

基于混沌神经网络的股票分析及其预测

张中华¹, 丁华福²

(1. 哈尔滨理工大学 计算机科学与技术学院, 黑龙江 哈尔滨 150080;

2. 哈尔滨工业大学 计算机科学与技术学院, 黑龙江 哈尔滨 150001)

摘要:混沌动力学理论提供了证券市场中股价波动的一种分析方法。为了考察中国证券市场的价格是否存在混沌行为,以1990.12.19到2008.4.24的上海证券交易所每天的收盘数据,分析了价格波动的非线性特征,通过重构相空间方法重构了1990年到2008年上证指数时间序列的奇怪吸引子,计算其关联维数,并求出其Lyapunov指数为正,从而确认了上证指数时间序列的混沌行为。神经网络是一个非线性系统,通过学习,可以实现非线性函数逼近,从而能更好地实现股票的走向预测。

关键词:混沌行为;重构相空间;关联维数;Lyapunov指数;神经网络

中图分类号:TP183

文献标识码:A

文章编号:1673-629X(2009)03-0185-04

Analysis and Prediction of Stock Based on Chaotic Neural Networks

ZHANG Zhong-hua¹, DING Hua-fu²

(1. Sch. of Computer and Techn., Haerbin Univ. of Sci. and Techn., Haerbin 150080, China;

2. Sch. of Computer and Techn., Haerbin Univ. of Industry, Haerbin 150001, China)

Abstract: The chaotic dynamics theory has provided an analytic method to stock price volatility in the stock market. To find out whether the price of stock has the chaotic behavior, has analyzed nonlinear character of the stock price, based on the close index of stock, from the 19th of Dec in 1990 to 24th of Apr in 2008, and have reconstructed the strange attractor of the time sequence by the method of reconstructing phase space, have calculated correlated dimension, and have proved that the exponent of Lyapunov is positive, then affirmed that the chaotic behavior does exit in the time sequence of stock. The neural network is a nonlinear system, it can approach the nonlinear function by learning, by using the neural network, can forecast the stock's direction of the next day.

Key words: chaotic behavior; reconstructed phase space; correlated dimension; Lyapunov exponent; neural network

0 引言

混沌时间序列研究,始于Packard等在1980年提出的重构相空间理论。由于系统各个自由度之间的非线性相互作用,每一独立变量的时间演化,也包含了系统所有其他变量的长期演化信息。可通过某一单变量的时间序列来研究非线性系统整体的混沌行为^[1,2]。而动力系统在相空间中的吸引子的不变量,如关联维数、Lyapunov指数可以定量地表征系统的混沌性质,其中, Lyapunov指数是对于初始邻近轨道的指数发散速率的量化,可以从整体上反映动力系统的混沌程度。因此,基于时间序列相空间重构方法的Lyapunov指数

计算对于一个无法写出其微分方程的动力系统是否存在混沌行为的判断尤其重要^[3,4]。

到目前为止,混沌时间序列的预测已被应用于很多领域,如河水的径流量预测^[5]、城市用水量的预测以及安全库存的预测^[6]等,它们均是基于混沌时间序列理论的神经网络预测。如今,我国股票市场,跌宕起伏,为了认清股票市场是否健康发展,以及对将来的股市做出尽可能的准确走向预测,基于混沌理论,对股票进行分析,并用神经网络对其预测。

文中首先对上海证券交易所1990年到2008年每天的收盘指数数据分析了价格波动的非线性特征,通过重构相空间方法重构了1990年到2008年每天的收盘指数时间序列的奇怪吸引子,计算出其关联维数,并根据Wolf方法求出其Lyapunov指数,证实其最大Lyapunov指数为正,从而确认了上证指数时间序列的混沌行为。通过小波混沌神经网络对重构相空间的学习^[7],实现股票收盘指数的预测走向。

收稿日期:2008-07-10

基金项目:国家自然科学基金(60736014)

作者简介:张中华(1981-),男,山东人,硕士研究生,研究方向为神经网络;丁华福,博士,教授,硕士生导师,研究方向为自然语言处理、数据挖掘。

1 数据采集

采集了上证指数的从 1990 年到 2008 年的每天的收盘指数数据,一共有 4255 个数据,有了这样充足的数据,可以讨论上证指数的非线性特征,将原始数据制成折线图(如图 1 所示),这就是要分析的时间序列。

