

神经网络在股价预测中的应用

徐迪 马大军 李元熹

(复旦大学管理学院, 上海 200433)

摘要 利用时延神经网络模型(Time Delay Neural Network)对四川长虹的股价作了预测。股价的涨跌预报可视为高维空间的非线性分类问题, 本文使用增益可调的反向传播算法, 对其走势作了预报。借助前馈神经网络对非线性函数的逼近能力, 本文对四川长虹股价这个时间序列作了连续若干天的一步预测。最后, 我们采用不同形式的误差函数对预测结果作了比较。

关键词 神经网络 反向传播算法 时间序 预测

The Application of Neural Networks to Stock Price Forecasting

Xu Di Ma Dajun Li Yuanxi

(School of Management, Fudan University, Shanghai 200433)

Abstract This paper presents the application of time delay neural networks to the prediction of the Sichuan Changhong Stock Price. The rise/fall prediction of the stock price can be regarded as the classification problem in the highdimension space. Using the back-propagation learning algorithm with an adaptable slope of the activation functions, we report on the rise/fall prediction of the stock price. The nonlinear-functional approximation capabilities of feedforward neural networks lead us to forecast time series, and we have undertaken the one-step stock-price forecast. In the end, different error functions are compared with respect to their performance on prediction.

Keywords neural networks; back-propagation; time series; forecasting

股票市场是一个复杂的非线性动态系统, 利用传统的时间序列预测技术很难揭示其内在的规律。而近十几年发展起来的神经网络理论逐渐成为非线性动态系统预测与建模的强有力工具。本文利用推广的反向传播算法预测了四川长虹的股价, 对其结果作了细致的分析, 并与其他预测技术作了比较。

1 前馈神经网络模型和反向传播算法

一般的前馈神经网络模型及其学习算法可表示如图1, 其中 $g(w, x)$ 一般可表为1个输入层, 1~2个隐层, 1个输出层的结构(见图2):

其中 $x = y^0 \in R^n$, $y^1 \in R^{n_1}$, $y^2 \in R^{n_2}$, $y = y^3 \in R^m$

$$u^s = w^{s-1} \cdot y^{s-1} \quad s = 1, 2, 3$$

$$y_j^s = f(u_j^s) = \tanh(Y_j u_j^s) = \frac{1 - e^{-2Y_j u_j^s}}{1 + e^{-2Y_j u_j^s}}$$

本文于1997年4月21日收到

这样 $g(w, x)$ 完成了 $R^n \rightarrow R^m$ 上的映照。

假设有训练样本集 $\{(x^k, d^k)\} \{k=1, \dots, N\}$ ，学习算法的目的是在权空间 W 中找到 W^* ，使 $g(w^*, x^k) = d^k (k = 1, 2, \dots, N)$ 。

定义误差函数 $E = \frac{1}{2} \sum_{k=1}^N \sum_{i=1}^m (d_i^k - g_i(w, x^k))^2$ ，构造学习算法等价于优化问题: $\min_w E(w)$ 。标准反向传播算法 (Back-Propagation) 是基于梯度法的, 即:

$$\Delta w = - \nabla E(w), w \leftarrow w + \Delta w$$

具体可写为:

$$\delta_i^s = f'(u_i^s) \cdot (d_i^s - g_i(w, x^k))$$

$$k = 1, 2, \dots, N; i = 1, 2, \dots, m$$

$$\delta_i^2 = f'(u_i^2) \cdot \sum_{j=1}^m w_{ji}^2 \delta_j^3 \quad i = 1, 2, \dots, n_2$$

$$\delta_i^1 = f'(u_i^1) \cdot \sum_{j=1}^{n_2} w_{ji}^1 \delta_j^2 \quad i = 1, 2, \dots, n_1$$

