

文章编号: 1001-0920(2007)11-1235-06

自组织多模型逼近性能分析与改进

高大远, 祝晓才, 胡德文

(国防科技大学 机电工程与自动化学院, 长沙 410073)

摘要: 针对基于自组织映射神经网络的非线性函数逼近, 研究其方法和原理, 指出它与一般前向神经网络在逼近原理上的不同. 在此基础上, 进一步研究该方法的逼近性能, 分析其两个不足之处, 进而提出一种提高逼近性能的改进神经网络训练策略. 最后通过仿真实例验证了所得结论, 表明了改进方法的有效性.

关键词: 自组织映射; 神经网络; 多模型逼近

中图分类号: TP183 **文献标识码:** A

Performance analysis and improvement of multiple models approximation using self-organizing map neural network

GAO Da-yuan, ZHU Xiao-cai, HU De-wen

(College of Mechatronics and Automation, National University of Defense Technology, Changsha 410073, China.

Correspondent: HU De-wen, E-mail: dwhu@nudt.edu.cn)

Abstract: The method of multiple models approximation using self-organizing map neural network is analyzed. The mechanism of this method is different to that of approximation method using other kinds of feedforward neural networks. The disadvantages of this method are analyzed. Based on the analyses, an improved training method for the self-organizing map neural network is proposed to reduce the approximation error. Simulation examples illustrate the conclusions and show the effectiveness of the new training method.

Key words: Self-organizing map; Neural network; Multiple models approximation

1 引言

随着被控对象的日益复杂, 神经网络在非线性系统建模与控制中的应用得到越来越广泛的研究. 神经网络用于非线性系统的建模与控制, 是基于其对非线性函数的逼近. 已经证明^[1-3], 给定充分多的隐层神经元节点, 前向神经网络能以任意精度逼近紧集上的连续函数. 随着先导性研究工作^[4,5]的发表, 各种神经网络和相应的系统结构被提出并用于非线性系统的建模与控制^[6].

由于自组织映射(SOM)神经网络^[7]对样本空间具有很好的映射性质, 被广泛地用于非线性系统的建模与控制^[8]. 基于 SOM 的多模型逼近, 其思想是利用“分而治之”的方法, 在每一个小的区域对非线性系统进行逼近. 基于这一思想, SOM 已被用于混沌时间序列预测^[9,10], 机械手的手眼协调控

制^[11], 跨音速风洞的预测控制^[12]以及无人机的建模与控制^[13]等一些非线性特征较强的应用中. Principe 等^[9]指出, 用 SOM 局部动态模型对混沌时间序列建模, 所得到的预测精度优于用多层前向感知器神经网络进行全局建模所得到的精度.

虽然 SOM 已广泛地用于非线性系统的建模与控制, 但基于 SOM 的非线性函数逼近的原理和性能分析却很少见, 本文对此进行研究. 首先对 SOM 用于函数逼近的原理进行分析, 并与一般的前向神经网络逼近原理进行比较, 指出它们的不同之处; 然后分析 SOM 神经元节点分布对空间划分的特点及其对逼近性能的影响, 指出 SOM 的空间划分规则不能使逼近性能达到最优, 针对训练样本分布对逼近性能的影响, 提出一种改进的自组织多模型逼近方法; 最后通过仿真实例对文中的分析和方法进行

收稿日期: 2006-07-23; 修回日期: 2007-01-29.

基金项目: 国家自然科学基金重点项目(60736018); 国家自然科学基金项目(60675005); 高等学校优秀青年教师教学科研奖励计划项目.

作者简介: 高大远(1978—), 男, 新疆石河子人, 博士生, 从事非线性控制、神经网络控制等研究; 胡德文(1963—), 男, 湖南岳阳人, 教授, 博士生导师, 从事系统辨识、神经网络等研究.

了说明.

