

## 基于 TSA- BP 神经网络的企业产品市场占有率预测模型

陈新辉 乔 忠 曾广晟 刘石球

(中国农业大学管理工程学院)

**摘 要** 在分析了现有的各种市场占有率预测模型的基础上, 把产品市场占有率的时间序列作为 BP 神经网络的输入与输出, 建立了预测模型, 并对 1999 年部分品牌彩电的市场占有率进行了实证分析, 结果表明, 所建模型具有很高的预测精度, 是一种值得推广的预测方法。

**关键词** 市场占有率; 时间序列分析; 人工神经网络

**分类号** O 233

## A Forecasting Model on the Market Share of the Enterprise Product Based on Time Series Analysis-Back Propagation Neural Network

Chen Xinhui Qiao Zhong Zeng Guangsheng Liu Shiqiu

(College of Management Engineering, CAU)

**Abstract** Based on the analysis of all existed models of forecasting market share. A back propagation neural network model is provided which is characterized by the input and output of time series of market share. At last the model is tested with the data on market share of some color TV brands in 1999. The result shows that the forecasting precision of this model is very high. And this model is worth spreading.

**Key words** market share; time series analysis; artificial neural network

企业产品市场占有率是指在一定时期内、一定区域市场上, 企业某种品牌产品的实际销量与该市场对同类产品需求总量的比值<sup>[1]</sup>。它是衡量企业产品销售状况的一个重要指标, 企业经营者只有通过一定的方法, 随时跟踪、预测本企业产品的市场占有率, 才能及时发现问题, 采取有效的市场营销策略。然而, 市场占有率是一个受许多因素影响的变量, 所以该指标值具有很强的不确定性。目前, 许多涉及市场占有率测算模型的市场营销文献, 均引入了市场营销力的概念<sup>[2]</sup>。事实上, 这些通过计算市场营销力来测算市场占有率的模型无法预测短期内企业产品的市场占有率, 这是因为企业市场营销力对产品市场占有率的影响具有滞后效应, 计算所需原始数据难以获得, 且各部分市场营销力的弹性系数一般只能靠人员估计, 主观随意性很大。笔者认为, 从系统的角度看, 某一既定时期内在某一既定市场上, 竞争企业产品市场占有率值在空间上的分布, 构成了一个竞争均衡模式。这种模式在时间上的分布形成了企业产品竞争模式的时间序列, 而且这种时间序列, 由于受诸多因素的影响, 经常发生随机波动。

近年来, 在市场研究中神经网络的应用还不多见, 但是它是一种新兴的数学建模方法, 具

收稿日期: 2000-06-02

国家自然科学基金资助项目

陈新辉, 北京清华东路 17 号 中国农业大学(东校区)211 信箱, 100083

有处理系统非线性的功能<sup>[3]</sup>。本文中把企业产品竞争均衡模式的时间序列作为BP神经网络模型的输入模式和期望输出模式,构建了企业产品市场占有率的预测模型,并且根据全国百家商城各个品牌彩电的市场占有率数据进行了实证分析。

## 1 基于 TSA-BP 神经网络的企业产品市场占有率预测模型

### 1.1 时间序列分析模型

ARMA 模型是时间序列分析中一种最基本、应用较广的模型,其模型结构可通过方程<sup>[4]</sup>

$$Y_t = \sum_{i=1}^n \varphi_i Y_{t-i} - \sum_{j=1}^m \theta_j a_{t-j} + a_t \quad (1)$$

