪声是“白色”的或是随机的。这类噪声是在听空白磁带时的“嘶嘶”声。因为它没有固有的规模变化，所以不管磁带的速度如何，嘶嘶声听起来总是相通的。它的被积函数被称作“布朗”噪声。布朗噪声就是白噪声的运行总和。看起来似乎有某种东西在那里，但是，没有什么信息实际存在于布朗噪声里。

这些噪声能够被它们的功率谱所特征化。它们遵循简单的逆幂法则。功率谱是通过傅立叶变换计算出来的，它通常叫做谱分析。傅立叶变换一个时间序列成为由频率所定义的函数。它假定任何时间序列可由不同频率和无限区间的正弦或余弦波的总和来代表。在许多频率或时间增量上，傅立叶函数的系数以与光有谱的相同的方式定义了“谱”。在有尖陡峰的频率上，在原始的时间序列，存在一个周期性要素。这样，谱分析假定：(1) 研究中的时间序列一般是周期性的。(2) 而且循环实际上也是周期性的。

然而，当分形是现实时，功率谱是无特色的并且依逆幂法则而规模变化。这些逆幂法则是一个频率函数  $f$ ，并遵循  $f^{-b}$  的形式。由于在研究中的系统的自相似本质，功率谱遵循逆幂法则。这些规模变化因子，或谱指数  $b$ ，其变化范围可以由 0 到 4。对于白噪声， $b=0$ ，即白噪声功率谱与频率无关。在所有的频率上，白噪声保持不变，那就是磁带上的嘶嘶声为什么在任何速度（频率）下听起来都一样。在相空间里白噪声分行维的计算是相似的。白噪声充满了它置身其中的嵌入维（它在某种情形时是一个频率）。这里不存在规模变化法则。当白噪声被积分时，对于布朗噪声的幂谱， $b=2$ 。因而，布朗噪声有  $1/f^2$  的形式。

正如在大多数随机过程中，这个规模变化算子是一个平方。

对于  $b$  来说，还存在一些其他的值。例如  $0 < b < 2$ ，我们有一个粉红噪声。粉红噪声常常被认作  $1/f$  噪声。粉红噪声似乎一般是广泛分布，而且在模拟湍流时有用，特别当  $b$  假定是 1 与 2 之间的分数值时。在布朗噪声以外，当  $b > 2$  时，存在黑噪声。黑噪声一只用来模拟持续系统，它们以存在突然崩塌而著名。现在有一个分数维噪声与赫斯特过程之间的关系：
$$b = 2 * H + 1 \quad (6.1)$$

这里， $b$  = 谱指数； $H$  = 赫斯特指数。

黑噪声涉及到长期记忆效果 ( $H > 0.50, 2.00 < b \leq 4.00$ )；粉红噪声涉及反持续性 ( $H < 0.50, 1 \geq b > 2$ )。

## 6.1.2 粉红噪声

长期以来,  $0 < H < 0.50$  一直被认为是“不怎么令人感兴趣”的情形。然而, 事实并非如此, 反持续性可能相当重要。易变性与湍流之间的关系将大大提高我们对市场理解的兴趣。它还将减少大量的对于形而下系统和市场之间的误解。

等式(6.1)显示了湍流与赫斯特指数之间的联系。H 的反持续性与粉红噪声相符。理解粉红将提高对反持续性和易变性的理解。

### 1. 松弛过程

$1/f$  噪声与松弛过程密切相关。实际上,  $1/f$  造声势在许多不同频率发生的大量平行松弛过程的集合。

在分形市场的假说理论下, 具有不同投资起点的不同的投资者, 更有可能对具有多重松弛时间的信息有所反应: 那就是, 信息对于不同的投资者影响也不同, 而这些信息依赖于它们的投资起点。

模拟  $1/f$  噪声的公式是:

$$x_{n+1} = \rho * x_n + \sqrt{1 - \rho^2} * r_n \quad (6.2)$$

这里,  $x_0 = 0$ ;  $r =$  唯一随机数;  $\rho =$  确定的时间关系。

$\rho$  与松弛事件相关,  $t$  遵循下列关系:

$$\rho = \exp(-1/t) \quad (6.3)$$

这里,  $t$  是松弛时间。如果选择  $t$  的三个值, 它们均匀地散布在  $\log$  空间里, 就会产生三个序列  $x$ 。

由于分形市场假说假定的市场结构, 多重平行松弛结构就有可能存在。每个投资起点(或频率)有其自身的概率结构。这个自相似的概率结构意味着, 在短期内的规模调整之后, 每个投资起点面临着相同的风险水平。每个投资起点都有相同的不稳定的易变性结构。

这些不稳定易变性的和, 是具有  $b=1.56$ , 或  $H=0.44$  特征指数  $1/f$  噪声。

### 2. 间歇性

在实际中,  $b = 1.67$  或  $H = 0.33$  的特征值经常出现。Kolmogorov (1941) 预测, 湍流速度的改变, 应该是  $b = 5/3$ 。由 Kida (1991) 和 Schmitt (1992) 已表示过, 湍流的实际指数非常接近其预测值。H 的持续值逐渐趋向 0.70, 反持续性值趋向 0.33。这暗示, 湍流和市场结构之间可能存在一种关系。大部分人把湍流等同于股票市场, 他们考虑的是价格的变化, 相反, 湍流可能较好的模拟易变性, 这易变性可以通过期权市场而被买卖。

湍流被认为是小瀑布现象。它被一个能量特征化，这个能量正由较大规模结构转换为较小规模结构。这个自相似性小瀑布结构，是头一批动态分形映像之一。市场更像逆幂法则现象，大规模变化是小规模变化的总和（放大过程）。这个放大过程构筑了长期记忆过程的基础。在易变性里，可能是这样的：

