

非线性多输入系统的近似动态规划 H_∞ 控制

吕永峰^{1†}, 田建艳¹, 菅 堃¹, 任雪梅²

(1. 太原理工大学 电气与动力工程学院, 山西 太原 030024; 2. 北京理工大学 自动化学院, 北京 100081)

摘要: 实际工程中存在各种多输入系统, 比如用于大型雷达和火炮的多驱动伺服系统、多自由度机械臂系统等, 针对这些系统的 H_∞ 控制研究具有重要意义. 同时, 近似动态规划方法已被广泛用于求解各类最优控制问题, 但并未涉及多输入系统的 H_∞ 控制. 本文应用近似动态规划方法, 设计多输入非线性系统的 H_∞ 控制器. 应用基于强化学习的神经网络在线逼近非线性Hamilton–Jacobi–Isaacs (HJI)方程的解, 引进一种新的自适应律更新神经网络权值, 然后直接用于 H_∞ 控制器的设计, 并证明了权值的收敛性和系统的闭环稳定性, 保证了多输入系统受到外界未知干扰时的良好性能. 最后应用仿真实例验证所提方法的正确性和有效性.

关键词: H_∞ 控制; 近似动态规划; 多输入系统; 神经网络; 参数估计

引用格式: 吕永峰, 田建艳, 菅堃, 等. 非线性多输入系统的近似动态规划 H_∞ 控制. 控制理论与应用, 2021, 38(10): 1662 – 1670

DOI: 10.7641/CTA.2021.00559

Approximate-dynamic-programming H_∞ controls for multi-input nonlinear system

LÜ Yong-feng^{1†}, TIAN Jian-yan¹, JIAN Long¹, REN Xue-mei²

(1. College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan Shanxi 030024, China;
2. School of Automation, Beijing Institute of Technology, Beijing 100081, China)

Abstract: There exist all kinds of multi-input systems in practical engineering, such as multi-driven servo system for large radar and artillery, multi-degree manipulator system and so on. Although the approximate dynamic programming (ADP) has been used to solve the various optimal control problems, it has not involved in the designation of the H_∞ controls for multi-input system. In this paper, the ADP method is used to design the H_∞ controllers of the multi-input nonlinear system. The neural network (NN) based on reinforcement learning is applied to study the solution of nonlinear Hamilton–Jacobi–Isaacs (HJI) equation. A new adaptive law is introduced to update the NN weights, which is then directly used to design the H_∞ controllers. The convergence of weights and the stability of the system are proved, which guarantees the good performance of the multi-input system under the unknown disturbance. Finally, a simulation example is given to illustrate the correctness and effectiveness of the proposed method.

Key words: H_∞ control; approximate dynamic programming; multi-input systems; neural networks; parameter estimation

Citation: LÜ Yongfeng, TIAN Jianyan, JIAN Long, et al. Approximate-dynamic-programming H_∞ controls for multi-input nonlinear system. *Control Theory & Applications*, 2021, 38(10): 1662 – 1670

1 引言

近似动态规划方法(approximate dynamic programming, ADP)作为一种新型有效的最优控制求解方法, 已经得到了广泛的研究^[1–2]. 在强化学习(reinforcement learning, RL)算法的基础上, Werbos^[3]首次提出近似动态规划方法用于求解离散时间系统的最优控制问题. 近年来, 这种方法得到进一步的发展, 可以

近似地求解连续系统的动态规划最优控制问题, 为未知非线性系统实时在线的最优控制求解提供了一种有效的解决方案^[4–6]. 但是现有基于近似动态规划的研究大多只用于求解无扰动系统的最优控制^[7–10]或最优跟踪控制问题^[11–15], 也得到许多推广应用研究^[16–17], 比如制导^[16], 电网能源管理控制^[18]等.

在实际工程中, 系统通常会受到未知干扰, 需要设

收稿日期: 2020–08–25; 录用日期: 2021–02–16.

[†]通信作者. E-mail: lvyilian1989@foxmail.com; Tel.: +86 18213853441

本文责任编辑: 徐胜元.

山西省重点研发计划项目(201903D121062), 国家自然科学基金项目(61973036)资助.

Supported by the Key Research and Development Projects of Shanxi Province of China (201903D121062) and the National Natural Science Foundation of China (61973036).

计 H_∞ 控制器, 它可以看作是一个两人零和博弈模型^[19-21], 这一问题在过去几年得到广泛的研究^[22]. 学者们提出许多控制方案^[19, 23-24]通过获得预定性能指标的鞍点获得线性和非线性系统的 H_∞ 控制. 一般来说, H_∞ 控制器需要求解非线性系统的 Hamilton-Jacobi-Isaacs (HJI) 方程和线性系统的代数黎卡提方程^[25]. 然而, 由非线性系统导出的 HJI 方程是一个非线性偏微分方程, 当系统较大时会出现“维数灾难”问题而难以求解^[26]. 为了解决这一难题, 学者们应用近似动态规划方法学习 H_∞ 控制^[27]. 张等^[28]、Abu Khalaf 等^[29]利用离线迭代强化学习算法得到 H_∞ 控制器; 文献[19]提出在线同步强化学习算法解决这一问题. 然而, 这些方法要求系统动态完全已知, 李等在文献[30]中使用积分强化学习算法学习动态未知线性系统 HJI 方程的解. 文献[29]进一步提出高效的异步策略强化学习算法学习 HJI 方程的解.

但是上述 H_∞ 控制大部分只涉及单输入系统, 并未考虑多输入系统. 现有针对多输入系统的研究大部分仅仅基于非零和博弈求解最优控制, 并未考虑系统外部干扰^[4, 31]. 少数涉及多输入系统或者多智能体 H_∞ 控制的研究成果^[32-33], 文献[32]求解多智能体的 H_∞ 控制, 其中每个智能体涉及一个外部扰动; 文献[33]虽然求解混合博弈下多输入系统的均衡解, 却只考虑最后一个输入的抗干扰性能; 文献[24]应用 Q 学习方法设计线性多输入系统的 H_∞ 控制, 但未考虑非线性多输入系统. 非线性多输入系统的抗干扰控制具有许多工程应用, 比如多驱动负载系统, 机械臂系统和许多其他的工控系统^[24]. 因此针对多输入非线性系统的 H_∞ 控制研究具有重要意义. 总之, 先前研究成果大部分使用标准的执行-评价 ADP 结构, 而且并未考虑多输入系统的 H_∞ 控制器设计. 因此针对受外部干扰的多输入非线性系统, ADP 方法的实现仍然会面对很大挑战.

