第4期
2004年4月

电 子 学 报
ACTA ELECTRONICA SINICA

Vol.32 No.4
Apr. 2004

扩展交替投影神经网络
——具备联想记忆功能的充分必要条件

王金根[1],龚沈光[2],陈世福[1]
(1.南京大学计算机软件新技术国家重点实验室,江苏南京 210093;2.海军工程大学,湖北武汉 430033)

摘 要: 对交替投影神经网络(APNN)的连接权矩阵进行修改,将其应用范围从实数域拓展到复数域,从而得到一种新的神经网络——扩展交替投影神经网络(Extended Alternating Projection Neural Networks).对EAPNN网络进行深入研究后,给出了网络稳态值的通用数学表达式,并从表达式中推出了网络具备联想记忆功能的充分必要条件.最后设计仿真实验对文中的理论分析结果进行了验证.

关键词: 交替投影;神经网络;联想记忆;收敛条件
中图分类号: Q811.1;TN911.72;O235 **文献标识码**: A **文章编号**: 0372-2112(2004)04-0596-05

Sufficient and Necessary Condition of the Extended Alternating Projection Neural Network Configured as a Content Addressable Memory

WANG Jin-gen[1], GONG Shen-guang[2], CHEN Shi-fu[1]
(1. *State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, Jiangsu* 210093, *China*;
2. *Naval University of Engineering, Wuhan, Hubei* 430033, *China*)

Abstract: The paper extends the original Alternating Projection Neural Network(APNN) and proposes an Extended Alternating Projection Neural Network (EAPNN) which functions in the field of complex numbers. An improved weight-learning approach, which permits linear dependence of complex patterns, has been presented. A general mathematical expression of the EAPNN steady-state solution has been obtained. From the general mathematical expression we have derived the sufficient and necessary condition of the EAPNN configured as a content addressable memory. In addition, simulation experiments have been designed to verify the theoretical analyses in the paper. Finally it is pointed out that the EAPNN has been applied to signal processing such as band-limited signal extrapolation, notch filter and weak-signal separation.

Key words: alternating projection; neural networks; content addressable memory; condition of convergence

1 引言

交替投影神经网络(Alternating Projection Neural Networks)是由美国Washington大学Marks Ⅱ等人提出的.它是利用凸集投影的概念,在向量空间中建立的一种神经网络模型.APNN具有如下特征:

(1)APNN可完成两个或多个约束集合之间的交替投影,与传统的能量尺度法[2~8]不同.能量尺度法仅能给出能量的下界及表明随着每一次迭代能量不断减小,而APNN的稳态解对同步或异步工作模式很容易保证.

(2)许多齐次神经网络[6,9~12]并不具有良好的标度性:当神经元数目加倍时,存储容量小于2倍[13,14];而APNN存储容量的数目大约等于输入和隐层神经元的数目.

(3)与BP网络不同[15,16],它的每次训练迭代都以准确形式存储,不存在计算不准确性,网络对训练数据的分类是完备的.这一特性在实际中是很有用的,如当大量准确数据库使用的记忆推理[17].传统的并行结构不适于求解这类问题,因为处理器数目的增加将增加计算时间.

(4)APNN权值学习过程和联想过程都极为简单,易于并行计算和VLSI实现.

Marks Ⅱ等人在讨论APNN网时,首先假定了网络用于学习的N个库模式是线性独立的,于是连接权矩阵可以用这样的方法获得,即$T=F(F^TF)^{-1}F^T$;但是实际中N个库模式极可能不是线性独立的,这样F^TF可能是奇异阵或者严重病态,前述获取T的方法不再有效.为此,需要研究一种新的权值学习方法,这种方法允许N个库模式不是线性独立的.

由于APNN收敛条件及其证明是建立在N个库模式线性独立这一前提条件基础上的,而新方法却考虑了线性相关情

收稿日期:2001-11-09;修回日期:2003-12-24

况,那么这种新方法获得的连接权矩阵具体是什么?还能不能保证网络稳定?若能,网络具备联想记忆功能的充分必要条件是什么?这是值得研究的一个问题.另外,围绕 APNN 的一切工作都是在实数域内展开的,为了使 APNN 能够广泛用于信号处理领域,需将其应用范围从实数域拓展到复数域.能不能直接拓展?这又是一个值得研究的问题.

针对上述问题,本文对原 APNN 网进行扩展,提出一种扩展交替投影神经网络(Extended Alternating Projection Neural Network).EAPNN 网与原 APNN 网相比,具有如下改进之处:(1) EAPNN 网的作用范围为复数域而非实数域,这样更利于其在信号处理方面的应用;(2)给出的改进权值学习方法不要求网络用于学习的复模式一定要线性独立;(3)得到了网络稳态值的数学表达式,并从表达式中推出了网络具备联想记忆功能的充分必要条件,而不仅仅是充分条件.

以上的改进工作经过了严格的数学证明,并得到了实例的验证.

2 扩展交替投影神经网络(EAPNN)

EAPNN 是一种全互连网络,该网络的作用域为复数域,其网络结构如图 1 所示.假设它由 L 个神经元组成.则任意神经元 i 与神经元 j 之间是双向连接的.其连接权值 t_{ij} 与 t_{ji} 相等.依据神经元状态是固定的还是浮动的将其称为箝位神经元或浮动神经元.任意浮动神经元 i 在 m 时刻的状态 $s_i(m)=\sum_{p=1}^{L}t_{pi}s_p(m-1)$,任意箝位神经元 j 在 m 时刻的状态 $s_j(m)=s_j(m-1)=\cdots=s_j(0)$ 等于初始状态.

