

基于 ADP 的状态约束互联非线性系统的分散镇定^{*}

赵 博[†] 杜文千 袁郭玲 孔 杰

(北京师范大学系统科学学院, 100875, 北京)

摘要 针对一类含有常数型状态约束的互联非线性系统, 提出一种基于自适应动态规划(adaptive dynamic programming, ADP)的分散镇定方法. 引入边界函数对原系统进行坐标变换, 将状态约束系统转化为无约束系统. 对转化后的系统构造独立子系统和改进的代价函数, 将鲁棒分散镇定问题转化为最优调节问题. 构建局部评判神经网络并采用策略迭代算法求解哈密顿-雅可比-贝尔曼(Hamilton-Jacobi-Bellman, HJB)方程, 进而得到近似最优镇定律. 通过李雅普诺夫稳定性理论证明了本文所提方法可使闭环互联系统和局部评判神经网络估计误差动态最终一致有界. 数值仿真结果验证了所提出分散镇定方法的有效性.

关键词 自适应动态规划; 强化学习; 状态约束; 互联非线性系统; 分散镇定; 最优控制

中图分类号 TP13

DOI: 10.12202/j.0476-0301.2023069

0 引言

系统科学是一门研究系统行为、结构、环境和功能之间普适关系, 进而分析与设计系统演化与调控规律的交叉性学科. 控制论作为系统科学的重要分支, 旨在研究如何设计和优化控制策略, 以使系统在不确定性和变化的环境中能够达到满意的稳定性、准确性、快速性等预设的性能要求^[1]. 在现实世界中, 许多系统是由多个动态互联的子系统组成的, 此类系统广泛存在于工程、电力、交通、经济、社会等领域^[2]. 由于其规模庞大, 且不同子系统的状态之间存在关联, 难以对整个互联系统设计集中控制器. 因此, 针对各个子系统分别设计控制律, 构成一组控制器控制整个大规模系统^[3-4]的分散控制方式得到了广泛应用.

随着系统规模的不断扩大, 人们对于这类系统的控制精度、响应速度、优化性能等指标均有更高的要求. PID 控制等经典控制方法难以满足这类系统稳定性和性能指标最优性的要求. 自适应动态规划(adaptive dynamic programming, ADP)作为一种基于动态规划、强化学习、神经网络等方法的智能控制理论, 能够处理系统的非线性特性和不确定性, 因此对解决非线性系统的最优控制问题非常适用^[5]. 目前, ADP 理论与方法已广泛应用于系统的状态调节、输出调节、轨迹跟踪、容错控制等问题^[5-9]. 在实际应用方面, ADP 已成功用于解决电网控制^[10]、电机调速^[11]、

地铁调度^[12]等问题. 此外, ADP 在解决含约束系统的控制问题上也做了一些有益尝试.

由于系统固有组件自身特性的限制和系统运行安全的要求, 系统的运行状态以及执行器输出通常会受到约束^[13-14], 这增大了控制系统设计的难度. 模型预测控制是一种解决含状态约束系统控制问题的有效方法^[15], 但这种方法具有很高的计算复杂度. 边界李雅普诺夫函数法也可用于处理系统的状态约束^[16], 但这种方法通常对非线性系统结构有特定要求, 如必须是严格反馈形式的非线性系统. 在最优控制理论中, 松弛函数法可处理控制过程的约束条件, 因此也能够用于解决状态约束系统的控制问题. Fan 等^[17]引入松弛函数, 将含状态约束的非线性系统控制问题转化为无约束非线性系统控制问题, 并在此基础上实现原系统的容错控制; Xie 等^[18]引入边界李雅普诺夫函数, 研究了具有参数不确定性、外部干扰和执行器故障的状态约束线性系统的容错控制问题; Yang 等^[19]构造了一种用于处理全状态约束的边界函数, 实现了在含有输入约束和全状态约束的情况下对非线性系统的鲁棒控制; Song 等^[20]运用边界函数处理状态约束, 并引入事件触发机制, 实现了同时包含输入与状态约束的非线性系统的 H_∞ 控制. 然而, 以上方法考虑的均为单一系统, 未考虑互联系统中各子系统状态之间存在的耦合性.

^{*} 国家自然科学基金资助项目(61973330), 北京市自然科学基金资助项目(4212038)

[†] 通信作者: 赵博(1987—), 男, 博士, 副教授. 研究方向: 自适应动态规划与强化学习. E-mail: zhaobo@bnu.edu.cn

收稿日期: 2023-04-29

本文引入边界函数,通过坐标变换将含有常数型状态约束的互联非线性系统镇定问题转化为无约束互联非线性系统的镇定问题,对变换后的系统构建不含耦合项的独立子系统.针对各个独立子系统分别设计含辅助控制律的值函数和哈密顿-雅可比-贝尔曼(Hamilton-Jacobi-Bellman, HJB)方程,并通过局部评判神经网络求解得到近似最优镇定律.本文采用的 2 个仿真实例将该方法与未处理状态约束的 ADP 镇定方法进行了对比,验证了其有效性.

1 问题描述与系统变换

含 N 个子系统的互联非线性系统为

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

式中: $i = 1, 2, \dots, N$; $\mathbf{x}_i(t) \in \mathbb{R}^{n_i}$; $\mathbf{u}_i(t) \in \mathbb{R}^{m_i}$ 分别是第 i 个子系统的状态和控制输入,第 i 个子系统的第 j 维状态 x_{ij} 满足 $c_{ij} < x_{ij} < d_{ij}$ ($c_{ij} < 0$, $d_{ij} > 0$, $j = 1, 2, \dots, n_i$); c_{ij} 、 d_{ij} 分别为 x_{ij} 的下界和上界约束. $\mathbf{f}_i(\mathbf{x}_i) \in \mathbb{R}^{n_i}$ 、 $\mathbf{g}_i(\mathbf{x}_i) \in \mathbb{R}^{n_i \times m_i}$ 均为光滑的非线性函数,分别表示系统动态和控制输入矩阵, $\mathbf{z}_i(\mathbf{x}) \in \mathbb{R}^{n_i}$ 是第 i 个子系统与其他子系统之间的耦合项.