针对以上所述现存关于多输入受干扰非线性系统的 H_∞ 控制问题, 本文提出一种改进的在线 ADP 方法求解动态部分未知多输入非线性连续系统的 H_∞ 控制. 在该方案中, 避免了文献[8]中的执行神经网络, 从而简化 ADP 框架. 此外, 还应用一种新的自适应律^[28]在线更新神经网络权重. 引入评价神经网络在线学习 HJI 方程的解(即最优的性能指标函数), 利用逼近的神经网络计算多输入系统 H_∞ 控制. 最后, 证明在所设计 H_∞ 控制作用下闭环系统的稳定性, 并保证近似的 H_∞ 控制收敛到最优策略附近的紧集. 仿真结果验证了所提方法的有效性.

2 多输入系统 H_∞ 控制问题描述

考虑以下带干扰的多输入非线性系统:

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^N g_i(x)u_i(t) + k(x)d(t), \quad (1)$$

其中: $x \in \mathbb{R}^n$ 为系统状态, $u_i \in \mathbb{R}^{m_i}$ ($i = 1, \dots, N$) 表示 N 个系统输入, $d \in \mathbb{R}^q$ 为系统有界未知的外部干扰. $f(x) \in \mathbb{R}^n$ 系统非线性动态, $g_i(x) \in \mathbb{R}^{n \times m_i}$ 和 $k(x) \in \mathbb{R}^{n \times q}$ 分别表示系统输入动态和干扰动态. 假设 $f(0) = 0$, 且 $f(x) + \sum_{i=1}^N g_i(x)u_i(t) + k(x)d(t)$ 在实数集 Ω 上 Lipschitz 连续且稳定.

本文结合二人零和微分博弈理论^[18]求解多输入非线性系统的 H_∞ 控制, 其中求解的控制 $u_i(t)$ 要最小化给定多输入系统的性能指标, 求解的干扰 $d(t)$ 为使性能指标最坏的情况下的外部干扰, 即控制对 $(u_1^*, \dots, u_N^*, d^*)$ 满足零和博弈下的鞍点平衡条件. 在求解到最坏的干扰之后, 多输入系统的 H_∞ 控制器 $u_i(t)$ 各参数已经调整到相应的值, 在之后系统受到未知的外部干扰, 能使性能指标的值更小, 从而更加满足 H_∞ 性能指标条件. 多输入非线性系统的 H_∞ 性能指标定义为

$$\begin{aligned} V(x, u_1, \dots, u_N, d) = & \int_t^\infty r(x(\tau), u_1(\tau), \dots, u_N(\tau), d(\tau))d\tau, \\ V(0) = 0, \quad i = 1, \dots, N, \end{aligned} \quad (2)$$

其中: $r(x(\tau), u_i(\tau), d(\tau)) = x^T Q x + \sum_{i=1}^N u_i^T R_i u_i - \gamma^2 \|d\|^2$ 表示 H_∞ 性能指标的效用函数, Q 和 R_i 为维数相匹配的正定对称矩阵, γ 是 H_∞ 性能指标的扰动衰减因子. $u_i(t)$ 和 $d(t)$ 分别表示容许控制和外部干扰.

定义 1 如果求得的 H_∞ 控制能使多输入系统的性能指标函数 $V(x(t))$ 满足以下不等式

$$\begin{aligned} V(x, u_1^*, \dots, u_i^*, \dots, u_N^*, d) & \leq \\ V(x, u_1^*, \dots, u_i^*, \dots, u_N^*, d^*) & \leq \\ V(x, u_1^*, \dots, u_i, \dots, u_N^*, d^*), \end{aligned} \quad (3)$$

则 $(u_1^*, \dots, u_N^*, d^*)$ 达到了纳什均衡下的鞍点. 其中 $\{u_1^*, \dots, u_N^*\}$ 为多输入系统最优抗扰动控制, d^* 为使性能指标情况最坏的外部扰动.

需要指出, 本文所提方法适用于多输入系统具有一个 H_∞ 性能指标的情况. 当多输入系统每个输入都具有一个 H_∞ 性能指标时, 需要应用更加复杂的方法进行求解.

最优的性能指标函数表示为^[34-35]

$$\begin{aligned} V^*(x(0)) = & \min_{u_i} \max_d \int_t^\infty r(x(\tau), u_1(\tau), \dots, u_N(\tau), d(\tau))d\tau. \end{aligned} \quad (4)$$

假设以上性能指标的鞍点存在, 则系统(1)存在唯一的一组 H_∞ 控制^[36]. 为了获得 H_∞ 控制器, 哈密尔顿

函数给定为

$$H(x, \nabla V, u_i, d) = r(x, u_1, \dots, u_N, d) + (\nabla V)^T (f(x) + \sum_{i=1}^N g_i(x)u_i + k(x)d), \quad (5)$$

其中 $\nabla V \triangleq \frac{\partial V}{\partial x}$ 表示性能指标函数 V 对 x 的偏微分, $r(x, u_1, \dots, u_N, d) = x^T Q x + \sum_{i=1}^N u_i^T R_i u_i - \gamma^2 |d|^2$ 表示效用函数. 令

$$\frac{\partial H(x, \nabla V, u_i, d)}{\partial u_i} = 0 \text{ 和 } \frac{\partial H(x, \nabla V, u_i, d)}{\partial d} = 0,$$

推导可得 $2R_i u_i + (\nabla V)^T g_i(x) = 0$ 和 $-2\gamma^2 d + (\nabla V)^T k(x) = 0$, 可得多输入系统零和博弈下的 H_∞ 控制为

$$u_i^* = -\frac{1}{2} R_i^{-1} g_i^T(x) \nabla V^*, \quad i = 1, \dots, N, \quad (6)$$

$$d^* = \frac{1}{2\gamma^2} k^T(x) \nabla V^*. \quad (7)$$

如文献[34–35]所示, HJI方程可以表示为

$$0 = H(x, \nabla V^*, u_i^*, d^*) = xQ^T x + \nabla V^{*T}(x) f(x) - \frac{1}{4} \sum_{i=1}^N \nabla V^{*T}(x) g_i(x) R_i^{-1} g_i^T(x) \nabla V^*(x) + \frac{1}{4\gamma^2} \nabla V^{*T}(x) k k^T \nabla V^*(x). \quad (8)$$

给出以下引理^[34–35].

引理 1 1) 如果衰减因子 $\gamma > 0$, $V(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ 光滑连续, 则 $V^*(x) \geq 0$ 是 HJI 方程(8)的最优解. 2) 如果系统(1)状态可观, 则所求得的最优 H_∞ 控制 $u_i^*(t)$ 和 $d^*(t)$ 满足性能指标函数(4), 且 $(u_1^*, \dots, u_i^*, \dots, u_N^*, d^*)$ 为在 $[L_2, \infty)$ 上的鞍点, 能使多输入非线性系统(1)稳定.

注 1 不同于一般的最优控制, H_∞ 控制要求在鞍点存在的情况下得到性能指标函数的鞍点, 在系统受到其他外部干扰时, 能够使系统状态稳定且满足 H_∞ 性能指标, 具有较强的抗干扰能力. 为了避免直接求解 HJI 方程, 本文使用神经网络逼近性能指标函数, 避免了文献[27]中的执行结构, 优化了近似动态规划方法.

