

引用格式:戴姣,刘春生,孙景亮.基于自适应动态规划的一类非线性系统的容错控制器设计[J].电光与控制,2018,25(10):84-88. DAI J, LIU C S, SUN J L. Fault tolerant controller design based on adaptive dynamic programming for a class of nonlinear systems[J]. Electronics Optics & Control, 2018, 25(10):84-88.

## 基于自适应动态规划的一类非线性系统的容错控制器设计

戴 姣, 刘春生, 孙景亮  
(南京航空航天大学,南京 211106)

**摘要:**针对存在执行器故障的一类仿射非线性系统,基于自适应动态规划方法,提出了一种新型的容错控制器。利用故障观测器估计执行器故障,并利用故障信息构建一个改进型的性能指标函数,将容错控制问题转化为最优控制问题。同时使用策略迭代(PI)算法,通过构造评价神经网络来求解HJB方程,获得近似最优容错控制律,并且基于李雅普诺夫函数,证明该容错控制器可以确保闭环系统渐近稳定。最后,通过仿真验证了该方法的有效性。

**关键词:**故障容错控制; 自适应动态规划; 神经网络; 故障估计

中图分类号: V249.1 文献标志码: A doi:10.3969/j.issn.1671-637X.2018.10.017

## Fault Tolerant Controller Design Based on Adaptive Dynamic Programming for A Class of Nonlinear Systems

DAI Jiao, LIU Chun-sheng, SUN Jing-liang  
(Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

**Abstract:** A novel Fault Tolerant Control (FTC) scheme is developed based on Adaptive Dynamic Programming (ADP) for a class of affine nonlinear systems with actuator failures. The estimated actuator failure from a fault observer is utilized to construct an improved performance index function, so the FTC problem can be transformed into an optimal control problem. By using Policy Iteration (PI) algorithm, the Hamilton-Jacobi-Bellman equation can be solved by constructing a critic neural network. Then, the approximate optimal controller can be derived directly. The closed-loop system is guaranteed to be asymptotically stable via the Lyapunov stability theorem. Finally, an example is presented to demonstrate the effectiveness of the proposed method.

**Key words:** Fault-Tolerant Control (FTC); Adaptive Dynamic Programming (ADP); neural network; fault estimation

### 0 引言

随着科学技术的发展,控制系统的规模和复杂性不断增加,系统出现故障的机率也随之增大。任何类型故障的发生都有可能对整个系统性能下降,甚至影响系统稳定性,造成不可预期的损失<sup>[1]</sup>。因此,提高控制系统的安全性和可靠性显得尤为重要。

根据故障处理的方式可将容错控制方法进行分类。1985年,文献[2]将容错控制分为两大类:主动容

错控制(Active FTC)和被动容错控制(Passive FTC)。被动容错控制针对的是特定的故障类型,不需要故障检测环节,其不改变控制器的参数和结构,设计方法相对简单<sup>[3]</sup>,与被动容错控制相比,主动容错控制具有更高的灵活性<sup>[4]</sup>。其在系统的故障发生之后,利用在线获得的故障信息,通过控制器的重构或者调整来保持系统的稳定性。

近年来,主动容错控制技术与鲁棒控制、最优控制、自适应控制<sup>[5]</sup>相结合,得到迅速发展。文献[6]针对执行器故障的不确定系统,利用神经网络估计系统的不确定性,并采用故障估计算法,研究了满足H<sub>2</sub>性能要求的容错控制器;文献[7]研究了一类时变故障的不确定非线性系统的H<sub>∞</sub>容错控制问题,在线估计时变故障,利用状态信息和故障信息构成控制律以补偿故障造成的影响,使H<sub>∞</sub>性能指标最小化。

收稿日期:2017-10-12

修回日期:2017-11-20

基金项目:中央高校基本科研业务费专项资金资助;南京航空航天大学研究生创新基地(实验室)开放基金(kfj20170304,kfj20170320)

