

基于自适应动态规划的非线性鲁棒近似最优跟踪控制

洪成文, 富月[†]

(东北大学 流程工业综合自动化国家重点实验室, 辽宁 沈阳 110819)

摘要: 为克服现有近似最优跟踪控制方法只能跟踪连续可微参考输入的局限, 本文针对一类具有未知动态的连续时间非线性时不变仿射系统, 提出了一种新的基于自适应动态规划的鲁棒近似最优跟踪控制方法. 首先采用递归神经网络建立系统模型, 然后建立评价神经网络对最优性能指标进行估计, 从而得到最优性能指标偏导数的估计值, 进而得到近似最优跟踪控制器, 最后利用系统输出与参考输入之间的跟踪误差设计鲁棒项对神经网络建模误差进行补偿. 分别针对两个非线性系统进行仿真实验, 仿真结果表明了所提方法的有效性和优越性.

关键词: 自适应动态规划; 最优跟踪; 未知非线性系统

引用格式: 洪成文, 富月. 基于自适应动态规划的非线性鲁棒近似最优跟踪控制. 控制理论与应用, 2018, 35(9): 1285 – 1292

中图分类号: TP273 **文献标识码:** A

Nonlinear robust approximate optimal tracking control based on adaptive dynamic programming

HONG Cheng-wen, FU Yue[†]

(State Key Laboratory of Process Industry Automation, Northeastern University, Shenyang Liaoning 110819, China)

Abstract: In order to overcome the limitation of the existing approximate optimal tracking control method that can only track continuously differentiable reference inputs, aiming at a kind of continuous time nonlinear invariant affine systems with unknown dynamics, a new robust approximation optimal tracking control method based on adaptive dynamic programming is proposed. Firstly, the system model is established by using the recurrent neural network. Then, the optimal performance index is estimated by the established critic neural network. The estimated value of the partial derivative of the optimal performance index, and the approximate optimal tracking controller can consequently be obtained. Finally, The error between the output and the reference input is used to designed the robust term to compensate for the modeling error of the neural network. Simulation experiments are conducted for two nonlinear systems respectively. The simulation results show the effectiveness and superiority of the proposed method.

Key words: adaptive dynamic programming; optimal track; unknown nonlinear system

Citation: HONG Chengwen, FU Yue. Nonlinear robust approximate optimal tracking control based on adaptive dynamic programming. *Control Theory & Applications*, 2018, 35(9): 1285 – 1292

1 引言(Introduction)

最优控制是一种寻找最节能最经济的控制策略, 很多控制问题都必须从最优控制的角度去进行研究和设计^[1]. 动态规划是求解最优控制问题的主要方法, 它将最优化问题分多级讨论, 寻求每一级的最优策略, 从而达到全局最优. 然而, 随着系统维数和运行时间的增加, 其计算量和存储量急剧增大, 具有“维数灾”问题^[2].

为避免“维数灾”的出现, 文[3]提出了自适应动态规划的概念, 它的实施无需已知被控对象的数学模

型, 因此适用于不确定系统, 近年来得到了很多控制理论专家和工程师的广泛关注^[4-5]. 文[6]针对一类动态未知的连续时间线性系统, 提出了一种新的自适应最优控制方法, 通过自适应动态规划技术迭代求解代数Riccati方程, 从而得到每一步的最优解. 文[7]针对一类连续时间仿射非线性系统, 提出了一种基于在线逼近的最优控制器, 解决了一类连续时间仿射非线性系统的最优控制问题. 文[8]提出的一种在线自适应增强学习方法, 解决了一类未知非线性系统的最优控制问题. 文[9-10]针对一类未知非线性连续时间系统,

收稿日期: 2018-01-23; 录用日期: 2018-03-23.

[†]通信作者. E-mail: fuyue@mail.neu.edu.cn.

本文责任编辑: 贾英民.

国家自然科学基金项目(61573090), 高校基本科研业务费项目(N160801001)资助.

Supported by the National Natural Science Foundation of China (61573090) and the Fundamental Research Funds for the Central Universities (N160801001).

基于自适应动态规划的方法,设计了一种鲁棒控制器,解决了一类未知非线性系统的最优控制问题.文[11]针对一类未知非线性连续时间系统,采用基于干扰观测器的自适应动态规划方法,设计了一种最优控制器,解决了一类未知非线性系统的最优控制问题.文[12]针对一类具有匹配不确定性的连续时间仿射非线性系统,应用自适应动态规划的框架,解决了一类连续时间仿射非线性系统的最优控制问题.文[13]针对一类连续时间非线性系统,通过基于自适应动态规划的方法,提出了两种新的在线方案用于合成控制器而不构建或假定系统的模型,从而解决了一类连续时间非线性系统的最优控制问题.

上述文献所提出的近似最优控制方法都属于调节问题.对于跟踪问题,文[14]针对一类连续时间非线性系统,提出了一种数据驱动的自适应跟踪控制方法,解决了一类连续时间非线性系统的跟踪控制问题.文[15]针对一类连续时间线性系统,应用增强学习技术设计了一种自适应最优跟踪控制器.文[16–17]针对一类未知非线性系统,根据自适应动态规划的方法设计最优跟踪控制器,其中前馈控制器是通过动态反演获得的,这要求输入矩阵的可逆性和系统的动力学知识.文[18–21]中设计的基于自适应动态规划的在线最优跟踪器,用于解决具有未知动态的线性连续时间系统的跟踪问题.文[22–23]针对一类连续时间未知非线性系统,基于哈密顿方程的近似解,设计了一类连续时间未知非线性系统的自适应跟踪控制器,解决了一类未知非线性连续时间系统的跟踪控制问题.文[24]针对一类具有未知动态的连续时间仿射非线性系统,提出了一种基于强化学习技术的最优跟踪控制方法.上述近似跟踪控制方法要么假设参考输入是连续可微的,要么假设参考输入由一个渐近稳定的外部系统产生.

针对现有近似最优跟踪控制方法对参考输入限制的限制性,本文提出一种新的基于自适应动态规划的近似最优跟踪控制方法,不要求参考轨迹连续可导,当参考轨迹具有不连续点时仍具有良好的跟踪效果.