受文献 [19–20] 启发,本文采用边界函数(2),将原互联系统的约束状态 x_{ij} 转化为不含约束的状态 p_{ij} , 如

$$p_{ij} = F_b(x_{ij}; c_{ij}, d_{ij}) = \ln\left(\frac{d_{ij} c_{ij} - x_{ij}}{c_{ij} d_{ij} - x_{ij}}\right), \quad (2)$$

其中 $\forall x_{ij} \in (c_{ij}, d_{ij})$, 则式(2)的反函数为

$$x_{ij} = F_b^{-1}(p_{ij}; c_{ij}, d_{ij}) = \frac{c_{ij}d_{ij}(e^{p_{ij}} - 1)}{c_{ij}e^{p_{ij}} - d_{ij}}, \quad (3)$$

其中 $\forall p_{ij} \in \mathbb{R}$. 将 $F_b^{-1}(p_{ij}; c_{ij}, d_{ij})$ 简记为 $F_b^{-1}(p_{ij})$, 将 $\mathbf{f}_i(\mathbf{x}_i)$ 、 $\mathbf{g}_i(\mathbf{x}_i)$ 、 $\mathbf{z}_i(\mathbf{x})$ 的第 j 行元素分别记为 $f_{ij}(\mathbf{x}_i)$ 、 $g_{ij}(\mathbf{x}_i)$ 、 $z_{ij}(\mathbf{x})$. 通过式(3)将原系统(1)化为

$$\begin{bmatrix} \dot{F}_b^{-1}(p_{i1}) \\ \vdots \\ \dot{F}_b^{-1}(p_{in_i}) \end{bmatrix} = \mathbf{F}_{si}(\mathbf{p}_i) + \mathbf{G}_{si}(\mathbf{p}_i)\mathbf{u}_i + \mathbf{Z}_{si}(\mathbf{p}), \quad (4)$$

式中, $\mathbf{p}$ 为式(1)中 $\mathbf{x}(t)$ 经变换后所得的状态.

$$\mathbf{F}_{si}(\mathbf{p}_i) = \begin{bmatrix} f_{i1}([F_b^{-1}(p_{i1}), \dots, F_b^{-1}(p_{in_i})]^T) \\ \vdots \\ f_{in_i}([F_b^{-1}(p_{i1}), \dots, F_b^{-1}(p_{in_i})]^T) \end{bmatrix},$$

$$\mathbf{G}_{si}(\mathbf{p}_i) = \begin{bmatrix} g_{i1}([F_b^{-1}(p_{i1}), \dots, F_b^{-1}(p_{in_i})]^T) \\ \vdots \\ g_{in_i}([F_b^{-1}(p_{i1}), \dots, F_b^{-1}(p_{in_i})]^T) \end{bmatrix},$$

$$\mathbf{Z}_{si}(\mathbf{p}) = \begin{bmatrix} z_{i1}([F_b^{-1}(p_{i1}), \dots, F_b^{-1}(p_{Nn_N})]^T) \\ \vdots \\ z_{in_i}([F_b^{-1}(p_{i1}), \dots, F_b^{-1}(p_{Nn_N})]^T) \end{bmatrix}.$$

将式(3)对 p_{ij} 求偏导可得

$$\frac{dF_b^{-1}(p_{ij}; c_{ij}, d_{ij})}{dp_{ij}} = \frac{d_{ij}c_{ij}^2 - c_{ij}d_{ij}^2}{c_{ij}^2e^{p_{ij}} + d_{ij}^2e^{-p_{ij}} - 2c_{ij}d_{ij}}. \quad (5)$$

记 $F_{bp_{ij}} = \frac{c_{ij}^2e^{p_{ij}} + d_{ij}^2e^{-p_{ij}} - 2c_{ij}d_{ij}}{d_{ij}c_{ij}^2 - c_{ij}d_{ij}^2}$. 因此, 经坐标变换后的系统状态可表示为

$$\begin{aligned} \dot{p}_{ij} &= \frac{\dot{x}_{ij}}{\frac{dF_b^{-1}(p_{ij}; c_{ij}, d_{ij})}{dp_{ij}}} = F_{bp_{ij}}\dot{x}_{ij} = F_{bp_{ij}} \cdot F_b^{-1}(p_{ij}) = \\ & f_{ij}(\mathbf{p}_i) + g_{ij}(\mathbf{p}_i)\mathbf{u}_i(t) + z_{ij}(\mathbf{p}). \end{aligned} \quad (6)$$

通过式(6)可将式(4)进一步化为

$$\dot{\mathbf{p}}_i(t) = \mathbf{F}_i(\mathbf{p}_i(t)) + \mathbf{G}_i(\mathbf{p}_i(t))\mathbf{u}_i(t) + \mathbf{Z}_i(\mathbf{p}), \quad (7)$$

式中:

$$\mathbf{p}_i = [p_{i1}, p_{i2}, \dots, p_{in_i}]^T,$$

$$\mathbf{F}_i(\mathbf{p}_i) = \text{diag}\{F_{bp_{i1}}, \dots, F_{bp_{in_i}}\} \cdot \mathbf{F}_{si}(\mathbf{p}_i),$$

$$\mathbf{G}_i(\mathbf{p}_i) = \text{diag}\{F_{bp_{i1}}, \dots, F_{bp_{in_i}}\} \cdot \mathbf{G}_{si}(\mathbf{p}_i),$$

$$\mathbf{Z}_i(\mathbf{p}) = \text{diag}\{F_{bp_{i1}}, \dots, F_{bp_{in_i}}\} \cdot \mathbf{Z}_{si}(\mathbf{p}).$$

为便于后续进行分散镇定律设计和稳定性分析,对变换后的互联系统(7),定义如下假设.

假设 1 $\mathbf{F}_i(\mathbf{p}_i)$ 和 $\mathbf{G}_i(\mathbf{p}_i)$ 是 Lipschitz 连续的, 且 $\mathbf{p}_i = 0$ 为系统(7)的平衡点.

假设 2 耦合项 $\mathbf{Z}_i(\mathbf{p})$ 满足非匹配条件

$$\mathbf{Z}_i(\mathbf{p}) = \mathbf{K}_i(\mathbf{p}_i)\mathbf{D}_i(\mathbf{p}) \quad (\mathbf{K}_i(\mathbf{p}_i) \neq \mathbf{G}_i(\mathbf{p}_i)), \quad (8)$$

式中: $\mathbf{K}_i(\mathbf{p}_i) \in \mathbb{R}^{n_i \times w_i}$ 是已知的光滑函数; $\mathbf{D}_i(\mathbf{p}) \in \mathbb{R}^{w_i}$ 为不确定项, 其上界满足

$$\|\mathbf{D}_i(\mathbf{p})\| \leq a_i \sum_{m=1}^N \alpha_m(\|\mathbf{p}_m\|), m = 1, 2, \dots, N, \quad (9)$$

式中: $a_i \geq 0$ ($i = 1, 2, \dots, N$) 为非负常数; $\alpha_m(\cdot)$ 为半正定函数, 且 $\alpha_m(0) = 0$, $m = 1, 2, \dots, N$.

假设 3 存在函数 $\beta_l(\cdot)$, $l = 1, 2, \dots, N$. 使

$$\|\mathbf{G}_i^+(\mathbf{p}_i)\mathbf{Z}_i(\mathbf{p})\| \leq \sum_{l=1}^N e_{il}\beta_l(\|\mathbf{p}_l\|), \quad (10)$$

式中: $\mathbf{G}_i^+(\mathbf{p}_i)$ 为 $\mathbf{G}_i(\mathbf{p}_i)$ 的 Moore-Penrose 广义逆矩阵; e_{il} ($l = 1, 2, \dots, N$) 为非负常数.

针对各个独立子系统构建辅助镇定律和对应的代价函数, 并通过局部评判神经网络执行策略迭代算法, 从而找到一组反馈镇定律 $u_1(p_1), u_2(p_2), \dots, u_N(p_N)$, 使原互联系统(1)渐近稳定, 且在控制过程中系统状态始终保持在预设的约束区间内.