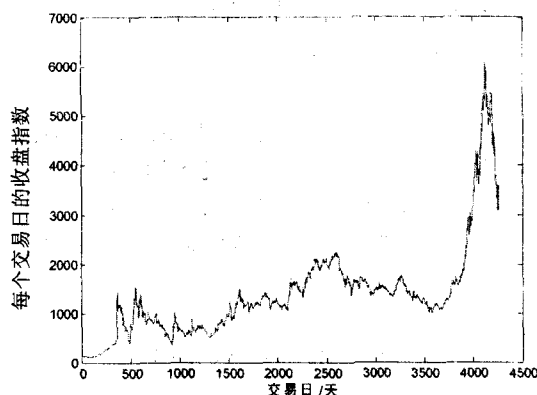


图 1 从 1990.12.19~2008.4.28 每天的收盘指数

2 相空间重构

根据观察到的上述时间序列 $x(t)$ 重构相空间中的运动轨道,或者说是吸引子。常用的方法是时间坐标延迟法。即用 $x(t)$ 作第一个坐标, $x(t + \tau)$ 作第二个坐标, $\dots, x(t + (m - 1)\tau)$ 作第 m 个坐标,也是最后一个坐标,这样来重构完全的 m 维状态矢量 $\mathbf{X}(t)$ 。显然应用该方法需要解决两个重要问题^[7]: 如何确定时间延迟间隔 τ 和吸引子的嵌入维 m 。

这里,采用自相关法求时间延迟 τ ,自相关法也称自相关函数法。对一个时间序列 x_1, x_2, x_3, \dots , 序列跨度为 $j\tau$ 的自相关函数为:

$$R_{xx}(j\tau) = \frac{1}{N} \sum_{t=0}^{N-1} x(t)x(t+j\tau) \quad (1)$$

由此可固定 j , 做出自相关函数关于时间的 τ ($\tau = 1, 2, \dots$) 函数图像,则自相关函数下降到初始值的 $(1 - 1/e)$ 倍时,所得到的时间 τ 即是重构相空间的时间延迟 τ_0 。这里取 $\tau = 1, 2, 3, \dots, 4000$, 做出自相关函数 R_{xx} 关于时间 τ 的函数图像,见图 2。

从图 2 看出在 $\tau = 306$ 时,相关函数下降到初始值 $(2.911e + 006)$ 的 $(1 - 1/e)$, 即 0.6321 倍,因此最佳延迟时间为 306 个交易日。

然后,采用由 Grassberger 和 Procaccia 提出的 G-P 算法计算关联维 d ,并根据 Takens 定理得到相应的嵌入维 m 。Takens 定理说:如果延迟坐标的维数 $m \geq 2d + 1$,则在这个嵌入空间里可以恢复吸引子的相空间结构。

具体算法的主要步骤如下:

(1) 对于时间序列 x_1, x_2, x_3, \dots 先尝试取一个较

小的嵌入维数值 m_0 ,则在相应的重构相空间中有:

$$\mathbf{X}(t_i) = [x(t_i), x(t_i + \tau), x(t_i + 2\tau), \dots, x(t_i + (m - 1)\tau)], i = 1, 2, \dots \quad (2)$$

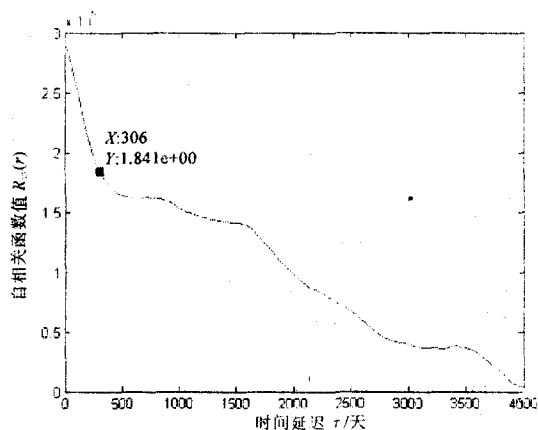


图 2 自相关函数 R_{xx} 关于时间 τ 的图像

(2) 计算关联积分。

$$C(r) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i,j=1}^N \theta(r - |X(t_i) - X(t_j)|) \quad (3)$$

其中 $|X(t_i) - X(t_j)|$ 表示相点 $X(t_i)$ 和 $X(t_j)$ 的距离, $\theta(z)$ 是 Heaviside 函数:

$$\theta(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases} \quad (4)$$

$C(r)$ 是一个关联积分,表示相空间中吸引子上两点之间距离小于 r 的概率。

(3) 对于关联球半径 r 的某个适当范围,吸引子的关联积分 $C(r)$ 应该随 r 的变化呈现幂函数标度律,即在双对数坐标图上, $C(r)$ 随 r 的变化曲线呈现为一段直线,此直线的斜率给出吸引子维数: $d(m) = \ln C(r) / \ln r$ 。对于所有 $(r, C(r))$ 点进行拟合,从而求出对应于 m_0 的关联维数估计值 $d(m_0)$ 。