2 问题描述(Problem description)

考虑如下连续时间非线性时不变仿射系统:

$$\dot{x} = f(x) + g(x)u(t); x(0) = x_0, \quad (1a)$$

$$y = Cx, \quad (1b)$$

其中: $x(t) = [x_1(t) \ x_2(t) \ \cdots \ x_n(t)] \in \mathbb{R}^n$ 是状态向量, $u(t) = [u_1(t) \ u_2(t) \ \cdots \ u_m(t)] \in \mathbb{R}^m$ 是输入向量, $y(t) = [y_1(t) \ y_2(t) \ \cdots \ y_l(t)] \in \mathbb{R}^l$ 是输出向量, $0 \leq l \leq m \leq n$; $C \in \mathbb{R}^{l \times n}$ 是已知的常数矩阵, $f(x) \in \mathbb{R}^n$, $g(x) \in \mathbb{R}^{n \times m}$ 是关于 $x(t)$ 的未知矩阵函数,满足 i) $f(0) = 0$; ii) $f(x) + g(x)u(t)$ 在包含原点的

集合 Ω 内可镇定,即存在连续的控制输入 $u(t) \in U$ 使得系统(1)在 Ω 上是渐近稳定的.

本文的目标是,针对具有未知动态的非线性系统(1),设计近似最优跟踪控制器,使系统输出 $y(t)$ 跟踪任意参考输入 $y_d(t)$ 的变化,并使如下性能指标最小:

$$J(t) = \frac{1}{2} \int_0^\infty [e^T(\tau)Qe(\tau) + u^T(\tau)Ru(\tau)]d\tau, \quad (2)$$

其中: $Q = Q^T > 0$, $R = R^T > 0$ 为具有适当维数的对称正定加权矩阵, $e(t) = y_d(t) - y(t)$ 为跟踪误差.

当系统的动态 $f(x), g(x)$ 未知时,文[17]提出了由稳态控制器、动态控制器和鲁棒补偿器组成的近似最优跟踪控制方法.其中,稳态控制器由连续可微的参考轨迹所满足的动态模型进行设计;动态控制器根据参考轨迹所满足的动态模型和辨识后的被控对象模型之间的误差系统模型进行设计,属于近似最优调节器,该调节器的实现包括两个神经网络,即评价网络和行为网络;鲁棒补偿器根据上述误差系统模型进行设计.上述近似最优跟踪控制方法假设参考输入 $y_d(t)$ 是连续可微的,当参考轨迹具有不连续点,例如当参考轨迹为方波时,无法实现良好的跟踪控制.

为克服这一局限,本文提出了一种新的基于自适应动态规划的鲁棒近似最优跟踪控制方法,该方法由动态控制器和鲁棒补偿器两部分组成.由于参考轨迹不是连续可微的,无法得到误差系统模型,所以动态控制器不能采用文献[17]的控制器设计方法.本文的动态控制器直接根据辨识后的系统模型进行设计,属于近似最优跟踪控制器,该控制器的实现只需要一个神经网络,即评价网络.鲁棒补偿器是根据系统输出和参考输入之间的误差进行设计的,可以同步补偿建模误差和评价网络近似误差.

3 基于动态规划的最优跟踪控制(Optimal tracking control based on dynamic programming)

为使所提出的方法更直观,首先以线性系统为例,设计基于动态规划的线性最优跟踪控制.

3.1 基于动态规划的线性最优跟踪控制(Linear optimal tracking control based on dynamic programming)

考虑如下连续时间时不变线性系统:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (3a)$$

$$y = Cx, \quad (3b)$$

其中: A, B, C 为相应维数的常数矩阵, $\{A, B\}$ 完全可控, $\{A, C\}$ 完全可观.

假设 $u^*(x, y_d)$ 为线性系统(3)的最优跟踪控制器,相应的最优跟踪指标为 $J^*(x, y_d)$, 则哈密顿函数满足

$$H(x, u^*, J_x^*) = J_x^{*T}(Ax + Bu^*) +$$

$$e^T Q e + u^{*T} R u^* = 0 \quad (4)$$

其中 $J_x^* = \frac{\partial J^*}{\partial x}$ 为最优性能指标关于 x 的偏导数. 令

$$\frac{\partial H}{\partial u^*} = R u^*(t) + B^T J_x^* = 0, \quad (5)$$

得到最优跟踪控制器

$$u^* = -R^{-1} B^T J_x^*. \quad (6)$$

由文[18]可知, 最优性能指标

$$J^*(x, y_d) = \frac{1}{2} x^T P x - x^T g + w(t), \quad (7)$$

其中 P, g, w 由下式确定:

$$0 = A^T P + P A - P B R^{-1} B^T P + C^T Q C, \quad (8)$$

$$-\dot{g} = (A - B R^{-1} B^T P)^T g + C Q y_d, \quad (9)$$

$$-\dot{w} = \frac{1}{2} y_d^T Q y_d - \frac{1}{2} g^T B R^{-1} B^T g. \quad (10)$$

因此

$$J_x^* = P x(t) - g(t), \quad (11)$$

将式(11)代入式(6)得

$$u^* = -R^{-1} B^T P x + R^{-1} B^T g, \quad (12)$$

这与采用极小值原理得到的最优跟踪控制器一致^[18], 因此采用动态规划的方法设计最优跟踪控制器具有一定的合理性.

3.2 基于动态规划的非线性最优跟踪控制(Non-linear optimal tracking control based on dynamic programming)

基于以上启发, 假设 $u^*(x, y_d)$ 为非线性系统(1)的最优跟踪控制器, 相应的最优跟踪指标为 $J^*(x, y_d)$, 则哈密顿函数满足

$$\begin{aligned} H(x, u^*, J_x^*) = \\ J_x^{*T} (f(x) + g(x)u^*(t)) + e^T Q e + u^{*T} R u^* = 0, \end{aligned} \quad (13)$$

令

$$\frac{\partial H}{\partial u^*} = R u^*(t) + g^T(x) J_x^* = 0, \quad (14)$$

得到最优跟踪控制器

$$u^* = -R^{-1} g^T(x) J_x^*. \quad (15)$$

众所周知, 当被控对象为非线性系统时, 根据式(13)求解最优性能指标 J^* 较为困难, 即使 $f(x), g(x)$ 已知也很难得到最优跟踪控制器的显示形式. 为此下面本文采用基于自适应动态规划的方法设计最优跟踪控制器. 首先采用递归神经网络估计系统动态 $f(x), g(x)$, 然后建立评价神经网络估计最优性能指标 J^* , 从而得到 J_x^* , 进而根据式(15)得到近似最优跟踪控制器.

