

人工神经网络在股票市场预测中的应用

冯俊秋 李 挥

(深圳大学工程技术学院,深圳 518060;第一作者 22岁,男,98届毕业生)

摘 要 运用人工神经网络理论,对股票市场进行动态建模和仿真预测实验,仿真实验结果表明,低阶多层改进 BP模型能预测深圳股市趋势且取得良好结果.

关键词 神经网络;非线性系统;误差反传(BP)算法;金融预测

中图分类号 TP387

引 言

目前预测股票市场趋势的主要方法是传统的技术分析法,如:运用 K线图、移动平均线和 OBV 线^[1,2].技术分析立足于股票交易中历史事件重现的假设上,以人为判断和将过去的股价图相比较得到最有可能的结果.如图 1,我们可以看出股市上股价的大致变化,其走势是一个受政治、经济等多方面因素影响^[2]的离散时间序列(图 2),用简单的统计数学模型来拟合很困难.另则,技术分析存在着一定的时滞现象^[3].中国股市是尚未完全规范的市场,其投机行为多,理性投资少,用理想假设进行预测往往不能得到预期结果.

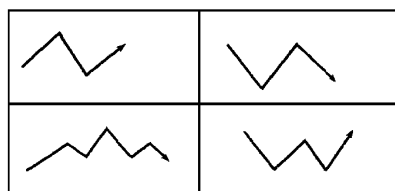


图 1 股票市场的 4种基本趋势

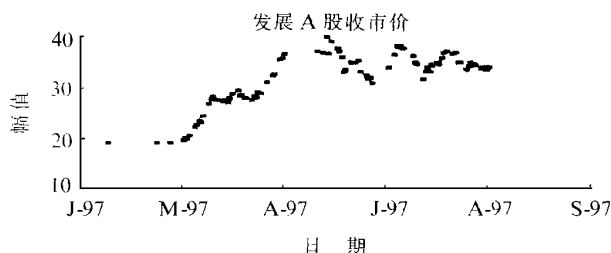


图 2 股票市场是离散时间序列

神经网络具有较强大的自我学习修正误差的能力,理论上证明神经网络能够逼近任何非线性系统^[4],为此,神经网络在国际上已广泛用于金融分析和预测,取得较好效果^[5,6].基于上述事实,目前国内已有一些学者采用神经网络进行股票市场预测的研究.本文试图运用改进多层低阶 BP (Back Propagation)神经网络模型对深圳股市进行预测.仿真实验结果表明 3层低阶改进 BP神经网络能有效地预测中国股市.

1 神经网络模型

1.1 BP算法

Rumelhart 1985年发展了基于多层感知器的前馈神经网络模型 - BP模型,它实现了输入到输出的一个高度非线性映射, $F: R^n \xrightarrow{f(\cdot)} R^m \quad Y = F(x)$, 拓扑如图

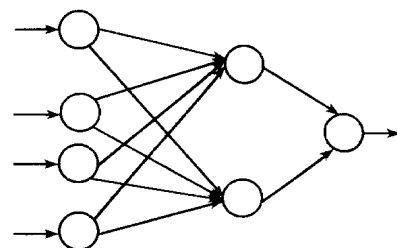


图 3 三层 BP 神经网络拓扑结构

3. BP算法是目前应用最为广泛的神经网络模型,为简明起见,有关算法详见文 [7].

1.2 BP算法的特点

BP算法的优点是具有很强的非线性映射能力和柔性的网络结构. BP网络能够通过学习带正确答案的实例集自动提取“合理的”求解规则.

在实际应用中误差逆传播神经网络存在以下主要不足:

1.2.1 学习收敛速度缓慢

即使一个简单的问题,网络也需要几百次甚至上万次的学习才能收敛. BP算法本质上是梯度下降法,其最小化目标函数非常复杂(其等值面远不是超球面),因此必然会出现“锯齿形现象”,这决定了 BP算法效率较低.

1.2.2 不能保证收敛到全局最小点

由于 BP学习规则采用了按误差函数梯度下降的方向进行收敛,这就意味着由能量函数构成的连接权空间不仅只有一个极小点的抛物面,而且存有多个局部极小点的超曲面.因此, BP网络的收敛过程很可能在遇到局部极小极值时便被“冻结”,而无法最终收敛到全局最小点.