(4) 增加嵌入维数,取 $m_1 > m_0$,重复计算步骤(2)和(3),直到相应的维数估计值 $d(m)$ 不再随 m 的增长而增长(在一定误差范围内)。此时得到的 $d(m)$ 即为吸引子的关联维数, m 即为嵌入维数。如果 d 随 m 的增长而增长,并不收敛于一个稳定的值,则表明所考虑的系统是一个相应于随机时间序列的随机系统,它并不是少自由度的决定论性非线性动力学系统,它在有限维数的相空间中不存在吸引子。

这里取 $\tau = 306, r = 2000, 1800, 1600, 1400, 1200, 1000, 800, 600, 400, 200, 100, 50$, 嵌入维数从 $m = 2$ 开始尝试,分别得到: $m = 2, d = 1.3234$; $m = 3, d = 1.5436$; $m = 4, d = 1.6321$; $m = 5, d = 1.6714$; $m = 6, d = 1.7013$; $m = 7, d = 1.7097$; $m = 8, d = 1.7031$; $m = 9, d = 1.6785$; $\ln C(r)$ 与 $\ln(r)$ 的关系如

图3所示。

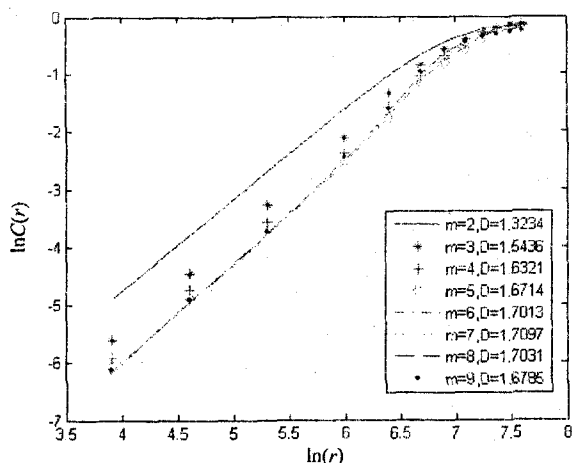


图3 $\ln C(r)$ 与 $\ln(r)$ 的关系图

从图3中可以看出, $m = 4, 5, 6, 7, 8, 9$ 时的曲线几乎重合, 说明 d 随着 m 的增长收敛在 1.7 附近, 表明系统不是一个随机系统, 根据 $m \geq 2d + 1$, 上证收盘指数时间序列数的计算给出吸引子的嵌入维数为 $m = 4$, 关联维数为 $d = 1.70$ 。

从以上的分析可以看出, 我国股市的混沌吸引子是存在的, 1990年12月19日到2008年4月24日上证收盘指数时间序列数据对应于一个关联维数为 $d = 1.70$ 的混沌吸引子。此维数反映了此段时间内股票波动的复杂程度。关联维数值越大, 股票波动越复杂, 其市场越活跃, 所以, 股票指数的关联维数可用来测定股市在某一段时期内的活跃程度。

3 Lyapunov 指数

Wolf (1985) 给出最大 Lyapunov 指数的计算方法。这个方法测度了重构相空间中相邻两点之间的距离随时间的发散速率。Wolf 算法如下^[7]: 对于混沌时间序列 $x_1, x_2, \dots, x_n, \dots$ 在取定嵌入维数 m , 延迟时间 τ 之后, 重构相空间中的 m 维矢量由(2)式表示。

取初始点 $X(t_0)$, 设其与最近邻点 $X_0(t_0)$ 的距离为 L_0 。追踪这两点的时间演化, 直到 t_1 时刻, 其间距超过某给定值 $\epsilon > 0$, $L'_0 = |X(t_1) - X_0(t_1)| > \epsilon$, 保留 $X(t_1)$, 并在 $X(t_1)$ 邻近另找一个点 $X_1(t_1)$, 使得 $L_1 = |X(t_1) - X_1(t_1)| < \epsilon$ 并且与之夹角尽可能的小, 继续上述过程, 直至 $M(t)$ 到达时间序列的终点 M , 这时追踪演化过程总的迭代次数为 $t_M - t_0$, 则最大 Lyapunov 指数为

