简报

神经网络中伴随着搜索过程的联想

刘 刚 梁明理

(武汉大学物理系，武汉 430072)

关键词 神经网络，联想，学习律，突触联系强度，相关性

1982年 J. J. Hopfield 提出了他的神经网络模型[1]，并发现它具有许多集合运算特性[2]. 由于这一模型与凝聚态理论的自旋玻璃模型有惊人的相似[3]，可在研究中引入物理学的诸如统计力学等科学方法来考察神经网络的特性，其中对网络稳态条件的研究提出了网络综合的观点，即怎样决定突触矩阵和阈值等网络参数，才能保证记忆态是稳定的且能成为吸引子. 现在我们在 Hopfield 模型的基础上，提出一个联想记忆模型，在输入一个新图样时，网络能通过一种算法，自动找到与它强相关的记忆图样. 由于该模型是网络稳态条件的解，因而系统的动力学能保证该联想能作为一个稳态吸引子被记忆.

网络的稳态条件

McCulloch-Pitts 模型指出[4]，神经元是一个两态的、有几个输入一个输出的阈值部件. 在时刻 t 神经元 i 的状态

$$\lambda_i = \begin{cases} H = 1 & (\text{兴奋}) \\ L = -1 & (\text{抑制}) \end{cases}$$

对于一个由 n 个神经元组成的网络，已经记忆了 M 个图样，神经元之间的突触系数矩阵为 $C^{(M)}$，神经元 i 的自旋状态为 $\sigma_i(t)$，则网络的稳态条件可表示为[5]

$$\sum_{j=1}^{n}(C_{ij}\lambda_j - \theta_i)\sigma_i \geqslant 0 \quad (i = 1,2,\cdots,n)$$

式中 θ_i 为神经元的阈值. 对于一组给定的 P 个记忆态 $\{\lambda^K\}$，可用 $\vec{\lambda}^K$ 表示，若要求矢量 $\vec{\lambda}^K$ 表示的态稳定，则

$$C\vec{\lambda}^K - \vec{\theta} = A^K\vec{\sigma}^K \ (K = 1,2,\cdots,p) \tag{1}$$

式中 A 为任意非对角阵. 如果适当选择 A^K 和 $\vec{\theta}$，则稳态条件变为[5]

$$C\lambda^K = \lambda^K \tag{2}$$

式中

$$C = A^I A \tag{3}$$

这里 A 是由矢量 $\vec{\lambda}^K$ 的元素组成的矩阵，A^I 为 A 的伪变换[6].

若记忆矢量线性独立，方程(3)的非平常解为

$$C = \Sigma\Sigma^I$$

式中 Σ 是由矢量 $\vec{\sigma}^K$ 组成的矩阵. 若记忆矢量是正交的，则 $(\Sigma^I\Sigma)^{-1}$ 可简化为 $(1/n)I$，以致矩阵 C 就是 L. Cooper 提出的学习规则

$$C = (1/n)\Sigma\Sigma^T$$

在许多文章中，总使得记忆或突触矩阵从零开始，这不符合生物学规律，人脑的神经网络似乎是从混沌的到清晰的(有序的)，用数学的语言来说是由一个随机矩阵发展起来的. 对 P 个记忆态，由(2)式可得

$$C\Sigma = \Sigma \qquad \Sigma = (\sigma^1, \sigma^2, \cdots, \sigma^K, \cdots, \sigma^P) \tag{4}$$

上式总是有解的，解的一般形式为[5]

$$C = \Sigma\Sigma^I + B(I - \Sigma\Sigma^I) \tag{5}$$

式中 B 是一任意 (n, n) 矩阵，Σ^I 是 Σ 的 Moore-Penrose 变换[6]. $B = 0$ 时，$C = \Sigma\Sigma^I$，与经典的 Hebb 学习律一致. 显然，即使矩阵 B 是对称的，突触矩阵 C 在学习后一般不再是对称的，与 Hopfield 模型相比，(5)式更接近于实际情况.

若原记忆态矢量是正交的，则 $\Sigma^I\Sigma = (1/n)I$，I 为单位矩阵，于是 $\Sigma^I = (\Sigma^T\Sigma)^{-1}\Sigma^T$ 可简化为 $\Sigma^I = \Sigma^T$，代入(5)式得

$$C = B + \frac{1}{n}(B - I)\Sigma\Sigma^T \tag{6}$$

于是我们得到了网络稳定条件的解，它与 Hebb 学习律一起，产生一新的突触强度矩阵 $\{C_{ij}^{new}\}$.

一般来说，学习是个渐进过程，每次学习一个新图样，突触矩阵经历一个变化，所有的矩阵也经历了一个变化，因而(6)式应改为迭代形式.

