

一类反馈神经网络对非线性连续系统近似能力的证明

李晓东, 梁晓波

(宁夏大学物理与电气信息工程系, 银川 750021)

摘 要: 本文从多层前馈神经网络的一般近似定理出发, 证明了带有输入的非线性连续系统在有限时间段内的输出轨迹可以被一类反馈神经网络输出神经元的状态向量近似到任何程度。

关键词: 非线性连续系统; 反馈神经网络; 近似

中图分类号: TP18 **文献标识码:** A **文章编号:** 0372-2112 (2001) 01-0103-03

Proof of the Approximation Ability of Recurrent Neural Networks to Nonlinear Continuous-Time System with Input

LI Xiao-dong, LIANG Xiao-bo

(Dept. of Physics & Electrical Information Engineering, Ningxia University, Yinchuan 750021, China)

Abstract: Starting from the universal approximation theorem of multilayered feed forward neural networks, this paper proves that the finite time trajectory of nonlinear continuous-time system with input can be approximated by the state vector of the output units of a class of recurrent neural networks.

Key words: nonlinear continuous-time system; recurrent neural networks; approximation

1 引言

近些年, 人工神经网络理论在控制、滤波和信号处理等领域得到了广泛的应用^[1,2]. 这些应用的理论基础之一就是人工神经网络对非线性系统的近似能力. 文章^[3]证明了反馈神经网络对不带输入的非线性连续系统 $\dot{x} = F(x)$ 的近似能力. 本文将把文^[3]的结论推广到更一般的情形. 无疑, 这对丰富神经网络的近似理论具有重大的意义.

2 一类连续型反馈神经网络

本文考虑含有 L 个神经元的如下形式的一类反馈神经网络

$$\dot{z} = -\alpha z + W_1 \sigma(z + W_2 u) \quad (1)$$

其中 $z \in R^L$ 和 $u \in R^m$ 分别是 L 个神经元的状态向量和该神经网络的输入向量; $W_1 \in R^{L \times L}$ 是各神经元之间的互连接权矩阵; $W_2 \in R^{L \times m}$ 是输入信号对各神经元的连接权矩阵; $0 < \alpha < 1$ 是神经元状态的延迟因子; $\sigma(\cdot)$ 是一个 C^1 -sigmoid 向量函数.

本文将证明: 神经网络式(1)的权矩阵经合适选取后, 非线性连续系统 $\dot{x} = F(x, u)$ 在有限时间段内的输出轨迹可以被神经网络式(1)的输出神经元的状态任意逼近.

3 反馈神经网络对非线性连续系统的近似

引理: 令 $S \subset R^n$ 和 $U \subset R^m$ 是开集, $F, \tilde{F}: S \times U \rightarrow R^n$ 是 Lipschitz 连续映射, L 是 $F(x, u)$ 在 $S \times U$ 上对 x 的 Lipschitz 常数, 且对于所有的 $x \in S$ 和 $u \in U$ 满足

$$F(x, u) - \tilde{F}(x, u) < \varepsilon \quad (2)$$

如果 $x(t)$ 和 $\tilde{x}(t)$ 分别是方程

$$\dot{x} = F(x, u) \text{ 和 } \dot{\tilde{x}} = \tilde{F}(\tilde{x}, u) \quad (3)$$

的在区间 J 上满足初始条件 $x(t_0) = \tilde{x}(t_0)$ 的解, 那么

$$|x(t) - \tilde{x}(t)| \leq \frac{\varepsilon}{L} (\exp L|t - t_0| - 1) \quad t \in J \quad (4)$$

该引理的证明相似于文献^[4]中第十五章中定理3的证明.