$$\lambda_1 = \frac{1}{t_M - t_0} \sum_{i=0}^M \ln \frac{L'_i}{L_i} \quad (5)$$

运用此方法, 求得上证收盘指数这一时间序列的最大 Lyapunov 指数 $\lambda_{\max} = 0.0050$ 。

4 基于神经网络股票预测

使用前馈网络进行股票的混沌时间序列预测的方法是通过前馈神经网络的非线性逼近能力来实现预测原点和预测时域之间的定量关系^[8], 假设混沌时间序列通过延迟坐标重构获得如下的延迟坐标向量:

$$X(t) = \{x(t), x(t - \tau), \dots, x[t - (m - 1)\tau]\} \quad (6)$$

式中: m 和 τ 分别为嵌入维数和延迟时间。

假设预测步数为 h , 需要前馈神经网络来逼近非线性函数:

$$x(t + h) = F_h[X(t)] \quad (7)$$

从以上可以知道, 前馈网络的作用是通过学习样本构建延迟向量 $X(t)$ 与 $x(t + h)$ 之间的定量关系, 网络训练完毕后, 就可以直接进行 h 步的预测^[9]。设置预测时域为不同的值, 可以训练得到不同的神经网络模型, 也就是说每一个 h 需要一个神经网络来逼近。

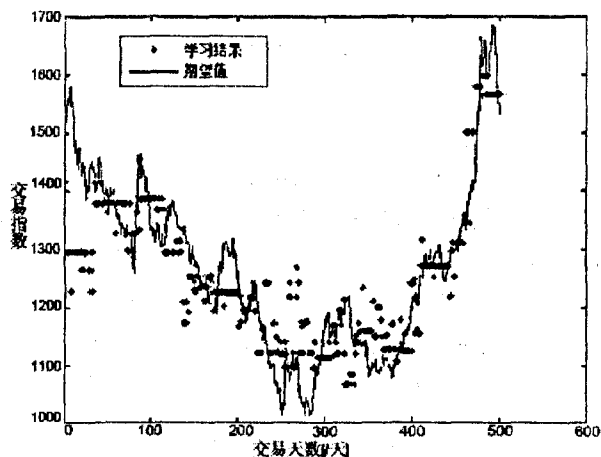


图4 前向神经网络学习结果

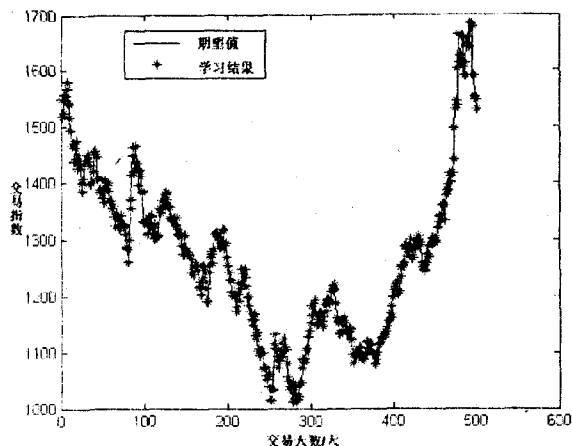


图5 径向基神经网络学习结果

下面分别用前向(BP)神经网络和径向基(RBF)神经网络分别进行学习, 从中取出 500 个样本, 取 $h = 200$, 其中, 前向神经网络的结构为 4-64-1, 即输入层神经元为 4 个, 它是根据嵌入维数确定的, 隐藏神经