2 SOM 多模型逼近方法分析

2.1 SOM 神经网络

SOM 神经网络由芬兰 Kohonen^[7]教授提出,它是模拟生物神经系统的自组织形成过程而发展起来的一种人工神经网络.一个基本的 SOM 神经网络有 n 个输入,对应于 n 维输入数据;神经元节点个数为 N ;各神经元节点相互连接,构成规则的网格,网格维数可以是一维也可以是高维.对于每个神经元节点 j ,有一个参考向量 $\mathbf{r}_j = [r_{j1}, r_{j2}, \dots, r_{jn}]^T$ 与之对应,它表示此节点在输入空间中的位置. SOM 对输入样本空间具有很好的映射性质.对于一维 SOM 网络和一维输入数据,已经证明^[7],在一定条件下,神经元节点的分布可以形成一种有序的拓扑结构,并且神经元节点分布的点概率密度是训练样本数据分布概率密度的单增函数.对于高维数据,这个性质还没有严格的理论证明,但其原理已经清楚,并已通过大量的仿真和应用实例得到验证.

2.2 SOM 多模型逼近原理

在 SOM 多模型逼近过程中, SOM 通过其神经元节点的分布,实现对工作区域的划分,而对应于每个神经元节点(即每个子区域),都有一个局部模型(通常是线性模型)负责对相应子区域内的函数进行逼近.

对于非线性函数 $y = f(x), x \in D \subset R^n, y \in R^m$,用含有 N 个神经元节点的 SOM 神经网络在区域 D 上对其进行逼近.对于每个神经元节点 j ,有一个局部模型 $M_j: \hat{y} = \hat{f}_j(x)$ 与之对应.假设样本数据集 X 和 Y 已经得到,逼近的一般步骤为:

- 1) 用 X 中的样本对 SOM 神经网络进行训练;
- 2) 根据所得 SOM,将数据集 X 和 Y 分成子集合

$$\begin{aligned} X_j &= \{x \in X \mid x \text{ 被映射到神经元 } j\}, \\ Y_j &= \{y \in Y \mid y = f(x), x \in X_j\}, \\ & \quad j = 1, 2, \dots, N; \end{aligned}$$

- 3) 用 X_j 和 Y_j 中的数据对 M_j 中的参数进行辨识.

上面步骤中,当 SOM 训练完成后,就自然地将区域 D 划分成多个子区域 D_j

$$D_j = \{x \in D \mid \|x - \mathbf{r}_j\| < \|x - \mathbf{r}_k\|, k \neq j\}, \quad j = 1, 2, \dots, N. \quad (1)$$

即 D_j 中的点到神经元节点 j 的距离小于到其他神经元节点的距离.而样本子集 X_j 和 Y_j 可以看成是从这些子区域中分别选取的.

在这样一个基本步骤中, SOM 的训练和局部模型的辨识是显式的、分开的,而实际上它们可以同时进行.在文献[14]中,当每一步获胜神经元节点被选中后,不仅各节点的参考向量被更新,节点对应的局部模型的参数也同时被更新,参数更新的方式与节点更新的方式一样,都被邻域作用函数和学习速率函数加权.随着 SOM 网络趋于稳定,参数的辨识也逐渐收敛.这样做的好处是可以使建模在线进行,但与上面步骤相比,并没有提高逼近精度.

2.3 SOM 多模型逼近分析

为简化分析,假设区域 D 是一个单连通紧集, $y = f(x)$ 在 D 上连续,由 SOM 划分的子区域 D_j 也都是单连通的紧集.在每个子区域 D_j 上,局部模型是连续的.每个区域上的逼近性能用下式度量:

$$J_j = \int_{D_j} [\hat{f}_j(x) - f(x)]^T [\hat{f}_j(x) - f(x)] dx, \quad (2)$$

总的逼近性能为 $J = \sum_{j=1}^N J_j$. 令

$$e_j(x) = [\hat{f}_j(x) - f(x)]^T [\hat{f}_j(x) - f(x)] \geq 0,$$

表示局部模型 M_j 在子区域 D_j 内某一点的逼近误差,则它在 D_j 上连续,由积分中值定理,有

$$\begin{aligned} J_j &= \int_{D_j} e_j(x) dx = e_j(\xi_j) \int_{D_j} dx = \\ & \quad e_j(\xi_j) \text{Vol}(D_j) \triangleq \bar{e}_j V_j. \end{aligned} \quad (3)$$