3 多输入系统 H_∞ 控制设计

先前研究成果^[37]应用执行-评价神经网络结构求解非线性系统的 H_∞ 控制对 (u^*, d^*) , d^* 表示使给定性能指标表现情况最坏的外部扰动, 如何服系统运行过程中使给定性能指标函数最大的外部力矩, 飞行器飞行过程中的风力扰动和机械臂运行过程中的未知外部阻力等, 其中神经网络权值使用最小二乘法或迭代自适应算法更新. 本文提出一种新的自适应神经网络

逼近最优性能指标函数, 直接用于多输入系统 H_∞ 控制的求解, 避免了在设计 H_∞ 控制器过程中使用执行神经网络结构. 并利用基于估计误差的自适应律估计性能指标的神经网络权值, 在线计算出多输入系统最优 H_∞ 控制 $(u_1^*, \dots, u_N^*, d^*)$.

3.1 近似动态规划神经网络设计

为了设计 H_∞ 控制式(6)和式(7), 需要求解非线性 HJI 方程(8)获得最优的值函数 $V^*(x)$. 因为非线性 HJI 方程(8)难以求解, 所以使用神经网络逼近最优的性能指标函数 $V^*(x)$. 假设性能指标函数 $V^*(x)$ 在实数集 Ω 上光滑连续, 则可用单层神经网络逼近为

$$V^*(x) = W^T \phi(x) + \varepsilon_v. \quad (9)$$

其关于 x 的偏微分可以表示为

$$\frac{\partial V^*(x)}{\partial x} = \nabla \phi^T W + \nabla \varepsilon_v, \quad (10)$$

其中: $W \in \mathbb{R}^l$ 指逼近性能指标后的神经网络权重, $\phi(x) \in \mathbb{R}^{l \times n}$ 是激活函数, ε_v 为神经网络逼近误差, l 指神经网络节点数. $\nabla \phi = \frac{\partial \phi}{\partial x}$ 和 $\nabla \varepsilon_v = \frac{\partial \varepsilon_v}{\partial x}$ 分别表示激活函数 ϕ 和逼近误差 ε_v 关于 x 的偏微分. 为了进行稳定性分析, 给出以下关于神经网络逼近方法的假设^[38].

假设 1 理想的神经网络权重 W , 激活函数 ϕ 及其偏微分 $\nabla \phi$, 逼近误差 ε_v 及其偏微分 $\nabla \varepsilon_v$ 有界, 即满足 $\|W\| \leq W_N$, $\|\nabla \phi\| \leq \phi_M$, $\|\nabla \varepsilon_v\| \leq \phi_\varepsilon$.

在实际应用中选择激活函数 $\phi(x)$ 为独立基函数, 使用假设 1 和万能逼近定理^[39], 应用式(9)–(10)能够逼近性能指标函数 $V^*(x)$ 及其偏微分 $\frac{\partial V^*(x)}{\partial x}$, 且当 $l \rightarrow +\infty$ 时, 如文献[8, 38]所示逼近误差 $\varepsilon_v, \nabla \varepsilon_v \rightarrow 0$.

近似的性能指标函数 $\hat{V}(x)$ 可以用神经网络逼近为

$$\hat{V}(x) = \hat{W}^T \phi(x). \quad (11)$$

则根据式(6)–(7)和式(11), 可以得到近似的 H_∞ 控制 (u_1, \dots, u_N, d) 为

$$u_i = -\frac{1}{2} R_i^{-1} [g_i(x)]^T \frac{\partial \hat{V}(x)}{\partial x} = -\frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T(x) \hat{W}, \quad i = 1, \dots, N, \quad (12)$$

$$d = \frac{1}{2\gamma^2} [k(x)]^T \frac{\partial \hat{V}(x)}{\partial x} = \frac{1}{2\gamma^2} [k(x)]^T \nabla \phi^T(x) \hat{W}, \quad (13)$$

其中: \hat{W} 是理想神经网络权值 W 的估计值; $\{u_1, \dots, u_N\}$ 为多输入系统近似最优抗扰动控制, d 为使性能指标(2)情况最坏的近似外部扰动.

现有大多数H_∞控制器只考虑单输入系统, 并未考虑多输入系统和给出多输入系统具体的H_∞控制器设计方法. 虽然文献[24]设计了H_∞控制器, 但是只考虑了线性系统. 本文应用近似动态规划方法解决多输入非线性系统的H_∞控制器设计问题, 为多输入受干扰系统的H_∞控制设计提出一种新而有效的解决方案.

最后设计自适应律更新神经网络权值 \hat{W} , 使其收敛于真值 W 附近. 由于逼近的神经网络要使HJI方程趋近于零, 可根据HJI方程设计权值 W 的自适应律. 则HJI方程可用神经网络结构(10)重新表述为

$$0 = H(x, \nabla V, u_i, d) = W^T \nabla \phi[(f(x) + \sum_{i=1}^N g_i(x)u_i + k(x)d)] + x^T Qx + \sum_{i=1}^N u_i^T R_i u_i - \gamma^2 \|d\|^2 + \varepsilon_{\text{HJI}}, \quad (14)$$

其中 $\varepsilon_{\text{HJI}} = \nabla \varepsilon_v[(f(x) + \sum_{i=1}^N g_i(x)u_i + k(x)d)]$ 为有界的HJI冗余误差, 当神经网络节点趋近于无穷时, 冗余误差将趋近于零^[18, 38], 即 $l \rightarrow +\infty, \varepsilon_{\text{HJI}} \rightarrow 0$. 为了设计自适应律, 令

$$\Xi = \nabla \phi[(f(x) + \sum_{i=1}^N g_i(x)u_i + k(x)d)]$$

和

$$\Theta = x^T Qx + \sum_{i=1}^N u_i^T R_i u_i - \gamma^2 \|d\|^2,$$

则HJI方程(14)可以简化为

$$\Theta = -W^T \Xi - \varepsilon_{\text{HJI}}. \quad (15)$$

由式(15)可知, 应用简化的HJI方程将神经网络权重 W 表示成线性化的形式, 因此可以用 Θ 和 Ξ 设计自适应律更新神经网络权值, 首先设计滤波辅助矩阵 $P \in \mathbb{R}^{l \times l}$ 和 $Q \in \mathbb{R}^l$ 为

$$\begin{cases} \dot{P} = -\ell P + \Xi \Xi^T, & P(0) = 0, \\ \dot{Q} = -\ell Q + \Xi \Theta, & Q(0) = 0, \end{cases} \quad (16)$$

其中: ℓ 为滤波因子, P 和 Q 分别为 $\Xi \Xi$ 和 $\Xi \Theta$ 的滤波矩阵. 可以得到

$$\begin{cases} P(t) = \int_0^t e^{-\ell(t-r)} \Xi \Xi^T dr, \\ Q(t) = \int_0^t e^{-\ell(t-r)} \Xi \Theta dr. \end{cases} \quad (17)$$

设计另一个辅助矩阵 $M \in \mathbb{R}^l$ 为

$$M = P\hat{W} + Q. \quad (18)$$

由式(15)和式(17)可得

$$Q = -PW + v, \quad v = -\int_0^t e^{-\ell(t-r)} \varepsilon_{\text{HJI}}(r) \Xi^T(r) dr.$$

且存在正实数 ε_v 满足 $\|v\| \leq \varepsilon_v$. 根据式(16)–(18)可得

$$M = P\hat{W} + Q = -P\tilde{W} + v, \quad (19)$$

其中 $\tilde{W} = W - \hat{W}$ 表示神经网络估计误差. 最后, 自适应律可以设计为

$$\dot{\hat{W}} = -\Gamma M, \quad (20)$$

其中 Γ 是学习增益.

