

自组织前向神经网络与非线性动态系统模化

申 弢 韩守木 黄树红 刘德昌
(华中理工大学动力工程系, 武汉, 430074)

TP18 TP271

摘要: 将自组织学习过程引入到前向网络的训练中, 提出了一种新的三层前向神经网络的训练方法. 训练过程首先利用自组织分簇算法确定隐含层结点的数目以及权值, 然后通过求解线性最小二乘问题估计输出层权值. 自组织过程产生的激活权值对输入数据具有一种特征变换的功能. 利用该方法训练的网络可以称之为自组织前向网络(SOFN). 文中通过实际非线性动态系统建模的例子, 说明了 SOFN 网络具有良好的性能.

关键词: 神经网络; 训练算法; 非线性系统; 自组织

文献标识码: A

Self-Organizing Feedforward Neural Network and Modeling of Nonlinear Dynamical System

SHEN Tao, HAN Shoumu, HUANG Shuhong and LIU Dechang

(Department of Power Engineering, Huazhong University of Science and Technology, Wuhan, 430074, P. R. China)

Abstract: In this paper a new learning procedure of MLP is presented which named as self-organizing feedforward neural Network (SOFN). The optimization of weights is implemented layer by layer. At the stage of training hidden weights, an unsupervised self-organizing clustering is introduced, then the weights of output layer are estimated by supervised least-square algorithms. With self-organizing stage, the number of hidden nodes can be determined automatically, furthermore, the hidden-layer weights created by clustering work as a feature transformation matrix for input vectors. Two examples are given to show the feasibility and advantages of the approach, which is particularly suitable for modeling of nonlinear dynamical system.

Key words: neural network; training algorithm; nonlinear system; self-organization

1 引言 (Introduction)

确定神经元数目以及相应权值, 是用神经网络模化非线性过程问题的关键. 通常前向网络中的权值采用后向传播 (BP) 算法估算, 但 BP 算法速度慢而且收敛性能差. 另外一些研究表明, 基于参数估计思想的网络训练速度可以显著提高^[1], 但这些方法多是采用径向基函数网络而不是标准的线性输入函数. 随机激活权值网络 (RAWN) 虽然可以快速有效地估算权值^[2], 但随机选取的隐含层权值意义不明显, 而且对于隐层结点数目的选择, 也只是采用反复尝试的方法.

本文将自组织学习的思想引入到前向网络的训练过程中, 提出一种新的三层前向网络学习方法——自组织前向网络 (SOFN). 网络的学习分两点进行, 其中隐含层结点的数目及输入层至隐含层的连接权值通过自组织的方式学习获得. 输出层权值通过线性最小二乘估计得出. 自组织学习产生的隐含层权值相当于一个特征变换矩阵. 对于输入、输出已

定的非线性系统的模化, SOFN 可以自动选择合适的隐含结点的数目, 因此更适于实际应用.

2 自组织前向神经网络 (Self-organizing feedforward neural network)

2.1 前向网络的分步训练 (Stepwise training of feedforward neural network)

单隐含层前向神经网络的映射关系可以利用矩阵形式描述为:

$$Z = XW^h, U = f(Z), Y = UW^o \quad (1)$$

其中 $X \in \mathbb{R}^{N_s \times (N_i + 1)}$ 为输入矩阵, $Y \in \mathbb{R}^{N_s \times N_o}$ 为输出矩阵, $W^h \in \mathbb{R}^{(N_i + 1) \times N_h}$ 为隐含权值矩阵, $W^o \in \mathbb{R}^{N_h \times N_o}$ 为输出权值矩阵, $Z, U \in \mathbb{R}^{N_s \times N_h}$ 为中间结果, N_s 为样本总数, N_i 为输入结点数, N_o 为输出结点数, $f(\cdot)$ 为非线性激活函数, 本文的研究中取为 tanh 函数.