4 鲁棒近似最优跟踪控制 (Robust approximate optimal tracking control)

4.1 系统模型辨识(System model identification)

由文[17]和文[25]可知, 系统(1a)可重新表示为如下递归神经网络的形式:

$$\dot{x}(t) = F^*(x) + G^*(x)u + \varepsilon(t), \quad (16)$$

其中: $F^*(x) = A_1^* x(t) + A_2^* \gamma(x(t)) + A_3^*$, $G^*(x) = B^*$, $\varepsilon(t)$ 为有界的建模误差, A_1^*, A_2^*, A_3^*, B^* 是理想的未知常值矩阵, $\gamma(x)$ 为单调递增的激励函数, 对于任意的 $x \geq y \in \mathbb{R}$ 满足 $0 \leq \gamma(x) - \gamma(y) \leq k(x - y)$, k 为正常数. 基于式(16), 建立如下神经网络估计模型:

$$\dot{\hat{x}}(t) = \hat{F}(\hat{x}) + \hat{G}(\hat{x})u - v(t), \quad (17)$$

$$v(t) = S e_m(t) + \frac{\hat{\lambda} e_m(t)}{e_m^T(t) e_m(t) + \eta}, \quad (18)$$

其中: $\hat{x}(t) \in \mathbb{R}^n$ 为估计模型的状态, $e_m(t) = x(t) - \hat{x}(t)$ 是估计误差, $S \in \mathbb{R}^{n \times n}$ 是可选的常数矩阵, $\hat{\lambda}(t) \in \mathbb{R}$ 是可调参数, $\eta > 1$ 是常数; $\hat{F}(\hat{x}) = \hat{A}_1 \hat{x}(t) + \hat{A}_2 \gamma(\hat{x}(t)) + \hat{A}_3$, $\hat{G}(\hat{x}) = \hat{B}$; $\hat{A}_1, \hat{A}_2, \hat{A}_3, \hat{B}$ 是 A_1^*, A_2^*, A_3^*, B^* 的估计, 根据下式更新:

$$\dot{\hat{A}}_1^T(t) = \Gamma_1 \hat{x}(t) e_m^T(t), \quad (19a)$$

$$\dot{\hat{B}}^T(t) = \Gamma_2 u(t) e_m^T(t), \quad (19b)$$

$$\dot{\hat{A}}_2^T(t) = \Gamma_3 \gamma(\hat{x}) e_m^T(t), \quad (19c)$$

$$\dot{\hat{A}}_3^T(t) = \Gamma_4 e_m^T(t), \quad (19d)$$

$$\dot{\hat{\lambda}}(t) = \Gamma_5 \frac{e_m^T(t) e_m(t)}{e_m^T(t) e_m(t) + \eta}, \quad (19e)$$

其中 $\Gamma_i, i = 1, 2, \dots, 5$ 是正定矩阵.

由文[17]可知, 当 $t \rightarrow \infty$ 时, $e_m \rightarrow 0$, 从而 $v(t) \rightarrow 0$; 同时 $\hat{A}_1(t), \hat{A}_2(t), \hat{A}_3(t), \hat{B}(t)$ 都趋于常值矩阵, 分别记为 A_1, A_2, A_3, B . 因此, 神经网络估计模型可重新写为

$$\dot{x}(t) = F(x) + G(x)u, \quad (20)$$

其中 $F(x) = A_1 x(t) + A_2 \gamma(x(t)) + A_3$, $G(x) = B$.

4.2 评价神经网络及近似最优跟踪控制 (Critic neural network and approximate optimal tracking control)

假设 1 最优性能指标 $J_x^*(x, y_d)$ 是连续可微的.

由文[26]可知, 若假设1成立, 则基于Weierstrass高阶近似理论, 存在完全独立的基函数集 $\{\varphi_i(x, y_d)\}$, 使得最优性能指标 $J^*(x, y_d)$ 和它的偏导数 J_x^* 可以通过如下神经网络近似

$$J^*(x, y_d) = W^T \phi(x, y_d) + \varepsilon_1(x, y_d), \quad (21a)$$

$$J_x^* = \nabla \phi^T W + \nabla \varepsilon_1, \quad (21b)$$

其中: $W = [w_1 \ w_2 \ \dots \ w_N]^T \in \mathbb{R}^N$ 为理想神经网络

络权值向量, $\phi(x, y_d) = [\varphi_1(x, y_d) \ \varphi_2(x, y_d) \ \cdots \ \varphi_N(x, y_d)]^T$ 为激励函数向量, N 为神经网络隐层节点数, $\varepsilon_1(x, y_d)$ 为有界神经网络逼近误差, 其对 x 的一阶偏导也是有界的, $\nabla\phi = \frac{\partial\phi}{\partial x}$, $\nabla\varepsilon_1 = \frac{\partial\varepsilon_1}{\partial x}$. 由文[26]可知, 当隐层节点数 $N \rightarrow \infty$ 时, 近似误差 $\varepsilon_1 \rightarrow 0$, $\nabla\varepsilon_1 \rightarrow 0$.

