高校应用数学学报
APPLIED MATHEMATICS A
JOURNAL OF CHINESE UNIVERSITIES
1999年 第14卷 第1期 Vol.14 No.1 1999

时间离散Hopfield神经网络系统的若干问题

沈世镒

摘要 如果一个Hopfield神经网络系统(以下简记为HNNS)是时间离散、状态连续的，就称之为时间离散的HNNS(以下简记为TD-HNNS).对这种系统，如果在它的算子的作用下，状态的能量函数具有固定的增、减趋势，那么就称之为单向的TD-HNNS.本文讨论并给出了这种模型的一系列性质，如运动轨迹的稳定性、收敛性和稳定解的唯一性等，并由此给出了它在优化计算中的一系列应用.

关键词 时间离散的Hopfield神经网络系统(TD-HNNS)，运动轨迹的稳定性、收敛性与稳定解的唯一性，优化计算.

分类号 (中图) O23; (1991MR)94A15.

SOME PROBLEMS OF THE TIME DISCRETE HOPFIELD NEURAL NETWORK SYSTEM

Shen Shiyi
(Dept. of Math.,Nankai University,Tianjin,300071)

Abstract A Hopfield neural network system (HNNS) is said to be the time discrete HNNS (TD－HNNS),if it is time discrete and state continuous.A TD－HNNS is said to be unilateral,if its energy functions are monotone increasing or monotone decreasing for the HNNS operator.In this paper, a sequence of properties for the unilateral TD－HNNS such as the stability of motion,convergence,and the unicity of the stable solution,etc. are obtained.And some applications in the optimizing calculation are given.

Keywords Time Discrete Hopfield Neural Network System (TD－HNNS), Stability, Convergence, and Unicity of the Stable Solution for the Motion Locus, Optimizing Calculation.

Subject Classification (CL)O23;(1991MR)94A15.

1 引 言

HNNS模型是人们所熟悉的，自1982年J.J.Hopfield在［1］文中提出以来，曾一度引起人们的强烈兴趣，从而导致80年代末、90年代初神经网络系统理论发展的高潮.但是，随着理论研究的深入发展，人们发现HNNS理论存在严重的缺陷.其主要原因是HNNS模型存在大量的假吸引点(见［2］文)，从而使许多优化与识别问题无法使用.另外，现有的HNNS的运动方程，在权矩阵是对称正定条件下，其运动状态的能量函数是一个递减函数(见［3］文)，这与许多常用的优化问题并不一致，因此无法利用HNNS的运动方程来得到最优解.因此，近年来人们对HNNS提出了许多质疑与改进意见，但还未得到满意的解决，从而大大影响了神经网络系统理论的应用与发展.

为了解决上述问题，本文提出了TD-HNNS模型和它的运动轨迹与方向问题.TD-HNNS模型就是指时间离散、状态连续的HNNS模型，这种模型具有以下特点：

1.以差分的形式建立相应神经网络系统的运算子，因此它与离散型HNNS有本质的区别.

2.对其运动轨迹提出了方向性问题，这时其运动状态在相应的算子作用下，其能量函数具有固定的增减变化关系，如果它运动状态的能量函数是递减的，则称之为是正向的，反之，如果它运动状态的能量函数是递增的，则称之为是反向的.在现有的HNNS模型中，一般只研究正向情形，但大量优化计算问题都是反向的.因此，本文讨论的反向模型有重要的实际应用背景.

3.为寻找反向的TD-HNNS，作者对激励函数做更广泛的讨论，并给出了激励函数对神经网络系统状态运动影响的若干性质.

本文针对上述TD-HNNS，给出了它的一系列基本性质，如稳定性、运动轨迹的收敛性、稳定解的唯一性等.本文还给出了一系列应用问题.由此可见，TD-HNNS不仅具有完美的理论性质，而且还有广泛的应用范围，是一个值得重视的神经网络发展领域.

