

线性可分结构系对二进神经元的覆盖问题

杨 娟 陆 阳 黄镇谨

(合肥工业大学计算机与信息学院 合肥 230009)

摘 要 二进神经网络中每个二进神经元等价于一个线性可分函数,但每个二进神经元所表达的线性可分函数的逻辑意义仍不完全清楚。对此,首先分析了已有的几种线性可分结构系;其次,讨论了其是否覆盖了所有的二进神经元;最后,指出阈值在某些范围内二进神经元所对应的线性可分函数的逻辑意义仍不清楚,这为进一步完善二进神经元的覆盖问题指明了方向。

关键词 二进神经网络,线性可分结构系,覆盖问题,逻辑意义

中图法分类号 TP18 文献标识码 A

Cover Problem of Linear Separable Structures to Binary Neurons

YANG Juan LU Yang HUANG Zhen-jin

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

Abstract In binary neural networks, every neuron is equivalent to a linear separable function, however, the logical meaning of every linear separable function which is expressed by the binary neuron is still not clear. This paper firstly analysed the known linear separable structures, then discussed whether these known structures cover the whole binary neurons, finally pointed out when the threshold value is in some range, the logical meaning of the linear separable function is still unclear. This result provides the way of the cover problem of binary neurons.

Keywords Binary neural networks, Linear separable structures, Cover problem, Logical meaning

1 概述

二进神经网络是使用硬限幅函数作为激发函数的三层前向神经网络。只要网络中隐层具有足够的神经元,二进神经网络就可以实现任意给定的布尔函数^[1-4],而实现任意给定的一个布尔函数所使用的二进神经网络结构的最小规模仍不能确定^[5-7]。问题的关键在于布尔函数随着维数的增加,其个数以级数方式增加,如9维的布尔函数,共有 $2^{2^9} \approx 2^{512} \approx 1.34078 \times 10^{154}$ 个,如此巨大的数量比宇宙的年龄都要大^[8]。

布尔函数从本质上可分为线性可分函数与非线性可分函数。线性可分函数仅占整个布尔函数的一小部分,对于较小的维数 $n \leq 9$ 时,文献[9]通过枚举法给出了线性可分函数的个数。该类函数可用一个二进神经元来表达,每个二进神经元等价于一个线性可分函数,但二进神经元的表达式以分类超平面来描述,使得函数的信息融合在超平面的权值及阈值中,从机器学习的角度考虑,其并不能反映出函数的本质,缺乏透明性,在应用上受到限制,故需在二进神经元和线性可分函数之间建立联系的充要条件。研究表明,某些线性可分函数具有相似的逻辑结构和知识内涵,组成一类线性可分结构系。另外,非线性可分函数占了布尔函数的绝大部分^[10],通常实现该类函数是将其分解为一些线性可分函数之间的逻辑运算^[11,12]。一般来讲,二进神经网络中隐层神经元的个数

和所分解的线性可分函数的个数是相同的,故寻找最短的非线性可分函数的分解方法是优化网络结构的关键问题^[13-15]。如果二进神经网络中的每个神经元都对应于某种线性可分结构,则可以将任一非线性可分函数有目标地划分为某几种线性可分结构之间的逻辑运算。故研究线性可分结构系对二进神经元的覆盖问题及网络结构的优化都有着重要意义。

本文旨在对目前已经存在的各种线性可分结构系进行全面的分析,并讨论已知的线性可分结构系与二进神经元的覆盖问题之间的关系,以期为进一步提出线性可分结构系及解决二进神经元的覆盖问题提供借鉴。

2 已知的线性可分结构系

B_2 表示 $\{0,1\}$ 二元域, F 是 $B_2^n \rightarrow B_2$ 的任意映射,称 F 为 n 元布尔函数, B_2^n 构成了 n 维布尔函数样本空间。对于所有 $X=(x_1, x_2, \dots, x_n) \in B_2^n$,以 $F(x_1, x_2, \dots, x_n)$ 表示 F 在 X 处的取值,简记为 $F(X)$,称 X 为 B_2^n 中的样本, $F(X)$ 为样本值, $F(X) \in B_2$ 。

根据布尔空间线性可分的概念,设 $F(x_1, x_2, \dots, x_n)$ 是定义于 $\{0,1\}$ 的线性可分布尔函数, $\sum_{i=1}^n w_i x_i - \theta = 0$ 是 F 的分类超平面,则 $F(x_1, x_2, \dots, x_n) = 1$ 时, $\sum_{i=1}^n w_i x_i - \theta \geq 0$; $F(x_1, x_2, \dots, x_n) = 0$ 时, $\sum_{i=1}^n w_i x_i - \theta < 0$ 。对于如下结构的二进神经元

到稿日期:2011-09-13 返修日期:2011-11-23 本文受国家自然科学基金(60873195,61070220),高等学校博士点基金(20090111110002)资助。

