第9卷第3期
1998年3月

软 件 学 报
JOURNAL OF SOFTWARE

Vol. 9, No. 3
Mar. 1998

基于任意给定训练集的离散型 Hopfield 网学习算法*

孟祥武

（北京邮电大学计算机工程系 北京 100876）

程 虎

（中国科学院软件研究所 北京 100080）

摘要 本文提出了一个离散型 Hopfield 网联想记忆学习算法，该算法增加了训练样本的维数，因而能存储任意给定的训练模式集. 实验结果也证明了该方法的有效性.

关键词 Hopfield 网，联想记忆，学习算法，神经网络，人工智能.

中图法分类号 TP18

离散型 Hopfield 网是一个重要的神经网络模型[1]，主要用于联想记忆. 另外，在光学上还容易实现它的并行运算，[2,3]这使得 Hopfield 网广泛应用于复杂问题成为可能.

但 Hopfield 网主要又存在以下几个缺点：(1) 记忆的样本个数少. 对于随机样本，Hopfield 网的样本容量 $m \leqslant 0.15n$，n 为神经元个数.[1]由于系统规模的限制，包括软件系统或器件上的限制，神经元的个数不能太多. 从而限制了存储样本的个数. (2) 不能学习任意给定的训练集数据. Hopfield 网仅能记忆相关性弱的样本，不能记忆相关性强的样本.

已有的研究工作主要集中于扩大 Hopfield 网的记忆容量.[4~8]文献[8]还提出了一个建立 Hopfield 网的一般方法——利用线性规划方法求解 Hopfield 网的权值. 它可以解决任意连接的网络的训练，并扩大了网络的容量，为 Hopfield 网的广泛应用打下了基础，但该方法也不能学习任意给定的训练集数据.

本文针对上述缺点，提出了一个基于任意给定训练集的离散型 Hopfield 网联想记忆学习算法. 该算法通过增加训练样本的维数(分量)，减弱了样本间的相关性，从神经网络角度看，就是增加神经元个数，结果提高了 Hopfield 网的存储和联想容错能力. 所提算法能有效地克服上述缺点，并且简单易行，便于实现. 从而可解决离散型 Hopfield 网联想记忆学习问题. 实验结果也证明了该算法的有效性.

1 离散型 Hopfield 网

离散型 Hopfield 网有 n 个神经元，权矩阵 W 为 $n\times n$ 阶实数矩阵. 每个神经元的状态为 1 或 -1. 称 n 维矢量训练模式对被存入 Hopfield 网中，若满足

$$Y=\mathrm{sgn}(WX-\theta),\ X\in\{-1,1\}^n,\ Y\in\{-1,1\}^n,\ W\in R^{n\times n},\ \theta\in R^n.$$

$$\mathrm{sgn}(x)=\begin{cases}1, & \text{当 } x\geq 0;\\ -1, & \text{当 } x<0.\end{cases}$$

在学习过程中，学习算法的目的是确定一组权值 W_{ij}，使得 m 个待记忆的二值样本$[X(u),Y(u)]$存入 Hopfield 网中，$u=1,2,\ldots,m$；W_{ij}表示神经元 i 与神经元 j 之间的连接权，$i,j=1,2,\ldots,n$. 每个样本为一对 n 维矢量.

$$[X(u),Y(u)]=[(X_1(u),X_2(u),\ldots,X_n(u),Y_1(u),Y_2(u),\ldots,Y_n(u)]$$

$$X_i(u)\in\{-1,1\},\quad Y_i(u)\in\{-1,1\}.$$

2 基于任意给定训练集的离散型 Hopfield 网联想记忆学习

学习算法的目的是寻找一组权值 W_{ij}，$i,j=1,2,\ldots,n$，其中 n 为神经元数.

* 作者孟祥武，1966 年生，博士，讲师，主要研究领域为神经网络，进化算法. 程虎，1938 年生，研究员，博士导师，主要研究领域为程序编译，软件工程，人工智能和神经网络.

本文通讯联系人：孟祥武，北京 100876，北京邮电大学计算机工程系 135 信箱

本文 1997-01-23 收到原稿，1997-04-11 收到修改稿

m 个二值训练样本：$[X(u),Y(u)]=[X_1(u),X_2(u),\ldots,X_n(u),Y_1(u),Y_2(u),\ldots,Y_n(u)]$，$u=1,2,\ldots,m$；其系统稳定状态的充分条件是

$$\Big[\sum_{j=1}^{n}W_{ij}X_j(u)\Big]Y_i(u)>0,\qquad X_i(u),Y_i(u)\in\{-1,1\}$$

为求满足上述条件的 W_{ij}，Hopfield 网的学习问题可转化为求解 n 组线性不等式组

$$X_1(u)Y_i(u)W_{i1}+X_2(u)Y_i(u)W_{i2}+\ldots+X_n(u)Y_i(u)W_{in}>0,$$

其阈值 θ 均为 0. 其中每组中含有 m 个不等式.