1、在股票、债券及通货市场，易变性以快于时间方根的速度增加。一个投资起点对于另外一个投资起点的关系，放大了较小投资起点的效果，可能是易变性具有幂指数规模变化特征的动态学原因。每一次市场的分形结构（那就是，具有不同投资起点的同时交易的投资者）是一个放大过程的快照。

2、股票与债券市场的确有最大规模变化，表现为记忆效果就如同湍流中的能量一样驱散。然而，通货没有这个特征与能量放大，而且记忆也不会永远继续。对于湍流具有相似值  $b$  的易变性，应是这样被模式化了的。

### 6.1.3 黑噪声

赫斯特过程，大致是讲一个黑噪声的过程。如同粉红噪声，黑噪声过程在自然界里似乎很丰富。像湍流一样，粉红噪声在松弛过程中出现。就像江河的流量，太阳的黑子，树的年轮厚度，以及股票市场价格的变化，黑噪声在长期运行的循环记录中出现。赫斯特过程可能是对出现黑噪声的一种解释，但是，有另外的理由解释在时间序列里存在着持续性。下面将考察分形布朗运动。

#### 1. 约瑟效应

分形布朗运动（FBM）是布朗运动的一般化，而布朗运动长期被当作一个“违约”的发散过程。布朗运动最初是作为漂浮在液体中的微小粒子的无规则运动而进行研究的，这个不规则运动是液体自身的性质，不规则运动归咎于水分子同悬浮微粒的碰撞。下面将解释这个过程。

取  $X(t)$  为随机分子在时间  $t$  的位置，让  $(e)$  作为具有 0 均值和单位方差的高斯随机过程，包含着以  $e$  为标志的随机数。随机分子从时间  $t_0$  到时间  $t$  的位置改变由下式给出：

$$X(t) - X(t_0) \approx e * |t - t_0|^H, t > t_0 \quad (6.4)$$

这里，对于布朗运动， $H = 0.50$ 。

对于分形布朗运动，我们将  $H$  一般化以便它能从 0 变化到 1。如果我们现在将  $B_H(t)$  定为 FBM 中粒子的位置，那么，位置变化方差以如下时间进行规模变化：

$$V(t - t_0) \approx |t - t_0|^{2 * H} \quad (6.5)$$

对于  $H = 0.50$ ，这降到了经典的高斯情形。方差随时间呈线性增长，或标准差以时间

的方根增长。然而，分形布朗运动 FBM，具有较快于布朗运动方差变化速度的规模变化方差。此时， $0.5 < H < 1$ 。根据等式 (6.3)，标准差应该以等同于  $H$  的速率增长。

在增量之间的相关关系  $C(t)$ ，定义如下：

$$C(t) = 2^{(2 \cdot H - 1)} - 1 \quad (6.6)$$

这个等式表达了经过  $t$  时间的过程，且具有或先或后的  $t$  时间所有增量，位置变化的相关关系。用市场的话来讲，它将是具有将来和过去日收益的所有日受益。它也可以运用于具有过去和将来 5 日收益的所有 5 日收益。从理论上讲，它可应用于所有的时间增量。它是一个长期记忆结果强度的测度，并且它覆盖了所有时间的规模变化。

当过程是一个布朗运动时， $H=0.50$ ，而且  $C(t)$  为零。没有长期记忆效果。当  $0 < H < 0.50$  时， $C(t)$  为负。存在反转性效果，并就多重时间规模变化出现。在反持续性，粉红噪声过程看到了相似的效果。然而，当过程是黑噪声时， $0.5 < H < 1.0$  我们有无数的长期运行的相关关系；那就是，我们有一个发生在多重时间规模变化，或在资本市场的投资起点里的长期记忆效果。

定义 FBM 的等式使用这个无限记忆的效果：

$$B_H(t) = [1/\Gamma(H + 0.50)] * \left[ \int_{-\infty}^0 (|t - t'|)^{H-0.50} - |t'|^{H-0.50} \right) dB(t') + \int_0^t (|t - t'|)^{H-0.50} dB(t') \right] \quad (6.7)$$

当  $H=0.50$ ，等式 (6.7) 降至最初的布朗运动。等式 (6.7) 可以简化变成一个比较容易理解的形式。

$$B_H(t) - B_H(t-1) = [n^{-H} / \Gamma(H + 0.50)] * \left[ \sum_{i=1}^{n+1} i^{H-0.50} * r_{(1+n)(M+i)-i} + \sum_{i=1}^{n+(M-1)} (n+1)^{H-0.50} - i^{H-0.50} * r_{1+n(M-1+i)-i} \right] \quad (6.8)$$

这里， $r$  是一个  $M$  个高斯随机变量的序列。等式 (6.8) 是等式 (6.7) 的离散形式。本质上，它们说的是同一件事，用求和代替积分。

在这个基本形式里，黑噪声序列的时间序列（或“时间图样”）变成比较广华，有较高的  $H$  或  $b$ 。就模拟而言，光滑是移动平均的产物。在理论上，它由观测之中增加的相关关系所引起。长期记忆效果导致趋势和循环的出现。

## 2. 诺亚效应

正如图 6.1 所示，等式 (6.8) 产生了一个具有适合的  $H$  值或恰当数量的无规则形状的时间图样；那就是它复制了时间图样的分形维、约瑟效应或长期记忆效果。黑噪声具有附加的效果：灾变。等式 (6.7) 和 (6.8) 因其是分形高斯噪声不能推力灾变。他们仅仅解释了黑噪声的一个方面：长期记忆。