1.2.3 网络隐含层数及隐含层单元数的选取尚无统一的理论指导

网络隐含层数及各隐含层单元数的选取一般根据经验确定.因此,网络往往具有很大的冗余性,也增加了网络学习的时间.

1.3 BP算法的改进

如何加速 BP网络的收敛速度和尽量避免陷入局部最小点是目前研究的热点问题.限于篇幅,这里仅介绍本文所采用的改进算法.

1.3.1 动量校正法

每一次对连接权或输出阈值进行校正时,按一定比例加上前一次学习时的校正量,即动量项,由此加速网络学习的收敛速度.具体作法为

$$w_{ij}(k+1) = w_{ij}(k) + \eta \Delta w_{ij}(k) + \alpha \Delta w_{ij}(k-1) \quad (1)$$

式中, $w_{ij}(k+1)$ 为本次校正量, $w_{ij}(k)$ 为前次校正量, α 为动量因子 ($0 < \alpha < 1$). 由 (1) 式可知,当前一次的校正量过调时,动量项与本次计算所得误差校正项符号相反,使得本次实际校正量减小,起到减小振荡的作用.而当前一次校正量欠调时,动量项与本次计算所得误差校正项符号相同,本次实际校正量增加,起到加速校正的作用.

1.3.2 改进动量校正法

分析表明,较大的 α 虽可改善网络收敛速度,但对提高网络的收敛精度和摆脱局部极小没有任何积极效果.为此,文 [6] 提出了如下改进方法.

考虑到动量因子对 BP 网络的加速收敛和平滑作用,同时考虑动量误调噪声的随机特性,可将 (1) 式中的动量因子 T 设计成为误差的函数,并使其成为具有衰特性的随机变量从而随着算法的收敛也趋于零.这样设计的动量因子不仅可保留动量的平滑作用,且在算法收敛后它已变为零,从而消除了动量误调噪声,改变后的权系数更新公式为

$$w_{ij}(k+1) = w_{ij}(k) + \eta \nabla W_j(k) x_j(k) + T_j(k) \Delta w_{ij}(k) \quad (2)$$

式中,因子 $T_j(k)$ 为一衰减的随机变量,可以推出 $T_j(k) = \bar{W}_j(k) / \sum_{m=1}^{k-1} \bar{W}_j(m)$,当算法收敛时,有 $\lim_{k \rightarrow \infty} T_j(k) = 0$,说明算法收敛到稳定点附近时 $T_j(k)$ 已变为 0,从而完全消除了动量因子引入的误调噪声.

1.4 BP 算法的实现

根据 BP 模型理论,作者采用 C 语言进行计算机动态仿真建模,3 层 BP 神经网络输入层神经元数为 L ,隐含层神经元数为 M ,输出层神经元数为 N ,训练模式为 P 对.下面给出了整个 BP 网络学习过程的具体算法步骤:

- ① 给各个连接权 $\{w_{ij}\}$ 、 $\{w_{jk}\}$ 及阈值 $\{\theta_j\}$ 、 $\{\theta_k\}$ 赋予 $(-1, +1)$ 间的初始化随机值;
- ② 取一模式对 $A_n = (a_{n1}, \dots, a_{nL})$, $d_n = (d_{n1}, \dots, d_{nN})$ 提供给网络;
- ③ 用输入模式 $A_n = (a_{n1}, \dots, a_{nL})$ 连接权 $\{w_{ij}\}$ 和阈值 $\{\theta_j\}$ 计算中间层各单元的输入 s_j ; 通过 S 函数计算输入层各单元的输出 h_j

$$s_j = \sum_{i=1}^L w_{ij} \cdot a_{ni} - \theta_j \quad j = 1, 2, \dots, M \quad (3)$$

$$h_j = f(s_j) \quad n = 1, 2, \dots, P \quad (4)$$

- ④ 用输入层的输出 h_j 连接权 $\{w_{jk}\}$ 和阈值 $\{\theta_k\}$ 计算输出层各单元的输入 l_k , 通过 S 函数计算输出层各单元的响应 y_k

$$l_k = \sum_{j=1}^M w_{jk} h_j - \theta_k \quad j = 1, 2, \dots, M \quad (5)$$

$$y_k = f(l_k) \quad k = 1, 2, \dots, N \quad (6)$$

- ⑤ 用期望输出模式 $d_n = (d_{n1}, \dots, d_{nN})$ 网络实际输出 y_k 计算输出层的各单元的一般化误差 X_k ,

$$X_k = (y_k - d_{nk}) \cdot d_{nk} (1 - d_{nk}) \quad k = 1, 2, \dots, N \quad (7)$$

