

基于改进的 Elman 神经网络的股价预测模型

余 健, 郭 平

(北京师范大学 信息科学与技术学院, 北京 100875)

摘 要: Elman 神经网络是一种典型的回归神经网络, 比前向神经网络具有更强的计算能力, 具有适应时变特性的能力, 因而非常适用于对股市这一类极其复杂的非线性动力学系统进行预测。文中以深市 A 股中的个股中集集团(股票代码: 000039)的共 180 天的实际收盘价的时间序列作为预测对象, 提出基于改进的 Elman 神经网络的个股价格预测模型, 实验结果取得较高的预测精度、较为稳定的预测效果和较快的收敛速度。这表明该预测模型对于个股价格的短期预测是可行和有效的。

关键词: Elman; 神经网络; 预测

中图分类号: TP183

文献标识码: A

文章编号: 1673-629X(2008)03-0043-03

Stock Price Forecasting Model Based on Improved Elman Neural Network

YU Jian, GUO Ping

(College of Information Science and Technology, Beijing Normal University, Beijing 100875, China)

Abstract: Elman neural network is a classical kind of recurrent neural network. It has greater calculating capability than feed forward neural network. And its characteristic of adapting time variability especially adapts to complicated nonlinear dynamics system forecasting like stock market. Uses the market price of Zhongji company (No. 000039) in Shenzhen stock market A index as forecasting object and improved Elman neural network as forecasting model. The experiment results get higher precision, steadier forecasting effect and more rapid convergence speed. These show that model is feasible and efficient to forecast short term stock price.

Key words: Elman; neural network; forecasting model

0 引言

Elman 神经网络是一种典型的回归神经网络, 它是 Elman 于 1990 年提出, 该模型是在前馈神经网络的隐层中增加一个承接层, 作为一步延时算子, 达到记忆的目的, 从而使系统具有适应时变特性的能力^[1,2]。因而它是一种反馈动力学系统, 能直接反映动态过程系统的特性, 比前向神经网络具有更强的计算能力, 非常适用于对股市这一类极其复杂的非线性动力学系统建立时间序列的预测模型。

1 Elman 神经网络股价预测模型

1.1 Elman 神经网络结构

Elman 神经网络一般分为 4 层: 输入层、隐层、输

出层和承接层(又称上下文层), 如图 1 所示。其输入层、隐层和输出层的连接类似于前馈网络, 输入层的单元仅起信号传输作用, 输出层单元起线性加权作用, 隐层单元的传递函数可采用线性或非线性函数, 上下文层用来记忆隐层单元前一时刻的输出值, 可以认为是一步延时算子^[3]。

Elman 型神经网络的特点是隐层的输出通过承接层的延迟与存储, 自联到隐层的输入, 这种自联方式使对历史状态的数据具有敏感性, 内部反馈网络的加入增加了网络本身处理动态信息的能力, 从而达到动态建模的目的。

改进的 Elman 神经网络^[3,4]是在基本的 Elman 神经网络的承接层单元上增加了一个固定增益为 a 的自反馈连接, 使承接层单元在 k 时刻的输出等于隐层在 $k-1$ 时刻的输出加上承接层在 $k-1$ 时刻输出值的 a 倍。其结构图如图 1 所示。

其中, a 为自反馈增益因子。在第 k 个时刻, Elman 网络的非线性状态空间表达式为:

收稿日期: 2007-06-24

作者简介: 余 健(1977-), 男, 广东潮州人, 讲师, 硕士, 研究方向为神经网络及其应用、智能计算; 郭 平, 教授, 博士生导师, IEEE 高级会员, 研究方向为神经网络、模式识别、智能信息处理等。