网络训练的目的就是针对一定的目标函数寻找优化的 W^h, W^o . 由于需要同时优化两组参数, 而且中间还有非线性激活函数的作用, 因此是一个高度

非线性的优化问题. 对于该问题的求解, 最优解与求解的速度往往是互相矛盾的. 基于 RAWN 网络的分步训练思想^[2], 本文提出了一种通过无导师学习确定隐含层结点数目以及激活权值的 SOFN 网络学习方法, 其在满足一定的精度要求下, 训练速度是相当快的. 训练分两个步骤: 首先对输入向量形成的模式空间进行自组织分簇(聚类), 分簇结果生成一个“标准”模式矩阵 $E = [e_1, e_2, \dots, e_n]$, 其中 $e_k = [e_{k1}, e_{k2}, \dots, e_{kN}]^T$ 为簇中心矢量. 选择簇中心的数目为隐含层结点数, 而且将簇中心矢量作为对应隐含结点的激活权值. 然后通过求解最小二乘问题得到优化的 W^0 . 这是因为, 如果分簇算法是收敛的、稳定的, 则不可能产生两个相同的簇中心. 另外, 假定 X 阵的元素不全部为零且样本数大于输入特征数. 则矩阵 Z 中出现两列相同的概率为:

$$P[z_i = z_j] \Big|_{i \neq j} = P\left[\bigcap_{k=1}^{N_e} [x_{ki}e_i = x_{kj}e_j]\right] \Big|_{i \neq j}.$$

其中 $P\left[\bigcap_{i=1}^n A_i\right]$ 为 n 个事件同时发生的概率. 由假设条件不难得出 $P[z_i = z_j] \Big|_{i \neq j} = 0$. 同时非线性激活函数的作用可以使 U 阵中也不可能有完全相同的列, 因此 U 就是一个满阶矩阵, 则 W^0 的优化就是一个可以求解的最小二乘问题. 另外由于隐含层权值为分簇后的簇中心矢量, 它对输入模式具有一种特征变换的功能, 因此 SOFN 网络的分步优化方法在某种程度上要优越于 RAWN 网络. 文中将通过例子说明 SOFN 网络的这种优越性.

2.2 分步训练的过程(Procedures of stepwise training)

在 SOFN 分步训练中, 我们选择 DIGNET 分簇算法^[3]对隐含层权值进行自组织学习. DIGNET 算法能够根据用户定义的相似性测度(本文定义为矢量之间的内积), 自动地提取输入数据的内在特征以产生新的簇中心. 因此自组织过程不仅可以给出优化的隐层权值, 而且隐层结点的数目也可以随之自动确定. 有关算法的详细说明参见文献^[3].

隐含层权值的自组织学习过程可以归结如下: 首先设置算法有关的参数, 如活动周期长度 SA 、簇强度阈值、相似性测度的阈值等. 设 x_n 为第 n 次输入矢量, 则:

1) 计算规范化后输入的矢量 x_n 与已存储的簇中心 e_{n-1} 之间的相似度.

2) 经过竞争选择后, 确定每个簇中心的更新状态 c_n . 如果输入矢量落入某个簇内且该簇在竞争中获胜, 则 $c_n = 1$; 没有输入矢量落入的簇, 其 $c_n = 0$;

如果输入矢量落入某个簇内但该簇在竞争中失败了, 则 $c_n = -1$.

3) a) 如果所有簇的更新状态 c_n 均为 0, 则将 x_n 初始化为一个新的簇. 否则, b) 接下式更新簇的强度 d_n 与中心 e_n : $d_n = d_{n-1} + c_n$; $e_n = \frac{c_n}{d_n}x_n + \frac{d_{n-1}}{d_n}e_{n-1}$. 初始时 $d_0 = 0, e_0 = 0$.

4) 如果活动周期(SA)到, 则检查每个簇的强度. 删除不必要的簇.

5) 接受下一个输入矢量, 重复 1)~4).

