

Hopfield 型连续神经网络指数收敛速度的估计*

韦 岗 田传俊

(华南理工大学电子与信息学院·广州, 510640)

摘要: 讨论了 Hopfield 型连续神经网络的指数收敛速度问题, 从理论上得到了新的局部指数收敛条件. 与以往的结果相比, 本文导出指数收敛的阶更大, 指数收敛的速度可更快. 文中还给出了数值举例.

关键词: 神经网络; Hopfield 网络; 指数收敛速度

文献标识码: A

Estimate of Exponential Convergence Rate for Hopfield-Type Continuous Neural Networks

WEI Gang and TIAN Chuanjun

(College of Electronic & Information Engineering, South China University of Technology, Guangzhou, 510640, P.R. China)

Abstract: This paper is concerned with the estimating of exponential convergence rate for Hopfield-type neural networks, and establishing a new sufficient condition for local exponential convergence. The larger order the exponential convergence, the faster the convergence rate may be. An example is presented also.

Key words: neural networks; Hopfield network; exponential convergence rate

1 引言(Introduction)

Hopfield 型连续神经网络

$$\frac{dx_i(t)}{dt} = a_i x_i(t) + \sum_{j=1}^n T_{ij} g_j(x_j(t)) + c_i, \\ t \geq 0, i = 1, \dots, n \quad (1)$$

的稳定性和指数收敛性等问题已被不少作者研究过, 获得了许多结果^[1~8]. 在(1)中, $X = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, $a_i > 0$ 和 $c_i (i = 1, \dots, n)$ 是常数, 连接矩阵 $T = (T_{ij})_{n \times n}$ 是常数矩阵, 函数 $g_i: \mathbb{R} \rightarrow \mathbb{R}, i = 1, \dots, n$, 表示神经元的特性函数, 并满足局部 Lipschitz 条件. 显然有 $D^- g_i(s) \leq D^+ g_i(s) < +\infty$, 其中符号 D^+ 和 D^- 表示 Dini 导数, 即对任意连续函数 $f: \mathbb{R} \rightarrow \mathbb{R}$, 其 Dini 导数定义为

$$D^+ f(t) = \limsup_{h \rightarrow 0^+} \frac{f(t+h) - f(t)}{h},$$

$$D^- f(t) = \liminf_{h \rightarrow 0^+} \frac{f(t+h) - f(t)}{h}.$$

(1) 式可写成如下矩阵形式

$$X' = -AX + TG(X) + C,$$

其中

$$A = \text{diag}(a_1, \dots, a_n),$$

$$G(X) = (g_1(x_1), \dots, g_n(x_n))^T$$

和

$$C = (c_1, \dots, c_n)^T.$$

当特性函数 $g_i(x)$ 分别取不同的函数如 Sigmoid 函数或 $g_i(s) = (|s+1| + |s-1|)/2$ 时, (1) 分别表示 Hopfield 网络^[6]和细胞网络^[7]等. 目前连续的 Hopfield 网络和细胞网络有许多实际应用, 如可解决某些最优化、联想记忆等问题.

若 $X^* = (x_1^*, \dots, x_n^*)^T$ 满足 $-AX^* + TG(X^*) + C = 0$, 则称 X^* 为(1)的平衡点. 对于方程(1), 联想记忆表现为对给定的初值 $X(0)$, 其解将趋于方程的某个平衡点, 趋于平衡点的快慢可决定联想记忆的速度. 本文就来对(1)的解的局部指数收敛速度问题进行研究, 对已有的结果作一些改进.

2 局部指数收敛速度(Local exponential convergence rate)

在本文中, 假设 $g_i: \mathbb{R} \rightarrow \mathbb{R}$ 是单调增加的, 且满足

$$0 \leq D^- g_i(s) \leq D^+ g_i(s) \leq M_i, \\ i = 1, \dots, n, s \in D \subset \mathbb{R} = (-\infty, +\infty), \quad (2)$$

其中, D 是某个凸子集. 显然, 对一切 $u \neq v \in D \subset$

* 基金项目: 霍英东青年教师基金和国家自然科学基金(69772027)资助项目.

收稿日期: 2000-07-21; 收修改稿日期: 2001-01-15

$\mathbb{R}, (g_i(u) - g_i(v))/(u - v) \leq M_i$. 众所周知, Hopfield 神经网络可用于联想记忆, 其平衡点可不止一个. 为了讨论 Hopfield 型神经网络(1)在某个平衡点处解的局部指数收敛速度问题, 首先需要引入描述指数收敛速度的阶的概念.