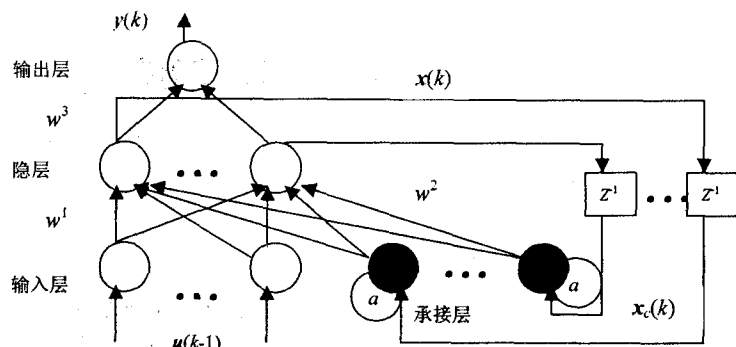


图 1 改进的 Elman 神经网络结构

$$x(k) = f(w^1(u(k-1)) + w^2 x_c(k))$$

$$x_c(k) = x(k-1) + a * x_c(k-1)$$

$$y(k) = g(w^3 x(k))$$

1.2 改进的 Elman 神经网络的学习过程

以图 1 为例, Z^{-1} 是一步延时算子, u 表示 r 维网络输入向量, x 表示 n 维隐层单元输出向量, x_c 表示 n 维反馈状态向量, y 表示 Elman 网络 m 维输出向量。 w^1 为输入层到隐层的连接权值, w^2 为上下文层到输出层的连接权值, w^3 为隐层到输出层的连接权值。 $f(\cdot)$ 为隐层单元的传递函数, 常采用 S 型函数, $g(\cdot)$ 为输出单元的线性传递函数, 是隐层输出的线性组合。

Elman 神经网络实质上是在 BP 神经网络的基础上, 通过引入反馈信号, 存储内部状态使其具备映射动态特征的功能, 从而使系统具有适应时变特性的能力, 使它的时间序列的预测性能较 BP 神经网络更好。对于网络权值的修正, Elman 网络采用 BP 算法, 学习指标函数采用误差平方和函数^[5]:

$$E(w) = \sum_{k=1}^n [y_k(w) - \hat{y}_k(w)]^2$$

2 仿真实验

2.1 样本的生成

在股市中, 影响股市交易的因素有很多, 衡量股票价格变化的指标也很多: 如大盘指数、开盘价、收盘价、升跌幅度、成交量等^[6]。为了研究方便, 这里, 选用深证 A 股中的个股——中集集团(000039)2006 年 1 月 6 日起至 12 月 5 日, 共 180 个股票交易日的收盘价的时间序列作为学习的数据样本和预测对象, 即共有 180 个数据点。

2.2 归一化处理

为使输入输出符合神经网络的要求, 在训练网络之前必须对数据进行归一化处理, 即将数据处理为区间 $[0, 1]$ 之间的数据, 归一化方法有多种形式, 这里采用如下公式:

$$\hat{X} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

其中, X_{\min} 和 X_{\max} 分别为样本中股价最小值和最大值, X 为某个股价, \hat{X} 为得到的股价归一化值。

在实际的股市投资中, 当得到神经网络对股价的预测值之后, 需要进行反归一化处理, 得到需要的股价:

$$X = \hat{X} * (X_{\max} - X_{\min}) + X_{\min}$$

2.3 数据的训练和预测方式

将股中的数据看做一个时间序列进行处理, 这里假定股市中有时间序列^[6,7]: $X = \{x_i \in R, i = 1, 2, \dots, L\}$

现在希望通过序列的前 N 天的值, 预测出后 M 天的值。这里采用滚动预测法:

设预测阶数为 n , 用 $X_{i-1}, X_{i-2}, \dots, X_{i-n}$ 的信息数据来预测 i 时刻的值, 则将 $X_{i-1}, X_{i-2}, \dots, X_{i-n}$ 作为数据信息输入。而 X_i 的值作为预测的期望值。这样就可以构造如下的学习样本或检验样本:

输入: (X_1, X_2, \dots, X_n) ; 输出: X_{n+1}

输入: $(X_2, X_3, \dots, X_{n+1})$; 输出: X_{n+2}

...

