

基于自适应动态规划的约束非线性系统自触发间歇最优控制

陈浩, 王维峰*

(中南民族大学 数学与统计学学院, 武汉 430074)

摘要 通过自触发间歇控制机制研究了带有输入约束的连续非线性系统的最优控制问题. 将设计出的控制器应用到系统中, 并克服系统内部未知扰动的影响; 然后, 采用评价神经网络和执行神经网络分别逼近最优成本函数和最优控制输入, 得到最优控制策略, 再利用半全局实际有限时间稳定的推广引理保证系统的稳定性分析; 最后, 通过数值仿真验证了所提理论的可行性.

关键词 自适应动态规划; 连续非线性系统; 自触发间歇最优控制; 输入约束; 神经网络

中图分类号 O232 文献标志码 A 文章编号 1672-4321(2026)01-0110-09

doi:10.20056/j.cnki.ZNMDZK.20250825

Self-triggered intermittent optimal control of constrained nonlinear systems via adaptive dynamic programming

CHEN Hao, WANG Weifeng*

(College of Mathematics and Statistics, South-Central Minzu University, Wuhan 430074, China)

Abstract The optimal control problem of continuous nonlinear systems are investigated with input constraints using a self-triggered intermittent control mechanism. The designed controller is applied to the system to overcome the influence of unknown internal disturbances. Then, the critic and action neural networks are employed to approximate the optimal cost function and the optimal control input, respectively, to obtain the optimal control strategy. The stability of the system is analyzed by using a generalized lemma on semi-global practical finite-time stability. Finally, numerical simulations are conducted to verify the feasibility of the proposed theory.

Keywords adaptive dynamic programming; continuous nonlinear systems; self-triggered intermittent optimal control; input constraint; neural network

在过去的几十年, 自适应动态规划(ADP)作为一种有效的智能控制方法, 在寻找最优控制^[1-2]的解决方案方面发挥了重要作用. 为了克服传统动态规划方法导致的“维数灾”问题, WERBOS在1977年提出了该方法. 通常, 对于线性系统可以使用Riccati方程求解^[3], 而对于非线性系统来说, 其对应的哈密顿-雅可比-贝尔曼(HJB)方程是一个非线性偏微分方程, 其解析解很难得到. 因此, 运用ADP方法^[4-5]利用函数近似结构来逼近满足HJB方程的最优成本

函数和最优控制输入, 可以有效地解决此问题, 例如使用神经网络(NN)结构.

最优控制作为现代控制系统设计的核心理念, 致力于确保在满足系统性能指标的前提下, 寻求期望性能指标和可用控制资源之间的最优平衡. 对于传统的控制策略^[6], 控制器随着时间连续更新, 消耗了大量的通信资源. 为缓解该问题, 学界内提出了事件触发机制^[7], 该机制是一种非周期性的方式, 虽然在一定程度上减少了采样数量, 但需要连续检测

收稿日期 2025-04-24

* 通信作者 王维峰(1985-), 男, 副教授, 研究方向: 随机动力系统, E-mail: wwf87487643@163.com

基金项目 湖北省自然科学基金资助项目(2023AFC006)

触发条件,消耗了大量的计算资源.在此基础上,有学者提出了自触发机制^[8]和动态事件触发机制.自触发机制在当前时刻计算出下一触发时刻,有效避免连续检测触发条件,降低了计算的复杂度;相较于传统的事件触发机制,动态事件触发机制引入了一个额外的内部动态变量,可以动态调整触发阈值,使得采样间隔进一步增大.近年来,为了降低控制器的采样数量,且提高采样效率,很多学者将间歇控制方法^[9-11]引入最优控制问题,间歇控制方法将时间区间划分为工作区间和休息区间,在工作区间进行控制,而在休息区间停止控制,这种方式不仅能确保系统性能,还提高了资源的利用率.如 LIU 等^[12]将事件触发机制和间歇控制结合,提出了事件触发间歇最优控制方法,有效地减少了采样次数;ZHANG 等^[13]将动态事件触发机制和间歇控制结合,提出了动态事件触发间歇最优控制方法,使理论得到进一步扩展.在其他领域,该方法同样彰显优越性,如 MEI 等^[14]将其考虑在空调系统中,体现了间歇控制方法的节能优势.基于上述分析,本文通过将自触发机制和间歇控制方法相结合的方式去解决最优控制问题,具有重要的理论价值和实际意义.

近年来,带有输入约束的控制问题在工业界引起广泛关注^[15],因为在实际应用中,系统不可避免地存在影响内部特性的约束条件,一定程度上阻碍了控制的实行,甚至严重影响到具体方案的实际效果.本文通过引入一种非二次型函数解决此问题,以避免资源的过度消耗.此外,实际的系统中经常出现模型不确定性和外部扰动等因素,为降低这类因素对系统的影响,业内学者们对鲁棒控制问题进行了大量的研究工作,如 LIU 等^[16]通过设计合适的成本函数实现不确定系统和标称系统之间的等价性;LIU 等^[17]将鲁棒控制问题转换为最优控制问题进行求解,极大地保证了系统的性能.在实际的工程背景下,以上描述的限制条件和影响因素需要有效克服,本文综合考虑了输入约束和扰动因素的影响,以满足实际中的真实需求.

基于上述分析,本文基于 ADP 方法,利用自触发间歇控制机制对带有输入约束的连续非线性系统的最优控制问题进行研究.首先,将原控制问题通过一个带有修改后成本函数的标称系统转换为最优控制问题;另外,由于稳定性理论的进一步发展,本文介绍了半全局实际有限时间稳定性理论^[18-19](SGPFS);然后,采用两个 NN 分别逼近最优成本函数和最优控制输入,保证了系统的稳定性;

最后,通过仿真验证该方法的有效性.

符号说明: \mathbb{R}^n 代表 n 维列向量, $\mathbb{R}^{n \times m}$ 代表 $n \times m$ 阶矩阵, $\|\cdot\|$ 代表向量或矩阵范数, $\nabla = \partial/\partial x$ 代表梯度, $\lambda_{\min}(\cdot)$ 和 $\lambda_{\max}(\cdot)$ 分别代表矩阵的最小和最大特征值.

1 问题描述

考虑如下连续时间非线性仿射系统:

$$\dot{x}(t) = f(x(t)) + g(x(t))(u(t) + \Delta(x(t))), \quad (1)$$

其中 $x(t) \in \mathbb{R}^n$ 为系统的状态, $u(t) = [u_1(t), \dots, u_i(t), \dots, u_m(t)]^T \in \mathbb{R}^m$ 为控制输入,满足 $|u_i(t)| \leq \lambda$, $\lambda > 0$; $f(\cdot)$ 和 $g(\cdot)$ 都为连续可微函数, $f(\cdot) \in \mathbb{R}^n$ 是有界闭集 Ω 上的 Lipschitz 函数, $f(0) = 0$, $g(\cdot) \in \mathbb{R}^{n \times m}$ 是 Ω 上的有界函数; $\Delta(x) \in \mathbb{R}^n$ 是未知的有界扰动,满足 $\|\Delta(x)\| \leq \Delta_M(x)$, $\Delta_M(x)$ 是一个已知函数且 $\Delta_M(0) = 0$. 假设 $\Delta(0) = 0$, 使得 $x = 0$ 是 Ω 上的唯一平衡点,初始状态为 $x(0) = x_0$.