(1)的一个解在平衡点 $X^* = (x_1^*, \dots, x_n^*)^T$ 处称为至少是 λ 阶局部指数收敛的, 是指对任意正数 $\epsilon > 0$, 都存在 $\delta > 0$, 对一切初始条件 $|x_i(0) - x_i^*| < \delta (i = 1, \dots, n)$, 有

$$|x_i(t) - x_i^*| < \epsilon e^{-\lambda t}, \text{ 对一切 } t \geq 0 \text{ 和 } i.$$

显然, 由以上定义, 若阶 λ 大于零, 则解 $X(t)$ 一定是局部指数收敛的. 并且指数收敛解 $X(t)$ 的 λ 阶数越大, 则随着 t 的增加, 它收敛到平衡点处的速度就越快.

下面的两个引理取自于文献[3].

引理 1 设 $X^* = (x_1^*, \dots, x_n^*)^T$ 是(1)属于闭集

$$D = \{(x_1, \dots, x_n)^T \mid |x_i - x_i^*| \leq \alpha, i = 1, \dots, n\}$$

的一个平衡点, 其中常数 $\alpha > 0$. 若条件(2)成立, 并存在

$$\sigma \in (0, \min\{a_1, a_2, \dots, a_n\}), \quad (3)$$

使得

$$W = \left[-\frac{a_i - \sigma}{M_i} \delta_{ij} + \frac{T_{ij} + T_{ji}}{2} \right]_{n \times n},$$

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

是半负定矩阵, 那么(1)满足初始条件 $X(0) = (x_1(0), \dots, x_n(0))^T \in Q$ 的解 $X(t)$ 都属于 D , 其中 $Q = \{(x_1, \dots, x_n)^T \mid |x_i - x_i^*| < \alpha/\beta, i = 1, \dots, n\} \subset D$ 和

$$\beta = \max_{1 \leq i \leq n} \left\{ 1 + \frac{2}{2a_i - \sigma} \cdot \sum_{j=1}^n |T_{ij}| \sqrt{2M_j \sum_{r=1}^n M_r} \right\},$$

并且对一切 $t \geq 0$, 有

$$|x_i(t) - x_i^*| \leq$$

$$\|X(0) - X^*\| [e^{-\sigma t} +$$

$$\frac{2}{2a_i - \sigma} \sum_{j=1}^n |T_{ij}| \sqrt{2M_j \sum_{r=1}^n M_r} e^{-\sigma t/2}], \quad (4)$$

其中 $\|y\| = \max\{|y_1|, \dots, |y_n|\}$.

由(4)可知, 解 $X(t)$ 在平衡点 $X^* = (x_1^*, \dots, x_n^*)^T$ 处至少是 $\min(a_1, \dots, a_n, \sigma/2)$ 阶的局部指数收敛的. 因 $\sigma \in (0, \min_{1 \leq i \leq n} (a_i))$, 故上述引理 1 只是给出了解在平衡点处指数收敛的最大阶为

$$\mu_1 < \min(a_1, \dots, a_n)/2. \quad (5)$$

引理 2 设 $D \subset \mathbb{R}^n$ 是一个凸子集, $(x_1^*, \dots, x_n^*)^T$ 是某个固定点. 如果存在常数 $M_i > 0, i = 1, \dots, n$, 使得对任意 $(x_1, \dots, x_n)^T \in D$, 有 $0 \leq D^+ g_i(x_i) \leq M_i$, 则

$$\int_{x_i^*}^{x_i} (g_i(s) - g_i(x_i^*)) ds \geq (g_i(x_i) - g_i(x_i^*))^2 / 2M_i,$$

对一切

$$(x_1, \dots, x_n)^T \in D.$$

本文将给出新的局部指数收敛条件形式, 并指出条件(3)不是必要条件, 因而(5)可以不成立, 即(1)的解的指数收敛的阶可以大于 $\min(a_1, \dots, a_n)/2$.

定理 1 设 $X^* = (x_1^*, \dots, x_n^*)^T$ 是(1)属于某个开邻域 D 的一个平衡点和(2)成立. 如果存在正常数 λ 和对角矩阵 $B = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 使

$$\begin{aligned} a_i - \lambda_i M_i &\geq \lambda > 0, \text{ 对 } \lambda_i \geq 0; \\ a_i - \lambda_i D^- g_i(x_i^*) &> \lambda, \text{ 对 } \lambda_i < 0, \end{aligned} \quad (6)$$

且矩阵

$$\bar{W} = \left[-\lambda \delta_{ij} + \frac{T_{ij} + T_{ji}}{2} \right]_{n \times n},$$

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

为半负定矩阵, 那么(1)的解在平衡点 X^* 处是至少 $\min(a_1, \dots, a_n, \lambda/2)$ 阶局部指数收敛的.

