




## 第四章 自组织神经网络

XI'AN JIAOTONG UNIVERSITY

- 引言
- 横向抑制网络MAXNET
- ART1的工作原理
- 竞争学习机制和自稳学习机制
- ART2神经网络
- 小结

2004-8-8      《神经网络导论》——自组织神经网络      4-1




## § 4.1 引言

XI'AN JIAOTONG UNIVERSITY

前两章讨论的前向网络和反馈网络是应用最广泛，研究最多的两类型网络。但在某些方面也存在一些共同的问题，这两种网络在构成网络时未能充分借鉴人脑工作的特点。对前向网络来说，这些缺点和不足主要体现在以下几个方面：

- (1) 只适用于平稳的环境，也即要求观察矢量具有统计平稳性，但现实世界不是或不完全是这样的；
- (2) 学习是有监督的学习，不能“自学”，学习与工作状态截然分开，不能边学习边工作（要完成新工作，必须重新学习）；
- (3) 学习过程中，所有的权系数都要调整，而且调整的步幅不能太大，否则会产生振荡，步幅小学习速度又会非常慢；
- (4) 网络分类的类别数要事先确定，超出这个类别数，学习不可能给出正确结果；

2004-8-8      《神经网络导论》——自组织神经网络      4-2



## § 4.1 引言

XI'AN JIAOTONG UNIVERSITY

- (5) 有可能陷入局部极小点，从而给出错误的结果；
- (6) 误差准则是固定的，不能随环境的变化而进行相应的调整或改变。

相比之下，人脑的优越性就极其明显了。人脑神经系统既能够牢固地记住所学得的各种知识又能适应各种复杂多变的环境，能够通过“自学”来认识未学习过的新事物并解决不熟悉的新问题。人脑神经系统有如下主要特点：

- (1) 人脑的学习方式是自主的，人可以在一个复杂的、非平稳的、有“干扰”的环境中学习。人脑在大多数情况下是无监督的、无导师的“自学”，而且学习和工作分不开，是“半工半读”的；
- (2) 人脑的工作方式和信息存储方式具有明显的自组织的特点。人在与环境交互作用的过程中逐渐建立起大脑的信息处理能力（信息的存储、检索等），这种建立方式是自组织的，外界不可能对脑神经系统进行干预；
- (3) 人脑的学习和记忆能力既有很大的刚性，又有很大的弹性和可塑性；

2004-8-8      《神经网络导论》——自组织神经网络      4-3



## § 4.1 引言

XI'AN JIAOTONG UNIVERSITY

- (4) 人对外的输入信号响应时有两种渠道：“由底向上”（bottom-up）和“由顶向下”（top-down）。其中“由顶向下”涉及到人的集中注意力和忽略能力。人脑在一定的环境下会估计到可能出现的情况，预期会遇到、听到或看到的各种内容。

自适应谐振理论ART(Adaptive Resonance Theory)研究的出发点就是充分借鉴人脑工作时的特点。

本章我们主要讨论两个ART模型：ART1和ART2。其中ART1是针对二值信号的，ART2是针对任意模拟信号的。作为ART的预备内容，下一节先讨论一种横向抑制网络MAXNET。

2004-8-8      《神经网络导论》——自组织神经网络      4-4

XIAN JIAOTONG UNIVERSITY

### § 4.2 横向抑制网络MAXNET

一、汉明距离

两个二值向量（向量的各分量只取两种值0或1，或者-1和1）之间的汉明距离定义为两个向量的对应分量不相同的数目。设  $\mathbf{X}, \mathbf{U}_j \in \{-1, 1\}^N$ ，则它们之间的汉明距离可以表示如下：

$$\text{Hamming distance}(\mathbf{X}, \mathbf{U}_j) = N - \sum_{i=1}^N f_h(x_i u_{ji})$$

其中  $N$  为向量的维数， $x_i, u_{ji}$  分别表示向量  $\mathbf{X}, \mathbf{U}_j$  的第  $i$  个分量（元素）。

二、MAXNET

对于一个各分量为二值的模式，MAXNET分类器将按照输入模式  $\mathbf{X}$  与某个类别的典型样本  $\mathbf{U}_j$  之间的汉明距离确定这一模式是否属于  $C_j$  类。即确定  $\mathbf{X} \in C_j$  的充要条件为：

$$\text{Hamming distance}(\mathbf{X}, \mathbf{U}_j) < \text{Hamming distance}(\mathbf{X}, \mathbf{U}_k)$$

$$k = 1, 2, \dots, M \text{ 且 } k \neq j$$

2004-8-8 《神经网络导论》——自组织神经网络 4-5

XIAN JIAOTONG UNIVERSITY

### § 4.2 横向抑制网络MAXNET

其中， $\mathbf{U}_j$  是第  $j$  类的典型样本（样板），也具有二值特征。

实际上， $\{-1, 1\}^N$  的汉明距离定义式中的第二项  $\sum_{i=1}^N f_h(x_i u_{ji})$  就表示了两个向量中对应分量相同的个数。此项值越大，表示两个向量越相似，而越小表示差异越大，如果去掉硬限幅函数，余下的和式即为两个向量的内积，也有这样的性质，两者有等价作用。MAXNET就利用了这一点。

MAXNET的网络结构如右图所示。

图4-1Z 横向抑制网络MAXNET的结构

2004-8-8 《神经网络导论》——自组织神经网络 4-6

XIAN JIAOTONG UNIVERSITY

### § 4.2 横向抑制网络MAXNET

在横向抑制子网络中，设  $t_{jk}$  是从第  $j$  个节点到第  $k$  个节点的连接权，我们可以取：

$$t_{jk} = \begin{cases} 1 & j = k \\ -\epsilon & j \neq k \end{cases} \dots \dots \dots \text{EQ4.1Z}$$

$$\epsilon < \frac{1}{M}, k, j = 1, 2, \dots, M$$

子网络中各节点的迭代公式如下： 初值  $y_j(0) = \sum_i x_i u_{ji}$

$$y_j(t+1) = f_i \left( y_j(t) - \epsilon \sum_{\substack{k=1 \\ k \neq j}}^M y_k(t) \right) \quad j = 1, 2, \dots, M \quad \dots \dots \text{EQ4.2Z}$$

其中

$$f_i(\alpha) = \begin{cases} C\alpha & \alpha > 0 \\ 0 & \alpha \leq 0 \end{cases} \quad C \geq 1$$