2 分散镇定系统设计

将耦合项 $Z_i(p)$ 沿 $G_i(p_i)$ 方向投影分解为匹配与非匹配的两项, 如

$$Z_i(p) = G_i(p_i)G_i^+(p_i)K_i(p_i)D_i(p) + (I_{n_i} - G_i(p_i)G_i^+(p_i))K_i(p_i)D_i(p), \quad (11)$$

式中: $G_i^+(p_i)K_i(p_i)D_i(p)$ 为 $Z_i(p)$ 的匹配分量; $I_{n_i} \in \mathbb{R}^{n_i \times n_i}$ 为单位矩阵, 系统(7)对应的独立子系统为

$$\dot{p}_i = F_i(p_i) + G_i(p_i)u_i + C_i(p_i)v_i, \quad (12)$$

式中 $C_i(p_i) = (I_{n_i} - G_i(p_i)G_i^+(p_i))K_i(p_i)$.

辅助镇定律 v_i 用于处理系统(7)中 $Z_i(p)$ 的不匹配分量. 独立子系统(12)对应的值函数设为

$$V_i(p_i(t)) = \int_t^\infty r_i(p_i(\tau), u_i(\tau), v_i(\tau)) + \eta_i T_i^2(p_i(\tau)) d\tau, \quad (13)$$

式中: $r_i(p_i(t), u_i(t), v_i(t)) = p_i^T Q_i p_i + u_i^T R_i u_i + \varepsilon_i v_i^T v_i$; $T_i(p_i) \geq \max\{\alpha_i(\|p_i\|), \beta_i(\|p_i\|)\}$, $i = 1, 2, \dots, N$; ε_i 和 η_i 为可调参数, $Q_i \in \mathbb{R}^{n_i \times n_i}$ 和 $R_i \in \mathbb{R}^{m_i \times m_i}$ 为正定对称矩阵. 代价函数(13)中 $T_i(p_i)$ 考虑耦合项 $Z_i(p)$ 的范数上界, 且 $Z_i(p)$ 不出现在独立子系统(12)中. 将系统(7)的分散镇定问题转化为系统(12)的最优控制问题, 后续将通过最小化(13)补偿耦合项的影响.

引入独立子系统后, 需要寻找一对增广镇定律 (u_i^*, v_i^*) , 使得独立子系统(12)稳定, 并最小化值函数(13), 即

$$(u_i^*, v_i^*) = \arg \min_{\substack{u_i \in \mathcal{U}(\Omega_i) \\ v_i \in \mathcal{V}(\Omega_i)}} V_i(p_i(t)), \quad (14)$$

式中 $\mathcal{U}(\Omega_i)$ 和 $\mathcal{V}(\Omega_i)$ 分别为控制律 u_i 和 v_i 的可容许镇定控制集合. 当代价函数 $V_i(p_i(t))$ 达到最小, 即

$$V_i^*(p_i(t)) = \int_t^\infty [r_i(p_i^*(\tau), u_i^*(\tau), v_i^*(\tau)) + \eta_i T_i^2(p_i^*(\tau))] d\tau, \quad (15)$$

其满足

$$\nabla V_i^{*T}(p_i)(F_i(p_i) + G_i(p_i)u_i + C_i(p_i)v_i) + r_i(p_i, u_i, v_i) + \eta_i T_i^2(p_i) = 0. \quad (16)$$

独立子系统(12)的哈密顿函数为

$$H(\nabla V_i^T, p_i, u_i, v_i) = \nabla V_i^T(p_i)(F_i(p_i) + G_i(p_i)u_i + C_i(p_i)v_i) + r_i(p_i, u_i, v_i) + \eta_i T_i^2(p_i). \quad (17)$$

根据贝尔曼最优性原理, 可知

$$\min_{\substack{u_i \in \mathcal{U}(\Omega_i) \\ v_i \in \mathcal{V}(\Omega_i)}} H(\nabla V_i^T, p_i, u_i, v_i) = 0. \quad (18)$$

由式(13)和(18)可得, 独立子系统(12)的最优反馈控制律 (u_i^*, v_i^*) 表达式为

$$u_i^* = -\frac{1}{2}R_i^{-1}G_i^T(p_i)\nabla V_i^*(p_i), \quad (19)$$

$$v_i^* = -\frac{1}{2\varepsilon_i}C_i^T(p_i)\nabla V_i^*(p_i). \quad (20)$$

3 神经网络求解

3.1 策略迭代算法 由于 HJB 方程中存在偏微分项 $\nabla V_i^*(p_i)$, 使其难以直接求解, 本节将采用策略迭代算法对其进行近似求解. 研究证明, 该算法在迭代次数趋于无穷时, 代价函数和镇定律均收敛到最优^[21].

记

$$\mathcal{G}_i(p_i) = [G_i(p_i), C_i(p_i)], \quad (21)$$

$$\mu_i(p_i) = [u_i^T(p_i), v_i^T(p_i)]^T. \quad (22)$$

则独立子系统(12)可表示为

$$\dot{p}_i = F_i(p_i) + \mathcal{G}_i(p_i)\mu_i(p_i). \quad (23)$$

式(13)中的 $r_i(p_i(t), u_i(t), v_i(t))$ 可表示为

$$r_i(p_i(t), \mu_i(t)) = p_i^T Q_i p_i + \mu_i^T \mathcal{R}_i \mu_i, \quad (24)$$

式中 $\mathcal{R}_i = \text{diag}\{R_i, \varepsilon_i I_{n_i}\}$.

下面, 采用策略迭代算法逼近各个独立子系统的最优镇定律. 针对第 i 个独立子系统选择一个初始可容许镇定律 $\mu_i^{(0)} = [u_i^{(0)T}(p_i), v_i^{(0)T}(p_i)]^T$, 其中 $u_i^{(0)} \in \mathcal{U}(\Omega_i)$, $v_i^{(0)} \in \mathcal{V}(\Omega_i)$. 然后, 通过下面两式依次重复进行策略评价和策略更新, 即

$$\begin{aligned} & (\nabla V_i^{(q+1)}(p_i))^T (F_i(p_i) + \mathcal{G}_i(p_i)\mu_i^{(q)}(p_i)) + \\ & r_i(p_i, \mu_i^{(q)}(p_i)) + \eta_i T_i^2(p_i) = 0, \end{aligned} \quad (25)$$

$$\mu_i^{(q+1)} = -\frac{1}{2}\mathcal{R}_i^{-1}(p_i)\mathcal{G}_i^T(p_i)\nabla V_i^{(q+1)}(p_i). \quad (26)$$

直到迭代至第 q 步时代价函数收敛到设定的误差范围 ϑ , 迭代过程如图1所示.