黑噪声也由时间图样里的非连续性所刻画：存在突然的向上或向下的不连续运动。这些不连续突变导致了黑噪声过程的频率分布在均值具有陡峰和胖尾。图 6.1 表明了由于运

用 FBM 产生的变化频率分布。我们能够明白 (1) FBM 模拟算法不必捕捉所有我们寻找的特征, 而且 (2) R/S 分析的一个很大的不足: R/S 分析不能区分分形维高斯噪声和分形维非高斯噪声。单单就 R/S 分析而言, 它不足以计算系统是黑噪声。我们也需要一个陡峰、胖尾的频率分布。甚至于, 存在噪声混沌的第三种可能性。

诺亚效应, 作为黑噪声的一个重要方面, 对于分析而言, 它附加了复杂性的其他层次, 通常被漏看。因为大量的事件在系统中被放大, 所以该效应发生; 那就是, 发生的那些事导致重复的反馈环, 很像逻辑斯蒂克方程里。在逻辑斯蒂克方程里, 突变经常发生, 正如同粉红噪声过程中经常发生突变。在黑噪声里, 它很少发生; 系统保持持续而并非反持续。

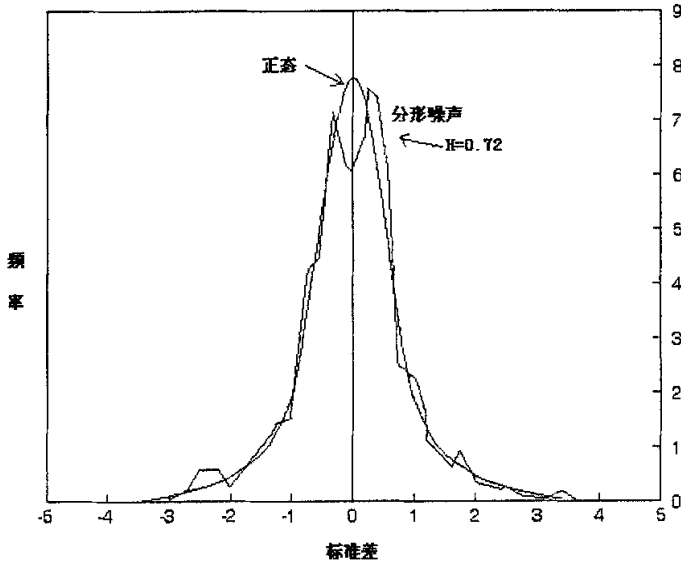


图 6.1

### 6.1.4 镜子效应

粉红噪声和黑噪声在自然界里极易被人发现, 可是这两者之间存在着一个关系吗? 找到一个噪声是否必然导出另一个? 在  $1/f$  噪声谱里, 这种情形极可能发生。

布朗噪声是白噪声的被积函数, 也就是说, 布朗噪声是白噪声的简单运行总和。布朗噪声的易变性和导数都是白噪声。因此, 在  $1/f$  标绘图, 白噪声序列经由这类“镜子”效应将轻易地转变为布朗噪声。

在等式(7.1)里, 谱指数  $b$ , 等于  $2 * H + 1$ 。对于 FBM 的衍生物, 其特征指数是  $2 * H - 1$ 。这样,  $0.50 < H < 1.00$  的持续性序列将有一个大于 2.0 的谱指数, 特指黑噪声过程。然

而，黑噪声过程的衍生物  $b < 1.0$  时，就变成了粉红噪声。

因此，我们并不奇怪，股票市场价格的易变性是反持续的。而正如我们发现的那样，市场收益是黑噪声过程，因而它们的加速度或易变性就应该是粉红噪声。当进一步考察粉红噪声和黑噪声之间的关系时，我们会测定第二个变化——变化中的变化（通过 R/S）分析。根据这一关系，假如第一个差别是黑噪声，那么第二个差别就应该是粉红噪声。

## 6.2 噪声混沌与 R/S 分析

分形布朗运动（FBM）作为一个可能的市场收益的模型，有许多重要的特性，与分形市场假说一致。其中有一些是在时间上的统计自相似、持续性，它产生了趋势与循环。统计自相似性与收益频率分布一致。在外形上，它们是相似的，在不同时间的规模变化上。持续性则与在不同投资起点，不规则地吸收信息的见解一致。市场收益显示了一个黑噪声，而易变性是一个粉红噪声，最后的结论是这个结果由这两个带色的噪声之间理论上的关系组成。

### 6.2.1 信息与投资者之间的关系

投资者们接纳了的各种不同的信息模式有很多。最简单的那些版本假定所有投资起点信息的瞬间性和同质性的解释。这导致了一个在所有时间的“公平”价格，而且是有效市场假说的基本原则。在定价结构上，以及胖尾上解释不连续性，假定信息是在“波浪起伏”不连续的状态下到达的。投资者仍旧同质地反映信息，但是，信息得到达是不连续的。这种理论为尺度理性的假定，对有效市场假说是重要的，但是，收入频率分布的形状和定价结构不连续性的辨别不被考虑。然而，两种理论都蔑视这么一个事实：人们不能以这种方式作决定。

对于在每个投资起点的投资者来说，一条特定的消息未必重要。当一条重要消息已经被明显指出，然后市场才能做出快速决定。假如数据是噪声，噪声可归因于结构性原因的特殊指标的易变性，或者是测度问题。二者有助于对市场无能的信息的一致评价。

有另外一种可能性：新的信息有助于增加不确定性的水平，而不是增加知识的水平。一般而言，经济学家认作新信息是一个积极的发展。新信息增加当前情况的知识有助于判断将来。我们增加的知识导致比较公正的证券价格。然而，也有增加不确定性的信息，否定我们认为我们已经知道的。套利等价理论以变量中的未预期到的变化而涉及这一问题，可是这些未预期的变化的影响未被考虑。随着新信息到达时的不确定性的上升，可能实际上导致增加关于“公平”价格的不确定性，而不是增加自动的价格合作。我们可能得到了增加了的易变性，或仅仅是一个噪声的颤抖。这种噪声经常极可能产生在高频率中，那里，市场试图计算出它所到达的当前信息的值。