2004-8-8 《神经网络导论》——自组织神经网络 4-7

XIAN JIAOTONG UNIVERSITY

### § 4.2 横向抑制网络MAXNET

这说明，在MAXNET子网中，每个节点都试图保持自己的值而抑制其它节点，从而通过竞争选出最大值。

需要讨论的几点问题：


(1) 仿真时，可能会出现两个初始值  $y_i(0) = y_j(0)$  且是最大值的情况，这时就会产生竞争，竞争的结果是所有输出都变为零或者两个最大值同时达到最大值。解决办法是加一个小的扰动：

$$y_j(t+1) = f_i \left( y_j(t) - \epsilon \sum_{k \neq j} y_k(t) + \epsilon_i \right), \quad \epsilon_i \leq \frac{i}{N^3}$$

(2)  $f_i(\alpha)$  的函数形式，如果  $\alpha > 0$  时采用线性函数且  $C=1$ ，那么会发生收敛较慢并且很难达到一个饱和值。建议采用：

$$f_i(\alpha) = \begin{cases} A & \alpha \geq A/C \\ C\alpha & 0 < \alpha < A/C \\ 0 & \alpha \leq 0 \end{cases}$$

2004-8-8 《神经网络导论》——自组织神经网络 4-8



XI'AN JIAOTONG UNIVERSITY

### § 4.2 横向抑制网络MAXNET


(3) 如果采用所建议的  $f_i(\alpha)$ ，那么  $A$  值应选得足够大，使竞争开始时不会有任何值达到  $A$  值的水平；对  $C$  值的选取，一方面要考虑  $A$  值，不能有两个值同时进入饱和区，在此条件下大一些可使收敛加快。

(4) 自反馈系数应该大于等于1，这样可以确保收敛较快。可以看出，该网络具有“中心激励、侧向抑制”的功能，这种现象称为中心分离效应（On center, off surround）。

(5) MAXNET 在使用过程中，时间十分重要。汉明网络运算时间短，而 MAXNET 需要的处理时间较长，必须在结构上保证，只有 MAXNET 收敛并已经为新的输入做好准备后，才能允许汉明网络开始输出。

(6) 对于模式分类，这种汉明网络明显优于 Hopfield 网络。当码元误差随机独立时，汉明网络是一种最佳的最小误差分类器。同时，汉明网络连接权少。当模式矢量维数为100，类别数为10时，汉明网络需1000个连接权，而 Hopfield 网络需要10000个。

2004-8-8
《神经网络导论》——自组织神经网络
4-9



XI'AN JIAOTONG UNIVERSITY

### § 4.3 ART1的工作原理

#### 一、聚类的含义


由于可以将  $N$  维模式样本看作  $N$  维特征空间中的点，所以有理由认为：类别相同或某些特征相似的模式在  $N$  维空间中比较靠近。

前面讨论的分类，主要就是根据这一特点进行的。所谓的学习（训练）就是通过对已知类别的样本的学习训练出类别之间的分界线，从而形成判决准则。这是有监督的学习。

但是，有时我们需要发现模式集合在模式空间中的分布情况，假如这些模式在空间中以某种明显的集群状分布，确定各集群的位置与分布的方法是有意义的。

所谓聚类，就是要按照某种度量准则，自动地发现、划分模式类别，并将观测样本分类。下面要讨论的 ART1 就可以用于聚类。

2004-8-8
《神经网络导论》——自组织神经网络
4-10



XI'AN JIAOTONG UNIVERSITY

### § 4.3 ART1的工作原理

#### 二、ART1神经网络

ART1 的网络结构有多种形式，这里讨论的只是其中的一种。网络结构如右图所示。输入应该为二值矢量模式， $b_{ji}$  是“由底向上”的连接权， $t_{ji}$  是“自顶向下”的连接权， $i=1, 2, \dots, N$  而  $j=1, 2, \dots, M$ 。

网络算法如下：

- ① 输入任一模式

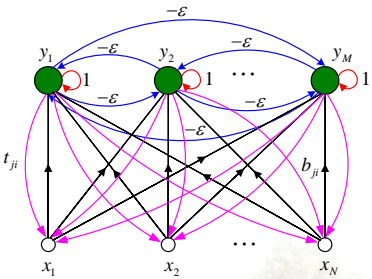

$$\mathbf{X} = \{x_1, x_2, \dots, x_N\} \quad x_i \in \{-1, 1\}, i=1, 2, \dots, N$$


图4-2Z ART1神经网络结构示意图

2004-8-8
《神经网络导论》——自组织神经网络
4-11




XI'AN JIAOTONG UNIVERSITY

### § 4.3 ART1的工作原理

- ② 进行“自下而上”（bottom-up processing）的处理，求出
 
$$y_j = \sum_{i=1}^N b_{ji} x_i \quad j=1, 2, \dots, M$$
- ③ 运用 MAXNET 找出最大值节点  $y_j$ 。
- ④ 执行“自上而下”（top-down processing）的处理，验证  $\mathbf{X}$  是否确实属于第  $j$  类。计算
 
$$\frac{\sum_{i=1}^N t_{ji} x_i}{\sum_i |x_i|} > \rho$$
 如果
 
$$\frac{\sum_{i=1}^N t_{ji} x_i}{\sum_i |x_i|} > \rho$$
 则  $\mathbf{X}$  属于第  $j$  类，转第 ⑤ 步；否则转第 ⑥ 步。其中  $\rho$  是警戒参量。

2004-8-8
《神经网络导论》——自组织神经网络
4-12



XI'AN JIAOTONG UNIVERSITY

### § 4.3 ART1的工作原理

⑤ 对于第  $j$  类, 更新  $b_{ji}$  和  $t_{ji}$ ,  $i=1, 2, \dots, N$ . 返回第①步。

⑥  $\mathbf{X}$  不属于第  $j$  类, 所以要抑制该节点, 然后返回第②步, 执行另一个聚类中心的处理过程。


ART1中的初值和更新计算公式如下:

$$t_{ji}(0) = 1 \quad b_{ji}(0) = \frac{1}{1+N}$$

$$t_{ji}(n+1) = t_{ji}(n) - \alpha t_{ji}(n) \operatorname{sgn}[-t_{ji}x_i]$$

$$b_{ji}(n+1) = \frac{t_{ji}(n+1)}{0.5 + \sum_{i=1}^N |t_{ji}(n)x_i|} \quad \begin{matrix} i = 1, 2, \dots, N \\ j = 1, 2, \dots, M \end{matrix}$$

