

# Hopfield 神经网络模型的恢复特性\*

马余强 张玥明<sup>†</sup> 龚昌德

南京大学物理系, 南京 210008;

中国高等科学技术中心, 北京 100080

<sup>†</sup> 南京医学院物理系, 南京 210029

1992年10月5日收到

通过引入不同概率的双峰无规神经激活阈分布, 来考虑对神经网络“记忆”恢复特性的影响, 结果表明即使储存模式数超过孤立 Hopfield 模型的临界值  $\alpha_c$  时系统仍然能成功地恢复储存信息.

PACC: 7510H; 8710; 6460; 0550

## 一、引 言

近年来, 神经网络的类自旋玻璃模型的研究日益受到重视<sup>[1-9]</sup>. 神经系统中的集体行为为已成功地解释了大脑记忆、学习、信息储存和恢复特性. 其中讨论的大多模型是由  $N$  个神经子组成的联想网络, 通过调节神经子之间的突触使得信息易于储存和恢复. 通常突触服从一定的记忆规则使得系统的状态在自身动力学演变过程中成为固定吸引子.

联想记忆的最简单一类系统——Hopfield 模型<sup>[2]</sup>, 是由  $N$  个 Ising 二分变量自旋的神经子组成的联想网络, 其系统状态由所有自旋变量  $\mathbf{S} \equiv \{S_1, S_2, \dots, S_N\}$  的构形来描述. 若  $P (= \alpha N)$  个独立“记忆”模式  $\{\epsilon_i^\mu\} (i=1, \dots, N, \mu=1, \dots, P)$  按照 Hebb 规则<sup>[10]</sup> 被编码入突触矩阵  $J_{ij}$

$$J_{ij} = \frac{1}{N} \sum_{\mu} \epsilon_i^{\mu} \epsilon_j^{\mu} \quad (1)$$

中, 这里  $\epsilon_i^{\mu}$  分别以等概率取  $+1$  或  $-1$ , 于是, 这些独立模式为储存在网络中的“记忆”态. 在一定的初始状态下, 当系统自身动力学演变到与某一特定记忆模式(如模式1)关联的定态构形时, 该记忆模式的恢复得以实现. 由于 Hopfield 模型中突触矩阵  $J_{ij}$  的对称性 ( $J_{ij} = J_{ji}$ ), 平衡态统计力学的自旋玻璃理论已成功用来研究这一网络的恢复特性<sup>[3,4]</sup>. Amit 等人应用复制(Replica)方法得到了 Hopfield 模型中最大的储存容量  $\alpha$  为  $\alpha_c = 0.14$ . 为了改善储存恢复性能, 许多作者试图修改 Hebb 记忆规则<sup>[11,12]</sup> 或引入神经激活阈使得存储容量  $\alpha$  超过临界存储  $\alpha_c$  时联想网络仍有恢复信息能力. 业已表明神经激活阈值(类似于磁性系统中的外加磁场)的存在能大大改善网络的“记忆”恢复功能.

在文献<sup>[13]</sup>中, 我们应用动力学方法考虑了无规高斯噪声和反对称突触对神经网络

\* 国家自然科学基金资助的课题.

恢复特性的影响,发现无规涨落和反对称性都缩小了恢复态的区域. 本文将避免冗长的压制方法,使用一种新的微扰方法,讨论具有不同概率分布的双峰无规神经阈对网络恢复特性的影响,以揭示恢复增强的潜在能力.

## 二、模型与方法

Hopfield 模型的哈密顿量表示为

$$H = - \sum_{(ij)} J_{ij} S_i S_j - \sum_i h_i^{\text{ext}} S_i, \quad (2)$$

式中  $J_{ij}$  代表最近邻神经子之间的相互作用,  $h_i^{\text{ext}}$  为无规神经阈值. 与磁学中的铁磁态对应,“记忆”恢复态可由一组决定恢复态序参数  $m_i$  的迭代方程

$$m_i = \tanh\beta \left( \sum_j J_{ij} m_j + h_i^{\text{ext}} \right) = \tanh\beta h_i, \quad i = 1, \dots, N \quad (3)$$