噪声的问题并不简单。测量误差不是产生噪声的唯一原因。它会是系统自身的一部分。

这两种噪声都可能存在。

测量噪声（也指那些可以观测的噪声），到目前为止，使经济数据问题中最普通的问题。测量经济活动是一个由于数据收集问题而做得很不精密的科学。作为结果，在事实之后，我们通常不知道衰退什么时候终止或开始。

当指标自身是易变性时，第二类噪声发生。最广泛跟随的经济指标是价格，它在领悟价格膨胀趋势之后才开始使用。商品价格自身服从他们自身的市场商品律。消费者价格指数（CPI）通常被移动的“易变的食品和能源”部分所分解。结果小于易变的膨胀数字被称作“核心比率”。甚至在 CPI 里的一个变化能够被解释成许多不同的样子。在 CPI 里，市场似乎反映当前趋势，以及相似的易变指标，而不是每月发布的变化，除非觉察到趋势已改变。除非趋势在早些时候已经发生变化，否则人们觉察不到他已经改变了。

在较长的频率里，市场以非线性的方式反映经济和基础信息。除此之外，假定市场与经济相关不无道理。这暗示，非线性动态系统模型其相互作用，会是一个恰当的方式，满足由分形布朗运动遗留给分析市场假说的、未解决的方面。非线性动态系统使它们自己趋于非周期循环以及有限集，称作吸引子。系统自身落入混沌系统分类之下。但是，为了能被称为混沌，必须满足非常特殊的条件。

## 6.2.2 混沌存在的必要条件

混沌系统是典型的反馈性系统，它们服从不稳定行为。事件的放大以及不连续性，对于被认为是混沌的系统，存在两个基本的必要条件：（1）分形维的存在，以及（2）北城所初始条件敏感性依赖的特征。在分形噪声和噪声混沌彼此之间进行区别是很困难的，特别是在检验经济数据时。R/S 分析是区别它们之间关系很稳妥的方法。在实验性数据里发现混沌是非常沮丧的事。大部分方法对于混沌不稳妥，相对而言，R/S 分析不仅对于混沌方面稳妥，而且由此繁荣发达。对于不仅仅是市场分析工具箱，R/S 分析是有用的补充，它还是科学家研究混沌现象的有力工具。

## 6.2.3 R/S 分析应用于混沌

### 1. 噪声指数

对于图 4.3， $H=0.92$ 。混沌流的连续、光滑的性质使得赫斯特指数非常高。因为重叠机制或经常在这一等式里的时间痕迹发生的转向使它不等于 1。在图 4.4，我们对系统补充一个白色的，均匀噪声的标准差。这使赫斯特指数降至 0.72 并且被显示在对噪声混沌的第一个 R/S 分析的应用：作为噪声指数使用赫斯特指数。

假如你希望检验逐月要素指标，而且计划运用麦凯-格拉斯等式去测验这些等式，并指导，对于每月数据有 0.72 的值，你就做一个逼真的现实模拟，加上噪声的一个标准差。在这个方式下，你就可以看到，就噪声而言，你的技术指数是否稳健。假如要考察混沌行



为，有一个能从随机行为里区分混沌的特殊试验，为使这个检验实用，你必须就噪声表明，它是稳健的。因为大部分观察到的时间序列，具有接近 0.70 的赫斯特值，你将需要足够的噪声是你的测试系列具有  $H=0.70$ 。或者你可以逐步地添加噪声，并观察在你的测试显示不确定时  $H$  的水平。

图 6.2 显示，当增加的噪声被补充到麦凯—格拉斯等式里时赫斯特的值。赫斯特指数迅速跌落至 0.70，而且逐步降至 0.60。这样，在添加了两个噪声标准差之后， $H$  仍然逼近 0.60。这就是说， $H=0.70$  的频率值如此迷惑 Hurst，它可能归于这样一个事实，即对非线性动态系统添加噪声，迅速使  $H$  值降至 0.70。另一方面，在市场被发现 0.65 以下的  $H$  值的读数很可能不仅仅是由于对一个混沌吸引子添加测度或可加性噪声而产生的，相反，可能是由于分形噪声所引起的。这或许进一步支持这个看法：在短期内市场是分形噪声，而在长期内市场则是混沌噪声。

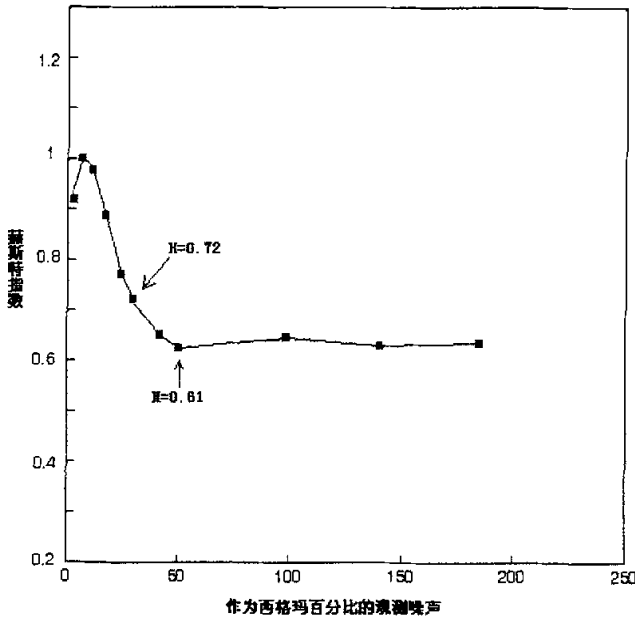


图 6.2 麦凯—格拉斯等式，对噪声敏感的赫斯特指数

## 2. 系统噪声

除了将要测试的可加性噪声，还存在另一种被称作系统噪声的噪声类型。当一个重复的系统随着噪声而成为衰退输出时，系统噪声发生，但这个重复系统不能从纯粹的信号里辨别出噪声信号，而是作为输入使用噪声信号，以作为下一个重复。这寓意观测到的噪声很不相同。因为，观测者将有测量过程的困难，所以该噪声发生。过程不顾我们的问题而继续下去。这样，对系统噪声而言，噪声侵入了系统自身。因为敏感依赖初始条件的问题，系统噪声增加了预测的问题。