证 因当 $\lambda_i < 0$ 时, $-\lambda_i \geq -\lambda_i D^- g_i(x_i^*)/M_i \geq -(a_i - \lambda)/M_i$; 当 $\lambda_i \geq 0$ 时, $-\lambda_i \geq -(a_i - \lambda)/M_i$, 故由矩阵 $\bar{W}$ 的半负定性, 知矩阵

$$\left[-\frac{a_i - \lambda}{M_i} \delta_{ij} + \frac{T_{ij} + T_{ji}}{2} \right]_{n \times n}$$

是半负定的. 取满足(3)的一个正常数 σ , 令 $\eta = \min(\sigma, \lambda)$, 则 η 满足(3)式, 并且

$$\left[-\frac{a_i - \eta}{M_i} \delta_{ij} + \frac{T_{ij} + T_{ji}}{2} \right]_{n \times n}$$

也是半负定的. 因函数

$$e^{-\eta t} + \frac{2}{2a_i - \eta} \sum_{j=1}^n |T_{ij}| \sqrt{2M_j \sum_{r=1}^n M_r} e^{-\eta t/2}$$

在区间 $(0, +\infty)$ 上是有界的, 故由引理 1 知, 存在 $\bar{\delta} > 0$, 使得当 $|x_i(0) - x_i^*| < \bar{\delta}$ 时, 对一切 $t \geq 0$, $x(t)$ 都在 D 内, 且

$$a_i - \lambda_i \frac{g_i(x_i(t)) - g_i(x_i^*)}{x_i(t) - x_i^*} \geq \lambda.$$

定义一个函数如下:

$$V(t) = \sum_{i=1}^n \int_{x_i^*}^{x_i(t)} (g_i(s) - g_i(x_i^*)) ds, \quad t \geq 0. \quad (7)$$

显然, $V(t) \geq 0$, 对其求导, 当 $|x_i(0) - x_i^*| \leq \bar{\delta}$ 时, 有

$$\begin{aligned} \frac{dV(t)}{dt} = & \sum_{i=1}^n a_i [g_i(x_i(t)) - g_i(x_i^*)] (x_i(t) - x_i^* + \\ & \sum_{j=1}^n T_{ij} [g_j(x_j(t)) - g_j(x_j^*)] - \\ & g_i(x_i^*)) [g_j(t) - g_j(x_j^*)] = \\ & - \sum_{i=1}^n a_i [g_i(x_i(t)) - g_i(x_i^*)] (x_i(t) - x_i^*) + \\ & \sum_{i=1}^n \lambda_i [g_i(x_i(t)) - g_i(x_i^*)]^2 + \\ & \sum_{i=1}^n \sum_{j=1}^n (-\lambda_i \delta_i + \frac{T_{ij} + T_{ji}}{2}) [g_i(x_i(t)) - \\ & g_i(x_i^*)] [g_j(t) - g_j(x_j^*)] \leq \\ & - \sum_{i=1}^n (a_i - \lambda_i \frac{g_i(x_i(t)) - g_i(x_i^*)}{(x_i(t) - x_i^*)}) \cdot \\ & [g_i(x_i(t)) - g_i(x_i^*)] (x_i(t) - x_i^*) \leq \\ & - \lambda \sum_{i=1}^n [g_i(x_i(t)) - g_i(x_i^*)] (x_i(t) - x_i^*) \leq \\ & - \lambda \sum_{i=1}^n \int_{x_i^*}^{x_i(t)} [g_i(x_i(s)) - g_i(x_i^*)] ds = \\ & - \lambda V(t). \end{aligned} \quad (8)$$

因此对一切 $t \geq 0$, 有 $V(t) \leq V(0)e^{-\lambda t}$. 由 $V(t)$ 的定义, 得

$$\begin{aligned} V(0) = & \sum_{i=1}^n \int_{x_i^*}^{x_i(0)} [g_i(s) - g_i(x_i^*)] ds \leq \\ & \sum_{i=1}^n [g_i(x_i(0)) - g_i(x_i^*)] (x_i(0) - x_i^*) \leq \\ & \sum_{i=1}^n M_i (x_i(0) - x_i^*)^2 \leq \\ & \|X(0) - X^*\|^2 \sum_{i=1}^n M_i. \end{aligned}$$