由式(21)可知, 针对式(20)的哈密顿函数为

$$H(x, u, W) = (W^T \nabla\phi + \nabla\varepsilon_1)(F(x) + G(x)u) + e^T Qe + u^T Ru, \quad (22)$$

令 \hat{W} 为 W 的估计, 则 J_x^* 的估计 \hat{J}_x^* 可表示为

$$\hat{J}_x^* = \nabla\phi^T \hat{W}. \quad (23)$$

由式(22)和式(23), 得到如下近似哈密顿函数:

$$H(x, u, \hat{W}) = \hat{W}^T \nabla\phi(F(x) + G(x)u) + e^T Qe + u^T Ru = e_1, \quad (24)$$

于是近似最优跟踪控制器方程为

$$\hat{u} = -R^{-1}G^T(x)\nabla\phi^T \hat{W}. \quad (25)$$

定义目标函数

$$E(\hat{W}) = e_1^T e_1, \quad (26)$$

采用梯度下降法最小化式(26)得到 \hat{W} 的更新算法

$$\dot{\hat{W}} = -\alpha_c h_1 (h_1^T \hat{W} + e^T Qe + \hat{u}^T R\hat{u}), \quad (27)$$

其中: $\alpha_c > 0$ 为神经网络的学习率, $h_1 = h_0/(h_0^T h_0 + 1)$, $h_0 = \nabla\phi(F(x) + G(x)\hat{u})$.

定理 1 假设 \hat{u} 为任意可行有界的控制策略, 并且评价网络中的数据向量 ϕ 是持续激励函数向量, 则:

1) 当 $\varepsilon_{\text{HJB}} = 0$ 时, 权重估计误差 \tilde{W} 收敛到 0;

2) 当 ε_{HJB} 有界时, \tilde{W} 是一致最终有界的, 其中 $\varepsilon_{\text{HJB}} = \nabla\varepsilon_1(F(x) + G(x)u)$, $\tilde{W} = \hat{W} - W$.

证 该定理的证明思路参考文[27]中的定理 2. 首先定义李雅普诺夫函数

$$L(t) = \frac{1}{2\alpha_c} \text{tr}\{\tilde{W}^T \tilde{W}\}. \quad (28)$$

由式(22)–(27)可知,

$$\dot{\tilde{W}} = -\alpha_c h_1 (h_1^T \tilde{W} - \varepsilon_{\text{HJB}}), \quad (29)$$

因此

$$\begin{aligned} \dot{L}(t) &= \frac{1}{\alpha_c} \text{tr}\{\tilde{W}^T \dot{\tilde{W}}\} = \\ &= \frac{1}{\alpha_c} \text{tr}\{\tilde{W}^T [-\alpha_c h_1 (h_1^T \tilde{W} - \varepsilon_{\text{HJB}})]\}. \end{aligned} \quad (30)$$

1) 当 $\varepsilon_{\text{HJB}} = 0$ 时, 式(30)可以简化为

$$\begin{aligned} \dot{L}(t) &= -\text{tr}(h_1^T \tilde{W})^T h_1^T \tilde{W} = \\ &= -\|h_1^T \tilde{W}\|^2 \leq 0. \end{aligned} \quad (31)$$

根据李雅普诺夫定理, 权重估计误差 \tilde{W} 将收敛于 0. 收

敛速度决定于激励函数 ϕ 和学习增益 α_c .

2) 当 ε_{HJB} 有界时, 定义 $h_1 h_1^T = P$, 因为 ϕ 是持续激励, 所以 P 是正定矩阵, 即存在正常数 σ , 使得 $\lambda_{\min}(P) > \sigma > 0$. 因此式(30)满足

$$\begin{aligned} \dot{L}(t) &= \text{tr}\{-\tilde{W}^T h_1 h_1^T \tilde{W} + \tilde{W}^T h_1 \varepsilon_{\text{HJB}}\} = \\ &= \text{tr}\{-\tilde{W}^T P \tilde{W} + \tilde{W}^T h_1 \varepsilon_{\text{HJB}}\} \leq \\ &= -\sigma \|\tilde{W}^T\|^2 + \text{tr}\{\tilde{W}^T h_1 \varepsilon_{\text{HJB}}\} \leq \\ &= -\|\tilde{W}\|(\sigma \|\tilde{W}\| - \|h_1 \varepsilon_{\text{HJB}}\|), \end{aligned} \quad (32)$$

由式(32)可知, $\dot{L}(t)$ 在紧集 Ω 的外面是负定的.

$$\Omega = \{\tilde{W} \mid \|\tilde{W}\| \leq \frac{\|h_1 \varepsilon_{\text{HJB}}\|}{\sigma}\}, \quad (33)$$

又根据 h_1 的定义, 存在一个正常数 σ_{1M} , 使得 $\|h_1\| \leq \sigma_{1M}$. 由于 ε_{HJB} 有界, 故权值估计误差是一致最终有界的. 证毕.

4.3 鲁棒项补偿器(Robust compensator)

虽然第 4.1 节介绍的递归神经网络能够对未知的系统动态进行有效的估计, 但是只有在稳态时建模误差才趋于 0. 为克服建模误差对控制系统动态性能的影响, 建立如下鲁棒项 u_r 对其进行补偿:

$$u_r = \frac{a_1 e}{e^T e + a_2}, \quad (34)$$

其中 $e = y_d - y$, a_1, a_2 为可调参数.

注 1 文[17]提出的近似最优跟踪控制方法中, 假设参考输入是连续可微的, 可以建立被控对象的神经网络模型和参考模型之间的误差系统模型, 所提出的鲁棒补偿器是根据误差系统模型进行设计的. 本文建立的鲁棒补偿器直接根据系统输出和参考输入之间的误差进行设计, 可以同步补偿建模误差和评价网络近似误差, 具有计算量小、补偿效果好等优点.

注 2 本文所提出的近似最优跟踪控制器只采用了一个评价网络. 文献[8]中的注释 1 指出: 只使用一个评价神经网络设计控制器是可行的, 但是系统的稳定性分析较为困难. 关于系统稳定性的研究将是下一步研究的重点.

综上所述, 由式(25)和式(34)可知, 本文提出的鲁棒近似最优跟踪控制为

$$\begin{aligned} u_t &= \hat{u} + u_r = \\ &= -R^{-1}G(x)^T \nabla\phi^T \hat{W} + \frac{a_1 e_r}{e_r^T e_r + a_2}. \end{aligned} \quad (35)$$

5 仿真实验(Simulation)

为了验证本文所提方法的有效性, 分别针对单输入单输出系统和双输入双输出系统进行仿真实验.