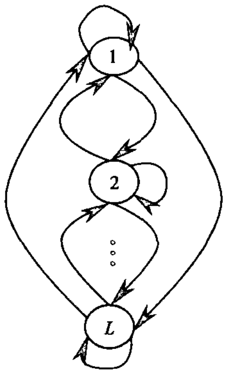


图 1　扩展交替投影神经网络结构示意图

这里我们给出它的一种实用权值学习方法如下:

(1)将初始连接权矩阵 $\boldsymbol{T}$ 置 $\boldsymbol{0}$, $i\leftarrow 1$.

(2) $\boldsymbol{\varepsilon}_i=(\boldsymbol{I}-\boldsymbol{T})\boldsymbol{f}_i$,此处 $\boldsymbol{I}$ 为 $L\times L$ 单位阵,$\boldsymbol{f}_i$ 为网络用于学习的复模式.

(3)若 $\|\boldsymbol{\varepsilon}_i\|=0$,则认为 $\boldsymbol{f}_i$ 已经记忆在网络中,转到(4);否则,$\boldsymbol{T}\leftarrow\boldsymbol{T}+\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^H/\boldsymbol{\varepsilon}_i^H\boldsymbol{\varepsilon}_i$.

(4) $i\leftarrow i+1$,如果 $i>N$ 则结束,否则转(2).

$\boldsymbol{T}$ 中的每个元素 t_{ij} 即为神经元 i 与神经元 j 的连接权值.若令矩阵 $\boldsymbol{F}=[\boldsymbol{f}_1\,\boldsymbol{f}_2\cdots\boldsymbol{f}_N]$,则采用上述权值学习方法得到的连接权矩阵 $\boldsymbol{T}=\boldsymbol{F}\boldsymbol{F}^+$,其中 $\boldsymbol{F}^+$ 为 $\boldsymbol{F}$ 的广义逆.

当权值学习完之后,首先确定哪些神经元是箝位神经元,哪些神经元是浮动神经元.不失一般性,假定 EAPNN 网中前 P 个神经元为箝位神经元,余下 $Q=L-P$ 个神经元为浮动神经元,在 $m=0$ 时刻给箝位神经元输入初值.若设网络在 m 时刻的状态向量为 $\boldsymbol{S}(m)=[s_1(m)\ s_2(m)\ \cdots\ s_L(m)]^{\mathrm{T}}$,其中箝位神经元在 m 时刻的状态组成的向量为 $\boldsymbol{S}^P(m)$,浮动神经元在 m 时刻的状态组成的向量为 $\boldsymbol{S}^Q(m)$,则 $m+1$ 时刻网络的状态向量为

$$\boldsymbol{S}(m+1)=\eta\boldsymbol{T}\boldsymbol{S}(m)=\begin{bmatrix}\boldsymbol{S}^P(m+1)\\ \boldsymbol{S}^Q(m+1)\end{bmatrix}=\eta\begin{bmatrix}T_2 & T_1\\ T_3 & T_4\end{bmatrix}\cdot\begin{bmatrix}\boldsymbol{S}^P(m)\\ \boldsymbol{S}^Q(m)\end{bmatrix}$$

$$=\begin{bmatrix}\boldsymbol{S}^P(m)\\ T_3\boldsymbol{S}^P(m)+T_4\boldsymbol{S}^Q(m)\end{bmatrix}$$

其中 $\boldsymbol{T}$ 为连接权矩阵,算子 η 将把由于 $\boldsymbol{T}$ 的作用而改变了状态的箝位神经元的状态置为原态.这样 EAPNN 的联想过程就是 $\boldsymbol{T}$ 算子和 η 算子交替作用的结果.

3 扩展交替投影神经网络的稳态值

既然在文献[18]中已经证明了网络是稳定的,那么它的稳态值是什么呢?这就是下面要证明的定理 1.在证明该定理之前,首先需证明一个引理.

引理 1　设 $\boldsymbol{F}$ 的前 P 行组成的矩阵为 $\boldsymbol{F}_P$,余下 Q 行组成的矩阵为 $\boldsymbol{F}_Q$,则对 $\forall\ \boldsymbol{Y}\in C^N$,有

(1) $(\boldsymbol{I}-\boldsymbol{T}_4)\cdot(\boldsymbol{F}_Q-\boldsymbol{F}_P^+\boldsymbol{F}_P)\boldsymbol{Y}=0$;

(2) $\boldsymbol{T}_3^H\cdot(\boldsymbol{F}_Q-\boldsymbol{F}_Q\boldsymbol{F}_P^+\boldsymbol{F}_P)\boldsymbol{Y}=0$.

证明　设 $\mathrm{rank}(\boldsymbol{F})=r$,对 $\boldsymbol{F}$ 进行满秩分解有:

$$\boldsymbol{F}=\begin{bmatrix}\boldsymbol{C}_1\\ \boldsymbol{C}_2\end{bmatrix}\cdot\boldsymbol{D}=\boldsymbol{C}_0\cdot\boldsymbol{D}$$

其中 $\boldsymbol{C}_0\in C^{L\times r}$, $\boldsymbol{D}\in C^{r\times N}$, $\boldsymbol{C}_1\in C^{P\times r}$, $\boldsymbol{C}_2\in C^{Q\times r}$, $\boldsymbol{F}_P=\boldsymbol{C}_1\cdot\boldsymbol{D}$, $\boldsymbol{F}_Q=\boldsymbol{C}_2\cdot\boldsymbol{D}$.