其中: $\xi_j \in D_j, V_j = \text{Vol}(D_j)$ 表示区域 D_j 的“体积”.即每个子区域上的逼近性能,可由局部模型在此区域上的平均逼近误差与区域“体积”的乘积得到.由于 D_j 是紧集,且 $e_j(x)$ 在 D_j 上连续,从而有界.设对于 $\forall x \in D_j, e_j(x) \leq e_{j\max}$, 则

$$J_j = e_j(\xi_j) V_j \leq e_{j\max} V_j \triangleq J_{j\max}. \quad (4)$$

于是总的逼近性能指标为

$$J = \sum_{j=1}^N J_j \leq \sum_{j=1}^N e_{j\max} V_j \leq e_{\max} \sum_{j=1}^N V_j = e_{\max} V.$$

其中: $V = \sum_{j=1}^N V_j$ 为区域 D 总的“体积”, $e_{\max} = \max_{1 \leq j \leq N} \{e_{j\max}\}$ 表示所有局部模型逼近误差的最大值.对于任意小的正实数 $\epsilon > 0$,只要令 $e_{\max} < \epsilon/V$,即可得到 $J < \epsilon$.假设每个子区域上都可以达到最优逼近,通过增加神经元个数,增加子区域个数,减小子区域的范围,可以使得在每一个子区域上的逼近误差界 $e_{j\max}$ 任意小,从而实现 $J < \epsilon$.

从上面分析可以看出,基于 SOM 的自组织多模型逼近,与其他前向神经网络对非线性函数的逼近相比,具有相同的结果,但在原理上二者有本质的区别.一般的前向神经网络在对非线性函数逼近时,是

对所有(隐层)神经元节点的活动函数进行线性或非线性的加权,这样所得到的逼近是在整个区域 D 上进行的,形成的模型是连续的,采用的逼近原理是 Stone-Weierstrass 定理;而基于 SOM 的逼近,它和对应的一组局部模型一起实现对非线性函数的逼近,所得到的模型是分段连续的,其逼近原理是基于连续函数在紧集上的有界性.

3 SOM 多模型逼近性能分析

由于 SOM 神经网络的神经元节点不能无限多,总希望通过给定的神经元数量实现尽可能高的逼近精度.如果给定非线性函数和局部模型结构,则在子区域上的最小逼近误差只与子区域有关,从而在每个局部模型都达到子区域上的最优逼近的假设下,总体逼近性能就变成只与子区域划分有关的函数.一般而言,给定子区域个数 N ,总存在一个最优划分,使得性能指标最优.但对于 SOM 神经网络而言,这样的划分不一定能够通过其神经元节点分布得到.

为了更清楚地理解这一点,以一个一维非线性函数的多线性模型逼近为例.对于连续非线性函数 $y = f(x), x \in [a, b]$,将区间 $[a, b]$ 划分为 N 个子区间 $[x_{j-1}, x_j], j = 1, 2, \dots, N$,其中

$$a = x_0 < x_1 < \dots < x_{N-1} < x_N = b.$$

在每一个区间上用线性函数 $\hat{y} = a_j x + b_j$ 对原函数进行逼近.局部性能指标为

$$J_j = \int_{x_{j-1}}^{x_j} [a_j x + b_j - f(x)]^2 dx = \frac{1}{3} a_j^2 (x_j^3 - x_{j-1}^3) + b_j^2 (x_j - x_{j-1}) + a_j b_j (x_j^2 - x_{j-1}^2) + F_{2j} - 2a_j F_{1j} - 2b_j F_{1j}.$$