杨娟(1983-),女,博士生,主要研究方向为人工智能、神经网络,E-mail: yangjuan6985@163.com;陆阳(1967-),男,教授,博士生导师,主要研究方向为人工智能、计算机控制;黄镇谨(1975-),男,博士生,讲师,主要研究方向为人工智能、计算机控制。

$$O=U(\sum_{i=1}^n w_i x_i - \theta), U(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

则称二进神经元 O 表达线性可分函数 $F(x_1, x_2, \dots, x_n)$, 记 O 为 $O(X, W, \theta)$, 其中 $X = (x_1, x_2, \dots, x_n)$, $W = (w_1, w_2, \dots, w_n)$ 。

首先介绍几类超立方体、汉明球、汉明球突、SP 函数等线性可分结构系。研究表明该类结构都是线性可分结构, 并且这些线性可分结构系都具有明确的逻辑意义。

定义 1 在二进神经网络中, 设 x_i 的权值为 w_i , 定义

$$x_i^* = \begin{cases} x_i, & w_i > 0 \\ -x_i, & w_i < 0 \end{cases}$$

定义 2 集合 $G = \{e_1, e_2, \dots, e_k\}$

$$\textcircled{1} C(G) = \begin{cases} k, & G \neq \phi \\ 0, & G = \phi \end{cases}$$

$$\textcircled{2} AS(G) = \begin{cases} \sum_{i=1}^k |e_i|, & G \neq \phi \\ 0, & G = \phi \end{cases}$$

$$\textcircled{3} G_{[m]} \subseteq G, C(G_{[m]}) = m$$

$$\textcircled{4} G_{[m]}^L \subseteq G, C(G_{[m]}^L) = m, \text{ 对于 } \forall G_{[m]}, AS(G_{[m]}^L) \leq AS(G_{[m]})$$

$$\textcircled{5} G_{[m]}^H \subseteq G, C(G_{[m]}^H) = m, \text{ 对于 } \forall G_{[m]}, AS(G_{[m]}^H) \geq AS(G_{[m]})$$

定义 3 B_2^n 空间中, 函数 $F(X) = x_{i_1}^* \wedge x_{i_2}^* \wedge \dots \wedge x_{i_m}^*$, $m \leq n$, 则由 $F^{-1}(1)$ 构成的空间结构称为正超立方体结构; 所有正超立方体组成的集合称为正超立方体系。

定义 4 B_2^n 空间中, 函数 $F(X) = x_{i_1}^* \wedge x_{i_2}^* \wedge \dots \wedge x_{i_l}^* \wedge (x_{i_1}^* \vee x_{i_2}^* \vee \dots \vee x_{i_h}^*)$, $l+h \leq n, h \neq 0$, 则由 $F^{-1}(1)$ 构成的空间结构称为凹超立方体结构; 所有凹超立方体组成的集合称为凹超立方体系。

定义 5 B_2^n 空间中, 函数 $F(X) = x_{i_1}^* \vee x_{i_2}^* \vee \dots \vee x_{i_l}^* \vee (x_{i_1}^* \wedge x_{i_2}^* \wedge \dots \wedge x_{i_h}^*)$, $l+h \leq n, h \neq 0$, 则由 $F^{-1}(1)$ 构成的空间结构称为凸超立方体结构; 所有凸超立方体组成的集合称为凸超立方体系。

定义 6 在 B_2^n 中, 对于样本 $X^c = (x_1^c, x_2^c, \dots, x_n^c)$, 如果存在集合 $U(d) = \{X = (x_1, x_2, \dots, x_n) \in F^{-1}(1) \mid d_H(X^c, X) \leq d\}$, 使得 $U(d) = F^{-1}(1)$, $\overline{U(d)} = F^{-1}(0)$, 其中 d 是位于区间 $[0, n]$ 中的整数, 则称 $U(d)$ 为以 X^c 为中心、 d 为半径的汉明球, $d_H(X^c, X)$ 为 X^c 与 X 之间的汉明距离。

定义 7 在 B_2^n 中, 对于样本 $X^c = (x_1^c, x_2^c, \dots, x_n^c)$, 如果存在集合 $U(d) = \{X = (x_1, x_2, \dots, x_n) \in F^{-1}(1) \mid d_H(X^c, X) \leq d\}$, 及 $K(d+1) = \{X = (x_1, x_2, \dots, x_n) \in F^{-1}(1) \mid d_H(X^c, X) = d+1\}$, 使得 $U(d) \cup K(d+1) = F^{-1}(1)$, $\overline{U(d) \cup K(d+1)} = F^{-1}(0)$, 其中 d 是位于区间 $[0, n]$ 中的整数, $C(K(d+1)) = m$, 则称 $M(d, m) = U(d) \cup K(d+1)$ 为以 X^c 为中心、 d 为半径的 m -阶汉明球突, $K(d+1)$ 为突点集。

定义 8 在 B_2^n 中, 如果布尔函数 $F(x_1, x_2, \dots, x_n)$ 能够表达为如下形式: $F(x_1, x_2, \dots, x_n) = x_1^* O(x_2^* O(\dots O(x_{n-1}^* O(x_n^*)) \dots))$, 其中算子 O 是逻辑与 $\wedge$ 或者是逻辑或 $\vee$, 则称布尔函数 $F(x_1, x_2, \dots, x_n)$ 为 SP 函数。

