

BP 网络预测能力仿真与分析

杜黎¹, 陈陶², 杜焰³, 周庆新⁴

(1. 昆明理工大学 生物与化学工程学院, 云南 昆明 650224; 2. 昆明理工大学 理学院, 云南 昆明 650093;
3. 昆明理工大学 材料冶金工程学院, 云南 昆明 650093; 4. 成都信息工程学院, 四川 成都 610041)

摘要: 人工神经网络用于预测时有多高的可靠性和精度? 是一个还没有得到很好探讨的问题。本文以 BP 神经网络为对象, 从数值仿真和欧氏空间理论的角度, 分析了预测样本与训练样本之间的关系对网络预测能力的影响, 给出一种预测精度的估计方法。

关键词: BP 神经网络; 预测; 欧氏空间

中图分类号: TP183 **文献标识码:** A **文章编号:** 1007-855X(2003)05-0097-03

Analysis on the Simulation of BP Network Predictive Ability

DU Li¹, CHEN Tao², DU Yan³, ZHOU Qing-xing⁴

(1. Faculty of Biological and Chemical Engineering, Kunming University of Science and Technology, Kunming 650224, China;
2. Faculty of Science, Kunming University of Science and Technology, Kunming 650093, China;
3. Faculty of Materials and Metallurgical Engineering, Kunming University of Science and Technology, Kunming 650093, China;
4. Chengdu Institute of Information Engineering, Chengdu 610041, China)

Abstract: How much of the reliability and precision are there when artificial neural network is used to predict a complex system? That is one hot point still getting heated discussion. Taking BP network as an object, from the value simulation and Euclidean space theory, the relations between forecast sample and training sample for the influence of network predictive ability are analyzed in order to obtain the way to estimate the precision of prediction.

Key words: BP network; prediction; Euclidean space

0 引言

在复杂系统的预测方法的研究中, 现在研究较多的是传统统计学方法和人工神经网络方法。传统的统计学方法存在不适合动态系统、建模复杂等局限性。与之相比, 作为新生事物的人工神经网络, 被普遍认为具有大规模非线性并行处理能力、依据数据自身的内在联系建模、良好的适应性与自学习能力、较强的抗干扰能力等等优良性能。因此, 与传统统计学方法比较, 人工神经网络用于复杂系统的预测就似乎具有了明显优势。

我们曾经将多种类型的人工神经网络, 用于几种典型的复杂系统的预测研究, 发现问题并不那么简单, 一个显而易见的问题是预测时的输入向量与网络训练时的输入向量一般属于不同的集合, 在这种情况下怎样判定预测结果的可靠性?

在各种类型的人工神经网络仿真实验中, 我们发现, 如果要将神经网络用于复杂系统的预测, 至少要解决两个方面的问题:

1) 输入向量与输出向量之间是否存在因果关系? 从本质上说, 神经网络只是一种非线性映射算法, 它本身并不能“创造”出因果律, 相反它是以因果律为基础的。如果离开因果律预测结果就不可靠。

2) 预测样本与训练样本之间要满足什么样的条件, 才能保证预测的可靠性?

严格地说怎样保证输入向量与输出向量之间存在因果关系, 已经不只是人工神经网络研究的问题。所

收稿日期: 2002-11-08.

第一作者简介: 杜黎(1958~), 男, 讲师; 主要研究方向: 复杂系统自控技术。

以本文是在假设这种关系已经成立的条件下,从仿真实验和欧氏空间理论两个角度,对第二个问题进行研究,分析训练样本、训练样本与预测样本之间的关系对网络的预测能力的影响,以及怎样根据这种关系,对预测精度进行估计。

因为 BP 神经网络是目前应用得比较普遍的一种神经网络,所以本文的讨论是以这种网络为基础进行的。但结果对其他类型的多层网络同样有效。

1 BP 神经网络的结构

常见的 BP 神经网络的结构是由一个输入层、2 至 3 个隐层、一个输出层所组成。输入层含 p 个输入单元,输出层含 q 个输出单元。

如果单元激活函数采用单调递增连续函数: $\varphi(x) = \frac{1}{1 + \exp(x - \theta)}$

则网络输出为: $F(x_1, \dots, x_p) = \sum_{i=1}^m \alpha_i \varphi(\sum_{j=1}^p w_{ij} x_j - \theta_i)$ (1)

并且有被证明的定理:

$\varphi(\cdot)$ 为有界、非常量的单调递增连续函数, I_p 代表 p 维单位超立方体 $[0, 1]^p$, $C(I_p)$ 表示定义在 I_p 上的连续函数构成的集合, 则给定任何函数 $f \in C(I_p)$ 和 $\varepsilon > 0$, 存在整数 M 和一组实常数 α, θ, w_{ij} , 其中 $i = 1, K, m, j = 1, K, p$ 使得网络输出可以任意逼近一个非线性连续函数, 即

$$|F(x_1, x_2, \dots, x_p) - f(x_1, x_2, \dots, x_p)| < \varepsilon, \forall (x_1, x_2, \dots, x_p) \in I$$

这个结果说明: 输入——输出数据属于训练样本集时, 网络输出与实际值无限靠近, 误差极限为零^[1]。

现在的问题是, 如果将 BP 网络应用于复杂系统的预测时, 输入向量一般不属于训练样本集, 这时就应该根据输入向量与训练样本集的关系估计预测结果的可靠性和精度。

2 仿真实验

根据我们的研究, 一个实际的复杂系统存在对应的“原因”系统, 这个系统可以用三维空间的螺旋线表示^[2]。即实际的复杂系统有状态参量集合 $\{Y\}$ 和对应的“原因”系统集合 $\{L\}$, $\{Y\}$ 与 $\{L\}$ 之间存在因果关系, 如果网络的输入向量来自于 $\{L\}$, 输出向量来自于 $\{Y\}$, 则网络的输入和输出之间就满足了因果律要求。在这种情况下, 网络的预测能力取决于网络本身和训练样本与预测样本的关系。所以, 为了对 BP 网络的预测能力进行仿真研究, 我们这样准备样本数据:

将美国道·琼斯指数及其对应的螺旋系统进行如下的变换:

1) 取道·琼斯指数的个位数构成的时间序列为输出向量集合 $\{\psi\}$, 对应的螺旋系统的数据序列构成 $\{\xi\}$ 。

2) 对 $\{\psi\}$ 和 $\{\xi\}$ 同时进行低频采样, 得 $\{L\}$ 和 $\{Y\}$, 进一步使样本元素具有随机变量的特征^[3]。

3) 将 $\{L\}$ 和 $\{Y\}$ 一分为二, 一部分为训练样本, 用于对 BP 网络进行训练。另一部分为预测样本, 预测样本中的 $\{L\}$ 为仿真时的输入向量, $\{Y\}$ 用于与仿真结果进行比较, 以便对预测结果进行评价。

经过这样的处理后, 样本中已经散失了道·琼斯指数的原有信息。自相关分析证明 $\{L\}$ 和 $\{Y\}$ 都是随机序列, 互相关分析证明 $\{L\}$ 和 $\{Y\}$ 没有相关性, 这增加了预测的难度。

用训练样本对 BP 网络进行训练, 在收敛后, 用预测样本进行仿真, 仿真结果如图 1 所示:

图中表现出这样几个特点:

1) 虽然实际值是一些随机数, 但表现出一定的涨跌特性, 1 至 13 号预测数据基本上反映出这种特性; 而 13 号以后的数据就差得多。

2) 7、9、12、19 号数据有很高的精度; 而 15、17、20 号预测数据是另一个极端, 预测的误差极大; 其他数据也具有一定的预测意义。我们认为这里面包含了规律性的东西。这些规律对人工神经网络的预测能力具有决定性的影响。

3 BP 神经网络预测能力分析

当输入数据不属于训练样本集时, 网络输出与实际值之间的情况如何呢? 或者说预测精度如何?

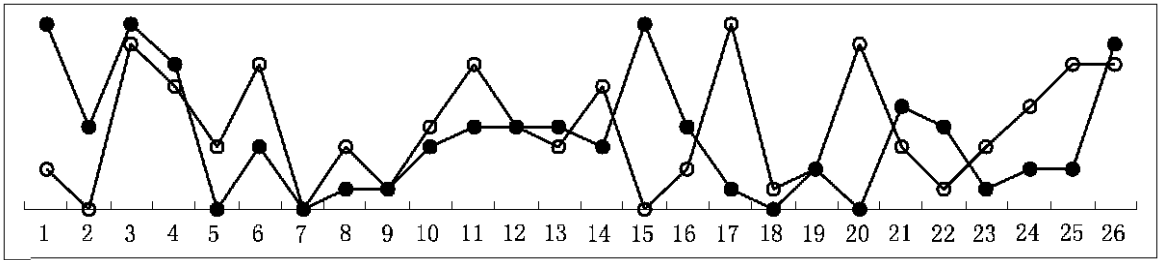


图 1 空心点为预测值, 实心点为实际值

设样本的输入为 p 维向量, 表示为 $x_i = \{x_{i1} \ x_{i2} \ \Lambda \ x_{ip}\}$, 相对应的实际值是 $Y_i = \{y_{i1} \ y_{i2} \ \Lambda \ y_{im}\}$, 其中 $x_{ij} \in R, y_{ij} \in R$. 输入向量可视为 p 维相空间中的点, 输出向量可视为 m 维相空间中的点.