例 1 考虑如下单输入单输出非线性仿射系统

$$\begin{cases} \dot{x}_1 = -x_1 + x_2, \\ \dot{x}_2 = -0.5x_1 - 0.5x_2(1 - (\cos(2x_1) + 2)^2) + \\ \quad (\cos(2x_1) + 2)u, \\ y = x_2. \end{cases} \quad (36)$$

本文的目标为针对具有未知动态的系统(36)设计鲁棒近似最优跟踪控制器, 使系统的输出 y 跟踪参考轨迹 y_d 的变化。

首先采用递归神经网络对系统(36)进行建模. 选择

$$S = -15I_2, \eta = 1.5, \Gamma_1 = [1, 0.1; 0.1, 1],$$

$$\Gamma_2 = [1, 0.2; 0.2, 1], \Gamma_3 = [1, 0.1; 0.1, 1],$$

$$\Gamma_4 = 0.2, \Gamma_5 = 0.1,$$

$$\gamma(x) = \tanh x, x = [x_1 \ x_2]^T,$$

利用式(19)建立如下神经网络模型:

$$\begin{cases} \dot{x} = F(x) + G(x)u(t), \\ y = x_2, \end{cases} \quad (37)$$

其中:

$$F(x) = A_1x(t) + A_2\gamma(x(t)) + A_3, G(x) = B,$$

$$A_1 = \begin{pmatrix} -0.8375 & 0.8955 \\ -2.368 & 0.4977 \end{pmatrix},$$

$$A_2 = \begin{pmatrix} -0.2785 & 0.2314 \\ 0.4097 & 2.8080 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 7.69 \times 10^{-5} \\ 9.03 \times 10^{-5} \end{pmatrix},$$

$$B = \begin{pmatrix} -0.0404 \\ 0.8325 \end{pmatrix}.$$

然后建立评价神经网络对最优性能指标进行估计. 令激励函数向量 $\phi(x, y_d)$ 和权值估计向量 \hat{W} 分别为

$$\phi = [(x_2 - y_d)^3 \ 0.5(x_2 - y_d)^2 \ 0.5(x_2 - y_d)^4]^T,$$

$$\hat{W} = [W_1 \ W_2 \ W_3],$$

选择神经网络学习率 $a_c = 1$, 加权常数 $Q = 1, R = 1$, 鲁棒项可调参数 $a_1 = 20, a_2 = 1$, 采样时间 $t = 0.01$ s. 图1为 \hat{W} 的收敛曲线, 图2为当参考输入 y_d 为如下方波时系统输出 y 与 y_d 之间的跟踪曲线,

$$y_d = \begin{cases} -0.3, & t \in [0, 2], [5, 8], [11, 14], [17, 18], \\ -0.5, & t \in [2, 5], [8, 11], [14, 17], \end{cases}$$

图3为参考输入 $y_d = \sin t$ 时系统输出与 y_d 之间的跟踪曲线。

由图2和图3可知, 当参考输入为连续的正弦函数和具有不连续点的方波时, 采用本文所提出的方法都

具有良好的跟踪效果。

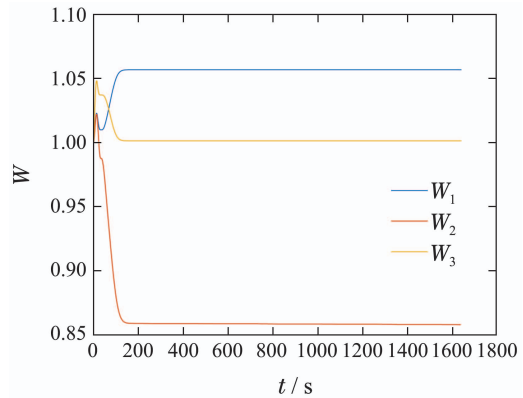


图 1 权值向量 \hat{W} 的收敛曲线

Fig. 1 Convergence curve of weight vector \hat{W}

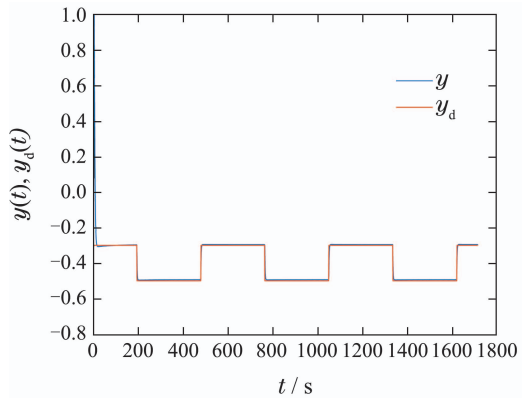


图 2 当参考输入 y_d 为方波时, 系统输出与参考输入之间的跟踪曲线

Fig. 2 The tracking curve between system output and reference input when the reference input y_d is a square wave

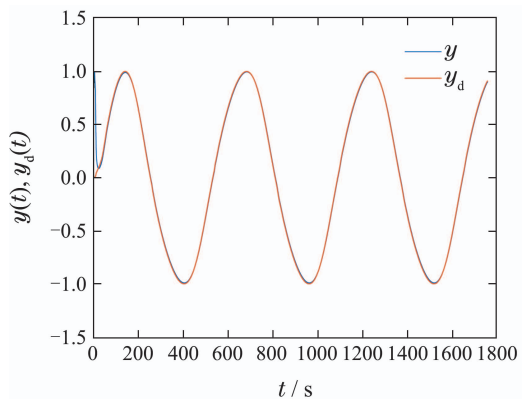


图 3 当参考输入 y_d 为正弦波时, 系统输出与参考输入之间的跟踪曲线

Fig. 3 The tracking curve between system output and reference input when the reference input y_d is a sine wave

为了突出本文所提方法的优越性, 采用文[17]所提出的鲁棒近似最优跟踪控制方法对系统(30)进行仿

真实实验. 图4为当参考输入 y_d 为上述方波时, 系统输出 y 与 y_d 之间的跟踪曲线; 图5为当参考输入 $y_d = \sin t$ 时, 系统输出 y 与 y_d 之间的跟踪曲线.