2 单向TD-HNNS模型

现有的HNNS模型大体上可分为两大类型，即离散型与连续型.一个HNNS的基本模型可用一个算子

$$x' = T(x) = T(W, h; x) \quad (1)$$

表示，其中$x=(x_1, x_2, \cdots, x_n)$，$x'=(x'_1, x'_2, \cdots, x'_n)$为该算子的输入与输出向量，$h=(h_1, h_2 2, \cdots, h_n)$为阈值向量，而

$$W = \begin{pmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,n} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,n} \\ \cdot & \cdot & \cdot & \cdot \\ w_{n,1} & w_{n,2} & \cdots & w_{n,n} \end{pmatrix} \quad (2)$$

为权矩阵.阈值向量h与权矩阵W都是HNNS的参变数，称(1)中的算子为HNNS算子.

定义1　一个HNNS算子T，如果它是$\{-1,+1\}^n \to \{-1,+1\}^n$的映射，且相应的运算结果为：

$$x'_i = \mathrm{Sgn}\Big[\sum_{j=1}^{n} w_{i,j} \cdot x_j - h_i\Big], \ i = 1, 2, \cdots, n, \quad (3)$$

那么称之为离散型的HNNS.

一个连续型HNNS模型可通过一个微分方程组模型描述，这时状态向量$x=(x_1, x_2, \cdots, x_n)$是n维欧氏空间R中的向量，而其运动方程为：

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = f\Big[\sum_{j=1}^{n} w_{i,j} \cdot x_j(t) - h_i\Big], \ i = 1, 2, \cdots, n, \quad (4)$$

其中$y=f(u)$为激励函数，在现有的HNNS理论中，一般取f(u),它是一个单增奇函数，且以$y=\pm 1$为渐近线.在本文中将取消这个限制.

离散型和连续型HNNS理论主要用于优化与识别问题，但是这些模型在理论上遇到了假吸引点问题与局部最优解问题的困难.也就是，现有的离散型HNNS模型，在它运动时会出现大量的假吸引点问题(见［2］中讨论)，而连续型HNNS在它运动时会出现大量的局部最优解，且其运动轨迹在许多情形下是发散的，这样最终还是无法得到真正的最优解，因此就大大影响了HNNS的应用范围.

本文选择的TD-HNNS是一个状态连续，时间离散的HNNS，其算子的基本运算形式与(1)相同，这时T是一个n维欧氏空间R中的一个映射，而相应的计算公式是：

$$x'_i = x_i + f\Big[\sum_{j=1}^{n} w_{i,j} \cdot x_j - h_i\Big], \ i = 1, 2, \cdots, n. \quad (5)$$

本文的主要目的就是讨论这种TD-HNNS的收敛性与稳定性问题，及它们在多种领域中的应用.构造TD－HNNS状态x的能量函数为：

$$E\{x\} = \sum_{i=1}^{n} h_i \cdot x_i - \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n} w_{i,j} \cdot x_i \cdot x_j. \quad (6)$$

定义TD－HNNS的吸引点:称x是一个吸引点，如果有关系式$x=T(x)$成立.

定义2　设T是一个TD－HNNS算子，如果对任何$x \in R$总有$E\{T(x)\} \leqslant E\{x\}$(或总有$E\{T(x)\} \geqslant E\{x\}$)成立，那么就称这种TD－HNNS是正向的(或反向的).对正向或反向的TD-HNNS总称之为单向的TD－HNNS.如果在定义的不等式中等号成立的充要条件是向量x是吸引点，那么分别称这种TD－HNNS是严格正向或严格反向的(总称是严格单向的).

3　TD－HNNS的稳定性问题

定义3　称$\{x(t),t=0,1,2,\cdots\}$为TD－HNNS的一个运动轨迹，如果

$$x(t+1)=[W,\ h;\ x(t)],\quad t=0,1,2,\cdots$$

成立.显然，这个TD－HNNS的运动轨迹由算子T与初始状态x(0)决定.

定义4　对一个固定的TD－HNNS，如果对任何初始状态x(0)，它的运动轨迹$\{x(t),t=0,1,2,\cdots\}$在R中总是一个收敛序列，那么称这个TD－HNNS是一个稳定系统.

定义5　一个TD－HNNS称其运算子是连续的，如果它的运算子T(x)关于变量x是连续的.