为了进行收敛性分析, 引入以下引理^[15].

引理 2 如果式(20)中的回归向量 Ξ 持续激励, 则矩阵式(16)中的滤波矩阵 P 正定, 即存在 $\sigma > 0$ 满足 $\lambda_{\min}(P) > \sigma > 0$.

下面给出被估神经网络权值的收敛性证明.

定理 1 针对逼近性能指标的神经网络(11), 应用自适应律(20), 如果式(15)中的回归向量 Ξ 持续激励, 那么神经网络权值估计误差 $\tilde{W} = W - \hat{W}$ 收敛于零附近的紧集, \hat{W} 收敛到其真值 W 附近.

证 考虑李雅普诺夫函数 $V = \frac{1}{2} \tilde{W}^T \Gamma^{-1} \tilde{W}$, 则由式(19)–(20)推导可得其偏微分 \dot{V} 为

$$\dot{V} = \tilde{W}^T \Gamma^{-1} \dot{\tilde{W}} = -\tilde{W}^T P \tilde{W} + \tilde{W}^T v. \quad (21)$$

根据Young不等式可得

$$\begin{aligned} \dot{V} &= -\tilde{W}^T P \tilde{W} + \tilde{W}^T v = \\ &= -\sigma \|\tilde{W}\|^2 + \tilde{W}^T v \leq \\ &= -\|\tilde{W}\|(\sigma \|\tilde{W}\| - \varepsilon_v). \end{aligned} \quad (22)$$

根据李雅普诺夫扩展定理可知, 权值估计误差 \tilde{W} 收敛于集合 $\Omega_w : \{\tilde{W} \mid \|\tilde{W}\| \leq \varepsilon_v / \sigma\}$, 它的界由近似误差 ε_v 和激励等级常数 σ 决定. 因为 $\tilde{W} = W - \hat{W}$, 根据假设1, $\|W\| \leq W_N$ 成立, 所以存在正实数 b_w 满足 $\|\hat{W}\| \leq \|W\| + \|\tilde{W}\| < b_w$. 证毕.

注 2 文献[28]应用迭代方法解决了两人零和博弈的最优解, 迭代值函数使其上界满足 $|\bar{V}^{i+1} - \bar{V}^i| < \varepsilon$, 下界满足 $|\underline{V}^{i+1} - \underline{V}^i| < \varepsilon$, 同时上下界满足 $|\bar{V}^{i+1} - \underline{V}^i| < \varepsilon$ 则停止迭代得出对应的H_∞控制对. 而本文直接用HJI方程的数据信息设计自适应律, 更新性能指标神经网络权值使HJI方程趋近于零, 得出最优的值函数用于求解多输入非线性系统的H_∞控制. 相比文献[28]中的迭代算法, 简化了程序设计步骤.

3.2 稳定性分析

为了分析系统的稳定性, 将H_∞控制式(12)–(13)代入系统(1)可以得到

$$\begin{aligned} \dot{x} &= f(x) + \sum_{i=1}^N g_i(x)u_i + k(x)d = \\ &= f(x) + \sum_{i=1}^N g_i(x) \left\{ -\frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T \hat{W} \right\} + \\ &= k(x) \left\{ \frac{1}{2\gamma^2} [k(x)]^T \nabla \phi^T \hat{W} \right\} = \end{aligned}$$

$$\begin{aligned}
 & f(x) + \sum_{i=1}^N g_i(x) \left\{ \frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T \tilde{W} \right\} - \\
 & \sum_{i=1}^N g_i(x) \left\{ -\frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T \nabla \varepsilon_v \right\} + \\
 & \sum_{i=1}^N g_i(x) u_i^* - k(x) \left\{ \frac{1}{2\gamma^2} [k(x)]^T \nabla \phi^T \tilde{W} \right\} - \\
 & k(x) \left\{ \frac{1}{2\gamma^2} [k(x)]^T \nabla \phi^T \nabla \varepsilon_v \right\} + k(x) d^*. \quad (23)
 \end{aligned}$$

给出以下关于系统动态有界的假设^[40-41].

假设 2^[40-41] 非线性多输入系统(1)中各动态满足 $\|f(x)\| \leq b_f \|x\|$, $\|g_i(x)\| \leq b_{g_i}$, $\|k(x)\| \leq b_k$, 其中 $b_f > 0$, $b_g > 0$, $b_k > 0$.

总结全文可以得出以下定理.

定理 2 针对多输入带干扰的非线性系统(1), 使用 H_∞ 控制式(12)–(13), 权值自适应律(20), 如果回归向量 ϕ 和 Ξ 持续激励, 则神经网络权值估计误差 \tilde{W} 一致最终有界, 式(12)中的 H_∞ 控制 u_i 收敛到式(6)中的真值 u_i^* 附近, 式(13)中的最坏干扰 d 收敛到式(7)中的真值 d^* 附近, 即存在正实数 ς_{ui} , ς_d 满足 $\|u_i - u_i^*\| \leq \varsigma_{ui}$, $\|d - d^*\| \leq \varsigma_d$.

