

时间序列神经网络预测方法

文新辉

(西安电子科技大学神经网络研究中心 西安 710071)

摘要 本文从信息论的角度出发,讨论了利用神经网络理论构造时间序列预测模型的可能性和关键问题,并在此基础上提出3种时间序列神经网络预测方法。它们是:神经网络非线性时间序列模型、神经网络多维时间序列模型和神经网络组合预测模型。将上述模型应用于实例的结果表明,在非线性信息的处理能力和预测精度方面都有很大提高。进一步,对今后智能信息预测方法的发展方向进行了探讨,提出了智能信息预测系统的结构模型。

关键词 信息论,信息处理,神经网络预测方法

1 引言

随着科学技术的发展,愈来愈多的事物进入人们认识和研究的领域成为科学的研究对象。而总括起来,构造所研究对象模型的方法有2种,即所谓唯理方法和唯象方法。

唯理方法着重于揭示支配事物生成的动力系统。此时事物必须呈现出“灰箱”特性,通过对于事物变化规律的不断了解,使得事物机制由“灰箱”向“白箱”转化。但当基础理论不存在或与之相差甚远时,事物系统呈现出“黑箱”特性。此时,我们对于事物的了解仅局限于输入值和输出值,而不知道其内部结构。只有通过分析输入值与输出值之间的联系,来推测未来的发展状况,这就是所谓唯象研究方法。时间序列分析方法实际上是一种唯象研究方法的特例^[1]。

设有1组随时间 t 变化的序列 $\{x_t\}(t=1,2,\dots,T)$,那么传统的一维时间序列模型可以广义地用映射表示为

$$x_{t+k-1} = \varphi(x_t, x_{t-1}, \dots, x_{t-l}). \quad (1)$$

如果将 x_t 替换为向量,那么相应地(1)式中的映射 φ 变为向量值映射,从而也就是多维时间序列问题。

如果构造 φ 是线性映射,那么(1)式称为线性时间序列模型,如著名的AR和ARIMA模型;相应地,当构造 φ 是非线性映射时,那么称(1)式为非线性时间序列模型,如TAR和SETAR模型。

从上面的分析可以看出,要使时间序列模型具有较高的拟合精度,关键在于映射 φ 与实际值的逼近程度。然而,由于数学工具的局限性,人们通常只能通过线性的方法来实现映射 φ 。因此,我们提出这样一个问题,能否采用其它技术来构造性地实现映射 φ ?

1992-12-22 收刊,1994-03-31 定稿

文新辉 男,1960年生,讲师,博士生,专业为系统工程。

人工神经网络是人脑神经元网络的一种物理模型。它借助于网络中各神经元之间的相互作用来完成网络的信息处理能力。由于其网络结构的物理特性,从而使得它具有非线性动力系统的共性,即吸引性、耗散性、非平衡性、高维性、广泛联结性和自适应性等。因此,本文旨在利用神经网络信息处理技术,构造性地实现映射 φ 。而利用神经网络系统理论构造的时间序列预测模型,具有并行处理、高速度、高可靠性的明显优越性。它能应用于经济、生产计划、销售、股票、人口统计、昆虫预测、库存控制等众多领域,对国民经济建设具有普遍的意义。

应用神经网络技术处理时间序列预测问题,要解决的关键问题是:

- (1) 如何针对不同类型的时间序列预测问题构造相应的神经网络预测模型;
- (2) 如何选择数据作预测;
- (3) 如何构造神经网络的训练数据库。

下面,我们分别建立了 3 种新的神经网络预测模型。它们是:神经网络非线性时间序列模型,神经网络多维时间序列模型和神经网络组合预测模型。

2 神经网络非线性时间序列模型^[2]