假定 $\mathrm{rank}(\boldsymbol{C}_1)=r_1$,对 $\boldsymbol{C}_1$ 进行满秩分解有:

$\boldsymbol{C}_1=\boldsymbol{D}_0\boldsymbol{D}_1$,其中 $\boldsymbol{D}_0\in C^{P\times r_1}$, $\boldsymbol{D}_1\in C^{r_1\times r}$.

这样就有 $\boldsymbol{F}_P=\boldsymbol{D}_0\boldsymbol{D}_1\boldsymbol{D}$

令 $\boldsymbol{D}_1\boldsymbol{D}=\boldsymbol{\Delta}$,则易证

$\mathrm{rank}(\Delta)=\mathrm{rank}(\boldsymbol{D}_1\boldsymbol{D})=r_1$,即 Δ 为行满秩阵.

$$\boldsymbol{T}=\boldsymbol{F}\boldsymbol{F}^+=\boldsymbol{C}_0\boldsymbol{D}\cdot\boldsymbol{D}^+\boldsymbol{C}_0^+$$

$$=\boldsymbol{C}_0\boldsymbol{D}\boldsymbol{D}^H(\boldsymbol{D}\boldsymbol{D}^H)^{-1}(\boldsymbol{C}_0^H\boldsymbol{C}_0)^{-1}\boldsymbol{C}_0^H$$

$$=\begin{bmatrix}\boldsymbol{C}_1\\ \boldsymbol{C}_2\end{bmatrix}\cdot(\boldsymbol{C}_2^H\boldsymbol{C}_2+\boldsymbol{C}_1^H\boldsymbol{C}_1)^{-1}\cdot[\boldsymbol{C}_1^H\ \boldsymbol{C}_2^H]$$

令 $\gamma=(\boldsymbol{C}_2^H\boldsymbol{C}_2+\boldsymbol{C}_1^H\boldsymbol{C}_1)^{-1}$,则

$$\boldsymbol{T}=\begin{bmatrix}\boldsymbol{C}_1\gamma\boldsymbol{C}_1^H & \boldsymbol{C}_1\gamma\boldsymbol{C}_2^H\\ \boldsymbol{C}_2\gamma\boldsymbol{C}_1^H & \boldsymbol{C}_2\gamma\boldsymbol{C}_2^H\end{bmatrix}$$

那么 $\boldsymbol{T}_4=\boldsymbol{C}_2\gamma\boldsymbol{C}_2^H$

$$\boldsymbol{T}_4\cdot(\boldsymbol{F}_Q-\boldsymbol{F}_Q\boldsymbol{F}_P^+\boldsymbol{F}_P)=\boldsymbol{C}_2\gamma\boldsymbol{C}_2^H\cdot\boldsymbol{C}_2\boldsymbol{D}\cdot(\boldsymbol{I}-\boldsymbol{\Delta}^+\boldsymbol{\Delta})$$

$$=\boldsymbol{C}_2\gamma(\boldsymbol{C}_2^H\boldsymbol{C}_2+\boldsymbol{C}_1^H\boldsymbol{C}_1-\boldsymbol{C}_1^H\boldsymbol{C}_1)\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{\Delta}^+\boldsymbol{\Delta})$$

$$=\boldsymbol{C}_2\gamma(\gamma^{-1}-\boldsymbol{C}_1^H\boldsymbol{C}_1)\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{\Delta}^+\boldsymbol{\Delta})$$

$$=\boldsymbol{C}_2\boldsymbol{D}(\boldsymbol{I}-\Delta^+\Delta)-\boldsymbol{C}_2\gamma\boldsymbol{C}_1^H\boldsymbol{D}_0\cdot\boldsymbol{\Delta}(\boldsymbol{I}-\boldsymbol{\Delta}^+\boldsymbol{\Delta})$$

$$=\boldsymbol{C}_2\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{\Delta}^+\boldsymbol{\Delta})=\boldsymbol{F}_Q-\boldsymbol{F}_Q\boldsymbol{F}_P^+\boldsymbol{F}_P$$

于是 $(\boldsymbol{I}-\boldsymbol{T}_4)\cdot(\boldsymbol{F}_Q-\boldsymbol{F}_Q\boldsymbol{F}_P^+\boldsymbol{F}_P)\boldsymbol{Y}=\boldsymbol{0}$　　(1)证毕.

$\boldsymbol{T}_3^H=\boldsymbol{C}_1\gamma\boldsymbol{C}_2^H$

$$\boldsymbol{T}_3^H\cdot(\boldsymbol{F}_Q-\boldsymbol{F}_Q\boldsymbol{F}_P^+\boldsymbol{F}_P)=\boldsymbol{C}_1\gamma\boldsymbol{C}_2^H\cdot\boldsymbol{C}_2\boldsymbol{D}\cdot(\boldsymbol{I}-\boldsymbol{\Delta}^+\boldsymbol{\Delta})$$

$$=\boldsymbol{C}_1\gamma(\gamma^{-1}-\boldsymbol{C}_1^H\boldsymbol{C}_1)\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{\Delta}^+\boldsymbol{\Delta})$$

$$=\boldsymbol{C}_1\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{\Delta}^+\boldsymbol{\Delta})-\boldsymbol{C}_1\gamma\boldsymbol{C}_1^H\boldsymbol{D}_0\cdot\boldsymbol{\Delta}(\boldsymbol{I}-\boldsymbol{\Delta}^+\boldsymbol{\Delta})$$

$$=\boldsymbol{F}_P-\boldsymbol{F}_P\boldsymbol{F}_P^+\boldsymbol{F}_P=\boldsymbol{0}$$

于是 $\boldsymbol{T}_3^H\cdot(\boldsymbol{F}_Q-\boldsymbol{F}_Q\boldsymbol{F}_P^+\boldsymbol{F}_P)\boldsymbol{Y}=\boldsymbol{0}$　　(2)证毕.

