70-72 层次双曲函数型网络 整体逼近理论 学习算法 (23)

计算机科学2000Vol. 27№ 8

层次双曲函数型网络的整体逼近理论及学习算法

Universal Approximate Theory and Learning Algorithm of Hierarchical Hyperboloid Function Link Aritifical Neural Networks

周永权 O 174.41
（广西民族学院数学与计算机科学系　南宁530006）

Abstract In this paper, a kind or function link artifical neural networks is designed and a new hierarchical function link artifical neural networks HFLANN is presented, and its universal approximate property is disscussed, the universal learning algorithm is presented. HFLANN is very suitable for application domains with hierarchical knowledge structures, and therefore, it has very important practical significance.

Keywords HFLANN, Hierarchical network, Universal learning algorithm

1 引言

函数型网络 FLANN（Function Link Artifical NN）最早由鲍·约翰提出[1]，其实质是将输入经函数变换展成为一组基函数，而输出则为基函数的线性组合，然后，再根据实际需要选取基函数，基函数可用三角函数，多项式函数等。由于输出是基函数的线性组合，所以，不存在局部极小，从网络的结构来看，FLANN 相比其它的典型函数网络，如：MLP 网（Multilayer Perceptron Network），PPN 网络（Polynomial Perceptron Network）有许多优点：MLP 是多层网络，所用的 BP 算法复杂度高，训练时间长，随着节点及层数增加，权值也增加，而 FLANN 抛弃了传统的网络所必需的隐层，采用单层结构来实现，所需的权值大大减少，从而提高了收敛速度。其次，虽然 PPN 网络为单层，但随着输入的维数增加和多项式次数的增加，权值的数目也急剧增加，而 FLANN 是在输入端引入高阶项来增加输入维数，其训练速度和精确度都有数量级提高[3]。

由于 FLANN 网络有独特的结构和优点，因而被广泛应用于工程计算[2]、系统辨识[3]等许多领域，并且取得了很好的学习效果，但在自然界存在着这样一种现象，许多应用领域的知识都呈现层次结构特征，变量对于系统来讲也存在层次关系，即是某个变量要等其它某些变量经过一定的映射得到输出后，再和输出一起进入该系统，依次进行，如果使用通常单一的 FLANN 网络来逼近整个系统，势必会消除系统本身固有的层次特性，达不到人们要求的学习效果，甚至可能得出错误的结论。

因此，建立依据系统层次结构特性的层次函数型网络 HFLANN 是一件很有意义的工作，这样既保留了 FLANN 网络的结构及优点，又能在实际应用中很好地逼近于整个目标系统。本文工作是首先设计出一类双曲型函数网络，建立一类基于双曲型函数的层次函数网络 HHFLANN，讨论 HHFLANN 网络的整体逼近特性，最后，给出 HHFLANN 网络学习算法。

2 HHFLANN 网络系统模型

层次函数型网络的基本思想是用低维函数神经网络的组合代替高维的函数网络，从而实现具有层次结构系统的神经网络的整体逼近，首先，我们提出一典型的双曲函数 HFLANN 结构模型，不失一般性，仅考察二元输入，单输出的情形（图1）。

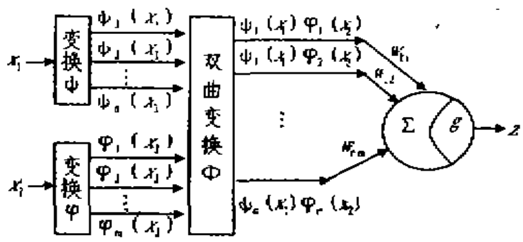

图1　一类双曲函数型网络 HFLANN 结构

若输入向量：$X=(x_1,x_2)$，那么该网络的输出：

$$z=g\left(\sum_{i=1}^{n}\sum_{j=1}^{m}w_{ij}\phi_i(x_1)\varphi_j(x_2)\right) \tag{1}$$

若激发函数 $g(\cdot)$ 取恒等函数，则有：

周永权　硕士，副教授，主要研究方向为神经网络，符号计算。

$$z=F(x_1,x_2)=\sum_{i=1}^{n}\sum_{j=1}^{m}w_{ij}\psi_i(x_1)\varphi_j(x_2) \qquad (2)$$

为了分析HHFLANN网络的逼近特性，在(2)中，以基函数$\psi_i(x_1)$、$\varphi_j(x_2)$均取三角型激发函数为例，即：

$$\psi_1(x_1)=\begin{cases}\frac{(\delta_2-x_1)}{(\delta_2-\delta_1)} & \delta_1\leqslant x_1\leqslant\delta_2\\ 0 & 其它\end{cases} \qquad (3)$$

$$\psi_i(x_1)=\begin{cases}\frac{(x_1-\delta_{i-1})}{(\delta_i-\delta_{i-1})} & \delta_{i-1}\leqslant x_1\leqslant\delta_i\\ \frac{(\delta_{i+1}-x_1)}{(\delta_{i+1}-\delta_i)} & \delta_i\leqslant x_1\leqslant\delta_{i+1}\\ 0 & 其它\end{cases} \qquad (4)$$

$$\psi_n(x_1)=\begin{cases}\frac{(x_1-\delta_{n-1})}{(\delta_n-\delta_{n-1})} & \delta_{n-1}\leqslant x_1\leqslant\delta_n\\ 0 & 其它\end{cases} \qquad (5)$$