根据定义 8 可知, 定义 3 至定义 5 中的各类超立方体即为特殊的 SP 函数。

下面介绍上述几类线性可分结构系的判别定理。

3 线性可分结构系的判别方法

以下几个判别定理, 均是建立在如下类型的神经元 $S(X, W, \theta)$ 上:

$$S = U(\sum_{i=1}^n w_i x_i - \theta)$$

$$U(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$

$$X = (x_1, x_2, \dots, x_n), x_i \in \{-1, 1\}, i = 1, 2, \dots, n$$

定理 1^[16] 若二进神经网络中节点 S 的结构为 $S(X, W, \theta)$, 则 $S(X, W, \theta)$ 表达逻辑关系 $S = x_1^* \wedge x_2^* \wedge \dots \wedge x_n^*$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (1)$$

式中, $\theta_1 = AS(W) - 2AS(W_{[1]}^L)$, $\theta_2 = AS(W)$ 。

定理 2^[16] 若二进神经网络中节点 S 的结构为 $S(X, W, \theta)$, 则 $S(X, W, \theta)$ 表达逻辑关系 $S = x_1^* \vee x_2^* \vee \dots \vee x_n^*$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (2)$$

式中, $\theta_1 = -AS(W)$, $\theta_2 = -AS(W) + 2AS(W_{[1]}^L)$ 。

定理 3^[16] 若二进神经网络中节点 S 的结构为 $S(X, W, \theta)$, $W = W_x \cup W_y$, $W_x \cap W_y = \phi$, $W_x = \{a_1, a_2, \dots, a_n\}$, $W_y = \{b_1, b_2, \dots, b_m\}$, x_i 是通过权值 a_i 与节点 S 相联的输入, y_j 是通过权值 b_j 与节点 S 相联的输入, 则 $S(X, W, \theta)$ 表达逻辑关系 $S = x_1^* \wedge x_2^* \wedge \dots \wedge x_n^* \wedge (y_1^* \vee y_2^* \vee \dots \vee y_m^*)$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (3)$$

式中, $\theta_1 = AS(W_x) + AS(W_y) - 2\min(AS(W_y), AS(W_{x[1]}^L))$, $\theta_2 = AS(W_x) - AS(W_y) + 2AS(W_{y[1]}^L)$ 。

定理 4^[16] 若二进神经网络中节点 S 的结构为 $S(X, W, \theta)$, $W = W_x \cup W_y$, $W_x \cap W_y = \phi$, $W_x = \{a_1, a_2, \dots, a_n\}$, $W_y = \{b_1, b_2, \dots, b_m\}$, x_i 是通过权值 a_i 与节点 S 相联的输入, y_j 是通过权值 b_j 与节点 S 相联的输入, 则 $S(X, W, \theta)$ 表达逻辑关系 $S = (x_1^* \wedge x_2^* \wedge \dots \wedge x_n^*) \vee y_1^* \vee y_2^* \vee \dots \vee y_m^*$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (4)$$

式中, $\theta_1 = AS(W_x) - AS(W_y) - 2AS(W_{x[1]}^L)$, $\theta_2 = -AS(W_x) - AS(W_y) + 2\min(AS(W_x), AS(W_{y[1]}^L))$ 。

定理 5^[17] 若二进神经网络中节点 S 的结构为 $S(X, W, \theta)$, d 是位于区间 $[0, n]$ 之间的整数, 则 $S(X, W, \theta)$ 表达汉明球 $U(d)$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (5)$$

式中, $\theta_1 = AS(W) - 2AS(W_{d+1}^L)$, $\theta_2 = AS(W) - 2AS(W_d^H)$, 并且该汉明球的中心是 $X^c = (x_1^c, x_2^c, \dots, x_n^c)$, x_i^c 满足

$$x_i^c = \begin{cases} 1, & w_i > 0 \\ -1, & w_i < 0 \end{cases}$$

定理 6^[18] 若二进神经网络中节点 S 的结构为 $S(X, W, \theta)$, d 是位于区间 $[0, n]$ 之间的整数, 则 $S(X, W, \theta)$ 表达线性可分的汉明球突 $M(d, k)$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (6)$$

式中, $\theta_1 = AS(W) - 2AS(W_{d+1}^H)$, $\theta_2 = AS(W) - 2AS(W_{d+1}^L)$, 且 $\theta \leq AS(W) - 2AS(W_{d+1}^H)$, $\theta > AS(W) - 2AS$

$(W_{[d+2]})$, 并且该汉明球突的中心是 $X^c = (x_1^c, x_2^c, \dots, x_n^c)$, x_i^c 满足

$$x_i^c = \begin{cases} 1, & w_i > 0 \\ -1, & w_i < 0 \end{cases}$$

定理 7^[19] 若 $F(X)$ 是线性可分函数, $\sum_{i=1}^n w_i x_i - \theta = 0$ 是 $F(X)$ 在 B_2^n 中的分类超平面, 则一定存在 θ_1 和 θ_2 , 满足 $\theta_1 < \theta \leq \theta_2$ 。