2004-8-8      《神经网络导论》——自组织神经网络      4-13



XI'AN JIAOTONG UNIVERSITY

### § 4.3 ART1的工作原理

三、关于ART1网络聚类原理的讨论

1. 连接权  $\{b_{ji}\}$  和  $\{t_{ji}\}$  本质上都相当于第  $j$  类的典型 (标准) 样本, 但功能不同,  $\{b_{ji}\}$  的功能是判断  $\mathbf{X}$  与哪一类最相似; 而  $\{t_{ji}\}$  的功能是验证  $\mathbf{X}$  是否正确属于这一类。
2. 
$$\left. \begin{matrix} \sum_i t_{ji}x_i & \leftarrow & \text{绝对相似度} \\ \sum_i |x_i| & \leftarrow & \text{元素的总数} \end{matrix} \right\} \text{相对相似度}$$
3. 警戒参数  $\rho$  的取值与聚类半径有关:
  - $\rho$  大, 聚类半径小, 对相似的程度要求高, 结果类别很多。
  - $\rho$  小, 聚类半径大, 对相似的程度要求低, 结果类别很少。

2004-8-8      《神经网络导论》——自组织神经网络      4-14




XI'AN JIAOTONG UNIVERSITY

### § 4.3 ART1的工作原理

4. 这里的特殊模型只适用于  $\{1, -1\}$  二值矢量模式的聚类, 而且只是其中的一种。对于连续取值的矢量模式聚类, 要用ART2型网络。
5. 对未启用的输出端要加标志位, 在第一次启用时不用警戒参数  $\rho$ 。

2004-8-8      《神经网络导论》——自组织神经网络      4-15



XI'AN JIAOTONG UNIVERSITY

### § 4.4 竞争学习机制和自稳学习机制

一、竞争学习机制

竞争学习的原理可借助图4-1来描述。

设输入  $N$  维观察矢量

$$\mathbf{X} = [x_0, x_1, \dots, x_{N-1}]$$

输出为  $M$  维矢量

$$\mathbf{Y} = [y_0, y_1, \dots, y_{M-1}]$$

其中,

$$y_i \in \{0, 1\}$$

网络共分为三层,  $F_2, F_1$  分别称为上、下短期记忆层, 记为STM (Short Time Memory), 两者之间为长期记忆层LTM (Long Time Memory)。

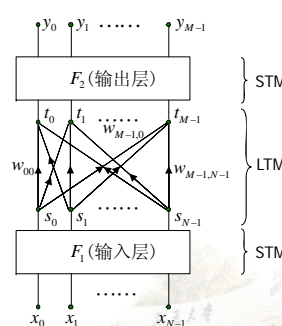


图4-1 竞争学习机制示意图

2004-8-8      《神经网络导论》——自组织神经网络      4-16

XIAN JIAOTONG UNIVERSITY

### § 4.4 竞争学习机制和自稳学习机制

各层的功能如下：

- $F_1$ 层 (STM)：使观测矢量规格化，规格化矢量  $\mathbf{S}$  的各分量表示如下

$$s_j = \frac{x_j}{\left[\sum_{i=0}^{N-1} x_i^2\right]^{1/2}}, \quad j = 0, 1, \dots, N-1$$

- $F_2$ 和  $F_1$ 之间的中间层 (LTM)：匹配运算

$$\mathbf{S} \in R^N \rightarrow \mathbf{T} \in R^M$$

$$t_i = \sum_{j=0}^{N-1} w_{ij}s_j, \quad i = 0, 1, \dots, M-1$$

权重系数  $w_{ij}$  的变化相对于  $F_2, F_1$  而言要慢很多，每输入一次观察矢量，它们只作微小的变化。

2004-8-8      《神经网络导论》——自组织神经网络      4-17

XIAN JIAOTONG UNIVERSITY

### § 4.4 竞争学习机制和自稳学习机制

- $F_2$ 层 (STM)：竞争抉择运算

选出矢量  $\mathbf{T}$  中的最大分量，即：

若  $t_i = \max\{t_i\}$ ，则  $y_i = \begin{cases} 1, & i = I \\ 0, & i \neq I \end{cases}$

此网络用于分类时的学习策略如下：

首先随机设置权重系数的初值  $w_{ij}(0)$ ，然后依次送入观察矢量  $\mathbf{X}(k)$ ，其中  $k = 0, 1, 2, \dots$  是时序变量，每送入一个观察矢量，就可以算出相应的输出矢量  $\mathbf{Y}(k)$ ，最后按下式调整权重系数。

$$w_{ij}(k+1) = w_{ij}(k) + \alpha \{s_j(k) - w_{ij}(k)\} y_i(k), \quad \begin{matrix} j = 0, 1, \dots, N-1 \\ i = 0, 1, \dots, M-1 \end{matrix}$$

其中， $\alpha$  是步幅值（小正数）。显然，由于  $y_i(k), i = 0, 1, \dots, M-1$  中只有一项为1，假设  $y_i(k) = 1$ ，那么只有与此非零项对应的权重系数才做调整。

2004-8-8      《神经网络导论》——自组织神经网络      4-18

XIAN JIAOTONG UNIVERSITY

### § 4.4 竞争学习机制和自稳学习机制

即只改变与竞争获胜者有关的各个权重系数。

$$w_{ij}(k+1) = w_{ij}(k) + \alpha \{s_j(k) - w_{ij}(k)\}, \quad j = 0, 1, \dots, N-1$$

设  $\mathbf{W}_i = [w_{i0}, w_{i1}, \dots, w_{i,N-1}]$ ，权重系数的调整策略是使  $w_{ij}$  与规格化矢量  $\mathbf{S}(k)$  的各分量趋于一致，即通过调整使  $\mathbf{W}_i \rightarrow \mathbf{S}(k)$ 。因为  $|\mathbf{S}(k)| = 1$ ，所以  $|\mathbf{W}_i| \rightarrow 1$ 。

由于这种算法只有竞争得胜者的对应权重才有机会进行学习，所以称之为**竞争学习算法**。如果通过学习，不同的观察矢量集合都找到了各自的获胜分量，那么根据获胜的编号就能自然地完成对它们进行分类。这种通过竞争学习完成分类的过程也称为聚类、分割、编码、正交化等。

如果每个类别中的观察矢量很“紧凑”，不同类别的观察矢量又相距很远，那么竞争学习机制可以趋于稳定并收到很好的聚类效果。

2004-8-8      《神经网络导论》——自组织神经网络      4-19

XIAN JIAOTONG UNIVERSITY


### § 4.4 竞争学习机制和自稳学习机制

但也会出现如下的情况，如果在相隔较远的两个时间点上两次输入同一观察矢量，而在其间插入若干其它观察矢量，那么第二次分类的结果会出现与第一次分类结果不一致的情况。这说明第一次分类后新学习得到的记忆内容冲掉了原有的学习记忆内容，从而导致了第二次分类错误。所以只依靠竞争学习机制不能保证记忆的牢固性、可靠性和稳定性。