- ⑥ 用连接权 $\{w_{ij}\}$ 输出层的一般误差 X_k 中间层的输出 h_j 计算中间层各单元的一般化误差 e_j ,

$$e_j = \left[\sum_{k=1}^N X_k \cdot w_{jk} \right] h_j (1 - h_j) \quad j = 1, 2, \dots, M \quad (8)$$

- ⑦ 用输出层各单元的一般化误差 X_k 中间层各单元的输出 h_j 修正连接权 $\{w_{ij}\}$ 和阈值 $\{\theta_k\}$

$$w_{ij}(N+1) = w_{ij}(N) + \eta \nabla l_j + T_j \Delta w_{ij} \quad (9)$$

$$j = 1, 2, \dots, M \quad k = 1, 2, \dots, N \quad (0 < T < 1, \nabla = X_k)$$

$$\theta_k(N+1) = \theta_k[N] + T \cdot X_k \quad k = 1, 2, \dots, N \quad (10)$$

- ⑧ 用中间层各单元的一般化误差 e_j 输入层各单元的输入 h_j 修正连接权 $\{w_{ij}\}$ 和阈值 $\{\theta_j\}$

$$w_{ij}(N+1) = w_{ij}(N) + \eta \nabla h_j + T_j \Delta w_{ij} \quad (11)$$

$$i = 1, 2, \dots, L \quad j = 1, 2, \dots, M \quad (0 < T < 1, \nabla = e_j)$$

$$\theta_j(N+1) = \theta_j[N] + T \cdot e_j \quad j = 1, 2, \dots, M \quad (12)$$

⑨ 随机选取下一个学习模式对提供给网络,返回到步骤③,直至全部 n 个模式对训练完毕。

⑩ 重新从 n 个学习模式对中随机选取一个模式对,返回步骤③,直至网络全局误差 E 小于预先设定的一个极小值,即网络收敛。当学习回数大于预先设定的值,则网络无法收敛。

最后,结束学习。

在以上学习步骤中,③~⑥ 为输入学习模式的“顺传播过程”,⑦、⑧ 为网络误差的“逆传播过程”,步骤⑨、⑩ 则完成训练和收敛过程。

2 数据处理

2.1 预测目标的确定

恰当的预测目标是取得良好效果的前提。若想预测明天股票的收市价,可假设某股股价为 340 点,且每天的变化量大约是 2 点,那么变化量相应为 $0.58\% (2/340)$;如果我们希望能预测到变化了 0.4 点,那么需要 0.1176% 的精度。但若预测明天股价的变化而非绝对数值,则只需 $20\% (0.4/2)$ 的精度。误差允许范围由 0.1176% 上升到 20% ,可见,预测股票收市价的变化率比预测股票收市价要优越,如图 4。这里作者预测深圳股市较具特色的深圳发展 A 股 5 天后的趋势: 1 表示上升趋势, 0 表示下降趋势。

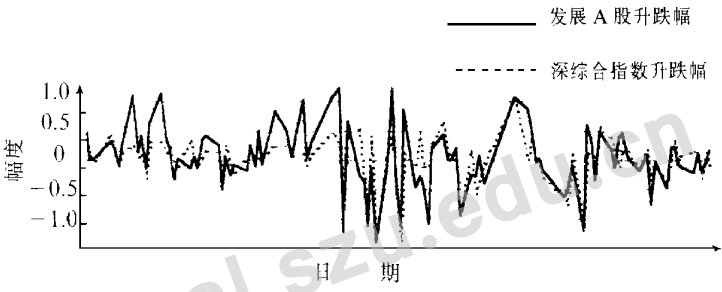


图 4 深综合指数升跌与深个股升跌的联系

2.2 训练数据的选择

由于目前对脑神经的生物学机制了解还不很清楚,人们所构造的神经网络模型只是人脑的一种功能模拟系统。因此,只能选择有限个输入信号,从而在实际应用时,人们可选择对于输出信息影响最大的关键因素作为输入变量。直接影响股市交易的因素有:

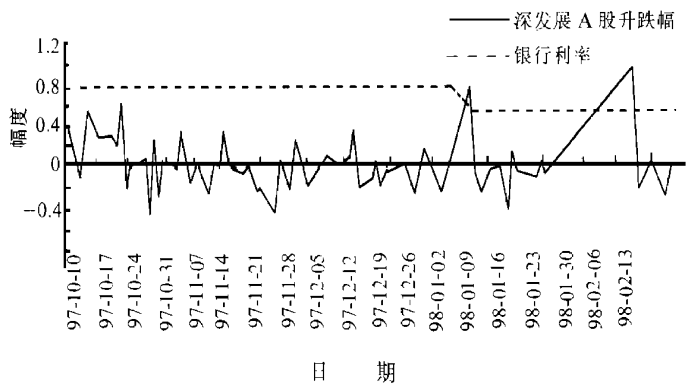


图 5 个股价升跌与银行利率的联系

银行利率、国家政策、物价指数、上市公司的业绩等,如图 5。本文选择有代表性的 5 个因素(参考各因素与个股联系图): 深交所发布的深发展 A 股 1997 年 3 月 5 日至 1997 年 8 月 9 日总共 100 天的历史数据(包括收市价、升跌幅度、深圳综合指数、深圳综合指数涨跌幅等),期间的银行利率作为训练模式 (Patterns),如图 6。

当预测目标是长期时,用周数据比用日数据噪声小.因此,为减小噪声,我们取股票收市价的 5 天移动平均值来预测后 5 天股价的变化量.

2.3 训练数据的规范化

输入数据通常规范在 $[-1, 1]$ 区间内.而当结点函数为曲线 $\text{Tanh}(x)$ 时,输出数据通常规范在 $[-0.8, 0.8]$; 当结点函数为 Sigmoid 函数时,其规范在 $[0.2, 0.8]$ 内.规范化的目的是为了帮助学习过程的顺利进行,本实验中使用一系列措施以达到把数据规范在 $[0, 1]$ 之间,且尽可能地平滑数据,以消除影响预测结果的噪声(参见图 7 和图 8).

