

文章编号: 1001-0920(2013)09-1423-04

## 基于单网络 ADP 的一类未知非线性系统的近似最优控制

崔黎黎, 刘杰, 张勇

(沈阳师范大学 科信软件学院, 沈阳 110034)

**摘要:** 针对一类未知的连续非线性系统, 提出一个基于单网络近似动态规划 (ADP) 的近似最优控制方案. 该方案通过设计一个新型的递归神经网络 (RNN) 辨识器放松了系统模型需已知或部分已知的要求, 并利用一个神经网络 (NN) 近似系统的性能指标函数消除了常规 ADP 方法中的控制网络. 通过 Lyapunov 理论分析严格证明了闭环系统内所有信号一致最终有界, 并且所获得的性能指标函数和控制输入分别收敛到最优性能指标函数和最优控制输入的小邻域内. 仿真结果验证了所提出控制方案的有效性.

**关键词:** 未知非线性系统; 递归神经网络; 近似动态规划; 自适应; 最优控制

中图分类号: TP273

文献标志码: A

## Near-optimal control of a class of unknown nonlinear systems based on single network ADP

CUI Li-li, LIU Jie, ZHANG Yong

(Kexin Software College, Shenyang Normal University, Shenyang 110034, China. Correspondent: CUI Li-li, E-mail: cuilili8396@163.com)

**Abstract:** The near-optimal control scheme is proposed for a class of unknown continuous-time nonlinear systems based on single network approximate dynamic programming (ADP). The proposed scheme relaxes the requirement of the system model being known or partly known by designing a novel recurrent neural network (RNN) identifier, and eliminates the action network of ordinary ADP methods by employing a neural network (NN) to approximate the performance index function. By Lyapunov theory, it is proved that all the signals in the closed-loop system are ultimately uniformly bounded and the obtained optimal performance index function and control input lie in small neighborhoods of the optimal performance index function and the optimal control input, respectively. Simulation results demonstrate the effectiveness of the proposed scheme.

**Key words:** unknown nonlinear system; recurrent neural network; approximate dynamic programming; adaptive; optimal control

### 0 引言

为了解决动态规划的“维数灾”问题, Werbos 于 1977 年首次提出了近似动态规划 (ADP) 方法. 其基本思想是利用函数近似结构逼近动态规划方程中的性能指标函数和控制策略, 以满足最优性原理从而获得最优性能指标函数和最优控制. 由于近似动态规划成功地解决了传统动态规划问题并避免了动态规划“维数灾”的产生, 近几年来受到广大学者越来越广泛的关注, 已取得了一定的理论研究成果<sup>[1-4]</sup>, 并在飞行器最优控制<sup>[5]</sup>、电力系统<sup>[6]</sup>等方面获得了成功应用. 然而值得注意的是, 基于近似动态规划方法的现有的大多数研究成果都要求系统模型已知, 这使得其应

用具有一定局限性. 因此, 开始有一些研究致力于在无法获得系统先验知识的情况下寻求最优控制, 如文献 [7-8]. 然而, 上述方法均难以解决未知的一般连续非线性系统的最优控制问题.

为了解决这个难题, 本文提出了一个基于单网络 ADP 的近似最优控制方案. 首先提出一个新型的递归神经网络 (RNN) 辨识器用于辨识未知的连续非线性系统, 并严格证明了所提出的 RNN 辨识器能够保证辨识误差渐近收敛到零, 放松了系统模型需已知或部分已知的要求; 然后利用辨识得到的 RNN 模型设计基于单网络 ADP 的近似最优控制器, 消除了常规 ADP 方法中的控制网络, 降低了计算复杂度; 最后,

收稿日期: 2012-04-16; 修回日期: 2012-12-13.

基金项目: 国家自然科学基金项目 (60970112).

作者简介: 崔黎黎 (1983—), 女, 讲师, 从事非线性系统、近似动态规划等研究; 刘杰 (1957—), 男, 教授, 从事工业控制的研究.

通过 Lyapunov 理论严格证明了所提出的控制方案能够保证闭环系统内所有信号一致最终有界, 并且所获得的性能指标函数和控制输入分别收敛到最优性能指标函数和最优控制输入的小邻域内.