来表示。其中  $Y_t$  是时间序列  $Y$  在  $t$  时刻的观察值,  $\varphi_i (i=1, 2, \dots, n)$  为自回归系数,  $\theta_j (j=1, 2, \dots, m)$  为滑动平均系数,  $a_t$  为残差项,  $n$  为自回归阶数,  $m$  为滑动平均阶数。式(1)是一个具有  $n$  阶自回归部分,  $m$  阶滑动平均部分的 ARMA 模型,故  $Y_t$  称为 APMA 时序或 ARMA 过程。可以看出,式(1)将  $Y_t$  分解成确定性部分与随机性部分,确定性部分为  $\sum_{i=1}^n \varphi_i Y_{t-i} - \sum_{j=1}^m \theta_j a_{t-j}$ 。因为  $t$  时刻,  $Y_{t-1}, Y_{t-2}, \dots$  均是已经确定的观察值,尽管  $a_t$  是随机干扰序列,是不可测的,但是  $t$  时刻以前所有的  $a_{t-1}, a_{t-2}, \dots$  都已经发生,因而该部分是确定的;  $Y_t$  的随机部分由不可观察的随机干扰  $a_t$  来确定。 $m, n$  的不同取值组成了 ARMA 的庞大体系;当  $m, n$  取较大值时,求解会有困难。本文中采用 BP 神经网络来进行 TSA 建模。

### 1.2 预测模型原理

在某一时刻,同一市场上,各个企业某品牌产品的市场占有率反映了各个企业营销的竞争结果,它必然与以前时刻的产品竞争结果有关。设  $Y_t$  为在  $t$  时刻,某一个区域市场,各竞争企业某品牌产品的市场占有率组成向量,则

$$Y_1 = \varphi_1 Y_0 + \sum_{i=2}^n \varphi_i Y_{1-i} - \sum_{j=0}^m \theta_j a_{1-j} \quad (2-a)$$

$$Y_2 = \varphi_1 Y_1 + \varphi_2 Y_0 + \sum_{i=3}^n \varphi_i Y_{2-i} - \sum_{j=0}^m \theta_j a_{2-j} \quad (2-b)$$

⋮

$$Y_N = \varphi_1 Y_{N-1} + \varphi_2 Y_{N-2} + \dots + \varphi_N Y_0 + \sum_{i=N+1}^n \varphi_i Y_{N-i} - \sum_{j=0}^m \theta_j a_{N-j}$$

将式(2-a)作变换可得  $Y_1$  关于  $Y_0$  的表达式,代入式(2-b)得  $Y_2$  关于  $Y_1$  的表达式, ..., 依次类推,可以得到  $Y_t = \text{TSA}(Y_{t-1}), t=1, 2, \dots, N$ 。按式(2)构造 BP 神经网络模型,即以  $Y_{t-1}$  作为输入模式,  $Y_t$  作为输出模式,将时间序列  $\{Y_0, Y_1, \dots, Y_N\}$  形成的  $N$  个样本对:  $(Y_0, Y_1), (Y_1, Y_2), \dots, (Y_{N-1}, Y_N)$ , 提供给 BP 网络学习,使之逐渐逼近 TSA( $\bullet$ ),从而使 BP 网络掌握企业综合竞争结果的动态变化特性。由  $N$  个样本训练成功的模型称为 TSA-BP 神经网络模型。

由于  $Y_0, Y_1, \dots, Y_N$  这  $N+1$  个数据的大小及其先后顺序蕴含了在既定市场上各个企业品牌产品竞争结果(产品市场占有率)变化的系统特性与状态变化信息,因而,将上述  $N+1$  个数据,按  $(Y_{t-1}, Y_t)$  构成  $N$  个样本,对 BP 网络模型进行训练, BP 网将信息凝聚在权空间  $w$  上,权空间  $w$  上的元素  $w$  取代了 ARMA 模型中的  $\varphi, \theta, a$ 。BP 神经网络模型的学习是自动的,当有新样本进入模型时, BP 网络将会进行新的学习;因此,可采用这个模型对企业产品市场占有率