其中

$$F_{2j} = \int_{x_j}^{x_{j-1}} f^2(x) dx,$$

$$F_{1j} = \int_{x_{j-1}}^{x_j} f(x) dx,$$

$$F_{sj} = \int_{x_{j-1}}^{x_j} x f(x) dx,$$

总体性能指标 $J = \sum_{j=1}^N J_j$.为使每个子区间上性能指标最优,令 $\frac{\partial J_i}{\partial a_i} = 0, \frac{\partial J_i}{\partial b_i} = 0$,则所求出的参数 a_j, b_j

只与区间端点 x_{j-1} 和 x_j 有关,即

$$a_j = \frac{12 F_{2j} - 12 c_j F_{1j}}{l_j^3},$$

$$b_j = \frac{(12 c_j^2 + l_j^2) F_{1j}}{l_j^3} - \frac{12 c_j F_{sj}}{l_j^3}.$$

其中 $c_j = \frac{x_{j-1} + x_j}{2}, l_j = x_j - x_{j-1}$.从而所得到的最

优性能指标 J_j^* 也只与 x_{j-1} 和 x_j 有关,

$$J_j^* = - \frac{12}{l_j^3} (c_j F_{1j} - F_{sj})^2 - \frac{1}{l_j} F_{1j}^2 + F_{2j}.$$

而总的性能指标则与所有子区间的划分有关,

$$J = \sum_{j=1}^N J_j^*(x_{j-1}, x_j) = J(x_1, x_2, \dots, x_{N-1}).$$

注意到这里 x_0 和 x_N 是固定的.假设存在一个区间的划分, $a = x_0 < x_1^* < \dots < x_{N-1}^* < x_N = b$,使得总性能指标最小.

对于一维 SOM 神经网络,设神经元节点参考值为 $x_j (j = 1, 2, \dots, N)$,神经元节点有序分布,即 x_j 按 j 的降序或升序排列.不失一般性,设

$$a < x_1 < x_2 < \dots < x_{N-1} < x_N = b,$$

则这 N 个神经元节点将区间 $[a, b]$ 划分为 N 个子区间 $[x_{j-1}, x_j]$,由 SOM 神经元节点划分区间的规则(1)可知, x_j 为 x_{j-1} 和 x_{j+1} 的中点,即

$$x_j = \frac{x_{j-1} + x_{j+1}}{2}, j = 1, 2, \dots, N - 1. \quad (5)$$

下面考虑这一规则对所划分区间的约束.考虑相邻的 3 个区间 $[x_j, x_{j+1}], [x_{j+1}, x_{j+2}], [x_{j+2}, x_{j+3}]$,由式(5)有

$$x_{j+2} = 2x_{j+1} - x_j,$$

$$x_{j+3} = 2x_{j+2} - x_{j+1} = 2x_{j+2} - 2x_{j+1} + x_j.$$

注意到 $x_{j+1} > x_j$, 并且 $x_{j+3} < x_{j+2}$, 从而有

$$x_{j+3} > 2x_{j+2} - 2x_{j+1} + x_j > 2x_{j+2} - 2x_{j+1} + x_j,$$

即

$$(x_{j+3} - x_{j+2}) + (x_{j+1} - x_j) > (x_{j+2} - x_{j+1}). \quad (6)$$

因此, $[x_{j-1}, x_j]$ 族是由 SOM 神经元节点划分而成的必要条件是在任意 3 个相邻的子区间中,两端的子区间长度之和必须大于中间子区间长度.但使得总性能指标最优的划分 $[x_{j-1}^*, x_j^*]$ 并不一定满足这个条件,所以,它不一定能由 SOM 神经网络划分实现.对于高维网络而言,也有类似的情况.因此,在神经元分布有序性和子区域划分规则的约束下, SOM 自组织多模型逼近不一定能达到 N 个区域情况下的最优逼近.

上述分析是在 SOM 神经元节点可以任意分布的条件下(满足有序性)进行的,结论是 SOM 多模型逼近不能达到任意子区域划分条件下的最优性.另外,在实际应用中, SOM 的训练受样本数据分布的影响较大,其神经元分布满足概率密度匹配性质,因此,很难使 SOM 神经元节点的分布对于逼近性能达到最优.这一点,在后面的仿真可清楚地看到.