考虑系统(1)对应的标称系统:

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t). \quad (2)$$

因未知扰动的影响,受 LIU^[16]的启发,为实现系统(1)和标称系统(2)之间的等价性,且将鲁棒控制问题转换为最优控制问题,定义成本函数如下:

$$V(x(t)) = \int_t^\infty \left(\frac{1}{2} \|\Delta_M(x(s))\|^2 + r(x(s), u(s)) \right) ds, \quad (3)$$

其中 $r(x, u) = x^T Q x + \vartheta(u)$, $Q \in \mathbb{R}^{n \times n}$ 为对称正定矩阵,满足 $V(0) = 0$. 为简单起见,后续与时间 t 相关的变量或函数在表示时均省略 t . 对于输入约束问题, LIU^[17]引入一种非二次型函数 $\vartheta(u)$, 定义为:

$$\vartheta(u) = 2 \int_0^u \lambda \tanh^{-1} \left(\frac{\tau}{\lambda} \right) d\tau = 2 \sum_{i=1}^m \int_0^{u_i} \lambda \tanh^{-1} \left(\frac{\tau_i}{\lambda} \right) d\tau_i. \quad (4)$$

定义系统(1)的 Hamiltonian 如下:

$$H(x, u, \nabla V(x)) = \frac{1}{2} \|\Delta_M(x)\|^2 + r(x, u) + (\nabla V(x))^T (f(x) + g(x)u), \quad (5)$$

其中 $\nabla V(x) = \partial V(x)/\partial x$, 根据 Bellman 最优性原理可以得到连续时间 HJB 方程:

$$\min_{u \in \Lambda(\Omega)} H(x, u, \nabla V^*(x)) = 0, \quad (6)$$

其中 $\Lambda(\Omega)$ 为容许控制域, $V^*(x)$ 为最优成本函数, 对应时刻的最优控制输入 u^* 为:

$$u^* = \arg \min_{u \in \Lambda(\Omega)} H(x, u, \nabla V^*(x)). \quad (7)$$

已知最优控制满足一阶必要性条件,通过求解 $\partial H(x, u, \nabla V(x))/\partial u = 0$,最小化(3)式得:

$$u^* = -\lambda \tanh(v^*), \quad (8)$$

其中 $v^* = \frac{1}{2\lambda} g^T(x) \nabla V^*(x)$. 通过(5)和(8)式, HJB 方程(6)可以写成:

$$\frac{1}{2} \|\Delta_M(x)\|^2 + x^T Q x + 2 \int_0^u \lambda \tanh^{-1}\left(\frac{\tau}{\lambda}\right) d\tau + (\nabla V^*(x))^T (f(x) + g(x)u^*) = 0. \quad (9)$$

本文控制目标为:设计一个满足约束条件的控制策略 u^* 使带有扰动的闭环系统(1)是稳定的,同时最小化成本函数.

成本函数(3)的设计是为了克服扰动因素的影响和实现输入约束的限制,结合后续的稳定性的证明,其首项评估每一时刻的扰动;而 $\vartheta(u)$ 的设计有效地将控制输入约束到指定范围. 受 LIU^[16] 的启发,本文设计的成本函数可以确保在标称系统中实现与原系统的等价性,将鲁棒控制问题转换为最优控制问题,再结合现有理论实现本文的控制目标.

2 自触发间歇控制器设计

本节首先介绍一种周期性间歇控制方法,通过逼近连续时间 HJB 方程,设计对应的周期性间歇控制器,且保证系统的稳定性. 间歇控制的目的在于减少通信资源的浪费,在系统稳定的前提下,减少采样量的同时最小化成本,而周期性提高了采样效率. 控制器以间歇方式运行,当系统处于工作区间,控制器运行;当系统处于休息区间,控制器关闭. 为了设计周期性间歇控制器,将 $[0, +\infty)$ 划分为一系列子区间 $[T_i, T_{i+1})$, $i = 0, 1, \dots$, 定义一个周期的长度为 $T_\Delta = T_{2i+2} - T_{2i}$, 一个周期内工作区间和休息区间的长度分别为 $T_1 = \delta T_\Delta$, $T_2 = (1 - \delta) T_\Delta$, 其中 $0 < \delta < 1$, 周期性间歇控制器的工作时间为 $t \in T^w = \bigcup_{i=0}^{\infty} [T_{2i}, T_{2i+1})$, 休息时间为 $t \in T^r = \bigcup_{i=0}^{\infty} [T_{2i+1}, T_{2(i+1)})$. 初始时刻为 $T_0 = 0$, 满足 $0 = T_0 < T_1 < T_2 < T_3 \dots$. 而后引入一种自触发机制,其触发时刻序列为 $\{t_k^i\}_{i \in \mathbb{N}, k \in \mathbb{N}}$, 满足 $t_0^i = T_{2i}$, $t_k^i \in \mathbb{R}^+$. 定义误差:

$$e = x_k - x(t), \quad t \in [t_k^i, t_{k+1}^i),$$

其中 $x_k = x(t_k^i)$.

考虑如下间歇最优控制器:

$$\bar{u}^* = \begin{cases} -\lambda \tanh(v_k^*), & t \in [t_k^i, t_{k+1}^i) \cap [T_{2i}, T_{2i+1}) \\ 0, & t \in [T_{2i+1}, T_{2(i+1)}) \end{cases}, \quad (10)$$

其中 $v_k^* = \frac{1}{2\lambda} g^T(x_k) \nabla V^*(x_k)$, 根据间歇控制思想将系统(2)写为:

$$\dot{x} = \begin{cases} f(x) + g(x)u_k^*, & t \in [t_k^i, t_{k+1}^i) \cap [T_{2i}, T_{2i+1}) \\ f(x), & t \in [T_{2i+1}, T_{2(i+1)}) \end{cases}, \quad (11)$$

其中 $u_k^* = u^*(x_k)$.

接下来给出半全局实际有限时间稳定性的推广引理:

引理 1^[11] 对于连续时间非线性系统(1), 若存在连续可微的正定函数 $V(t)$, 满足以下条件:

$$\dot{V}(t) \leq \begin{cases} -\beta_1 V^\alpha(t) + \gamma, & t \in [T_{2i}, T_{2i+1}) \\ \beta_2 V(t) + \gamma, & t \in [T_{2i+1}, T_{2(i+1)}) \end{cases}, \quad (12)$$

其中 $\beta_1, \beta_2, \gamma > 0, \alpha, \varpi \in (0, 1)$, 则以下不等式成立:

$$V^{1-\alpha}(t) \leq 1 + V^{1-\alpha}(0) e^{(\beta_1 + \beta_2)(1-\alpha)\delta t} - \varpi \beta_1 (1-\alpha)(1-\delta)t, \quad 0 \leq t \leq T,$$

其中 $\delta \in (0, 1)$ 为一个周期内工作区间的长度, T 是一个常数. 若满足 $e^{1 + \frac{(\beta_1 + \beta_2)\delta}{\varpi \beta_1 (1-\alpha)}} V^{1-\alpha}(0) \leq \frac{\varpi \beta_1 (1-\alpha)}{(\beta_1 + \beta_2)\delta}$,

设定的时间为 $T(x_0) = \frac{1}{(\beta_1 + \beta_2)(1-\alpha)\delta} + \frac{1}{\varpi \beta_1 (1-\alpha)(1-\delta)}$, 那么系统(1)是 SGPFs.