## 1 问题描述

考虑如下的一般连续非线性系统:

$$\dot{x} = h(x, u). \quad (1)$$

其中:  $x = (x_1, x_2, \dots, x_n)^T \in \mathbf{R}^n$  为状态向量,  $u = (u_1, u_2, \dots, u_m)^T \in \mathbf{R}^m$  为输入向量, 连续非线性光滑函数  $h(x, u)$  为系统未知的非线性动态. 本文的主要目的是为系统(1)设计一个最优控制  $u^*$  最小化如下的无限时间性能指标函数:

$$V(x) = \int_t^\infty r(x(\tau), u(\tau)) d\tau. \quad (2)$$

其中:  $r(x, u) = Q(x) + u^T R u$  为效用函数,  $Q(x) > 0$ ,  $R$  为对称正定矩阵. 常规的 ADP 方法均要求被控对象模型完全已知或部分已知, 难以解决未知非线性系统(1)的最优控制问题. 为了解决这个难题, 本文提出了一个新型的基于单网络 ADP 的近似最优控制方案.

## 2 基于单网络 ADP 的近似最优控制

### 2.1 RNN 辨识器设计

根据神经网络的万能逼近性, 系统动态(1)可重写为如下的 RNN 形式:

$$\dot{x} = C^{*T} x + A^{*T} f(x) + C_u^{*T} u + A_u^{*T} \varepsilon. \quad (3)$$

其中:  $C^{*T}$ ,  $A^{*T}$ ,  $C_u^{*T}$  和  $A_u^{*T}$  为理想权值矩阵;  $\varepsilon$  为近似误差; 激活函数  $f(x)$  为一类单调递增的有界函数, 对任意  $x, y \in \mathbf{R}$  和  $x \geq y, k > 0$ , 满足  $0 \leq f(x) - f(y) \leq k(x - y)$ , 如  $f(x) = \tanh(x)$ .

**假设 1**<sup>[11]</sup>  $\varepsilon^T \varepsilon \leq \varepsilon_M = \lambda^* e_m^T e_m$ , 其中  $\lambda^* > 0$ .

基于式(3), 构造如下的 RNN 辨识器:

$$\begin{aligned} \dot{\hat{x}} &= \hat{C}^T \hat{x} + \hat{A}^T f(\hat{x}) + \hat{C}_u^T u + \\ &\hat{A}_u^T - S e_m - \frac{\hat{\lambda} e_m}{e_m^T e_m + \eta}. \end{aligned} \quad (4)$$

其中:  $\hat{x}$  为  $x$  的估计;  $\hat{C}, \hat{A}, \hat{C}_u$  和  $\hat{A}_u$  分别为  $C^*, A^*, C_u^*$  和  $A_u^*$  的估计;  $e_m = x - \hat{x}$  为系统建模误差;  $S$  为设计矩阵;  $\hat{\lambda}$  为自适应参数;  $\eta > 1$  为设计参数. 设计权值及自适应参数的调节律为

$$\begin{aligned} \dot{\hat{C}} &= \Gamma_1 \hat{x} e_m^T, \quad \dot{\hat{A}} = \Gamma_2 f(\hat{x}) e_m^T, \quad \dot{\hat{C}}_u = \Gamma_3 u e_m^T, \\ \dot{\hat{A}}_u &= \Gamma_4 e_m^T, \quad \dot{\hat{\lambda}} = -\Gamma_5 \frac{e_m^T e_m}{e_m^T e_m + \eta}, \end{aligned} \quad (5)$$

其中  $\Gamma_i (i = 1, 2, \dots, 5)$  为正定对称矩阵.

由式(3)和(4), 可得建模误差动态如下:

$$\begin{aligned} \dot{e}_m &= \\ C^{*T} e_m + \tilde{C}^T \hat{x} + A^{*T} \tilde{f}(e_m) + \tilde{A}^T f(\hat{x}) + \tilde{C}_u^T u + \end{aligned}$$

$$\tilde{A}_u^T + \varepsilon + S e_m - \frac{\tilde{\lambda} e_m}{e_m^T e_m + \eta} + \frac{\lambda^* e_m}{e_m^T e_m + \eta}. \quad (6)$$