4 改进的 SOM 多模型逼近方法

前节分析了 SOM 多模型逼近中的两个问题.



第 1 个问题是由 SOM 神经元节点分布的有序性和空间划分的原则所致,而这两点正是 SOM 的基本特点,很难进行改变.因此,一般情况下是在神经元分布的约束下寻找最优逼近.第 2 个问题则由 SOM 网络的映射性质决定.虽然基本的 SOM 学习是一个无监督的过程,没有显式的目标函数,但其隐含实现的目标是网络对样本空间的划分使得样本的聚类(量化)误差达到最小.这与函数逼近的目标显然不一致,因此直接采用 SOM 与多模型的结合,导致了所得结果不能达到最优.

对此,目前的一些改进集中于增加神经元个数,采用非线性局部模型,使用一些变形模型,如神经气模型^[14]、增长节点模型^[15]、扩散 SOM 模型^[16]、同时采用不同的优化算法^[17]等.然而,这些方法不仅增加了模型的复杂程度,而且并未有效地改变神经元的分布,因此改进效果并不理想.

考虑下面一种方法.对于非线性函数 $y = f(x), x \in D \subset R^n, y \in R^m$, SOM 神经元个数为 N , 对每个神经元节点 j , 有一个局部模型 $M_j: \hat{y} = \hat{f}_j(x)$ 与之对应.函数的输入输出样本数据集为 X 和 Y . 首先进行基本的训练,即用 X 中的数据对 SOM 进行训练,得到训练子集 X_j 和 Y_j , 并对局部模型 M_j 进行辨识.然后,分别计算各局部模型逼近的均方误差

$$MSE_j = \frac{1}{K_j} \sum_{x \in X_j} (\hat{f}_j(x) - f(x))^2, \quad j = 1, 2, \dots, N.$$

其中: K_j 是子集 X_j 的样本个数, $K = \sum_{j=1}^N K_j$ 为总样本个数.令

$$j = \frac{MSE_i}{\sum_{i=1}^N MSE_i}, \quad K_j = \text{round}(jK),$$

其中 $\text{round}(\cdot)$ 表示四舍五入取整.从 X_j 中随机取出 K_j 个样本,形成新的样本子集 X_j , 这些样本子集的合集构成新的样本集 $X = \sum_{j=1}^N X_j$. 用此样本集继续进行下一次的训练,这样形成一个多批次的训练.设在新一批次的训练中,训练样本 x 从 X 中随机抽取,则 x 位于上一次训练得到的子区域 D_j 的概率为

$$P\{x \in D_j | x \in X\} = \frac{K_j}{K}$$

这是因为,由 X_j 和 X 的构成方法可知, X 中所有位于 D_j 的样本都属于子集合 X_j , 并且其他子集合不含有位于 D_j 的样本.通过这种样本选择方法,在新的训练样本集里,前一批次训练中逼近误差较大

的子区域(j 较大)内,样本出现的概率也较大,则在当前训练中,较多的神经元也会分布于此区域.

这样做的好处有两个:1) 子区域范围减小,平均逼近误差也减小;2) 区域的体积减小,从式(4)可知,子区域上的总误差会降低.对于那些子区域范围扩大的地方,因其平均逼近误差较小,故其总误差不会增加太多.这样总体逼近性能得到提高.需要注意的是, SOM 神经元分布对训练数据分布的依赖性只是一个定性的结果,而且局部最优逼近性能对于子区域的变化较为敏感,因此上面的过程只是一个近似分析,具体的定量计算很难进行.总的目标是使更多的局部模型向比较难以逼近的区域分布,以提高总体逼近性能.

上述方法的一个很好的解释可以从人学习的角度进行分析.人在学习时,对于简单容易的东西,只需要少数几次就可以掌握,而对于复杂难学的东西,则需要大量的学习.而每一次的学习,并不需要对所有的样本学习,只要根据前面学习的效果,对那些较难掌握的样本进行重点学习即可.本文的改进训练方法正是这一过程的反映.