在实际应用中, 集合 $\{X_i\}$ 中的元素是有限的.

向量的模为非负数 $\|X_i\| = \sqrt{x_{i1}^2 + x_{i2}^2 + \Lambda + x_{ip}^2}$, 最长、最短的模分别记为 $\|X\|_\infty$ 和 $\|X\|_0$, 样本 x_i 与 x_j 之间的距离为:

$$d_{i,j} = \|X_i - X_j\|, \text{ 其中 } i = 1, 2, \dots, n; j = 1, 2, 3, \dots, n; i \neq j; \text{ 显然有 } d_{i,j} = d_{j,i}.$$

设 x 为一个训练样本集以外的一个输入向量, y 为对应的网络输出, 与 x 距离最短的训练样本集中的样本是 x_i , 与 x 的距离次短的训练样本集中的样本 x_j , 由 x, x_i, x_j 组成相空间中的一个三角形.

根据式(1)的连续性, 可以有这样的推论:

推论 1 存在训练样本集中的 X_i, X_j 和 δ , 当 $\|X_i - X_j\| \leq \delta, \|Y_i - Y_j\| < \epsilon$

对于预测输入向量 X 有以下推理:

推论 2 当 $\|X_i - x_j\| \leq \|X_i - X_j\|$, 有 $\|Y_i - y\| < \|Y_i - Y_j\|$. 在这种情况下, BP 网络可以用于预测, 但还要看训练样本的“密度”, 因为只有训练样本的“密度”达到一定的程度, δ 的值才能足够小, 预测的精度才能达到实用的程度.

推论 3 $\|X_i - x\| > \|x_i - x_j\|$, 有 $\|Y_i - y\| > \|Y_i - Y_j\|$. 这时 BP 网络不能用于预测.

以上的讨论是在欧氏空间中进行的, 而(1)式的定义域是 p 维单位超立方体 $[0, 1]^p$, 所以需要在集合 $\{X_i\}$ 和 $[0, 1]^p$ 之间建立映射 A :

$$x_i \rightarrow \begin{cases} \frac{|x_j|}{|x_{i1}| + |x_{i2}| + \Lambda + |x_{ip}|} & |x_{i1}| + |x_{i2}| + \Lambda + |x_{ip}| \neq 0 \\ & |x_{i1}| + |x_{i2}| + \Lambda + |x_{ip}| = 0 \end{cases}$$

其中 $j = 1, 2, 3, \Lambda \Lambda p, i = 1, 2, 3, \Lambda \Lambda n$,

用此映射完成数据的归一化处理.

图一中显示的仿真结果证明了上述推论, 从图中可看出 7、9、12、19 号数据有很高的精度; 而与这几个数对应的输入样本正好满足推论 2 的要求. 15、17、20 号预测数据的误差极大, 这三个数对应的输入样本满足推论 3. 其他介于两者之间.

4 结论

在输入向量与输出向量之间存在因果关系的条件下, 如果满足:

- 1) 训练样本的“密度”足够地大, 这时存在足够小的 δ , 使得 $\|X_i - X_j\| \leq \delta$.
- 2) $\|X_i - x\| \leq \|x_i - x_j\|$.

BP 神经网络可以用于复杂系统的预测. 否则预测结果是不可靠的.

预测精度是由训练样本的多少、训练样本与预测样本之间的关系决定的. 预测精度的估计值为:

$$\|Y_i - y\| < \|Y_i - Y_j\|$$

参考文献:

- [1] 阎平凡, 张长水. 人工神经网络与模拟进化计算[M]. 北京: 清华大学出版社, 2000. 13~ 15.
- [2] 杜黎, 陈陶. 跨学科基本规律的螺旋表示[J]. 昆明理工大学学报(理工版), 2002, 27(4): 123.
- [3] 陈玉宏, 胡学敏. 自动控制原理[M]. 重庆: 重庆大学出版社, 1998. 155~ 156.