非线性时间序列分析是目前一个迅速发展的方向。这是因为在现实世界中许多现象都不能用线性时间序列模型去很好地解决。

本文利用神经网络技术进行非线性信号处理,从而构造了一种新的神经网络非线性时间序列模型。

2.1 神经网络非线性时间序列结构模型

用 x_t 表示在观察时间 t 的实值随机变量。

定义 1 时间序列 $\{x_t\}$ 是一个 $t \in Z$ 的实值随机变量,其中 Z 表示整数集。

我们首先构造网络的拓扑结构如图 1 所示。其结构是一个具有 Sigmoid 激励函数和平方误差准则的多层前馈网络。

2.1.1 输入单元 对于任意时间 t , L 是相关的时间延迟,即 $\{x_{(t-i)}, 1 \leq i \leq L\}$ 给出 L 个实值输入单元。神经网络的第 k 步预测因子 $x^*(t+k-1)$ 是由下式定义的:

$$x_{(t+k-1)}^* = F_{\omega}(x_{(t-1)}, x_{(t-2)}, \dots, x_{(t-L)}, x_{(t+k-1)-12}, \dots, x_{(t+k-1)-12p}), \quad (2)$$

其中 F_{ω} 表示网络的变换函数。

对于纯粹的自回归模型常常是很难反应其具有很强季节性的序列。因此,有必要对网络提供进一步的信息。一种直接的办法是增加一组 12 个二进制的月份单元。它利用一组较粗略的月份编码来表示当前月份。例如: $U_1 = 1, U_2 = 0 = U_3 = \dots = U_{12}$ 表示一月份 (Jan.), 而用 $U_1 = 0, U_2 = 1, U_3 = 0 = U_4 = \dots = U_{12}$ 表示二月份 (Feb.) 等等。

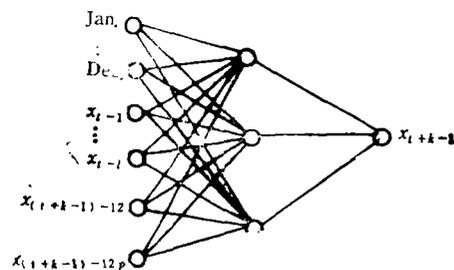


图 1

另用 P 个输入单元 $\{x_{(t+k-1)-12p}, 1 \leq p \leq P\}$, 表示关于所预测时间的 P 年延迟. 我们用年份单元来表示长数据的序列周期性.

2.1.2 输出单元 由于我们只限于一维非线性时间序列过程, 从而仅需要一个输出单元 $x_{(t+k-1)}$ 表示第 k 步预测值.

2.1.3 隐层单元 在我们的网络模型中, 采用一个或多个隐层, 其节点个数由实际问题的复杂度和规模决定.

网络模型的执行分为 2 个阶段: 学习阶段和预测阶段. 学习阶段是通过对历史数据的学习, 使已有模式以权系数的方式存储于网络之中; 预测过程是将所需预测的数据输入网络, 然后网络对此进行识别, 给出相应预测值. 学习过程采用 BP 学习算法.

2.2 实例分析

我们运用上节给出的非线性时间序列神经网络模型对太阳黑子年平均数问题进行了预测.

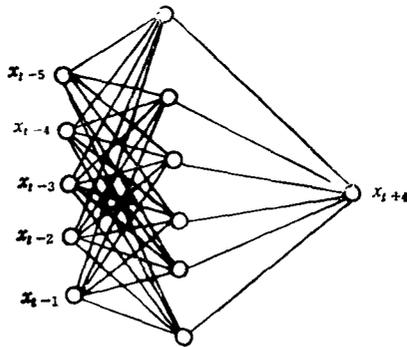


图 2

对第 2.1 节中的网络模型, 取 $L = 5, k = 5$, 则就有太阳黑子预测的网络拓扑结构图 2.

我们以 1868—1978 年为拟合区间, 原始数据共有 110 个, 由此可构造网络样本点 100 个. 它是原始数据 t_k , 经过变换

$$\left. \begin{aligned} v_k &= \log(t_k + 1), \\ x_k &= v_k / \log(200) \end{aligned} \right\} \quad (3)$$

得到的. 样本点结构为

$$(x_{t-5}, x_{t-4}, x_{t-3}, x_{t-2}, x_{t-1}, x_{t+4}), \\ (t = 6, \dots, 106).$$

取一个隐层, 其节点数为 28. 经过 18000 次迭代, 网络的总体误差为 0.00174. 用此网络模型对 1980—1987 年的太阳黑子年平均数进行了预测, 其结果如表 1 所示.

由表 1 可以看出, 神经网络模型要比传统预测方法好, 其最小均方差 (MSE) 要比线性预测和 SETAR 方法小得多. 这是因为神经网络方法更具有非线性特性, 更适合于识别非线性映射的变化特性.