5 仿真例子

本节给出两个例子.例 1 针对 SOM 的逼近性能分析,对第 3 节中的分析进行说明,并验证了改进训练方法的效果;例 2 通过对一个动态系统进行多模型建模,说明采用改进的训练方法对逼近性能的提高.

例 1 考虑一维高斯函数

$$y = f(x) = \exp(-x^2/10).$$

将区间 $[-10, 10]$ 划分为 20 个子区间,在每个子区间上用线性函数 $\hat{y} = a_j x + b_j$ 对原函数进行逼近.设在每个子区间上,以式(2)为性能指标,局部模型总能达到最优逼近.

表 1 给出了各种情况下子区间的端点以及逼近性能(连续逼近性能指标 J 和样本逼近均方差 MSE).注意到第 1 列中以 x_5, x_6, x_7 和 x_8 为端点的子区间长度分别为 0.603 3, 1.520 5 和 0.558 9, 不满足关系式(6),因此不能由 SOM 神经元节点得到.第 2 列为在 SOM 神经元节点有序分布的情况下,所能实现的一个最优分布.可以看出,其中已有几个区间范围很小,导致相应的神经元非常接近,说明这已是所能达到的极限情况.用样本对 SOM 进行训练,结果如第 3 列所示.可以看出,由于样本均匀分布,神经元也均匀分布,所得子区间也是均匀分布的,逼近误差比前两种划分方式大得多.采用改进的训练算法,结果如第 4 列所示.可以看出,经过训练后,虽

然所得逼近性能指标不是最优的,但相比于一般的训练结果,已改进了很多,逼近误差也接近最优值(第 2 列).

表 1 各种区间划分情况的端点和逼近性能

	最优子 区间划分	最优神经元 节点划分	实际训练 神经元划分	改进训练 神经元划分
x_1	- 7.171 2	- 7.398 7	- 8.935 8	- 8.403 0
x_2	- 5.975 2	- 6.232 1	- 7.916 5	- 6.981 3
x_3	- 5.116 7	- 5.426 6	- 6.938 4	- 5.813 7
x_4	- 4.393 3	- 4.741 3	- 5.945 9	- 4.843 6
x_5	- 3.721 1	- 4.119 7	- 4.940 2	- 3.998 1
x_6	- 3.027 8	- 3.531 1	- 3.952 5	- 3.178 6
x_7	- 1.507 3	- 2.803 1	- 2.990 6	- 2.238 1
x_8	- 0.948 4	- 1.515 4	- 2.028 6	- 1.383 5
x_9	- 0.461 6	- 0.825 27	- 1.028 7	- 0.758 49
x_{10}	- 0.001 0	- 0.277 37	- 0.014 91	- 0.171 03
x_{11}	0.459 0	0.325 07	0.990 87	0.444 91
x_{12}	0.944 7	0.840 41	1.996 1	1.101 2
x_{13}	1.505 2	1.564 0	2.981 1	1.875 7
x_{14}	3.028 6	2.798 7	3.959 4	2.818 2
x_{15}	3.722 6	3.640 3	4.925 3	3.725 3
x_{16}	4.394 5	4.366 1	5.880 2	4.555 7
x_{17}	5.117 7	5.059 4	6.841 4	5.432 4
x_{18}	5.977 3	5.895	7.817 2	6.382 3
x_{19}	7.173 9	7.140 0	8.861 1	7.905 6
J	2.827e-5	3.545e-5	1.622e-4	5.123e-5
MSE	0.001 2	0.001 4	0.002 8	0.001 6

例 2 考虑二维逼近的例子^[18],这是一个用于神经网络建模的基准性问题. 系统为

$$y(t+1) = f(y(t), u(t)) = \frac{y(t)}{1+y^2(t)} + u^3(t),$$

用多个局部模型

$\hat{y}(t+1) = a_j \hat{y}(t) + b_j u(t) + d_j, j = 1, 2, \dots, N,$
 对原系统进行建模. 样本数据可由均匀分布于 $[-2, 2]$ 区间上的随机控制输入 $u(t)$ 输入系统得到. 所构成的样本集为

$$X = \{[y(t), u(t)] / t = 1, 2, \dots, K\},$$

$$Y = \{y(t+1) / t = 1, 2, \dots, K\}.$$

取 10 000 个数据,即 $K = 10\ 000$. 用 20×9 的 SOM 神经网络和 $N = 180$ 个线性局部模型对原系统建模.