由引理 2, 对一切 $t \geq 0$, 有

$$\begin{aligned} V(t) \geq & \int_{x_i^*}^{x_i(t)} [g_i(s) - g_i(x_i^*)] ds \geq \\ & \frac{1}{2M_i} [g_i(x_i(t)) - g_i(x_i^*)]^2, \quad i = 1, \dots, n. \end{aligned}$$

于是, 当 $t \geq 0$ 时, 有

$$\begin{aligned} |g_i(x_i(t)) - g_i(x_i^*)| \leq \\ \|X(0) - X^*\| \sqrt{2M_i \sum_{j=1}^n M_j} e^{-\lambda t/2}, \quad (9) \\ i = 1, \dots, n. \end{aligned}$$

将(1)改写成

$$\frac{d(x_i(t) - x_i^*)}{dt} = -a_i(x_i(t) - x_i^*) + \sum_{j=1}^n T_{ij} [g_j(x_j(t)) - g_j(x_j^*)],$$

因此由(9)得

$$\begin{aligned} D^+ |x_i(t) - x_i^*| \leq \\ -a_i |x_i(t) - x_i^*| + \\ \sum_{j=1}^n |T_{ij}| |g_j(x_j(t)) - g_j(x_j^*)| \\ \leq -a_i |x_i(t) - x_i^*| + \\ \|X(0) - X^*\| \sum_{j=1}^n |T_{ij}| \sqrt{2M_j \sum_{r=1}^n M_r} e^{-\lambda t/2}. \end{aligned}$$

故对一切 $t \geq 0$, 都有

$$\begin{aligned} |x_i(t) - x_i^*| \leq \\ |x_i(t) - x_i^*| e^{-a_i t} + \|X(0) - X^*\| \cdot \\ \sum_{j=1}^n |T_{ij}| \sqrt{2M_j \sum_{r=1}^n M_r} \int_0^t e^{-a_i(t-s)} e^{-\lambda s/2} ds \leq \\ \|X(0) - X^*\| \cdot \\ [e^{-a_i t} + \frac{2}{2a_i - \lambda} \sum_{j=1}^n |T_{ij}| \sqrt{2M_j \sum_{r=1}^n M_r} e^{-\lambda t/2}]. \end{aligned} \quad (10)$$

令 $\mu_2 = \min(a_1, \dots, a_n, \lambda/2)$. 因函数

$$e^{(\mu_2 - a_i)t} + \frac{2}{2a_i - \lambda} \sum_{j=1}^n |T_{ij}| \sqrt{2M_j \sum_{r=1}^n M_r} e^{(\mu_2 - \lambda/2)t}.$$

在区间 $(0, +\infty)$ 上是有界的, 故由(10), 对任意 $\varepsilon > 0$, 都存在正数 $\delta < \bar{\delta}$, 使得当 $|x_i(0) - x_i^*| < \delta$ 时, 有

$$\begin{aligned} \|X(0) - X^*\| [e^{(\mu_2 - a_i)t} + \\ \frac{2}{2a_i - \lambda} \sum_{j=1}^n |T_{ij}| \sqrt{2M_j \sum_{r=1}^n M_r} e^{(\mu_2 - \lambda/2)t}] < \varepsilon. \end{aligned}$$

故由(10)知, 对一切 $t \geq 0$, 有 $|x_i(t) - x_i(0)| < \varepsilon e^{-\mu_2 t}$, 即(1)的解在平衡点 X^* 处的指数收敛的阶至少为 $\mu_2 = \min(a_1, \dots, a_n, \lambda/2)$. 证毕.

注 若引理 1 的条件成立, 即存在 $\sigma \in (0, \min\{a_1, \dots, a_n\})$, 使 $\bar{W}$ 为半负定矩阵, 则存在 $\lambda_i \in [0, (a_i - \sigma)/M_i]$, $i = 1, \dots, n$, 使 $a_i - \lambda_i M_i \geq \sigma > 0$, 故一定存在 $\lambda \geq \sigma$, 使(6)式成立, 且 $\bar{W}$ 为半负定矩阵. 因此由定理 1 知, (1)的解的指数收敛的阶至

少是

$$\begin{aligned}\mu_2 &= \min(a_1, \dots, a_n, \lambda/2) \geq \\ \min(a_1, \dots, a_n, \sigma/2) &= \mu_1.\end{aligned}\quad (11)$$

这样本文结果较文献[3]中的所得(1)的指数收敛的阶总不会小一些.

3 数值举例(Numerical results)

下面通过例子说明在定理 1 的(6)式中,常数 λ 要比引理 1 的(3)中的常数 σ 的取值范围要广一些,因而由本文结果可知,方程(1)的解的指数收敛的阶可以更大一些,即(11)可以为严格不等式.