因为这是聚类，还有一个聚类半径的问题。竞争机制只保证在已开启的几类中，输入样本离哪一类最近，并没有考虑验证输入是否确实属于这一类（前向分类有可能陷入局部极小点）或者是否需要新开辟一类。所以仅依靠竞争学习机制，获胜者就改变权重系数，会使以前这一类的模式在改变权重系数后再输入就不属于这一类。这是一种不应该发生的不稳定现象。

我们可以在竞争学习机制的基础上引入“由顶向下”的自稳机制来解决这一矛盾。

2004-8-8      《神经网络导论》——自组织神经网络      4-20



XI'AN JIAOTONG UNIVERSITY

### § 4.4 竞争学习机制和自稳学习机制

二、自稳学习机制的工作原理——自上而下的处理

1. 竞争选择原理不变。设竞争获胜者为  $y_i(k)$ 。
2. 从输出矢量  $\mathbf{Y}(k)$  产生一个由顶向下的矢量  $\mathbf{Z}(k)$ :
 
$$\mathbf{Z}(k) = [z_0(k), z_1(k), \dots, z_{N-1}(k)]$$


其中,  $z_j(k) = \sum_{i=0}^{M-1} w'_{ji}(k) y_i(k) \quad j = 0, 1, \dots, N-1$

$w'_{ji}$  是“由顶向下”的权系数, 是在时序  $k$  以前的各次学习中“记忆”下来的。记忆算法是: 当  $I$  端为获胜端时, 使  $w'_{ji}$  与当时的规格化矢量  $\mathbf{S}$  的各分量  $s_j$  趋于一致。即  $\mathbf{W}'_j$  应与  $\mathbf{S}$  相似。

由于所有的  $y_i(k)$  中只有  $y_l(k)$  为1, 其它均为0, 所以

$$z_j(k) = w'_{jl}(k), \quad j = 0, 1, \dots, N-1$$

2004-8-8      《神经网络导论》——自组织神经网络      4-21



XI'AN JIAOTONG UNIVERSITY

### § 4.4 竞争学习机制和自稳学习机制

即  $\mathbf{Z}(k) = \mathbf{W}'_l$ , 所以  $\mathbf{Z}(k)$  与  $\mathbf{S}$  相似。

现在, 我们通过比较  $\mathbf{S}(k)$  与  $\mathbf{Z}(k)$  的相似度, 就可以判断本次输入的规格化矢量  $\mathbf{S}(k)$  与以前  $I$  输出端获胜时的规格化输入矢量  $\mathbf{S}$  是否相似。


相似度衡量方法:

$$\eta = \frac{\langle \mathbf{S}(k), \mathbf{Z}(k) \rangle}{\|\mathbf{S}(k)\| \|\mathbf{Z}(k)\|} \quad \text{或} \quad \|\mathbf{S}(k) - \mathbf{Z}(k)\|$$

根据相似度  $\eta$  的不同, 网络采取不同的学习策略。

- (a) 如果相似度很高,  $\eta > \rho$ , 则转3。
- (b) 如果相似度不够高,  $\eta < \rho$ , 设有  $K$  个已启用的类别, 那么就抑制  $I$  端后回到1重新竞争, 在  $K-1$  个输出端中选出优胜者。直到选出的优胜者的相似程度足够高 ( $\eta > \rho$ ) 才转到3。

2004-8-8      《神经网络导论》——自组织神经网络      4-22



XI'AN JIAOTONG UNIVERSITY

### § 4.4 竞争学习机制和自稳学习机制

(c) 如果所有的  $K$  个已启用端的  $\eta$  均小于  $\rho$  时, 就需要开辟新的类别。这时令  $K = K + 1$ , 当  $K + 1 = M$  时, 表示网络容量已满, 就不能再开辟新类别了。


3. 进入本步的条件:  $\eta > \rho$  或开辟新类别。对获胜端或新开辟端对应的权系数进行调整, 设此端标号为  $L$ , 调整公式如下:

$$\begin{cases} w_{lj}(k+1) = w_{lj}(k) + \alpha \{s_j(k) - w_{lj}(k)\} \\ w'_{lj}(k+1) = w'_{lj}(k) + \alpha \{s_j(k) - w'_{lj}(k)\} \end{cases} \quad j = 0, 1, \dots, N-1$$

其中  $\alpha$  为步幅, 是一个小的正数。

从以上算法可以看到, 只有当新输入矢量与已存入记忆中的某个老矢量足够相似时, 两者才能互相融合, 才对有关的权系数进行调整, 从而使长期记忆得以改变。这就产生了一种所谓的“自适应 (Adaptive Resonance) 谐振”状

2004-8-8      《神经网络导论》——自组织神经网络      4-23



XI'AN JIAOTONG UNIVERSITY

### § 4.4 竞争学习机制和自稳学习机制

态, 这就是ART的来源。


$F_1$ 、 $F_2$  层对每一输入均发生变化, 变化快, 称为STM; 而LTM层的权系数只是在共振时才得以调整。这和人脑的记忆过程类似。

三、按ART构成的神经网络有如下特点

1. “半工半读”, 即学习和工作是分不开的, 可以是非平稳输入;
2. 学习是自治和自组织的, 无需导师指导;
3. 学习过程受由顶向下传送的模式矢量指导, 可以形成“集中注意”的状态, 可以有选择的学习, 把注意力集中于某些特定的内容。权系数的调整只涉及少量系数, 与前向神经网络相比有更高的学习效率;
4. 可以避免陷入局部极小点。

2004-8-8      《神经网络导论》——自组织神经网络      4-24





XI'AN JIAOTONG UNIVERSITY


### § 4.5 ART2神经网络

ART2神经网络可以完成对模拟输入矢量进行聚类（分类），采用了典型的竞争学习和自稳机制的原理。系统仍由  $F_1$  和  $F_2$  两个STM层及两者之间的LTM层组成。其中， $F_2$  层的功能也是选择竞争优胜者，与ART1没有区别。而  $F_1$  的设计比较特殊，也就是说  $F_1$  的结构和原理是我们讨论的核心。