采用一般的自组织多模型方法和改进的自组织多模型方法所得到的结果如图 1 和图 2 所示.

图 1 给出了训练样本和神经网络节点的分布. 可以看出,虽然 $u(t)$ 均匀分布,但由于系统的非线性, $y(t)$ 并不是均匀分布的,样本集中在轴 $y(t) = 0$

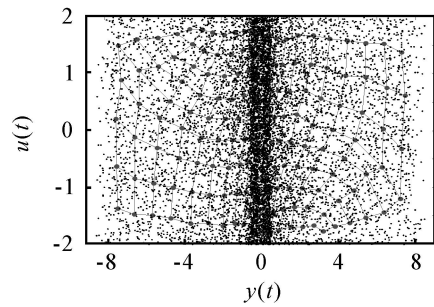


图 1 基本训练方法得到的神经网络分布图

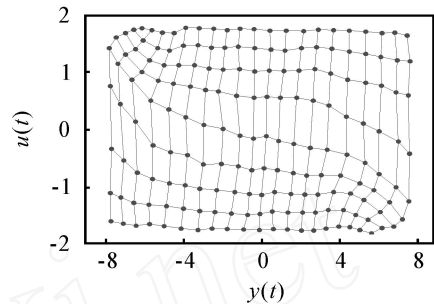


图 2 改进训练方法得到的神经网络分布图

附近,使得神经元分布也集中在此区域,从而导致总体的逼近误差较大. 从图 2 可以看出,改进的训练方法使神经元节点分布摆脱了样本分布的影响,使之更多地分布于边界和两个角落中,这些地方正是比较难以逼近的区域. 可见,新方法能有效地分配神经元节点的分布,改进了逼近效果.

在模型建立之后,采用输入

$$u(t) = \sin\left(\frac{2}{25}t\right) + \cos\left(\frac{1}{5}t\right), t = 1, 2, \dots, 100$$

对模型进行验证. 分别对原系统和两个模型进行仿真. 图 3 给出了两种模型的误差绝对值曲线,可以看出,采用改进方法建模,大大降低了误差范围. 图中那些较大的误差一般在系统运行到边界时发生,由于基本训练方法所导致的神经元分布,使得边界区域的逼近效果差,从而导致建模误差大. 而改进训练方法根据性能指标改变了神经元分布,减小了边界区域的逼近误差,降低了最大建模误差,提高了总体建模精度,使得模型输出误差的均方差从 0.103 7 降到了 0.062 2.

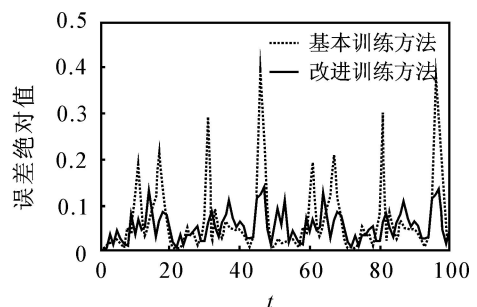


图 3 逼近误差曲线

6 结 语

SOM 神经网络是一种向量量化方法,它对输入数据的映射性质使其广泛地应用于非线性系统的建模与控制.但由于网络神经元节点的分布并不以提高逼近精度为目标,并且受网络本身结构的限制以及神经元节点划分空间规则的限制,使基于 SOM 神经网络的多模型逼近方法性能受到制约.本文对此进行研究,分析了逼近性能受到制约的原因,并根据人在学习时选择学习样本的方式,提出了一种改进的多模型逼近方法,有效地改进了神经元节点的分布,提高了逼近精度.通过仿真实例进一步验证了文中的分析,表明了所提方法的有效性.