定理 1 设 EAPNN 网络初始状态 $s(0)$ 的前 P 个元素组成的向量为 $s^P(0)$，余下 Q 个元素组成的向量为 $s^Q(0)$，F 的前 P 行组成的矩阵为 F_P，余下 Q 行组成的矩阵为 F_Q，$A=R(F)$ 是所有库模式张成的子空间，B 是前 P 个元素为 $s^P(0)$ 的所有向量组成的集合，则 $s(\infty)$ 在 B 内的值为

$$s_B(\infty)=\begin{bmatrix} s^P(0) \\ F_Q(I-F_P^+F_P)(F_Q(I-F_P^+F_P))^+(s^Q(0)-F_QF_P^+s^P(0))+F_QF_P^+s^P(0)\end{bmatrix}$$

在 A 内的值为

$$s_A(\infty)=\begin{bmatrix} F_PF_P^+s^P(0) \\ F_Q(I-F_P^+F_P)(F_Q(I-F_P^+F_P))^+(s^Q(0)-F_QF_P^+s^P(0))+F_QF_P^+s^P(0)\end{bmatrix}$$

证明 由文献[18]中网络稳定性的证明过程可知，网络最终稳定状态在 A、B 内的值 $s_A(\infty)$ 和 $s_B(\infty)$ 存在且唯一．由于 $s_A(\infty)=T\cdot s_B(\infty)$，这样只要求出 $s_B(\infty)$，就可得到 $s_A(\infty)$．$s_B(\infty)$ 的具体求解过程如下：

首先，让我们来观察 $s_A(\infty)$、$s_B(\infty)$ 具有哪些特性．

(1) T,η 在凸集 A,B 之间交替投影，由于是正交投影，每次投影的结果是 A,B 两空间内相应的点的距离减小，这样 $s_A(\infty)$、$s_B(\infty)$ 是 A,B 内两个距离最近的点；

(2) 设 $s_A(\infty)=\begin{pmatrix} s_A^P(\infty) \\ s_A^Q(\infty)\end{pmatrix}$，$s_B(\infty)=\begin{pmatrix} s_B^P(\infty) \\ s_B^Q(\infty)\end{pmatrix}$，则由 $\eta s_A(\infty)=s_B(\infty)$ 可知 $s_A^Q(\infty)=s_B^Q(\infty)$；

(3) 由 $s_B(\infty)\in B$，可得出 $s_B^P(\infty)=s^P(0)$．

设 B 内所有满足上述三个条件的点形成的空间为 Φ，显然 $s_B(\infty)\in\Phi$．下面给出 Φ 的显示表示式，

对 $\forall x\in\Phi$，A 内 Tx 与 x 之间的距离为 $d=\|Tx-x\|$ (d 为酉范数)．

令 $Tx=\begin{bmatrix} F_P \\ F_Q\end{bmatrix}\cdot X$，$X\in C^N$，则根据第(2)、(3)个条件可知 $x=\begin{bmatrix} s^P(0) \\ F_QX\end{bmatrix}$．

那么 $d=\|Tx-x\|=\|F_P\cdot X-s^P(0)\|$，使 d 取最小值的 $X=F_P^+s^P(0)+(I-F_P^+F_P)Y$，其中 $Y\in C^N$，于是有

$$x=\begin{bmatrix} s^P(0) \\ F_QX\end{bmatrix}=\begin{bmatrix} s^P(0) \\ F_QF_P^+s^P(0)+(F_Q-F_QF_P^+F_P)Y\end{bmatrix}$$

因此

$$\Phi=\left\{x\in C^L \mid x=\begin{bmatrix} s^P(0) \\ F_QF_P^+s^P(0)+(F_Q-F_QF_P^+F_P)Y\end{bmatrix},Y\in C^N\right\}$$

为了确定唯一的 $s_B(\infty)$，还需寻找一个约束条件．这个约束条件就是 $(s_B(\infty)-s(0))\perp\Phi$（或者说，$\|s_B(\infty)-s(0)\|$ 取最小值）．下面来证明这个约束条件．

设 $u(n)=\Gamma^n s(0)-\Gamma^{n-1}s(0)=(\Gamma-I)\cdot\Gamma^{n-1}s(0)$，其中 n 为自然数，$\Gamma=\eta\circ T$ 为 η 和 T 的合成算子．

对 $\forall x_1x_2\in\Phi$，有

$$x_1-x_2=\begin{bmatrix} 0 \\ (F_Q-F_QF_P^+F_P)(Y_1-Y_2)\end{bmatrix}$$

其中 $Y_1,Y_2\in C^N$，那么

$$u^H(n)\cdot(x_1-x_2)$$
$$=s^H(0)(\Gamma^{n-1})^H(\Gamma^H-I^H)\cdot\begin{bmatrix} 0 \\ (F_Q-F_QF_P^+F_P)(Y_1-Y_2)\end{bmatrix}$$
$$=s^H(0)(\Gamma^{n-1})^H\begin{bmatrix} 0 & T_3^H \\ 0 & T_4^H-I\end{bmatrix}\cdot\begin{bmatrix} 0 \\ (F_Q-F_QF_P^+F_P)(Y_1-Y_2)\end{bmatrix}$$