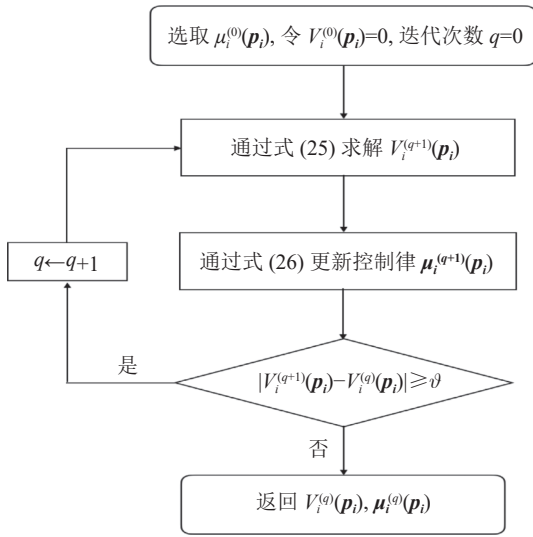


图 1 策略迭代算法

3.2 局部评判神经网络实现 神经网络具有通用近似能力, 因而可以用来估计子系统的值函数, 进而在线求解 HJB 方程(18). 本文系统模型已知, 因此可以舍弃执行神经网络以简化神经网络结构, 通过局部评判神经网络即可估计子系统的值函数(15). 最优代价函数

$$V_i^*(p_i) = \mathbf{W}_{ci}^T \boldsymbol{\sigma}_{ci}(p_i) + \zeta_{ci}(p_i), \quad (27)$$

式中: $\mathbf{W}_{ci} \in \mathbb{R}^{l_i}$ 是未知的理想权重; $\boldsymbol{\sigma}_{ci}(p_i) \in \mathbb{R}^{l_i}$ 是可导并连续的激活函数, 且 $\boldsymbol{\sigma}_{ci}(0) = 0$; l_i 是局部评判神经网络中神经元的数量; $\zeta_{ci}(p_i)$ 是重构误差. 将式(27)对 p_i 求梯度可得

$$\nabla V_i^*(p_i) = \nabla \boldsymbol{\sigma}_{ci}^T(p_i) \mathbf{W}_{ci} + \nabla \zeta_{ci}(p_i), \quad (28)$$

式中: $\nabla(\cdot)$ 表示求梯度运算. 由式(19)~(22)、(28), 可将增广控制律 $\mu_i^*(p_i)$ 重构为

$$\mu_i^*(p_i) = -\frac{1}{2} \mathcal{R}_i^{-1} \mathcal{G}_i^T(p_i) \nabla \boldsymbol{\sigma}_{ci}^T(p_i) \mathbf{W}_{ci} + \zeta_{\mu_i}(p_i), \quad (29)$$

式中 $\zeta_{\mu_i}(p_i) = -\frac{1}{2} \mathcal{R}_i^{-1} \mathcal{G}_i^T(p_i) \nabla \zeta_{ci}(p_i)$. HJB 方程(18)可化为

$$\eta_i T_i^2(p_i) + p_i^T Q_i p_i + \mu_i^{*T}(p_i) \mathcal{R}_i \mu_i^*(p_i) + (\nabla V_i^*(p_i))^T (F_i(p_i) + \mathcal{G}_i(p_i) \mu_i^*(p_i)) = 0. \quad (30)$$

由于局部评判神经网络的理想权重 $\mathbf{W}_{ci}$ 未知, 因此将其估计值记为 $\widehat{\mathbf{W}}_{ci}$, 因此最优代价函数的估计值记为 $\widehat{V}_i(p_i)$. 则

$$\widehat{V}_i(p_i) = \widehat{\mathbf{W}}_{ci}^T \boldsymbol{\sigma}_{ci}(p_i), \quad (31)$$

$$\widehat{\mu}_i(p_i) = -\frac{1}{2} \mathcal{R}_i^{-1} \mathcal{G}_i^T(p_i) \nabla \boldsymbol{\sigma}_{ci}^T(p_i) \widehat{\mathbf{W}}_{ci}. \quad (32)$$

由式(17)和(31)可知, 近似的哈密顿函数为

$$\widehat{H}(\nabla \widehat{V}_i, p_i, \widehat{\mu}_i) = \eta_i T_i^2(p_i) + r_i(p_i, \widehat{\mu}_i) + \widehat{\mathbf{W}}_{ci}^T \boldsymbol{\phi}_i, \quad (33)$$

式中 $\boldsymbol{\phi}_i = \nabla \boldsymbol{\sigma}_{ci}(p_i) (F_i(p_i) + \mathcal{G}_i(p_i) \widehat{\mu}_i(p_i))$.

定义哈密顿函数的估计误差为

$$e_{ci} = \widehat{H}_i(\nabla \widehat{V}_i, p_i, \widehat{\mu}_i) - H_i(\nabla V_i^*, p_i, \mu_i^*) = \eta_i T_i^2(p_i) + r_i(p_i, \widehat{\mu}_i) + \widehat{\mathbf{W}}_{ci}^T \boldsymbol{\phi}_i. \quad (34)$$

对局部评判神经网络权重 $\widehat{\mathbf{W}}_{ci}$ 训练的目标是使其收敛到理想权重 $\mathbf{W}_{ci}$. 因此, 定义目标函数 $E_{ci} = \frac{1}{2} \mathbf{e}_{ci}^T \mathbf{e}_{ci}$, 并通过梯度下降法(35)训练局部评判神经网络权重以最小化 E_{ci} , 则局部评判神经网络权重的更新律为

$$\dot{\widehat{\mathbf{W}}}_{ci} = -\frac{\gamma_{ci}}{(1 + \boldsymbol{\phi}_i^T \boldsymbol{\phi}_i)^2} \frac{\partial E_{ci}}{\partial \widehat{\mathbf{W}}_{ci}} = -\gamma_{ci} \frac{\boldsymbol{\phi}_i \mathbf{e}_{ci}}{(1 + \boldsymbol{\phi}_i^T \boldsymbol{\phi}_i)^2}, \quad (35)$$

式中 $\gamma_{ci} > 0$ 为学习率, $(1 + \boldsymbol{\phi}_i^T \boldsymbol{\phi}_i)^2$ 用于归一化.

定义局部评判神经网络权重误差为 $\widetilde{\mathbf{W}}_{ci} = \mathbf{W}_{ci} - \widehat{\mathbf{W}}_{ci}$. 将其对时间求导, 结合式(35)可得

$$\dot{\widetilde{\mathbf{W}}}_{ci} = -\gamma_{ci} \boldsymbol{\psi}_i \boldsymbol{\psi}_i^T \widetilde{\mathbf{W}}_{ci} + \frac{\gamma_{ci} \boldsymbol{\psi}_i \zeta_{Hi}(p_i)}{1 + \boldsymbol{\phi}_i^T \boldsymbol{\phi}_i}, \quad (36)$$

式中:

$$\boldsymbol{\psi}_i = \frac{\boldsymbol{\phi}_i}{1 + \boldsymbol{\phi}_i^T \boldsymbol{\phi}_i},$$

$$\zeta_{Hi}(p_i) = -\nabla \zeta_{ci}^T(p_i) (F_i(p_i) + \mathcal{G}_i(p_i) \widehat{\mu}_i(p_i)).$$

3.3 稳定性分析

假设 4 $\mathcal{G}_i(p_i)$ 、 $\boldsymbol{\sigma}_{ci}(p_i)$ 、 $\nabla \boldsymbol{\sigma}_{ci}(p_i)$ 、 $\zeta_{ci}(p_i)$ 、 $\nabla \zeta_{ci}(p_i)$ 、 $\zeta_{Hi}(p_i)$ 均为范数有界的, 即满足

$\|\mathcal{G}_i(p_i)\| \leq \lambda_1$, $\|\boldsymbol{\sigma}_{ci}(p_i)\| \leq \lambda_2$, $\|\nabla \boldsymbol{\sigma}_{ci}(p_i)\| \leq \lambda_3$, $\|\zeta_{ci}(p_i)\| \leq \lambda_4$, $\|\nabla \zeta_{ci}(p_i)\| \leq \lambda_5$, $\|\zeta_{Hi}(p_i)\| \leq \lambda_6$, 其中, $\lambda_1, \lambda_2, \dots, \lambda_6$ 均为正常数.

