

基于进化神经网络的股市预测研究

高 玮

(武汉工业学院 武汉430023)

摘 要 股票市场是一个受众多因素影响的动力学系统,其演化规律非常复杂。由于其对投资及管理的重要性,已提出了大量对其演化规律进行预测的方法,其中基于智能科学的人工神经网络预测方法以其独特的性能得到了人们的重视。基于此,这里提出一种基于免疫进化规划的进化神经网络股市预测方法,并用沪市的一段变化时序对其进行了验证,结果表明了所提方法的有效性。

关键词 股市预测,神经网络,进化神经网络

Stock Prediction Study Based on Evolutionary Neural Network

Gao Wei

(Wuhan Polytechnic University, Wuhan 430023)

Abstract The stock market is a very complicated non-linear dynamic system. Its variation is affected by many factors. Because its importance of stock prediction, there have been proposed many methods. In those methods, the intelligent method based on artificial neural network is more suitable and applied widely. Here, a new method for stock prediction based on new evolutionary neural network from immune evolutionary programming is proposed here. At last, using the real data of Shanghai stock market, our algorithm is tested.

Keywords Stock prediction, Neural network, Evolutionary neural network

1 引言

针对股票市场表现出的不同特点,人们已提出了多种多样的预测方法^[1,2]。但是,股票市场作为一种影响因素众多、各种不确定性共同作用的复杂巨系统,其价格波动往往表现出较强的非线性特征,且股市各因素间关系错综复杂,主次关系变化不定,数量关系难以提取及定量分析。正是由于其复杂的非线性动力学特性,使得关于股市的预测往往难如人意。

要进行股市预测,我们首先必须承认股市存在着某种可重复性,是有一定的基本规律可以遵循的。这些规律完全隐藏在历史数据中,从数学的角度来讲,这种规律表现为函数关系。预测的目的就是要找出这些规律、利用这些规律。研究已表明,基于神经网络的非线性时间序列预测方法是目前较好的一种预测方法^[3,4]。尽管已有人进行过人工神经网络在股市预测中应用的研究^[5,8],但目前的研究尚令人满意,因此,本文提出一种预测股市变化的进化神经网络新模型,并通过一个实例对其可行性进行研究。

2 股市预测的数学描述

由于股市变化的复杂性,其动态演化过程为一

个复杂的非线性动力学系统。从而,其时序预测建模也为该动力学系统的识别过程。以下对这一识别过程从数学上给予说明。

假设股市的演化过程可由一个 p 阶常微分方程描述,

$$\frac{d^p x}{dx^p} = f(x, x', \dots, x^{(p-1)})$$

如果这一过程的一个状态变量是该常微分方程的解,则根据这一变量在等时间间隔的观测值,可建立该动力学模型。为讨论方便,以二阶常微分方程为例进行说明,

$$\frac{d^2 x}{dt^2} = f(x, \frac{dx}{dt})$$

设上述方程的解 x 为某时序在等时步长 h 上的取值 $x(j)$, $j=1 \sim n$,则可利用常微分方程的数值解建立系统的模型。

用差分代替上述方程中的微分在 j 处的取值,并把差分式代入常微分方程中,可得到,

$$x(j+1) = h^2 f[x(j), \frac{x(j) - x(j-1)}{h}] + 2x(j) - x(j-1) + o(h^3)$$

去掉高阶小量,并把上式写成通式,并把它推广到 p 阶方程,得到下式,

$$x(j+1) = F[x(j), x(j-1), \dots, x(j-p+1)],$$

$h]$

可见,股市的动力演化模型可由上式表达。而上式表明,该系统后期的状态可由前期的若干状态来表达,因此确定函数关系 F 是问题的关键。

3 股市预测的进化神经网络模型

对上节给出的模型,神经网络的输入节点数由系统动力模型来确定,但由 n 阶微分方程描述的股市动力系统是一个未定问题,它涉及到对系统的演化规律的了解程度,对实际市场来说,这是一个非常困难的问题。转化到神经网络模型中就是时序延迟时步的确定问题,这是一个从时序构造网络训练样本的关键参数,对网络建模有举足轻重的作用,它严重地影响模型的预测能力。但目前却没有一个确定该参数的好方法,一般均由经验确定。