进行动态预测。

### 1.3 预测模型的结构和算法

基于 TSA-BP 神经网络的企业产品市场占有率预测模型由数据预处理器(对输入模式和目标模式的预处理有助于网络收敛和稳定)、BP 网络和数据后处理器(将网络输出模式还原)组成(图 1)。 $w_{ij}, v_{jk}$  分别为不同层次的连接权。网络隐层的学习函数采用双曲正切 Sigmoid 函数, 输出层的学习函数采用线性函数  $f(x) = x$ , 网络输入层的神经元个数  $n$  即为向量  $Y_t$  的维数, 也是本文中研究的既定市场上主要竞争企业的个数, 输出层的神经元个数为  $m$ 。显然, 本模型中  $n$  与  $m$  相等。隐层的神经元个数  $l$  由输入层神经元个数  $n$  与输出层神经元个数  $m$  决定,  $l = 3(nm)^{1/2}/2 = 3n/2$ ; 设输出层到隐层的连接权为  $w_{ij}$ , 隐层到输出层的连接权为  $v_{jk}$ ,  $\gamma_j$  为隐层的阈值,  $\delta_k$  为输出层的阈值。其中  $i = 1, 2, \dots, n; j = 1, 2, \dots, l; k = 1, 2, \dots, m$ 。

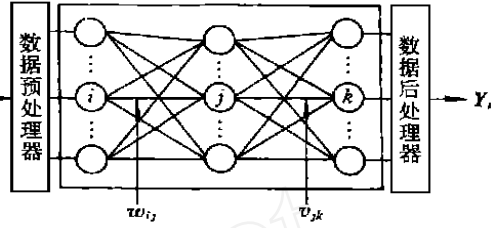


图 1 预测模型结构图

模型的基本算法如下: 1) 输入时间序列为  $X = [Y_{t-1}, Y_t | t = 1, 2, \dots, N]$ , 其中  $Y_t = (x_1^t, x_2^t, \dots, x_m^t)$ 。  $x_1^t$  代表  $t$  时刻第 1 个企业品牌产品的市场占有率,  $x_2^t$  为  $t$  时刻第 2 个企业品牌产品的市场占有率, .....; 依次类推,  $m$  为在即定市场上的竞争企业个数,  $N$  为网络训练模式对个数。 2) 对数据进行预处理, 将原输入数据等比例压缩到  $[-1, 1]$  区间。 3) 初始化 BP 网络的各层连接权值和阈值, 将各连接权  $w_{ij}, v_{jk}$  及阈值  $\gamma_j, \delta_k$  赋于  $[-1, 1]$  之间的随机值。 4) 选取一个输入模式  $Y_t = (x_1^t, x_2^t, \dots, x_m^t)$  进行循环: 计算隐层输入值  $s_j$  和输出响应值  $b_j$ , 输出层的输入值  $L_k$  和输出响应值  $C_k$ 。

$$s_j = \sum_{i=1}^n w_{ij} x_i^t - \gamma_j, b_j = \frac{e^{s_j} - e^{-s_j}}{e^{s_j} + e^{-s_j}}, L_k = \sum_{j=1}^m v_{jk} b_j - \delta_k, C_k = L_k$$

5) 计算输出层误差  $d_k$  和隐含层误差  $e_j^t$ 。  $d_k = (Y_k^t - C_k^t)$ ,  $e_j^t = (d_k^t v_{jk}) (1 + b_j) (1 - b_j)$ 。 6) 判断是否循环至样本总数, 否则返回 4)。 7) 计算各个样本的误差总和  $E$ , 并判断  $E$  是否满足精度要求, 若  $E$  小于精度值, 则学习停止, 转 8); 否则, 按式 (3) 修正连接权值和阈值, 并返回 4)。

$$\left. \begin{aligned} \Delta w_{ij}(p+1) &= \beta e_j^t x_i^t + \eta \Delta w_{ij}(p) \\ \Delta v_{jk}(p+1) &= \alpha d_k^t b_j + \eta \Delta v_{jk}(p) \\ \Delta \gamma_j(p+1) &= \beta e_j^t + \eta \Delta \gamma_j(p) \\ \Delta \delta_k(p+1) &= \alpha d_k^t + \eta \Delta \delta_k(p) \end{aligned} \right\} \quad (3)$$

式中  $p$  为最大学习次数的循环控制变量,  $\alpha, \beta, \eta$  是在  $(0, 1)$  区间的系数, 统称为学习系数。判断  $p$  是否大于最大学习次数, 若  $p$  大于最大学习次数, 转 8), 否则返回 4)。 8) 结束网络学习。 9) 利用网络进行预测(已知  $Y_{t-1}$  来预测  $Y_t$ )。

