第 32 卷　第 19 期　　计　算　机　工　程　　2006 年 10 月
Vol.32　　№ 19　　Computer Engineering　　October 2006

·人工智能及识别技术·　文章编号：1000—3428(2006)19—0196—03　文献标识码：A　中图分类号：TP183

新型 Sigma-Pi 泛函网络模型

周永权[1]，陈东用[2]，李陶深[2]

(1. 广西民族学院计算机与信息科学学院，南宁 530006；2. 广西大学计算机与电子信息学院，南宁 530003)

摘　要: 将泛函神经元结构变形，建立 Sigma-Pi 泛函网络模型，给出 Sigma-Pi 泛函网络学习算法。采用数值分析的方法，将 Sigma-Pi 泛函网络应用于异或问题，结果表明，该网络对于某些问题具有很强的分类能力。该方法的优点在于利用一元函数作为基函数来实现高维函数的逼近，在函数逼近技术上，有着重要的应用价值。

关键词: 泛函神经元；泛函网络；Sigma-Pi 泛函网络；基函数簇；异或问题

New Sigma-Pi Functional Networks Model

ZHOU Yongquan[1], CHENG Dongyong[2], LI Taosheng[2]

(1. College of Computer and Information Science, Guangxi University for Nationalities, Nanning 530006;
2. College of Computer and Electronic Information, Guangxi University, Nanning 530003)

【Abstract】This paper converts a structure of a functional neuron, presents a Sigma-Pi functional network(SPFN) structure, and proposes the Sigma-Pi functional networks learning algorithm. Using numerical analysis method, the Sigma-Pi functional network is applied to XOR problem. The results demonstrate that the functional network has powerful classification capability. The method has the advantages of using a variable function for multi-dimensional function approximation and important practical significance.

【Key words】Functional neuron; Functional networks; Sigma-Pi functional networks; Base functions; XOR problem

1 概述

人工神经网络(ANN)的研究自从 20 世纪 80 年代形成热潮，至今已有 20 多年了，其应用范围已扩展到许多领域，ANN 对信息的存储及处理表现在各神经元的连接强度上，ANN 学习的目的就是修正网络的权值，而网络的其它特性(如网络的拓扑结构、激活函数)在网络学习之前已确定，由于这些信息在学习阶段无法修改，在一定程度上限制了 ANN 的应用范围。

吴佑寿等人提出一种可调激活函数神经元模型[1]，在解决某些问题时，根据问题的先验知识，选用激活函数使它和待解问题相适应，但先验知识不易得到，可能要经过多次的试验，并且不适应所有问题。针对这些问题，文献[2]指出：虽然许多研究者认为ANN近年来在应用上取得很大进展，但在理论与新模型方面却进展不大。如何根据特定问题选择合适的激活函数，并对网络的结构根据特定问题进行修剪，设计出新的网络模型，这正是泛函网络的出发点所在。1998 年 Castillo E提出了泛函网络[3,4]。

泛函网络处理的是一般的泛函模型，泛函网络的各神经元之间没有连接权值，神经元函数不是不变的，而是可学习的，在学习的过程中可根据问题的需要选取基函数簇(如多项式、三角函数等)，这样使得神经元函数选择更加灵活，能够根据逼近函数的先验知识对神经元函数加以选择。泛函网络已经被成功地应用于非线性问题的辨识和预测、微分、差分方程求解、非线性回归等问题[4]，因此泛函网络是一个值得研究的方向。本文结合泛函网络的特点，提出一种Sigma-Pi泛函网络(Sigma-Pi Functional Network, SPFN)结构。SPFN的输出具有的形式为

$$\prod_{i=1}^{n}\sum_{j=1}^{m} f_{ij}(x_j)$$

其中，$f_{ij}(x_j)$ 是任意给定基函数簇或基函数簇的线性组合。因此把它称为 Sigma-Pi 泛函网络。SPFN 具有网络结构简单、基函数可学习、泛化能力好、学习算法简单等特点。

2 泛函网络简介

图 1 给出了一个典型的泛函网络拓扑结构模型，其主要组成部分如下：

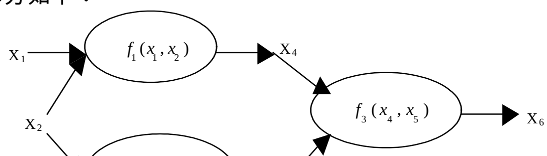

图 1　一个泛函网络结构模型

(1)一个输入单元层。其功能是输入信息，在图 1 中输入层包含 $\{x_1, x_2, x_3\}$。

(2)一层或多层泛函神经单元。其中每个泛函神经元是一个计算单元，对一个或多个输入的信息进行处理产生输出。在图 1 中，有两层泛函计算元 $\{f_1, f_2\}$ 和 $\{f_3\}$。