根据引理 1 可得到 $u^H(n)\cdot(x_1-x_2)=0$．

由于 $(s_B(\infty)-s(0))=\sum_{j=1}^{\infty}u(j)$

则 $(s_B(\infty)-s(0))^H\cdot(x_1-x_2)=\left(\sum_{j=1}^{\infty}u^H(j)\right)\cdot(x_1-x_2)$

$$=\sum_{j=1}^{\infty}u^H(j)\cdot(x_1-x_2)=0$$

由于 x_1,x_2 的任意性，则 $(s_B(\infty)-s(0))\perp\Phi$，也就是 $s_B(\infty)$ 与 $s(0)$ 之间的距离最小．

令 $d_0=\|x-s(0)\|$，其中 $x\in\Phi$，则

$$d_0=\|F_Q(I-F_P^+F_P)Y-(s^Q(0)-F_QF_P^+s^P(0))\|$$

令 $W=F_Q(I-F_P^+F_P)$，则当 $Y=W^+(s^Q(0)-F_QF_P^+s^P(0))+(I-W^+W)Z$，$Z\in C^N$ 时，d_0 取最小值，此时的 x 即为 $s_B(\infty)$，那么

$$\begin{aligned} & s_B^Q(\infty) \\ &=W(W^+(s^Q(0)-F_QF_P^+s^P(0))+(I-W^+W)Z) \\ &\quad +F_QF_P^+s^P(0) \\ &=WW^+(s^Q(0)-F_QF_P^+s^P(0))+F_QF_P^+s^P(0) \\ &=F_Q(I-F_P^+F_P)(F_Q(I-F_P^+F_P))^+(s^Q(0)-F_QF_P^+s^P(0)) \\ &\quad +F_QF_P^+s^P(0)\end{aligned}$$

下面求 $s_A(\infty)$，由第(1)个条件可知 $s_A^Q(\infty)=s_B^Q(\infty)$，又 $s_A^P(\infty)=F_P\cdot X=F_PF_P^+s^P(0)+F_P(I-F_P^+F_P)Y=F_PF_P^+s^P(0)$

证毕．

4 扩展交替投影神经网络具备联想记忆功能的充分必要条件

定理 1 得到了一个涵盖各种情况的通式．如果从两个空间几何关系角度考虑，它涵盖两个空间不交、交于一正维集、交于一点等三种情况．若从库矩阵秩的角度考虑，它涵盖了 $\mathrm{rank}(F_P)=\mathrm{rank}(F)$ 和 $\mathrm{rank}(F_P)\neq\mathrm{rank}(F)$ 两种情况．当网络用于联想记忆，或者说需要根据部分信息联想出另一部分信息时，要求网络稳态值不受浮动神经元初始状态影响，只与箝位神经元初态有关．而当什么条件满足时，定理 1 中的通式才与 $s^Q(0)$ 无关，即满足网络用于联想记忆时提出的要求．这就是下面要证明的定理 2．

定理 2 扩展交替投影神经网络具备联想记忆功能的充分必要条件为 $\mathrm{rank}(F_P)=\mathrm{rank}(F)$．

证明 当扩展交替投影神经网络具有联想记忆功能时，要求其稳态值不受浮动神经元初始状态影响，只与箝位神经元初态有关．观察定理 1 中的通式，再运用 $s^P(0)$ 与 $s^Q(0)$ 之间的独立性以及 $s^Q(0)$ 的任意性可知定理 2 与下面的命题 1 等价．

命题 1　$F_Q(I-F_P^+F_P)=0 \Leftrightarrow \text{rank}(F_P)=\text{rank}(F)$.

首先证明:$F_Q(I-F_P^+F_P)=0 \Rightarrow \text{rank}(F_P)=\text{rank}(F)$.

由 $F_Q(I-F_P^+F_P)=0$,可得到 $F(I-F_P^+F_P)=0$.又 $F_P^+F_P=(F_P^+F_P)^H=F_P^H(F_P^H)^+$,从文献[18]中网络稳定性的证明过程可得到 $\text{rank}(F_P^H(F_P^H)^+)=\text{rank}(F_P^H)$,于是就有 $\text{rank}(F_P^H)=\text{rank}(F_P^+F_P)=\text{rank}(F_P)$,那么 $\text{rank}(I-F_P^+F_P)=N-\text{rank}(F_P)$

由 $\text{rank}(F_P)\leqslant\text{rank}(F)$,可知 $\text{rank}(I-F_P^+F_P)\geqslant N-\text{rank}(F)$

但方程 $F\cdot x=0$ 解空间的维数为 $N-\text{rank}(F)$.于是可得到 $\text{rank}(F_P)=\text{rank}(F)$.

再来证明:$\text{rank}(F_P)=\text{rank}(F) \Rightarrow F_Q(I-F_P^+F_P)=0$

设 $\text{rank}(F_P)=\text{rank}(F)=r$,对 F 进行满秩分解有:

$F=\begin{bmatrix}C_1\\C_2\end{bmatrix}\cdot D=C_0\cdot D$,其中 $C_0=\begin{bmatrix}C_1\\C_2\end{bmatrix}\in C^{L\times r},D\in C^{r\times N},C_1\in C^{P\times r},C_2\in C^{Q\times r},F_P=C_1D,F_Q=C_2D$

$\text{rank}(F_P)=\text{rank}(F)=r=\text{rank}(C_1D)\leqslant\text{rank}(C_1)$,

$$\text{即}\quad \text{rank}(C_1)\geqslant r \qquad (a)$$

又 $r=\text{rank}(C_1D)\geqslant\text{rank}(C_1)+\text{rank}(D)-r=\text{rank}(C_1)$,

$$\text{即}\quad \text{rank}(C_1)\leqslant r \qquad (b)$$

由(a)、(b)可知 $\text{rank}(C_1)=r$,这样就有 $F_P^+=(C_1D)^+=D^+C_1)^+$,那么可得到

$$\begin{aligned}F_Q(I-F_P^+F_P)&=F_Q-F_QF_P^+F_P\\&=C_2D-C_2DD^+(C_1^HC_1)^{-1}C_1^HC_1D\\&=C_2D-C_2DD^+D=C_2D-C_2D=0\end{aligned}$$

从而定理 2 获证.