基于间歇控制思想和引理 1, 定理 1 给出对应的稳定性证明. 在此之前, 先介绍如下两个假设:

假设 1^[13] 假设 $f(x)$ 在包含原点的有界闭集 $\Omega \subseteq \mathbb{R}^n$ 上是 Lipschitz 连续的, $f(0) = 0$, $g(x)$ 在 Ω 上是有界函数, u^* 和 v^* 是 Lipschitz 连续的, 且 $\lambda - \tanh(\cdot)$ 是 λ -Lipschitz 连续函数, 即满足:

$$\|u^* - u_k^*\| \leq L_u \|x - x_k\| = L_u \|e\|,$$

$$\|f(x) + g(x)u_k^*\| \leq L_f \|x\| + L_g \|e\|,$$

$$\|\lambda \tanh(v^*) - \lambda \tanh(v_k^*)\| \leq \lambda \|v^* - v_k^*\| \leq \lambda L_v \|e\|,$$

其中 L_u, L_f, L_g 和 L_v 都是正数.

假设 2^[13] 假设存在 3 个正数 $\alpha_1, \alpha_2, \alpha_3$ 使得最优成本函数及其导数满足:

$$\alpha_1 \|x\|^2 \leq V^*(x) \leq \alpha_2 \|x\|^2,$$

$$\|\nabla V^*(x)\| \leq \alpha_3 \|x\|.$$

定理 1 考虑标称系统(2). 在假设 1 和假设 2

下,若存在一个满足 HJB 方程(9)的最优成本函数的导数 $\nabla V^*(x)$,且有正数 $\theta, \tau \in (0, 1)$ 满足:

$$\|e\|^2 \leq \frac{(1-\theta^2)(1-\tau^2)\lambda_{\min}(Q)\|x_k\|^2}{L_a^2 + \left(\frac{1}{\tau^2} - 1\right)\lambda_{\max}(Q)} = e_r, \quad (13)$$

其中 $L_a^2 = 2\lambda L_u L_v + L_g^2 \alpha_3^2$, 那么系统(1)在间歇最优控制器(10)的作用下是 SGPFS.

证明 选取 Lyapunov 函数 L_1 :

$$L_1 = V^*(x),$$

当 $t \in [t_k^i, t_{k+1}^i) \cap [T_{2i}, T_{2i+1})$ 时,可得:

$$\dot{L}_1 = (\nabla V^*(x))^T (f(x) + g(x)(u_k^* + \Delta(x))). \quad (14)$$

根据(8)和(9)式可知满足:

$$(\nabla V^*(x))^T g(x) = -2\lambda \tanh^{-1}\left(\frac{u^*}{\lambda}\right)^T, \quad (15)$$

$$\begin{aligned} (\nabla V^*(x))^T (f(x) + g(x)u^*) &= -\frac{1}{2}\Delta_M^2(x) - x^T Qx - \\ &2 \int_0^{u^*} \lambda \tanh^{-1}\left(\frac{\tau}{\lambda}\right) d\tau, \end{aligned} \quad (16)$$

另外,由误差定义有:

$$\begin{aligned} x^T Qx &= x_k^T Qx_k - 2x_k^T Qe + e^T Qe = (1-\tau^2)x_k^T Qx_k - \\ &\left(\frac{1}{\tau^2} - 1\right)e^T Qe + \left(\tau x_k - \frac{1}{\tau}e\right)^T Q \left(\tau x_k - \frac{1}{\tau}e\right) \geq \\ &(1-\tau^2)x_k^T Qx_k - \left(\frac{1}{\tau^2} - 1\right)e^T Qe, \end{aligned}$$

将(15)、(16)式代入(14)式,且满足 $-\vartheta(u_k^*) \leq 0$, 结合(13)式可知:

$$\begin{aligned} \dot{L}_1 &= -\frac{1}{2}\|\Delta_M(x)\|^2 - x^T Qx - 2 \int_0^{u^*} \lambda \tanh^{-1}\left(\frac{\tau}{\lambda}\right) d\tau - \\ &2\lambda \tanh^{-1}\left(\frac{u^*}{\lambda}\right)^T (u_k^* - u^*) + (\nabla V^*(x))^T g(x)\Delta(x) \leq \\ &-\frac{1}{2}\|\Delta_M(x)\|^2 - (1-\tau^2)x_k^T Qx_k + \left(\frac{1}{\tau^2} - 1\right)e^T Qe + \\ &2\lambda L_u L_v \|e\|^2 - 2 \int_0^{u^*} \lambda \tanh^{-1}\left(\frac{\tau}{\lambda}\right) d\tau + \\ &\frac{1}{2}(\nabla V^*(x))^T g(x)g(x)^T (\nabla V^*(x)) + \\ &\frac{1}{2}\Delta^2(x) \leq -[(1-\theta^2)(1-\tau^2)\lambda_{\min}(Q) - \\ &L_g^2 \alpha_3^2]\|x_k\|^2 + \left[L_a^2 + \left(\frac{1}{\tau^2} - 1\right)\lambda_{\max}(Q)\right]\|e\|^2 - \\ &\theta^2(1-\tau^2)\lambda_{\min}(Q)\|x_k\|^2 - \vartheta(u_k^*) + \frac{1}{2}\Delta^2(x) - \\ &\frac{1}{2}\|\Delta_M(x)\|^2 \leq -\eta_1 L_1 + \eta_1^\alpha L_1^\alpha - \eta_1^\alpha L_1^\alpha \leq -\eta_1^\alpha L_1^\alpha + \gamma_1, \end{aligned} \quad (17)$$

其中 $\eta_1 = \alpha_1^{-1}(\theta^2(1-\tau^2)\lambda_{\min}(Q) - L_g^2 \alpha_3^2)$, $\gamma_1 = (1 -$

$\alpha)\alpha^{\frac{\alpha}{(1-\alpha)}}$, $\alpha > 0$. 当 $t \in [T_{2i+1}, T_{2(i+1)})$ 时,控制器停止运行,根据系统(11)可得:

$$\begin{aligned} \dot{L}_1 &= (\nabla V^*(x))^T f(x) \leq \\ L_f \alpha_3 \|x\|^2 &\leq \eta_2 L_1 + \gamma_1, \end{aligned} \quad (18)$$

其中 $\eta_2 = \alpha_1^{-1}(L_f \alpha_3)$.

因此,基于引理 1,有 $\beta_1 = \eta_1^\alpha, \beta_2 = \eta_2, \gamma = \gamma_1$, 则选取的 Lyapunov 函数 L_1 是满足条件的,结合两种情形可知系统(1)是 SGPFS.