作者简介:戴 姣(1995—),女,湖南常德人,硕士生,研究方向为故障诊断与容错控制、自适应动态规划。

自适应动态规划(Adaptive Dynamic Programming, ADP)是最优控制领域新兴起的一种近似最优方法,是当前研究热点。ADP方法通过函数近似结构来近似哈密顿-雅可比-贝尔曼(Hamilton-Jacobi-Bellman, HJB)方程的解,采用离线迭代<sup>[8]</sup>或者在线更新<sup>[9]</sup>的方法,来获得系统的近似最优控制策略,能够有效地解决非线性系统的优化控制问题。近年来,ADP算法用于解决连续时间和离散时间系统的时间延迟<sup>[10]</sup>、外部干扰<sup>[11]</sup>、控制受限<sup>[12]</sup>以及协同控制<sup>[13-14]</sup>、轨迹跟踪<sup>[15]</sup>等控制问题。但是有关ADP容错控制的理论研究成果很少<sup>[16]</sup>。文献[17]针对一类具有执行器故障的仿射非线性系统,提出了一种基于策略迭代算法(PI)的新型在线故障补偿方案。控制律由PI算法和故障补偿器两部分组成。文献[18]针对一类具有执行器增益故障和偏转故障的仿射非线性系统,构建积分滑模函数,利用自适应动态规划方法,得到一种滑模容错控制方案,通过构建执行网-评价网结构,得到近似最优容错控制律。

本文基于自适应动态规划方法,提出一种新的主动容错控制方案,主要创新点可概括如下:

1) 通过设计故障观测器,采用估计的执行器故障来构造性能指标函数,使其能同时反映执行器故障、状态和控制律,从而将容错控制问题转化为最优控制问题;

2) 容错控制器可以仅依赖于评价神经网络而得到,不再需要执行网络的训练;

3) 利用ADP方法来解决FTC问题,这是一种针对执行器故障的新的容错控制方法。

## 1 系统故障模型

考虑一类连续时间仿射非线性系统为

$$\dot{x} = f(x) + g(x)(u - u_a) \quad (1)$$

式中: $x(t) \in \mathbf{R}^n$  是系统的状态; $u(t) = [u_1 \cdots u_m]^T \in \mathbf{R}^m$  是控制输入; $u_a \in \mathbf{R}^m$  为未知的执行器故障; $f(x(t))$  和  $g(x(t))$  是连续可微的系统矩阵,假定  $f(\cdot)$  和  $g(\cdot)$  Lipchitz 连续,且  $f(0) = 0$ 。

**假设 1**  $f(\cdot)$  和  $g(\cdot)$  都是有界的。

**假设 2** 执行器故障  $u_a$  未知但是满足范数有界  $\|u_a\| \leq \delta_1 < +\infty$ ,  $\delta_1$  是常数,  $\delta_1 > 0$ 。

针对式(1)系统,定义性能指标函数为

$$V(x_0) = \int_0^\infty (\rho \hat{u}_a^T(\tau) \hat{u}_a(\tau) + L(x(\tau), u(\tau))) d\tau \quad (2)$$

式中: $\rho > 0$ ;  $L(x, u) = x^T Q x + u^T R u$ ,  $Q \in \mathbf{R}^{n \times n} > 0$ ,  $R \in \mathbf{R}^{m \times m} > 0$ ;  $\hat{u}_a \in \mathbf{R}^m$  是执行器故障  $u_a$  的估计值。

根据极小值原理,哈密顿方程可以定义为

$$H(x, u, \nabla V(x)) = \rho \hat{u}_a^T \hat{u}_a + L(x, u) + (\nabla V(x))^T (f(x) + g(x)u) \quad (3)$$

式中,  $\nabla V(x) = \partial V(x) / \partial x$ 。

则最优性能指标为

$$V^*(x_0) = \min \int_0^\infty (\rho \hat{u}_a^T(\tau) \hat{u}_a(\tau) + L(x(\tau), u(\tau))) d\tau \quad (4)$$