如上所述,建立股市时序的神经网络模型时,神经网络的结构确定是非常复杂的问题,而且也是一个关键的问题。因为,对此问题而言,网络结构的确定不但包括隐层结构的确定,而且还包括输入层结构的确定,这更加大了网络结构确定的难度。因此,网络结构的确定只能采用进化的方法进行。由于这里网络结构的确定是一个复杂的问题,作为初步研究,这里仅提出一种网络结构进化的进化神经网络算法来解决时序建模问题。算法中,网络的训练采用 MBP 算法(带惯量项修正的 BP 算法)进行。为了简化问题,并尽可能加大网络的泛化能力,这里只针对三层网络进行研究,从而网络结构的确定变成对输入节点数及隐层节点数的确定。由于 MBP 算法中包括两个算法参数,迭代步长 η 及惯量参数 α 需要确定,且它们的确定对 MBP 算法的影响很大。因此,对 η 、 α 两个参数也采用进化的方法进行确定。从而,这里的进化神经网络算法中,有四个量(输入节点数、隐层节点数、迭代步长及惯量参数)需要由进化确定。为了达到尽可能好的进化效果,本算法采用作者提出的免疫进化规划^[9]进行网络进化研究。为了研究方便、直观,这里采用进化神经网络的间接编码方式的进化结构进化神经网络模型。

以下给出新型进化神经网络模型的算法。

(1)分别给出网络的输入节点数、隐层节点数的搜索范围,并确定 MBP 算法的两个参数 η 、 α 的搜索区间。并给定进化神经网络的进化代数阈值,每代群体的个体数目,进化终止的误差阈值,神经网络的输出节点数,MBP 算法的迭代总数阈值及其迭代误差限。

(2)在输入节点及隐层节点的搜索范围中分别取出一个数,代表一个网络结构。再在 η 、 α 两个算法参数的搜索区间分别随机取一个阈值,代表一个 MBP 算法。把四个参数合并在一起,代表一个 MBP

算法下的一个网络结构,即一个个体。

(3)对群体中的每个个体,执行下列学习训练过程,以得到其适值:

①以其输入节点数及输出节点数为基础,把整个股市变化时序列划分成该结构下网络的训练样本,并记录样本总数;②把整个训练样本分成两部分,一部分作为训练学习样本,以得到非线性映射网络,另一部分作为测试样本,来测试网络的泛化能力;③初始化该网络个体的连接权值及节点阈值。执行下列各步,以得到该网络个体的适值。④设置 MBP 算法的迭代次数 $j=1$;⑤以测试样本集训练该网络个体,计算网络的均方误差 $E(j)$,并设它为整个训练过程的最小误差 $\min E=E(j)$ 。如果 $\min E$ 小于进化终止误差阈值,则把 $\min E$ 作为该网络个体的适值,转(3)处理下一个个体;⑥以训练样本集对该网络进行训练,如其训练误差小于设定的 MBP 算法的误差限,则以 $\min E$ 作为该网络个体的适值,转(3)处理下一个个体;⑦对该网络个体的权值及阈值用 MBP 算法进行一次调整;⑧ $j=j+1$,转回⑤;⑨如果 j 大于所设定的算法迭代总数阈值,则以 $\min E$ 作为该网络个体的适值,转(3)处理下一个个体。

(4)如果进化代数达到要求或终止误差阈值达到,则算法过程结束,以最后一代中的最佳个体作为搜索结果;否则,转(5)。

(5)个体变异。对群体中的每个个体进行变异处理,由于个体为具有不同数据类型结构型个体,从而对个体中每个参数做不同形式的变异。输入节点数及隐层节点数为整型数据,采用均匀型变异;而 MBP 算法的控制参数 η 、 α 为实型数据,采用自适应柯西变异。父代个体变异后形成子代群体。

(6)对于子代群体中的每个个体采用(3)中方法进行适值计算。

(7)对于子代群体及父代群体形成的总群体集合,采用基于浓度的选择方法进行选择,以形成新的子代群体。

(8)群体进化代数加1,并转回(4)。

4 实例应用

这里以上海股市1996年3月中旬至6月初一段行情为例进行说明,数据来源于上海股市行情,总收集到53组收市指数及成交量的时间序列,具体序列见图1、2所示。

把收集得到的收市指数及成交量时间序列代入提出的进化神经网络模型进行训练,其中48组用于计算,另外5组用于预测测试,48组计算序列中30组序列作为训练学习样本,余下的18组序列作为测试样本。经过训练及测试可以得到预测收市指数的网