3 神经网络的实现

非线性 HJB 方程是一种偏微分方程,通常难以求解,本节通过应用 ADP 方法的评价网络和执行网络分别逼近最优成本函数 $V^*(x)$ 和最优控制 u^* .

(1)评价网络表示为:

$$V^*(x) = W_c^T \phi_c(x) + \varepsilon_c,$$

其中 $W_c \in \mathbb{R}^{N_1}$ 为评价网络的理想权值, $\phi_c(x) \in \mathbb{R}^{N_1}$ 为评价网络的激活函数, N_1 为隐藏层的神经元数量, ε_c 为评价网络的重构误差.

最优成本函数的导数表示为:

$$\nabla V^* = \nabla \phi_c^T W_c + \nabla \varepsilon_c,$$

其中 $\nabla V^* = \partial V^*(x)/\partial x, \nabla \phi_c = \partial \phi_c(x)/\partial x, \nabla \varepsilon_c = \partial \varepsilon_c(x)/\partial x$ 将上式代入(8)式,触发时刻的控制器可以表示为:

$$u_{ck}^* = -\lambda \tanh(v_k^*), \quad (19)$$

其中 $v_k^* = \frac{1}{2\lambda} g^T(x_k)(\nabla \phi_c^T W_c + \nabla \varepsilon_c)$. 利用评价网络逼近 $V^*(x)$:

$$\hat{V}^*(x) = \hat{W}_c^T \phi_c(x),$$

其中 \hat{W}_c 为评价网络的估计权值, $\hat{V}^*(x)$ 是 $V^*(x)$ 的估计.

因此,间歇最优控制器(10)的评价网络形式可以表示为:

$$\hat{u}^* = \begin{cases} -\lambda \tanh(\hat{v}_k^*), & t \in [t_k^i, t_{k+1}^i) \cap [T_{2i}, T_{2i+1}) \\ 0, & t \in [T_{2i+1}, T_{2(i+1)}) \end{cases}, \quad (20)$$

其中 $\hat{v}_k^* = \frac{1}{2\lambda} g^T(x_k) \nabla \phi_c^T \hat{W}_c$. 由上述形式可知, HJB 方程的近似形式可以写成:

$$\begin{aligned} H(x, \hat{u}^*, \hat{W}_c) &= \hat{W}_c^T \nabla \phi_c (f(x) + g(x)\hat{u}^*) + \\ x^T Qx + \vartheta(\hat{u}^*) &= e_c. \end{aligned}$$

以下采用梯度下降法最小化目标函数 $E_c =$

$\frac{1}{2}e_c^T e_c$, 得到评价网络权值更新律. 如果控制器处于休息区间, 评价网络权值的更新律为 $\dot{\hat{W}}_c = 0$, 当控制器工作时, 评价网络权值的更新律为:

$$\hat{W}_c^+ = \hat{W}_c - \alpha_c \sigma \left(\sigma^T \hat{W}_c + x^T Q x + \vartheta(\hat{u}^*) \right),$$

其中 $\alpha_c > 0$ 是评价网络的学习率, $\sigma = \sigma_1 / (\sigma_1^T \sigma_1 + 1)$, $\sigma_1 = \nabla \phi_c(f(x) + g(x)\hat{u}^*)$, 且满足 $\sigma_m \leq \sigma \leq \sigma_M$.

定义评价网络权值的估计误差为 $\tilde{W}_c = W_c - \hat{W}_c$, 有:

$$H(x, \bar{u}^*, W_c) = W_c^T \nabla \phi_c(f(x) + g(x)\bar{u}^*) + x^T Q x + \vartheta(\bar{u}^*) = \varepsilon_H,$$

其中 $\varepsilon_H = -\nabla \varepsilon_c(f(x) + g(x)\bar{u}^*)$, 由此可以得出:

$$\begin{cases} \dot{\tilde{W}}_c = 0, & t \neq t_k^i \\ \tilde{W}_c^+ = \tilde{W}_c - \alpha_c \sigma (\sigma^T \tilde{W}_c + \vartheta(\bar{u}^*) - \vartheta(\hat{u}^*) + \varepsilon_H), & t = t_k^i \end{cases} \quad (21)$$

(2) 执行网络表示为:

$$u^* = W_a^T \phi_a(x) + \varepsilon_a,$$

其中 $W_a \in \mathbb{R}^{N_2}$ 为执行网络的理想权值, $\phi_a(x) \in \mathbb{R}^{N_2}$ 为执行网络的激活函数, N_2 为隐藏层的神经元数量, ε_a 为执行网络的重构误差.

利用执行网络逼近 u^* :

$$\hat{u}^* = \hat{W}_a^T \phi_a(x), \quad (22)$$

其中 \hat{W}_a 为执行网络的估计权值, \hat{u}^* 是 u^* 的估计.

因此, 执行网络的近似误差可以表示为:

$$e_a = \hat{W}_a^T \phi_a(x) + \lambda \tanh\left(\frac{1}{2\lambda} g^T(x) \nabla \phi_c^T \hat{W}_c\right).$$

以下采用梯度下降法最小化目标函数 $E_a = \frac{1}{2}e_a^T e_a$, 得到执行网络权值更新律. 如果控制器处于休息区间, 执行网络权值的更新律为 $\dot{\hat{W}}_a = 0$, 当控制器工作时, 执行网络权值的更新律为:

$$\hat{W}_a^+ = \hat{W}_a - \alpha_a \phi_a \left(\hat{W}_a^T \phi_a(x_k) - \hat{u}_{ck}^* \right)^T,$$

其中 $\hat{u}_{ck}^* = -\lambda \tanh\left(\frac{1}{2\lambda} g^T(x_k) \nabla \phi_c^T \hat{W}_c\right)$, $\alpha_a > 0$ 是执行网络的学习率, 执行网络权值的估计误差为 $\tilde{W}_a = W_a - \hat{W}_a$, 则有:

$$u_k^* - u_{ck}^* = W_a^T \phi_a(x_k) + \varepsilon_a + \lambda \tanh\left(\frac{1}{2\lambda} g^T(x_k) (\nabla \phi_c^T W_c + \nabla \varepsilon_c)\right) = 0.$$

由此可得:

$$\begin{cases} \dot{\tilde{W}}_a = 0, & t \neq t_k^i \\ \tilde{W}_a^+ = \tilde{W}_a - \alpha_a \phi_a \left(\tilde{W}_a^T \phi_a(x_k) + (u_{ck}^* - \hat{u}_{ck}^*) + \varepsilon_a \right)^T, & t = t_k^i \end{cases} \quad (23)$$

4 稳定性分析

假设 3^[4] 假设评价网络的理想权值 W_c 、激活函数 ϕ_c 、激活函数的梯度 $\nabla \phi_c$ 、重构误差 ε_c 、重构误差的梯度 $\nabla \varepsilon_c$ 、残差 ε_H 分别满足 $\|W_c\| \leq W_{cM}$ 、 $\|\phi_c\| \leq \phi_{cM}$ 、 $\|\nabla \phi_c\| \leq \phi_{cM}$ 、 $\|\varepsilon_c\| \leq \varepsilon_{cM}$ 、 $\|\nabla \varepsilon_c\| \leq \varepsilon_{cM}$ 、 $\|\varepsilon_H\| \leq \varepsilon_{HM}$; 执行网络的参数满足 $\|W_a\| \leq W_{aM}$ 、 $\|\phi_a\| \leq \phi_{aM}$ 、 $\|\varepsilon_a\| \leq \varepsilon_{aM}$, 即均有界.