$F_1$  的设计要达到下面的要求：

- (1) 要用神经元完成两模拟矢量之间的相似度计算，从而完成输入矢量和由顶向下的矢量之间的匹配计算。
- (2) 当两者相似度足够高时，将两者的共同特征“融合”后记入LTM，每送入一次输入矢量样本，就可以将其全部重要特征存入长期记忆之中，无需反复多次学习。还有一个需要解决的问题是，在LTM系数的学习过程中，不能由于这些系数的变化造成优胜者换位或错误重置现象，即学习应该是自稳的。当两者相似度不够高时， $F_1$  应向  $F_2$  发出正确的重置信号。

2004-8-8
《神经网络导论》——自组织神经网络
4-25



XI'AN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

- (3) 当启用一个新的输出端时，不能发生反复重置。
- (4) 当警戒参数选得较高时，两个矢量之间的非本质微小差异也会导致重置的发生，应该消除这种现象。方法如下：A. 各处理层中保持矢量的规格化；B. 压低背景噪声；C. 突出重要特征。


为了满足上述要求，ART2中的  $F_1$  层采用了三层结构，如图4-5所示，其中包括了正反馈、规格化和非线性变换等。为了避免混淆，把  $F_1$  和  $F_2$  分别称为“特征表示场”和“类别表示场”。要说明的是，这一结构及其中参数的确定是通过大量实验的筛选而获得的。此外，图4-5的结构只是可供选择的几种结构中的一种。

下面结合图4-5来说明ART2的工作原理。

#### 一、场中第一个处理单元的描述

输入 维模拟观察矢量  $\mathbf{X} = [x_0, x_1, \dots, x_{N-1}]$ 。在  $F_1$  中有相应的  $N$  个处理单元。

2004-8-8
《神经网络导论》——自组织神经网络
4-26



XI'AN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

理单元，每个单元都有上、中、下三层，图中只画出了第一个处理单元的结构。其中每一层都包括两种功能不同的神经元，一种是小空心圆，另一种是大实心圆。

1. 空心圆
  - 空心圆有两种输入：
    - a. 兴奋激励，用空心箭头表示；
    - b. 抑制激励，用实心箭头表示。
  - 设神经元的输出活动电位为  $V$ ，所有兴奋激励的总和为  $J^+$ ，所有抑制激励的总和为  $J^-$ 。

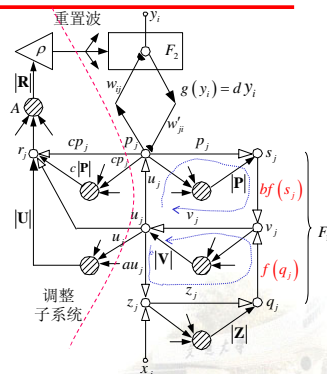



图4-5 ART2神经网络结构示意图

2004-8-8
《神经网络导论》——自组织神经网络
4-27



XI'AN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

神经生理学研究表明， $V$  满足如下的微分方程：

$$\varepsilon \frac{dV}{dt} = -AV + (1 - BV)J^+ - (C + DV)J^-$$

其中， $0 < \varepsilon \ll A \ll 1$   $B \ll 1$   $C \ll D \leq 1$

稳定时的近似解为  $V = \frac{J^+}{A + DJ^-}$

又因为  $D$  接近于1，而  $A$  是一个远小于1的正实数，所以上式可近似为：


$$V = \frac{J^+}{e + J^-}$$

其中  $e$  为远小于1的正实数。

2. 实心圆（求输入矢量的模）
 

例如图4-5中与  $p_j$  相连的实心圆，其输入为  $p_0, p_1, \dots, p_{N-1}$ ，写成向量形式  $\mathbf{P} = [p_0, p_1, \dots, p_{N-1}]$ ，实心圆的输出就等于  $\mathbf{P}$  的模  $|\mathbf{P}|$ ，即

2004-8-8
《神经网络导论》——自组织神经网络
4-28



XI'AN JIAOTONG UNIVERSITY

## § 4.5 ART2神经网络

$$|\mathbf{P}| = \{p_0^2 + p_1^2 + \dots + p_{N-1}^2\}^{1/2}$$

实心圆输出的都是抑制激励。

3.  $F_1$ 场中各层的方程

A. 第一层（下层）：

$$z_j = x_j + au_j$$


$$q_j = \frac{z_j}{e + |\mathbf{Z}|} \quad \text{其中 } e \text{ 是很小的正实数}$$

B. 第二层（中层）：

$$v_j = f(q_j) + bf(s_j)$$

$$u_j = \frac{v_j}{e + |\mathbf{V}|}$$

2004-8-8      《神经网络导论》——自组织神经网络      4-29



XI'AN JIAOTONG UNIVERSITY

## § 4.5 ART2神经网络

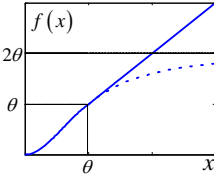
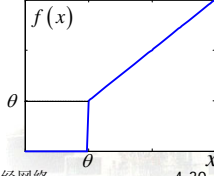
其中,  $f(x)$  是非线性变换, 可采取如下两种形式

①  $f(x) = \begin{cases} \frac{2\theta x^2}{(x^2 + \theta^2)}, & 0 \leq x \leq \theta \\ x, & x > \theta \end{cases}$


②  $f(x) = \begin{cases} 0, & 0 \leq x \leq \theta \\ x, & x > \theta \end{cases}$

C. 第三层（上层）

$$s_j = \frac{p_j}{e + |\mathbf{P}|}$$

$$p_j = u_j + \sum_{i=0}^{M-1} g(y_i) w'_{ji} = u_j + dw'_{ji}$$



2004-8-8      《神经网络导论》——自组织神经网络      4-30



XI'AN JIAOTONG UNIVERSITY

## § 4.5 ART2神经网络

$F_1$ 的下层和中层构成一个闭合的正反馈回路, 其中神经元  $z_j$  接受输入信号  $x_j$ , 而神经元  $v_j$  接受上层送来的信号  $bf(s_j)$ ;  $F_1$ 的中层和上层构成另一个闭合正反馈回路, 神经元  $p_j$  既接受中层送来的信号  $u_j$  又接受  $F_2$ 场送来的信号。这两个正反馈回路各包含两次规格化运算和一次非线性变换。非线性变换的作用是对小幅度信号进行抑制。其中的参数  $a$ 、 $b$  和  $\theta$  是通过大量的实验确定的。