理论分析表明^[3], 上述的分簇聚类过程是稳定的、收敛的, 因此产生的簇中心矢量, 可以直接作为前向神经网络的隐含层权值, 而且利用常规的最小二乘求解方法可以很方便地计算出输出层权值的估计值, 具体方法在此不再进行论述.

3 非线性动态系统建模的实例考核(Examples of modeling for nonlinear dynamic system)

3.1 实例 1(Example 1)

在实例 1 中, 我们通过文献^[2]中给出的用以测试 RAWN 网络的例子, 说明 SOFN 网络具有的良好性能. 考虑辨识下述方程描述的一个实际非线性系统的模型:

$$\begin{aligned} y(k) = & \\ & [y(k-1)y(k-2)y(k-3)u(k-2)(y(k-3)-1) + \\ & u(k-1)]/[1+y^2(k-2)+y^2(k-3)]+w(k). \quad (2) \end{aligned}$$

其中 w 为 $\mu = 0, \delta = 0.02$ 的白噪声. 假设测试时的系统输入为一个多频率成分的正弦信号

$$\begin{aligned} u(k) = & 0.5\sin\frac{2k\pi}{300} + 0.4\sin\frac{2k\pi}{200} + 0.3\sin\frac{2k\pi}{100} + \\ & 0.2\sin\frac{2k\pi}{20} + 0.1\sin\frac{2k\pi}{10}. \quad (3) \end{aligned}$$

训练神经网络时的输入为:

$$\begin{cases} u(k) = \sin\frac{2k\pi}{500} (k \leq 500), \\ u(k) = 0.8\sin\frac{2k\pi}{250} + 0.2\sin\frac{2k\pi}{25}, \\ 500 < k \leq 600. \end{cases} \quad (4)$$

文献^[2]采用 5-20-1 的网络结构. 经过训练后, RAWN 网络可以对测试信号进行“自由”预测, 即测试时网络的输入取自网络自身的输出而不是实际的系统输出. 我们利用文中提出的方法进行同样的试验, 训练及测试结果如图 1 所示.