基金项目：国家自然科学基金资助项目(60461001)；广西自然科学基金资助项目(0542048)

作者简介：周永权(1962－)，男，教授、博士，主研方向：神经网络，计算智能；陈东用，硕士生；李陶深，教授、博士

收稿日期：2005-12-01　**E-mail**：Zhou_yongquan@163.com

(3)若干个中间存贮单元层，它存储由泛函神经元产生的信息。在图 1 中，只有一个中间存储单元层 $\{x_4, x_5\}$。

(4)一个输出层。包含输出信息，图 1 中 $\{x_6\}$。

(5)一些直接连接。它们连接输入层、中间层神经元和输出层。图 1 中的箭头方向表示信息流动的方向。

3 Sigma-Pi 泛函网络结构

3.1 泛函神经元结构及其变形

设泛函神经元 $f_i(\bullet)$ 是一个计算单元，对任意一输入向量 x，Castillo E 提出的泛函神经元结构如图 2 所示。

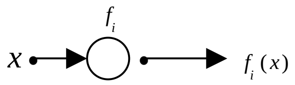

图 2 泛函神经元模型

由于泛函神经元 $f_i(\bullet)$ 都可表示成一个已知基函数的线性组合的形式，根据这一特性，可将泛函神经元模型结构进行适当的变形，变形后的结构见图 3。

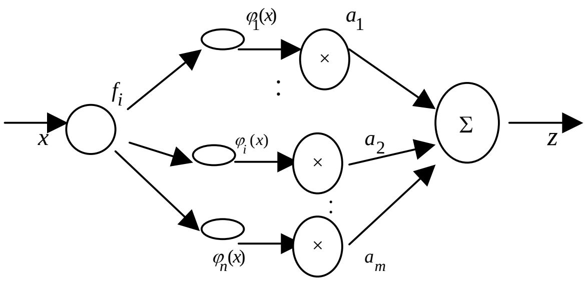

图 3 变形后的泛函神经元模型

在图 3 中，基函数 $\varphi_i(x)$ 是关于输入信号矢量 x 的基函数的线性组合，变形后的泛函模型结构与变形前模型结构的本质上是一致的，变形后的泛函神经元的输出为

$$z = f_i(x) = \sum_{i=1}^{m} a_i \varphi_i(x) \tag{1}$$

其中，a_i 是泛函参数，基函数 $\varphi_i(x)$ 是可变且可学习的。变形后的泛函神经元模型为 Sigma 泛函网络。

3.2 Sigma-Pi 泛函网络结构

SPFN 构造的方法是以多输入、单输出泛函神经元为基本处理单元，通过给定不同的基函数集，形成 k 个不同的 Σ 块，再经过∏运算而形成 Sigma-Pi 泛函网络(图 4)。

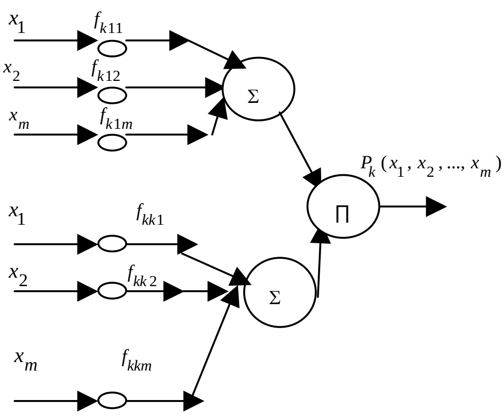

图 4 k 块 Sigma-Pi 泛函网络结构

PSFN 泛函网络是由 k 块 Sigma-Pi 泛函网络构成，该网络的输出为

$$P_k(x_1, x_2, ..., x_m) = \prod_{i=1}^{k} \sum_{j=1}^{m} f_{kij}(x_j) \tag{2}$$

4 Sigma-Pi 泛函网络学习算法

在图 4 所示网络的结构中，其中每一个泛函神经元可以是图 2 所示的模型，也可以是变形后图 3 所示的模型。以下分两种情形：

(1)若泛函神经元是图 3 所示的模型，事先给定一组基函数为

$$\{\varphi_{lij}(x_j) \mid i = 1, 2, ..., k,\ j = 1, ..., m, l = 1, 2, ..., K\} \tag{3}$$

于是，在图 3 所示的模型中，每一个泛函神经元为

$$f_{lij}(x_j) = \sum_{l=1}^{K} a_{lij} \varphi_{lij}(x_j) \tag{4}$$

则式(2)可写成

$$P_k(x_1, x_2, ..., x_m) = \prod_{i=1}^{k} \sum_{j=1}^{m} \sum_{l=1}^{K} a_{lij} f_{lij}(x_j) \tag{5}$$