例 考虑如下 Hopfield 神经网络

$$\begin{cases} \dot{x}_1 = -2x_1 - 2\ln 2g(x_1) + \ln 2g(x_2) + (17\ln 2)/6, \\ \dot{x}_2 = -0.5x_2 + \ln 2g(x_1) - 6\ln 2g(x_2) + (7\ln 2)/3. \end{cases}\quad (12)$$

其中 $g(x) = 1/(1 + e^{-x})$. 易知 $g'(x) = e^{-x}/(1 + e^{-x})^2 \in (0, 1/4]$, 故 $M = 1/4$. 显然 $a_1 = 2$ 和 $a_2 = 0.5$. 因向量 $X^* = (x_1^*, x_2^*)^T = (\ln 2, 0)^T$ 是方程(12)的一个平衡点,故 $g'(x_1^*) = g'(\ln 2) = 2/9$, $g'(x_2^*) = g'(0) = 1/4$. 矩阵

$$T = \begin{bmatrix} -2\ln 2 & \ln 2 \\ \ln 2 & -6\ln 2 \end{bmatrix} = \ln 2 \cdot \begin{bmatrix} -2 & 1 \\ 1 & -6 \end{bmatrix}$$

和

$$\left[\frac{T_{ij} + T_{ji}}{2} \right] = \ln 2 \cdot \begin{bmatrix} -2 & 1 \\ 1 & -6 \end{bmatrix}.$$

取 $\lambda_1 = \ln 2$ 和 $\lambda_2 = -5.6\ln 2$, 则容易验证矩阵

$$W = \left[-\lambda_i \delta_{ij} + \frac{T_{ij} + T_{ji}}{2} \right] = \ln 2 \cdot \begin{bmatrix} -3 & 1 \\ 1 & -0.4 \end{bmatrix}$$

是负定矩阵. 取 $\lambda = 1.4$, 则

$$a_1 - \lambda_1 M = 2 - \ln 2/4 > \lambda,$$

$$a_2 - \lambda_2 g'(x_2^*) = 0.5 + 5.6\ln 2/4 > \lambda.$$

由定理 1 知, (12)的解在平衡点 X^* 处是至少为 0.7 阶的局部指数收敛的, 而由引理 1 所能得到的局部指数收敛的阶不超过 0.25.

4 小结(Conclusion)

本文讨论了 Hopfield 型神经网络的局部指数收敛速度问题, 给出了新的指数收敛条件. 较之以往的结果, 本文得到指数收敛的阶更大, 指数收敛的速度更快. 因此, 对于分析 Hopfield 型连续神经网络的收敛问题有一定理论指导意义.

参考文献(References)

- [1] Liao Xiaoxin. On the stability of Hopfield neural networks[J]. Science in China, Series A, 1993, 23: 1025 - 1035
- [2] Liao Xiaoxin. Stability of Hopfield-type neural networks(I) [J]. Science in China, Series A, 1997, 40(8): 813 - 816
- [3] Zhang Y, Heng P A and Fu W C. Estimate of exponential convergence rate and exponential stability for neural networks[J]. IEEE Trans. Neural Networks, 1999, 10(6): 1487 - 1493
- [4] Fori M. On global asymptotic stability of a class of nonlinear systems arising in neural networks theory[J]. J. Differential Equations, 1994, 113(1): 146 - 164
- [5] Zhang Y. Global exponential stability and periodic solutions of delay Hopfield neural networks[J]. Int. J. System Science, 1996, 27(2): 227 - 331
- [6] Tank D W and Hopfield J J. Simple neural optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit[J]. IEEE Trans. Circuits System, 1986, 33(5): 533 - 541
- [7] Chua L O and Yang L. Cellular neural networks: Theory[J]. IEEE Trans. Circuits System, 1988, 35(10): 1257 - 1272
- [8] Yang H and Dillon T S. Exponential stability and oscillation of Hopfield graded response neural network[J]. IEEE Trans. Neural Networks, 1994, 5(5): 719 - 729

本文作者简介

韦 岗 1963 年生. 教授, 博士生导师. 国家自然科学基金委员会电子与信息学科评委, 中国电子学会通信学分会、电子线路与系统分会及信息论分会委员, 中国通信学会通信理论委员会委员. 主要研究方向为通信与信号处理, 神经网络等.

田俊俊 1964 年生. 副教授, 在读博士生. 主要研究方向为神经网络、微分和差分方程稳定性等.