络结构为7—10—1,而预测成交量的网络结构为9—18—3。其收市指数及成交量序列的预测结果见图1、2所示。

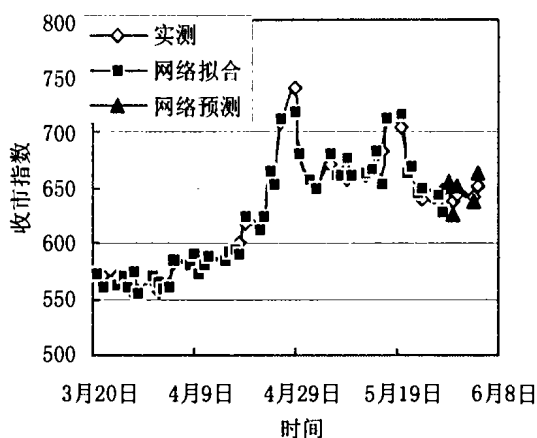


图1 股市收市指数实测与预测结果

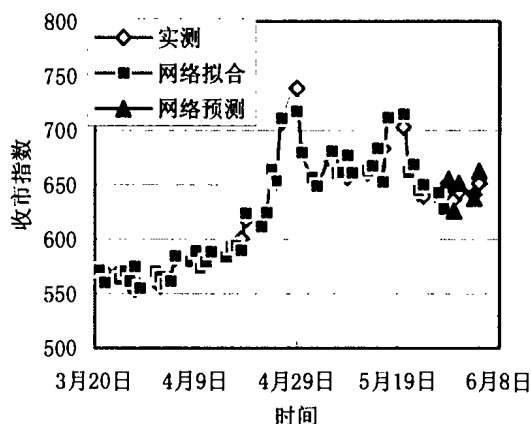


图2 股市成交量实测及预测结果

由上图1、2可见,本文提出的进化神经网络模型进行股市预测不但拟合效果良好,而且预测结果也很理想,是进行股市预测的一种较理想的方法。

结语 股票市场是一个受众多因素影响的非线性动态系统,对其演化规律进行预测对于证券投资

及市场管理均具有重大的理论及现实意义。本文通过对股市动态演化的数学描述,提出了一种进行股市预测的进化神经网络方法,并用沪市1996年的一段时间序列进行了方法验证,结果证明所提方法不但拟合效果好,而且预测结果也很理想,是进行股市预测的一种较理想的方法。

当然,任何模型、任何方法都只是从某些方面来反映或描述事物的特征,不可避免地具有一定的局限性。而且,由于我国股市建立时间比较短,市场法规尚不完善,会计制度、审计制度及上市公司信息披露等方面与国际惯例相差甚远。股市的不成熟常常使得国家政策、人为因素、有关消息及其他因素极易导致股市的不规则突变,这大大增加了技术预测的难度。因而,技术预测必须与基本面、政策面、消息面、资金面等相结合才能更有效地规避操作风险。同时,我们相信,传统的基本分析和技术分析与非线性科学、现代高新技术相结合将是证券投资分析方法发展的必然趋势。

参考文献

- 1 吴奉,陈宏立. 经济预测与决策方法. 广州:暨南大学出版社,1991
- 2 陈茂申. 股市技术分析. 合肥:安徽人民出版社,1995
- 3 高玮. 非线性时序建模预测方法探讨. 清华大学学报,2000,40(S1):6~10
- 4 高玮. 非线性信号预测模型研究. 计算机研究与发展,2000,37(S):254~258
- 5 Refenes A N, Zaprana A, Francis G. Stock performance modeling using neural networks: a comparative with regression model. Neural Network, 1994, 7(2): 375~388
- 6 Schomeburg. Stock Price Predication Using Neural Networks: An Empirical Test. Neurocomputing, 1991, 2(1): 34~38
- 7 赵建, 郭永革, 黄炯. 基于神经网络的股市预测. 计算机研究与发展, 1996, 33(9): 692~697
- 8 刘新勇, 贺江峰, 孟祥泽, 等. 基于神经网络的股市预测. 南开大学学报, 1998, 31(3): 39~44
- 9 Gao Wei. Study on new evolutionary neural network. International Conference on Machine Learning and Cybernetics, 2003. 1287~1292

(上接第173页)

的规模。因为应用框架分割了模块之间的耦合性,各个开发者只需要专注于各自的领域就可以了,有效地提高了开发效率。软件的规模越大,框架带来的优越性也越明显。

如前面所述,在设计阶段就完成的一个应用框架不会是完善的,但是早期合理的抽象才会使应用框架具有普适性,才能避免在实现阶段或者 QA 阶段时底层框架的频繁大规模修改。因此评价 Web 应用框架是否成功,大约要从两方面着手,一方面是复用度,另一方是应用框架的历次修改对开发造成的影响。

结论 本文提出了针对 J2EE 平台的多层应用框架,抽象出了一个企业业务相关的用户框架层,把

用户框架层分为通用 Web 视图框架和企业业务相关的构件模型,进而给出了框架的实现方式。从开发角度说明了本应用框架模型给应用开发流程带来的变化,框架本身不但可以提供良好的复用性,而且使开发流程更顺畅。

参考文献

- 1 Gellersen H-W, Gaedke M. Object-Oriented Web application development. IEEE Internet Computing, 1999, 3(1): 60~68
- 2 Hoque R, Sharma T. Programming Web Component. New York: McGraw-Hill, 1998
- 3 Johnson RE, Foote B. Designing reusable classes. Journal of Object-Oriented Programming, 1988, 1(2): 22~35
- 4 张波, 冯玉琳, 黄涛. 基于对象视图模型 WebView 的 Web 应用框架. 软件学报, 2002, 13(10): 1985~1990
- 5 王柏, 王红樱, 邹华. 分布计算环境. 北京: 北京邮电大学出版社, 2000