(2)若图 3 所示的模型中每一个泛函神经元本身就是给定的基函数为

$$\{\phi_{ij}(x_j) \mid i = 1, 2, ..., k,\ j = 1, ..., m\} \tag{6}$$

图 3 所示的泛函网络的输出为

$$P_k(x_1, x_2, ..., x_m) = \prod_{i=1}^{k} \sum_{j=1}^{m} a_{ij} \phi_{ij}(x_j) \tag{7}$$

不论哪种情形，系数代表是泛函网络的参数。

在第 2 种情形下，给出 Sigma-Pi 泛函网络学习算法，考察误差代价函数

$$e_p = y_p - P_k(x_1, x_2, ..., x_m) = y_p - \prod_{i=1}^{k} \sum_{j=1}^{m} a_{ij} \phi_{ij}(x_{jp}) \tag{8}$$

其中，P 为当前训练模式数；y_p 是第 P 个训练模式的期望输出。为了找到最优的泛函参数，最小化误差平方和为

$$E_0 = \frac{1}{2} \sum_{p=1}^{P} e_p = \frac{1}{2} \sum_{p=1}^{P} [y_p - \prod_{i=1}^{k} \sum_{j=1}^{m} a_{ij} \phi_{ij}(x_{jp})]^2 \tag{9}$$

为了保证泛函网络表达式的唯一性，需要给出网络的一些初始值。在这里设网络的初始值为

$$f_{ij}(x_{j0}) = u_{j0}, \quad j = 1, 2, ..., m \tag{10}$$

其中，u_{j0} 是给定的一些常数。

在很多情况下并不需要 m 个所有的初始值，但为了找到一般的学习算法，还是假设给出所有初始条件，式(9)经变换后，得到如下目标函数

$$\begin{aligned} E &= E_0 + \frac{\lambda}{2} \sum_{j=1}^{m} (f_{ij}(x_{j0}) - u_{j0})^2 \\ &= \frac{1}{2} \sum_{p=1}^{P} [y_p - \prod_{i=1}^{k} \sum_{j=1}^{m} a_{ij} \phi_{ij}(x_j)]^2 + \frac{\lambda}{2} \sum_{j=1}^{m} (f_{ij}(x_{j0}) - u_{j0})^2 \end{aligned} \tag{11}$$

分别对泛函参数 a_{ij} 和参数 λ 求偏导数。

$$\frac{\partial E}{\partial a_{tr}} = -\sum_{p=1}^{P} e_p [\prod_{t \neq i=1}^{k} \sum_{r \neq j=1}^{m} a_{ij}\phi_{ij}(x_j)]\phi_{tr}(x_r) \tag{12}$$

$$+ \lambda(f_{ij}(x_{j0}) - u_{j0})\phi_{tr}(x_{j0})$$

其中，$t = 1, 2, ..., k; r = 1, 2, ..., m$，可通过下式来自适应地调节泛函参数。

$$a_{tr}(n+1) = a_{tr}(n) + \mu(-\frac{\partial E}{\partial a_{tr}}) \tag{13}$$

式中，μ 是学习率。

5 算例

从上面的学习可看出，学习算法采用梯度下降法求解，与神经网络学习相同，只是采用的网络模型不同而已，本文不再赘述。本节从数值分析的方法入手，分析 Sigma-Pi 泛函网络如何应用于异或分类问题

例 图 5 是一简单的 Sigma-Pi 泛函网络，其中 f_1、f_2 为泛函神经元。

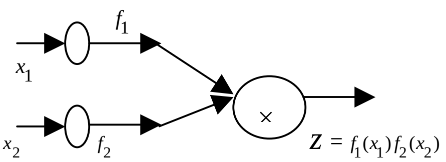

图 5 用于解决异或问题的 Sigma-Pi 泛函网络

该网络的输出为

$$z = f_1(x_1)f_2(x_2) \tag{14}$$

在式(14)中，若选取基函数分别是 $\{1, x_1\}, \{1, x_2\}$，那么，对每一 f_1、f_2 可表示成

$$f_1(x_1) = a_1 + a_2 x_1 \quad f_2(x_2) = b_1 + b_2 x_2 \tag{15}$$

将式(15)代入式(14)式有

$$z = f_1(x_1)f_2(x_2) = c_1 + c_2 x_1 + c_3 x_2 + c_4 x_1 x_2 \tag{16}$$

它可用于模拟异或函数。事实上，取泛函参数 c_i 的值分别为 $c_1 = -1, c_2 = c_3 = 1, c_4 = -2$。

式(16)可转化为

$$z = x_1 + x_2 - 2x_1x_2 - 1 \tag{17}$$

45 ° 坐标旋转后，新坐标为 $x_1^{'}$、$x_2^{'}$，则二次曲线 $x_1 + x_2 - 2x_1x_2 = 1$ 的标准形式为

$$-\frac{(x_1' - \sqrt{2}/2)^2}{(\sqrt{2}/2)^2} + \frac{(x_2^{'})^2}{(\frac{\sqrt{2}}{2})^2} = 1$$