定理 1. 设有 m 个二值训练样本：$[X(u),Y(u)]=[X_1(u),X_2(u),\ldots,X_n(u),Y_1(u),Y_2(u),\ldots,Y_n(u)]$，$(u=1,2,\ldots,m)$. 若对于任意 i 值 $(i=1,2,\ldots,n)$ 都至少存在一个 j 值 $(j=1,2,\ldots,n)$，使得 $X_j(1)Y_i(1)=X_j(2)Y_i(2)=\ldots=X_j(m)Y_i(m)$ 成立 $(X_j(u),Y_i(u)\in\{-1,1\})$，则这 m 个样本为系统的稳定状态.

证明：Hopfield 网的学习问题可转化为求解 n 组线性不等式组的问题. 对于任意第 i 组不等式 $(i=1,2,\ldots,n)$，因为有 m 个样本，故要同时满足 m 个不等式.

$$\begin{cases}X_1(1)Y_i(1)W_{i1}+X_2(1)Y_i(1)W_{i2}+\ldots+X_n(1)Y_i(1)W_{in}>0\\ X_1(2)Y_i(2)W_{i1}+X_2(2)Y_i(2)W_{i2}+\ldots+X_n(2)Y_i(2)W_{in}>0\\ \ldots\ldots\\ X_1(m)Y_i(m)W_{i1}+X_2(m)Y_i(m)W_{i2}+\ldots+X_n(m)Y_i(m)W_{in}>0\end{cases}$$

其阈值 θ 均为 0.

对于任意 j，$j=1,2,\ldots,n$；当 $X_j(1)Y_i(1)=X_j(2)Y_i(2)=\ldots=X_j(m)Y_i(m)$ 成立时，令 W_{ij} 与 $X_j(1)Y_i(1)$ 同符号，即 $W_{ij}\,X_j(1)Y_i(1)>0$；其余 $W_{ik}=0$，当 $k\neq j$.

因为对于任意 i 值 $(i=1,2,\ldots,n)$ 都至少存在一个 j 值 $(j=1,2,\ldots,n)$，使得 $X_j(1)Y_i(1)=X_j(2)Y_i(2)=\ldots=X_j(m)Y_i(m)$ 成立. 故对于任意样本 u，$u=1,2,\ldots,m$，

$$\Big[\sum_{k=1,k\neq j}^{n}W_{ik}X_k(u)\Big]Y_i(u)=0,\quad W_{ij}X_j(u)Y_i(u)>0,\quad \Big[\sum_{k=1}^{n}W_{ik}X_k(u)\Big]Y_i(u)>0$$

因为是对于任意 i 的值，$(i=1,2,\ldots,n)$，即对于 n 组不等式组中的每一组不等式组，上式均成立.

所以，这 m 个样本为系统的稳定状态. □

定理 2. 若 m 个二值训练样本：$[X(u),Y(u)]=[X_1(u),X_2(u),\ldots,X_n(u),Y_1(u),Y_2(u),\ldots,Y_n(u)]$ $(u=1,2,\ldots,m)$ 不能都成为系统的稳定状态，则通过增加样本的维数（分量），可以使样本都变成系统的稳定状态.

证明：设存在 i 值，$i=1,2,\ldots,n$，且不存在一个 j 值，$j=1,2,\ldots,n$，使得 $X_j(1)Y_i(1)=X_j(2)Y_i(2)=\ldots=X_j(m)Y_i(m)$ 成立.

则对于所有的样本 u，$u=1,2,\ldots,m$. 令 $X_{n+1}(u)=Y_i(u)$，故

$$X_{n+1}(1)Y_i(1)=X_{n+1}(2)Y_i(2)=\ldots=X_{n+1}(m)Y_i(m).$$

此时，令 $n=n+1$，若还存在上述 i 值，重复上述过程，使得对于任意 i 值，均有

$$X_j(1)Y_i(1)=X_j(2)Y_i(2)=\ldots=X_j(m)Y_i(m)\quad j=1,2,\ldots,n$$

因此，根据定理 1，此时这 m 个样本为系统的稳定状态，即样本增加维数（分量）后，可以使样本都变成系统的稳定状态. □

定理 3. 所有样本至多增加 n 维（个分量），就可学习成功.

证明：若对于每一组不等式，均有 $W_{ij}=0$，$i,j=1,2,\ldots,n$，则针对每一组不等式，对每个样本增加一维（分量），最多只有 n 组不等式. 故最多增加 n 维（个分量）. □

定理 3 说明：增加训练样本的维数（分量）是有限的，从神经网络角度看，就是增加有限个神经元.

基于任意给定训练集的离散型 Hopfield 网联想记忆学习算法：

(1) 利用给定训练集（样本），产生相应的线性不等式组，$i=1$；

(2) 对于第 i 组不等式 $(i=1,2,\ldots,n)$，判断是否至少存在一个 j 值 $(j=1,2,\ldots,n)$，使得 $X_j(1)Y_i(1)=X_j(2)Y_i(2)=\ldots=X_j(m)Y_i(m)$ 成立. 若存在这样的 j 值，令 W_{ij} 与 $X_j(1)Y_i(1)$ 同符号，即 $W_{ij}\,X_j(1)Y_i(1)>0$，对于 $k=1,2,\ldots,n$，且 $k\neq j$，$W_{ik}=0$，其阈值 θ 均为 0. 转(5)；

(3) 对每一样本增加一维（分量），令 $X_{n+1}(u)=Y_i(u)$，$Y_{n+1}(u)=Y_i(u)$，$u=1,2,\ldots,m$；

(4) $n=n+1$，转(1)；

(5) $i=i+1$，若 $i>n$，结束；否则，转(2).