证 考虑以下李雅普诺夫函数

$$V = V_1 + V_2 = \frac{1}{2} \tilde{W}^T \Gamma^{-1} \tilde{W} + \Pi x^T x + KV^*, \quad (24)$$

其中: $K > 0$, $\Gamma > 0$ 为正实数. 考虑不等式 $ab \leq a^2\eta/2 + b^2/2\eta$, ($\eta > 0$), 根据式(22)可以得到

$$\begin{aligned}
 \dot{V}_1 &= -\tilde{W}^T P \tilde{W} + \tilde{W}^T v \leq -\sigma \|\tilde{W}\|^2 + \|\tilde{W}^T v\| \leq \\
 & -\left(\sigma - \frac{1}{2\eta}\right) \|\tilde{W}\|^2 + \frac{\eta \|v\|^2}{2}. \quad (25)
 \end{aligned}$$

根据式(2)和式(23)可得

$$\begin{aligned}
 \dot{V}_2 &= 2\Pi x^T \dot{x} + K(-x^T Qx - \\
 & \sum_{i=1}^N u_i^{*T} R_i u_i^* + \gamma^2 \|d\|^2) = 2\Pi x^T \{f(x) + \\
 & \sum_{i=1}^N g_i(x) \left\{ \frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T \tilde{W} \right\} - \\
 & \sum_{i=1}^N g_i(x) \left\{ -\frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T \nabla \varepsilon_v \right\} + \\
 & \sum_{i=1}^N g_i(x) u_i^* - k(x) \left\{ \frac{1}{2\gamma^2} [k(x)]^T \nabla \phi^T \tilde{W} \right\} - \\
 & k(x) \left\{ \frac{1}{2\gamma^2} [k(x)]^T \nabla \phi^T \nabla \varepsilon_v \right\} + k(x) d^*\} + \\
 & K(-x^T Qx - u^{*T} R u^* + \gamma^2 \|d\|^2) \leq \\
 & -[K\lambda_{\min}(Q) - \\
 & 2\Pi b_f - 2 \sum_{i=1}^N \eta b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) -
 \end{aligned}$$

$$\begin{aligned}
 & \frac{2\eta b_k^2 \phi_M}{\gamma^2} - 2] \|x\|^2 + \\
 & \left[\frac{1}{4\eta} \sum_{i=1}^N \Pi^2 b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) + \right. \\
 & \left. \frac{1}{4\eta\gamma^2} \Pi^2 b_k^2 \phi_M \|\tilde{W}\|^2 + \right. \\
 & \left. \left[\frac{1}{4\eta} \sum_{i=1}^N \Pi^2 b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) + \right. \right. \\
 & \left. \left. \frac{1}{4\eta\gamma^2} \Pi^2 b_k^2 \phi_M \|\nabla \varepsilon_v\|^2 - \right. \right. \\
 & \left. \left. \sum_{i=1}^N [K\lambda_{\min}(R_i) - \Pi^2 b_{g_i}^2] \|u_i^*\|^2 + \right. \right. \\
 & \left. \left. [K\gamma^2 + \Pi^2 b_k^2] \|d^*\|^2 \right]. \quad (26)
 \end{aligned}$$

由式(25)–(26)可得

$$\begin{aligned}
 \dot{V} &= \dot{V}_1 + \dot{V}_2 \leq \\
 & -[K\lambda_{\min}(Q) - 2\Pi b_f - \\
 & 2 \sum_{i=1}^N \eta b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) - \\
 & \frac{2\eta b_k^2 \phi_M}{\gamma^2} - 2] \|x\|^2 - \\
 & \left[\sigma - \frac{1}{2\eta} - \frac{1}{4\eta} \sum_{i=1}^N \Pi^2 b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) - \right. \\
 & \left. \frac{1}{4\eta\gamma^2} \Pi^2 b_k^2 \phi_M \|\tilde{W}\|^2 + \right. \\
 & \left. \left[\frac{1}{4\eta} \sum_{i=1}^N \Pi^2 b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) + \right. \right. \\
 & \left. \left. \frac{1}{4\eta\gamma^2} \Pi^2 b_k^2 \phi_M \|\nabla \varepsilon_v\|^2 - \right. \right. \\
 & \left. \left. \sum_{i=1}^N [K\lambda_{\min}(R_i) - \Pi^2 b_{g_i}^2] \|u_i^*\|^2 + \right. \right. \\
 & \left. \left. [K\gamma^2 + \Pi^2 b_k^2] \|d^*\|^2 + \frac{\eta \|v\|^2}{2} \right]. \quad (27)
 \end{aligned}$$

选择参数 K, Π, σ 满足

$$\begin{aligned}
 K &> \max \left\{ [2\Pi b_f + 2 \sum_{i=1}^N \eta b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) + \right. \\
 & \left. \frac{2\eta b_k^2 \phi_M}{\gamma^2}] / \lambda_{\min}(Q), \sqrt{\lambda_{\min}(R_i) \Pi^2 / b_{g_i}^2} \right\}, \\
 \sigma &> \max \left\{ \frac{1}{2\eta} - \frac{1}{4\eta} \sum_{i=1}^N \Pi^2 b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) - \right. \\
 & \left. \frac{1}{4\eta\gamma^2} \Pi^2 b_k^2 \phi_M \right\}, \Pi > 0, i = 1, \dots, N
 \end{aligned}$$

时, 式(27)可以进一步简化为

$$\dot{V} \leq -a_1 \|\tilde{W}\|^2 - a_2 \|x\|^2 + \chi, \quad (28)$$

其中 a_1, a_2 和 χ 指

$$a_1 = \sigma - \frac{1}{2\eta} - \frac{1}{4\eta} \sum_{i=1}^N \Pi^2 b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) -$$

$$\frac{1}{4\eta\gamma^2} \Pi^2 b_k^2 \phi_M,$$

$$a_2 = K\lambda_{\min}(Q) - 2\Pi b_f - 2 \sum_{i=1}^N \eta b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) - \frac{2\eta b_k^2 \phi_M}{\gamma^2} - 2,$$

$$\chi = \left[\frac{1}{4\eta} \sum_{i=1}^N \Pi^2 b_{g_i}^2 \phi_M \lambda_{\max}(R_i^{-1}) + \frac{1}{4\eta\gamma^2} \Pi^2 b_k^2 \phi_M \|\nabla \varepsilon_v\|^2 + [K\gamma^2 + \Pi^2 b_k^2] \|d^*\|^2 + \frac{\eta \|v\|^2}{2} \right].$$

由式(28)可知, 如果各变量满足 $\|\hat{W}\| > \sqrt{\chi/a_1}$ 和 $\|x\| > \sqrt{\chi/a_2}$, 则 $\dot{V} < 0$ 成立. 这表示闭环系统有界, 值函数有界.