定理: 令 $S \subset R^n$ 和 $U \subset R^m$ 是开集, $X \subset S$ 和 $D_U \subset U$ 是紧集, $F: S \times U \rightarrow R^n$ 是 C^1 类向量函数. 非线性连续系统

$$\dot{x}(t) = F(x(t), u(t)), \quad x \in S, u \in U, \quad t \in [0, T], (0 < T < \infty) \quad (5)$$

只要其初始条件 $x(0) \in X$, 则对于任意的 $\varepsilon > 0$ 和任意的有界输入 $u(t) \in D_U, t \in [0, T]$, 都存在着一个整数 N 和一个具有适当初始条件 $z(0)$ 的反馈神经网络式(1)使得

$$\max_{t \in [0, T]} |x - p| < \varepsilon \quad (6)$$

这里 $z = (p, h)^T$, $p \in R^n$, $h \in R^N$, p 是反馈神经网络式(1)输出神经元的状态向量.

证明: $F: S \times U \rightarrow R^n$ 是 C^1 类向量函数. 对于任何初始条件 $x(0) = X$, 由于非线性连续系统式(5)的输出轨迹的连续性和有界性, 在区间 $[0, T]$ 上其输出轨迹点的集合 D_S 是 S 的一个紧子集.

令 λ_1 是紧集 D_S 和开集 S 的边界 $\partial(S)$ 间的距离, λ_2 是紧集 D_U 和开集 U 的边界 $\partial(U)$ 间的距离. 对于给定的 $\varepsilon > 0$, 取 η 满足 $0 < \eta < \min\{\varepsilon, \lambda_1, \lambda_2\}$. 再令

$$\begin{aligned} S_\eta &= \{x \in R^n \mid \exists x_1 \in D_S, \|x - x_1\| \leq \eta\} \\ U_\eta &= \{u \in R^m \mid \exists u_1 \in D_U, \|u - u_1\| \leq \eta\} \end{aligned} \quad (7)$$

则 S_η 和 U_η 也是紧集, 并且 $D_S \subset S_\eta \subset S$, $D_U \subset U_\eta \subset U$. $F(x, u)$ 在 $S_\eta \times U_\eta$ 上对 x 是 Lipschitz 的. 令 L_F 是 F 在 $S_\eta \times U_\eta$ 上关于 x 的 Lipschitz 常数. 取 ε_1 满足

$$0 < \varepsilon_1 < \frac{\eta L_F}{2(\exp(L_F T) - 1)} \quad (8)$$

在 $S_\eta \times U_\eta$ 上本文将证明这个定理成立.

根据多层前馈神经网络的一般近似定理^[5], 对于任意的常数 α , 存在着整数 N, N 维阈值向量 θ 和矩阵 $A \in R^{n \times N}$, $B_1 \in R^{N \times n}$, $B_2 \in R^{N \times m}$ 使下式成立

$$\max_{x \in S_\eta, u \in U_\eta} F(x, u) + \alpha x - A\sigma(B_1 x + B_2 u + \theta) < \varepsilon_1 \quad (9)$$

这里 $\sigma(\cdot)$ 是一个 C^1 -sigmoid 向量函数, 特意选取 α 满足下式:

$$\alpha\theta < \frac{\eta L_{\tilde{G}}}{2(\exp(L_{\tilde{G}} T) - 1)} \text{ 和 } |\alpha| < L_{\tilde{G}}/2 \quad (10)$$

上式中的 $L_{\tilde{G}}$ 是以后要定义的映射 $\tilde{G}: R^{n+N} \times R^m \rightarrow R^{n+N}$ 的 Lipschitz 常数.

令 x 和 $\tilde{x}$ 分别为方程 $\dot{x} = F(x, u)$ 和

$$\dot{\tilde{x}} = -\alpha\tilde{x} + A\sigma(B_1\tilde{x} + B_2u + \theta) \quad (11)$$

的满足初始条件 $x(0) = \tilde{x}(0) = x_0 \in X$ 的解. 根据前面的引理及式(8), 有

$$\max_{t \in [0, T]} \|x(t) - \tilde{x}(t)\| \leq \frac{\varepsilon_1}{L_F} (\exp(L_F T) - 1) < \frac{\eta}{2} \quad (12)$$

令 $\tilde{y} = B_1\tilde{x} + \theta$, $\tilde{z} = (\tilde{x}, \tilde{y})^T \in R^{n+N}$, 则由式(11)可推出

$$\dot{\tilde{z}} = \tilde{G}(\tilde{z}, u) = -\alpha\tilde{z} + W_1\sigma(\tilde{z} + W_2u) + \alpha\theta_1 \quad (13)$$