其中

$$\begin{aligned} \tilde{C} &= C^* - \hat{C}, \quad \tilde{A} = A^* - \hat{A}, \quad \tilde{C}_u = C_u^* - \hat{C}_u, \\ \tilde{A}_u &= A_u^* - \hat{A}_u, \quad \tilde{f}(e_m) = f(x) - f(\hat{x}), \quad \tilde{\lambda} = \lambda^* - \hat{\lambda}. \end{aligned}$$

**定理 1** 对于 RNN 辨识器(4), 如果权值调节律为式(5), 设计矩阵  $S$  满足

$$S < -\left(C^{*T} + \frac{1}{2} A^{*T} A^* + \left(\frac{1}{2} + \frac{3}{2} \lambda^* + \frac{1}{2} k^2\right) I_n\right), \quad (7)$$

则当  $t \rightarrow \infty$  时, 建模误差  $e_m$  渐近收敛到零, 其中  $I_n$  为  $n \times n$  维单位阵.

**证明** 选取如下的 Lyapunov 函数:

$$\begin{aligned} J &= \frac{1}{2} e_m^T e_m + \frac{1}{2} \text{tr} \{ \tilde{C}^T \Gamma_1^{-1} \tilde{C} + \tilde{A}^T \Gamma_2^{-1} \tilde{A} + \\ &\tilde{C}_u^T \Gamma_3^{-1} \tilde{C}_u + \tilde{A}_u^T \Gamma_4^{-1} \tilde{A}_u + \tilde{\lambda}^T \Gamma_5^{-1} \tilde{\lambda} \}, \end{aligned} \quad (8)$$

对  $J$  沿式(6)求导, 代入式(5)和(6), 并利用激活函数的性质和假设 1, 可得

$$\begin{aligned} \dot{J} &\leq e_m^T C^{*T} e_m + \frac{1}{2} e_m^T A^{*T} A^* e_m + \frac{e_m^T \lambda^* e_m}{e_m^T e_m + \eta} + \\ &e_m^T \left( \left( \frac{1}{2} + \frac{1}{2} \lambda^* + \frac{1}{2} k^2 \right) I_n + S \right) e_m \leq \\ &e_m^T \Xi e_m. \end{aligned} \quad (9)$$

如果条件(7)满足, 则有  $\dot{J} < 0$ . 因为  $J > 0$ , 由文献[9]可得, 当  $t \rightarrow \infty$  时,  $e_m \rightarrow 0$ .  $\square$

根据定理 1, 当  $t \rightarrow \infty$  时,  $e_m \rightarrow 0$ , 相应地有  $\hat{C}, \hat{A}, \hat{C}_u$  和  $\hat{A}_u$  分别趋向于常数矩阵  $C, A, C_u$  和  $A_u$ . 根据式(4), 非线性系统(1)可重写为

$$\dot{x} = C^T x + A^T f(x) + C_u^T u + A_u^T. \quad (10)$$

此时问题可以转化为: 针对未知非线性系统(1), 利用辨识器辨识得到的系统模型(10), 设计最优控制  $u^*$  最小化性能指标函数(2).

### 2.2 基于单网络 ADP 的近似最优控制器设计

定义如下的 Hamiltonian 函数:

$$\begin{aligned} H(x, u, V_x) &= V_x^T (C^T x + A^T f(x) + C_u^T u + A_u^T) + \\ &Q(x) + u^T R u, \end{aligned} \quad (11)$$

其中  $V_x = \partial V(x) / \partial x$ .

根据最优控制的一阶必要条件  $\partial H(x, u, V_x) / \partial u = 0$ , 可求得最优控制  $u^*$  为

$$u^* = -\frac{1}{2} R^{-1} C_u V_x^*, \quad (12)$$

其中  $V_x^* = \partial V^*(x) / \partial x$ ,  $V^*$  为最优性能指标函数. 将式(12)代入(11)可得如下的 HJB (Hamilton-Jacobi-Bellman) 方程:

$$\begin{aligned} V_x^{*T} (C^T x + A^T f(x) + A_u^T) + Q(x) - \\ V_x^{*T} C_u^T R^{-1} C_u V_x^* / 4 = 0. \end{aligned}$$

如果  $V^*(x)$  能够从HJB方程中获得, 则根据式(12)可直接计算最优控制  $u^*$ . 然而HJB方程是非线性偏微分方程, 很难直接求解, 因此本文提出利用单网络ADP方法近似求解HJB方程, 然后根据式(12)获得近似最优控制.