最后可以得到最优H_∞控制和近似值之间的误差为

$$\lim_{l \rightarrow +\infty} \|\hat{u}_i - u_i^*\| = \left\| -\frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T(x) \hat{W} + \frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T(x) (\nabla \phi^T W + \nabla \varepsilon_v) \right\| = \left\| \frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T(x) \tilde{W} + \frac{1}{2} R_i^{-1} [g_i(x)]^T \nabla \phi^T(x) \nabla \varepsilon_v \right\| \leq \frac{1}{2} \lambda_{\max}(R_i^{-1}) b_{g_i} \phi_M \|\tilde{W}\| + \frac{1}{2} \lambda_{\max}(R_i^{-1}) b_{g_i} \phi_M \|\nabla \varepsilon_v\| \leq \varsigma_{ui}, \quad (29)$$

$$\lim_{l \rightarrow +\infty} \|\hat{d} - d^*\| = \left\| \frac{1}{2\gamma^2} [k(x)]^T \nabla \phi^T(x) \hat{W} - \frac{1}{2\gamma^2} [k(x)]^T \nabla \phi^T(x) (\nabla \phi^T W_2 + \nabla \varepsilon_v) \right\| = \left\| -\frac{1}{2\gamma^2} [k(x)]^T \nabla \phi^T(x) \tilde{W} + \frac{1}{2\gamma^2} [k(x)]^T \nabla \varepsilon_v \right\| \leq \frac{1}{2\gamma^2} b_k \phi_M \|\tilde{W}\| + \frac{1}{2\gamma^2} b_k \phi_M \|\nabla \varepsilon_v\| \leq \varsigma_d, \quad (30)$$

其中: $\varsigma_{ui} > 0$, $\varsigma_d > 0$ 是由神经网络逼近误差和权值估计误差导致的误差因子, 当神经网络节点 $l \rightarrow \infty$, 误差因子将收敛到零, 非线性多输入系统的近似H_∞控制收敛到最优值. 证毕.

4 仿真实验

考虑以下非线性系统^[36]

$$\dot{x} = f(x) + g(x)u + k(x)d, x \in \mathbb{R}^2, \quad (31)$$

其中:

$$f(x) = \begin{bmatrix} x_2 - x_1 \\ -x_2 - 0.5x_1 - 0.25x_2(\cos(2x_1) + 2)^2 + \\ 0.25x_2(\sin(4x_1) + 2)^2 \end{bmatrix},$$

$$g_1(x) = \begin{bmatrix} 0 \\ \cos(2x_1) + 2 \end{bmatrix},$$

$$xg_2(x) = \begin{bmatrix} 0 \\ \sin(4x_1^2) + 2 \end{bmatrix},$$

$$k(x) = \begin{bmatrix} 0 \\ \sin(4x_1) + 2 \end{bmatrix}.$$

H_∞性能指标给定为

$$V(x) = \int_0^\infty (x^2 + u_1^2 + u_2^2 - \gamma^2 d^2) dt, \quad (32)$$

其中 $\gamma = 8$ 为衰减因子.

仿真中, 设置系统初值为 $[4 \ -4]^T$, 应用性能指标函数的逼近神经网络(11), 激活函数设置为 $\phi(x) = [x_1^2 \ x_1 x_2 \ x_2^2]^T$, 使用权值自适应律(20), 自适应律中各参数设置为 $\hat{W}(0) = 0$, $Q = \text{diag}\{[1 \ 1]\}$, $R_1 = R_2 = 1$, $\ell = 1$, $\Gamma = \text{diag}\{[30 \ 30 \ 30]\}$. 图1为神经网络权值 \hat{W} 的收敛效果, 根据式(12)–(13), 可得图2的H_∞控制, 图3为系统状态. 可以看出, 应用所设计的H_∞控制能够使多输入系统性能指标收敛, 且系统稳定.

为了验证本文所设计H_∞控制器的有效性, 在仿真时间为 $t = 15$ s时, 加入一个常值 $d = 0.1$ s的外部干扰. H_∞控制的变化结果和系统状态如图4所示; 另外在 $t = 15$ s时加入另一个有界时变扰动 $d = 0.1 \sin t$, H_∞控制的运行结果和系统状态如图5所示.

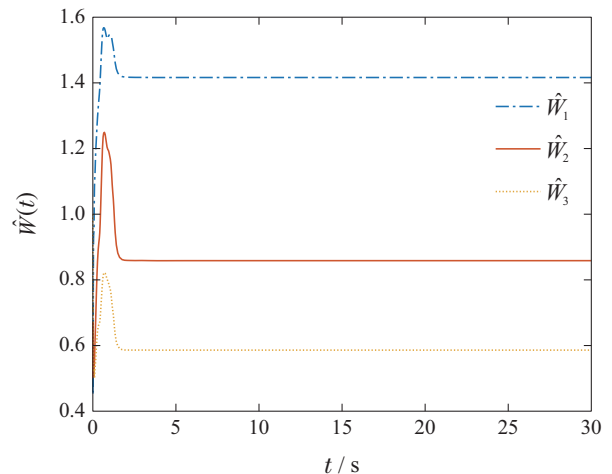
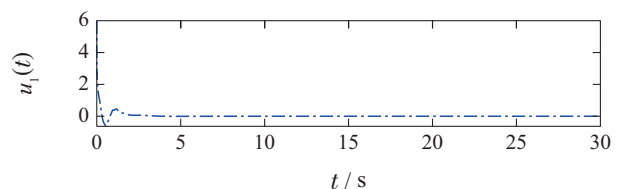