假设 5 $\boldsymbol{\psi}_i$ 满足持续激励条件, 即存在 $\varrho_2 \geq \varrho_1 > 0$, $T > 0$, 使

$$\varrho_1 \mathbf{I}_{l_i} \leq \int_t^{t+T} \boldsymbol{\psi}_i(\tau) \boldsymbol{\psi}_i^T(\tau) d\tau \leq \varrho_2 \mathbf{I}_{l_i}, \quad \forall t \geq 0.$$

定理 1 对于第 i 个独立子系统(12), 若假设 1~5 成立, 初始镇定律满足 $\mathbf{u}_i^{(0)} \in \mathcal{U}(\Omega)$, $\mathbf{v}_i^{(0)} \in \mathcal{B}(\Omega)$, 局部评判神经网络的权重由式(35)更新, 且满足

$$\gamma_{ci} \geq \frac{\lambda_1^2 \lambda_3^2}{\lambda_{\min}(\boldsymbol{\psi}_i \boldsymbol{\psi}_i^T) \lambda_{\min}(\mathcal{R}_i)}, \quad (37)$$

则独立子系统状态 p_i 和权重估计误差 $\widetilde{\mathbf{W}}_{ci}$ 是最终一致有界的.

证明: 选取 Lyapunov 函数

$$\mathcal{L}_i(t) = V_i^*(p_i(t)) + \frac{1}{2} \widetilde{\mathbf{W}}_{ci}^T \widetilde{\mathbf{W}}_{ci},$$

记 $\mathcal{L}_{i1}(t) = V_i^*(p_i(t))$, $\mathcal{L}_{i2}(t) = \frac{1}{2} \tilde{\mathbf{W}}_{ci}^T \tilde{\mathbf{W}}_{ci}$. 将 $\hat{\mu}_i$ 代入式(23), 得到 $\dot{p}_i = F_i(p_i) + \mathcal{G}_i(p_i) \hat{\mu}_i$, 其代价函数导数表示为

$$\dot{V}_i^*(p_i) = \nabla V_i^{*T}(p_i)(F_i(p_i) + \mathcal{G}_i(p_i) \hat{\mu}_i). \quad (38)$$

结合式(29)、(30)和(38), 可得

$$\begin{aligned} \dot{\mathcal{L}}_{i1}(t) &= \nabla V_i^{*T}(p_i)(F_i(p_i) + \mathcal{G}_i(p_i) \hat{\mu}_i) = \\ &= -\eta_i T_i^2(p_i) - p_i^T Q_i p_i + \mu_i^{*T} \mathcal{R}_i \mu_i^* - 2\mu_i^{*T} \mathcal{R}_i \hat{\mu}_i = \\ &= \|\mathcal{R}_i^{-\frac{1}{2}}(\mu_i^* - \hat{\mu}_i)\|^2 - \eta_i T_i^2(p_i) - r_i(p_i, \hat{\mu}_i). \end{aligned} \quad (39)$$

记 $\Gamma_{i1} = \|\mathcal{R}_i^{-\frac{1}{2}}(\mu_i^* - \hat{\mu}_i)\|^2$, 将式(29)、(32)代入式(39), 并结合不等式 $\frac{\|a+b\|^2}{2} \leq \|a\|^2 + \|b\|^2$, $\forall a, b \in \mathbb{R}$, 可得

$$\begin{aligned} \Gamma_{i1} &= \left\| \frac{1}{2} \mathcal{R}_i^{-\frac{1}{2}} \mathcal{G}_i^T(p_i) \nabla \sigma_{ci}^T(p_i) \tilde{\mathbf{W}}_{ci}(p_i) - \mathcal{R}_i^{-\frac{1}{2}} \zeta_{\mu_i}(p_i) \right\|^2 \leq \\ &= \frac{1}{2} \|\mathcal{R}_i^{-\frac{1}{2}} \mathcal{G}_i^T(p_i) \nabla \sigma_{ci}^T(p_i) \tilde{\mathbf{W}}_{ci}(p_i)\|^2 + 2\|\mathcal{R}_i^{-\frac{1}{2}} \zeta_{\mu_i}(p_i)\|^2. \end{aligned} \quad (40)$$

因此, 结合假设4, 式(39)可进一步化为

$$\begin{aligned} \dot{\mathcal{L}}_{i1}(t) &\leq -\eta_i T_i^2(p_i) - r_i(p_i, \hat{\mu}_i) + 2\|\mathcal{R}_i^{-\frac{1}{2}} \zeta_{\mu_i}(p_i)\|^2 + \\ &= \frac{1}{2} \|\mathcal{R}_i^{-\frac{1}{2}} \mathcal{G}_i^T(p_i) \nabla \sigma_{ci}^T(p_i) \tilde{\mathbf{W}}_{ci}(p_i)\|^2 \leq \\ &= -\lambda_{\min}(Q_i) \|p_i\|^2 + \frac{\lambda_1^2 \lambda_5^2}{2\lambda_{\min}(\mathcal{R}_i)} + \frac{\lambda_1^2 \lambda_3^2}{2\lambda_{\min}(\mathcal{R}_i)} \|\tilde{\mathbf{W}}_{ci}(p_i)\|^2. \end{aligned} \quad (41)$$

由式(36)可得

$$\begin{aligned} \dot{\mathcal{L}}_{i2}(t) &= -\gamma_{ci} \tilde{\mathbf{W}}_{ci}^T(p_i) \psi_i \psi_i^T \tilde{\mathbf{W}}_{ci}(p_i) + \\ &= \gamma_{ci} \frac{\tilde{\mathbf{W}}_{ci}^T(p_i) \psi_i}{1 + \varphi_i^T \varphi_i} \zeta_{H_i}(p_i), \end{aligned} \quad (42)$$

式中

$$\begin{aligned} \gamma_{ci} \frac{\tilde{\mathbf{W}}_{ci}^T(p_i) \psi_i}{1 + \varphi_i^T \varphi_i} \zeta_{H_i}(p_i) &\leq |\gamma_{ci} \tilde{\mathbf{W}}_{ci}^T(p_i) \psi_i \zeta_{H_i}(p_i)| \leq \\ &= \frac{\gamma_{ci}}{2} \tilde{\mathbf{W}}_{ci}^T(p_i) \psi_i \psi_i^T \tilde{\mathbf{W}}_{ci}(p_i) + \frac{\gamma_{ci}}{2} \zeta_{H_i}^T \zeta_{H_i}, \end{aligned}$$