当网络用于学习的各库模式线性独立时,此时由于 F 变为列满秩阵,那么相应地网络具备联想记忆功能的充分必要条件就变为 F_P 为列满秩阵.

5　仿真实例

为了说明上述理论分析的正确性,下面设计两个仿真实例进行验证.

例 1　取 $L=5,N=3$,复模式:

$$f_1=\begin{bmatrix}1 & 2+\frac{i}{2} & 6 & 9+3i & 8\end{bmatrix}^{\mathrm{T}},$$

$$f_2=[2\ \ {-4}+i\ \ 9\ \ 8\ \ 7]^{\mathrm{T}},$$

$$f_3=\begin{bmatrix}13-12i & 32-\frac{35i}{2} & 63-51i & 61-29i & 59-33i\end{bmatrix}^{\mathrm{T}}$$

将其训练到 EAPNN 网中.设 EAPNN 网经过 n 次联想得到浮动神经元的状态向量为 $s^Q(n)$,直接运用定理 1 的结论得到浮动神经元的稳定状态向量为 $s_B^Q(\infty)$.令 $s^P(0)=F_P\cdot[c_1\quad c_2\quad c_3]^{\mathrm{T}}$,其中 $c_1=(1+3i)/\sqrt{10},c_2=2+i,c_3=e^{i\pi/5}$..令 $s^Q(0)=0,y=F_Q\cdot[c_1\quad c_2\quad c_3]^{\mathrm{T}}$,则当 $P=2,3$ 时的计算结果如表 1 所示.

表 1　例 1 在 $S^Q(0)=0$ 时的计算结果

P	n	$s^Q(n)$	$s_B^Q(\infty)$	$s_B^Q(\infty)-y$	$s^Q(n)-y$	$s^Q(n)-s_B^Q(\infty)$
2	5	$\begin{pmatrix}58.5829+7.9256i\\-6.1063-21.1168i\\1.3350+16.7505i\end{pmatrix}$	$\begin{pmatrix}73.5147+9.9456i\\-7.6627-26.4991i\\1.6753+21.0199i\end{pmatrix}$	$\begin{pmatrix}-27.3278-0.5171i\\-90.0585-56.3794i\\-81.9835-1.5514i\end{pmatrix}$	$\begin{pmatrix}-42.2596-2.5372i\\-88.5021-50.9971i\\-82.3237-5.8208i\end{pmatrix}$	$\begin{pmatrix}-14.9318-2.0201i\\1.5564+5.3823i\\-0.3403-4.2694i\end{pmatrix}$
	15	$\begin{pmatrix}72.8987+9.8622i\\-7.5985-26.2771i\\1.6613+20.8437i\end{pmatrix}$			$\begin{pmatrix}-27.9438-0.6005i\\-89.9943-56.1573i\\-81.9975-1.7275i\end{pmatrix}$	$\begin{pmatrix}-0.6160-0.0833i\\0.0642+0.2220i\\-0.0140-0.1761i\end{pmatrix}$
	25	$\begin{pmatrix}73.4893+9.9421i\\-7.6600-26.4900i\\1.6747+21.0126i\end{pmatrix}$			$\begin{pmatrix}-27.3532-0.5206i\\-90.0558-56.3702i\\-81.9840-1.5586i\end{pmatrix}$	$\begin{pmatrix}-0.0254-0.0034i\\0.0026+0.0092i\\-0.0006-0.0073i\end{pmatrix}$
3	100	$\begin{pmatrix}58.6141+14.3300i\\63.1627+23.8628i\end{pmatrix}$	$\begin{pmatrix}82.3958+29.8802i\\83.6587+22.5712i\end{pmatrix}$	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}-23.7817-15.5502i\\-20.4960+1.2915i\end{pmatrix}$	$\begin{pmatrix}-23.7817-15.5502i\\-20.4960+1.2915i\end{pmatrix}$
	300	$\begin{pmatrix}80.3198+28.5228i\\81.8695+22.6840i\end{pmatrix}$			$\begin{pmatrix}-2.0760-1.3575i\\-1.7892+0.1127i\end{pmatrix}$	$\begin{pmatrix}-2.0760-1.3575i\\-1.7892+0.1127i\end{pmatrix}$
	600	$\begin{pmatrix}82.3423+29.8452i\\83.6126+22.5741i\end{pmatrix}$			$\begin{pmatrix}-0.0535-0.0350i\\-0.0461+0.0029i\end{pmatrix}$	$\begin{pmatrix}-0.0535-0.0350i\\-0.0461+0.0029i\end{pmatrix}$

从表 1 可以看出,$s^Q(n)-s_B^Q(\infty)$随着 n 的增加趋近于 0,这验证了定理 1.

当 $P=2$ 时,$\text{rank}(F_P)=1\neq\text{rank}(F)=2$,不论 n 为何值,$s_B^Q(\infty)-y$、$s^Q(n)-y$ 的值皆不为 0,这说明了网络此时不具备联想记忆功能.