采用如下的NN(即评价网)近似  $V(x)$ , 即

$$V(x) = W_1^T \varphi_1(x) + \varepsilon_1. \quad (13)$$

其中:  $W_1$  为理想权值,  $\varphi_1(x)$  为激活函数向量,  $\varepsilon_1$  为近似误差. 将式(13)对  $x$  求导可得  $V_x = \nabla \varphi_1^T W_1 + \nabla \varepsilon_1$ ,  $\nabla \varphi_1 \triangleq \partial \varphi_1(x) / \partial x$ ,  $\nabla \varepsilon_1 \triangleq \partial \varepsilon_1 / \partial x$ . 利用式(12)可得

$$u^* = -\frac{1}{2} R^{-1} C_u \nabla \varphi_1^T W_1 - \frac{1}{2} R^{-1} C_u \nabla \varepsilon_1. \quad (14)$$

将式(14)代入HJB方程, 可得

$$W_1^T \nabla \varphi_1 (C^T x + A^T f(x) + A_u^T) + Q(x) - \frac{1}{4} W_1^T \nabla \varphi_1 C_u^T R^{-1} C_u \nabla \varphi_1^T W_1 + \varepsilon_{HJB} = 0. \quad (15)$$

其中

$$\varepsilon_{HJB} = \nabla \varepsilon_1^T \dot{x}_1^* + \nabla \varepsilon_1^T C_u R^{-1} C_u \nabla \varepsilon_1 / 4, \\ \dot{x}_1^* = C^T x + A^T f + C_u u^* + A_u^T.$$

令  $\hat{W}_1$  为  $W_1$  的估计, NN的实际输出  $\hat{V}(x) = \hat{W}_1^T \varphi_1(x)$ . 最优控制  $u^*$  的估计为

$$\hat{u} = -\frac{1}{2} R^{-1} C_u \nabla \varphi_1^T \hat{W}_1. \quad (16)$$

将式(16)代入(10), 可得

$$\dot{x} = C^T x + A^T f(x) - \frac{1}{2} C_u^T R^{-1} C_u \nabla \varphi_1^T \hat{W}_1 + A_u^T. \quad (17)$$

设计NN权值调节律为

$$\dot{\hat{W}}_1 = -\alpha_1 \frac{\sigma}{(\sigma^T \sigma + 1)^2} \left( \hat{W}_1^T \nabla \varphi_1 (C^T x + A^T f(x) + A_u^T) + Q(x) - \frac{1}{4} W_1^T \nabla \varphi_1 C_u^T R^{-1} C_u \nabla \varphi_1^T W_1 \right) + N(x, \hat{u}) \frac{\alpha_2}{2} \nabla \varphi_1 C_u^T R^{-1} C_u L_x. \quad (18)$$

其中

$\sigma = \nabla \varphi_1 (C^T x + A^T f(x) + A_u^T - C_u^T R^{-1} C_u \nabla \varphi_1^T \hat{W}_1 / 2)$ ;  $\alpha_1 > 0, \alpha_2 > 0$  为设计参数;  $L_x = \partial L(x) / \partial x$ ,  $L(x)$  为径向无界且连续可导的Lyapunov函数, 并满足  $\dot{L}(x) = L_x \dot{x}^* < 0$ ;

$$N(x, \hat{u}) = \begin{cases} 0, & L_x^T (C^T x + A^T f(x) + A_u^T - C_u^T R^{-1} C_u \nabla \varphi_1^T \hat{W}_1 / 2) < 0; \\ 1, & \text{else.} \end{cases} \quad (19)$$