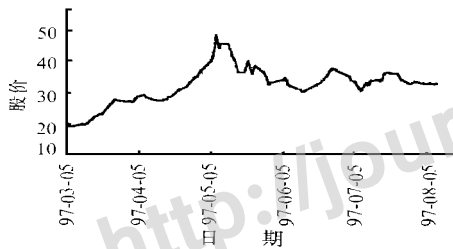


图 7 深发展 A 股收市价原始数据

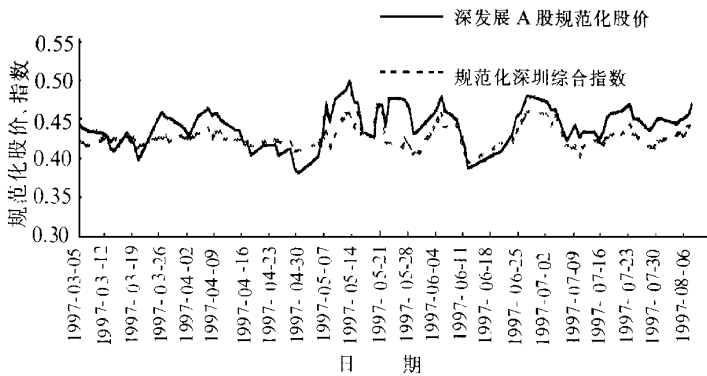


图 6 规范化股价与综合指数的联系

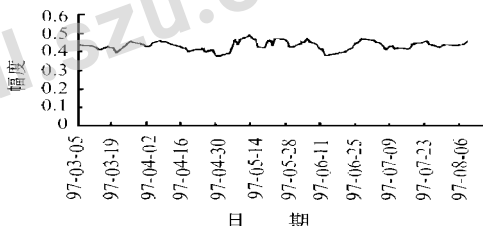


图 8 深发展 A 股收市价规范化数据

3 实例分析

3.1 激活函数对预测结果的影响

实验 1 使用 $\text{Sigmoid}(x)$ 函数作为隐含层和输出层的激活函数,对深圳发展 A 股采用 5 个输入神经元(收市价、升跌幅、深圳综合指数及其升跌幅、当时的银行利率等经过规范化的因素),一个输出神经元(1 表示 5 天后股价将上升;0 表示 5 天后股价将下降),训练模式对选用 1997 年 3 月 5 日至 1997 年 8 月 8 日,共 100 对,隐含层学习率 $\mu = 0.7$;输出层学习率 $\mu = 0.7$,结果见表 1.实验 2 在相同的层数、学习率下,仅仅是隐含层和输出层激活函数采用 $\text{Tanh}(x)$.结果证明:在 BP 模型中采用 $\text{Tanh}(x)$ 作为隐含层和输出层的激活函数,不仅能

表 1 神经元作用函数采用 Sigmoid

序号	日期	收市价	期望值	训练结果
1	1997- 03- 12	20. 04	1. 000000	0. 832781
2	1997- 03- 13	19. 88	1. 000000	0. 822857
3	1997- 03- 14	20. 34	1. 000000	0. 829651
4	1997- 03- 17	22. 33	1. 000000	0. 837898
5	1997- 03- 18	22. 59	1. 000000	0. 831298
6	1997- 03- 19	23. 40	1. 000000	0. 831572
7	1997- 03- 20	23. 10	1. 000000	0. 811842
8	1997- 03- 21	24. 40	1. 000000	0. 829451
9	1997- 03- 24	26. 84	1. 000000	0. 837646
10	1997- 03- 25	27. 75	0. 000000	0. 835108
11	1997- 03- 26	28. 17	0. 000000	0. 837884
13	1997- 03- 27	27. 50	0. 000000	0. 830541
14	1997- 03- 28	27. 56	1. 000000	0. 828335
15	1997- 03- 31	27. 30	1. 000000	0. 837861
16	1997- 04- 01	27. 45	1. 000000	0. 832832
17	1997- 04- 02	27. 18	1. 000000	0. 828755
平均误差				0. 068090

使学习输出准确度比采用 $\text{Sigmoid}(x)$ 作为激活函数的 BP模型高出许多 (参见表 2).

3.2 网络层数对预测结果的影响

实验 3中,本文使用只含有 1个隐含层和 4个隐含层 BP网络模型进行仿真实验对比,结果表明:只含有一个隐含层的 BP模型在精确度和收敛速度上都比包含 4个隐含层的模型要好,因此可以认为对于类似深圳股市这样的离散非线性时间序列,采用具有 3个以下隐含层的 BP神经网络可取得较为理想的结果 (见表 2和表 3).

表 2 隐含层为 1时的结果				
序号	日期	收市价	期望值	训练结果
1	1997- 03- 12	20. 04	1. 00000	0. 999513
2	1997- 03- 13	19. 88	1. 00000	0. 999891
3	1997- 03- 14	20. 34	1. 00000	0. 999907
4	1997- 03- 17	22. 33	1. 00000	0. 953872
5	1997- 03- 18	22. 59	1. 00000	0. 999679
6	1997- 03- 19	23. 40	1. 00000	0. 999782
7	1997- 03- 20	23. 10	1. 00000	0. 999898
8	1997- 03- 21	24. 40	1. 00000	0. 999920
9	1997- 03- 24	26. 84	1. 00000	0. 991207
10	1997- 03- 25	27. 75	0. 00000	- 0. 147525
11	1997- 03- 26	28. 17	0. 00000	0. 003683
13	1997- 03- 27	27. 50	0. 00000	0. 150665
14	1997- 03- 28	27. 56	1. 00000	0. 838536
15	1997- 03- 31	27. 30	1. 00000	0. 889625
16	1997- 04- 01	27. 45	1. 00000	0. 939695
17	1997- 04- 02	27. 18	1. 00000	0. 998588
平均误差				0. 009108