引理1　一个TD-HNNS有以下性质：

(1) 在(5)式中，如果激励函数f(u)是一个连续函数，那么这个TD－HNNS的运算子一定是连续的.

(2) 一个稳定、连续的TD－HNNS，它的任何一个运动轨迹必收敛于一个吸引点.

(3) 在(5)式中，如果激励函数在零点的解是唯一的，也就是f(u)=0的解是唯一的，那么x是这个TD－HNNS吸引点的充要条件是x是方程组

$$\sum_{j=1}^{n} w_{i,j}\cdot x_j = h_i,\ i=1,2,\cdots,n \tag{7}$$

之解.由此可得，如果矩阵W是满秩的那么它的吸引点是唯一的，否则就无吸引点或有无穷多个吸引点.

引理1的证明是显然的.

定理1　如果TD-HNNS(5)的权矩阵W是对称正定的，且在激励函数y=f(u)中，y与u同号，那么这个TD\|HNNS是正向的.这时对任何$x\in R$总有$E\{T(x)\}\leqslant E\{x\}$成立，如果激励函数在零点的解是唯一的，那么等号成立的充要条件是x为吸引点.

定义6　T是一个TD－HNNS的算子，称之为是有界的，如果它对任何初始状态向量x(0),由这个TD－HNNS算子T与这个初始状态x(0)决定的运动轨迹x(t),t=0,1,2,…，总是一个有界序列.

定理2　在定理1的条件下，如果TDHNNS的算子T是有界的，它的激励函数是连续的，且在零点的解是唯一的，那么它一定是稳定的，且有唯一的一个吸引点.这个吸引点就是能量函数的最小值.

在此需要说明的一点是正向有界的TD－HNNS算子在一般情形下是不存在的.

例1　如果取$w_{i,j}=\delta_{i,j}$,当i=j时，$\delta_{i,j}=1$,否则$\delta_{i,j}=0$，而$h_i=0$,取激励函数f(u)是一个与u同号的连续函数，f(u)=0的根是唯一的.那么易证这个TD－HNNS算子是无界的.

以下讨论反向TD－HNNS的稳定性问题.设λ_i,i=1,2,…,n是权矩阵W的n个特征根(可以是重根)，而记

$$\lambda_{max}=\max\{|\lambda_i|,\ i=1,\cdots,n\},\ \lambda_{min}=\min\{|\lambda_i|,\ i=1,\cdots,n\}.\tag{8}$$

这时取满足以下条件的激励函数$f_0(u)=-f(u)$：

J-1.　$f_0(u)$是一个连续的奇函数，且对任何$u>0,f_0(u)>0$.

J-2. 对任何$u>0,f_0(u)\leqslant u/(\lambda_{max}+1)$.

这时以下定理成立：

定理3　如果TD－HNNS的权矩阵是对称的，且其激励函y=f(u)满足条件J-1与J-2，那么这个TD-HNNS一定是反向的.这时对任何$x\in R$总有$E\{T(x)\}\geqslant E\{x\}$成立，且等号成立的充要条件是x为吸引点.

在定理3中对权矩阵W只作对称性要求，而激励函数f(u)是可以适当选择的，因此，定理3的条件是较为广泛的.

定理4　在定理3的条件下，如果权矩阵W是正定的，那么这个TD－HNNS一定是稳定的，且有唯一的一个吸引点.这个吸引点就是能量函数的最大值.

定理3与定理4的证明从略.

4　在某些约束条件下的TD－HNNS

所谓约束条件就是指HNNS的状态向量x在一定的区域范围D中变化，常见的区域D有：

超平面：$D=\{x:\ \sum_{j=1}^{n} a_{i,j}\cdot x_j=c_i,i=1,2,\cdots,m\}$.

凸区域：$D=\{x:\ \sum_{j=1}^{n} b_{i,j}\cdot x_j\geqslant d_i,i=1,2,\cdots,m\}$.