表 1 1980—1987 年太阳黑子数预测结果

| 年份 | 观察值 | 线性预测 | 误差 | SETAR 预测 | 误差 | 神经网络预测 | 误差 |
|------|-------|-------|-------|----------|-------|--------|-------|
| 1980 | 154.7 | 162.8 | 8.1 | 174.7 | 20.0 | 143.0 | -11.7 |
| 1981 | 140.5 | 133.5 | -7.0 | 147.5 | 7.0 | 128.1 | -12.5 |
| 1982 | 115.9 | 91.5 | -24.4 | 92.4 | -23.5 | 104.8 | -11.1 |
| 1983 | 66.6 | 53.3 | -13.3 | 62.9 | -3.7 | 60.1 | -6.6 |
| 1984 | 45.9 | 23.1 | -22.8 | 24.1 | -21.8 | 43.3 | -2.6 |
| 1985 | 17.9 | 4.6 | -13.3 | 12.6 | -5.3 | 20.5 | 2.6 |
| 1986 | 13.4 | 2.0 | -11.4 | 0.2 | -13.2 | 13.3 | -0.1 |
| 1987 | 29.2 | 21.1 | 8.1 | 10.8 | 18.4 | 18.6 | -10.6 |
| MSE | | | 222.4 | | 253.9 | | 73.23 |

3 神经网络多维时间序列模型^[3]

本节提出的多维时间序列神经网络模型能够解决用传统方法难以预测或预测效果不佳的问题。

3.1 网络模型的拓扑结构

设 $x(t) = (x_1(t), \dots, x_r(t))^T$ 是 r 维时间序列, 那么 p 阶模型可一般表示为

$$x(t) = \Phi(x(t-1), \dots, x(t-p)), \quad (4)$$

其中 Φ 可以是线性或非线性函数。

构造 p 阶 r 维时间序列神经网络模型如图 3 所示。它完成了 $x(t-1)$ 到 $x(t)$ 之间的一个非线性映射。下面的定理保证了网络映射的精度。

定理 1 设有 $x(t) = (x_1(t), \dots, x_r(t))^T$, 具有如下映射关系:

$$x(t) = \Phi(x(t-1)) \quad (5)$$

且 $x(t), x(t-1) \in E^r$, 则由图 3 确定的三层神经网络映射 Φ' 可以与 Φ 无限接近。

证 这可由 Kolmogorov 定理直接推得, 详见文献[4]

网络的学习采用改进的 BP 算法, 采用分阶段重置权系数的方法, 可缩短学习时间。

3.2 应用实例

我们应用上述多维时间序列神经网络模型对新疆玛纳斯棉田的害虫进行了预测。基础数据包括 1980 年—1981 年在玛纳斯棉田每年 5—9 月期间, 每隔 5 天, 害虫烟蓟马、天敌纹蓟马的种群密度以及此时的旬平均温雨系数。在此, 我们之所以采用每隔 5 天采样一次是由植保学理论及农艺师的经验决定的。

我们采用的神经网络拓扑结构如图 4 所示。应用改进的 BP 算法进行 10000 次学

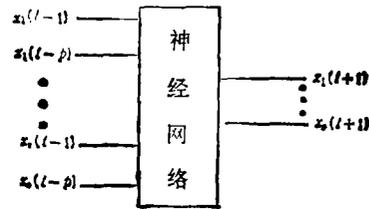


图 3

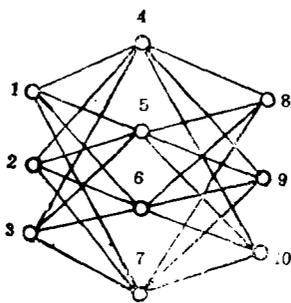
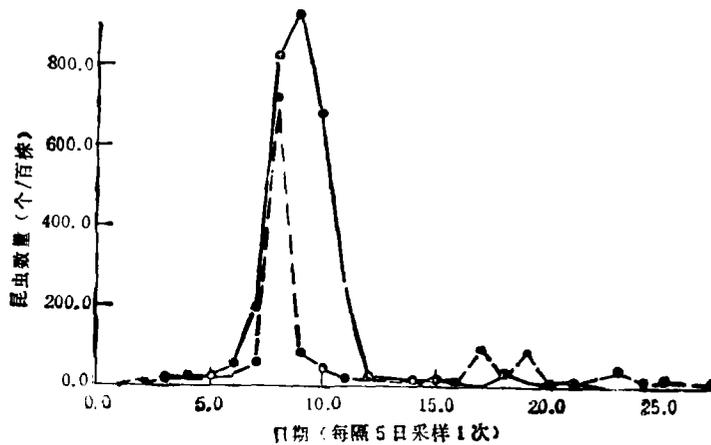