$$\Delta w_{ji}^s = \eta \delta_i^{s+1} \delta_j^s \quad s = 1, 2, 3$$

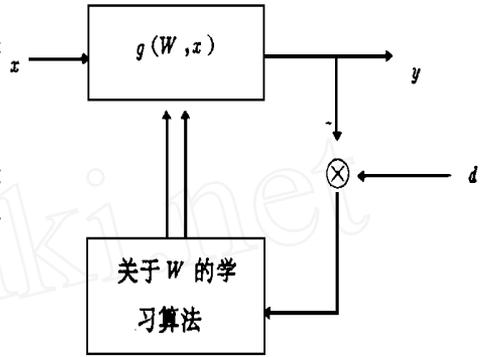


图 1 前馈神经网络模型及其学习算法

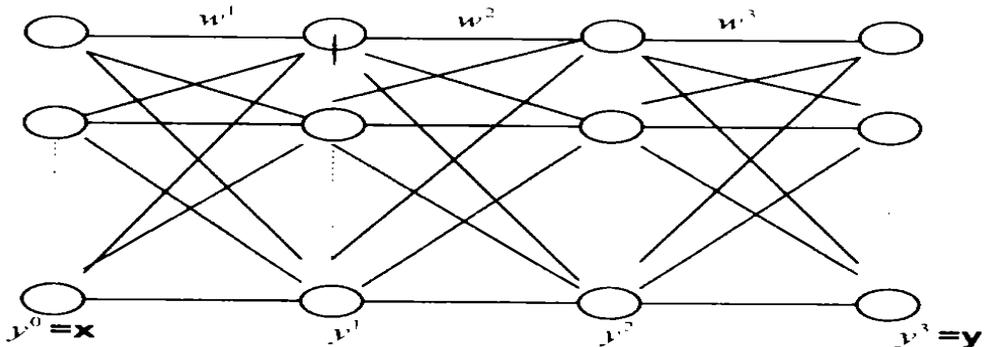


图 2 具体的神经网络实现

BP 学习算法可分为在线(增量)方式和批处理方式。在线方式中,对每一个样本学习后,权值便改变;批处理方式中,只有对所有样本都学习后,权值才改变,进入下一轮迭代。(注:批处理方式是严格意义上的最速下降法。但在线方式这种启发式搜索能够加速学习。)

为了加速BP 算法的收敛,人们提出了各种BP 改进算法,最常见的是增添动量项,即:

$$\Delta w_{ji}(k+1) = - \eta \frac{\partial E}{\partial w_{ji}} + \rho \Delta w_{ji}(k) \quad 0 < \rho < 1, \text{一般取 } \rho = 0.9$$

本文中,我们采用增益可调的BP 算法,见 § 4。

2 时间序列预测和神经网络

股票的价格可视为一离散时间序列。设第 k 天的股票价格(收盘价)为 x_k , 则:

$$x_k = g(x^k) = g(x_{k-1}, \dots, x_{k-l})$$

其中 $x^k = (x_{k-1}, \dots, x_{k-l})^T$, $g(\cdot)$ 为一非线性连续函数。 $g(\cdot)$ 刻划了生成这一时间序列的动态机制。只要找到合适的 $g(\cdot)$, 就能对这一时间序列作出适当的预测。然而,由于股票市场的复杂性,刻划其具体的动态过程是极其困难的。

对时间序列的预测一般采用 AP 或者 ARMA 模型。但它们对非线性问题的处理并不能得到很好的效果。而前馈神经网络对非线性函数的逼近特别适合, 因此利用它对股票市场的动态过程进行建模是合适的。我们采用时延神经网络模型, 即图 1 中的输入是 $x^k = (x_{k-1}, \dots, x_{k-l})^T$, 输出端的教师信号为 $d = x_{k+1}$ 。

选取一些匹配对 (x^k, x_{k+1}) 作为训练样本, 学习后的 $g(w, x)$ 便可作为一预测器, 对未来的股价作出一步预测, 见 § 4。(注: 多步预测的建模方式是找 $g(\cdot)$, 使 $x_{k+p} = g(x^k)$ 。

3 对股价涨跌的预测

首先我们对四川长虹的股价涨跌作出预报。

股价的涨跌可视为一个高维空间的非线性分类问题。此处将连续 5 天的收盘价作为神经网络的输入, 也就是分类的对象, 输出的状态作为分类的结果。一般将输出“+ 1”和“- 1”分别看作是未来股价的涨和跌(注 1: 因为“+ 1”和“- 1”分别是神经元输出的饱和态, 所以文中将输入“+ 0.8”和“- 0.8”分别看作涨和跌, 作为教师信号训练神经网络; 注 2: 神经网络作一步预测的处理方法见 § 4)。将新的 x^k 输入训练后的神经网络, 若输出正值, 则可视为第 k 天股价为涨; 若输出负值, 则可视为第 k 天股价为跌。应该注意的是, 输出值越大, 预报的可信度也越大。