满足 HJB 方程

$$H(x, u^*, \nabla V^*(x)) = 0 \quad (5)$$

则得最优控制律为

$$u^*(x) = -\frac{1}{2} R^{-1} g^T(x) \nabla V^*(x) \quad (6)$$

由式(6)可知,欲得最优控制策略,必须求解 HJB 方程。但是,直接求解式(5)的非线性 HJB 方程是十分困难的,几乎不可能。因此,本文利用自适应动态规划(ADP)方法求解 HJB 方程。

## 2 基于 ADP 的容错控制器设计

### 2.1 故障估计

针对式(1)的故障系统,设计故障观测器为

$$\dot{\hat{x}} = f(\hat{x}) + g(\hat{x})(u - \hat{u}_a) + L_1(x - \hat{x}) \quad (7)$$

式中: $\hat{x}$  是系统状态  $x$  的估计值; $L_1$  是正定的观测器增益矩阵; $\hat{u}_a$  是执行器故障的估计值,其在线更新律为

$$\dot{\hat{u}}_a = -L_2 g^T(\hat{x}) e \quad (8)$$

式中: $L_2$  是正定矩阵; $e = x - \hat{x}$  代表系统状态的观测误差。由式(1)和式(7)可得

$$\begin{aligned} \dot{e} = & f(x) - f(\hat{x}) + (g(x) - g(\hat{x}))(u - u_a) - \\ & g(\hat{x})(u_a - \hat{u}_a) - L_1 e = \\ & e_f + e_g(u - u_a) - g(\hat{x})(u_a - \hat{u}_a) - L_1 e \end{aligned} \quad (9)$$

式中: $e_f = f(x) - f(\hat{x})$ ;  $e_g = g(x) - g(\hat{x})$ 。记  $e_1 = e_f + e_g(u - u_a)$ 。

**假设 3**  $e_1$  范数有界:  $\|e_1\| \leq \delta_2$ ,  $\delta_2$  是一个正数。

**注 1** 因为  $f(x(t))$  和  $g(x(t))$  都是有界的,所以  $e_f$  和  $e_g$  均范数有界,而且,  $u$  是紧密集  $S$  上的容许控制,所以  $u - u_a$  也有界,因此,假设 3 合理。

**定理 1** 对于式(1)的故障系统,在假设 1 和假设 2 成立的情况下,故障观测器(7)能够保证故障的观测误差最终一致有界稳定。

**证明** 选取 Lyapunov 函数

$$\Gamma_1 = \frac{1}{2} e^T e + \frac{1}{2} \tilde{u}_a^T L_2^{-1} \tilde{u}_a \quad (10)$$

式中,  $\tilde{u}_a = u_a - \hat{u}_a$  代表故障的估计误差。对式(10)求导有

$$\begin{aligned} \dot{\Gamma}_1 &= e^T \dot{e} - \hat{u}_a^T L_2^{-1} \tilde{u}_a = \\ e^T (e_f + e_g(u - u_a) - g(\hat{x})(u_a - \hat{u}_a) - L_1 e) - \hat{u}_a^T L_2^{-1} u_a &\leq \\ \delta_2 e - e^T g(\hat{x}) \tilde{u}_a - \lambda_{\min}(L_1) \|e\|^2 - \hat{u}_a^T L_2^{-1} \tilde{u}_a = \\ -(\lambda_{\min}(L_1) \|e\| - \delta_2) \|e\| - (e^T g(\hat{x}) + \hat{u}_a^T L_2^{-1}) \tilde{u}_a \end{aligned} \quad (11)$$

将式(8)自适应更新律代入式(11)有

$$\dot{\Gamma}_1 \leq -(\lambda_{\min}(L_1) \|e\| - \delta_2) \|e\| \quad (12)$$

因此,当  $\|e\| > \delta_2 / \lambda_{\min}(L_1)$  时,  $\dot{\Gamma}_1 < 0$ 。根据李雅普诺夫稳定性原理,故障观测误差满足最终一致有界稳定。

## 2.2 容错控制器设计

**定理 2** 对于式(1)的非线性故障系统,当发生执行器故障时,如果  $Q, R$  和  $\rho$  满足  $\lambda_{\min}(Q) \geq c_f + 1/2 + c_g^2/2, \lambda_{\min}(R) \geq c_g^2/2 + \|R\|^2, \rho \geq 3/2$ , 那么式(6)最优控制器能够保证系统渐近稳定。