因此，这时的TD－HNNS为：

$x' = T(W, h; x), \quad x, x' \in D$, (9)

其中算子T由(1)与(5)式定义.以下只讨论在约束条件

$$D = \left\{ x : x_i \geq 0,\ i = 1, 2, \cdots, n;\ \sum_{i=1}^{n} x_i = 1 \right\} \tag{10}$$

下的TD－HNNS运动问题.为了使得(9)式中的运动算子T与约束条件D保持一致，引进在约束条件(10)下的运动算子$x' = T(W, h; x)$由以下步骤组成：

T－1. 定义：$v_i = x_i + f(u_i),\ i = 1, 2, \cdots, n$，其中 $u_i = \sum_{j=1}^{n} w_{i,j} \cdot x_j - h_i$.

T－2. 取 $x_i'' = \begin{cases} v_i & v_i \geq 0, \\ 0 & v_i < 0. \end{cases}$

T－3. 取 $x_i' = \dfrac{x_i''}{\sum_{j=1}^{n} x_j''}$.

由T-1，T-3式确定的运算为在约束条件(10)下的TD－HNNS的运算子.记之为

$x' = T_D(x) = T_D(W, h; x)$. (11)

对运算子$x' = T_D(x)$同样可以定义与讨论它的稳定性问题.

5 若干应用问题

反向TD-HNNS模型应用问题的基本模式是在约束条件(10)下求二次型(6)的最大值.其基本算法是利用TD－HNNS的运动算子(5)或(T-1)，(T-3)，其状态函数趋向于求二次型(6)的最大值.这个模型的应用范围十分广泛，举以下情形给以说明.

1.求凸区域与原点的最小距离

记Q是一个在m维空间R^m中的由向量组$q_i = (q_{i,1}, q_{i,2}, \cdots, q_{i,m})$, i=1, 2, …, n, 决定的凸区域.这时

$$Q = \left\{ z = \sum_{i=1}^{n} x_i q_i : x_i \geq 0, i = 1, 2, \cdots, n; \sum_{i=1}^{n} x_i = 1 \right\}. \tag{12}$$

求集合Q与原点的最小距离是求

$$z^2 = E\{x\} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j w_{i,j} \tag{13}$$

在约束条件(10)下的最小值，其中$w_{i,j}$是向量q_i与q_j的内积：

$$w_{i,j} = (q_i, q_j) = \sum_{k=1}^{m} q_{i,k} q_{j,k}, \quad i, j = 1, 2, \cdots, n.$$

这时W矩阵是对称正定的，只要适当选择激励函数f(u)就可使TD－HNNS的运动轨迹收敛于最优解.一个凸区域与原点最小距离的值，在感知器学习算法的复杂度理论中有十分重要的意义，有关这方面的理论见文［4］论述.

2.多元线性回归系数的估计

一个多元线性回归分析问题的基本模型是：

$$Y=\sum_{j=1}^{n} a_j \cdot U_j+\varepsilon, \quad (14)$$

其中a_j, j=1, 2, …, n是回归系数，而ε是随机误差.一个多元线性回归分析的基本问题是由已知观察样本：

$$(u_i, y_i),\ i=1,2,\cdots,m \quad (15)$$

来估计回归系数a_j, j=1, 2, …, n, 其中

$$u_3=(u_{i,1}, u_{i,2}, \cdots, u_{i,n}),\ i=1,2,\cdots,m.$$

记$x=(x_1, x_2, \cdots, x_n)$是回归系数$a=(a_1, a_2, \cdots, a_n)$的估计量，这时估计量x的均方误差是

$$E\{x\}=\sum_{i=1}^{m}\left(y_i-\sum_{j=1}^{n} x_j u_{i,j}\right)^2. \quad (16)$$

因此，回归系数a的最优估计是使E{x}为最小的x, 其中$\{u_{i,j}, y_i, i=1,2,\cdots,m; j=1,2,\cdots, n\}$是已知数据.展开(16)式就可得到

$$E\{x\}=h_0-\sum_{j=1}^{n} h_j x_j+\sum_{j=1}^{n}\sum_{k=1}^{n} x_j x_k w_{j,k}, \quad (17)$$

其中

$$h_0=\sum_{i=1}^{m} y_i^2, \quad h_j=2\sum_{i=1}^{m} y_i u_{i,j}, \quad w_{j,k}=\sum_{i=1}^{m} u_{i,j} u_{i,k}.$$

在(17)式中，权矩阵W显然是正定的，因此，同样可用TD－HNNS来求回归系数的最优估计.