表 3 隐含层为 4时的结果				
序号	日期	收市价	期望值	训练结果
1	1997- 03- 12	20. 04	1. 000000	0. 997811
2	1997- 03- 13	19. 88	1. 000000	0. 999231
3	1997- 03- 14	20. 34	1. 000000	0. 999192
4	1997- 03- 17	22. 33	1. 000000	0. 950786
5	1997- 03- 18	22. 59	1. 000000	0. 997995
6	1997- 03- 19	23. 40	1. 000000	0. 998428
7	1997- 03- 20	23. 10	1. 000000	0. 999284
8	1997- 03- 21	24. 40	1. 000000	0. 999395
9	1997- 03- 24	26. 84	1. 000000	0. 987548
10	1997- 03- 25	27. 75	0. 000000	0. 156109
11	1997- 03- 26	28. 17	0. 000000	0. 015812
13	1997- 03- 27	27. 50	0. 000000	0. 166474
14	1997- 03- 28	27. 56	0. 000000	0. 809097
15	1997- 03- 31	27. 30	1. 000000	0. 886633
16	1997- 04- 01	27. 45	1. 000000	0. 944545
17	1997- 04- 02	27. 18	1. 000000	0. 996559
平均误差				0. 010716

3.3 学习率 μ 对预测结果的影响

实验 4采用学习率 $\mu=0.87$ 和 $\mu=0.70$ 以及 $\mu=0.15$ 作对比,仿真结果表明,采用 $\mu=0.87$ 的 BP模型能够较快地收敛且准确度相当高,因此我们可以认为 $\mu=0.87$ 是一个较为合理的参数 (见图 9).

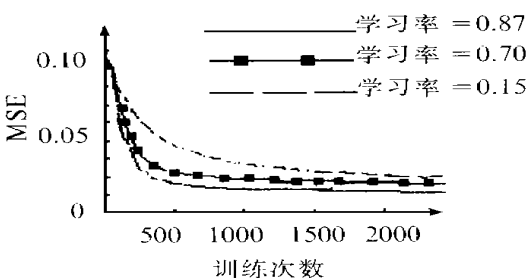


图 9 学习率对预测结果的影响

结 论

综上可得出如下结论:

① 低阶多层 BP神经网络模型有很强的自我学习修正误差的能力;

② 低阶多层 BP神经网络模型适合于逼近类似像中国深圳股市这样一种离散非线性时间序列;

③ 选取 BP模型参数时要综合考虑各种因素的影响,既要有快速性,准确度又不能太低. 本文通过仿真认为 3层 BP神经网络模型, 1个隐含层, 3个隐含层神经元(本文训练模式有 5 个输入神经元, 1个输出神经元, 且设输出神经元 \leq 隐含层神经元 \leq 输入神经), 参数惯性率 $\alpha = 0.85$, 学习率 $\mu = 0.87$;

④ 一旦网络模型确定下来, 数据处理就是关键, 经验表明: 神经网络应用的主要工作在于建模和组织数据. 选择合理的训练模式, 则数据经规范化后网络能较快地收敛. 本文采用的训练模式经规范化后均方误差 (MSE) 可以较快地达到 < 0.007 .

参考文献

- 1 李春常, 盛 然, 刘 波. 证券投资原理. 上海: 上海科技出版社, 1992
- 2 罗 兴. 股票、债券入门. 合肥: 安徽人民出版社, 1992
- 3 陈之大, 贺学会. 证券投资技术分析. 成都: 西南财经大学出版社, 1996
- 4 Tim H, Leorey Marquez, Marcus. Internation Journal of Forecasting. 1994, 10 25~ 30
- 5 Dagli C, Schierholt K. Stock Market Prediction Using Different Neural Network Classification Ardhitecs. IEEE/IAFE 1996 Conference on Computational Intelligence for Financial Engineering, 1996, 72~ 78
- 6 Ormoneit D, Neuneier R. Experiments in Predicting the German Stock Index DAX with Density Estimating Neural Networks. QIEEE/IAFE 1996 Conference on Computational Intelligence for Financial Engineering, 1996, 66~ 71
- 7 焦李成. 神经网络系统理论. 西安: 西安电子科技大学出版社, 1995. 34~ 41
- 8 杨大力, 刘泽民. 多层前向神经网络中 BP算法的误调分析及其改进算法. 电子学报, 1995, 23(1): 117~ 120

【责任编辑 肖 祺】

Application of Artificial Neural Network to Chinese Stock Market Forecast

Feng Junqiu and Li Hui

School of Engineering & Technology
Shenzhen University
Shenzhen 518060
P. R. China

Abstract

Some simulations are conducted with artificial neural network theory in the presented work. Results show that the modified BP method is effective to forecast the trend of Chinese Stock Market.

Key words neural network; nonlinear; system; BP algorithm; financial forecast