由式(42)和假设4可得

$$\dot{\mathcal{L}}_{i2}(t) \leq -\frac{\gamma_{ci}}{2} \lambda_{\min}(\psi_i \psi_i^T) \|\tilde{\mathbf{W}}_{ci}(p_i)\|^2 + \frac{\gamma_{ci}}{2} \lambda_6^2. \quad (43)$$

联立式(41)和(43)可得

$$\begin{aligned} \dot{\mathcal{L}}_i(t) &= \dot{\mathcal{L}}_{i1}(t) + \dot{\mathcal{L}}_{i2}(t) \leq \\ &= -\lambda_{\min}(Q_i) \|p_i\|^2 + \frac{\lambda_1^2 \lambda_5^2}{2\lambda_{\min}(\mathcal{R}_i)} + \frac{\gamma_{ci} \lambda_6^2}{2} - \\ &= \left(\frac{\gamma_{ci}}{2} \lambda_{\min}(\psi_i \psi_i^T) - \frac{\lambda_1^2 \lambda_3^2}{2\lambda_{\min}(\mathcal{R}_i)} \right) \|\tilde{\mathbf{W}}_{ci}(p_i)\|^2. \end{aligned}$$

因此, 根据 Lyapunov 稳定性定理和式(37), 当

$$\|p_i\| > \sqrt{\frac{\lambda_1^2 \lambda_5^2 / \lambda_{\min}(\mathcal{R}_i) + \gamma_{ci} \lambda_6^2}{2\lambda_{\min}(Q_i)}}$$

或

$$\|\tilde{\mathbf{W}}_{ci}(p_i)\| > \sqrt{\frac{\lambda_1^2 \lambda_5^2 + \gamma_{ci} \lambda_6^2 \lambda_{\min}(\mathcal{R}_i)}{\gamma_{ci} \lambda_{\min}(\psi_i \psi_i^T) \lambda_{\min}(\mathcal{R}_i) - \lambda_1^2 \lambda_3^2}} \triangleq \lambda_7$$

其中任意一式成立, 独立子系统状态 p_i 和局部评判神经网络权重估计误差 $\tilde{\mathbf{W}}_{ci}(p_i)$ 均为最终一致有界的. 证毕.

由定理1和式(27)、(31)可得, $\|V_i^*(p_i) - \hat{V}_i(p_i)\| \leq \lambda_2 \lambda_7 + \lambda_4$. 通过选取恰当的学习率 γ_{ci} 可以对 λ_7 进行调整. 因此, $\hat{V}_i(p_i)$ 能够收敛到子系统最优代价函数 $V_i^*(p_i)$ 的邻域内.

4 仿真研究

本章将采用2个仿真算例来验证所提出针对状态约束互联非线性系统的分散镇定方法的有效性.

4.1 数值非线性系统 考虑互联非线性系统, 其动力学模型为

$$\dot{\mathbf{x}}_1 = \begin{bmatrix} -x_{11} + x_{12} \\ -0.5(x_{11} + x_{12}) + 0.5x_{12}(2 + \cos(2x_{11}))^2 \\ 0 \\ 2 + \cos(2x_{11}) \end{bmatrix} \mathbf{u}_1 + \mathbf{z}_1(\mathbf{x}), \quad (44)$$

$$\dot{\mathbf{x}}_2 = \begin{bmatrix} 0.5x_{22} \\ -x_{21} - 0.5x_{22} + 0.5x_{21}^2 x_{22} \\ 0 \\ x_{21} \end{bmatrix} \mathbf{u}_2 + \mathbf{z}_2(\mathbf{x}). \quad (45)$$

2个子系统的耦合项分别设为

$$\mathbf{z}_1(\mathbf{x}) = \begin{bmatrix} 0.25 \\ 0.25 \end{bmatrix} (x_{11} + x_{22}) \sin^2 x_{22} \cos x_{21},$$

$$\mathbf{z}_2(\mathbf{x}) = \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} (x_{12} + x_{22}) \cos(e^{x_{21}}).$$

在此例中, 初始状态分别选择为 $\mathbf{x}_1(0) = [1, 0]^T$, $\mathbf{x}_2(0) = [1, -1]^T$. 状态约束为 $x_{11}, x_{12} \in (-0.2, 2)$, $x_{21} \in (-1.5, 1.5)$, $x_{22} \in (-1.4, 1.5)$.

下面利用坐标变换构造无约束的系统. 经过坐标变换后的初始状态为 $\mathbf{p}_1(0) = [F_b(1; -0.2, 2), F_b(0; -0.2, 2)]^T$, $\mathbf{p}_2(0) = [F_b(1; -1.5, 1.5), F_b(-1; -1.4, 1.5)]^T$. 独立子系统的代价函数中选取 $\mathbf{Q}_1 = \mathbf{Q}_2 = \text{diag}\{6, 6\}$, $\mathcal{R}_1 = \mathcal{R}_2 = \text{diag}\{1, 0.3\}$.

构造2个局部评判神经网络分别估计2个独立子系统镇定律. 激活函数选为 $\sigma_{ci}(\mathbf{p}_i) = [p_{i1}^2, p_{i2}^2, p_{i1} p_{i2}]^T$, $i = 1, 2$. 根据子系统的初始可容许镇定律, 局部评判神经网络初值权重选取为 $\hat{\mathbf{W}}_{c1}^{(0)} = [0, 1, 0]^T$, $\hat{\mathbf{W}}_{c2}^{(0)} = [1, 2, 1]^T$.

学习率选取为 $\gamma_{c1} = \gamma_{c2} = 0.1$.

如图 2 所示, 采用式 (35) 作为更新律训练 2 个独立子系统的局部评判神经网络 150 s 后, 2 个独立子系统的局部评判神经网络权重分别收敛于

$$\hat{W}_{c1} = [0.157\ 4, 1.059\ 5, -0.239\ 0]^T,$$

$$\hat{W}_{c2} = [1.623\ 4, 2.486\ 4, 0.110\ 8]^T.$$

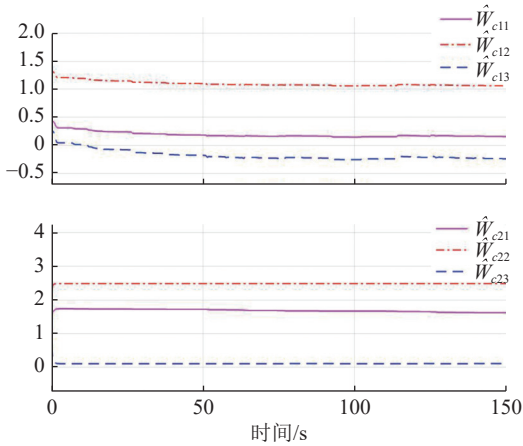


图 2 独立子系统 1、2 局部评判神经网络权重演化过程

图 3 所示为分散镇定方法的作用下, 原互联子系统的 (44)、(45) 的状态演化过程. 由此可见, 原系统的 2 个互联子系统状态能够收敛到平衡点附近的小邻域内, 即最终一致有界, 且满足状态约束.