由于泛函网络的泛函神经元函数它可学习的、是可变的，因此在图 5 的模型中，若取基函数集 $\{1, x_1, x_1^2\}$、$\{1, x_2, x_2^2\}$，那么 f_1、f_2 可表示为

$$f_2(x_2) = b_1 + b_2 x_2 + b_3 x_2^2 \tag{18}$$

将式(18)代入式(14)有

$$z = c_1 + c_2x_1 + c_3x_2 + c_4x_1x_2 + c_5x_1^2 + c_6x_2^2 + c_7x_1x_2^2 + c_8x_2x_1^2 \tag{19}$$

它可将 $\{(0,0),(1,1)\}$ $\{(1,0),(0,1)\}$ 2 组输入分成两类，如图 6 所示。

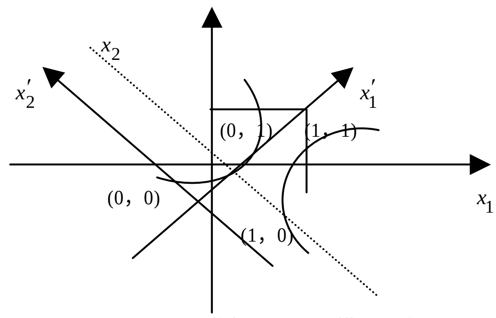

图 6 双曲函数把异或问题模式分类

同样的方法，适当地选取泛函参数 c_i，$i = 1, 2, ..., 8$。它也能模拟异或函数(见图 7)。

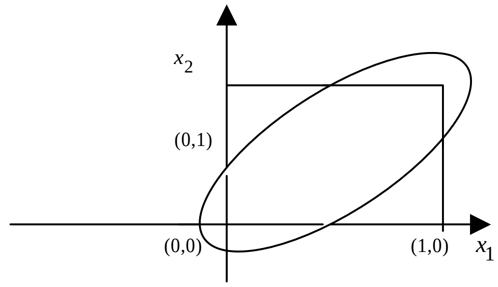

图 7 高阶函数把异或问题模式分类

在解决实际问题，通过对泛函神经元函数的学习，通过选取不同类型的基函数来实现，这给问题的处理带来很大方便。

从数值分析的角度来看，Sigma-Pi 泛函网络处理异或问题方法与神经网络比较，可得出以下结论：若网络收敛时，则泛函参数 $c_2 = c_3$，这一点和神经网络在异或问题中得到的权值 $w_1 = w_2$ 是一致的，在异或问题中，2 个输入 x_1 和 x_2 作用是相同的，因而，相应的权值相等是合理的、必然的，这一点可从图 6、图 7 可直观地看到。

6 结论

文中提出一种 Sigma-Pi 泛函网络新模型，给出了 Sigma-Pi 泛函网络学习算法，从数值分析的方法，分析了 Sigma-Pi 泛函网络在异或问题中应用，Sigma-Pi 泛函网络是利用一元函数作为基函数、或一元基函数的线性组合来实现多维函数的逼近。

本方法解决了目前函数逼近中存在的 2 个基本问题：(1)如何选择基函数；(2)如何基于基函数构造近似函数，使其具有与目标函数尽可能小的偏差。泛函网络是最近提出的一种新的神经网络模型，其有些理论和应用基础有待进一步完善，因此泛函网络是计算智能一个值得研究的新方向。

参考文献

1 吴佑寿，赵明生，丁晓青. 一种激励函数可调的新人工神经网络及应用[J]. 中国科学, 1997, 27(1): 55-60.

2 阎平凡. 对多层前向网络研究的几点看法[J]. 自动化学报, 1997, 23(1): 129-135.

3 Castillo E. ISSN1370-4621, Neural Processing Letters[S]. 1998.

4 Castillo E, Cobo A, Gutierrez J M. Functional Networks with Applications[M]. Kluwer Academic Publishers, 1999.

5 王守觉，王柏南. 人工神经网络的多维空间几何分析及其理论[J]. 电子学报, 2002, 30(1): 1-4.