表 1 是训练后的神经网络对 1996 年 1 月 8 日到 1996 年 1 月 19 日共 10 天(两周)的涨跌情况所作的预报。若将输出值高于 0.5 视为预报是可信的; 低于 0.5 视为预报是不可信的, 那么 10 天中能作出预报的有 8 次, 其中 6 次是正确的。

表 1 01/08/96- 01/19/96 四川长虹股价预测

日期	实际涨跌	预测结果
01/08/96		0.771
01/09/96		0.305
01/10/96		- 0.768
01/11/96		- 0.728
01/12/96		- 0.728
01/15/96		0.536
02/16/96		- 0.473
02/17/96		- 0.798
02/18/96		0.853
02/19/96		0.867

4 对四川长虹股价的预测

首先我们对数据进行预处理。设 μ 为 $x_k (k = 1, 2, \dots)$ 的均值, σ^2 为 x_k 的方差。 $x_k \leftarrow \frac{x_k - \mu}{\sigma}$, 预处理的目的是使学习能够收敛并得到加速。我们采用增益可调的 BP 算法, 即神经元输入-输出关系 $y_j = f(u_j) = \tanh(Y_j/u_j)$ 中, Y_j 是可变的; $g(w, x)$ 可视为 $g(w, Y, x)$ 。 Y 的迭代公式可作如下推导:

$$-\frac{\partial E}{\partial Y_j} = -\frac{\partial E}{\partial u_j} \cdot \frac{\partial u_j}{\partial Y_j} = -\frac{\partial E}{\partial u_j} \cdot \frac{\partial u_j}{\partial Y_j} \cdot \frac{\partial Y_j}{\partial Y_j}$$

其中, $-\frac{\partial E}{\partial u_j} = \delta_j$, $\frac{\partial u_j}{\partial Y_j} = \frac{1}{(1 - (y_j^2)^2) \cdot Y_j}$, $\frac{\partial Y_j}{\partial Y_j} = (1 - (y_j^2)^2) \cdot u_j^2$

所以 $-\frac{\partial E}{\partial Y_j} = \frac{u_j^2 \cdot \delta_j}{Y_j}$, 即 $\Delta Y_j = \eta \cdot \frac{u_j^2 \cdot \delta_j}{Y_j}$

Δw_{ji} 同 § 1 中所述。使用上述形式的 BP 算法可使收敛得到较大改善。

我们选取四川长虹连续 5 天的股价作为输入, 第 6 天的股价作为输出构成匹配对, 即 $x^k = (x_{k-1}, \dots, x_{k-5})^T, x_k$, 用 $S = \{(x^k, x_k) | k = l, l+1, \dots, l+9\}$ 共 10 个样本作为样本集, 训练神经网络。然后用 x^{l+10} 作为训练后的神经网络的输入, 其输出 y_{l+10} 便是对第 $l+10$ 天的预测 \hat{x}_{l+10} 。我们对 1995 年 12 月 14 日到 1996 年 1 月 4 日共 15 天的股价作了预测。这儿所作的都是一步预测, 即当我们得到 x_{l+10} 的实际值后, 样本集 S 作如下改变:

$$S \leftarrow S \setminus \{(x^{l+10}, x_{l+10})\} \cup \{(x^l, x_l)\}$$

$$l \leftarrow l + 1$$

然后对新的样本集 s 作又一轮的学习。预测结果如表 2:

表 2 12/14/95- 01/04/96 四川长虹股价预测

日期	实际值	预测值		
		神经网络	指数平滑	自回归
1995- 12- 14	7.66	7.60	7.69	7.63
1995- 12- 15	7.61	7.61	7.67	7.7
1995- 12- 18	7.5	7.60	7.63	7.58
1995- 12- 19	7.58	7.60	7.54	7.62
1995- 12- 20	7.82	7.58	7.57	7.44
1995- 12- 21	7.65	7.62	7.74	7.83
1995- 12- 22	7.7	7.62	7.68	7.74
1995- 12- 25	7.52	7.64	7.69	7.68
1995- 12- 26	7.56	7.62	7.57	7.46
1995- 12- 27	7.74	7.62	7.56	7.58
1995- 12- 28	7.63	7.69	7.69	7.64
1995- 12- 29	7.59	7.71	7.65	7.69
1996- 01- 02	7.52	7.66	7.61	7.61
1996- 01- 03	7.66	7.66	7.55	7.55
1996- 01- 04	7.76	7.70	7.63	7.66