**证明** 选取李雅普诺夫函数

$$L = \frac{1}{2} x^T x + V^*(x) \quad (13)$$

然后有

$$\begin{aligned} \dot{L} &= x^T \dot{x} + \nabla V^{*T}(x) \dot{x} = \\ x^T (f(x) + g(x)(u^* - u_a)) + \\ \nabla V^{*T}(x) (f(x) + g(x)(u^* - u_a)) &= \\ x^T f(x) + x^T g(x) u^* - x^T g(x) u_a - \nabla V^{*T}(x) g(x) u_a + \\ \nabla V^{*T}(x) (f(x) + g(x) u^*) \end{aligned} \quad (14)$$

基于式(5),可得

$$\nabla V^{*T}(x) (f(x) + g(x) u^*) = -\rho \hat{u}_a^T \hat{u}_a - x^T Q x - u^{*T} R u^* \quad (15)$$

由式(6)易得

$$-2R u^*(x) = g^T(x) \nabla V^*(x) \quad (16)$$

因此有

$$\begin{aligned} \dot{L} &= x^T f(x) + x^T g(x) u^* - x^T g(x) u_a + \\ 2u^{*T} R u_a - \rho \hat{u}_a^T \hat{u}_a - x^T Q x - u^{*T} R u &\leq \\ \|x\| \|f(x)\| + \frac{1}{2} \|x\|^2 + \frac{1}{2} \|g(x) u^*\|^2 + \\ \frac{1}{2} \|u_a\|^2 + \frac{1}{2} \|x^T g(x)\|^2 + \|u_a\|^2 + \|R\|^2 \|u^*\|^2 - \\ \rho \hat{u}_a^T \hat{u}_a - \lambda_{\min}(Q) \|x\|^2 - \lambda_{\min}(R) \|u^*\|^2 \end{aligned} \quad (17)$$

根据假设 1 可得

$$\begin{aligned} \dot{L} &\leq c_f \|x\|^2 + \frac{1}{2} \|x\|^2 + \frac{1}{2} c_g^2 \|u^*\|^2 + \\ \frac{1}{2} \|u_a\|^2 + \frac{1}{2} c_g^2 \|x\|^2 - \lambda_{\min}(Q) \|x\|^2 - \\ \lambda_{\min}(R) \|u^*\|^2 + \|u_a\|^2 + \|R\|^2 \|u^*\|^2 - \rho \hat{u}_a^T \hat{u}_a &= \\ -(\lambda_{\min}(Q) - c_f - \frac{1}{2} - \frac{1}{2} c_g^2) \|x\|^2 - \end{aligned}$$

$$\begin{aligned} (\lambda_{\min}(R) - \frac{1}{2} c_g^2 - \|R\|^2) \|u^*\|^2 + \frac{3}{2} \|u_a\|^2 - \rho \hat{u}_a^T \hat{u}_a &\leq \\ -(\lambda_{\min}(Q) - c_f - \frac{1}{2} - \frac{1}{2} c_g^2) \|x\|^2 - \\ (\lambda_{\min}(R) - \frac{1}{2} c_g^2 - \|R\|^2) \|u^*\|^2 - (\rho - \frac{3}{2}) \|u_a\|^2 \end{aligned} \quad (18)$$

因此,当  $Q, R$  和  $\rho$  满足条件

$$\begin{cases} \lambda_{\min}(Q) \geq c_f + \frac{1}{2} + \frac{1}{2} c_g^2 \\ \lambda_{\min}(R) \geq \frac{1}{2} c_g^2 + \|R\|^2 \\ \rho \geq \frac{3}{2} \end{cases} \quad (19)$$