这里 $W_1 = \begin{pmatrix} 0 & A \\ 0 & B_1 A \end{pmatrix} \in R^{(n+N) \times (n+N)}$, $W_2 = \begin{pmatrix} 0 \\ B_2 \end{pmatrix} \in R^{(n+N) \times m}$, $\theta_1 = \begin{pmatrix} 0 \\ \theta \end{pmatrix} \in R^{n+N}$. 方程(13)的解向量的前 n 个元素等价于方程(11)的解.

因为 $\sigma(\cdot)$ 是一个 C^1 -sigmoid 函数, 映射 $W_1\sigma(\tilde{z} + W_2u)$ 在 $R^{n+N} \times R^m$ 上对 $\tilde{z}$ 是 Lipschitz 的. 设 $L_{\tilde{G}}/2$ 是 $W_1\sigma(\tilde{z} + W_2u)$ 的 Lipschitz 常数. 另外, 由关于常数 α 的条件(10), $L_{\tilde{G}}/2$ 也是 $-\alpha\tilde{z}$ 的 Lipschitz 常数. 所以, $L_{\tilde{G}}$ 是 $\tilde{G}(\tilde{z}, u) = -\alpha\tilde{z} + W_1\sigma(\tilde{z} + W_2u) + \alpha\theta_1$ 在 $R^{n+N} \times R^m$ 上的对 $\tilde{z}$ 的 Lipschitz 常数.

考虑反馈神经网络系统

$$\dot{z} = G(z, u) = -\alpha z + W_1\sigma(z + W_2u) \quad (14)$$

其中 $z = (p, h)^T$, $p \in R^n$ 是其 n 个输出神经元的内部状态; $h \in R^N$ 是其 N 个隐藏神经元的内部状态.

利用 $\tilde{G}(\tilde{z}, u)$, $G(z, u)$ 的表达式(13), (14) 和关于 α 的条件式(10), 对于任何 $z \in R^{n+N}$ 有

$$\|\tilde{G}(z, u) - G(z, u)\| \leq \alpha\theta < \frac{\eta L_{\tilde{G}}}{2(\exp(L_{\tilde{G}} T) - 1)} \quad (15)$$

因此, 设 z 和 $\tilde{z}$ 分别为反馈神经网络(14)和系统(13)对于输入 u 的解, 取反馈神经网络(14)的初始条件和系统(13)的初始条件相等, 即

$$p(0) = x(0) \text{ 和 } h(0) = B_1 x(0) + \theta \quad (16)$$

那么由引理和式(15), 得

$$\max_{t \in [0, T]} \|\tilde{x} - p\| \leq \max_{t \in [0, T]} \|\tilde{z} - z\| \leq \eta/2 \quad (17)$$

x 是非线性系统 $\dot{x} = F(x, u)$ 在区间 $[0, T]$ 上的输出轨迹, 由式(12)和(17), 得

$$\begin{aligned} \max_{t \in [0, T]} \|x - p\| &\leq \max_{t \in [0, T]} \|x - \tilde{x}\| + \max_{t \in [0, T]} \|\tilde{x} - p\| \\ &\leq \eta/2 + \eta/2 = \eta < \varepsilon \end{aligned} \quad (18)$$

证毕

4 两个推论

当非线性连续系统 $\dot{x} = F(x, u)$ 中的 $F(x, u)$ 具有一些特殊结构时, 它的输出轨迹可以用更简单结构形式的反馈神经网络去逼近.

推论 1: 令 $S \subset R^n$ 是开集, $X \subset S$ 是紧集, $F: S \times R^m \rightarrow R^n$ 是 C^1 类向量函数. 非线性连续系统

$$\dot{x}(t) = F(x(t), t), x \in S, t \in [0, T], (0 < T < \infty) \quad (19)$$

只要其初始条件 $x(0) \in X$, 则对于任意的 $\varepsilon > 0$, 都存在着一个整数 N 和一个具有适当初始条件 $z(0)$ 的如下形式的反馈神经网络

$$\dot{z} = -\alpha z + W_1\sigma(z), z \in R^l \quad (20)$$

使得

$$\max_{t \in [0, T]} \|x - p\| < \varepsilon \quad (21)$$

这里 $z = (p, h)^T$, $p \in R^n$, $h \in R^N$, p 是反馈神经网络式(20)输出神经元的状态向量.