定理 2 考虑标称系统(2)和HJB方程(9). 在假设 1~假设 3 下, 采用间歇最优控制律(20)和权值更新律(21)、(23), 且触发条件满足:

$$\|e\|^2 \leq \frac{(1-\theta^2)(1-\tau^2)\lambda_{\min}(Q)\|x_k\|^2}{L_b^2 + \left(\frac{1}{\tau^2} - 1\right)\lambda_{\max}(Q)} = e_T, \quad (24)$$

其中 $L_b^2 = 4\lambda^2 L_u^2 L_v^2$, 则系统(2)是SGPFS.

证明 首先, 分析权值误差 \tilde{W}_c 和 \tilde{W}_a 的收敛性, 选取 Lyapunov 函数 L :

$$L = L_2 + L_3,$$

其中 $L_2 = (\alpha_c)^{-1} \text{tr}\{\tilde{W}_c^T \tilde{W}_c\}$, $L_3 = (\alpha_a)^{-1} \text{tr}\{\tilde{W}_a^T \tilde{W}_a\}$. 当 $t = t_k^i$ 时, 可得:

$$\Delta L = \Delta L_2 + \Delta L_3. \quad (25)$$

可知 $\|\vartheta(\bar{u}^*) - \vartheta(\hat{u}^*)\|$ 满足不等式:

$$\|\vartheta(\bar{u}^*) - \vartheta(\hat{u}^*)\| = \left\| 2 \int_{\hat{u}^*}^{\bar{u}^*} \lambda \tanh^{-1}\left(\frac{\tau}{\lambda}\right) d\tau \right\| \leq$$

$$2\lambda \|v_k^*\| \|\bar{u}^* - \hat{u}^*\| \leq L_g \phi_M W_{cM} \|\bar{u}^* - \hat{u}^*\| \leq$$

$$\lambda L_g \phi_M W_{cM} \|v^* - v_k^*\| \leq \frac{1}{2} L_g^2 \phi_M^2 W_{cM} (\|\tilde{W}_c\| + \varepsilon_M),$$

由此可得:

$$\begin{aligned} \Delta L_2 &= (\alpha_c)^{-1} \Delta \text{tr}\{\tilde{W}_c^T \tilde{W}_c\} = (\alpha_c)^{-1} \left(\text{tr}\{\tilde{W}_c^{*T} \tilde{W}_c^+\} - \text{tr}\{\tilde{W}_c^T \tilde{W}_c\} \right) \leq -2\sigma \left(\sigma^T \tilde{W}_c + \vartheta(\bar{u}^*) - \vartheta(\hat{u}^*) + \varepsilon_H \right) \tilde{W}_c + \\ &\quad \alpha_c \sigma^2 \left(\sigma^T \tilde{W}_c + \vartheta(\bar{u}^*) - \vartheta(\hat{u}^*) + \varepsilon_H \right)^2 \leq \\ &\quad - \left(1 + \frac{3}{2} \alpha_c b_1^2 + b_1^2 - \frac{b_1}{\|\sigma\|} - 3\alpha_c \|\sigma\|^2 \right) \|\sigma\|^2 \|\tilde{W}_c\|^2 + \\ &\quad (3\alpha_c \|\sigma\|^2 + 1) \varepsilon_{HM}^2 + \left(\frac{3}{2} \alpha_c b_1^2 \|\sigma\|^2 + 1 \right) \varepsilon_M^2, \quad (26) \end{aligned}$$

$$\Delta L_3 = (\alpha_a)^{-1} \Delta \text{tr}\{\tilde{W}_a^T \tilde{W}_a\} = (\alpha_a)^{-1} \left(\text{tr}\{\tilde{W}_a^{*T} \tilde{W}_a^+\} - \text{tr}\{\tilde{W}_a^T \tilde{W}_a\} \right) \leq -2\phi_a \left(\tilde{W}_a \phi_a + (u_{ck}^* - \hat{u}_{ck}^*) + \varepsilon_a \right) \tilde{W}_a +$$

$$\alpha_a \phi_a^2 \left(\tilde{W}_a \phi_a + (u_{ck}^* - \hat{u}_{ck}^*) + \varepsilon_a \right)^2 \leq -$$

$$\left(1 - 3\alpha_a \|\phi_a\|^2 \right) \|\phi_a\|^2 \|\tilde{W}_a\|^2 + \frac{3}{2} \alpha_a b_2^2 \|\tilde{W}_c\|^2 +$$

$$b_2 \|\tilde{W}_c\| \|\tilde{W}_a\| + \left(3\alpha_a \|\phi_a\|^2 + 1 \right) \varepsilon_{aM}^2 + \frac{3}{2} \alpha_a b_2^2 \varepsilon_M^2, \quad (27)$$

其中 $b_1 = L_g^2 \phi_M^2 W_{cM}$, $b_2 = L_g \phi_M \phi_{aM}$. 根据上述不等式可得:

$$\begin{aligned} \Delta L \leq & - \left(1 + \frac{3}{2} \alpha_c b_1^2 + b_1^2 - \frac{b_1}{\|\sigma\|} - 3\alpha_c \|\sigma\|^2 \right) \|\sigma\|^2 \|\tilde{W}_c\|^2 + \\ & (3\alpha_c \|\sigma\|^2 + 1) \varepsilon_{HM}^2 + \left(\frac{3}{2} \alpha_c b_1^2 \|\sigma\|^2 + 1 \right) \varepsilon_M^2 - \\ & (1 - 3\alpha_a \|\phi_a\|^2) \|\phi_a\|^2 \|\tilde{W}_a\|^2 + \frac{3}{2} \alpha_a b_2^2 \|\tilde{W}_c\|^2 + \\ & b_2 \|\tilde{W}_c\| \|\tilde{W}_a\| + (3\alpha_a \|\phi_a\|^2 + 1) \varepsilon_{aM}^2 + \frac{3}{2} \alpha_a b_2^2 \varepsilon_M^2 \leq \\ & \left[- \left(1 + \left(\frac{3}{2} \alpha_c + 1 \right) b_1^2 - \frac{b_1}{\|\sigma\|} - 3\alpha_c \|\sigma\|^2 \right) \|\sigma\|^2 + \right. \\ & \left. \frac{3}{2} \alpha_a b_1^2 \right] \|\tilde{W}_c\|^2 - (1 - 3\alpha_a \|\phi_a\|^2) \|\phi_a\|^2 \|\tilde{W}_a\|^2 + \\ & b_2 \|\tilde{W}_c\| \|\tilde{W}_a\| + (3\alpha_c \|\sigma\|^2 + 1) \varepsilon_{HM}^2 + (3\alpha_a \|\phi_a\|^2 + 1) \varepsilon_{aM}^2 + \\ & \left(\frac{3}{2} \alpha_c b_1^2 \|\sigma\|^2 + \frac{3}{2} \alpha_a b_2^2 + 1 \right) \varepsilon_M^2 \leq -\chi^T A \chi + \varepsilon^2, \end{aligned} \quad (28)$$