3.最优投资决策(Portfolios)问题的神经网络计算

一个最优投资决策(Portfolios)问题模型在文［5］和［6］中有详细论述，一个金融或股票的数据行情可以用下列多维随机微分方程表示.记$P_i(t)$为第i支股票在时刻t的价格，$Y_i(t)$ $(t)\triangleq\ln[P_i(t)],\ i=1,2,\cdots,n.$ 这时$P_i(t)$与$Y_i(t)$都是随机变量，在经济学的数学模型中，$Y_i(t)$满足以下随机微分方程(见［7］ 文)：

$$\mathrm{d}Y_i(t)=h_i\cdot \mathrm{d}t-Z_i\sqrt{\mathrm{d}t},\ i=1,2,\cdots,n, \quad (18)$$

其中Z_i, i=1, 2, …, n是一组正态随机变量，均值为零，协方差矩阵为$W=[\varpi_{ij}]_{n\times n}$.以下记$b=(b_1, b_2, \cdots, b_n)$为一个投资决策(Portfolio)，其中$b_i$表示投资人在第i个项目上的资金投入比例，因此，

$$b_i\geqslant 0,\ i=1,2,\cdots,n.\ \sum_{i=1}^{n} b_i=1. \quad (19)$$

这时由［8］的证明可知，决策b在时刻t的收益率为：

$$\frac{\mathrm{d}E\{b,t\}}{\mathrm{d}t} = \sum_{i=1}^{n} h_i \cdot b_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j} \cdot b_i \cdot b_j, \tag{20}$$

其中dg(t)表示变量t的微分记号，而

$$\mathrm{d}E\{b,t\} = E^*\left\{\sum_{i=1}^{n} b_i \cdot \mathrm{d}[Y_i(t)]\right\} \tag{21}$$

为决策b的收益率(见［5］).在(21)式等式的右边，E*{V}表示随机变量V的数学期望.

由以上讨论可知，为求市场模型(17)的最优投资决策(Portfolio)可化为求(20)式中二次型的最大值，其中权矩阵W显然是对称正定的，因此可用TD-HNNS来进行优化计算.

以上三个例子说明了利用TD-HNNS的优化计算有着十分广泛的应用前景，其主要优点是它可以进行直接计算，因此收敛速度快，程序编写也很统一简单.由于篇幅所限，有关模拟计算结果与分析在此就不一一说明了.

(本文作者通讯地址：天津市南开大学数学系　邮编　300071)

本文得到国家自然科学基金(19671048)、国家教委博士点基金(9105509)、金融数学重大项目和国家攀登计划认知科学(神经网络)重大关键项目(79013079790130)资助.

作者单位：沈世镒　(南开大学数学科学学院)

参考文献

1　Hopfield,J.J.,Neural Network and Physical Systems with Emergent Collective Computational Abilities,Proc.Nat.Acad.Sci.USA,1982,79:2554～2558.

2　张承福，王心强，赵凯华，神经网络动力学性质初探，中国神经网络首届学术大会论文集，北京，1990，334～337.

3　Bruck,J.and Goodman,J.W.,A generalized convergence theorem for neural networks,IEEE Trans.Inform.Theory,1988,34:1089～1092.

4　沈世镒，神经网络系统理论及其应用，即将出版.

5　Cover Thomas M.and Thomas,Joy A.,Information Theory,John Wiley and Sons,Inc.,New York,1991.

6　Sharpe J.E.,Investments,Prentice－Hall,Englewood Cliffs,NJ,1985.

7　Robert,A.Jarrow and Andrew Rudd,Option Pricing,Richard D.Irwin,INC.,1983.

8　Shen Shiyi and Lam Kin,Some Calculation Formulas of the Optimum Porfolios on the Multi－Stocks Prices Model,to the publisher.

收稿：1997-02-23.修回：1997-07-09.