定理 8 定理 7 中的 θ_1 即为当函数 $F(X) < 0$ 时 $\sum_{i=1}^n w_i x_i$ 的最大值; θ_2 即为当函数 $F(X) \geq 0$ 时 $\sum_{i=1}^n w_i x_i$ 的最小值, 即

$$\theta_1 = \max\{\sum_{i=1}^n w_i x_i \mid F(X) < 0, X = (x_1, x_2, \dots, x_n)\}$$

$$\theta_2 = \min\{\sum_{i=1}^n w_i x_i \mid F(X) \geq 0, X = (x_1, x_2, \dots, x_n)\}$$

证明: $F(X)$ 是线性可分函数, 则分类超平面 $\sum_{i=1}^n w_i x_i - \theta = 0$, 且 $\theta_1 < \theta \leq \theta_2$, 当 $F(X) \geq 0$ 时, 对于任意使得 $F(X) \geq 0$ 的 X 都有 $\sum_{i=1}^n w_i x_i - \theta \geq \sum_{i=1}^n w_i x_i - \theta_2 \geq 0$, 故 $\theta_2 \leq \min\{\sum_{i=1}^n w_i x_i\}$; 当 $F(X) < 0$ 时, 对于任意使得 $F(X) < 0$ 的 X 都有 $\sum_{i=1}^n w_i x_i - \theta < \sum_{i=1}^n w_i x_i - \theta_1 < 0$, 故 $\theta_1 \geq \max\{\sum_{i=1}^n w_i x_i\}$ 。

定理 7 及定理 8 指出了线性可分结构系的共性, 为提出新的线性可分结构系及其判别方法提供了研究方法。

对于定理 1 至定理 6, 其中所涉及到的样本 $X = (x_1, x_2, \dots, x_n)$, $x_i \in \{-1, 1\}$, $i = 1, 2, \dots, n$, 采用类似文献[18]推论 1 的证明过程或利用定理 8, 同样可得到对于如下神经元的 $O(X, W, \theta)$:

$$O = U(\sum_{i=1}^n w_i x_i - \theta)$$

$$U(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$x_i \in \{0, 1\}$, $i = 1, 2, \dots, n$ 。

下面给出超立方体、汉明球的判别方法。

推论 1 若二进神经网络中节点 O 的结构为 $O(X, W, \theta)$, 则 $O(X, W, \theta)$ 表达逻辑关系 $O = x_1^* \wedge x_2^* \wedge \dots \wedge x_n^*$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (7)$$

式中, $\theta_1 = AS(W) - AS(W_{[1]}^L)$, $\theta_2 = AS(W)$ 。

推论 2 若二进神经网络中节点 O 的结构为 $O(X, W, \theta)$, 则 $O(X, W, \theta)$ 表达逻辑关系 $O = x_1^* \vee x_2^* \vee \dots \vee x_n^*$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (8)$$

式中, $\theta_1 = 0$, $\theta_2 = AS(W_{[1]}^L)$ 。

推论 3 若二进神经网络中节点 O 的结构为 $O(X, W, \theta)$, $W = W_x \cup W_y$, $W_x \cap W_y = \phi$, $W_x = \{a_1, a_2, \dots, a_l\}$, $W_y = \{b_1, b_2, \dots, b_m\}$, x_i 是通过权值 a_i 与节点 O 相联的输入, y_j 是通过权值 b_j 与节点 O 相联的输入, 则 $O(X, W, \theta)$ 表达逻辑关系 $O = x_1^* \wedge x_2^* \wedge \dots \wedge x_l^* \wedge (y_1^* \vee y_2^* \vee \dots \vee y_m^*)$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (9)$$

式中, $\theta_1 = AS(W_x) + \max(AS(W_y) - AS(W_{[1]}^L), 0)$, $\theta_2 = AS(W_x) + AS(W_{[1]}^L)$ 。

推论 4 若二进神经网络中节点 O 的结构为 $O(X, W,$

$\theta)$, $W = W_x \cup W_y$, $W_x \cap W_y = \phi$, $W_x = \{a_1, a_2, \dots, a_l\}$, $W_y = \{b_1, b_2, \dots, b_m\}$, x_i 是通过权值 a_i 与节点 O 相联的输入, y_j 是通过权值 b_j 与节点 O 相联的输入, 则 $O(X, W, \theta)$ 表达逻辑关系 $O = (x_1^* \wedge x_2^* \wedge \dots \wedge x_l^*) \vee y_1^* \vee y_2^* \vee \dots \vee y_m^*$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (10)$$

式中, $\theta_1 = AS(W_x) - AS(W_{[1]}^L)$, $\theta_2 = \min(AS(W_x), AS(W_{[1]}^L))$ 。