其中 $A = \begin{bmatrix} D_3 & -b_1 \\ 0 & (1 - 3\alpha_a \|\phi_a\|^2) \|\phi_a\|^2 \end{bmatrix}$, $D_3 = \left(1 + \left(\frac{3}{2} \alpha_c + 1 \right) b_1^2 - \frac{b_1}{\|\sigma\|} - 3\alpha_c \|\sigma\|^2 \right) \|\sigma\|^2 + \frac{3}{2} \alpha_a b_1^2$, $\chi = [\|\tilde{W}_c\|, \|\tilde{W}_a\|]^T$, $\varepsilon^2 = (3\alpha_c \|\sigma\|^2 + 1) \varepsilon_{HM}^2 + (3\alpha_a \|\phi_a\|^2 + 1) \varepsilon_{aM}^2 + \left(\frac{3}{2} \alpha_c b_1^2 \|\sigma\|^2 + \frac{3}{2} \alpha_a b_2^2 + 1 \right) \varepsilon_M^2$. 若 A 的特征值都为正数, 且当 $\|\chi\| \geq \sqrt{\varepsilon^2 / \lambda_{\max}(A)}$ 时, $\Delta L \leq 0$. 另外, 当 $t \neq t_k^i$ 时, 权值向量保持不变, 有 $\dot{L} = 0$, 因此, 当满足 A 的特征值都为正数时, 权值误差 \tilde{W}_c 和 \tilde{W}_a 是 UUB.

接下来, 分析系统状态的收敛性, 选取 Lyapunov 函数为定理 1 中的 L_1 , 分情况讨论.

情形 1: 当 $t \in [T_{2i}, T_{2i+1})$ 时, 由假设 1 和杨不等式 $2a^T b \leq qa^T a + (1/q)b^T b$, $q > 0$, 另借助 (19) 式和 (20) 式可知:

$$\begin{aligned} \|\mathbf{v}^* - \hat{\mathbf{v}}_k^*\|^2 & \leq \|\mathbf{v}^* - \mathbf{v}_k^* + \mathbf{v}_k^* - \hat{\mathbf{v}}_k^*\|^2 \leq \\ 2\|\mathbf{v}^* - \mathbf{v}_k^*\|^2 + 2\|\mathbf{v}_k^* - \hat{\mathbf{v}}_k^*\|^2 & \leq 2L_v^2 \|e\|^2 + \\ \frac{1}{2\lambda^2} \left\| g^T(x_k) (\nabla \phi_c^T \tilde{W}_c + \nabla \varepsilon_c) \right\|^2 & \leq \\ 2L_v^2 \|e\|^2 + \lambda^{-2} L_g^2 \phi_M^2 \|\tilde{W}_c\|^2 + \lambda^{-2} L_g^2 \varepsilon_M^2. \end{aligned} \quad (29)$$

已知 $-\vartheta(\hat{u}^*) \leq 0$, 满足 (24) 式和 (25) 式, 则可得:

$$\begin{aligned} \dot{L}_1 \leq & -\frac{1}{2} \|\Delta_M(x)\|^2 - x^T Q x - 2 \int_0^u \lambda \tanh^{-1} \left(\frac{\tau}{\lambda} \right) d\tau - \\ & 2\lambda \tanh^{-1} \left(\frac{u^*}{\lambda} \right)^T (\hat{u}^* - u^*) \leq -\frac{1}{2} \|\Delta_M(x)\|^2 - x^T Q x - \\ & 2 \int_0^{\hat{u}^*} \lambda \tanh^{-1} \left(\frac{\tau}{\lambda} \right) d\tau - 2 \int_{\hat{u}^*}^u \lambda \tanh^{-1} \left(\frac{\tau}{\lambda} \right) d\tau + \\ & 2 \int_{\hat{u}^*}^u \lambda \tanh^{-1} \left(\frac{u^*}{\lambda} \right) d\tau \leq -\frac{1}{2} \|\Delta_M(x)\|^2 - \vartheta(\hat{u}^*) - \\ & (1 - \tau^2) x_k^T Q x_k + \left(\frac{1}{\tau^2} - 1 \right) e^T Q e + L_b^2 \|e\|^2 + \\ & 2L_u^2 L_g^2 \phi_M^2 \|\tilde{W}_c\|^2 + 2L_g^2 \varepsilon_M^2 \leq -(1 - \tau^2) \beta^2 \lambda_{\min}(Q) \|x_k\|^2 - \\ & \vartheta(\hat{u}^*) - \frac{1}{2} \|\Delta_M(x)\|^2 + 2L_u^2 L_g^2 \phi_M^2 \|\tilde{W}_c\|^2 + 2L_g^2 \varepsilon_M^2 \leq -\theta_1 L_1 + \\ & \theta_1^\alpha L_1^\alpha - \theta_1^\alpha L_1^\alpha + 2L_u^2 L_g^2 \phi_M^2 W_M^2 + 2L_g^2 \varepsilon_M^2 \leq -\theta_1 L_1 + \psi_1, \end{aligned} \quad (30)$$

其中 $\theta_1 = \alpha_1^{-1} \left((1 - \tau^2) \beta^2 \lambda_{\min}(Q) \right)$, $\psi_1 = 2L_u^2 L_g^2 \phi_M^2 W_M^2 + 2L_g^2 \varepsilon_M^2 + (1 - \alpha) \alpha^{\frac{\alpha}{1-\alpha}}$, $\|\tilde{W}_c\| \leq W_M$.

情形 2: 当 $t \in [T_{2i+1}, T_{2(i+1)})$ 时, 接下来有:

$$\begin{aligned} \dot{L}_1 & \leq (\nabla V^*(x))^T f(x) \leq \\ \alpha_1^{-1} (L_f \alpha_3) V^*(x) & \leq \theta_2 L + \psi_2, \end{aligned} \quad (31)$$

其中 $\theta_2 = \alpha_1^{-1} (L_f \alpha_3)$, $\psi_2 = \gamma_1$.

综合对上述两种情况的分析和引理 1, 有 $\beta_1 = \theta_1^\alpha$, $\beta_2 = \theta_2$, $\gamma = \max\{\psi_1, \psi_2\}$, 则选取的 Lyapunov 函数 L_1 是满足条件的, 可知间歇最优控制律 (20) 能确保系统 (2) 是 SGPFs, 定理 2 得证.

接下来, 给出自触发机制的触发规则, 并且表明芝诺行为是可以避免的.

$$\begin{aligned} \|\dot{x}\| & \leq \|f(x)\| + \|g(x)(u_k^* + \Delta(x))\| \leq \\ L_f \|x\| + L_g L_u (\|e\| + \lambda) + L_g \Delta_M & \leq \\ l_1 \|x\| + l_2 \|e\| + l_3, \end{aligned}$$

其中 $l_1 = L_f$, $l_2 = L_g L_u$, $l_3 = L_g L_u \lambda + L_g \Delta_M$, 且满足 $\|u_k^*\| \leq \|u^* - u_k^*\| + \|u^*\| \leq L_u \|e\| + \lambda$.