## 2 预测模型的应用

全国彩电市场是家电市场的重要组成部分, 由于近几年彩电进一步普及, 市场竞争愈发激烈; 因此彩电企业在研究自身品牌产品市场占有率预测时, 不应仅看到本企业的品牌产品市场占有率随时间的变化, 而更应着眼于市场内各竞争品牌产品市场占有率的协同变化。在进行产

品市场占有率预测时,应把所有竞争企业定义为一个竞争集合,作为一个系统进行预测。

本文中选取康佳、长虹、TCL等10个全国知名品牌的彩电作为研究对象,这些品牌彩电的销量约占全国彩电市场80%的份额,因而它们反映了我国整个彩电市场的销售情况。表1中的数据取自全国35个城市106个商城的统计资料<sup>[5]</sup>,样本采集的时间范围为1999年1月至2000年1月。设康佳、长虹、TCL、海尔、厦华、海信、创维、索尼、熊猫、金星等10个品牌的市场占有率实际值依次为 $x_1, x_2, \dots, x_{10}$ ;预测值分别为 $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_{10}$ 。采用MATLAB语言编程,进行模型算法的实现。最大学习次数 $p = 5000$ ,学习系数 $\alpha, \beta, \eta$ 都取值0.08。总误差精度 $\mu = 0.05$ ,将第 $t$ 个模式记为 $X_t = (x_1^t, x_2^t, \dots, x_{10}^t)$ 。平均平方误差即总误差 $E$ 的学习曲线见图2。经过训练,模型的输出值见表2。

表1 市场占有率的实际值<sup>[5]</sup>

日期	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$
1999-01	17.5	10.9	9.2	8.7	6.6	6.6	3.3	3.5	6.3	4.2
1999-02	17.7	12.6	9.5	8.5	5.4	6.9	3.8	3.7	5.0	4.0
1999-03	16.2	9.9	9.5	8.2	11.4	5.6	4.0	4.5	3.1	3.8
1999-04	18.6	13.2	10.3	8.3	4.9	7.1	4.6	4.4	3.3	3.3
1999-05	17.5	16.5	10.7	9.1	5.2	8.3	6.4	3.3	3.1	3.1
1999-06	14.8	14.1	11.0	8.1	7.1	8.0	5.8	3.2	3.7	2.8
1999-07	16.9	13.4	12.4	9.2	6.4	9.2	4.5	2.9	3.8	2.9
1999-08	16.3	14.4	12.0	7.9	6.9	10.4	4.7	3.3	2.2	2.3
1999-09	14.7	13.7	11.4	6.9	6.4	9.1	6.0	4.5	2.9	3.6
1999-10	14.1	13.9	10.9	7.3	5.5	8.8	4.6	4.3	3.4	4.1
1999-11	15.4	13.2	11.6	7.3	5.5	9.0	4.6	3.7	2.9	3.0
1999-12	12.2	11.7	12.1	4.5	8.1	10.3	4.7	4.5	3.9	4.1
2000-01	12.9	15.3	13.0	3.9	7.3	11.0	5.2	3.9	2.6	3.5

说明: 当月某品牌产品的市场占有率= 当月该品牌实际销售量/当月全部品牌产品的销售量。

其中平均平方误差  $E = \{ [\sum_{t=1}^N \sum_{j=1}^m (\hat{y}_j^t - y_j^t)^2] / mN \}^{1/2}$ , 式中 $m$ 为输出层神经元个数, $N$ 为样本模式对个数, $\hat{y}_j^t$ 为预测值,而 $y_j^t$ 为实际值; $E$ 是衡量神经网络模型是否收敛的重要指标之一,它一般随学习次数的增加而减小。