在市场里，系统噪声而非观测噪声，更像是一个问题。面对它：我们毫无疑问直到最后一个交易值，可是我们不知道它的值公平与否。或许销售者是在危机中而必须在可能的价格出售以做到收益要求。我们反映这个“噪声”输出，但不知道它的真实值。假如系统噪声是难处理的，那么预测就变得更为困难而且检验将依此而调整。

系统噪声在赫斯特指数上的影响类似于可加性噪声，并且在图 6.3 中表示出来。

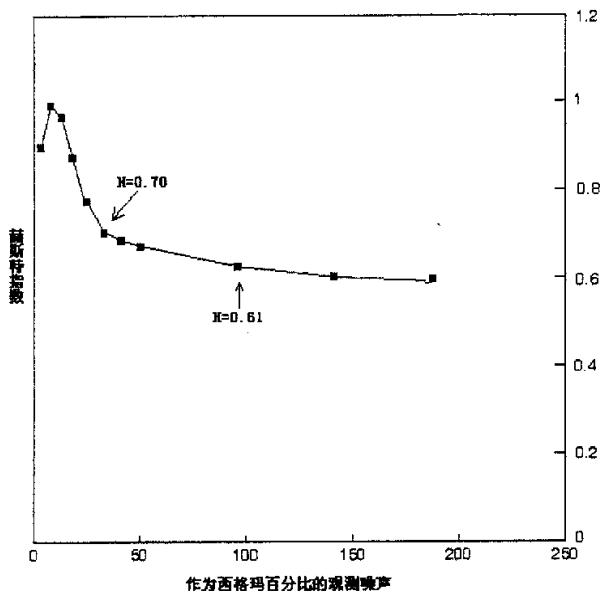


图 6.3 麦凯-格拉斯等式，对噪声敏感的赫斯特指数

### 3. 循环

在第 4 章讨论了 R/S 分析如何能够区分一个循环，即便是在观察噪声的一个标准查处现实。赫斯特指数实际上是同一的 ( $H=0.72$ )，而且 50 个观测循环仍是可以觉察到的。

当  $\log/\log$  的标绘图的斜率穿过一个随机游走的时候，意味着什么？有两个可能的解释：

1. 记忆过程可能是个具有长期但是有限记忆的分形布朗运动。对于有限记忆没有任何其他原因的解释，所以，它可能是一个观测数的函数。因为对于大的  $n$  值足够的观察值不存在，所以规模变化通常停止。

2. 系统是一个噪声混沌系统，而且有限记忆长度测量了吸引子的重叠。在向空间里相邻轨道的分叉意味着在轨道周期之后它们变得不相关了。这样，在轨道性循环之后记忆过程停止。基本上，有限记忆长度成为使得系统忘掉它的初始条件的的时间长度。

当系统通过一个轨道时，它遍历吸引子的长度；当它覆盖吸引子的长度时，极差就不会增长的更大，这是因为，吸引子是有限集。分形噪声过程不是优先级，而且这样的极差将不会停止增长。吸引子的这种形而下的性质，也适合重标极差的性质。

两种解释好像都有道理，特别是我们使用短期数据集时，我们如何在两者之间定夺呢？

#### 4.区分来自分形噪声的噪声混沌

假如，对数比对数标绘图里的断裂真是一个循环，而不是一个统计制品，那么，它将独立于 R/S 分析里使用过的时间增量。假如循环在样本上不独立，那么我们就能恰当确定正在考虑的噪声混沌而非分形噪声。

假如我们面临的是一个以小数据集为开始的问题，那么，我们将使用循环长度估计作为一致证据的另一段，但是，它自身却是不明确的。例如，假定我们对 500 个观测值数据集使用 R/S 分析，就会发现明显的赫斯特指数 ( $0.50 < H < 1.0$ ) 以及 50 个观测值的循环长度，这暗示，我们有 10 个观测循环，每个循环有 50 个观测值。我们计算了最大李亚普诺夫指数的估计，假如这个指数是正的，对于我们做出系统是混沌的结论有好处。假如，最大李亚普诺夫指数的反转更近似与循环长度，那么我们有更大的确定性，在长期里是混沌的。

### 6.3 BDS 统计的概念

BDS 统计，是一个在相关维上的变体，它基本上用来测量相关维计算的统计显著性。对于从确定性混沌或从非线性随机系统里区分随机系统来说。它是强有力的工具。但是，它不能区分非线性确定性系统和非线性随机系统之间的关系。一般来讲，它发现了非线性依赖。联合其他的对于混沌的测试方法与 R/S 分析并用，它会是很有用的。

根据 BDS 测试，相关积分将是正态分布。相关积分是概率，那就是，任意两个在相空间里，在给定的长度  $e$  之内分开的点，随着  $e$  的增加，根据相空间得分形维，概率呈现规模增长。相关积分是根据下列等式计算的：

$$C_m(e) = (1/N^2) * \sum_{i,j=1}^T Z(e - |X_i - X_j|) \quad i \neq j \quad (6.9)$$

