

文章编号:1672-058X(2012)05-0050-05

线性神经网络模型在新上证综指的应用研究^{*}

王 奎

(重庆工商大学 数学与统计学院,重庆 400067)

摘 要:利用线性神经网络模型对“新上证综指(000017)”进行拟合预测,选取从“新上证综指”开始发行月份(2006 年 1 月)开始到 2011 年 6 月的月度数据,共计 66 个,用前 62 个做训练组,最后 4 个数据做预测组,通过比较不同滞后窗口模型的误差平方和,选择适当的窗口数为最优模型,为了提高模拟的效果,对模型的初始数据进行优化,然后进行预测分析;结果显示,拟合效果很好,除 6 月份股市波动稍大,其他月份拟合误差不到 3%,阐释了股票市场的短期可预测性。

关键词:新上证指数;线性神经网络模型;最优模型

中图分类号:F224.9

文献标志码:A

上证指数包含的成分股为上交所所有的股票,新综指的成分股是指上交所已经完成股改的股票,就是 G 股(A 股中已经完成股改的股票)。由于 A 股股权分置的原因,综合指数最大的缺陷就是真实性的缺失^[1],所以综指在严格意义上,对于判断市场的真实走势是有一定误导性。股市预测是经济预测的一个研究分支,股市价格的未来走势是否具有预测性与市场的有效性密切相关。如果股票市场是符合有效市场假说的,也即股票价格服从随机游走模型,则股票价格的预测就毫无意义。随着我国证券市场的发展,广大经济工作者对我国股票价格的预测也进行了深入的研究。先后有许多学者对我证券市场的有效性和可预测性进行了研究。从目前学者们对我国股票市场特征的研究来看,诸多学者认为我国股票市场还没有达到弱势有效^[2],也就是说用股票交易的历史信息在一定程度上可以预测其未来的走势。特别是在短期内,是具有一定的可预测性的。目前证券投资已经成为社会生活的一个重要部分,证券市场的运行状况对我国实体经济的影响逐渐加深,正在真正地发挥着国民经济“晴雨表”和“报警器”的作用。深刻理解股票市场的运行规律,预测分析股票价格的未来走势,无论是对广大投资者降低投资风险还是宏观经济管理部门的宏观调控,保障我国证券市场的健康持续发展,都有着十分重要的意义^[3]。

1 模型介绍

1.1 线性神经网络的定义^[4]

线性神经网络是最简单的一种神经元网络,由一个或多个线性神经元构成,线性神经网络每个神经元

收稿日期:2011-11-22;修回日期:2011-12-08.

* 基金项目:国家“十一五”科技支撑计划重大项目(2006BAJ05A06);重庆市科委重点攻关项目(2008AC0043);重庆工商大学创新型项目(yjscxx2011-7);重庆市南岸区科技计划项目:重庆市南岸区产业竞争力决策分析系统开发及应用示范.

作者简介:王奎(1986-),男,湖北襄阳人,硕士研究生,从事统计学研究.

的传递函数为线性函数,因此,线性神经网络的输出可取任意值。

1.2 线性神经网络的训练学习过程

线性神经网络的训练过程一般分为3个阶段,一是根据给定的输入向量计算网络的输出向量 $Y = WX + B$,与期望输出向量之间的误差 E ,其中, Y 为输出列向量, X 为输入列向量, W 为权重矩阵, B 为阈值列向量。若用 T 表示网络的目标输出向量,则线性网络的输出误差函数定义为:

$$E(W, B) = (T - Y)^2 / 2 = (T - WX - B)^2 / 2 \quad (1)$$

由式(1)可知,当网络的输入和目标输出值给定时,其输出误差取决于网络的权重和阈值,而且其误差是一个具有抛物面型的结构,只有一个误差最小值。二是将网络输出误差的平方与其期望误差相比较,如果其值小于期望误差,或训练已经达到事先设定的最大训练次数,则终止训练,否则继续进行训练。三是采用 Window-Hoff 学习规则,计算新的权值和阈值,并返回到第一步。对于线性神经网络的学习规则,可以采用 Window-Hoff 或者 LMS 来训练网络,调整网络的权重和阈值。它采用梯度下降法,不断修正权重和阈值,使输出误差达到最小值,按照这种方法训练时,网络的下一个权重向量等于现在的权重(或阈值)向量加一个正比于均方误差梯度负值的变化量,为了便于实现,取当前位置上的误差函数的梯度作为均方误差梯度的估计,则 Window-Hoff 学习规则的权重和阈值的修正公式可分别表示为:

$$W_{k+1} = W_k - \eta * \partial E_k(W, B) / \partial W = W_k + \eta(T_k - Y_k)X_k^T \quad (2)$$

$$B_{k+1} = B_k - \eta * \partial E_k(W, B) / \partial B = B_k + \eta(T_k - Y_k) \quad (3)$$