时,系统保持渐近稳定。

## 3 神经网络实现

由于性能指标函数通常是高度非线性的,采用神经网络近似它<sup>[19]</sup>。在本文中,采用单层神经网络来逼近紧密集  $S$  上的假定可微性能指标函数。

$$V(x) = w_c^T \sigma(x) + \varepsilon_c(x) \quad (20)$$

式中:  $w_c \in \mathbf{R}^l$  是理想的权值向量,  $l$  代表隐含层中的神经元个数;  $\sigma(x) \in \mathbf{R}^l$  是激活函数,  $\varepsilon_c(x)$  表示神经网络的近似误差。式(20)中  $V(x)$  梯度为

$$\nabla V(x) = (\nabla \sigma(x))^T w_c + \nabla \varepsilon_c(x) \quad (21)$$

将式(8)代入式(5)中有

$$\rho \hat{u}_a^T \hat{u}_a + L(x, u) + (w_c^T \nabla \sigma(x) \nabla \varepsilon_c(x)) \dot{x} = 0 \quad (22)$$

因此,哈密顿方程可以表示为

$$\rho \hat{u}_a^T \hat{u}_a + L(x, u) + w_c^T \nabla \sigma(x) \dot{x} = -\nabla \varepsilon_c(x) \dot{x} \triangleq e_{\text{HBB}} \quad (23)$$

式中,  $e_{\text{HBB}}$  表示由神经网络近似带来的残余误差。

因为理想权值向量  $w_c$  是未知的,因此,评价神经网络可以近似表示为

$$\hat{V}(x) = \hat{w}_c^T \sigma(x) \quad (24)$$

$\hat{V}(x)$  的梯度可以表示为

$$\nabla \hat{V}(x) = (\nabla \sigma(x))^T \hat{w}_c \quad (25)$$

因此,近似哈密顿方程可以表示为

$$\rho \hat{u}_a^T \hat{u}_a + L(x, u) + w_c^T \nabla \sigma(x) \dot{x} \triangleq e_c \quad (26)$$

**假设 4** 评价神经网络的理想权值是有界的:  $\|w_c\| \leq w_1$ 。神经网络近似误差  $\varepsilon_c$  有界:  $\|\varepsilon_c\| \leq \varepsilon_1$ 。激活函数  $\sigma(x)$  及其微分是有界的。

定义目标函数  $E(t)$  为

$$E(t) = \frac{1}{2} e_c(t)^2 \quad (27)$$

则目标函数关于权值估计  $\hat{w}_c$  求偏微分有

$$\frac{\partial E}{\partial \hat{w}_c} = e_c \frac{\partial e_c}{\partial \hat{w}_c} = e_c \dot{x}^T (\nabla \sigma(x))^T \quad (28)$$

因此定义评价神经网络权值更新律为

$$\dot{\hat{w}}_c = -\gamma \cdot \text{proj} \left\{ \frac{\partial E}{\partial \hat{w}_c} \right\} = -\gamma \cdot \text{proj} \{ e_c \dot{x}^T (\nabla \sigma(x))^T \} \quad (29)$$

式中： $\gamma > 0$  是学习率； $\text{proj}$  表示投影算子。

因此，得到近似最优控制律为

$$\hat{u}(x) = -\frac{1}{2} R^{-1} g^T(x) (\sigma(x))^T \hat{w}_c \quad (30)$$

### 4 仿真分析

本文引用具有执行器故障的非线性系统模型<sup>[20]</sup>：

$$\dot{x} = \begin{bmatrix} -x_1 + x_2 \\ -0.5(x_1 + x_2) + 0.5x_1^2 x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ x_1 \end{bmatrix} (u - u_a), \text{ 其中,}$$

$$u_a = \begin{cases} 0 & 0 \leq t < 4 \text{ s} \\ 10 & 4 \text{ s} \leq t \leq 20 \text{ s} \end{cases}$$

选取设计参数  $Q = I_{2 \times 2}, R = 1, \rho = 4$ ，系统初始状态为  $x(0) = [1, -1]^T$ 。神经网络的权值向量定义为  $\hat{w}_c = [\hat{w}_{c1}, \hat{w}_{c2}, \hat{w}_{c3}]^T$ ，其初始值为  $\hat{w}_{c0} = [0.4, 0.4, 0.4]^T$ ，选取其激活函数为  $\sigma(x) = [x_1^2, x_1 x_2, x_2^2]$ ，学习率  $\gamma = 0.01$ ，故障观测器的增益矩阵  $L_1 = 30I_2, L_2 = 50I_1, \hat{x}(0) = [0, -0.5]^T, \hat{u}_a(0) = 0$ 。