设神经网络已学习了 $K-1$ 个图样，在学习了新图样 σ^K 后，突触矩阵 $C(K-1)$ 随之改变，给出一个新矩阵

$$C(K)=C(K-1)+\frac{1}{n}(B-I)\sigma^K\sigma^{KT}$$

可以证明[7]

$$C(K-1)\sigma^K=B\sigma^K$$

若先前图样记忆矢量是正交的，可得学习规则的迭代形式为

$$C(K)=C(K-1)+\frac{1}{n}[C(K-1)-I]\sigma^K\sigma^{KT} \quad (7)$$

在学习第一个图样时，$K=1$，则

$$C(0)=B$$

上二关系表明，学习过程是从由 B 产生的初始图样逐次迭代进行的.

具备搜寻过程的联想学习规则

突触强度的改变 $\triangle C_{ij}$ 能简单地写作类似 Hebb 学习律的形式，由(7)式可得

$$\triangle C_{ij}=\frac{1}{n}\sigma_j^K\sum_{r=1}^{n}C_{ir}(K-1)\sigma_r^K-\frac{1}{n}\sigma_i^K\sigma_j^K \quad (8)$$

式中 $C_{ir}(K-1)$ 是学习新图样 σ^K 前的矩阵. 上式第一项中的求和因子之值是在一个新图样 σ^K 输入到网络 $C(K-1)$ 时，神经细胞 i 的膜电位，它包含与 σ^K 相关图样的信息，体现了记忆的搜寻作用. 第二项表示神经细胞 i 和 j 的局部作用，表达了对新图样的抑制作用. 因此 $\triangle C_{ij}$ 的变化依赖于神经细胞 j 的状态和神经细胞 i 的膜电位，它吸取了所有与神经细胞 i 相关的全部突触影响. 表明新的学习规则具有选择性联想的性质.

数值模拟和结果

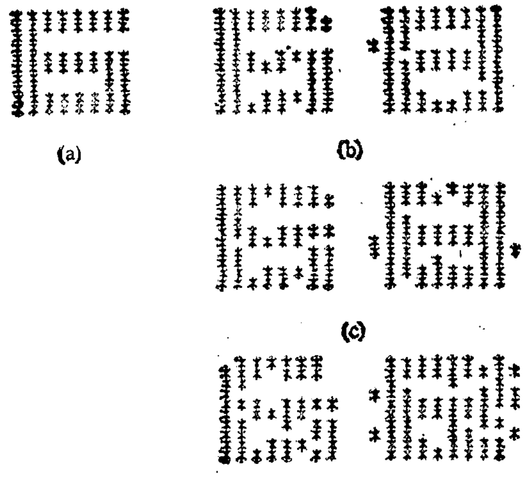
我们采用 IBM-XT 计算机，对新的学习规则(8)式进行了模拟检验. 设置神经网络由 100 个神经元构成，记忆 10 个图样，其中 9 个为随机图样，一个为指定图样，如图 1 所示.

(a) (b)

图 1 设置神经网络中记忆的 10 个图样

(a) 随机图样之一 (b) 指定图样(字符 8)

在输入一个与字符"8"相关的字符"6"后，得到记忆图样"8"，如图 2 所示.



(a) (b) (c) (d)

图 2 计算机模拟结果，输入为相关字符"6"，联想结果为记忆图样"8"

(a) 输入新图样"6"(b)输入带 5% 噪音的图样"6"，联想结果为 96% 重迭率的图样"8" (c)输入带 10% 噪音的图样"6"，联想结果为 92% 重迭率的图样"8" (d) 输入带 15% 噪音的图样"6"，联想结果为 83% 重迭率的图样"8"

上述模拟结果表明，新的学习律能完成相关图样的自动搜索，并联想出来，而且具备一定的容错能力. 但是随着噪音的进一步加大，联想结果的重迭率迅速下降，以至输出伪态.

参考文献

1 Hopfield J J. *Proc Natl Acad Sci USA*, 1982; **79**: 2554

2 Personnaz L *et al*. *Phys Review A*, 1986; **34** (5): 4217

3 Amit D J *et al*. *Phys Review A*, 1985; **32**(2):1007

4 Shinomoto S. *Biol Cybern*, 1987; **57**: 197

5 Personnaz L *et al*. *J Physique Lett*, 1985; **46**: 359

6 Albert A. *In regression and the moore-penrose pseudoinverse*. New York: Acad Press, 1972

7 Personnaz L *et al*. *J Statistical Phys*, 1986; **43**(5): 411

[本文于 1990 年 5 月 28 日收到，1991 年 1 月 3 日修回]