第 18 卷第 2 期
2002 年 5 月

广西科学院学报
Journal of Guangxi Academy of Sciences

Vol. 18, No. 2
May 2002

代数超曲面神经元的空间几何分析及其代数表示理论*

Algebra Expression Method and Space Geometry Analysis of Algebra Hyper-curve Neurons

周永权
Zhou Yongquan

戴祯杰
Dai Zhenjie

(广西民族学院数学与计算机科学系 南宁 530006)
(Dept. of Math. and Comp. Sci., Guangxi University for Nationalities, Nanning, 530006)

(广西教育学院数学系 南宁 530023)
(Dept. of Math., Guangxi Pedagogical College, Nanning, 530023)

摘要 在研究 M-P 神经元模型的几何意义基础上，从代数簇的观点出发，分析 M-P 神经元模型的代数本质，提出 1 种代数超曲面神经元模型；从多维代数和空间几何分析的观点出发，刻画和描述出代数超曲面神经元模型的数学实质，给人们研究高阶神经网络系统的空间几何理论及其多维代数表示理论奠定了基础．

关键词 神经元 代数超曲面神经元 空间几何 代数表示

中图法分类号 TP181;TP183 A

Abstract On the base of the geometrical presentation of M-P neurons model combining algebra cluster knowledge, a new algebra hyper-curve neurons model is released to depict the algebra essence of M-P neurons model. The algebra hyper-curve neurons model would take more contribution to the development of the neural networks and the research of nonlinear neural networks and expression of high-dimensional algebra theoretically.

Key words neurons, algebra hyper-curve neurons, geometrical space, algebra expression

人工神经网络由于它具有极强的函数逼近能力和学习能力、容错能力以及高超的分类能力等优点而广泛用于解决各类实际问题，但由于每个神经元的数学模型都是一个多变量非线性函数方程，组成的神经网络以后的庞大多变量非线性方程组更难以进行求解．因此，它限制着人工神经网络从应用角度的深入发展．

本文对神经网络的行为的认识和研究中的“黑匣子”式难以理解的状态，提出一种用多维代数簇分支理论认识分析的方法，在作者已发表的代数神经网络新模型及学习算法[1,2]的结果

2002-02-02 收稿，2002-03-18 修回．

* 广西自然科学基金资助项目[桂科基 0141034]．

中，已经反映了这种分析方法的优越性. 现将进一步讨论代数超曲面神经元的空间几何对应分析及代数表示理论.

1 *M-P* 神经元模型的几何意义

1943 年，Meculloch 和 Pitts 第一次提出了神经元的数学模型，此模型一直沿用至今，归纳起来，*M-P* 神经元是一个有几个输入和一个输出的运算系统，最常用神经元的运算函数为：

$$Z = f(\sum_{i=1}^{n} w_i x_i - \theta), \tag{1}$$

其中 x_i，w_i 分别是第 i 个神经元状态与神经元联接权系数，θ 为阀值，f 为激励函数，一般取符号函数，令

$$g(X) = \sum_{i=1}^{n} w_i x_i - \theta. \tag{2}$$

那么，可以看出，神经元的运算函数是由 2 个函数复合而成，其中一个函数：$g(X) = \sum_{i=1}^{n} w_i x_i - \theta = 0$. 这个方程在 n 维空间表示一个超平面 P，当 $g(X) > 0$ 时，表示点 X 落在超平面 P 的正半空间内，此时有 $f(g(X)) = 1$；当 $g(X) < 0$ 时，表示点 X 落在超平面 P 的负半空间，此时有 $f(g(X)) = -1$. 那么，一个 *M-P* 神经元可以看成一个由超平面划分空间位置的识别器[3].

显然，超平面的几何概念对于帮助人们对神经网络的行为的认识与分析是十分有效的，但又有局限性[3].