定义  $\tilde{W}_1 = W_1 - \hat{W}_1$ , 利用式(15)可得

$$\dot{\tilde{W}}_1 = -\alpha_1 \frac{1}{\sigma^2} \left( \nabla \varphi_1 \left( \dot{x}^* + \frac{C_u^T R^{-1} C_u \nabla \varepsilon_1}{2} \right) + \right.$$

$$\left. \frac{\nabla \varphi_1 C_u^T R^{-1} C_u \nabla \varphi_1^T \tilde{W}_1}{2} \right) \left( \tilde{W}_1^T \nabla \varphi_1 \left( \dot{x}^* + \frac{C_u^T R^{-1} C_u \nabla \varepsilon_1}{2} \right) + \frac{\tilde{W}_1^T \nabla \varphi_1 C_u^T R^{-1} C_u \nabla \varphi_1^T \tilde{W}_1}{4} \right) - N(x, \hat{u}) \frac{\alpha_2}{2} \nabla \varphi_1 C_u^T R^{-1} C_u L_x, \quad (20)$$

其中  $\bar{\sigma} = (\sigma^T \sigma + 1)$ .

**假设2**  $\|W_1\| \leq W_{1M}, \|\varepsilon_1\| \leq \varepsilon_{1M}, \|\nabla \varphi_1\| \leq \varphi_{1M}, \|\nabla \varepsilon_1\| \leq \varepsilon_{1M}, \|\varepsilon_{HJB}\| \leq \varepsilon_{HJBM}$ .

**假设3** 存在一个函数  $\beta(x)$  使得  $\|\dot{x}_1^*\| \leq \beta(x)$ .

本文取  $\beta(x) = \sqrt[4]{K^* \|L_x\|}$ , 其中  $K^*$  为正常数.

**引理1** 存在一个正定矩阵  $\bar{Q}(x)$  使得  $L_x^T \dot{x}^* = L_x^T \bar{Q}(x) L_x$ .

**定理2** 对于系统(10)和控制器(16), 如果NN权值调节律为(18), 并假设系统状态满足持续激励条件, 则闭环所有信号均一致最终有界, 且性能指标函数  $\hat{V}(x)$  收敛到最优性能指标函数  $V(x)$  的  $\varepsilon_V$  邻域内, 控制输入  $\hat{u}$  收敛到最优控制输入  $u^*$  的  $\varepsilon_u$  邻域内, 即当  $t \rightarrow \infty$  时,  $\|\hat{V}(x) - V(x)\| \leq \varepsilon_V, \|\hat{u} - u^*\| \leq \varepsilon_u$ .

**证明** 选取如下的Lyapunov函数:

$$L_n = \alpha_2 L(x) + \frac{1}{2} \text{tr} \{ \tilde{W}_1^T \tilde{W}_1 \}. \quad (21)$$

对式(21)沿(17)求导, 并代入(20), 整理可得

$$\dot{L}_n \leq \alpha_2 L_x^T (C^T x + A^T f(x) - \frac{1}{2} C_u^T R^{-1} C_u \nabla \varphi_1^T \hat{W}_1 + A_u^T) - N(x, \hat{u}) \frac{\alpha_2}{2} \tilde{W}_1^T \nabla \varphi_1 C_u^T R^{-1} C_u L_x - \frac{\alpha_1}{\bar{\sigma}^2} \|\tilde{W}_1\|^4 \delta_1 + \frac{\alpha_1}{\bar{\sigma}^2} \delta_2 + \frac{\alpha_1}{\bar{\sigma}^2} \delta_3 \beta^4(x). \quad (22)$$

其中

$$\delta_1 = \nabla \varphi_{1m}^4 \|C_u^T R^{-1} C_u\|^2 / 64, \\ \delta_2 = 1024 / \|C_u^T R^{-1} C_u\|^2 + 3/2, \\ \delta_3 = 64 / \|C_u^T R^{-1} C_u\|^2 + 3(\nabla^4 \varepsilon_{1M} + \nabla^4 \varepsilon_{1M} \|C_u^T R^{-1} C_u\|^2) / 2.$$

常数  $\nabla \varphi_{1m}$  满足  $0 < \nabla \varphi_{1m} \leq \|\nabla \varphi_1\|$ , 其存在性可由系统状态满足持续激励条件保证. 根据式(19)中  $N(x, \hat{u})$  的定义, 存在以下两种情况:

1) 当  $N(x, \hat{u}) = 0$  时, 由式(19)可知式(22)中的第1项小于0, 根据假设2和  $\|1/\bar{\sigma}^2\| \leq 1$ , 可得

$$\dot{L}_n \leq -(\alpha_2 \dot{x}_m - \alpha_1 \delta_3 K^*) \|L_x\| - \alpha_1 \|\tilde{W}_1\|^4 \delta_1 / \bar{\sigma}^2 + \alpha_1 \delta_2 / \bar{\sigma}^2.$$