对于 $t \in [t_k^i, t_{k+1}^i) \cap [T_{2i}, T_{2i+1})$, 有:

$$\|\dot{e}\| = \|\dot{x}\| \leq l_1 \|x_k\| + l_2 \|e\| + l_3,$$

由比较原理可知:

$$\|e\| \leq \frac{l_1 \|x_k\| + l_3}{l_1 + l_2} \left(e^{(l_1 + l_2)(t - t_k^i)} - 1 \right),$$

在触发时刻 $t = t_{k+1}^i$, 满足:

$$\frac{l_1 \|x_k\| + l_3}{l_1 + l_2} \left(e^{(l_1 + l_2)(t - t_k^i)} - 1 \right) \geq e(t_{k+1}^i) > \sqrt{e_T},$$

那么可以得到:

$$t_{k+1}^i - t_k^i > \frac{1}{l_1 + l_2} \ln \left(1 + \frac{l_1 + l_2}{l_1 \|x_k\| + l_3} \sqrt{e_T} \right) > 0. \quad (32)$$

基于上述分析,设计的触发规则如下:

$$t_{k+1}^i = t_k^i + \frac{1}{l_1 + l_2} \ln \left(1 + \frac{l_1 + l_2}{l_1 \|x_k\| + l_3} \sqrt{e_T} \right). \quad (33)$$

由设计的自触发时刻的关系可知,自触发机制在当前时刻计算下一触发时刻,避免了计算资源的消耗,其触发间隔相较于事件触发机制更加保守,由定理 2 可知,系统在自触发机制下是满足稳定性要求的,且通过(32)式可知两个相邻触发时刻的差大于零,因此,该机制避免了芝诺行为.以下给出的定理 3 表明系统在自触发机制下的稳定性,具体证明可通过定理 2 得出,故省略.另外,图 1 展示了该方案的框架,增强了理论的可实现性.

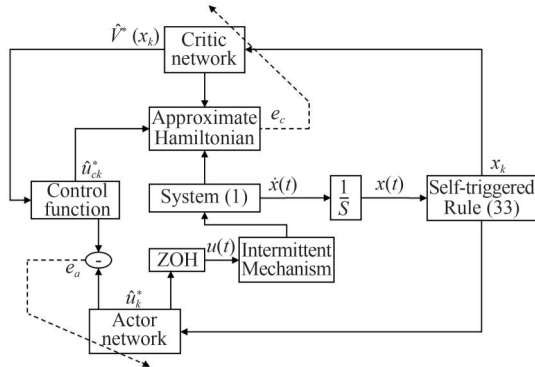


图 1 基于自适应动态规划的自触发间歇控制方案框架图

Fig. 1 The block diagram of the ADP-based self-triggered intermittent control scheme

定理 3 考虑标称系统(2)和 HJB 方程(9).在假设 1~假设 3 下,采用间歇最优控制律(20)和权值更新律(21)、(23),自触发规则设置为(33)式,则系统(2)是 SGPFPS.

5 仿真结果

考虑如下非线性系统:

$$\dot{x} = f(x) + g(x)(u + \Delta(x)),$$

$$\text{其中 } x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, f(x) = \begin{bmatrix} -x_1 + x_2 \\ -0.5x_1 - 0.5x_2 + 0.5x_2((\cos(2x_1) + 2))^2 \end{bmatrix}, g(x) = \begin{bmatrix} 0 \\ \cos(2x_1) + 2 \end{bmatrix}, u \text{ 为控制输入,满足}$$

$|u| \leq 6.5, \Delta(x) = 0.5px_1 \sin(x_2), p \in [-1, 1]$.仿真中初始值设为 $x_0 = [1, -1]^T, Q = \text{diag}(2, 2), \lambda = 6.5, \Delta_M(x) = \|x\|$,选取周期 $T_\Delta = 10\text{s}$,工作区间长度 $T_1 =$

9.5s,在自触发间歇机制下得到最优控制策略,且应用到系统.在神经网络中,评价网络的激活函数选择为 $\phi_c = [x_1^2 \ x_1 \ x_2 \ x_2^2]^T$,定义其权值 $W_c = [W_{c1} \ W_{c2} \ W_{c3}]^T$ 且初始权值设置为 $[1, 1, 1]^T$.执行网络的激活函数选择 $\phi_a = [x_1^2 \ x_1 \ x_2 \ x_2^2]^T$,定义执行网络权值 $W_a = [W_{a1} \ W_{a2} \ W_{a3}]^T$ 且初始权值的设置使初始控制是容许的.另外,通过在仿真中对相关参数的优化调整,评价网络和执行网络的学习率分别设为 0.5 和 0.5,触发阈值的设置满足性能需求.

图 2 中展示了系统状态的收敛过程;带有输入约束的自触发间歇控制收敛过程见图 3,其有效地将控制值限制在指定范围,若没有考虑输入约束,控制值将会超过所设范围,就需要将控制值设置为上界;图 4 和图 5 分别呈现了评价网络和执行网络权值的变化过程,最终收敛到一个稳定的值;成本函数的变化过程见图 6,该图表明成本有限累积到一个固定值;图 7 对比了连续控制和自触发间歇控制的采样数量,连续控制需要的采样数量为 10000,而自触发间歇控制的采样数量为 3972,可知所提出的方法减少了约 60% 的通信资源.本文方法克服了系统中未知扰动的影响和输入约束的限制,减少了控制器的更新数量,提高了采样效率,体现出本文方法的优越性.