仿真结果如图 1 ~ 图 5 所示。

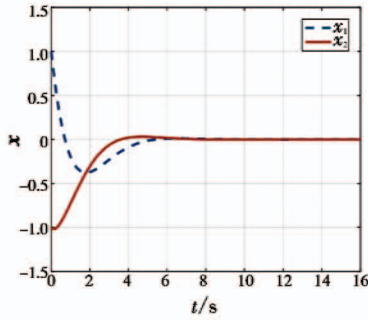


图 1 无故障时系统的动态响应

Fig.1 Dynamic responses with no fault

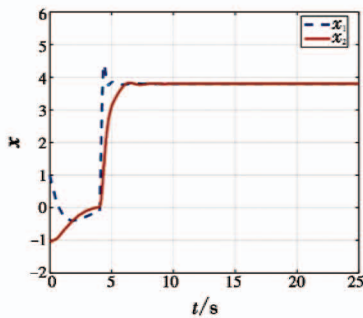


图 2 未引入容错控制器的故障系统的动态响应

Fig.2 Dynamic responses of the fault system without FTC

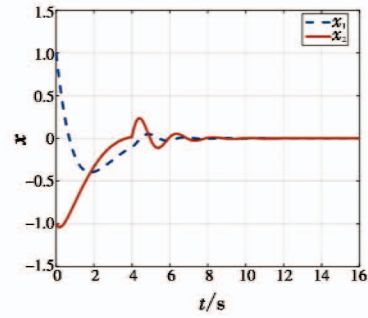


图 3 容错控制器作用下的故障系统的动态响应

Fig.3 Dynamic responses of the fault system under the FTC input

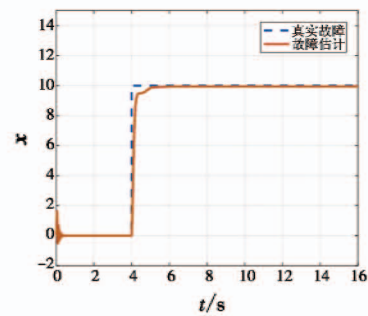


图 4 基于故障观测器的故障估计

Fig.4 The estimated fault based on the fault observer

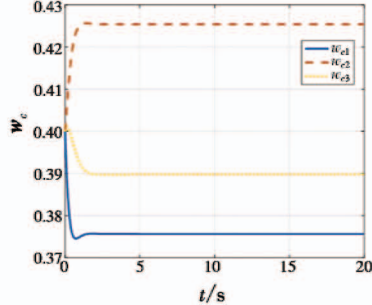


图 5 评价神经网络的权值

Fig.5 The weights of the critic neural network

从图 5 可以看出，评价神经网络的权值在设计的算法下最终收敛。执行器的故障估计是构造性能指标函数的关键，图 4 表明了故障观测器的故障估计值可靠有效。从图 2 和图 1 可以看出，当第 4 s 故障发生后，在没有引入容错控制的情况下，系统的状态受到影响并出现稳态误差，但是在图 3 容错控制输入的作用下，系统能很快趋于稳定，无误差。因此，仿真结果证明，本文所提出的主动容错控制方案是有效的。

### 5 结论

提出了一类自适应优化容错控制设计方法。为了实现主动容错控制，构造了自适应故障观测器，并将其用于容错控制律的设计，同时基于 ADP 算法获得近似

最优容错控制律,解决了容错控制的优化问题。本文的另一特点是采用单层神经网络近似性能指标,简化了算法。仿真结果验证了本文算法的有效性,对解决一类非线性系统的执行器故障容错控制有良好的参考价值。