图 1 近似神经网络权重 \hat{W}

Fig. 1 Approximate NN weights \hat{W}



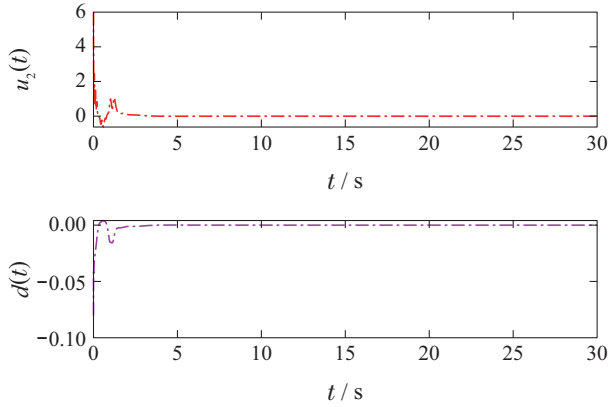


图2 H_∞ 控制 u_1, u_2 和最坏干扰 d
Fig. 2 H_∞ controls u_1, u_2 and d

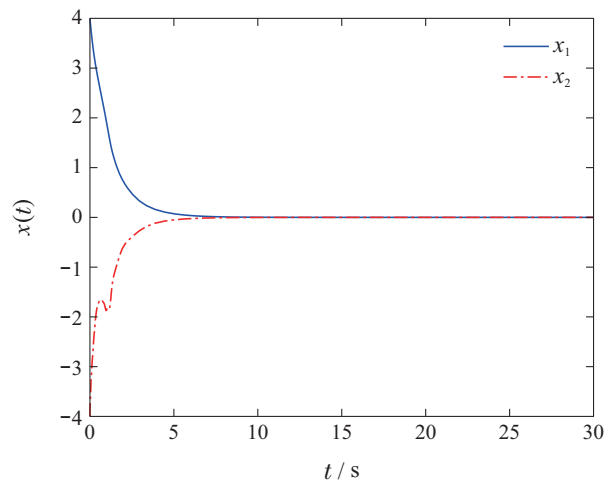


图3 系统状态
Fig. 3 System states

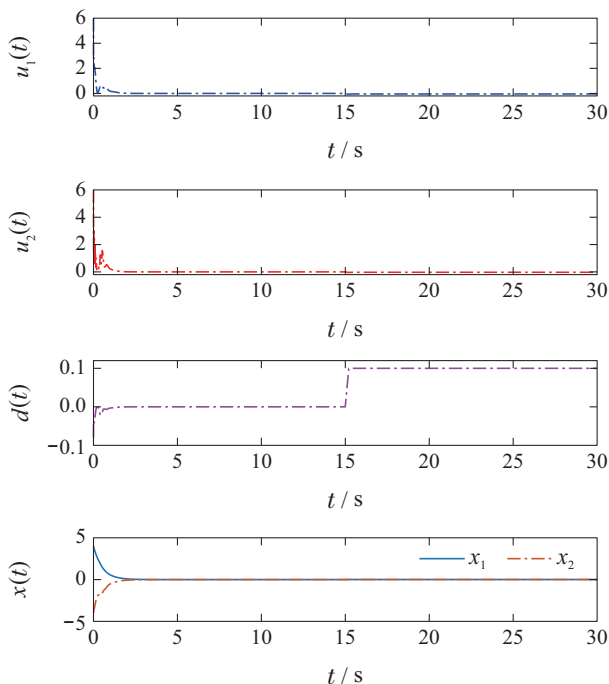


图4 $d = 0.1$ 下的 H_∞ 控制和系统状态
Fig. 4 H_∞ controls and system states with $d = 0.1$

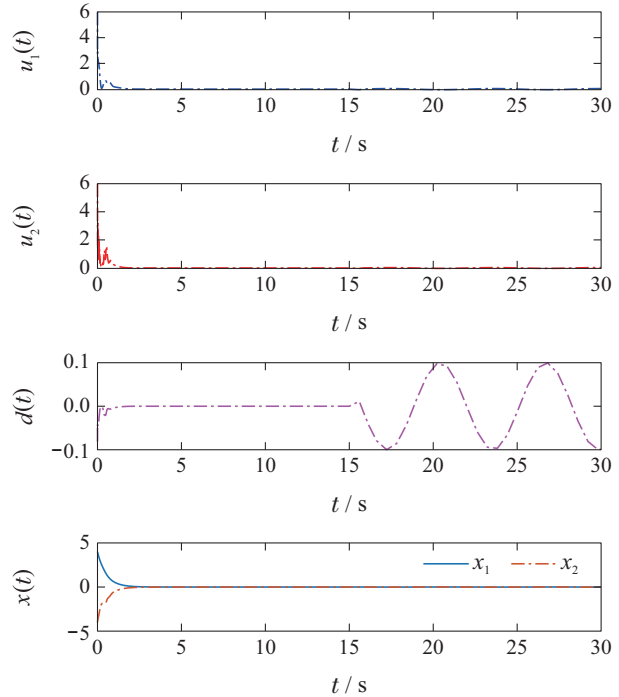


图5 $d = 0.1 \sin t$ 下的 H_∞ 控制和系统状态
Fig. 5 H_∞ controls and system states with $d = 0.1 \sin t$

可以看出, 在多输入系统 H_∞ 控制求解出来以后, 系统再收到其他外部常值或者有界时变干扰, 能够以更加满足 H_∞ 性能指标的形式稳定运行.

5 结论

针对多输入带干扰的非线性系统, 提出应用近似动态规划方法解决其 H_∞ 控制. 首先给出多输入系统 H_∞ 控制问题的描述, 定义纳什均衡解. 由于非线性的HJI方程难以求解, 应用神经网络逼近最优解, 引进一种新的自适应律更新神经网络权值, 直接用于 H_∞ 控制器的设计. 对权值的收敛性和在所设计 H_∞ 控制作用下的系统稳定性进行分析和证明. 最后用一个非线性多输入仿例证明所提出方法的正确性和有效性. 未来工作笔者会将这种 H_∞ 控制方法应用于非晶薄带的制造过程.

参考文献:

- [1] WANG Feiyue, WEI Qinglai. Intelligent control: from learning control to parallel control. *Control Theory & Applications*, 2018, 35(7): 939 – 948.
(王飞跃, 魏庆来. 智能控制: 从学习控制到平行控制. 控制理论与应用. 2018, 35(7): 939 – 948.)
- [2] CUI Xiaohong, LUO Yanhong, ZHANG Huaguang, et al. Finite-horizon optimal control for unknown systems with saturating control inputs. *Control Theory & Applications*, 2016, 33(5): 631 – 637.
(崔小红, 罗艳红, 张化光, 等. 未知饱和控制系统有穷域最优控制. 控制理论与应用, 2016, 33(5): 631 – 637.)
- [3] WERBOS P, MILLER W, SUTTON R. A menu of designs for reinforcement learning over time. *Neural Networks for Control*, 1990: 67 – 95.