证明: 令 $\bar{x} = (x, t)^T \in R^{n+1}$, 则非线性连续系统式(19)能被表示为

$$\dot{\bar{x}} = \bar{F}(\bar{x}) \quad (22)$$

这里 $\bar{F}(\bar{x}) = (F(x, t), 1)^T$, 初始条件为 $\bar{x}(0) = (x(0), 0)^T$. 当 $x(0) \in X$ 时, $\bar{x}(0)$ 属于 R^{n+1} 的紧子集. 系统式(22)为不带输入的非线性连续系统. 根据上面的定理, 用来近似式(22)的反馈神经网络(1)中的权矩阵 $W_2 = 0$, 即存在形如式(20)的反馈神经网络, 使得

$$\max_{t \in [0, T]} \|x - p\| < \varepsilon$$

推论 2: 令 $X \subset R^n$ 是紧集, $F: R \rightarrow R^n$ 是 C^1 类向量函数. 非线性连续系统

$$\dot{x}(t) = F(t), t \in [0, T], (0 < T < \infty) \quad (23)$$

只要其初始条件 $x(0) \in X$. 则对于任意的 $\varepsilon > 0$, 都存在着一个整数 N 和一个具有适当初始条件 $z(0)$ 的形如式(20)的反

神经网络使得 $\max_{t \in [0, T]} \|x - p\| < \varepsilon$.

证明:令 $\bar{x} = (x, t) \in R^{n+1}$, 则非线性连续系统式 (23) 能被表示为 $\dot{\bar{x}} = \bar{F}(\bar{x})$, 这里 $\bar{F}(\bar{x}) = (F(t), 1)'$, 初始条件为 $\bar{x}(0) = (x(0), 0)'$. 以下证明思路同推论 1.

5 结论

本文证明了带有输入的非线性连续系统在有限时间段内的输出轨迹可以被一类反馈神经网络输出神经元的状态向量任意逼近. 所用的证明方法是一种构造性的方法, 它从多层前馈神经网络的一般近似定理出发, 将一个低维的连续系统嵌入一个反馈神经网络, 本文得出的结果极大地丰富了神经网络的近似理论.

参考文献:

- [1] Kiyotoshi Matsuoka, Masahiro Ohya, and Mitsuru Kawamoto. A neural net for blind separation of nonstationary signals [J]. Neural Networks, 1995, 8(3): 411 - 419.
- [2] Tommy W. S. Chow, Xiao-Dong LI, and Yong FANG. A real-time learning control approach for nonlinear continuous-time system using recurrent neural networks [J]. IEEE Trans. on Industrial Electronics, 2000, 47(2): 478 - 486.
- [3] K. Funahashi and Y. Nakamura. Approximation of dynamical systems by continuous time recurrent neural networks [J]. Neural Networks, 1993, 6: 801 - 806.
- [4] M. W. Hirsch and S. Smale. Differential equations, dynamical systems, and linear algebra [M]. New York: Academic Press, 1974.
- [5] K. Hornik, M. Stinchcombe, and H. White. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks [J]. Neural Networks, 1990, 3(6): 551 - 560.

作者简介:



李晓东 1966 年生, 1990 年 2 月在南京理工大学自动控制专业获工学硕士学位. 现为宁夏大学物理与电气信息工程系副教授. 已在国内外学术刊物上发表论文十余篇. 主要研究方向为人工神经网络、2-D 线性系统理论、迭代学习控制等.



梁晓波 1964 年生, 1992 年 7 月在大连理工大学系统工程专业获工学硕士学位. 现为宁夏大学物理与电气信息工程系副教授. 主要研究方向为自动控制, 计算机应用等.