选取  $\alpha_1$  和  $\alpha_2$  使得  $\alpha_2 \dot{x}_m - \alpha_1 \delta_3 K^* > 0$ , 则只要不等式  $\|L_x\| > \alpha_1 \delta_2 / (\alpha_2 \dot{x}_m - \alpha_1 \delta_3 K^*) \equiv B_{L0}$  或  $\|\tilde{W}_1\| > \sqrt[4]{\delta_2 / \delta_1} \equiv B_{W0}$  成立, 就有  $\dot{L}_n < 0$ . 注意到, 系统状态满足持续激励条件和式(19), 能够保证存在一个常数  $\dot{x}_m$  使得  $0 < \dot{x}_m < \dot{x}$ . 根据标准Lyapunov扩展定理<sup>[10]</sup>,  $\|L_x\|$  和  $\|\tilde{W}_1\|$  均一致最终有界.

2) 当  $N(x, \hat{u})=1$  时, 对式 (22) 加减  $\alpha_2 L_x^T C_u^T R^{-1} C_u (\nabla \varphi_1^T W_1 + \nabla \varepsilon_1)$ , 并根据引理 1 可得

$$\begin{aligned} \dot{L}_n \leq & -\alpha_2 \bar{Q}_m \|L_x\|^2 / 2 - \alpha_1 \|\tilde{W}_1\|^4 \delta_1 / \bar{\sigma}^2 + \\ & \alpha_1 \delta_2 / \bar{\sigma}^2 + \alpha_2 \|C_u^T R^{-1} C_u\|^2 \nabla^2 \varepsilon_{1M} / (4 \bar{Q}_m), \end{aligned}$$

其中  $0 < \bar{Q}_m > |\bar{Q}(x)|$ . 只要不等式

$$\|L_x\| > \sqrt{\|C_u^T R^{-1} C_u\|^2 \nabla^2 \varepsilon_{1M} / 2 \bar{Q}_m^2} \equiv B_{L1},$$

或

$$\|\tilde{W}_1\| > \sqrt[4]{\delta_2 / \delta_1 + \alpha_1 \delta_3^2 K^{*2} / \delta_1 \alpha_2 \bar{Q}_m} \equiv B_{W1}$$

成立, 就有  $\dot{L}_n < 0$ . 根据标准 Lyapunov 扩展定理<sup>[10]</sup>,  $\|L_x\|$  和  $\|\tilde{W}_1\|$  均一致最终有界.

综上,  $\|L_x\|$  和  $\|\tilde{W}_1\|$  均一致最终有界, 满足  $\|L_x\| < \max(B_{L0}, B_{L1}) \equiv B_L$ ,  $\|\tilde{W}_1\| < \max(B_{W0}, B_{W1}) \equiv B_W$ . 因为  $L_x$  径向无界且连续可导, 所以  $\|x\|$  一致最终有界. 当  $t \rightarrow \infty$  时, 利用式 (13) 和  $\hat{V}_x$  的定义可得  $\|\hat{V}(x) - V(x)\| \leq B_W \nabla \varphi_{1M} + \varepsilon_{1M} \equiv \varepsilon_V$ . 类似地, 利用式 (14) 和 (16) 可得

$$\|\hat{u} - u^*\| \leq$$

$$\|R^{-1}\| \|C_u\| \nabla \varphi_{1M} B_W / 2 + \|R\|^{-1} \|C_u\| \nabla \varepsilon_{1M} / 2 \equiv \varepsilon_u.$$