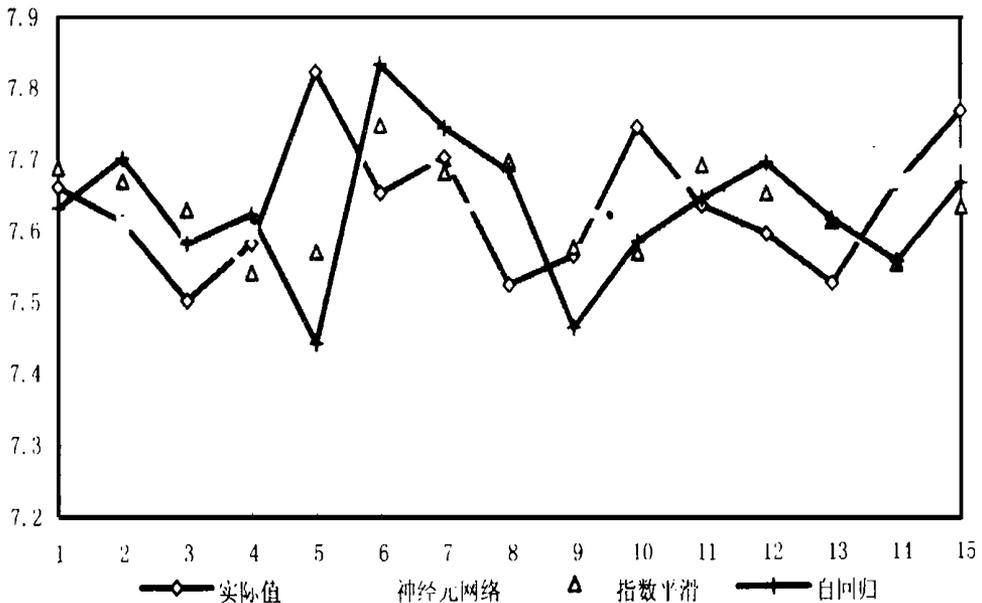


图 3 各预测方法对四川长虹股价(12/14/95# 01/04/96)的预测结果

表 2 与图 3 是分别利用神经网络, 指数平滑, 自回归这三种方法进行预测的结果。表 3 是对预测结果的分析。在几个指标上, 利用神经网络模型预测的结果优于其它两种方法。

表 3 不同预测方法的比较

预测方法	平均相对误差	相对误差的标准差	最大相对误差
神经网络	1.05%	0.80%	3.05%
指数平滑	1.25%	0.88%	3.23%
自回归	1.45%	1.14%	4.86%

值得注意的是,学习中止的控制误差越小,预测结果并非一定越好。表 4 比较了神经网络模型在不同控制误差下的预测结果($error < 0.08$ 表示当输出和实际值之间的平均误差不超过 0.08 元时,学习结束)。容易看出,控制误差为 0.08 时的预测结果最好。这种过度拟合的现象可以这样来解释:样本集 S 中的输入输出信号夹杂有噪声信号。学习开始时, (w, γ) 的调整主要是针对实际信号的特征所进行的;随着学习的继续, (w, γ) 的结构也反映了噪声的某些特征。因此过度学习将过多的噪声信息引入了 (w, γ) 的结构。

表 4 不同控制误差下的预测结果

控制误差	平均相对误差	相对误差的标准差	最大相对误差
$error < 0.090$	0.99%	0.79%	3.05%
$error < 0.085$	0.99%	0.79%	3.05%
$error < 0.080$	0.98%	0.80%	3.05%
$error < 0.075$	1.05%	0.80%	3.05%
$error < 0.070$	1.08%	0.82%	3.05%
$error < 0.065$	1.12%	0.82%	3.05%
$error < 0.060$	1.10%	0.82%	3.04%