这里，如果  $e - |X_i - X_j| > 0$ ,  $Z(x) = 0$ ；此外， $Z(x) = 1$ ； $T =$  观测数； $e =$  距离； $C_m =$  维数  $m$  的相关积分。函数  $Z$  计算了距离  $e$  内相互之间的点数。根据理论， $C_m$  将在比率  $e^D$  上增加，这里的  $D$  是相空间的相关维。

来自于等式 (6.9) 的相关积分，计算了在相空间的  $e$  单位分开的两个点的概率，这两个点是两条不同轨道的一部分。假定  $X_i$  在时间序列  $X$  (具有  $T$  个观测值) 是独立的。我们滞后这个序列到 “ $N$  历史”，那就是，使用时间延滞方法产生一个来自时间序列  $X$  的  $N$  维的相空间。使用等式 (6.9)，进而计算相关积分  $C_N(e, T)$ ，以  $T$  无穷方法表示：

$$C_N(e, T) \Rightarrow C_1(e)^N \quad \text{具有 100\% 的概率} \quad (6.10)$$

这是典型的随机过程规模变化的特征。相关积分简单地填满了无论维数被放置在什么里面的空间。 $|C_N(e, T) - C_1(e, T)^N| * \sqrt{T}$  是具有零均值的正态概率分布。BDS 统计  $W$ ，也

服从正态分布：

$$W_N(e, T) = |C_N(e, T) - C_1(e, T)| * \sqrt{T} / S_N(e, T)$$

(6.11)

这里， $S_N(e, T)$  = 相关积分的标准差。

这样，BDS 统计， $W$  具有正态分布。在它大于 2.0 的时候，我们能以 95% 的置信区间，研究下的系统是随机的零假设而拒绝它。当它大于 3.0 时，我们能以 99% 的置信区间拒绝它。并且 BDS 检验将在数据中发现线性和非线性依赖。这样的话，对于这个检验，需要采用 AR(1) 的残值，正如我们对 R/S 分析所做的一样。另外，像 R/S 分析，依赖性可能是随机的（诸如赫斯特过程，或 GARCH），或者它可能是确定性的（诸如混沌）。

为了做检验，必须选择半径，即  $e$  的值，以及嵌入维  $m$ 。假如  $e$  太小，将没有足够的店铺统计结构，假如  $e$  太大将有太多的点。我们选择一个使得导致向空间重构不要太稀疏也不要太拥挤的嵌入维。假如  $m$  太小，点将被紧密地包在一起，假如  $m$  太大，点将里的太远。

为了进行比较研究，我们构造了另一个比较的时间序列：首先，建立一个自回归模型去拟合原收益序列，然后，将残差序列随机地变化位置并且用线性模型的估计值与实际数据的相同的初始值重新构造一个时间序列。对收益率序列用线性模型拟合后的残差进行 BDS 统计。

## 6.4 BSD 统计的实证研究

为了消去线性依赖，我们还是使用检验 R/S 分析的时候所用的残差序列  $\{X_t\}$ ，为了检验价格是否服从随机游走模型，也就是说市场是否有效，我们就是要检验价格变动序列  $\{X_t\}$  是否是独立分布的。

人们通常习惯用一阶的非相关性来近似独立性，而对高阶的情况毫不关心，但是早在 1963 年分形几何的创始者 Mandelbort 受水文学中的“Hurst 现象”的启发，发现股票收益中也存在“波动集聚效应”，收益波动代表着风险，其数学度量为二阶距，这说明收益在二阶距上是相关的。Engle (1982)、Bollerslev (1986) 和 Nelson (1991) 等人发展的 ARCH 族模型，正是基此为探求收益和风险及风险自身数量关系而提出的。事实上收益和风险的关系也恰恰就是数理金融的核心领域之一。因此，验证随机游走的收益独立性时，不考虑高阶情况是不合理的。

这里，我们还是选用上证综指 1999 年 1 月 1 日至 2005 年 12 月 31 日的 1637 个日收盘价的数据，先进行拟合，产生序列  $\{X_t\}$ 。下面就是对  $\{X_t\}$  所作的二阶 BDS 检验，距离参

数分别取分别取  $\{X_t\}$  序列的 1 个标准差, 1.25 个标准差, 1.5 个标准差,  $e=1$ ,  $N$  分别取 2、4、6、8。得出的结果如 6.1 表所示:

表 6.1

N	1s.d				1.25s.d				1.5s.d			
	2	4	6	8	2	4	6	8	2	4	6	8
$C_N(e,T)$	0.521	0.240	0.101	0.039	0.641	0.373	0.189	0.100	0.733	0.490	0.312	0.183
$S_N$	0.099	0.234	0.293	0.295	0.084	0.255	0.403	0.507	0.067	0.246	0.460	0.677
BSD	-5.27	-5.32	-4.59	-3.70	-6.63	-6.51	-5.41	-4.03	-8.25	-7.68	-6.02	-4.87

由表 6.1 可见, 统计量 BDS 的绝对值均大于 5% 时标准正态分布的临界值 2, 因此拒绝  $\{X_t\}$  序列二阶独立的假设, 因此可以得出结论,  $\{X_t\}$  二阶相关, 这也从统计上验证了中国股市确实存在“波动集聚效应”, 收益  $\{X_t\}$  存在异方差, 这也说明中国股票市场不是有效的。这也进一步证明了上一章节论述的内容。

通过上面的实证我们可以看到, BDS 统计对于高斯零假设是稳健的。能如同 R/S 分析一样能够发现线性依赖型, 一旦线性依赖被过滤掉, BDS 统计依然是显著性的检验。但是, BDS 检验不能区分分形噪声和确定性混沌, 即便这样, 联合其他检验一起使用的话, BDS 将是个有力的工具。