输入: $(X_{n-m}, X_{n-m+1}, \dots, X_{n-1})$; 输出: X_n

即先进行单步预测, 然后将输出反馈给输入端作为网络输入的一部分, 预测未来 q 个时刻的取值)。

合理确定神经网络的结构是预测性能的基础。而其中, 最重要的是解决网络的隐节点个数问题。隐单元在学习过程中起着对输入向量进行特征提取的作用, 其个数的多少对网络的训练效果及网络的推广能力有着直接的影响。隐单元数目过少严重降低了网络的分类能力。但是当隐单元数目增加到一定程度后, 迭代步数并不随着隐单元数目的进一步增加而明显地减少。另一方面, 过多的隐单元数会增加网络负荷, 降低系统效率, 更严重的是还可能降低网络的推广能力。如何确定合适的隐层神经元的数目, 目前还没有明确而又广泛适用的定论^[8]。

目前多数的作法仍是通过试凑法来确定, 本文同样采用这种方法, 经过反复多次实验, 将隐层神经元的数目设为 11 个。网络的结构采用 3-11-1 形式, 即含有 3 个输入层节点, 11 个隐层节点, 1 个输出节点的网络结构。输入层节点的输入为连续 3 日的实际股价, 输出为第 4 日的预测股价。

这样, 就可以采用前 3 个数据预测下一个数据, 将数据集划分成 177 个样本, 前 167 个样本作为学习和训练样本, 后 10 个样本作为预测检验样本。

2.4 个股价预测

因此, 这里采用改进的 Elman 网络结构, 同时采用

2 种改进的 BP 算法来修正 Elman 的权值,进行股市时间序列预测,由于 BP 算法固有的随机性,所以基本上每一次的训练结果都是不同的,所以文中每一个实验都进行 10 次训练,取其 MSE 最小的预测结果。设置所有神经网络的训练次数均为 5000 次,目标误差设为 0.0001,采用滚动预测的方法,在配置为 Celeron4 2.0GHz,内存 512M 的 PC 机,得到的预测结果见表 1。

表 1 改进的 Elman 网络各算法预测效果对比表

实验序号	采用 BP 算法	MSE	SSE	MAE	运行时间(秒)
1	有动量和自适应 lr 的梯度下降法	0.0021	0.0207	0.0361	67.0523
2	弹性梯度下降法	0.0021	0.0212	0.0366	70.0274

(1)对于实验 1 中,预测检验曲线如图 2 所示。图 2 中实线表示实际值,虚线表示 Elman 网络输出的预测值,可以看到网络的预测值的走势与实际值较好地吻合。实验 1 真实值与预测值对比见表 2。

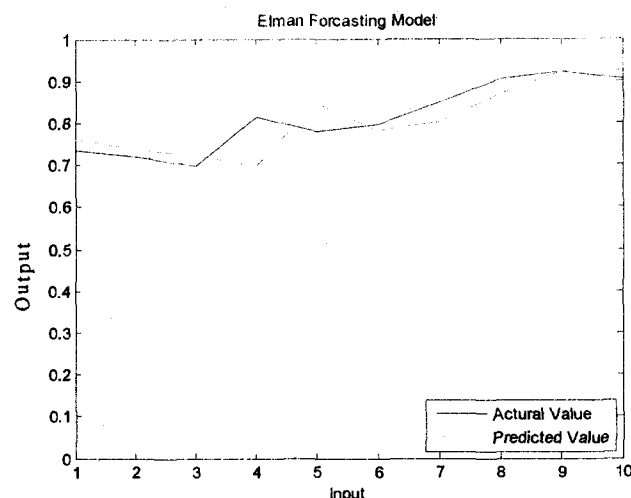


图 2 实验 1 的预测检验曲线图

表 2 实验 1 真实值与预测值对比表
(单位:元,年份:2006 年)