同理$\varphi_j(x_2)(j=1,2,\cdots,m)$也定义具有(3)～(5)的基函数，这样所定义的$n$个激发函数$\psi_i(x_1)(i=1,2,\cdots,n)$的图形如下：

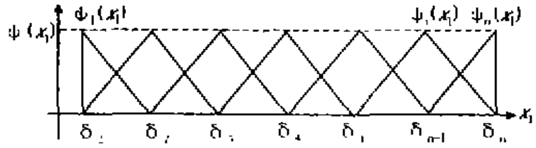

图2 x_1的n个三角型激发函数

在给定的基函数$\psi_i(x_1)$、$\varphi_j(x_2)$下，那么(2)式确定的函数是二元分片一次插值函数，以此函数型网络为基本非线性元件，来构造一类典型的层次函数神经网络(图3)。

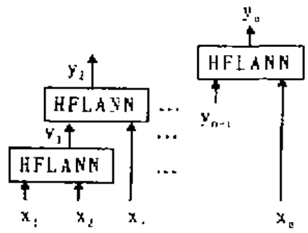

图3 一类典型层次函数神经网络(HHFLANN)

假设有n个输入变量$x_i(i=1,2,\cdots,n)$用$f_i(\vec{x}_i)$表示HHFLANN的第i个子HFLANN的输出，其中$\vec{x}_i$为第i个HFLANN的输入向量，其构造步骤为：

步1 置$i:=1$，假设第一层输入变量$x_1, x_2,\cdots,x_{n_1}$，那么第一个HFLANN的输出为：$y_1=f_1(\vec{x}_1)$，其中，$\vec{x}_1=(x_1,x_2,\cdots,x_{n_1})$。

步2 $i:=i+1$。第i个HFLANN输入变量有$n_i+1(n_i\geqslant1)$个，那么输出为：$f_i(x_{k_i+1},\cdots,x_{k_i+n_i},y_{i-1})$，其中：$y_{i-1}=f_{i-1}(\vec{x}_{i-1})$，$k_i=\sum_{j=1}^{i-1}n_j$。

步3 若$\sum_{j=1}^{i}n_j<n$，转步2，否则，系统结束。

3 HHFLANN整体逼近特性

在上面取定的三角型基函数下，单个HFLANN网络的输出为三角激发函数的线性组合，所以不存在局部极小，从函数复合的观点来看，HHFLANN网络可看作由HFLANN经若干次复合而成，因此，整个系统输出仍为新的基函数的线性组合，同样不存在局部极小。下面以二层HHFLANN为例，分析系统的逼近过程。

首先，设$x_i(i=1,2,3)\in[\alpha_i,\beta_i]$，第一个子HFLANN输入为$x_1$、$x_2$，输出为$y_1$，使得$y_1\in[\alpha_4,\beta_4]$，$x_1$经变换$\psi$后，得到输出$\psi_i(x_1)(i=1,2,\cdots,n)$，其节点数为$n$，我们沿$x_1$轴到区间$n$-1等份，在(3)～(5)中，置：

$\delta_1=\alpha_1,\delta_2=\alpha_1+\frac{\beta_1-\alpha_1}{n-1},\delta_j=\alpha_1+\frac{\beta_1-\alpha_1}{n-1}\cdot(j-1)$,

$\delta_{j+1}=\alpha_1+\frac{\beta_1-\alpha_1}{n-1}\cdot j,\delta_{j+1}=\alpha_1+\frac{\beta_1-\alpha_1}{n-1}\cdot(j+1),(j=2,3,\cdots,n-1)$

同理，我们沿x_2轴区间m-1等份，激发函数$\varphi_j(x_2)(j=2,3,\cdots,m-1)$使得满足(3)～(5)式，且$\delta_i$的取法同上，且将区间$[\alpha_4,\beta_4]$划分为n-1等份，并为$y_1$定义n个三角型激发函数$\gamma_j(y_1)(1\leqslant j\leqslant n)$，同样具有(3)～(5)式，图4是m=n=5的情形。

此时，第一个HFLANN的输出：

$$y_1=f_1(x_1,x_2)=\sum_{i=1}^{n}\sum_{j=1}^{m}w_{ij}\psi_i(x_1)\varphi_j(x_2) \qquad (6)$$