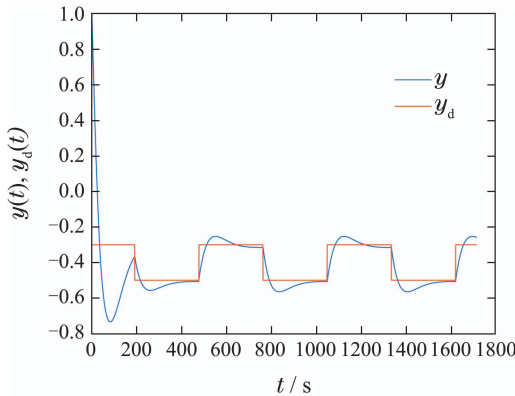


图4 当参考输入 y_d 为方波时, 系统输出与参考输入之间的跟踪曲线

Fig. 4 The tracking curve between system output and reference input when the reference input y_d is a square wave

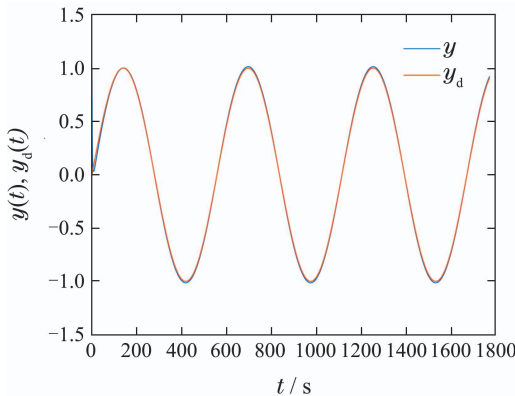


图5 当参考输入 y_d 为正弦波时, 系统输出与参考输入之间的跟踪曲线

Fig. 5 The tracking curve between system output and reference input when the reference input y_d is a square wave

由图4和图5可知, 虽然采用文[11]提出的鲁棒近似最优跟踪控制方法能够使系统的输出跟踪正弦函数曲线, 但是当参考输入为方波时, 无法实时有效的跟踪参考输入的变化.

例2 考虑如下双输入双输出非线性仿射系统:

$$\begin{cases} \dot{x}_1 = -x_1 - x_2 + \cos x_1 u_1, \\ \dot{x}_2 = -x_1^3 + x_1^3 + u_2, \\ y = x, \end{cases} \quad (38)$$

其中 $x = [x_1 \ x_2]^T$.

本文的目标为针对具有未知动态的系统(38)设计鲁棒近似最优跟踪控制器, 使系统的输出 $y = [y_1 \ y_2]^T$ 跟踪参考输入 $y_d = [y_{d1} \ y_{d2}]^T$ 的变化.

首先采用递归神经网络对系统(38)进行建模. 选择 $S = -10I_2$, $\eta = 1.5$, $\Gamma_1 = [1, 0.1; 0.1, 1]$, $\Gamma_2 = [1, 0.2; 0.2, 1]$, $\Gamma_3 = [1, 0.1; 0.1, 1]$, $\Gamma_4 = 0.2$, $\Gamma_5 = 0.1$, $\gamma(x) = \tanh x$, 建立如下神经网络模型:

$$\begin{cases} \dot{x} = F(x) + G(x)u(t), \\ y = [x_1 \ x_2]^T, \end{cases} \quad (39)$$

其中:

$$F(x) = A_1 x(t) + A_2 \gamma(x(t)) + A_3, \quad G(x) = B,$$

$$A_1 = \begin{pmatrix} -0.2888 & -0.5604 \\ -0.2507 & -0.5526 \end{pmatrix},$$

$$A_2 = \begin{pmatrix} -0.0331 & -0.0631 \\ -0.0579 & -0.1270 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 0.0011 \\ 0.0019 \end{pmatrix}, \quad B = \begin{pmatrix} 0.4625 & 0.3225 \\ 0.3025 & 0.3993 \end{pmatrix}.$$

然后建立评价神经网络对最优性能指标进行估计. 令激励函数向量 $\phi(x, y_d)$ 和权值估计向量 \hat{W} 分别为

$$\phi = [(x_1 - y_{d1})^2 \ (x_1 - y_{d1})(x_2 - y_{d2}) \ (x_2 - y_{d2})^2]^T,$$

$$\hat{W} = [W_1 \ W_2 \ W_3]^T,$$

选择神经网络学习率 $a_c = 1$, 加权阵 $Q = 10I_2$, $R = 0.01I_2$, 鲁棒项可调参数 $a_1 = 55$, $a_2 = 0.1$, 采样时间 $t = 0.01$ s.

图6为 \hat{W} 的收敛曲线, 图7为当参考输入 y_{d1} 和 y_{d2} 为如下方波时, 系统输出 y_1 与 y_{d1} 以及 y_2 与 y_{d2} 之间的跟踪曲线:

$$y_{d1} = \begin{cases} 0.75, & t \in [0, 3], [6, 9], [12, 15], [18, 25], \\ 0.8, & t \in [3, 6], [9, 12], [15, 18], \end{cases}$$

$$y_{d2} = \begin{cases} -0.3, & t \in [0, 2], [5, 8], [11, 14], [17, 25], \\ -0.5, & t \in [2, 5], [8, 11], [14, 17], \end{cases}$$

图8为当参考输入 $y_{d1} = \sin t$, $y_{d2} = \cos t$ 时的跟踪曲线.

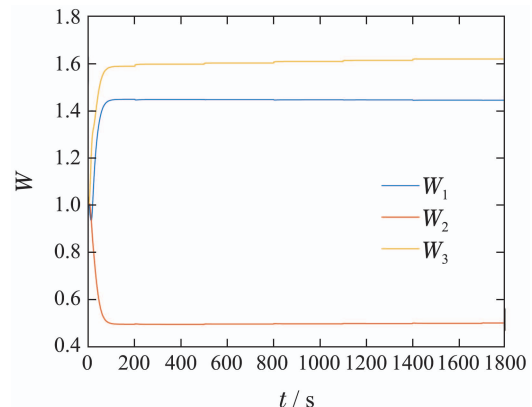


图6 权值向量 \hat{W} 的收敛曲线

Fig. 6 Convergence curve of weight vector \hat{W}

由图7和图8可知, 对于双输入双输出系统, 当参考输入为连续的正弦函数和具有不连续点的方波时, 采用本文所提出的方法都具有良好的跟踪效果. 图9为当参考输入为如上方波时, 采用文[17]所提方法时系统输出 y 与参考输入 y_d 之间的跟踪曲线, 图10为当参考输入 $y_{d1} = \sin t$, $y_{d2} = \cos t$ 时的跟踪曲线.