二、 $F_2$ 场中所完成运算的描述


$F_2$ 场的输入矢量  $\mathbf{T} = [t_0, t_1, \dots, t_{M-1}]$

$$\text{其中, } t_i = \sum_{j=0}^{N-1} w_{ij} p_j \quad i = 0, 1, \dots, M-1$$

$w_{ij}$  为  $F_1$ 至  $F_2$ 的LTM权重系数,  $p_j$  由  $F_1$ 的上层给出。按照竞争学习机制, 输出矢量  $\mathbf{Y} = [y_0, y_1, \dots, y_{M-1}]$  由下式确定。

如果  $t_i = \max\{t_i, i = 0, 1, \dots, M-1\}$ , 则

2004-8-8      《神经网络导论》——自组织神经网络      4-31



XI'AN JIAOTONG UNIVERSITY

## § 4.5 ART2神经网络

$$y_i = \begin{cases} 1 & i = I \\ 0 & i \neq I \end{cases}$$

我们选择  $g(y_i) = dy_i$ , 其中  $d$  是一个实常数。所以

$$p_j = u_j + dw'_{ji}, \quad j = 0, 1, \dots, N-1$$

当  $F_2$ 场尚未受到激励时, 所有  $y_j = 0$ , 此时  $p_j = u_j$ 。

三、LTM权重系数  $w_{ij}$  和  $w'_{ji}$  的学习

当由顶向下的矢量和输入观察矢量相似度足够高或开辟一个新的输出端时, 就进入LTM权重系数学习阶段。设竞争获胜端为  $y_j$ , 权重系数调整公式如下:

由顶向下:

$$w'_{ji}(k+1) = w'_{ji}(k) + g[y_i(k)] \{p_j(k) - w'_{ji}(k)\}, \quad \begin{matrix} i = 0, 1, \dots, M-1 \\ j = 0, 1, \dots, N-1 \end{matrix}$$

2004-8-8      《神经网络导论》——自组织神经网络      4-32



XIAN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

注意到只有  $g(y_i) = d$ ，其它均为零，可以得到下式

$$\begin{cases} w'_{ji}(k+1) = w'_{ji}(k) + d\{p_j(k) - w'_{ji}(k)\}, & j = 0, 1, \dots, N-1 \\ w'_{ji}(k+1) = w'_{ji}(k), & \text{当 } i \neq I \end{cases}$$

同理，可以得到由底向上的权值调整公式

$$\begin{cases} w_{ij}(k+1) = w_{ij}(k) + d\{p_j(k) - w_{ij}(k)\}, & j = 0, 1, \dots, N-1 \\ w_{ij}(k+1) = w_{ij}(k), & \text{当 } i \neq I \end{cases}$$

又因为  $p_j = u_j + dw'_{ji}$ ，将其代入由顶向下的公式可得：

$$w'_{ji}(k+1) = w'_{ji}(k) + d(1-d)\left\{\frac{u_j(k)}{1-d} - w'_{ji}(k)\right\}, \quad j = 0, 1, \dots, N-1$$

2004-8-8      《神经网络导论》——自组织神经网络      4-33

XIAN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

至此，我们已经讨论了  $F_1$  场的结构、 $F_1$  场和  $F_2$  场之间的相互作用以及两个场之间的LTM系数  $w_{ij}$  和  $w'_{ji}$  的学习调整等问题，并给出了相应的计算公式。下面，我们从物理定性的概念出发，讨论  $\mathbf{P}$ 、 $\mathbf{U}$ 、 $\mathbf{W}_i$ 、 $\mathbf{W}'_i$  等矢量间的关系。

1.  $F_1$  场有输入矢量  $\mathbf{X}$  的作用而  $F_2$  场尚未受到激励时

$$\begin{array}{c} \mathbf{P} = \mathbf{U} \longrightarrow \mathbf{S} = \frac{\mathbf{P}}{|\mathbf{P}|} \longrightarrow \\ \mathbf{U} = \frac{\mathbf{V}}{|\mathbf{V}|} \longleftarrow \mathbf{V} = bf(\mathbf{S}) + f(\mathbf{Q}) \\ \mathbf{Z} = \mathbf{X} + a\mathbf{U} \longrightarrow \mathbf{Q} = \frac{\mathbf{Z}}{|\mathbf{Z}|} \longrightarrow \end{array}$$

$F_1$  场中各矢量满足如下关系：

2004-8-8      《神经网络导论》——自组织神经网络      4-34

XIAN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

①  $\mathbf{S} = \mathbf{P} = \mathbf{U} = \mathbf{V}/|\mathbf{V}| \quad |\mathbf{S}| = |\mathbf{P}| = |\mathbf{U}| = 1$

②  $\mathbf{Z} = \mathbf{X} + a\mathbf{U} \quad \mathbf{Q} = \mathbf{Z}/|\mathbf{Z}| \quad |\mathbf{Q}| = 1$

③  $\mathbf{U}$  反映了  $\mathbf{X}$  的重要特征，抑制了它的基底噪声。

2. 当  $F_2$  受到激励并选出了一个已被占用过的输出端  $I$  时

$$\begin{array}{c} \mathbf{P} = \mathbf{U} + d\mathbf{W}'_I \longrightarrow \mathbf{S} = \frac{\mathbf{P}}{|\mathbf{P}|} \longrightarrow \\ \mathbf{U} = \frac{\mathbf{V}}{|\mathbf{V}|} \longleftarrow \mathbf{V} = bf(\mathbf{S}) + f(\mathbf{Q}) \\ \mathbf{Z} = \mathbf{X} + a\mathbf{U} \longrightarrow \mathbf{Q} = \frac{\mathbf{Z}}{|\mathbf{Z}|} \longrightarrow \end{array}$$

可以看到，在上、下两个正反馈回路的作用下， $\mathbf{U}$  融合了  $\mathbf{X}$  和  $\mathbf{W}'_I$  的重要特征，并抑制了  $\mathbf{X}$  的基底噪声。

2004-8-8      《神经网络导论》——自组织神经网络      4-35

XIAN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

值得注意的是，当  $\mathbf{X}$  和  $\mathbf{W}'_I$  足够相似时， $\mathbf{P}$  和  $\mathbf{U}$  也足够相似；反过来，如果  $\mathbf{X}$  和  $\mathbf{W}'_I$  不相似时， $\mathbf{P}$  和  $\mathbf{U}$  的差异也比较大。

因此，我们可以选择用  $\mathbf{P}$  和  $\mathbf{U}$  来比较相似度。这样比较有以下优点：

① 在启用新端时， $\mathbf{P}$  和  $\mathbf{U}$  保持一致，而  $\mathbf{X}$  和  $\mathbf{W}'_I$  差异极大造成重置。