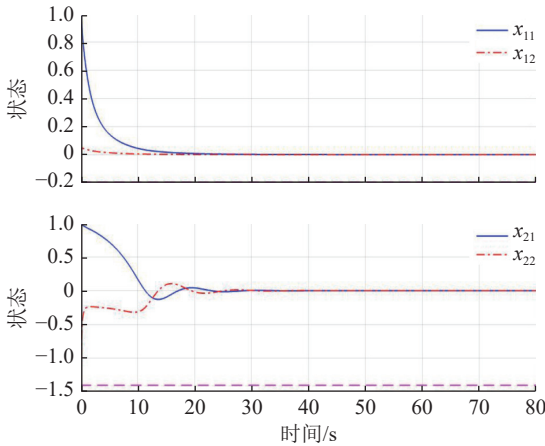


图 3 所提出分散镇定律作用下子系统 1、2 状态

为了验证系统状态约束对系统镇定性能的影响, 在不处理状态约束的情况下, 直接采用了现有基于 ADP 的分散镇定方法对系统 (44)、(45) 进行控制, 得到的系统状态演化过程如图 4 所示. 由仿真结果可见, 虽然该方法能获得更快的收敛速度, 但系统状态超出了约束区间.

4.2 双倒立摆系统 采用双倒立摆系统进行仿真^[22], 以进一步验证本文所提方法的有效性. 由弹簧连接

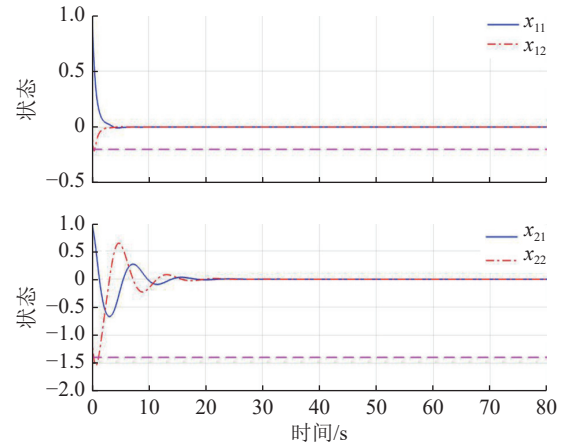


图 4 现有基于 ADP 的分散镇定律作用下子系统 1、2 状态(未处理状态约束)

的双倒立摆系统如图 5 所示, 可看作含 2 个子系统的互联非线性系统, 动力学模型可表示为

$$m_1 l_1^2 \ddot{\theta}_1 - m_1 g l_1 \sin \theta_1 + b_1 \dot{\theta}_1 - F a_1 \cos(\theta_1 - \beta) = \delta_1 u_1,$$

$$m_2 l_2^2 \ddot{\theta}_2 - m_2 g l_2 \sin \theta_2 + b_2 \dot{\theta}_2 - F a_2 \cos(\theta_2 - \beta) = \delta_2 u_2.$$

式中, b_1 和 b_2 分别为 2 个弹簧的阻尼系数,

$$F = k\{1 + A^2(l_k - l_0)^2\}(l_k - l_0),$$

$$|A(l_k - l_0)| < 1,$$

$$\beta = \arctan \frac{a_1 \cos \theta_1 - a_2 \cos \theta_2}{l_0 - a_1 \sin \theta_1 + a_2 \sin \theta_2},$$

$$l_k = \{(l_0 - a_1 \sin \theta_1 + a_2 \sin \theta_2)^2 + (a_1 \cos \theta_1 + a_2 \cos \theta_2)^2\}^{\frac{1}{2}}.$$

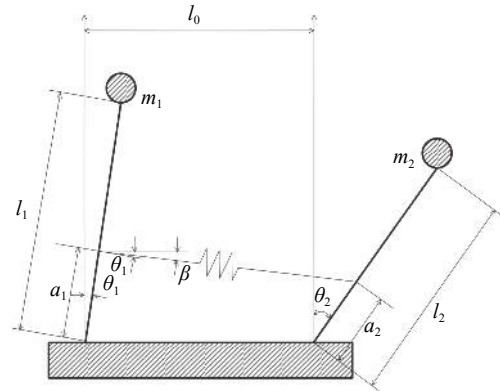


图 5 双倒立摆系统

系统参数选取为 $\delta_1 = \delta_2 = 1$, $m_1 = m_2 = 1\text{ kg}$, $l_1 = l_2 = 0.5\text{ m}$, $l_0 = 1\text{ m}$, $g = 9.8\text{ m}\cdot\text{s}^{-2}$, $b_1 = b_2 = 0.009$, $k = 30$, $A = 0.1$, 弹簧与 2 个倒立摆底端的距离 $a_1 = a_2 = 0.1$. 状态约束为 $\dot{\theta}_1 \in (-1, 1)$, $\dot{\theta}_2 \in (-1.3, 1.3)$.

令 $x_{i1} = \theta_i$, $x_{i2} = \dot{\theta}_i$, 系统状态的初始值为 $x_1(0) = [-2, 0.3]^T$, $x_2(0) = [1, 0]^T$, 其中 $i = 1, 2$. 2 个互联子系统的动力学模型为

$$f_i(x_i) = \begin{bmatrix} x_{i2} \\ 5.88 \sin x_{i1} - 0.036 x_{i2} \end{bmatrix},$$

$$\mathbf{g}_i(\mathbf{x}_i) = \begin{bmatrix} 0 \\ \delta_i \end{bmatrix}, \mathbf{k}_i(\mathbf{x}) = \begin{bmatrix} 0 \\ 4Fa_i \cos(x_{i1} - \beta) \end{bmatrix}.$$

对原系统进行坐标变换并构造独立子系统, 选取代价函数中的参数为 $\mathbf{Q}_1 = \text{diag}\{0.1, 0.1\}$, $\mathbf{Q}_2 = \text{diag}\{0.2, 0.2\}$, $\mathbf{R}_1 = \text{diag}\{0.01, 0.0025\}$, $\mathbf{R}_2 = \text{diag}\{0.02, 0.0025\}$.

构造 2 个局部评判神经网络分别估计两个独立子系统镇定律. 激活函数选为 $\sigma_{ci}(\mathbf{p}_i) = [p_{i1}^2, p_{i2}^2, p_{i1}p_{i2}]^T$, $i = 1, 2$. 根据子系统的初始可容许镇定律, 局部评判神经网络权重初值选取为 $\hat{\mathbf{W}}_{c1}^{(0)} = [1, 1.5, 1]^T$ 和 $\hat{\mathbf{W}}_{c2}^{(0)} = [0, 1, 2]^T$, 学习率选取为 $\gamma_{c1} = \gamma_{c2} = 0.1$.

图 6 所示为 2 个独立子系统的局部评判神经网络权重演化过程, 可见其分别收敛到

$$\begin{aligned} \hat{\mathbf{W}}_{c1} &= [1.0179, 1.3602, 0.9926]^T, \\ \hat{\mathbf{W}}_{c2} &= [0.0531, 0.9891, 1.2753]^T. \end{aligned}$$

图 7 所示为双倒立摆系统在本文所提的分散镇定方法作用下的状态演化过程. 由于倒立摆位置和角速度的初值不处于平衡状态, 因此在施加控制后的 0~1 s 内系统状态 x_{12} 和 x_{22} 有短暂的调节过程, 但系统状态始终未超出状态约束, 并达到最终一致有界.