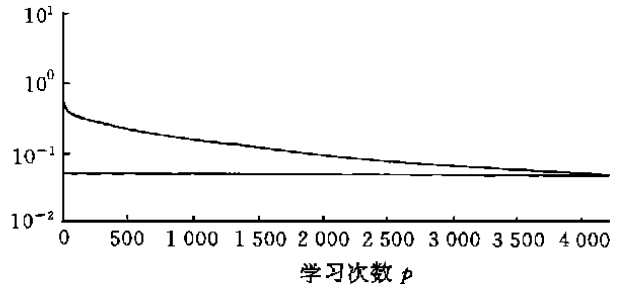


图2 平均平方误差的学习曲线

从图2可见,当学习次数 $p = 4207$ 时,网络以0.05的精度收敛。由于网络连接权的初始值和阈值是随机赋值的,所以平均平方误差曲线是不唯一的,本文中仅选取了其中的一条学习曲线。

比较预测值和实际值发现:

1) 基于TSA-BP神经网络的企业产品市场占有率预测模型的总预测效果比较理想,但对于个别品牌,预测效果不如人意。笔者认为,这与该品牌在市场中的地位有关,模型很难正确预

表 2 市场占有率的预测值

日期	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$
1999-02	17.8	11.2	9.7	9.1	7.3	6.4	4.1	3.9	4.4	3.7
1999-03	16.3	10.8	9.7	7.9	8.3	6.6	3.8	4.3	3.9	4.0
1999-04	18.5	13.3	10.2	8.3	5.5	7.0	5.0	4.3	3.4	3.4
1999-05	17.5	16.2	10.8	9.1	4.5	8.0	5.5	3.3	2.9	2.9
1999-06	15.5	14.2	11.2	8.6	6.5	8.3	5.8	3.3	3.7	3.1
1999-07	15.9	14.0	12.0	8.3	6.9	9.6	4.7	2.9	3.6	2.8
1999-08	16.9	14.3	12.2	8.4	6.6	10.2	4.8	3.3	2.5	2.4
1999-09	14.2	13.6	11.3	6.4	6.9	8.7	6.0	4.5	2.9	3.3
1999-10	13.8	13.9	10.9	7.6	5.7	9.2	4.9	4.2	3.4	4.0
1999-11	15.5	14.5	11.3	6.4	6.3	8.6	4.9	4.2	2.8	3.4
1999-12	12.4	11.8	12.1	4.5	8.7	10.0	4.3	4.4	3.6	4.0
2000-01	13.5	15.0	12.7	4.2	6.8	10.3	5.0	4.0	2.8	3.3
2000-02	13.8	15.4	12.9	4.1	6.7	10.4	5.3	4.0	2.5	3.2

测具有相对优势的中间品牌产品市场占有率的变化。

2) 预测趋势的正确度与该品牌的市场占有率有关。对于市场占有率有绝对优势的品牌(如长虹、康佳), 趋势预测正确度较高; 市场占有率相对较低的品牌(如金星), 趋势预测正确度也很高; 而市场占有率处于中间状态的品牌, 预测正确度相对较低。

3) 模型对市场占有率小幅随机波动的预测正确度很差, 例如夏华彩电在 1999 年 3 月到 8 月市场占有率发生微小变动, 但模型无法预测其正确的变化趋势, 这也许是模型为保证总体的预测精度而忽略了个别的微小变动。

### 3 结束语

电子商务的发展, 加速了市场一体化的进程, 各个企业在统一的大市场上会产生错综复杂的关系, 导致了各个企业竞争结果的相关性。企业经营环境的复杂性导致了竞争结果的随机性, 因此企业市场预测技术不能局限于单个企业。笔者从系统的角度提出的基于 TSA-BP 神经网络的企业产品市场占有率预测模型, 预测结果表明, 所建模型是有效的, 但主要适用于对竞争非常激烈而且市场一体化程度比较高的品牌产品的市场预测。

### 参 考 文 献

- 1 侯铁珊 市场预测原理与技术 北京: 中国商业出版社, 1985 35~ 56
- 2 郭国庆 市场营销原理- 理论与模型 中国人民大学出版社, 1996 134~ 135
- 3 Li X, Cheng L, Gray R. An intelligent business forecaster for strategic business planning J Forecast, 1998(18): 181~ 204
- 4 谭章禄 人工神经网络在矿业工程中应用 北京: 煤炭工业出版社, 1998 95~ 134
- 5 商务网站 <http://www.eea.com.cn/market>