②  $\mathbf{P}$  和  $\mathbf{U}$  比较的是去除了基底噪声的主要特征，不会因为非重要特征而引起重置。

四、调整子系统的工作原理、 $w_{ij}$  和  $w'_{ji}$  初值的选择、参数  $c$  和  $d$  的选择

图4-5的左侧是ART2网络的调整子系统，它主要用来完成相似度的比较。神经元  $r_j$  (空心圆) 的输出可以表示如下：

$$r_j = \frac{u_j + cp_j}{e + |\mathbf{U}| + c|\mathbf{P}|}, \quad j = 0, 1, \dots, N-1 \quad e \ll 1 \quad \dots \dots \dots \text{EQ4.29}$$

2004-8-8      《神经网络导论》——自组织神经网络      4-36

XI'AN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

由于  $e \ll 1$ ，下面的计算中将其略去。神经元A的输出是  $\mathbf{R}$  的模，即

$$|\mathbf{R}| = \left[ \sum_{j=0}^{N-1} r_j^2 \right]^{1/2} \dots \dots \dots \text{EQ4.30}$$

显然， $\mathbf{U}$  和  $\mathbf{P}$  的相似度越高， $|\mathbf{R}|$  越接近于1。这样，我们可以用  $|\mathbf{R}|$  和警戒参数  $\rho$  进行比较， $0 < \rho < 1$ ， $|\mathbf{R}| > \rho$  无需重置，反之要对  $F_2$  场重置。为了探讨启用新端或系数调整过程中是否会发生错误重置，需要研究  $|\mathbf{R}|$  在各种情况下是如何变化的。EQ4.29、EQ4.30可以改写如下的向量形式。

$$\mathbf{R} = \frac{\mathbf{U} + c\mathbf{P}}{1 + c|\mathbf{P}|} \dots \dots \dots \text{EQ4.31}$$

$$|\mathbf{R}| = \frac{[1 + 2c|\mathbf{P}|\cos(\mathbf{U}, \mathbf{P}) + c^2|\mathbf{P}|^2]^{1/2}}{1 + c|\mathbf{P}|} \dots \dots \dots \text{EQ4.32}$$

More...

2004-8-8      《神经网络导论》——自组织神经网络      4-37

XI'AN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

其中  $\cos(\mathbf{U}, \mathbf{P})$  表示  $\mathbf{U}$  和  $\mathbf{P}$  夹角的余数值，可以用下式计算：

$$\cos(\mathbf{U}, \mathbf{P}) = \frac{\langle \mathbf{U}, \mathbf{P} \rangle}{|\mathbf{P}||\mathbf{U}|} \Big|_{\mathbf{P}=\mathbf{U}+d\mathbf{W}_i'} = \frac{|\mathbf{U}|^2 + d\langle \mathbf{U}, \mathbf{W}_i' \rangle}{|\mathbf{P}||\mathbf{U}|} = \frac{1 + d\langle \mathbf{U}, \mathbf{W}_i' \rangle}{|\mathbf{P}|}$$

又因为：

$$\cos(\mathbf{U}, \mathbf{W}_i') = \frac{\langle \mathbf{U}, \mathbf{W}_i' \rangle}{|\mathbf{U}||\mathbf{W}_i'|} \Rightarrow \langle \mathbf{U}, \mathbf{W}_i' \rangle = |\mathbf{W}_i'| \cos(\mathbf{U}, \mathbf{W}_i')$$

所以：

$$|\mathbf{P}|\cos(\mathbf{U}, \mathbf{P}) = 1 + d|\mathbf{W}_i'| \cos(\mathbf{U}, \mathbf{W}_i') \dots \dots \dots \text{EQ4.33}$$

又由于  $\mathbf{P} = \mathbf{U} + d\mathbf{W}_i'$ ，所以

$$|\mathbf{P}| = [1 + 2d|\mathbf{W}_i'| \cos(\mathbf{U}, \mathbf{W}_i') + d^2|\mathbf{W}_i'|^2]^{1/2} \dots \dots \dots \text{EQ4.34}$$

把EQ4.33和EQ4.34代入EQ4.32，可得

2004-8-8      《神经网络导论》——自组织神经网络      4-38

XI'AN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

$$|\mathbf{R}| = \frac{[(1+c)^2 + 2(1+c)|cd\mathbf{W}_i'| \cos(\mathbf{U}, \mathbf{W}_i') + |cd\mathbf{W}_i'|^2]^{1/2}}{1 + [c^2 + 2c|cd\mathbf{W}_i'| \cos(\mathbf{U}, \mathbf{W}_i') + |cd\mathbf{W}_i'|^2]^{1/2}} \dots \dots \dots \text{EQ4.35}$$

$$|\mathbf{R}| = f(|cd\mathbf{W}_i'|) \Big|_{\cos(\mathbf{U}, \mathbf{W}_i')}$$

图4-6是  $c=0.1, d=0.9$  时的一簇曲线。下面我们讨论系数调整过程中可能出现的几种情况以及各参数的选择。

① 如果  $\mathbf{U}$  和  $\mathbf{W}_i'$  完全匹配，则  $\cos(\mathbf{U}, \mathbf{W}_i') = 1$ ，由图4-6，此时  $|\mathbf{R}| \equiv 1$ ，这样对任何警戒参数均不会有重置发生。

图4-6  $|\mathbf{R}|$  相对于  $|cd\mathbf{W}_i'|$  的变化曲线

2004-8-8      《神经网络导论》——自组织神经网络      4-39

XI'AN JIAOTONG UNIVERSITY


### § 4.5 ART2神经网络

② 新开辟一个输出端时，由于  $\mathbf{W}_i'$  的初值为零，所以  $|cd\mathbf{W}_i'| = 0$  由图4-6可见  $|\mathbf{R}| = 1$ ，因而不会产生重置。在学习过程中， $\mathbf{U}$  和  $\mathbf{W}_i'$  始终保持一致，因此不可能引起重置。这就保证了在启用新端时可以进行可靠的学习。

③ 若  $I$  是一个已开辟的输出端，但  $\mathbf{U}$  和  $\mathbf{W}_i'$  不完全匹配。如果匹配程度低，使得  $|\mathbf{R}| < \rho$ ，那么就向  $F_2$  场重置信号。如果匹配程度高，使得  $|\mathbf{R}| > \rho$ ，那么就要调整LTM层的系数，此时应保证在整个调整过程中不发生重置。