元为 64 个,输出层神经元为 1 个,学习的仿真结果如图 4 和图 5 所示,可见,径向基神经网络的学习效果要好于前向神经网络。

再取 30 个样本,对系统进行预测,预测仿真结果如图 6 所示,可见,混沌时间序列通过径向基神经网络,可以实现短期的收盘指数走向的预测。

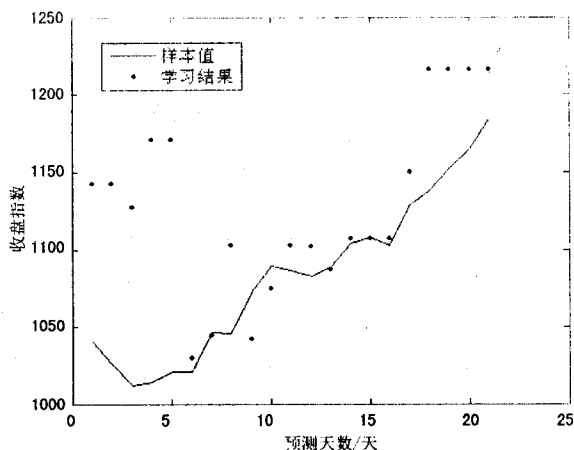


图 6 预测结果

5 结束语

采集了上证指数的从 1990 年到 2008-4-24 的收盘指数数据,经过分析发现收盘指数时间序列存在混沌现象,利用 Taken 定理,对股票数据进行相空间重构,还原了股票系统的力学特性,并计算出股票系统的奇怪吸引子的维数为 $D = 1.70$,嵌入维数为 $m = 4$,最大 Lyapunov 指数为正,因此可见,股票中存在混沌

现象。最后,通过求出的嵌入维数和时间延迟,设计出神经网络,神经网络通过对收盘指数样本的学习,实现了对收盘指数走向的预测。

参考文献:

- [1] 侯媛彬,杜京义,汪梅.神经网络[M].西安:西安电子科技大学出版社,2007:26-52.
- [2] 孙博文,张本祥.中国股市波动的混沌吸引子的测定与计算[J].哈尔滨理工大学学报,2001,6:361-367.
- [3] 吕金虎,陆君安,陈士华.混沌时间序列分析及其应用[M].武汉:武汉大学出版社,2002:118-178.
- [4] 黄润生,黄浩.混沌及其应用[M].武汉:武汉大学出版社,2005:366-390.
- [5] 蒋传文,权先璋.径流序列的混沌神经网络预测方法[J].水电能源科学,1999,17(2):9-11.
- [6] 何佳,王子牛,罗刚,等.基于混沌神经网络技术的安全库存预测研究[J].计算机技术与发展,2007,17(8):247-250.
- [7] 韩敏.混沌时间序列预测理论与方法[M].北京:中国水利水电出版社,2007:155-185.
- [8] 徐耀群,孙明.混沌神经网络时间序列的研究[C]//中国控制与决策学术年会论文集.沈阳:东北大学出版社,2006:397-402.
- [9] Jiang Jianguo, Shao Kuizhi, Wei Yuheng, et al. Chaotic Neural Network Model for Output Prediction of Polymer Flooding [C]//Proceedings of the 2007. IEEE, International Conference on Mechatronics and Automation. Harbin, Heilongjiang, China: IEEE, 2007:2347-2351.

(上接第 184 页)

总线周期以 ALE 信号有效的前沿(上升沿)后的第一个时钟开始计算,需要 5 个时钟周期来完成。

4 结束语

介绍了飞控设备中的 PC/104 总线,并根据实际需要研制了 PC/104 总线接口板,实现了 PC/104 总线的总线控制、终端仿真及数据监控功能。提出利用 FPGA 构建模块 IP 核技术来实现 PC/104 总线通信,并详细介绍了部分电路的实现方法。应用 FPGA 简化了系统结构,缩短了设计周期,提高了系统的性能和可扩展性。

参考文献:

- [1] Shannon L, Fort B, Parikh S, et al. Designing an FPGA SoC Using a Standardized IP Block Interface [C]//Brebner G J, Chakraborty S, Wong W F. Proceedings of the 2005 IEEE International Conference on Field-Programmable Technology.

Singapore: IEEE International, 2005:341-342.

- [2] Ristimäki T, Nurmi J. Reconfigurable IP Blocks: a Survey [C]//System-on-Chip, 2004. Proceedings. 2004 International Symposium. Finland: IEEE International, 2004.
- [3] 井新宇.基于 VHDL 设计有限状态机 FSM 的方法[J].信息技术与自动化,2005(4):29-30.
- [4] 罗朝霞.基于 CPLD 的序列信号检测器设计与实现[J].现代电子技术,2005,28(11):63-64.
- [5] Sjöholm S, Lindh L. 用 VHDL 设计电子线路[M].边计年,薛宏熙译.北京:清华大学出版社,2003:157-186.
- [6] 孔建,杨洪斌.基于有限状态自动机描述及综合[J].计算机工程,2003,29(15):82-83.
- [7] 邬杨波,王曙光,胡建平.有限状态机的 VHDL 设计及优化信息技术[J].信息技术,2004,28(1):75-78.
- [8] 廖日坤. CPLD/FPGA 嵌入式应用开发技术白金手册[M].北京:中国电力出版社,2005:98-99.
- [9] 吴继华,王诚. Altera FPGA/CPLD 设计(高级篇)[M].北京:人民邮电出版社,2005.