通过该算法获得的连接权矩阵是非对称的，即是一个非对称网络，这种采用权值的非对称连接，有利于表达信息成分间的不等同关系，具有一定的优越性.[9]

若连接权矩阵是对称的，精确地保证 $W_{ij}=W_{ji}$，硬件实现几乎是不可能的. 实际上，物理实现的神经网络是不对称的.[10]

3 实验结果

给定训练集（样本），构造相应的 Hopfield 网.

(1) XOR 问题.

x_1	x_2	y_1	y_2
−1	−1	−1	−1
−1	1	1	1
1	−1	1	1
1	1	−1	−1

转为相应不等式组：
$$\begin{cases} -W_{11}-W_{12}<0 \\ -W_{11}+W_{12}\geq 0 \\ W_{11}-W_{12}\geq 0 \\ W_{11}+W_{12}<0 \end{cases}$$

因为第 1 个不等式与第 4 个不等式矛盾，无法获得相应权值，即学习不成功. 增加样本的维数（分量）

x_1	x_2	x_3	y_1	y_2	y_3
−1	−1	−1	−1	−1	−1
−1	1	1	1	1	1
1	−1	1	1	1	1
1	1	−1	−1	−1	−1

则可获相应权矩阵：
$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$

学习成功.

(2) 选择函数，根据 x_1 值，选择输出 x_2 或 x_3.

x_1	x_2	x_3	y_1	y_2	y_3
−1	−1	−1	−1	−1	−1
−1	−1	1	1	1	1
−1	1	−1	−1	−1	−1
−1	1	1	1	1	1
1	−1	−1	−1	−1	−1
1	−1	1	−1	−1	−1
1	1	−1	1	1	1
1	1	1	1	1	1

根据以上样本，无法获得相应权值，即学习不成功. 增加样本的维数（分量），则可获相应权矩阵：
$$\begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

学习成功.

4 结束语

本文提出了一个新的离散型 Hopfield 网联想记忆学习算法. 对于任意给定的训练集数据，当学习不成功时，该算法通过增加训练样本的维数（分量）（从神经网络角度看，就是增加有限个神经元结点），能成功地学习所给训练集数据，并提高了网络的容量. 实验结果也表明了该算法的有效性.

参考文献

1 Hopfield J J. Neural networks and physical systems with emergent collective computational abilities. In: Proceedings of the National Academy of Science. USA, 1982,79:2554～2558

2 Psaltis D, Farhat N. Optical information processing based on an associative-memory model of neural nets with thresholding and feedback. Optics Letters, 1985,10(2):98～100

3 Farhat N *et al*. Optical implementation of the Hopfield model. Applied Optics, 1985,24(10):1469～1475

4 McEliece R J *et al*. The capacity of the Hopfield associative memory. IEEE Transactions on Information Theory, 1987,33(4):461～482

5 Amari S, Maginu K. Statistical neurodynamics of associative memory. Neural Networks, 1988,1(1):63～73

6 Wang J H *et al*. Determination of the Hopfield associative memory characteristics using a single parameter. Neural Networks, 1990,3(3):319～331
7 Chang Jiu Chen *et al*. High capacity for the Hopfield neural networks. IEEE International Conference on Neural Networks, 1994. 1175～1180
8 刘晓鸿,戴汝为.建立Hopfield网的一般方法.自动化学报,1996,22(3):301～307
(Liu Xiao-hong, Dai Ru-wei. General methods of construction of Hopfield neural networks. Acta Automatica Sinica, 1996,22(3):301～307)
9 张承福等.几种联想记忆神经网络模型的分析.模式识别与人工智能,1990,3(1):14～21
(Zhang Cheng-fu *et al*. The analysis of some associative memory neural network models. Pattern Reagnition and Artificial Intelligence, 1990,3(1):14～21)
10 Xu Z B *et al*. Asymmetric Hopfield-type networks: theory and applications. Neural Networks, 1996,9(3):483～501

A Learning Algorithm of Discrete-time Hopfield with any Given Training Set

MENG Xiang-wu

(*Department of Computer Engineering Beijing University of Posts and Telecommunications Beijing* 100876)

CHENG Hu

(*Institute of Software The Chinese Academy of Sciences Beijing* 100080)

Abstract This paper presents a learning algorithm of Hopfield's discrete-time associative memory. The algorithm adds the dimensions of training patterns, so it can store any given training patterns set. Experimental results also demonstrate the effectiveness of the approach.

Key words Hopfield neural networks, associative memory, learning algorithm, neural networks, artificial intelligence.

Class number TP18