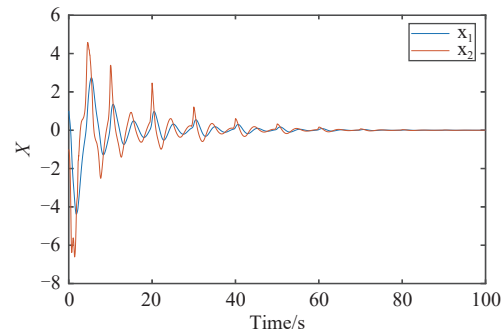


图 2 状态收敛过程图

Fig. 2 State convergence process diagram

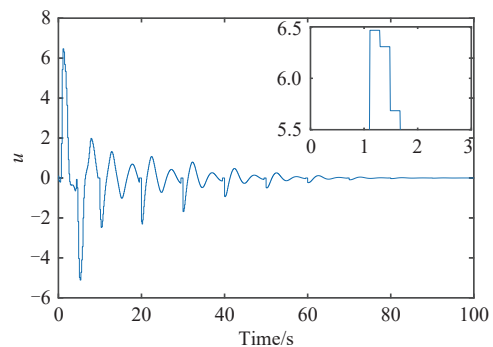


图 3 控制收敛过程图

Fig. 3 Control convergence process diagram

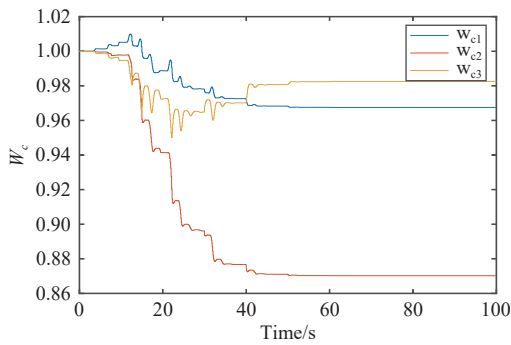


图 4 评价网络权值变化过程图

Fig. 4 Critic neural network weights evolution diagram

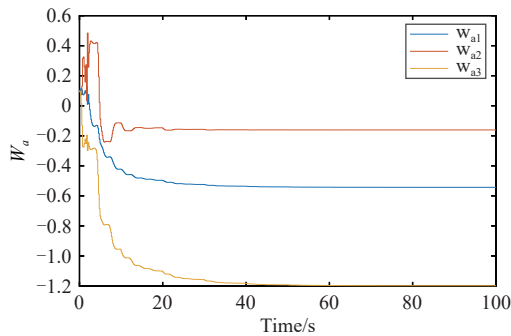


图 5 执行网络权值变化过程图

Fig. 5 Action neural network weights evolution diagram

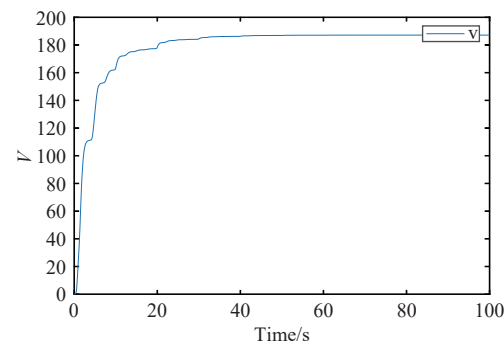


图 6 成本函数变化过程图

Fig. 6 Cost function evolution diagram

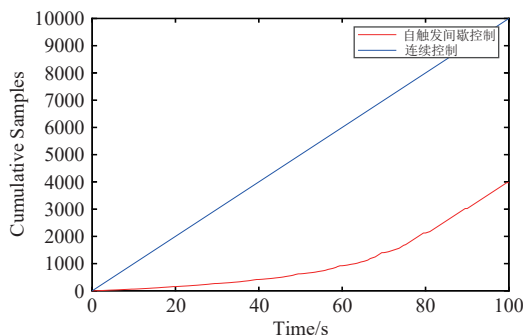


图 7 采样数量对比图

Fig. 7 Sample quantity comparison diagram

6 结论

本文运用 ADP 方法研究了带有输入约束的连

续时间非线性系统的最优控制问题. 采用自触发间歇控制器, 克服了系统内部扰动的影响, 有效地减少了采样数量, 提高了采样效率; 使用评价网络和执行网络分别逼近最优成本函数和最优控制输入, 且在有限时间内使系统稳定; 最后, 通过仿真测试验证了理论的可行性. 在接下来的研究中, 可以进一步探索间歇机制在控制理论中的应用, 优化自触发机制.

参 考 文 献

- [1] WANG D, GAO N, LIU D, et al. Recent progress in reinforcement learning and adaptive dynamic programming for advanced control applications [J]. IEEE/CAA Journal of Automatica Sinica, 2023, 11(1): 18-36.
- [2] WANG K, MU C, NI Z, et al. Safe reinforcement learning and adaptive optimal control with applications to obstacle avoidance problem [J]. IEEE Transactions on Automation Science and Engineering, 2023, 21(3): 4599-4612.
- [3] BOURDIN L, TRÉLAT E. Linear-quadratic optimal sampled-data control problems: Convergence result and Riccati theory [J]. Automatica, 2017, 79: 273-281.
- [4] 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.
- [5] XIANG Z, LI P, ZOU W. Event-triggered optimal control for a class of continuous-time switched nonlinear systems [J]. IEEE Transactions on Automation Science and Engineering, 2024, 22: 1620-1630.
- [6] FENG T, ZHANG H, LUO Y, et al. Stability analysis of heuristic dynamic programming algorithm for nonlinear systems [J]. Neurocomputing, 2015, 149: 1461-1468.
- [7] YANG D, LI T, ZHANG H, et al. Event-trigger-based robust control for nonlinear constrained-input systems using reinforcement learning method [J]. Neurocomputing, 2019, 340: 158-170.
- [8] MING Z, ZHANG H, YAN Y, et al. Self-triggered adaptive dynamic programming for model-free nonlinear systems via generalized fuzzy hyperbolic model [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 53(5): 2792-2801.
- [9] 宋广发, 梅俊. 基于 PMV 的多区域暖通空调事件触发间歇控制研究 [J]. 三峡大学学报(自然科学版), 2024, 46(3): 99-104.
- [10] LIU C, LIU L, CAO J, et al. Intermittent event-

- triggered optimal leader-following consensus for nonlinear multi-agent systems via actor-critic algorithm[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 34(8): 3992-4006.
- [11] WANG W, GU H, MEI J, et al. Output information-based intermittent optimal control for continuous-time nonlinear systems with unmatched uncertainties via adaptive dynamic programming[J]. *ISA Transactions*, 2024, 147: 163-175.
- [12] LIU C, LIU L, WU Z. Intermittent event-triggered optimal control for second-order delayed multiagent systems with input constraints[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2024, 54(5): 2698-2710.
- [13] ZHANG C, ZHANG X, XIAO F, et al. Intermittent dynamic event-triggered optimal control for networked control systems with input saturation[J]. *International Journal of Robust and Nonlinear Control*, 2025, 35(6): 1935-1949.
- [14] MEI J, LU Z, HU J, et al. Energy-efficient optimal guaranteed cost intermittent-switch control of a direct expansion air conditioning system[J]. *IEEE/CAA Journal of Automatica Sinica*, 2021, 8(11): 1852-1866.
- [15] HUO Y, WANG D, QIAO J, et al. Off-policy model-free learning for multi-player non-zero-sum games with constrained inputs[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2022, 70(2): 910-920.
- [16] LIU D, WANG D, WANG F Y, et al. Neural-network-based online HJB solution for optimal robust guaranteed cost control of continuous-time uncertain nonlinear systems[J]. *IEEE Transactions on Cybernetics*, 2014, 44(12): 2834-2847.
- [17] LIU S, NIU B, ZONG G, et al. Data-driven-based event-triggered optimal control of unknown nonlinear systems with input constraints[J]. *Nonlinear Dynamics*, 2022, 109(2): 891-909.
- [18] RUAN Z, HU J, MEI J. Robust optimal triple event-triggered intermittent control for uncertain input-constrained nonlinear systems[J]. *Communications in Nonlinear Science and Numerical Simulation*, 2024, 129: 107718.
- [19] LIU M, JIANG H, HU C. Finite-time synchronization of delayed dynamical networks via aperiodically intermittent control[J]. *Journal of the Franklin Institute*, 2017, 354(13): 5374-5397.

(责编 曹东,校对 雷建云)