日期	11月22日	11月23日	11月24日	11月27日	11月28日	11月29日	11月30日	12月1日	12月4日	12月5日
真实值	14.75	14.65	14.44	15.40	15.13	15.27	15.70	16.15	16.29	16.16
预测值	15.00	15.05	15.05	14.89	15.47	15.21	15.49	15.66	15.90	16.06

(2)对于实验 2 中,预测检验曲线如图 3 所示。图 3 中实线表示实际值,虚线表示 Elman 网络输出的预测值,可以看到网络的预测值的走势与实际值较好地吻合。实验 2 真实值与预测值对比见表 3。

表 3 实验 2 真实值与预测值对比表
(单位:元,年份:2006 年)

日期	11月22日	11月23日	11月24日	11月27日	11月28日	11月29日	11月30日	12月1日	12月4日	12月5日
真实值	14.75	14.65	14.44	15.40	15.13	15.27	15.70	16.15	16.29	16.16
预测值	15.02	15.05	15.02	14.86	15.50	15.21	15.48	15.69	15.93	16.04

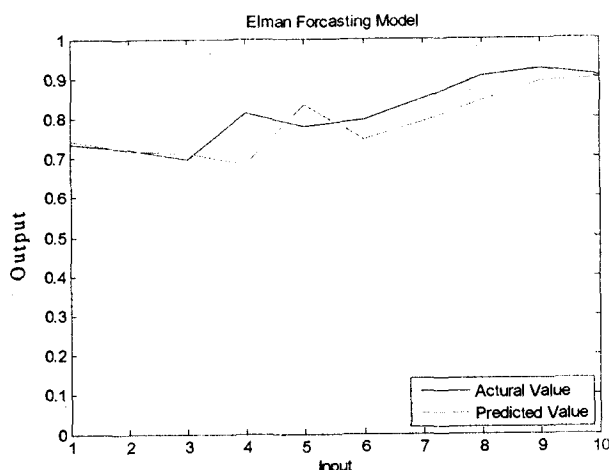


图 3 实验 2 的预测检验曲线图

3 结束语

文中研究了 Elman 神经网络的特点、改进结构、改进算法和学习过程,并选用了深市 A 股中的个股——中集集团(000039)共 180 个股票交易日的收盘价作为时间序列预测的样本。最后使用 Elman 神经网络进行股市预测实验,给出相应的实验结果和时序预测效果评价。Elman 网络的预测效果较为稳定,收敛速度也较快。

实验结果表明,基于改进的 Elman 神经网络的股价时间序列预测模型对于股市进行的短期预测是行之有效的。

参考文献:

- [1] 飞思科技产品研发中心. 神经网络理论与 MATLAB7 实现[M]. 北京:电子工业出版社,2005.
- [2] 程 球. 人工神经网络设计算法的研究[J]. 微机发展, 2005,15(12):61-62.
- [3] Gao X Z, Ovaska S J. Genetic Algorithm Training of Elman Neural Network in Motor Fault Detection[J]. Neural Comput & Applic, 2002, 11:37-44.
- [4] 周洪煜. 基于改进 Elman 网络的自适应预测函数控制[J]. 测控技术, 2003, 22(4):41-42.
- [5] 王 杰, 闫东伟. 提高预测精度的 ELMAN 和 SOM 神经网络组合[J]. 系统工程与电子技术, 2004, 26(12):1943-1945.
- [6] 杨 颖, 陈德华. 基于小波神经网络的时间序列流数据的研究[J]. 计算机技术与发展, 2006, 16(6):193-195.
- [7] 韩卫华, 宁佐贵. 时间序列预测的动态神经网络方法[J]. 微机发展, 2004, 14(9):40-41.
- [8] 张玉林. 神经网络在股市预测中的建模及应用[D]. 大连:大连理工大学, 2004.