的固定点来描述,固定点要求与某一“记忆”模式 ( $\mu$ ) 有较大的交叠  $q^\mu$ , 这里交叠  $q^\mu$  定义为

$$q^\mu = \frac{1}{N} \sum_i \epsilon_i^\mu m_i, \quad \mu = 1, \dots, \alpha N. \quad (4)$$

对于 Hebb 矩阵,作用在神经子  $i$  上的“外场”  $h_i$  可用  $q^\mu$  表示成

$$h_i = \sum_j J_{ij} m_j + h_i^{\text{ext}} = \sum_\mu \epsilon_i^\mu q^\mu + h_i^{\text{ext}}. \quad (5)$$

上式给出关于交叠  $q^\mu$  的  $\alpha N$  个固定点方程

$$q^\mu = \frac{1}{N} \sum_i \omega_i^\mu \tanh\beta(h_i), \quad \mu = 1, \dots, \alpha N, \quad (6)$$

即

$$\begin{aligned} q^\mu &= \frac{1}{N} \sum_i \tanh\beta(\epsilon_i^\mu h_i) = \frac{1}{N} \sum_i \tanh\beta(H_i^\mu) \\ &= \langle \tanh\beta(H^\mu) \rangle_{\text{av}}, \end{aligned} \quad (7)$$

式中  $H_i^\mu = \epsilon_i^\mu h_i$ ,  $\langle \dots \rangle_{\text{av}}$  表示对指标  $i$  的平均. 在  $N \rightarrow \infty$  极限下,平均可看作对局域场分布  $P(H^\mu)$  的积分

$$q^\mu = \int dH^\mu P(H^\mu) \tanh\beta(H^\mu). \quad (8)$$

为方便起见,将考虑仅与某一特定“记忆”模式(如模式1)交叠的恢复态,即  $q^1 \sim 1, q^\mu \sim O(N^{-\frac{1}{2}})$  ( $\mu > 1$ ), 于是模式1的局域场

$$H_i^1 = \epsilon_i^1 \sum_\nu \epsilon_i^\nu q^\nu + \epsilon_i^1 h_i^{\text{ext}} \quad (9)$$

可以分成两部分

$$H_i^1 = (q^1 + \epsilon_i^1 h_i^{\text{ext}}) + \epsilon_i^1 \sum_{\nu > 1} \epsilon_i^\nu q^\nu. \quad (10)$$

上式等号右边第二项可看作平均为零,分差  $\sigma^2$  的高斯噪声项,这里  $\sigma^2$  为

$$\sigma^2 = \sum_{\nu > 1} (q^\nu)^2. \quad (11)$$

根据(5),(6)两式,  $q^\mu$  可表示为

$$\begin{aligned}
 q^{\nu} &= \frac{1}{N} \sum_i \epsilon_i^{\nu} \tanh \beta \left[ \sum_{\rho} \epsilon_i^{\rho} q^{\rho} + h_i^{\text{ext}} \right] \\
 &= \frac{1}{N} \sum_i \epsilon_i^{\nu} \tanh \beta \left[ \sum_{\rho \neq \nu} \epsilon_i^{\rho} q^{\rho} + h_i^{\text{ext}} + \epsilon_i^{\nu} q^{\nu} \right], \quad (12)
 \end{aligned}$$

式中单一项( $\rho=\nu$ )与其它求和项( $\rho \neq \nu$ )相比很小,可看作微扰项,通过展开(12)式到 $q^{\nu}$ 的一阶,得到

$$\begin{aligned}
 q^{\nu} &= \frac{1}{N} \sum_i \epsilon_i^{\nu} \left[ \tanh \beta \left( \sum_{\rho \neq \nu} \epsilon_i^{\rho} q^{\rho} + h_i^{\text{ext}} \right) + \epsilon_i^{\nu} q^{\nu} \right. \\
 &\quad \left. \times \tanh' \beta \left( \sum_{\rho \neq \nu} \epsilon_i^{\rho} q^{\rho} + h_i^{\text{ext}} \right) \right]. \quad (13)
 \end{aligned}$$