图 4



习,误差 $E = 0.020088$ 。利用此网络模型,对 1982 年的烟蓟马种群密度进行了预测,结果由图 5 给出。在图 5 中,虚线表示烟蓟马种群的预测密度,而实线表示烟蓟马种群的实际密度。由于每隔 5 天采样一次,从而共有 27 个值。

由图 5 中可以看出,预测值较好地反映了数据峰值的变化状态,这在害虫预防工作中是十分重要的。

4 神经网络的组合预测模型^[5]

4.1 神经网络组合预测模型

为了尽可能多地利用全部有用信息, Bates 和 Granger 于 1969 年提出了“组合预测”思想,并得到了很大的发展。但传统的组合预测方法具有很大的局限性。因此,我们首先推广了传统组合思想为广义组合原理。

原理 1 设对事件 F 有 m 种预测方法,则可用映射 $x \in X \subset R^n \xrightarrow{\varphi_i} y \in Y \subset R$ 表示第 i 种预测方法。广义组合预测原理是说不同预测方法的非线性组合函数

$$y = W(\varphi_1, \varphi_2, \dots, \varphi_m), \quad (6)$$

其中 W 是非线性映射,在某种测度之下, W 的度量要比 $\varphi_i(x) (i = 1, 2, \dots, m)$ 优越。

在此,我们引入神经网络技术来确定非线性映射 W 。采用四层前馈神经网络,其拓扑结构如图 6 所示。事实上,每种预测方法都有一个信息通道,而网络隐层的作用就在于综合各信息通道的信息,从而得到较优的组合预测值。

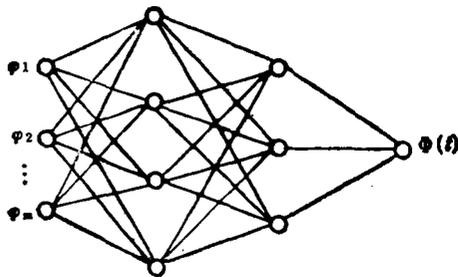


图 6

表 2 (单位: 亿元)

| 年 份 | 储蓄存款 | 最小平方法 | 三次指数法 | 神经网络法 |
|-----|--------|--------|--------|--------|
| 71 | 11.49 | 18.47 | 10.03 | 18.41 |
| 72 | 13.06 | 14.54 | 11.23 | 16.52 |
| 73 | 15.34 | 12.84 | 15.24 | 17.22 |
| 74 | 20.58 | 13.88 | 18.67 | 18.26 |
| 75 | 23.28 | 16.15 | 27.78 | 21.72 |
| 76 | 26.46 | 21.16 | 26.36 | 23.11 |
| 77 | 27.33 | 28.40 | 29.67 | 27.08 |
| 78 | 34.22 | 37.87 | 27.40 | 30.37 |
| 79 | 40.19 | 49.58 | 42.73 | 42.94 |
| 80 | 53.37 | 63.53 | 47.35 | 53.54 |
| 81 | 77.79 | 79.71 | 69.84 | 76.79 |
| 82 | 118.63 | 98.12 | 109.32 | 111.89 |
| MSE | | 553.97 | 188.87 | 107.49 |

4.2 应用实例

为说明我们所提出的非线性组合预测方法的有效性, 我们分别运用最小平方方法、三次指数平滑法和神经网络组合预测法对某种宏观经济指标——至上年末乡村居民标准储蓄存款余额进行了拟合, 数据来源于 1971—1982 年的国家统计年鉴, 拟合结果由表 2 给出。