2 代数超曲面神经元模型

虽然从几何的观点来理解 *M-P* 神经元模型相当直观，但是当 n，m 较大，n 维空间的 m 个超平面位置关系就很复杂，很难从几何意义出发直观理解神经网络的内在本质，而失去直观性. 若从多维代数的分支理论来分析，可使得 *M-P* 神经元的几何背景更加直观. 首先我们置：

$$V_0 = \{X \in R^n \mid g(X) = 0\}. \tag{3}$$

若从代数簇的观点来理解，则 V_0 表示 1 个代数超曲面. 特别当 $n = 2$ 时，R^2 中代数超曲面表示是 1 条平面代数曲线. 进一步，若记它输入状态向量与权系数向量分别为：$X = (x_1, x_2, \cdots, x_n)$，$W = (w_1, w_2, \cdots, w_n)$，那么(1) 式可写成

$$z = f(u(W, X) - \theta), \tag{4}$$

其中 $u(W, X)$ 为 W 的线性函数，且是 X 的多项式函数，那么(4) 式进一步表示为：

$$z = f\Big(\sum_{s=1}^{m} w_s u_s(X) - \theta\Big), \tag{5}$$

其中 $u_s(X)$ 为关于多元多项式函数的一个单项，记 $u_s = u_s(X) = \prod_{i=1}^{n} (x_i)^{j_i(s)}$，其中 $j_i(s)$ 为 $u_s(X)$ 中 x_i 变量的幂，记

$$j^*(s) = (j_1(s), j_2(s), \cdots, j_n(s)) \tag{6}$$

为项 $u_s(X)$ 中的幂向量，称向量簇：

$$J = \{j^*(s) \mid s = 1, 2, \cdots, m\} \tag{7}$$

为(5) 式的结构向量簇. 又记

$$j(s) = j_1(s) + \cdots + j_n(s) \tag{8}$$

为项 $u_s(X)$ 的阶数，而

$$j = \max\{j(s) \mid s = 1,2,\cdots,m\} \tag{9}$$

为高阶感知器(5) 式的阶数．

定义 1　我们称(5) 式的 M-P 神经元模型是一个具有 m 项，n 个神经元的代数超曲面神经元，其代数超曲面神经元结构向量族及阶数由(6) ～ (9) 式确定．

例 1　$z = f(w_1x_1 + w_2x_2 + w_3x_1x_2 + w_4x_2x_3 + w_5(x_1)^2 - \theta)$ 是一个具有 3 元，5 项，2 阶代数超曲面神经元．在 M-P 神经元模型的基础上，引进神经元模型的项及阶概念，使得 M-P 神经元变得更精细，在应用上更方便．我们就把这样的神经元称代数超曲面神经元．

3　代数超曲面神经元的空间几何对应及代数表示理论

由上面关于代数超曲面神经元的定义可以看出，代数超曲面神经元结构仍是多输入单输出的运算系统，对 $x = (x_1, x_2, \cdots, x_n) \in R^n$，每个输出状态由下式确定：

$$z = f(\sum_{i=1}^{n} w_i x_i + \sum_{1\leqslant i\leqslant j\leqslant n} w_{i,j} x_i y_j + \cdots + w_{1,2,\cdots,n} x_1 x_2 \cdots x_n - \theta), \tag{10}$$

其中 $(w_1, \cdots, w_2, \cdots, w_{1,2,\cdots,n})$ 为待定权系数，θ 为阈值，f 为激励函数．

为了方便，考虑对于二维空间 R^2 空间的 2 阶代数超曲面神经元时的情形，上述模型：

$$z = f(w_1x_1 + w_2x_2 + w_{1,2}x_1x_2 + w_3x_1^2 + w_4x_2^2 - \theta), \tag{11}$$

表示的是一个二次平面代数曲线．

例 2　R^2 代数曲线神经元：$z = f(x_1w_1 + x_2w_2 + x_1x_2w_3 - \theta)$ 来实现异或函数．

实际上，取权系数 $w_1 = w_2 = 1, w_3 = -2, \theta = 1$ 则上式可转化为

$$z = f(x_1 + x_2 - 2x_1x_2 - 1).$$

这时，$g(X) = x_1 + x_2 - 2x_1x_2 - 1 = 0$，做 45° 坐标旋转．新坐标为 x_1', x_2'，则二次平面曲线 $g(X) = 0$ 变成

$$g(x') = -\frac{(x'_1 - \sqrt{2}/2)^2}{(\sqrt{2}/2)^2} + \frac{(x_2')^2}{(\sqrt{2}/2)^2} - 1 = 0.$$

它表示一双曲线，它可将 2 组输入：$\{(0,0),(1,1)\}$，$\{(1,0),(0,1)\}$ 分成二类，见图 1.