推论 5 若二进神经网络中节点 O 的结构为 $O(X, W, \theta)$, d 是位于区间 $[0, n]$ 之间的整数, 则 $O(X, W, \theta)$ 表达汉明球 $U(d)$ 的充要条件是

$$\theta_1 < \theta \leq \theta_2 \quad (11)$$

式中, $\theta_1 = \frac{1}{2} AS(W) + \frac{1}{2} \sum_{i=1}^n w_i - AS(W_{[d+1]}^L)$, $\theta_2 = \frac{1}{2} AS(W) + \frac{1}{2} \sum_{i=1}^n w_i - AS(W_{[d]}^L)$, 并且该汉明球的中心是 $X^c = (x_1^c, x_2^c, \dots, x_n^c)$, x_i^c 满足

$$x_i^c = \begin{cases} 1, & w_i > 0 \\ 0, & w_i < 0 \end{cases}$$

推论 1 至推论 5 及文献[18]中的推论 1 给出了超立方体、汉明球、线性可分的汉明球突 3 类线性可分结构系的判别方法。其判别方法相似, 均是利用阈值的取值范围与权值之间的关系进行判别, 而 SP 函数目前没有类似上述直接通过阈值进行判别的方法, 而是通过权值与阈值分层进行判别^[20]。因此, 首先根据其阈值的取值范围, 讨论超立方体、汉明球、线性可分的汉明球突的阈值取值范围在实轴上的分布情况, 进一步讨论与二进神经元的覆盖情况; 接下来再利用汉明图, 考虑 SP 函数。

4 线性可分结构系与二进神经元的覆盖

对于 $W = \{w_1, w_2, \dots, w_n\}$, 记 $K = \frac{1}{2} AS(W) + \frac{1}{2} \sum_{i=1}^n w_i$, 将上述推论的结果汇总, 如表 1 所列。

表 1 阈值取值范围汇总表

线性可分结构系	阈值 θ 取值范围
$O = x_1^* \wedge x_2^* \wedge \dots \wedge x_n^*$	$AS(W) - AS(W_{[1]}^L) < \theta \leq AS(W)$
$O = x_1^* \vee x_2^* \vee \dots \vee x_n^*$	$0 < \theta \leq AS(W_{[1]}^L)$
$O = x_1^* \wedge x_2^* \wedge \dots \wedge x_l^* \wedge (y_1^* \vee y_2^* \vee \dots \vee y_m^*)$	$AS(W_x) + \max(AS(W_y) - AS(W_{[1]}^L), 0) < \theta \leq AS(W_x) + AS(W_{[1]}^L)$
$O = (x_1^* \wedge x_2^* \wedge \dots \wedge x_l^*) \vee y_1^* \vee y_2^* \vee \dots \vee y_m^*$	$AS(W_x) - AS(W_{[1]}^L) < \theta \leq \min(AS(W_x), AS(W_{[1]}^L))$
汉明球 $U(d)$	$K - AS(W_{[d+1]}^L) < \theta \leq K - AS(W_{[d]}^L)$
汉明球突 $M(d, k)$	$\begin{cases} K - AS(W_{[d+1]}^L) < \theta \leq K - AS(W_{[d+1]}^L) \\ K - AS(W_{[d+2]}^L) < \theta \leq K - AS(W_{[d]}^L) \end{cases}$

定理 9^[20] 二进神经元输入输出定义于 $\{0, 1\}$, $X = (x_1^*, x_2^*, \dots, x_j^*, \dots, x_n^*)$, $X' = (x_1^*, x_2^*, \dots, x_j^*, \dots, x_n^*)$, 如果对于布尔函数 $F(X)$, 用二进神经元 $O(X, W_1, \theta_1)$ 可以表达, 则对于布尔函数 $F(X')$, 可以用二进神经元 $O(X, W_2, \theta_2)$ 表达, 其中

$$\begin{cases} w_{2i} = w_{1i} & (i \neq j) \\ w_{2j} = -w_{1j} \\ \theta_2 = \theta_1 - w_{1j} \end{cases}$$

$$W_1 = (w_{11}, w_{12}, \dots, w_{1n}), W_2 = (w_{21}, w_{22}, \dots, w_{2n})$$

定理9表明,对于任意的二进神经元 $O(X, W, \theta)$, 可以将 $O(X, W, \theta)$ 中小于0的权系数乘以-1, 并相应调整阈值, 形成所有权系数都为正数的神经元 $O(X, W', \theta')$ 。故只需考虑所有权系数都为正数的神经元, $K = \frac{1}{2} AS(W) + \frac{1}{2} \sum_{i=1}^n w_i$ 可化简为 $AS(W)$, 将推论的结果精简, 如表2所列。