若参数  $c$  选得较小，那么  $|\mathbf{R}|$  曲线簇的最小值都在  $|cd\mathbf{W}_i'| = 1$  附近。根据由顶向下的权调整公式， $\mathbf{W}_i'$  在学习过程中将趋于  $\mathbf{U}/(1-d)$  并在结束时达到它，由于  $|\mathbf{U}| = 1$  所以学习开始和结束时  $|\mathbf{W}_i'|$  的值皆等于  $1/(1-d)$  而在学习过程中， $|\mathbf{W}_i'|$  下降然后升高。这样在学习过程中  $|cd\mathbf{W}_i'|$  将  $cd/(1-d)$  下降然后再回升到  $cd/(1-d)$  到，只要我们选择参数使得  $d < cd/(1-d) < 1$  那么就能保证在学习过程中不会发生重置。

2004-8-8      《神经网络导论》——自组织神经网络      4-40



XI'AN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

此外，在满足  $cd/(1-d) < 1$  条件的同时，应使其尽可能接近于1，因为在1附近， $|R|$  对  $U$  和  $W_i$  的失配最敏感。所以图中选  $c=0.1, d=0.9$  从而使得  $cd/(1-d)=0.9$ 。

④  $W_i$  的初值随机选择，但是不能为零，否则此端永远不会被占用。此外还应满足  $|W_i| < 1/(1-d)$ 。因为，如果  $|W_i| > 1/(1-d)$ ，那么在学习过程中  $|W_i|$  将从大于  $1/(1-d)$  的方向趋于  $1/(1-d)$ ，也就是说， $|W_i|$  在学习过程中逐渐减小，而在此过程中，其它  $W_i$  并没有改变，这样可能导致优胜者换位。


#### 五、参数 $\theta$ 、 $a$ 和 $b$ 的选择

首先  $\theta$  的选择和噪声有关，一般选  $\theta = 1/\sqrt{N}$ 。实验表明，参数  $a$  和  $b$  选得很小且  $\rho$  也很小时，分类结果不稳定。如果选

$$\theta = 1/\sqrt{N}, a = b = 10, c = 0.1, d = 0.9$$

那么  $\rho$  选各种数值的情况下都能得到稳定的分类结果。

2004-8-8                      《神经网络导论》——自组织神经网络                      4-41



XI'AN JIAOTONG UNIVERSITY


### § 4.5 ART2神经网络

#### 六、模拟实验



警戒参量较大时的聚类结果                      警戒参量较小时的聚类结果

2004-8-8                      《神经网络导论》——自组织神经网络                      4-42




XI'AN JIAOTONG UNIVERSITY

### § 4.5 ART2神经网络

#### 七、对ART2网络的评价

1. 采用了  $P$ 、 $U$  衡量相似度，实现了记忆的稳定性 and 弹性之间的折衷。
2. 对于属于已有类别的输入，系统可以很快检索；对于新类别输入，系统会反复搜索和重置，直至开辟一个新端。
3. 可以“快速学习”，输入  $X$  只呈现一次。
4. 规格化运算的实现比较简单。
5. 计算有相当一部分都是局部的，无需很多全局运算。
6.  $F$  场采用了独特的具有正反馈和非线性变换的三层结构，强化了特征分量，抑制了基底噪声。
7. 解决了匹配和重置的矛盾。
8. 只要参数选择适当就可以保证不产生误重置和优胜者换位。
9. 系统的学习是无监督的自组织学习，“半工半读”。

2004-8-8                      《神经网络导论》——自组织神经网络                      4-43



XI'AN JIAOTONG UNIVERSITY

### § 4.6 本章小结

本章首先在引言中指出了前向网络和Hopfield网络的缺点和不足，并列出了人脑神经系统的主要特点，由此引出了ART网络；接下来，讨论了ART网络的预备内容——横向抑制网络MAXNET；在此基础上介绍了聚类的含义并讨论了适用于二值输入矢量的ART1神经网络的原理；然后，我们讨论了ART网络中的竞争学习机制和自稳学习机制；最后，详细讨论了可以完成对模拟输入矢量进行聚类的ART2神经网络。

2004-8-8                      《神经网络导论》——自组织神经网络                      4-44

XIAN JIAOTONG UNIVERSITY

### 附录A. MAXNET算法的解释(1)

$$y_j(t+1) = f_i \left( y_j(t) - \varepsilon \sum_{\substack{k=1 \\ k \neq j}}^M y_k(t) \right) \quad j=1, 2, \dots, M$$

$$f_i(\alpha) = \begin{cases} C\alpha & \alpha > 0 \\ 0 & \alpha \leq 0 \end{cases} \quad C \geq 1$$

2005-6-2      《神经网络导论》——第四章附录      4A-45

XIAN JIAOTONG UNIVERSITY

### 附录A. MAXNET算法的解释(2)

2005-6-2      《神经网络导论》——第四章附录      4A-46

XIAN JIAOTONG UNIVERSITY

### 附录B. EQ4. 30到EQ4. 32的推导

$$|\mathbf{R}| = \left[ \sum_{j=0}^{N-1} r_j^2 \right]^{\frac{1}{2}} = \frac{\left[ \sum_{j=0}^{N-1} (u_j + cp_j)^2 \right]^{\frac{1}{2}}}{|\mathbf{U}| + c|\mathbf{P}|} = \frac{\left[ \sum_{j=0}^{N-1} (u_j^2 + 2cu_j p_j + c^2 p_j^2) \right]^{\frac{1}{2}}}{1 + c|\mathbf{P}|}$$

$$= \frac{\left[ \sum_{j=0}^{N-1} u_j^2 + 2c \sum_{j=0}^{N-1} u_j p_j + c^2 \sum_{j=0}^{N-1} p_j^2 \right]^{\frac{1}{2}}}{1 + c|\mathbf{P}|} = \frac{\left[ |\mathbf{U}|^2 + c^2 |\mathbf{P}|^2 + 2c \sum_{j=0}^{N-1} u_j p_j \right]^{\frac{1}{2}}}{1 + c|\mathbf{P}|}$$

$$\because \sum_{j=0}^{N-1} u_j p_j = \langle \mathbf{U}, \mathbf{P} \rangle = |\mathbf{P}| |\mathbf{U}| \cos(\mathbf{U}, \mathbf{P})$$

$$\therefore |\mathbf{R}| = \frac{\left[ 1 + c^2 |\mathbf{P}|^2 + 2c |\mathbf{P}| \cos(\mathbf{U}, \mathbf{P}) \right]^{\frac{1}{2}}}{1 + c|\mathbf{P}|}$$

2005-6-2      《神经网络导论》——第四章附录      4A-47