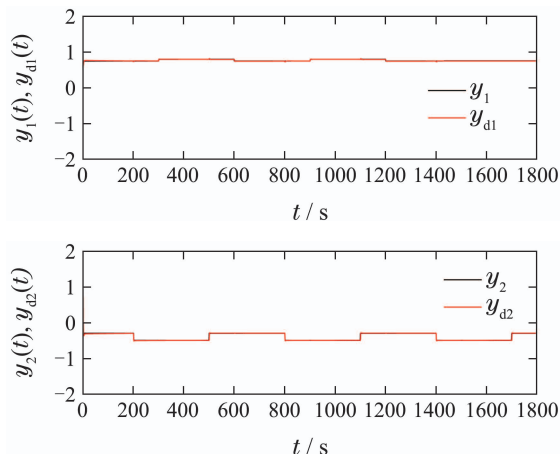


图 7 当参考输入 y_{d1} 和 y_{d2} 均为方波时, 系统输出与参考输入之间的跟踪曲线

Fig. 7 The tracking curve between system output and reference input when the reference input y_{d1} and y_{d2} are square wave

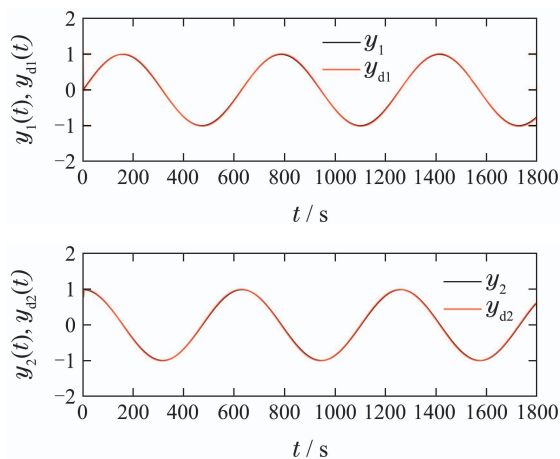


图 8 当参考输入 $y_{d1} = \sin t$, $y_{d2} = \cos t$ 时的跟踪曲线

Fig. 8 The tracking curve when reference $y_{d1} = \sin t$, $y_{d2} = \cos t$

由图9和图10可知, 采用文[17]所提方法时, 只有当参考输入为连续曲线时, 才具有良好的跟踪效果.

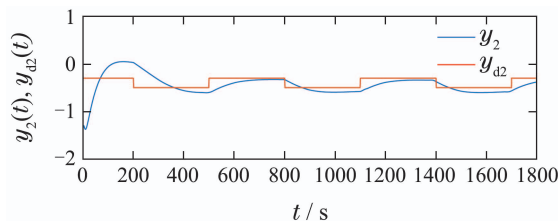
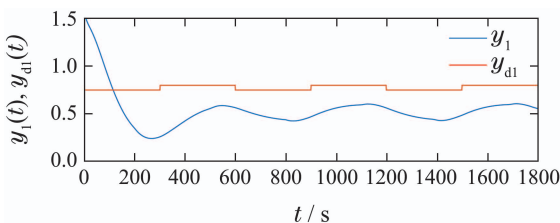


图 9 当参考输入为方波时, 采用文[17]所提方法的系统输出 y 与参考输入 y_d 之间的跟踪曲线

Fig. 9 The tracking curve between system output and reference input, using the method given in [17], When the reference input is a square wave

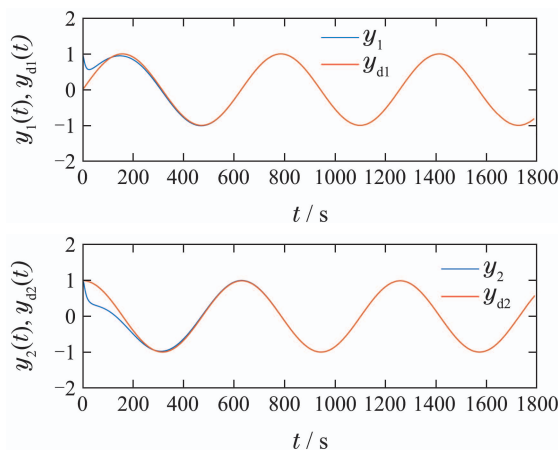


图 10 当参考输入 $y_{d1} = \sin t$, $y_{d2} = \cos t$ 时, 采用文[17]所提方法的系统输出 y 与参考输入 y_d 之间的跟踪曲线

Fig. 10 The tracking curve between system output and reference input, using the method given in [17], when reference $y_{d1} = \sin t$, $y_{d2} = \cos t$

6 总结(Summary)

本文针对一类具有未知动态的连续时间非线性时不变仿射系统, 提出了一种新的基于自适应动态规划的鲁棒近似最优跟踪控制方法, 该方法无需假设参考输入是连续可微的. 通过与现有方法进行仿真比较, 验证了本文所提出的方法可以跟踪具有不连续点的方波, 因此具有一定的优越性.

参考文献(References):

- [1] ZHEN Ziyang, JIANG Ju, WANG Zhisheng, et al. Finite-time information fusion for linear quadratic optimal control [J]. *Control Theory & Applications*, 2012, 29(2): 39 - 43. (甄子洋, 江驹, 王志胜, 等. 有限时间信息融合线性二次型最优控制 [J]. *控制理论与应用*, 2012, 29(2): 39 - 43.)
- [2] ZHANG Huaguang, ZHANG Xin, LUO Yanhong, et al. An overview of research on adaptive dynamic programming [J]. *Acta Automatica Sinica*, 2013, 39(4): 303 - 311. (张化光, 张欣, 罗艳红, 等. 自适应动态规划综述 [J]. *自动化学报*, 2013, 39(4): 303 - 311.)
- [3] WERBOS P.J. Advanced forecasting methods for global crisis warning and models of intelligence [J]. *General Systems Yearbook*, 1977, 22(6): 25 - 38.