图中含噪声的线为实际曲线, 光滑的为网络输入曲线. 残差自相关系数图中, 两条虚线间代表了

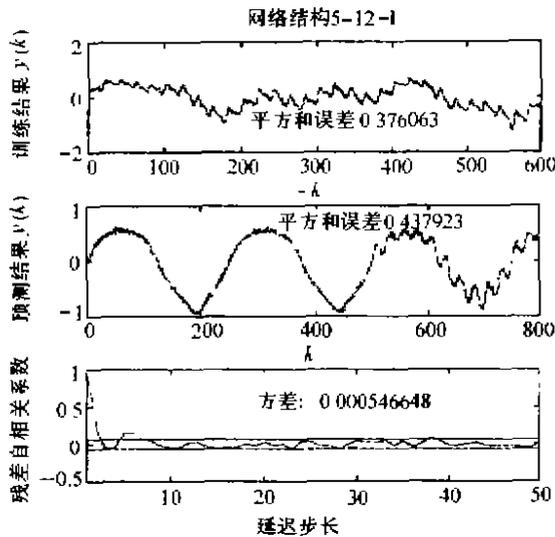


图1 一个非线性系统建模的试验结果

Fig. 1 Results of modeling for a nonlinear system

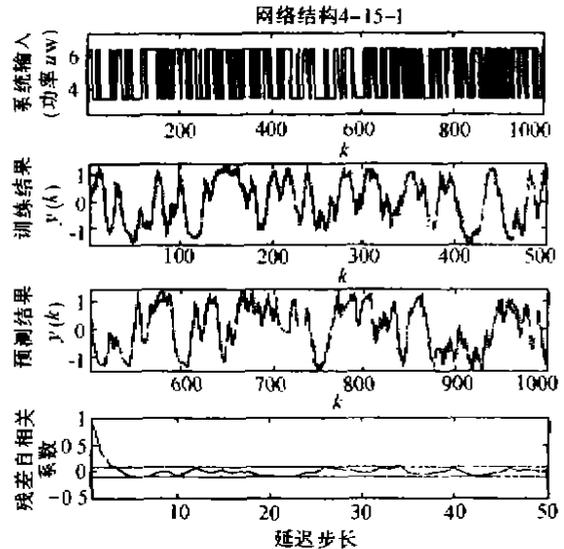


图2 一个实际的烘干机系统的建模结果

Fig. 2 Results of modeling for a real-life dryer

95%置信度区域。由图可以看出,训练与测试的误差以及方差均很小,因此,SOFN网络的模化结果是令人满意的。需要说明的是,在分步训练的无导师学习阶段,网络只生成了12个隐含结点。即用所提出的方法,可以采用较精简的网络结构对同样的系统进行成功的辨识。这种现象的原因是,通过无导师的学习阶段产生的隐含层权值,不是一个随机的数值,而是对输入矢量潜在结构提取的一种特征。经过隐含层的处理,输入数据相当于变换到隐含权值矢量张开的特征空间上,变换后的数据更能突出系统的潜在特征,因此对同一个非线性系统进行模化时,所提出的方法可以产生较精简的结构。SOFN网络的这种优点具有一定的普遍性。

3.2 实例2(Example 2)

在实例2中,我们采用了MATLAB软件提供的一个实测的烘干机工作过程中系统工作功率(输入 u)与出口处热风温度(输入 y)数据。总的采样点数为1000个,试验中取前一半为训练集,后一半为测试集。神经网络的输入、输出数据均经过去均值与归一化处理。其中温度值 y 带有 $\mu = 0, \delta = 0.02$ 的白噪声。根据MATLAB示例程序的建议,输入矢量取为 $(y(k-1), y(k-2), u(k-1), u(k-2))$,输出矢量取为当前的热风温度 $y(k)$ 。SOFN网络对该实际数据的模化结果如图2所示。图中的各标识含义与图1相同。由图可以看出,对于该实际的一个烘干机工作过程,SOFN网络通过采用4-15-1的结构可以给出满意的模化结果。

4 结论(Conclusion)

本文在前向网络的训练过程中引入了自组织学习环节,提出了一种新的三层前向神经网络的训练

方法——自组织前向网络(SOFN)。训练分两步进行,首先利用自组织分簇算法确定隐含层权值,然后将输出层权值的求解转化为一个最小二乘优化问题;SOFN网络不仅可以自动确定隐结点数,而且激活权值对输入数据可以起到一种特征变换的功能。因此在对非线性系统进行模化时,SOFN网络可以采用较精简的结构而取得满意的效果。文中通过与RAWN网络的对比以及实际的非线性动态系统的例子,说明了SOFN网络具有的良好性能。

参考文献(References)

- [1] Scalerto R S and Tependelenioglu N. A fast new algorithm for training feedforward neural network[J]. IEEE Trans. on Signal Processing, 1992, 40(1): 202-210
- [2] te Braake H A B and van Straten G. Random activation weight neural network (RAWN) for fast non-iterative training[J]. Engineering Applications of Artificial Intelligence, 1995, 8(1): 71-80
- [3] Thomopoulos C A and Bougoulis D K. Dignet: a self-organizing neural network for automatic pattern recognition and classification [A]. In: Proc. Int. Joint IEEE and International Neural Network Society Conf. on Neural Networks[C], Singapore, 1991, 2683-2692

本文作者简介

申 强 1973年生,华中理工大学动力工程系博士研究生,主要研究方向为:人工智能,神经网络等在机械故障诊断中的应用。

韩守木 1936年生,华中理工大学动力工程系教授,主要研究方向为:汽轮发电机组的振动分析,监测与诊断等。

黄树红 1960年生,华中理工大学动力工程系教授,主要研究方向为:汽轮发电机组的振动分析,监测与诊断等。

刘德昌 1934年生,华中理工大学煤燃烧国家重点实验室教授,博士生导师,主要研究方向为:洁净煤燃烧技术,循环流化床燃烧技术等。