## 第7章 总结和展望

### 7.1 总结

本论文运用 R/S 分析方法来研究中国股票市场的分形特征问题，以进一步揭示中国股票市场的内在特性。发现了股票市场的自相似性和非周期循环，并计算出了该非周期循环的长度，揭示了中国股市的分形特征。通过对上证综合指数的收益率数据进行了 R/S 分析，发现中国股市不服从正态分布，而是个具有长期记忆的有偏随机游走。通过对数据进行分段回归，发现股市存在一个非周期循环，并通过 V 统计计算出了该循环的长度。股市中存在许多种类型的噪声，将 R/S 分析应用于有噪声的序列以检验 R/S 分析方法是否依然有效，发现 R/S 分析依然能发现过程中的非线性并能得出存在的非周期循环，可见，R/S 分析对于噪声来讲也是非常稳健的。最后，通过使用 BDS 统计，进一步证实了中国股市并非独立过程，而是二阶相关的，中国股票市场并不是个有效的市场。本文得出的结论基本上能够如实地反映中国股票市场的基本分形特征，能给股票投资者提供了一个良好的分析方法和建议。

通过本论文的分析，相信大部分投资者能够明白什么是必需的。在短周期里，我们需要持续赫斯特指数与自相似性频率分布。在长周期里，我们需要持续赫斯特指数，长期有限记忆和非周期循环。短循环并没有呈现出稳定，长循环对于所有的已研究的时间周期是牢固和稳定的。对于股票市场和证券市场有如下建议：在短期里，市场由交易过程所统治，它们是分形噪声过程，近似于无限方差行为，在很长的周期里，市场由确定性非线性系统或确定性混沌所特征化。非周期循环从它们当中各种各样独立的资本市场因缘而起，也从各色独立的经济中而产生。市场基本上由不联系基础经济波动的交易者所统治，即便是在长期的情况下，也不会被确定性混沌所特征化，总体上是稳定列维特征。

### 7.2 展望

经过论文的研究工作，深深感到分形理论应用股票市场这一研究工作的内容还非常多，自己研究的内容在很多地方还需要不断地进行改进。就本论文研究工作来说，至少还有以下几个方面需要改进，或者有待进行深入地分析和研究：

- 1、在进行拟合的过程中，发现分组情况以及组内数据的多少将影响回归方程，进而影响赫斯特指数并可能影响结论，有必要进行进一步的完善。

- 2、在对数据进行拟合的时候采用的模型并不完美，至今完全令人满意的一个都没有，似乎每个模型都是提出了一个局部的市场性质，许多这样的局部性质似乎都与某些、而不是与所有的投资起点联系在一起。所以拟合度并没有达到很精确的程度，有待进一步的完善是与所有的投资起点联系在一起。所以拟合度并没有达到很精确的程度，有待进一步的完善。

3、在经济数据的序列中存在许多噪声，R/S 分析对于有噪声的系统也是很稳健，通过使用 BDS 统计，能够发现序列中噪声是线性的还是非线性的，但是却不能区分噪声是分形噪声还是混沌噪声，所以最好是将 BDS 统计联合其他的检验一起使用，这将使分析更有力。

## 参考文献

- [1] Eugene F.Fama.a, Efficient Capital market II ,[J].Journal of Finance, 1995,5:1575~1617.
- [2] Sharpe,W.F. ,Portfolio Theory and Capital Marke[M]t,New York:McGraw-Hill,1970.
- [3]Turner,A.L.,and Weigel,E, An Analysis of Stock Market Volatility[J],Russell Research Commentaries.Tacoma,WA:Frank Russell Co.,1990.
- [4]Peters,E.Fractal Structure in the Capital Market[J],Financial Analysts Journal,July/August 1989.
- [5]Perers,E.Chaos and Order in the Capital Marke[M]t.New York:Jhon Wiley&Sons,1991a.
- [6]Peters,E.A Chaotic Attractor for the S&P 500[J],Financial Analysts Journal,March/Arpil 1991b.
- [7]Peters,E.R/S Analysis using Logarithmic Returns[J]:A Technical Note,Financial Analysts Journal,November/December 1992.
- [8] Hurst,H.E.The Long-Term Storaga Capacity of Reservoirs[J],Transactions of the American Society of Cihil Engineers 116,1951.
- [9]Mandelbort, B.The Fractal Geometry of Nature[M].New York:W.H.Freeman,1982.
- [10]Mandelbort,B.Statistical Methodology for Non-Periodic Cycles:From the Covariance to R/S Analysis[J],Annals of Economic and Social Measurement 1,1972.
- [11]Vaga,T.The Coherent Market Hypothesis[J] , Financial Analysts Journal,December / January 1991.
- [12]Larrain,M.Protfolio Stock Adjustment and the Real Exchange Rate:The Dollar-Mark and the Mark-Sterling[M].Journal of Policy Modeling,Winter 1986.
- [13] Larrain,M.Empirical Test of Chaotic Behavior in a Nonlinear Interest Rate Model[J],Financial Analysts Journal 47,1991.
- [14] [美]比尔·威廉姆 (Bill Williams, Ph. D.) 著, 王宁 译, 证券交易新空间 面向 21 世纪的混沌操作获利指南[M], 地震出版社, 2004。
- [15] [美]埃德加·E·彼得斯 著, 储林海 殷勤 译, 刘建平 校, 分形市场分析—将混沌理论应用到投资和经济理论上[M], 经济科学出版社, 2002。
- [16] 埃德加·E·彼得斯 著, 王小东 译, 资本市场的混沌与秩序[M], 经济科学出版社, 1999。
- [17] [美]比尔·威廉姆 (Bill Williams, Ph. D.) , 证券混沌操作法低风险获利指南[M], 宇航出版社, 2002。
- [18] 陈平, 文明分岔、经济混沌和演化经济动力学[M], 北京大学出版社, 2004。
- [19] 刘式达, 梁福明, 刘式适, 辛国君, 自然科学中的混沌和分形[M], 北京大学出版社, 2003。
- [20] 邱沛光, 上证指数 Hurst 指数的测定及应用, 西北农业科技大学学报[J], 2004, 9: 92~95。
- [21] 马燕舞, 刘学华, 从有效市场假说到分形市场假说[J], 经纪人学报, 2005, 1: 63~