表2 权系数为正的阈值取值范围汇总表

线性可分结构系	阈值 θ 取值范围
$O = x_1^* \wedge x_2^* \wedge \cdots \wedge x_n^*$	$AS(W) - AS(W_{[1]}^L) < \theta \leq AS(W)$
$O = x_1^* \vee x_2^* \vee \cdots \vee x_n^*$	$0 < \theta \leq AS(W_{[1]}^L)$
$O = x_1^* \wedge x_2^* \wedge \cdots \wedge x_l^* \wedge (y_1^* \vee y_2^* \vee \cdots \vee y_m^*)$	$AS(W_x) + \max(AS(W_y) - AS(W_{x[1]}^L), 0) < \theta \leq AS(W_x) + AS(W_{y[1]}^L)$
$O = (x_1^* \wedge x_2^* \wedge \cdots \wedge x_l^*) \vee (y_1^* \vee y_2^* \vee \cdots \vee y_m^*)$	$AS(W_x) - AS(W_{x[1]}^L) < \theta \leq \min(AS(W_x), AS(W_{y[1]}^L))$
汉明球 $U(d)$	$AS(W_{[n-d-1]}^H) < \theta \leq AS(W_{[n-d]}^L)$
汉明球突 $M(d, k)$	$\begin{cases} AS(W_{[n-d-1]}^L) < \theta \leq AS(W_{[n-d-1]}^H) \\ AS(W_{[n-d-2]}^L) < \theta \leq AS(W_{[n-d]}^L) \end{cases}$

将阈值 θ 标注到实轴上, 由于权系数的不同, 往往不能确定 $AS(W_{[d]}^H)$ 与 $AS(W_{[d+1]}^L)$ 等大小, 故分为几部分分别讨论。图1给出了可以确定的4个线性可分结构, 在实轴上指出其阈值的分布情况; 根据权值的不同, 图2(a)至图2(e)将 $AS(W_{[n-d-2]}^H)$, $AS(W_{[n-d-1]}^L)$, $AS(W_{[n-d-1]}^H)$, $AS(W_{[n-d]}^L)$ 之间所有可能的大小关系排在数轴上。针对每种情况, 分别给出阈值取值范围所对应的线性可分结构系, 但其中的某些部分是未知结构, 接下来借助汉明图讨论未知部分是否为凹超立方体、凸超立方体或 SP 函数。

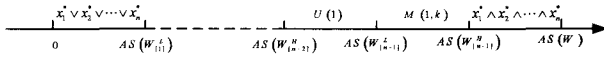


图1 $x_1^* \wedge x_2^* \wedge \cdots \wedge x_n^*$, $x_1^* \vee x_2^* \vee \cdots \vee x_n^*$, $U(1)$ 及 $M(1, k)$ 阈值分布

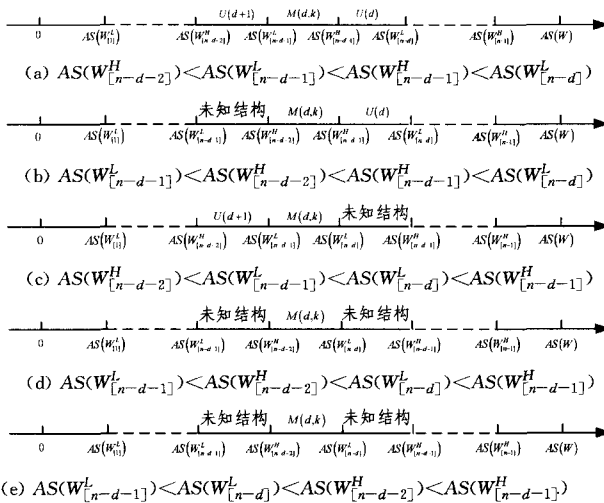


图2

下面介绍汉明图, 利用汉明图可以进一步判断图2(b)~2(e)中未知结构的可能情形。

定义9 n 维汉明图 $HG = [V, E]$ 是表示 n 维汉明几何空间的一个 n 次正则图。其节点集 V 有 2^n 个节点, 每个节点对应 n 元布尔函数 $F(x_1, x_2, \dots, x_n)$ 的一个样本 $X = (x_1, x_2, \dots, x_n)$, 即相应 n 维超立方体的一个顶点。若 $F(x_1, x_2, \dots, x_n) = 1$, 称 $X = (x_1, x_2, \dots, x_n)$ 为真节点; 若 $F(x_1, x_2, \dots, x_n) = 0$, 称

$X = (x_1, x_2, \dots, x_n)$ 为伪节点。 HG 的边集 E 有 2^{n+1} 条边, 每条边邻接两节点间的汉明距离均为1。

汉明图在核心节点未确定之前是无向图; 确定之后其是有向图, 并规定核心节点为第0层。图3是 $n=4$ 、核心节点为 $(0, 0, 0, 0)$ 的汉明图。将图中节点编码与所要求的不同核心节点的二进码按位相加, 即可得任意核心节点的有向汉明图, 如图3所示。

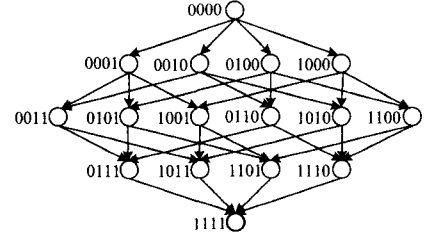


图3 $n=4$ 、核心节点为 $(0, 0, 0, 0)$ 的汉明图

对于中心为 $X^c = (x_1^c, x_2^c, \dots, x_n^c)$ 、半径为 d 的汉明球 $U(d)$ 对应的汉明图, 其核心节点为 $X^c = (x_1^c, x_2^c, \dots, x_n^c)$ 。从第0层一直到第 d 层节点都为真节点, 其余节点均为伪节点。