定义量 $C$

$$\begin{aligned}
 C &= \langle \tanh' \beta(h_i) \rangle_{\text{av}} = \frac{1}{N} \sum_i \tanh' \beta(h_i) \\
 &= \int P(H^1) dH^1 \tanh' \beta(H^1), \quad (14)
 \end{aligned}$$

式中

$$P(H^1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{(H^1 - q^1 - h^{\text{ext}})^2}{2\sigma^2} \right]. \quad (15)$$

于是(13)式为

$$q^{\nu}(1 - C) = \frac{1}{N} \sum_i \epsilon_i^{\nu} \tanh \beta \left( \sum_{\rho \neq \nu} \epsilon_i^{\rho} q^{\rho} + h_i^{\text{ext}} \right). \quad (16)$$

上式等号两边平方后得到

$$[(1 - C)q^{\nu}]^2 = \frac{1}{N^2} \sum_i \tanh^2 \beta \left( \sum_{\rho \neq \nu} \epsilon_i^{\rho} q^{\rho} + h_i^{\text{ext}} \right) \approx \frac{1}{N^2} \sum_i \tanh^2 \beta(h_i). \quad (17)$$

定义 Edwards-Anderson 序参数 $q$

$$\begin{aligned}
 q &= \langle \tanh^2 \beta(h) \rangle = \frac{1}{N} \sum_i \tanh^2 \beta(h_i) \\
 &= \int P(H^1) dH^1 \tanh^2 \beta(H^1), \quad (18)
 \end{aligned}$$

于是(17)式可写为

$$(q^{\nu})^2 = q/N(1 - C)^2. \quad (19)$$

由(11),(19)两式,局域场分布(15)式的分差 $\sigma^2$ 为

$$\sigma^2 = \alpha q / (1 - C)^2. \quad (20)$$

于是, $q^1, C$ 和 $q$ 的自洽方程可由下列方程给出:

$$q^1 = \int dH^1 P(H^1) \tanh \beta(H^1), \quad (21)$$

$$C = \int dH^1 P(H^1) \tanh^1 \beta(H^1), \quad (22)$$

$$q = \int dH^1 P(H^1) \tanh^2 \beta(H^1). \quad (23)$$

我们选择与模式1线性耦合的无规神经阈值的概率分布 $P(h_i^{\text{ext}})$ 为

$$P(h_i^{\text{ext}}) = P\delta(h_i^{\text{ext}} - h\epsilon_i^1) + (1 - P)\delta(h_i^{\text{ext}} + h\epsilon_i^1). \quad (24)$$

把(24)式代入方程组(21)–(23),并引入变量  $y = (H^1 - q^1 - h^{\text{ext}})/\sigma$  后,得到热噪声  $T \rightarrow 0$  时的自洽方程

$$q^1 = \text{Perf}\left(\frac{q^1 + h}{\sqrt{2\alpha\gamma}}\right) + (1 - P)\text{erf}\left(\frac{q^1 - h}{\sqrt{2\alpha\gamma}}\right), \quad (25)$$

$$C = \left(\frac{2}{\pi\alpha\gamma}\right)^{\frac{1}{2}} \{P\exp\left[-\frac{(q^1 + h)^2}{2\alpha\gamma}\right] + (1 - P)\exp\left[-\frac{(q^1 - h)^2}{2\alpha\gamma}\right]\}, \quad (26)$$

$$\gamma = (1 - C)^{-2}. \quad (27)$$

不难发现,上述方程中若  $h=0$ ,得到了 Amit 等人的结果<sup>[3]</sup>,另一方面,当  $P=1$  和  $P=\frac{1}{2}$  时,可分别得到外场平行于模式<sup>[4]</sup>和尖峰混乱状态下的方程.