其中, η 为学习速率,它决定网络的收敛速度,当 η 增大时,网络收敛速度加快,但是当 η 太大时,学习过程变得不稳定,而且误差会增大,因此要适当地选取学习速率(通常取 $0.01 \leq \eta \leq 1$)。Window-Hoff 学习规则训练网络时,针对每一对训练样本,按照式(2)、(3)修正权重和阈值,直到达到规定的误差精度或者规定的训练次数。在给定的数据中,将采用前117个样本点的数据进行训练,通过训练学习过程拟合模型。

2 实证分析

2.1 数据选取与处理

选择了从2006年1月4日到2011年6月14日的沪市交易所新综指(000017)的月度收盘价,共66个数据。其中前62数据作为训练组,进行模型的拟合;后面4个数据作为预测组,进行比较。

2.2 模型的拟合

利用下述程序,对训练组的数据进行模拟运算,通过计算误差平方和,发现在滞后窗口为5时的误差较小,且滞后窗口数比较适宜,所以选择滞后窗口为5的线性神经网络进行拟合。为提高拟合精度,对前5个数据的初始值进行处理,从第6个数据进行拟合,程序如下。

```
L = [数据];
Q = length(L);
P = zeros(5, Q);
T = L';
P(1, 2:Q) = T(1, 1:(Q-1));
P(2, 3:Q) = T(1, 1:(Q-2));
P(3, 4:Q) = T(1, 1:(Q-3));
P(4, 5:Q) = T(1, 1:(Q-4));
P(5, 6:Q) = T(1, 1:(Q-5));

figure(1);
plot((1:Q), T, (1:Q), T, 'd');
xlabel('时间顺序个数');
label('新综指月度收盘价')
P_1 = zeros(5, 61);
T_1 = zeros(1, 61);
for i = 1:61
    P_1(:, i) = P(:, i+5);
    T_1(:, i) = T(:, i+5);
end
```

```

end
net = newlind(P_1,T_1);
a = sim(net,P_1);
Q_1 = length(L) - 5
figure(2);
plot((1:Q_1),a,(1:Q_1),a,'d');
xlabel('时间顺序个数');
ylabel('模拟的新综指月度收盘价')
figure(3);
plot((1:Q_1),T_1,'g',(1:Q_1),T_1,'d',(1:Q_1),
      a,'r',(1:Q_1),a,'*');
      error_sum = 0;
for i = 1:61
      error_sum = error_sum + abs(T_1(1,i) - a(1,i))^2;
end
error_sum
PP = zeros(5,57);
TT = zeros(1,57);
for j = 1:57
      PP(:,j) = P_1(:,j);
      TT(:,j) = T_1(:,j);
end
net = newlind(PP,TT);
aa = sim(net,PP);
error_sum_57 = 0;
for i = 1:57
      error_sum_57 = error_sum_57 + abs(TT(1,i) - aa(1,
      i))^2;
end
error_sum_57
a4 = zeros(1,4);
a4(1,1) = sim(net,P_1(:,58));
a4(1,2) = sim(net,P_1(:,59));
a4(1,3) = sim(net,P_1(:,60));
a4(1,4) = sim(net,P_1(:,61));
percentage_1 = abs((a4(1,1) - T_1(1,58))/ T(1,
58)) * 100
percentage_2 = abs((a4(1,2) - T_1(1,59))/ T(1,
59)) * 100
percentage_3 = abs((a4(1,3) - T_1(1,60))/ T(1,
60)) * 100
percentage_4 = abs((a4(1,4) - T_1(1,61))/ T(1,
61)) * 100
L_pre = [aa a4];
figure(4);
plot((1:Q_1),T_1(1:Q_1),'g',(1:Q_1),L_pre(1:Q_
1),'r',(Q_1-3):Q_1,T_1((Q_1-3):Q_1),'d',(Q_1-
3):Q_1,L_pre((Q_1-3):Q_1),'*')