当 $P=3$ 时,$\text{rank}(F_P)=\text{rank}(F)=2,s_B^Q(\infty)-y=0$,随着 n 的增加 $s^Q(n)-y$ 趋近于 0,这说明了此时网络可用于联想记忆.

例 2　取 $L=50,N=2$,复模式:

$$f_1=\begin{bmatrix}2 & 3 & 4 & e^{i\frac{6\pi}{47}\times0} & e^{i\frac{6\pi}{47}\times1} & e^{i\frac{6\pi}{47}\times2} & \cdots & e^{i\frac{6\pi}{47}\times46}\end{bmatrix}^{\mathrm{T}}$$

$$f_2=\begin{bmatrix}4 & 6 & 8 & e^{i\frac{2\pi}{47}\times0} & e^{i\frac{2\pi}{47}\times1} & e^{i\frac{2\pi}{47}\times2} & \cdots & e^{i\frac{2\pi}{47}\times46}\end{bmatrix}^{\mathrm{T}}$$

显然 f_1 与 f_2 线性独立,令 $s^P(0)=F_P\cdot[c_1\quad c_2]^T$,其中,$c_1=(1+3i)/\sqrt{10}$,$c_2=2+i$,$s^Q(0)=0$,$y=F_Q\cdot[c_1\quad c_2]^T$.定义 $e(n)=\|s^Q(n)-y\|/\|y\|_2$ 为网络经过 n 次联想得到的联想结果 $s^Q(n)$ 与源信号相应部分 y 的相对偏差.

当 $P=3$ 时,$\mathrm{rank}(F_P)=1\neq\mathrm{rank}(F)=2$.当网络分别经过 $n=1,2,3,4,\cdots,98,99,100$ 次联想得到的 $e(n)\sim n$ 曲线如图 2 所示.从图中可以看出,网络很快收敛到稳态解,但此稳定解不是源信号相应部分 .这说明网络此时不具有联想功能.

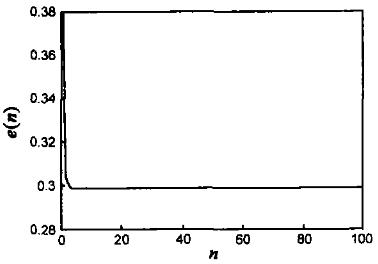


图 2　$P=3$ 时,$e(n)\sim n$ 曲线

当 $P=4$ 时,$\mathrm{rank}(F_P)=\mathrm{rank}(F)=2$.当网络分别经过 $n=50,100,150,\cdots,1150,1200$ 次联想得到的 $e(n)\sim n$ 曲线如图 3 所示.

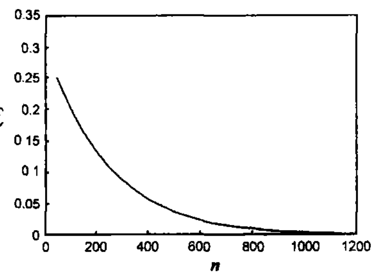


图 3　$P=4$ 时,$e(n)\sim n$ 曲线

从图 3 中可以看出,网络渐渐收敛到稳态解,此稳定解为 y.说明了网络具有联想功能.$P=3$ 时网络收敛得远比 $P=4$ 快是因为 $P=3$ 时 T_4 小于 1 的最大特征值($\lambda_2=0.2448$)比 $P=4$ 时 T_4 小于 1 的最大特征值($\lambda_1=0.9958$)小很多的缘故.

以上的仿真实验表明,不论网络用于学习的复模式是否线性独立,网络总能正确无误地工作,而且仿真结果与前面的理论分析结果完全一致.

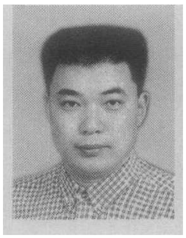
6　结束语

交替投影神经网络(APNN)是由美国 Washington 大学 Marks Ⅱ 等人提出的,他们在研究 APNN 时只给出了网络用于联想记忆时的充分条件,没有给出必要条件,而且所有的研究都是建立在这样两个前提条件基础上:(1)作用域为实数域;(2)网络用于学习的库模式线性独立.

本文根据实际应用的需要,在研究 APNN 网的基础上提出一种新的神经网络——扩展交替投影神经网络(Extended Alternating Projection Neural Network),并围绕该网络,做了如下改进工作:(1)拓宽了其应用范围,EAPNN 网的应用范围为复数域而不是实数域,这样更利于其在信号处理方面的应用;(2)给出了一种改进的权值学习方法,该方法不要求网络用于学习的复模式一定要线性独立;(3)得到了网络稳态值的通用数学表达式,并从表达式中推出了网络具备联想记忆功能的充分必要条件.另外,我们还设计了仿真实验,对文中的理论分析结果进行了验证.

在工程实际中我们已经利用 EAPNN 网络良好的根据部分信息联想出另一部分信息的功能成功地解决了信号处理中的带限信号外推问题[19]、选频和陷波问题[20]以及弱信号分离问题[21].其实,要想用 EAPNN 网络较好地解决现实问题,库模式的生成也是至关重要的.如何生成库模式,跟具体的问题有关.我们针对不同的具体问题采用了不同的方法来生成库模式,像谱分析的方法,非线性变换方法以及经验的方法等等.文献[1]为了用 APNN 网络来解决分类问题,采用了非线性变换的方法来生成库模式.

作者简介:



王金根　男,1972 年出生于安徽怀宁,博士,副教授,现在南京大学计算机科学与技术博士后流动站从事软计算及其应用方面的研究工作.E-mail:wjg1288@sina.com.