对于中心为 $X^c = (x_1^c, x_2^c, \dots, x_n^c)$ 、半径为 d 的汉明球突 $M(d, m)$ 对应的汉明图, 其核心节点为 $X^c = (x_1^c, x_2^c, \dots, x_n^c)$ 。从第0层一直到第 d 层节点都为真节点, 第 $d+1$ 层有 m 个节点为真节点, 其余的节点均为伪节点。

图4为中心在 $(0, 0, 0, 0)$ 、半径为2的汉明球 $U(2)$; 图5为中心在 $(0, 0, 0, 0)$ 、半径为1的汉明球突 $M(1, 3)$, 其中实心节点为真节点, 空心节点为伪节点。

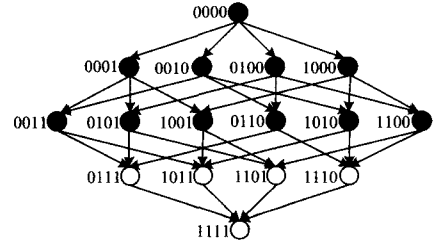


图4 中心为 $(0, 0, 0, 0)$ 的汉明图 $U(2)$

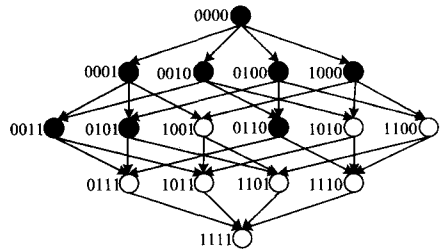


图5 中心为 $(0, 0, 0, 0)$ 的汉明球突 $M(1, 3)$

利用汉明图清晰的几何结构, 综合图2(b)~2(e)中的未知结构, 可将其归纳为如下5种情况:

(1) 若 $AS(W_{[n-d-1]}^L) < \theta \leq AS(W_{[n-d-2]}^H) < AS(W_{[n-d]}^L)$, 则第1层至第 d 层全部为真节点, 第 $d+1$ 层及第 $d+2$ 层有部分真节点, 第 $d+3$ 层及后面层的情况还需进一步判断 θ 与 $AS(W_{[n-d-3]}^L)$ 之间的大小关系;

(2) 若 $AS(W_{[n-d-1]}^L) < AS(W_{[n-d]}^L) < \theta \leq AS(W_{[n-d-1]}^H)$, 则第 d 层、第 $d+1$ 层及第 $d+2$ 层有部分真节点, 其他层的情况还需进一步判断;

(3) 若 $AS(W_{[n-d-1]}^L) < \theta \leq AS(W_{[n-d]}^L) < AS(W_{[n-d-2]}^H)$,

则第 1 层至第 d 层全部为真节点,第 $d+1$ 层及第 $d+2$ 层有部分真节点,第 $d+3$ 层及后面层的情况还需进一步判断 θ 与 $AS(W_{[n-d-3]}^L)$ 之间的大小关系;

(4)若 $AS(W_{[n-d]}^L) < \theta \leq AS(W_{[n-d-2]}^H)$,则第 d 层、第 $d+1$ 层及第 $d+2$ 层有部分真节点,其他层的情况还需进一步判断;

(5)若 $AS(W_{[n-d]}^L) < AS(W_{[n-d-2]}^H) < \theta \leq AS(W_{[n-d-1]}^H)$,则第 d 层及第 $d+1$ 层有部分真节点, $d+2$ 层及其后面无真节点,第 d 层之前的情况还需进一步判断。

对凹超立方体、凸超立方体或 SP 函数的结构进行分析,将其标注到汉明图上,可以发现其属于上述 5 种情况中的某一部分,但并不清楚是否完全覆盖这 5 种情况。另外,超立方体可归类为特殊的 SP 函数,而在已知的线性可分函数系中,SP 函数没有找到阈值与权值直接判别法,故需对 SP 函数的判别法提炼出直接判别法,这将会进一步解决线性可分函数系的覆盖问题。

结束语 已知的线性可分结构系中正超立方体、汉明球、线性可分的汉明球突的判别法较为明确简洁,而凹超立方体、凸超立方体及 SP 函数的判别方法较为复杂,有待进一步简化。另外,根据阈值在实轴上的取值范围,通过本文的分析,目前已知的几类线性可分结构系仍不能覆盖所有的二进神经元,即仍存在其他线性可分结构系,有待进一步研究。本文指出了未知的线性可分结构系阈值的取值范围,为提出新的线性可分结构系及判别法提供了研究方向。