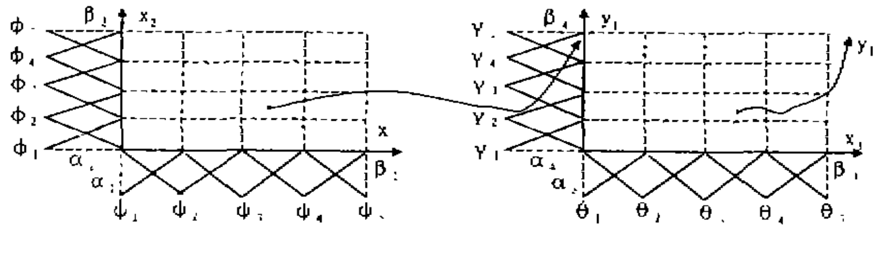

图4 输入变量数目为2的情形($n=m=5$)

第二个 HFLANN 的输出为:

$$y_2 = f_2(y_1, x_3) = \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} r_i(y_1) \theta_j(x_3) \tag{7}$$

依次进行,这样就可得到整个 HHFLANN 的输出

$$f(x_1, x_2, x_3) = f_2(f_1(x_1, x_2), x_3)) \tag{8}$$

若 $g(x_1, x_2, x_3)$是目标函数,给定一组训练样本通过学习,我们可确定权值 w_{ij}, w'_{ij},使得 $f(x_1, x_2, x_3)$逼近于目标函数 g。

4 HHFLANN 网络整体学习算法

先考察第一层 HFLANN 网络,随机地给定一组训练数据,$\{((x_1^{(k)}, x_2^{(k)}), y_1^{(k)}) | k=1,2,\cdots,p\}$ 代入(6)式,得到以 w_{ij}($i=1,2,\cdots,n$, $j=1,2,\cdots,m$)为未知变元的方程组:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \phi_i(x^{(k)}) \varphi_j(x_2^{(k)}) = y_1^{(k)} \tag{9}$$

令:$a_{kij} = \varphi_i(x_1^{(k)}) \varphi_j(x_2^{(k)})$,则(9)式转化为:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} a_{kij} = y_1^{(k)} \quad k=1,2,\cdots,p \tag{10}$$

这样,通过对权值 w_{ij}学习[2],得到满意的解 y_1。

我们继续考察第二层 HFLANN 网络,在第一层学习的基础上,添加新的学习样本:

$\{((y_1^{(k)}, x_3^{(k)}), y_2^{(k)}) | k=1,b,\cdots,p\}$

同理,我们可得到以 w_{ij}为未知变元方程组:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} w'_{ij} r_i(y_1^{(k)}) \theta_j(\chi_3^{(k)}) = y_2^{(k)} \tag{11}$$

同样的方法将(11)转化为:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} b_{kij} = y_2^{(k)} \quad k=1,2,\cdots,p \tag{12}$$

对权值 w_{ij}学习,得到满意的 y_2作为下一层输入,依次类推,最终完成对 HHFLANN 整体系统的学习。

实际上,我们进一步简化(10)式,若令:$u=(i-1)m+j$,则(10)式可写成:

$$\sum_{u=1}^{mn} a_{ku} w_u = y_1^{(k)} \quad k=1,2,\cdots,p \tag{13}$$

置 $v=mn$, $A=(a_{ku})$, $Y=(y_1^{(1)}, y_1^{(2)}, \cdots, y_1^{(p)})^T$, $W=(w_1, w_1, \cdots w_v)$,则(13)式可转化成:

$$AW=Y \tag{14}$$

通过求解矩阵方程(14),可方便地完成对 HHFLANN 网络的学习。

结语 文中首先设计出一类双曲函数型网络,提出一种层次函数网络模型,讨论了 HHFLANN 网络的整体逼近特性,给出了 HHFLANN 网络的整体学习算法。该算法不存在局部极小,更适合于具有层次知识结构的应用领域,因而,具有较高的理论价值和广泛的应用背景。

参 考 文 献

1 Pao Y H. Adaptive pattern recognition and neural networks. New-Wesley publish company Inc. 1989

2 Pao Y H, et al. Neural-net computing and intelligent control of system. Int. J. Control, 1992, 56(2): 263～289

3 Patra J C, et al. A functional link artifical neural networks for adaptive channel equalization. signal Processing, 1995 (43): 181～195

4 於东军,王士同,等. 层次径向神经网络的全局逼近理论. 计算机研究与发展, 1999, 36(11): 1329～1334

5 史忠植. 神经计算. 北京:电子工业出版社, 1993

(上接第74页)

上,基于线性规划的网络算法是 BP 算法的三分之一。这说明,原有的 BP 算法计算复杂,并且易于产生局部极小,使训练得到的网络不能正确反映预测系统的实际情况。而本文提出的网络算法有效地克服了这些问题。

参 考 文 献

1 Buckley J J, Hayashi Y. Numerical relationships between neural networks, continuous functiouns. Fuzzy Sets and Systems, 1993, 60: 1～8

2 Ishibuchi H, et al. Fuzzy neural networks with fuzzy weights and fuzzy biases In: Proc. IEEE Int. Conf. Neural Networks, San Francisco, 1993, 3: 1650～1655

3 Buckley J J, Hayashi Y. Fuzzy neural networks: A survey. Fuzzy Sets and Systems, 1994, 66: 1～13