参考文献(References)

- [1] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators [J]. *Neural Networks*, 1989, 2(3): 359-366.
- [2] Funahashi K. On the approximation realization of continuous mappings by neural networks [J]. *IEEE Trans on Neural Networks*, 1991, 2(1): 183-192.
- [3] Cybenko G. Approximations by superpositions of a sigmoidal function [J]. *Mathematics of Control, Signals, and System*, 1989, 2(3): 303-314.
- [4] Narendra K S, Parthasarathy K. Identification and control of dynamical systems using neural networks[J]. *IEEE Trans on Neural Networks*, 1990, 1(1): 4-27.
- [5] Hunt K J, Sbarbaro D, Zbikowski R, et al. Neural networks for control systems — A survey [J]. *Automatica*, 1992, 28(6): 1083-1112.
- [6] Narendra K S, Lewis F L. Introduction to the special issue on neural network feedback control [J]. *Automatica*, 2001, 37(8): 1147-1148.
- [7] Kohonen T. *Self-organizing maps*[M]. 2nd ed. Berlin: Springer-Verlag, 1997.
- [8] Principe J C, Wang L, Motter M A. Local dynamic modeling with self-organizing maps and applications to nonlinear system identification and control[J]. *Proc of the IEEE*, 1998, 86(11): 2240-2258.
- [9] Principe J C, Wang J. Nonlinear time series modeling with self-organizing features maps[C]. *IEEE Workshop on Neural Network for Signal Processing Cambridge*. 1995: 11-20.
- [10] 朱家元,虞建飞,张恒喜. 嵌入局部模型的 SOM 网络对混沌时间序列预测研究[J]. *控制与决策*, 2003, 18(1): 106-109.
- (Zhu Jia-yuan, Yu Jian-fei, Zhang Heng-xi. Prediction of chaotic time series based on self-organizing map embedded local models [J]. *Control and Decision*, 2003, 18(1): 106-109.)
- [11] Thomas M Matinetz, Ritter H J, Klaus J Schulten. Three-dimensional neural net for learning visumotor coordination of a robot arm[J]. *IEEE Trans on Neural Networks*, 1990, 1(1): 131-136.
- [12] Mottor M A, Principe J C. Predictive multiple model switching control with the self-organizing map[J]. *Int J of Robust and Nonlinear Control*, 2002, 17(12): 1029-1051.
- [13] Cho J, Principe J C, Erdogmus D, et al. Modeling and inverse controller design for an unmanned aerial vehicle based on the self-organizing map[J]. *IEEE Trans on Neural Networks*, 2006, 17(2): 445-460.
- [14] Martinetz T M, Berkovich S G, Schulten K J. Neural-gas network for vector quantization and its application to time-series prediction[J]. *IEEE Trans on Neural Networks*, 1993, 4(4): 558-569.
- [15] Fritzke B. Growing grid — A self-organizing network with constant neighborhood range and adaptation strength[J]. *Neural Processing Letters*, 1995, 2(5): 9-13.
- [16] Yin J S, Hu D W, Chen S, et al. DSOM: A novel self-organizing model based on no dynamic diffusing mechanism[J]. *Science in China Series F: Information Sciences*, 2005, 48(2): 247-262.
- [17] 吕强,俞金寿. 基于粒子群优化的自组织特征映射神经网络及应用[J]. *控制与决策*, 2005, 20(10): 1115-1119.
- (Lv Qiang, Yu Jin-shou. Self-organizing feature map neural network based on particle swarm optimizer and its application [J]. *Control and Decision*, 2005, 20(10): 1115-1119.)
- [18] Narendra K S, Parthasarathy K. Gradient methods for the optimization of dynamic system containing neural networks[J]. *IEEE Trans on Neural Networks*, 1991, 2(2): 252-262.