而对 1983—1986 年的预测情况如图 7 所示。由拟合与预测的情况可以看出, 神经网络组合预测方法的结果最好。

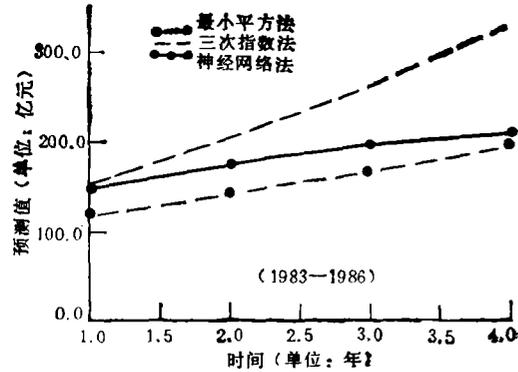


图 7

5 神经网络预测方法和智能信息处理

目前, 在范围广泛的经贸信息处理中, 神经网络技术可以解决许多用传统方法无法解决的问题。

但仍然有许多急待解决的问题, 如神经网络预测方法的基础理论的研究, 数据结构的研究以及如何将传统方法、专家知识库与人工神经网络进行有机的结合。

我们将传统预测方法和神经网络方法进行了对比, 各种性能特征如表 3 所示。从表中可以看出, 传统方法适合于抽象思维, 而神经网络更适合于形象思维。这与人思维机制的“左右脑”理论相吻合。

表 3

| 特 性 | 名 称 | 传 统 方 法 | 神 经 网 络 |
|-------|-----|---------|---------|
| 数值计算 | | * | 0 |
| 过程操作 | | * | 0 |
| 样本学习 | | 0 | * |
| 模式识别 | | 0 | * |
| 模糊数据 | | 0 | * |
| 可维护性 | | * | * |
| 适 应 性 | | * | 0 |
| 自适应性 | | 0 | * |
| 并行处理 | | 0 | * |
| 大规模性 | | 0 | * |
| 最大拟合 | | 0 | * |
| 高 精 度 | | * | 0 |
| 容 错 性 | | 0 | * |
| 解释结果 | | 0 | * |

注: * 为具有特性 0 为不具有特性

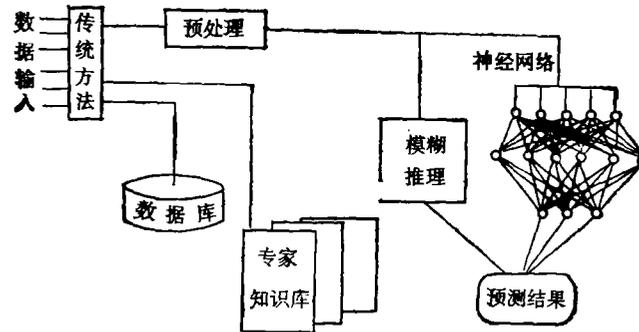


图 8

从而一种可能的智能信息预测系统可由图 8 表示。

可以预料,随着新的预测方法的产生,更加实用的现代智能预测系统的实现,将会给国民经济的发展带来巨大的效益。

致谢 作者在此十分感谢陈开周教授的悉心指导和帮助。

参 考 文 献

- [1] 翁文波. 预测论基础. 北京: 石油工业出版社, 1984, 第二章.
- [2] 文新辉, 陈开周. 西安电子科技大学学报, 1994, 21(1): 73—78.
- [3] 文新辉, 陈开周. 预测, 1993, 12(6): 48—51.
- [4] 文新辉, 牛明洁. 预测, 1992, 26(4): 58—61.
- [5] 文新辉, 陈开周. 神经网络在经济管理中的应用之二: 神经网络广义组合预测模型, 中国神经网络 1993 年学术大会, 西安, 1993, 1034—1042.
- [6] Bates J M, Granger CW J. Operations Research Quarterly, 1969, 20(2): 319—324.

TIME SERIES NEURAL NETWORK FORECASTING METHODS

Wen Xinhui

(The Central of Neural Network, Xidian University, Xi'an 710071)

Abstract This paper discussed the possibility and key problem of constructing the neural network time series model, and three time series neural network forecasting methods has been proposed. That is, the neural network nonlinear time series model, the neural network multi-dimension time model and the neural network combining predictive model. These three methods are applied to real problems. The results show that these methods are better than the traditional ones. Furthermore, the comparison of the neural network with the traditional methods and the constructed model of intellectual information forecasting system are given.

Key words Information theory, Information processing, Neural network forecasting method