现考虑一般平面代数曲线，对于代数超平面的神经元阶数有明确几何意义，它表示对任意一条一般直线(即不是 V_0 的一个分支的直线) 与代数簇 V_0 标交的交点个数[4](重复交点，重复计数)．这样一来，平面代数曲线对 R^2 划分具有明确几何意义．

从例 2 可以看出，代数超曲面神经元对异或问题具有划分能力．

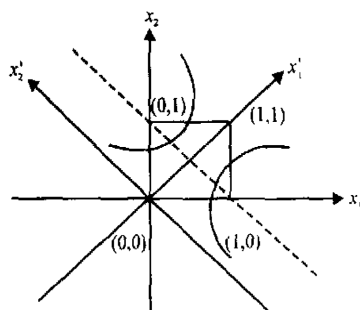

图 1　平面代数曲线实现异或分类

接着考虑 R^n 空间代数超曲面，对给定 R^n 中代数超曲面 V_0，我们可利用文献[1] 中的学习算法作近似因式分解，容易求出代数簇 V_0 的每一个不可约的分支，若不计非常数因子，则这种分解是唯一确定的，实质上，每一个不可约分支对应着空间一个子代数簇，而每一个子代数簇对应着一个子空间．

特别地，如果 $g(X)$ 本身是不可约的，则 V_0 为不可约超曲面．此时，代数超曲面对空间 R^n

划分具有明确的几何意义，则该超曲面神经元本身就构成了一个划分空间；若 $g(X)$ 可约，该超曲面的每一个不可约分支就构成一个划分子空间．即对一个数据聚类可以由一个超曲面进行分割，无论哪一种情形，我们完全可以用代数超曲面来完成对样本空间划分，调整权值 w 和值 θ，使得落入代数超曲面内的样本均属于同一种类型，这种具有明确的几何意义上分割对于我们以代数超曲面神经元为基本元件构建新的网络结构具有重要指导意义．

引理[1] 若超曲面函数：$g(x_1,x_2,\cdots,x_n)$ 能分解成：

$g(x_1,x_2,\cdots,x_n)=g_1(x_1,x_2,\cdots,x_n)g_2(x_1,x_2,\cdots,x_n)\cdots g_r(x_1,x_2,\cdots,x_n)$，则该曲面的每一个分支 $g_i(x_1,x_2,\cdots,x_n)(i=1,2,\cdots,r)$ 构成一个划分子空间．

4 结论

本文在深入研究 M-P 神经元模型的几何意义基础上，从空间几何分析和多维代数簇观点出发，探讨 M-P 神经元模型的数学本质，提出一种代数超曲面神经元新模型，该模型具有以下特点：① 一个代数超曲面神经元本身就构成一个划分子空间，即一个数据聚类可以由一个代数超曲面或若干个子曲面进行分割；② 从空间几何分析和多维代数分支观点出发，进一步理解 M-P 神经元模型的代数几何本质；③ 若以代数超曲面神经元为基本元件，构造出的网络结构清晰，系统物理意义明确；④ 代数超曲面神经元模型相比 M-P 神经元模型更精细，应用范围更大．

总之，关于非线性神经网络研究还是一个探讨性问题，它的理论基础需要进一步完善．本文提出的代数超曲面神经元新模型，给今后研究高阶非线性神经网络系统理论提供了几何背景和多维代数表示的方法．

参考文献

1 周永权．基于代数神经网络的多元多项式近似因式分解模型及学习算法．计算机研究与发展，1999，36(6)：668～674．

2 周永权．前向代数神经网络整体逼近理论及学习算法．计算机研究与发展，2000，37(3)：264～271．

3 张 铃，张 钹．M-P 神经元模型的几何意义及其应用．软件学报，1998，(5)：334～338．

4 P 格列菲斯著．代数曲线．北京：北京大学出版社，1985．

5 张 铃，张 钹．前向神经网络设计问题的回顾与探索．计算机工程与科学，1998，20(4)：1～10．

6 沈世镒著．神经网络系统理论及其应用．北京：北京科学技术出版社，1998．

（责任编辑：邓大玉）