定理得证.  $\square$

### 3 仿真研究

考虑连续搅拌反应釜系统

$$\dot{x}_1 = -x_1 - a + D_a(1 - x_1 - a) \exp\left(-\frac{1}{x_2 + b + \rho}\right),$$

$$\dot{x}_2 = -(1 + \varrho)(x_2 + b) + H D_a(1 - x_1 - a) \times$$

$$\exp\left(-\frac{1}{x_2 + b + \rho}\right) + \varrho u. \quad (23)$$

其中:  $a = 0.0642$ ,  $b = 0.3948$ ;  $D_a = 0.072$ ;  $H = 8$ ;  $\rho = 20$ ;  $\varrho = 0.3$ .  $Q(x) = x^T x$ ,  $R = 1$ . 首先, 采用 RNN 辨识器 (4) 辨识系统 (23), 其中  $S = -2I_2$ ,  $\eta = 1.5$ . 设计参数选取  $\Gamma_1 = [1, 0.1; 0.1, 1]$ ,  $\Gamma_2 = [1, 0.2; 0.2, 1]$ ,  $\Gamma_3 = [1, 0.1; 0.1, 1]$ ,  $\Gamma_4 = 0.2$ ,  $\Gamma_5 = 0.1$ . 建模误差轨迹如图 1 所示. 然后利用辨识到的 RNN 模型设计基于单网络 ADP 的近似最优控制器. 设计参数  $\alpha_1 = 200$ ,  $\alpha_2 = 0.1$ . 系统状态轨迹如图 2 所示, 评价网的权值收敛轨迹如图 3 所示. 由仿真结果可以看出, 本文提出的基于单网络 ADP 的近似最优跟踪控制方案不需要系统模型已知, 并且能够获得较好的控制性能.

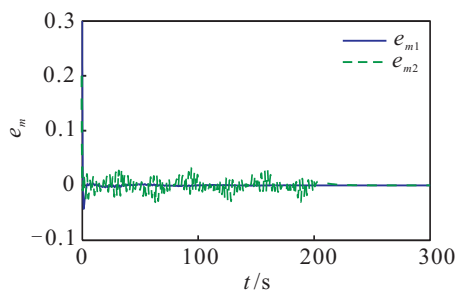


图 1 建模误差轨迹

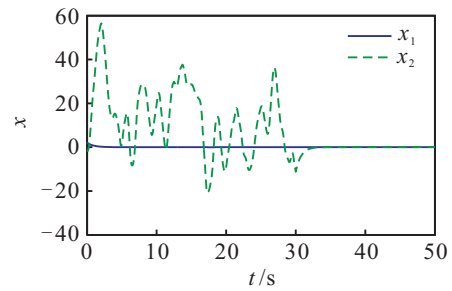


图 2 状态轨迹

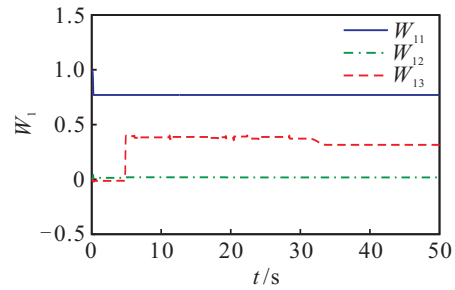


图 3 评价网权值收敛轨迹

### 4 结 论

本文针对一类未知的非线性系统提出了一个基于单网络 ADP 的近似最优控制方案. 所提出的控制方案不需要知道系统模型, 并且消除了常规 ADP 方法中的控制网络, 降低了计算复杂度. 仿真结果验证了本文所提出控制方案的有效性.

#### 参考文献(References)

- [1] Vamvoudakis K G, Lewis F L. Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem[J]. Automatica, 2010, 46(5): 878-888.
- [2] Zhang H G, Luo Y H, Liu D. Neural-network-based near-optimal control for a class of discrete-time affine nonlinear systems with control constraints[J]. IEEE Trans on Neural Networks, 2009, 20(9): 1490-1503.
- [3] Dierks T, Jagannathan S. Optimal Control of Affine Nonlinear Continuous-time Systems Using an Online Hamilton-Jacobi-Isaacs Formulation[C]. Proc of IEEE Conf on Decision and Control. New York: IEEE Press, 2010: 3047-3053.
- [4] Werbos P J. Approximate dynamic programming for real-time control and neural modeling[C]. Intelligent Control: Neural, Fuzzy, and Adaptive Approaches. New York: Van Nostrand Reinhold, 1992.
- [5] Liu D, Javaherian H, Tandale M D, et al. Adaptive critic learning technique for engine torque and air-fuel ratio control[J]. IEEE Trans on Systems, Man, and Cybernetics, Part B: Cybernetics, 2008, 38(4): 988-993.

(下转第 1430 页)