- [4] SONG R, LEWIS F, WEI Q. Off-policy integral reinforcement learning method to solve nonlinear continuous-time multiplayer nonzero-sum games. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(3): 704 – 713.
- [5] NA J, WANG B, LI G, et al. Nonlinear constrained optimal control of wave energy converters with adaptive dynamic programming. *IEEE Transactions on Industrial Electronics*, 2018, 66(10): 7904 – 7915.
- [6] XUE S, LUO B, LIU D. Event-triggered adaptive dynamic programming for zero-sum game of partially unknown continuous-time nonlinear systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, 50(9): 3189 – 3199.
- [7] LV Y, REN X, HU S, et al. Approximate optimal stabilization control of servo mechanisms based on reinforcement learning scheme. *International Journal of Control, Automation and Systems*, 2019, 17(10): 2655 – 2665.
- [8] VAMVOUDAKIS K, LEWIS F. Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica*, 2010, 46(5): 878 – 888.
- [9] WANG D, LIU D, WEI Q, et al. Optimal control of unknown non-affine nonlinear discrete-time systems based on adaptive dynamic programming. *Automatica*, 2012, 48(8): 1825 – 1832.
- [10] WANG Ding, MU Chaoxu, LIU Derong. Data-driven nonlinear near-optimal regulation based on iterative neural dynamic programming. *Acta Automatica Sinica*, 2017, 43(3): 366 – 375.
(王鼎, 穆朝絮, 刘德荣. 基于迭代神经动态规划的数据驱动非线性近似最优调节. *自动化学报*, 2017, 43(3): 366 – 375.)
- [11] MU C, NI Z, SUN C, et al. Data-driven tracking control with adaptive dynamic programming for a class of continuous-time nonlinear systems. *IEEE Transactions on Cybernetics*, 2016, 47(6): 1460 – 1470.
- [12] MODARES N, LEWIS F. Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. *Automatica*, 2014, 50(7): 1780 – 1792.
- [13] 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. *IEEE Transactions on Neural Networks*, 2011, 22(12): 2226 – 2236.
- [14] HONG Chengwen, FU Yue. Nonlinear robust approximate optimal tracking control based on adaptive dynamic programming. *Control Theory & Applications*, 2018, 35(9): 1285 – 1292.
(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.)
- [15] NA J, HERRMANN G. Online adaptive approximate optimal tracking control with simplified dual approximation structure for continuous-time unknown nonlinear systems. *IEEE/CAA Journal of Automatica Sinica*, 2014, 1(4): 412 – 422..
- [16] CHEN Yanni, LIU Chunsheng, SUN Jingliang. Finite-time differential guidance law based on adaptive optimal control. *Control Theory & Applications*, 2019, 36(6): 877 – 884.
(陈燕妮, 刘春生, 孙景亮. 基于自适应最优控制的有限时间微分对策制导律. *控制理论与应用*, 2019, 36(6): 877 – 884.)
- [17] WANG Ding. Research progress on learning-based robust adaptive critic control. *Acta Automatica Sinica*, 2019, 45(6): 1031 – 1043.
(王鼎. 基于学习的鲁棒自适应评判控制研究进展. *自动化学报*, 2019, 45(6): 1031 – 1043.)
- [18] YIN Linfei, YU Tao, ZHANG Zeyu, et al. Deep adaptive dynamic programming based integration algorithm for generation control and optimization of islanded active distribution network. *Control Theory & Applications*, 2018, 35(2): 169 – 183.
(殷林飞, 余涛, 张泽宇, 等. 基于深度自适应动态规划的孤岛主动配电网发电控制与优化一体化算法. *控制理论与应用*, 2018, 35(2): 169 – 183.)
- [19] ZHANG H, QIN C, JIANG B, et al. Online adaptive policy learning algorithm for H_∞ state feedback control of unknown affine nonlinear discrete-time systems. *IEEE Transactions on Cybernetics*, 2014, 44(12): 2706 – 2718.
- [20] ZHONG X, HE H, WANG D, et al. Model-free adaptive control for unknown nonlinear zero-sum differential game. *IEEE Transactions on Cybernetics*, 2017, 48(5): 1633 – 1646.
- [21] WEI Q, ZHANG H. A new approach to solve a class of continuous-time nonlinear quadratic zero-sum game using ADP. *IEEE International Conference on Networking, Sensing and Control*. Sanya: IEEE, 2008, 4: 129 – 134.
- [22] ZHU Qiang, WANG Kexin, SHAO Zhijiang. Distributed dynamic optimization in the framework of dynamic games. *Control Theory & Applications*, 2020, 37(6): 1185 – 1195.
(朱强, 王可心, 邵之江. 动态博弈框架下的分布式动态优化. *控制理论与应用*, 2020, 37(6): 1185 – 1195.)
- [23] LIU D, LI H, WANG D. Neural-network-based zero-sum game for discrete-time nonlinear systems via iterative adaptive dynamic programming algorithm. *Neurocomputing*, 2013, 110: 92 – 100.
- [24] LI J, XIAO Z. H_∞ control for discrete-time multi-player systems via off-policy q -learning. *IEEE Access*, 2020, 8: 28831 – 28846.
- [25] WU H, LUO B. Simultaneous policy update algorithms for learning the solution of linear continuous-time H_∞ state feedback control. *Information Sciences*, 2013, 222: 472 – 485.
- [26] QIN C, ZHANG H, WANG Y, et al. Neural network-based online H_∞ control for discrete-time affine nonlinear system using adaptive dynamic programming. *Neurocomputing*, 2016, 198: 91 – 99.
- [27] MODARES H, LEWIS F, JIANG Z. H_∞ Tracking control of completely unknown continuous-time systems via off-policy reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2015, 26(10): 2550 – 2562.
- [28] ZHANG H, WEI Q, LIU D. An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. *Automatica*, 2011, 47(1): 207 – 214.
- [29] ABU-KHALAF M, LEWIS F, HUANG J. Neurodynamic programming and zero-sum games for constrained control systems. *IEEE Transactions on Neural Networks*, 2008, 19(7): 1243 – 1252.
- [30] LI H, LIU D, WANG D. Integral reinforcement learning for linear continuous-time zero-sum games with completely unknown dynamics. *IEEE Transactions on Automation Science and Engineering*, 2014, 11(3): 706 – 714.
- [31] LV Y, REN X, NA J. Online optimal solutions for multi-player nonzero-sum game with completely unknown dynamics. *Neurocomputing*, 2018, 283: 87 – 97.
- [32] JIAO Q, MODARES H, XU S, et al. Multi-agent zero-sum differential graphical games for disturbance rejection in distributed control. *Automatica*, 2016, 69: 24 – 34.
- [33] LV Y, REN X. Approximate Nash solutions for multiplayer mixed-zero-sum game with reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019, 49(12): 2739 – 2750.
- [34] BASAR T, OLSDER G. Dynamic noncooperative game theory. *Society for Industrial and Applied Mathematics*, 1998.
- [35] BASAR T, BERNHARD P. H -infinity optimal control and related minimax design problems: a dynamic game approach. *Springer Science & Business Media*, 2008.
- [36] VAMVOUDAKIS K, LEWIS F. Online solution of nonlinear two-player zero-sum games using synchronous policy iteration. *International Journal of Robust and Nonlinear Control*, 2012, 22(13): 1460 – 1483.

- [37] LIU D, HUANG Y, WEI Q. Neural network H_∞ tracking control of nonlinear systems using GHJI method. *International Symposium on Neural Networks*. Berlin: Springer, 2013: 186 – 195.
- [38] ABU-KHALAF M, LEWIS F. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. *Automatica*, 2005, 41(5): 779 – 791.
- [39] KOSMATOPOULOS E, POLYCARPOU M, CHRISTODOULOU M, et al. High-order neural network structures for identification of dynamical systems. *IEEE Transactions on Neural Networks*, 1995, 6(2): 422 – 431.
- [40] HANSELMAN T, NOAKES L, ZAKNICH A. Continuous-time adaptive critics. *IEEE Transactions on Neural Networks*, 2007, 18(3): 631 – 647.
- [41] MODARES H, LEWIS F, NAGHIBI-SITANS M. Adaptive optimal control of unknown constrained-input systems using policy iteration

and neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2013, 24(10): 1513 – 1525.

作者简介:

吕永峰 讲师, 目前研究方向为近似动态规划、数据驱动建模与控制、博弈理论和伺服系统, E-mail: lvyilian1989@foxmail.com;

田建艳 教授, 目前研究方向为复杂系统建模、多源信息融合智能监控和智能机器人, E-mail: tut_tianjy@163.com;

菅 奎 讲师, 目前研究方向为多智能体、自适应控制, E-mail: longtu11@163.com;

任雪梅 教授, 目前研究方向为神经网络与自适应控制、系统辨识与非线性控制、多电机驱动控制, E-mail: xmren@bit.edu.cn.