### 三、结 果

通过数值求解(25)–(27)式,可以获得不同  $P$  值下的“记忆”恢复特性.图1给出  $\alpha=0.16$  时交叠  $q^1$  随激活阈强度  $h$  的变化关系.由于临界值  $\alpha_c=0.14 < 0.16$ ,对于所有  $P$  值,  $q^1(h=0)=0$ .随着  $h$  的增加,系统进入有序的恢复态( $q^1 \sim 1$ ).当  $h$  达到某些值时,交叠  $q^1$  有一最大值.如果  $h \rightarrow \infty$ ,则交叠  $q^1 \rightarrow (2P-1)$ .因此,  $h$  的中间值特别有利于“记忆”恢复.另一方面,当  $P$  较小时,没有一级相变发生;当  $P$  较大时,可观察到一级相变.  $P$  值越大,对“记忆”恢复越有利.另外,从方程(25)–(27)可看出,随着  $\alpha$  的增加,形成一级相变的  $P, h$  值也将增加,而“记忆”恢复的能力将逐渐衰退.

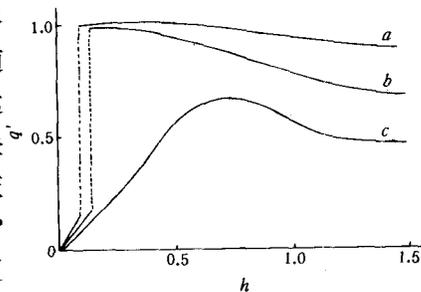


图1 网络交叠  $q^1$  随激活阈强度  $h$  的变化 (取  $\alpha=0.16$ ) 曲线  $a(P=0.95)$ ; 曲线  $b(P=0.85)$ ; 曲线  $c(P=0.75)$

### 四、结 论

至此,我们应用微扰方法研究了神经激活阈存在下的吸引子神经网络系统.目前日益引起理论工作者兴趣的 Hopfield 模型,多数的研究都忽视了神经激活阈对“记忆”恢复的影响,而这点在大脑模型中是不容忽略的.本文中提出神经激活阈取非对等概率的双峰分布将更接近于大脑模型的实际情形.解析和数值结果表明神经激活阈强度取适当的值,将有利于储存“记忆”的恢复.

[1] W. A. Little, *Math. Biosci.*, **19**(1974),101.

[2] J. J. Hopfield, *Proe. Natl. Acad. Sci, USA*, **79**(1982),2554.

- [ 3 ] D. J. Amit, H. Gutfreund and H. Sompolinsky, *Phys. Rev.*, **A32**(1985), 1007;  
*Phys. Rev. Lett.*, **55**(1985),1530.
- [ 4 ] D. J. Amit, H. Gutfreund and H. Sompolinsky, *Ann. Phys. (NY)*, **173**(1987),30.
- [ 5 ] H. Sompolinsky, A. Crisanti and H. J. Sommers, *Phys. Rev. Lett.*, **61**(1988),259.
- [ 6 ] E. Barkai, D. Hansel and I. Kanter, *Phys. Rev. Lett.*, **65**(1990),2312.
- [ 7 ] M. Y. Choi, *Phys. Rev. Lett.*, **61**(1988),2809.
- [ 8 ] T. L. H. Watkin and D. Sherrington, *J. Phys. A*, **24**(1991),5427.
- [ 9 ] H. Eissfeller and M. Opper, *Phys. Rev. Lett.*, **68**(1992),2094.
- [10] D. O. Hebb, *The Organization of Behavior* (Wiley, New York, 1949).
- [11] J. F. Fontanari and R. Koberk, *J. Phys. A*, **21**(1988),h253.
- [12] I. Kanter and H. Sompolinsky, *Phys. Rev.*, **A35**(1987),380.
- [13] Yu-qiang Ma and Chang-de Gong, *Phys. Rev.*, **B46**(1992),3436.

## RETRIEVAL PROPERTIES OF HOPFIELD NEURAL NETWORK MODELS

MA YU-QIANG    ZHANG YUE-MING    GONG CHANG-DE

*Department of Physics, Nanjing University, Nanjing 210008;*

*CCAST (World Laboratory), Beijing 100080*

<sup>†</sup>*Department of Physics, Nanjing Medical College, Nanjing 210029*

(Received 5 October 1992)

### ABSTRACT

In this paper, we propose a bimodal distribution of random neuronal activity threshold with different probabilities, to consider the influences on the retrieval properties of neural network. It is shown that the system successfully retrieves information even if the number of stored patterns exceeds the critical value of the pure Hopfield model.

**PACC:** 7510H; 8710; 6460; 0550