### 参 考 文 献

- [1] 周东华, DING X. 容错控制理论及其应用 [J]. 自动化学报, 2000, 26(6): 788-797.
- [2] ETERNO J S, WEISS J L, LOOZE D P, et al. Design issues for fault tolerant-restructurable aircraft control[C]// Proceedings of the 24th IEEE Conference on Decision and Control, 1985. doi: 10.1109/CDC.1985.268630.
- [3] SONG X Q, LIU C S, ZHANG S J. Adaptive active fault tolerant control for discrete-time systems with uncertainties[J]. Asian Journal of Control, 2015. doi: 10.1002/asjc.1225.
- [4] 毛海杰, 李炜, 冯小林. 非线性系统主动容错控制综述 [J]. 传感器与微系统, 2014(4): 6-9, 13.
- [5] CHEN F, JIANG B, TAO G. An intelligent self-repairing control for nonlinear MIMO systems via adaptive sliding mode control technology[J]. Journal of the Franklin Institute, 2014, 351(1): 399-411.
- [6] 刘春生, 姜斌. 一类非线性系统的 H<sub>2</sub> 容错控制器的设计及其在空间飞行器的应用 [J]. 自动化学报, 2013, 39(2): 188-196.
- [7] LIU C S, JIANG B.  $H_{\infty}$  fault-tolerant control for time-varied actuator fault of nonlinear system [J]. International Journal of Systems Science, 2014, 45(12): 2447-2457.
- [8] AL-TAMIMI A, LEWIS F L, ABU-KHALAF M. Discrete-time non-linear HJB solution using approximate dynamic programming: convergence proof [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2008, 38(4): 943-949.
- [9] YANG X, LIU D R, WEI Q L. Online approximate optimal control for affine non-linear systems with unknown internal dynamics using adaptive dynamic programming [J]. IET Control Theory & Applications, 2014, 8(16): 1676-1688.
- [10] ZHANG H, WANG Z, LIU D. Global asymptotic stability of recurrent neural networks with multiple time-varying delays [J]. IEEE Transactions on Neural Networks, 2008, 19(5): 855-873.
- [11] WANG D, LIU D, LI H. Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems [J]. IEEE Transactions on Automation Science and Engineering, 2014, 11(2): 627-632.
- [12] YANG X, LIU D, WANG D. Reinforcement learning for adaptive optimal control of unknown continuous-time nonlinear systems with input constraints [J]. International Journal of Control, 2014, 87(3): 553-566.
- [13] ZHANG H, ZHANG J, YANG G H, et al. Leader-based optimal coordination control for the consensus problem of multiagent differential games via fuzzy adaptive dynamic programming [J]. IEEE Transactions on Fuzzy Systems, 2015, 23(1): 152-163.
- [14] WANG Y, SONG Y, LEWIS F L. Robust adaptive fault-tolerant control of multiagent systems with uncertain non-identical dynamics and undetectable actuation failures [J]. IEEE Transactions on Industrial Electronics, 2015, 62(6): 3978-3988.
- [15] ZHANG H, SONG R, WEI Q, et al. Optimal tracking control for a class of nonlinear discrete-time systems with time delays based on heuristic dynamic programming [J]. IEEE Transactions on Neural Networks, 2011, 22(12): 1851-1862.
- [16] 孙景亮, 刘春生. 基于自适应动态规划的导弹制导律研究综述 [J]. 自动化学报, 2017, 43(7): 1101-1113.
- [17] ZHAO B, LIU D, LI Y. Online fault compensation control based on policy iteration algorithm for a class of affine non-linear systems with actuator failures [J]. IET Control Theory & Applications, 2016, 10(15): 1816-1823.
- [18] FAN Q Y, YANG G H. Adaptive fault-tolerant control for affine non-linear systems based on approximate dynamic programming [J]. IET Control Theory & Applications, 2016, 10(6): 655-663.
- [19] WANG D, LIU D, ZHANG Q, et al. Data-based adaptive critic designs for nonlinear robust optimal control with uncertain dynamics [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2016, 46(11): 1544-1555.
- [20] JIANG H, ZHANG H, LIU Y, et al. Neural-network-based control scheme for a class of nonlinear systems with actuator faults via data-driven reinforcement learning method [J]. Neurocomputing, 2017, 239: 1-8.