龚沈光　男,1939 年出生于上海,现为海军工程大学教授、博士生导师,主要研究领域为:现代信号处理,人工智能,非触发智能引信等.

陈世福　男,1938 年出生于安徽,现为南京大学计算机科学与技术系教授、博士生导师,主要研究领域为机器学习,知识工程,分布式人工智能.

参考文献:

[1] R J Marks Ⅱ,S Oh,L E Atlas. Alternating projection neural networks[J]. IEEE Trans CAS. 1989,36(6):846-857.

[2] J J Hopfield,D W Tank. Neural computation of decisions in optimization problem[J]. Biol. Cybern,1985,52:141-152.

[3] D W Tank,J J Hopfield. Simple neural optimization networks: An A/D converter, signal decision circuit and a linear programming circuit[J]. IEEE Trans. CAS,1986,33(5):533-541.

[4] M Takeda,J W Goodman. Neural networks for computation: Number representation and programming complexity[J]. Appl Opt,1986,25(18):3033-3046.

[5] S Geman,D Geman. Stochastic relaxation, Gibb's distributions, and the Bayesian restoration of images[J]. IEEE Trans PAMI,1984,6:721-741.

[6] J J Hopfield. Neural networks and physical systems with emergent collective computational abilities[J]. In Proc National Academy of Sciences, USA,1982,79:2554-2558.

[7] D H Ackley,G E Hinton,T J Sejnowski. A learning algorithm for Boltzmann machines[J]. Cognitive Sci,1985,9:147-169.

(下转第 605 页)

活)成为求解问题的策略:把经验性知识转换成为结构主义的策略;把正规性知识转换成为功能主义的策略;把常识性知识转换成为行为主义的策略;

第四,通过控制和显示,可以把求解问题的策略转变为相应的行为,在满足环境约束的条件下达到预定的目标,使问题得到解决.

可见,信息获取理论、全信息理论、知识理论、综合智能理论及控制理论之间相互贯通,构成了"信息-知识-策略-行为的转换与统一理论",构成了智能科学的核心和精髓.

作者相信,"信息-知识-策略-行为的转换与统一理论"是信息科学与智能科学发展到今天必然要产生的结果.信息时代"信息-知识-策略-行为的转换与统一理论"的科学意义可以同工业时代"能量转换与守恒定律"相媲美,因为它们分别揭示了能量资源和信息资源加工转换的根本规律,为人们利用能量和信息资源提供了理论的指导.自然,本文的结果仅仅是"信息-知识-策略-行为统一理论"的初步,更精彩的结果有待广大同仁共同努力去开发.

参考文献:

[1] 钟义信,等.智能理论与技术:人工智能与神经网络,邮电出版社,1992.

[2] 钟义信.信息科学原理(第三版).北京邮电大学出版社,2002.

[3] 钟义信.知识理论框架.中国工程科学,2000,9(2):50-64.

[4] 钟义信. Unified Theory of Information, Knowledge and Intelligence. Chinese Journal of Electronics,2003,6.

作者简介:

钟义信 男,1940年2月出生于江西赣州市,北京邮电大学教授,博士生导师.

(上接第600页)

[8] K F Cheung, R J Marks II, L E Atlas. Synchronous vs. asynchronous behaviour of Hopfield's CAM neural net[J]. Appl Opt, 1987, 26(22): 4808-4813.

[9] R P Lippman. An introduction to computing with neural nets[J]..IEEE ASSP Mag, Apr 1987, 4(2): 4-22.

[10] N Farhat, D Psaltis, E Paek. Optical implementation of the neural model [J]. Appl Opt, 1985, 24(10): 1469-1475.

[11] L E Atlas. Auditory coding in higher centers of the CNS[J]. IEEE Eng Med Biol Mag, June 1987, 29-32.

[12] R J Marks II, L E Atlas. Geometrical interpretation of Hopfield's content addressable memory neural nelwork [A]. In Proc. Northcon' 88 [C]. Seattle, WA, Vol. II Oct. 1988. 964-977.

[13] Y S Abu-Mostafa, J M St. Jaques. Information capacity of the Hopfield model[J]. IEEE Trans IT, 1985, IT-31: 461.

[14] R J McEliece, E C Posner, E R Rodemich, S S Venka-tesh. The capacity of the Hopfield associative memory[J]. IEEE Trans IT, 1987, IT-33: 461-482.

[15] D E Rumelhart, J L McClelland, the PDP Research Group. Parallel Distributed Processing [M]. vols. I and II. Cambridge, MA: Bradford Books, 1986. 318-362.

[16] D E Rumelhart, G E Hinton, R J Willians. Learning representations by back-propagation errors[J]. Nature, 1986, 323 (6088): 533-536.

[17] C Stanfill, D Waltz. Toward memory-based reasoning[J]. Commun. ACM, Dec. 1986, 29: 1213-1228.

[18] 王金根.航空磁探系统中目标信号检测与磁性目标定位研究[D].博士学位论文.武汉:海军工程大学,2001,5.

[19] 王金根,林春生,龚沈光.基于交替投影神经网络的带限信号外推算法[J].电子学报,2000,28(10):52-55.

[20] 王金根,龚沈光,林春生,唐劲飞,刘胜道.基于复交替投影神经网络的陷波器[J].数据采集与处理,2001,16(4):440-445.

[21] WANG Jin-gen, CHEN Shi-fu, GONG Shen-guang, CHEN Zhao-qian. An Extended Alternating Projection Neural Networks based weak-signal separation algorithm[A]. IEEE International Conference on Robotics. Intelligent Systems and Signal Processing[C]. Changsha, China, Oct. 2003. 554-558.