```

利用上述程序拟合,结果如图 1 所示,通过观察真实股票曲线和模拟的股票曲线,发现模拟序列与真实序列差别很小,模拟效果良好。

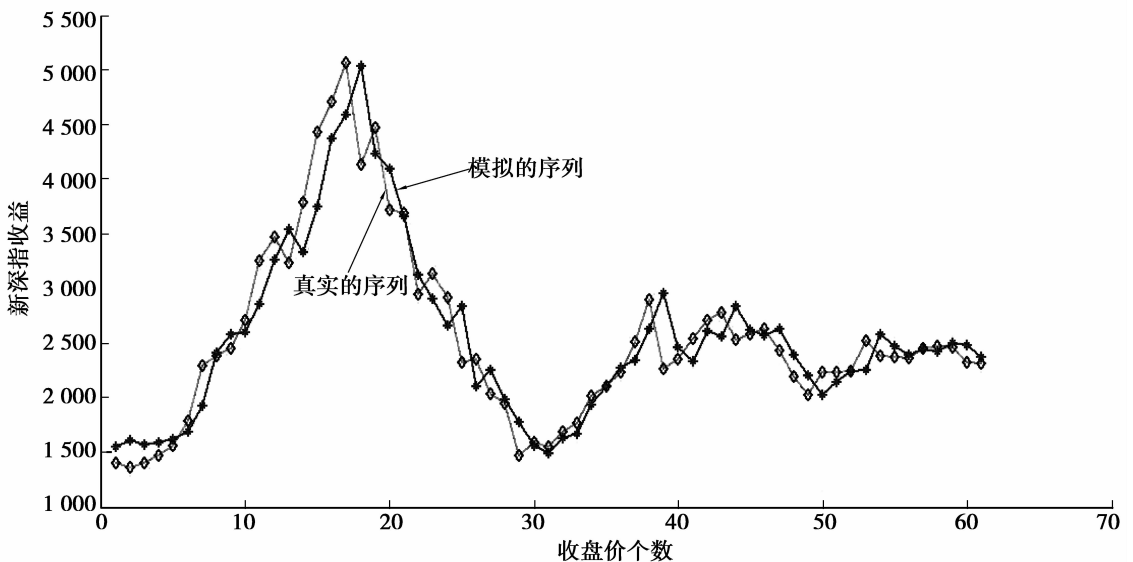


图 1 新综指原序列和模拟序列图

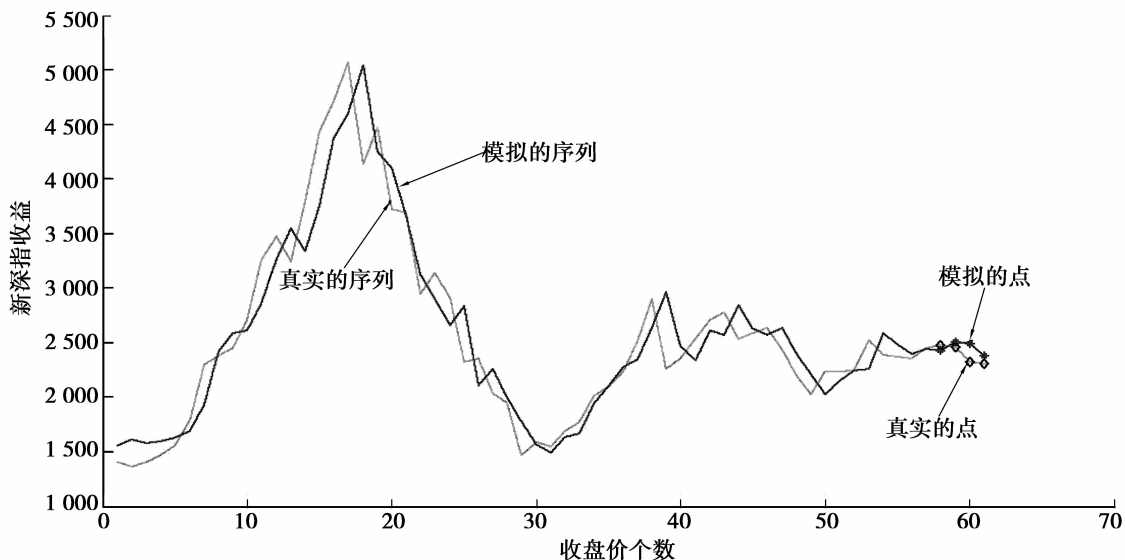


图 2 预测值与实际值的比较

然后对 2011 年 2 月至 2011 年 6 月这 4 个月度的新上证综指收盘价进行预测,并与实际值进行对比,以说明预测结果的效果,预测结果表明(图 2)。运用该模型对未来 4 个月度数据进行预测的结果与实际值比较吻合,误差控制在了一定的范围之类,说明预测结果较为理想。若进一步对未来的收盘价进行预测,预测结果与实际值的偏离有可能会增大,因为时间序列有其自身的特殊性,若预测时间太长,预测精度下降很快,参考价值就会降低。因此,虽然预测结果表明该模型的预测能力较好,但在选择预测的时间上仍然不宜太长,以降低预测误差的范围^[5]。

2.3 误差的分析

为了进一步具体说明线性神经网络模型预测能力的效果,计算最后 4 点的预测 MAPE(均值绝对百分比误差),计算公式

$$MAPE = \frac{1}{m} \sum_{i=1}^m \left| \frac{y_a - y_p}{y_a} \right| \times 100\% \tag{4}$$

表 1 预测值均值绝对百分比误差表

编号	实际(y _a)	预测(y _p)	误差值%
1	2 474. 71	2 426. 1	1. 929 2
2	2 460. 88	2 507. 8	1. 966 5
3	2 318. 97	2 486. 7	7. 073 9
4	2 308. 76	2 377. 7	2. 925 4

从 4 个点误差的计算结果可以看出(表 1),线性神经网络的预测误差最大也不只有 7%,即表明 2011 年 5 月份股市动荡较大,但经过调整,在 6 月份误差又降低到只有不到 3%,这与现实生活 2011 年 5、6 月股市动荡较大的一致。

3 结 论

综上所述,股票市场长期预测虽无定论,但短期股票交易的历史信息在一定程度上还是可以预测其未来走势的,也是上文所说的“我国股票市场还没有达到弱势有效”的印证,利用了沪市交易所新综指(000017)的月度收盘价实证了这一点。在非线性数据预测方面,神经网络模型相对于传统的计量经济模型更加有优势^[4],通过选择适合的滞后窗口数,拟合效果显著,除 6 月份股市波动稍大,其他月份拟合误差不到 3%,充分阐释了股票市场的短期可预测性。

参考文献:

- [1] 傅波,朱雪仪. 上证综合指数失真分析[J]. 中国证券期货,2009(07):4-7
- [2] 闻新,周露. MATLAB 神经网络仿真与应用[M]. 北京:科学出版社,2003
- [3] 贺本岚. 股票价格预测的最优选择模型[J]. 统计与决策,2008(06):135-136
- [4] 赵国顺. 基于时间序列分析的股票价格趋势预测研究[D]. 厦门:厦门大学,2009
- [5] 常松,何建敏. 基于小波包和神经网络的股票价格预测模型[J]. 中国管理科学,2001(08):8-15

Application of Linear Neural Network Model to New Shanghai Composite Index Predication

WANG Kui

(School of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China)

Abstract: Linear neural network model is used to make fitting predication for New Shanghai Composite Index (000017). By selecting total 66 data of the monthly data of New Shanghai Composite Index from January, 2006 to June, 2011, by using former 62 data as training group and latter 4 data as predication group, through comparing error square sum of the model in different time-lag windows, proper window numbers are selected as optimal model, the initial data of the model are optimized in order to promote simulation effect and then predication analysis is conducted. Results show that fitting effect is very good, except that the vibration of stock market is bigger in June, the monthly fitting error of other months is less than 3 percent. Short-term predication feasibility for stock market is elaborated.

Key words: New Shanghai Stock Index; linear neural network model; optimal model

责任编辑:李翠薇