参 考 文 献

- [1] Gray D L, Michel A N. A training algorithm for binary feed forward neural networks[J]. IEEE Trans. Neural Networks, 1992, 3(2): 176-194
- [2] Kim J H, Park S K. The geometrical learning of binary neural networks[J]. IEEE Trans. Neural Networks, 1995, 6(1): 237-247
- [3] Muselli M. On sequential construction of binary neural networks[J]. IEEE Trans Neural Networks, 1995, 6(3): 678-690
- [4] Yamamoto A, Saito T. A flexible learning algorithm for binary neural networks[J]. IEICE Trans. Fundamentals, 1998, E81-A(9): 1925-1930
- [5] Chen Fang-yue, Chen Guan-rong, He Guo-long. Universal Perceptron and DNA-Like Learning Algorithm for Binary Neural Networks;LSBF and PBF Implementations[J]. IEEE Transac-
- tions on Neural network, 2009, 20(10): 1645-1658
- [6] Chen Fang-yue, Chen Guan-rong, He Qin-bin. Universal Perceptron and DNA-Like Learning Algorithm for Binary Neural Networks; Non-LSBF Implementation[J]. IEEE Transactions on Neural network, 2009, 20(8): 1293-1301
- [7] Lu Yang, Yang Juan, Wang Qiang, et al. The upper bound of the minimal number of hidden nodes in parity problem by neural networks[J]. Science China Information Sciences, Oct. 2011
- [8] Chua L O. Visions of nonlinear science in the 21st century in CNN; A Paradigm for Complexity[M]. Singapore: World Scientific, 1999
- [9] Chen Fang-yue, He Guo-long, Chen Guan-rong. Realization of Boolean Functions via CNN; Mathematical Theory, LSBF and Template Design[J]. Circuits and Systems I: Regular Papers, IEEE Transactions on, 2006, 53(10): 2203-2213
- [10] Wegener I. The complexity of Boolean function[M]. New York: Wiley, 1987
- [11] Crounse K R, Fung E L, Chua L O. Efficient implementation of neighborhood logic for cellular automata via the cellular neural network universal machine[J]. IEEE Transactions on Circuit System I, Fundam. Theory Application, 1997, 4(3): 355-361
- [12] Nemes L, Chua L O, Roska T. Implementation of arbitrary Boolean function on a CNN universal machine[J]. Int. J. Circuit Theory Appl., 1998, 26(6): 593-610
- [13] Haykin S. Neural Networks: A comprehensive foundation (2nd ed) [M]. Englewood Cliffs, NJ: Prentice-Hall, 1999
- [14] Negnevitsky M. Artificial Intelligence: A guide to intelligent systems(2nd ed)[M]. New York: Addison-Wesley, 2004
- [15] Hassoun M H. Fundamentals of artificial neural networks[M]. New York: MIT Press, 1995
- [16] 陆阳, 韩江洪, 张维勇. 二进神经网络逻辑关系判据及等价性规则提取[J]. 模式识别与人工智能, 2001, 14(2): 171-176
- [17] 陆阳, 魏臻, 高隽, 等. 二进神经网络中汉明球的逻辑意义及一般判别方法[J]. 计算机研究与发展, 2002, 3(1): 79-86
- [18] 杨娟, 陆阳, 黄振谨, 等. 二进神经网络中的汉明球突及其线性可分性[J]. 自动化学报, 2011, 37(6): 737-745
- [19] 陆阳, 韩江洪, 张维勇. 二进神经网络中笛卡尔球的研究[J]. 模式识别与人工智能, 2004, 17(3): 368-373
- [20] Lu Yang, Han Jiang-hong, Wei Zhen. A general judging and constructing method of SP functions in binary neural networks[J]. Acta Automatica Sinica, 2003, 29(2): 234-241

(上接第 164 页)

- [7] Aggarwal C C, Li Y, Wang J, et al. Frequent pattern mining with uncertain data[C]// International Conference on Knowledge Discovery and Data Mining. 2009
- [8] Bernecker T, Kriegel H, Renz M, et al. Probabilistic frequent itemset mining in uncertain databases[C]// International Conference on Knowledge Discovery and Data Mining. 2009
- [9] Muntz R, Mining Y. Frequent Itemsets in Uncertain Datasets [R]. CSD-TR No. 030042
- [10] Tobji M A B, Yaghlane B B, Mellouli K. Frequent Itemset Mining from Databases Including One Evidential Attribute[C]// International Conference on Scalable Uncertainty Management. 2008
- [11] Hewawasam K K R G K, Premaratne K, Subasingha S P, et al. Rule Mining and Classification in Imperfect Databases[J]. IEEE Transactions on Systems, Man and Cybernetics, 2007
- [12] Tobji M A B, Yaghlane B B, Mellouli K. A New Algorithm for Mining Frequent Item-sets from Evidential Databases[C]// International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems. 2008
- [13] Zhang Q, Li F, Yi K. Finding frequent items in probabilistic data [C]// International Conference of SIGMOD. 2008
- [14] Leung C K, Hao B. Mining of frequent itemsets from streams of uncertain data[C]// IEEE International Conference on Data Engineering. 2009
- [15] Leung C K, Hao B, Jiang F. Constrained frequent itemset mining from uncertain data streams [C]// IEEE International Conference on Data Engineering. 2010