- [4] WEI Qinglai, ZHANG Huaguang, LIU Derong, et al. An optimal control scheme for a class of discrete-time nonlinear systems with time delays using adaptive dynamic programming [J]. *Acta Automatica Sinica*, 2010, 36(1): 121 – 129.
(魏庆来, 张化光, 刘德荣, 等. 基于自适应动态规划的一类带有时滞的离散时间非线性系统的最优控制策略 [J]. *自动化学报*, 2010, 36(1): 121 – 129.)
- [5] LI Xiaoli, LIU Dexin, JIA Chao, et al. Multiple set-points tracking control method based on adaptive dynamic programming [J]. *Control Theory & Applications*, 2013, 30(6): 709 – 716.
(李晓理, 刘德馨, 贾超, 等. 基于自适应动态规划的多设定值跟踪控制方法 [J]. *控制理论与应用*, 2013, 30(6): 709 – 716.)
- [6] JIANG Y, JIANG Z P. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics [J]. *Automatica*, 2012, 48(10): 2699 – 2704.
- [7] DIERKS T, JAGANNATHAN S. Optimal control of affine nonlinear continuous-time systems [C] // *American Control Conference*. Baltimore, MD, USA: IEEE, 2010: 1568 – 1573.
- [8] BHASIN S, KAMALAPURKAR R, JOHNSON M, et al. A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems [J]. *Automatica*, 2013, 49(1): 82 – 92.
- [9] YANG X, HE H B, LIU D R, et al. Adaptive dynamic programming for robust neural control of unknown continuous-time non-linear systems [J]. *IET Control Theory & Applications*, 2017, 11(14): 2307 – 2316.
- [10] WANG D, LIU D, LI H, et al. Neural-network-based robust optimal control design for a class of uncertain nonlinear systems via adaptive dynamic programming [J]. *Information Sciences An International Journal*, 2014, 282: 167 – 179.
- [11] LIU D, HUANG Y, DING W, et al. Neural-network-observer-based optimal control for unknown nonlinear systems using adaptive dynamic programming [J]. *International Journal of Control*, 2013, 86(9): 1554 – 1566.
- [12] WANG D, LI C, LIU D, et al. Data-based robust optimal control of continuous-time affine nonlinear systems with matched uncertainties [J]. *Information Sciences*, 2016, 366: 121 – 133.
- [13] ZHANG J, ZHANG H, LIU Z, et al. Model-free optimal controller design for continuous-time nonlinear systems by adaptive dynamic programming based on a precompensator [J]. *ISA Transactions*, 2015, 57: 63 – 70.
- [14] MU C, NI Z, SUN C, et al. Data-driven tracking control with adaptive dynamic programming for a class of continuous-time nonlinear systems [J]. *IEEE Transactions on Cybernetics*, 2017, 47(6): 1460 – 1470.
- [15] GAO W, JIANG Z P. Linear optimal tracking control: an adaptive dynamic programming approach [C] // *American Control Conference*. Chicago, IL, USA: IEEE, 2015: 4929 – 4934.
- [16] WANG D, LIU D, WEI Q. Finite-horizon neuro-optimal tracking control for a class of discrete-time nonlinear systems using adaptive dynamic programming approach [J]. *Neurocomputing*, 2012, 78(1): 1 – 22.
- [17] ZHANG H, CUI L, ZHANG X, et al. Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method [J]. *IEEE Transactions on Neural Networks*, 2011, 22(12): 2226 – 2236.
- [18] QIN C, ZHANG H, LUO Y. Online optimal tracking control of continuous-time linear systems with unknown dynamics by using adaptive dynamic programming [J]. *International Journal of Control*, 2014, 87(5): 1000 – 1009.
- [19] XIAO G, ZHANG H, LUO Y, et al. Optimal output tracking control for nonlinear systems via successive approximation approach [J]. *Nonlinear Analysis Theory Methods & Applications*, 2007, 66(6): 1365 – 1377.
- [20] YANG X, LIU D, WEI Q, et al. Guaranteed cost neural tracking control for a class of uncertain nonlinear systems using adaptive dynamic programming [J]. *Neurocomputing*, 2016, 198(S): 80 – 90.
- [21] JING N, HERRMANN G. Online adaptive approximate optimal tracking control with simplified dual approximation structure for continuous-time unknown nonlinear systems [J]. *IEEE/CAA Journal of Automatica Sinica*, 2014, 1(4): 412 – 422.
- [22] NA J, LV Y, WU X, et al. Approximate optimal tracking control for continuous-time unknown nonlinear systems [C] // *Control Conference*. Nanjing, China: IEEE, 2014: 8990 – 8995.
- [23] MU C, SUN C, WANG D, et al. Adaptive tracking control for a class of continuous-time uncertain nonlinear systems using the approximate solution of HJB equation [J]. *Neurocomputing*, 2017, 260: 432 – 442.
- [24] XIAO G Y, ZHANG H G, LUO Y H, et al. Data-driven optimal tracking control for a class of affine non-linear continuous-time systems with completely unknown dynamics [J]. *IET Control Theory & Applications*, 2016, 10(6): 700 – 710.
- [25] RUBIO J D J, YU W. Stability analysis of nonlinear system identification via delayed neural networks [J]. *IEEE Transactions on Circuits & Systems II Express Briefs*, 2007, 54(2): 161 – 165.
- [26] 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.
- [27] MODARES H, LEWIS F L, NAGHIBI-SISTANI M B. Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks [J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2013, 24(10): 1513 – 1525.

作者简介:

洪成文 (1994–), 男, 硕士研究生, 主要研究方向为自适应控制、解耦控制、最优控制, E-mail: 921706019@qq.com;

富月 (1978–), 女, 副教授, 主要研究方向为工业过程智能解耦控制、自适应最优控制, E-mail: fuyue@mail.neu.edu.cn.