65。

- [22] 莫馨, 马军海, 经济时间序列的非线性特征检验及其应用[J], 河北工业大学学报, 2004, 12: 13~18。
- [23] 叶中行, 曹奕剑, Hurst 指数在股票市场有效性分析中的应用[J], 系统工程, 2001, 3: 21~24。
- [24] 庄新田, 庄新路, 田莹, Hurst 指数及股市的分形结构[J], 东北大学学报, 2003, 9: 862~865
- [25] 黄诒蓉, 中国股市分形结构的 R/S 实证分析[J], 现代管理科学, 2005, 2: 53~55。
- [26] 杨俊凯, 基于长期记忆特征时序分析的中国股市效率性检验[J], 西安金融, 2005, 1: 37~38。
- [27] 宗兆昌, 田华, 中国股票市场分形特征的实证研究[J], 统计观察, 2004, 12: 85~87。
- [28] 解保华, 马征, 高荣兴, 中国股票市场有效性实证检验[J], 广东商学院学报, 2001, 5: 22~25。
- [29] 范英, 魏一鸣, 基于 R/S 分析的中国股票市场分形特征研究[J], 系统工程, 2004, 11: 46~51。
- [30] 马凤翔, 韦凯华, 基于重标极差法 (R/S) 对金融市场非线性动力学特征的解构[J], 金融论坛, 2006, 2: 60~63。
- [31] 黄诒蓉, 罗奕, 论 R/S 分析法与股票市场的分形结构[J], 现代管理科学, 2006, 1: 105~107。
- [32] 徐国栋, 邱世远, 中国股市分形现象与非周期循环[J], 统计与决策, 2004, 2: 89~90。
- [33] 李涛, 贺勇军, 刘志俭等, Matlab 工具箱应用指南—应用数学篇[M], 北京: 电子工业出版社, 2000。
- [34] 肖冬荣, 肖逸波, 黄崇起, 李涛, 股指预测的马尔柯夫过程模型[J], 天津商学院学报, 1997, 2: 29~33。
- [35] 徐龙炳, 缪铨生, 陆蓉, 上海股票市场分类指数的非线性研究[J], 镇江师专学报, 1999, 2: 86~88。
- [36] 宋家旺, 张世英, 洞悉中国股市的波动特征及其时间敏感性—R/S 分析方法的理论及其应用[J], 天津理工大学学报, 2005, 2: 5~9。
- [37] 谢鸿飞, 分形市场假说及其对沪市的实证分析[J], 惠州学院学报, 2005, 10: 18~24。
- [38] 耀仲诚, 郭华林, 分形市场假说在上海股票市场中的实证研究[J], 统计观察, 2004, 2: 80~81。
- [39] 肖冬荣, 朱京, 张辉, 基于决策指标分类的供需链合作伙伴选择算法[J], 辽宁工程技术大学学报, 2002, 10: 664~666。
- [40] 肖冬荣, 陆振宇, 鲁棒控制理论应用于宏观经济系统分析[J], 决策与控制, 2002, 9: 629~634。
- [41] 曹宏铎, 李昊, 经济系统分形机制与股票市场 R/S 分析[J], 系统工程理论与实践, 2003, 3: 9~13。
- [42] 郇元兴, 朱喜安, 李炳林, 证券市场风险度量方法的分析与评价[J], 郑州航空工业管理

学院学报, 2005, 3: 105~108。

[43] 姚中立, 应益荣, 上海股票市场分形特征检验[J], 商场现代化, 2005, 9: 77~78。

[44] 周鸿涛, 王宗军, 上海股市非线性特征: 一个基于 R/S 方法的实证分析[J], 管理学报, 2005, 9: 597~600。

[45] 陈昭, 梁静溪, 赫斯特指数的分析与应用[J], 中国软科学, 2005, 3: 134~138。

[46] 何兴强, 中国股票市场收益非线性相关结构的经验分析[J], 世界经济, 2004, 8: 60~67。

[47] 陈春晖, 雷旭辉, R/S 在深圳股票市场有效性分析中的应用[J], 金融与经济, 2005, 1: 37~39。

## 致 谢

在此论文完稿之际，首先向三年来给我悉心指导、精心培养的导师肖冬荣教授致以最衷心的感谢！在三年的学习生活中，导师以其严谨求实的治学态度、精神渊博的学术造诣以及忘我的工作精神给我留下了深刻的印象，这必将对我以后的学习、生活和工作产生积极而深远的影响。

感谢所有同门以及师弟、师妹在学习中对我的热情帮助，感谢宿舍舍友储晶、赵妍、李淑静对我的论文提出的宝贵意见。同时感谢多年来父母对我学业的支持，是他们在不断地支持我和鼓励我，使我顺利完成学业。

在研究生阶段，研究生部钱翼梅老师给予了我很多关心和帮助，在此向她表示感谢！

## 作者在学期间发表的论文清单

- 1、李元媛，肖冬荣，赵妍，混沌理论在股票市场分析上的应用，微计算机信息，2006，3-3：148~150。
- 2、赵妍，肖冬荣，李元媛，几种网络控制系统的比较及典型应用，微计算机信息，2006，4-1：40~42