在对新的样本集和原样本集学习时,由于输入值的移位,它们对应的误差函数相差很大。但是用原样本集学习后的权值,作为对新样本集学习的初始权值,学习收敛得很快(控制误差为 0.06 时,对第 1、9、10 个样本集,分别经过 430 次、282 次和 400 次迭代,学习中止;对其余样本集,最多经过 11 次迭代学习便中止);亦即 (w, γ) 的改变很小。这也从某种程度上说明了造成相邻若干天股市变化的动态机制是相似的。

我们还可以取不同的误差函数来考察预测结果。例如引进 Minkowski r 距离,将误差函数写作:

$$E = \frac{1}{r} \sum_{k=1}^N \sum_{i=1}^m |d_i^k - g_i(w, x^k)|^r$$

于是 § 1 中 δ_i^k 的迭代公式修改为:

$$\delta_i^k = f(u_i^k) \cdot \text{sgn}(d_i^k - g_i(w, x^k)) \cdot |d_i^k - g_i(w, x^k)|^{r-1}$$

从表 5 中可以看出, r 值比 2 小时,预测品质得到了不同程度上的改善。 r 值较小时,能减少训练样本集中异常点对整体性能的影响。

表 5 相同控制误差下(0.075),Minkowski r 距离对预测结果的影响

r	平均相对误差	相对误差的标准差	最大相对误差
$r=2.0$	1.05%	0.80%	3.05%
$r=1.9$	1.07%	0.79%	3.01%
$r=1.8$	1.01%	0.73%	3.02%
$r=1.7$	1.01%	0.73%	2.94%

5 结束语

由于股票市场的动态机制非常复杂,因此利用神经网络对股票市场建模是值得进一步探讨的。我们

可以考虑以下两点:

1) 本文只从误差控制方面对网络的外推有效性 (generalization) 作了初步探讨。在学习过程中, 对权进行裁剪 (pruning), 对网络进行调整, 可以提高网络的预测能力。

2) 本文只考虑了股票价格所构成的时间序列。股票价格不仅受前几天价格波动的影响, 还会受其它因素的影响。如何对相关因素 (如成交金额) 进行预处理和编码, 是很值得探讨的。

参考文献

- 1 杨璐, 黄梯云, 洪家荣. 一种基于神经网络的时间序列自适应建模和预测方法. 决策与决策支持系统, 1996, 6(2): 69~ 75
- 2 Cichocki A. and Unbehauen R. Neural Networks for Optimization and Signal Processing. John Wiley & Sons, Chichester, 1993
- 3 Hassoun M. Foundations of Artificial Neural Networks. MIT Press, Cambridge, 1995
- 4 Cottrell M., Girard B., Girard Y., M angeas M. and Muller C. Neural Modeling for Time Series: A Statistical Stepwise Method for Weight Elimination. IEEE Trans. Neural Networks 1995, 6(6): 1355~ 1363
- 5 Kamijo K and Tanigawa T. Stock Price Pattern Recognition. A Recurrent Neural Network Approach. IEEE Int. Conf. Neural Networks 1990, (1): 215~ 221

(上接第 60 页)

5 模型应用

淄博矿务局某矿一工作面, 平均煤厚 1.27m, 顶底板较稳定, 煤层硬度系数 $f = 1.8$, 煤厚变异系数为 0.06, 工作面无断层, 煤层无夹矸。

将上述条件输入网络得结果如下:

表 3 网络输出结果

A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
0.0281	0.9979	0.0307	0.0288	0.0316	0.0451	0.0341	0.0000

结果显示, 该工作面应采用 A_2 采煤方式, 即采煤机爬底式。

6 结语

本文建立了缓倾斜薄煤层采煤方式识别模型, 识别准确率高, 随着样本的更加丰富, 模型识别能力还可进一步完善扩充。

薄煤层开采中如采煤方式识别等一些问题存在较多的模糊因素, 用传统数学方法是很难准确描述的, 而人工神经网络具有学习能力, 可综合更多的专家知识, 较准确地解决这些问题。

参考文献

- 1 周继成. 人工神经网络. 科学普及出版社, 1993, 3: 64~ 70
- 2 王卫军, 陈良棚. 应用人工神经网络实现综采面设备选型. 煤炭学报, 1996, 3: 231~ 234
- 3 王文博. 复杂条件下薄煤层机械化开采. 山东煤炭科技, 1990, 4: 28~ 31