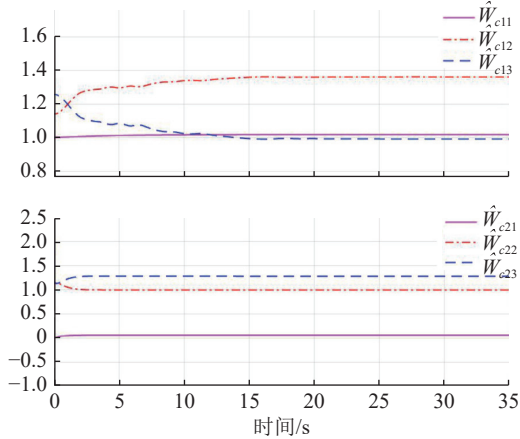


图 6 独立子系统 1、2 局部评判神经网络权重演化过程

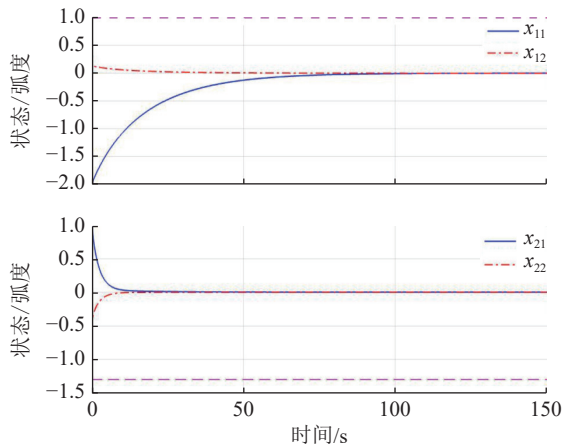


图 7 所提出分散镇定律作用下子系统 1、2 状态

作为对比, 本例也采用了未考虑状态约束的基于 ADP 的分散镇定方法进行仿真, 结果如图 8 所示. 在该方法作用下, 系统所用的调节时间与图 7 相比较短, x_{12} 和 x_{22} 在超出状态约束之后再收敛. 因此, 本文所提方法通过更长的调节时间换取了系统性能的保证.

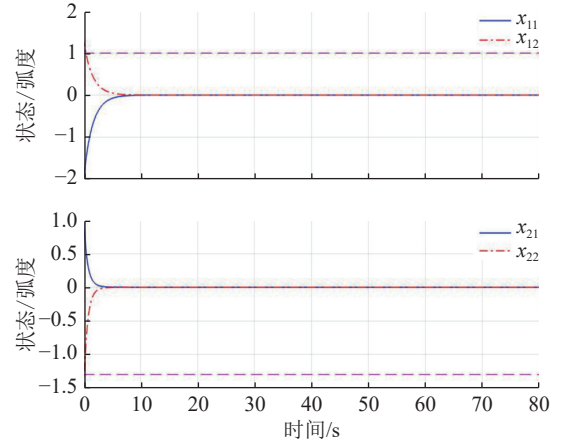


图 8 现有基于 ADP 的分散镇定律作用下子系统 1、2 状态 (未处理状态约束)

5 结论

本文针对一类含有常值型状态约束的互联非线性系统, 提出了一种基于 ADP 的分散镇定方法. 该方法首先引入边界函数进行坐标变换, 将含状态约束系统的镇定问题转化为无约束系统的镇定问题, 进而构造独立子系统, 并引入辅助控制构成增广镇定律, 以补偿耦合项的影响. 通过局部评判神经网络对各个独立子系统的增广镇定律进行近似求解, 并利用 Lyapunov 稳定性定理证明了系统状态与局部评判神经网络的权重估计误差均为最终一致有界的. 仿真结果表明, 所提出的分散镇定方法是有效的.

6 参考文献

- [1] 夏宏兵. 非线性系统的自学习优化安全调控方法研究 [D]. 北京: 北京师范大学, 2023
- [2] BIAN T, JIANG Y, JIANG Z P. Decentralized adaptive optimal control of large-scale systems with application to power systems[J]. *IEEE Transactions on Industrial Electronics*, 2015, 62(4): 2439
- [3] EL-ATTAR R, VIDYASAGAR M. Subsystem simplification in large-scale systems analysis[J]. *IEEE Transactions on Automatic Control*, 1979, 24(2): 321
- [4] YANG X, HE H B. Adaptive dynamic programming for decentralized stabilization of uncertain nonlinear large-scale systems with mismatched interconnections[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2015, 45(1): 1-12

- 2020, 50(8): 2870
- [5] WANG D, HE H B, LIU D R. Adaptive critic nonlinear robust control: a survey[J]. *IEEE Transactions on Cybernetics*, 2017, 47(10): 3429
- [6] ZHAO B, SHI G, LIU D R. Event-triggered local control for nonlinear interconnected systems through particle swarm optimization-based adaptive dynamic programming[J/OL]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, <http://dx.doi.org/10.1109/TSMC.2023.3298065>
- [7] ZHAO B, ZHANG Y W, LIU D R. Adaptive dynamic programming-based cooperative motion/force control for modular reconfigurable manipulators: a joint task assignment approach[J/OL]. *IEEE Transactions on Neural Networks and Learning Systems*, <http://dx.doi.org/10.1109/TNNLS.2022.3171828>
- [8] LEWIS F L, VRABIE D. Reinforcement learning and adaptive dynamic programming for feedback control[J]. *IEEE Circuits and Systems Magazine*, 2009, 9(3): 32
- [9] LEWIS F L, VRABIE D, VAMVOUDAKIS K G. Reinforcement learning and feedback control: using natural decision methods to design optimal adaptive controllers[J]. *IEEE Control Systems Magazine*, 2012, 32(6): 76
- [10] WANG D. Robust policy learning control of nonlinear plants with case studies for a power system application[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(3): 1733
- [11] KHIABANI A G, HEYDARI A. Optimal torque control of permanent magnet synchronous motors using adaptive dynamic programming[J]. *IET Power Electronics*, 2020, 13(12): 2442
- [12] LIU R M, LI S K, YANG L X, et al. Energy-efficient subway train scheduling design with time-dependent demand based on an approximate dynamic programming approach[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, 50(7): 2475
- [13] ZHAO B, LIU D R, LUO C M. Reinforcement learning-based optimal stabilization for unknown nonlinear systems subject to inputs with uncertain constraints[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(10): 4330
- [14] LIU Y J, TONG S C, LI D J, et al. Fuzzy adaptive control with state observer for a class of nonlinear discrete-time systems with input constraint[J]. *IEEE Transactions on Fuzzy Systems*, 2016, 24(5): 1147
- [15] ZHANG L G, LIU X J. The synchronization between two discrete-time chaotic systems using active robust model predictive control[J]. *Nonlinear Dynamics*, 2013, 74(4): 905
- [16] TEE K P, GE S S, TAY E H. Barrier Lyapunov Functions for the control of output-constrained nonlinear systems[J]. *Automatica*, 2009, 45(4): 918
- [17] FAN Q Y, YANG G H. Nearly optimal sliding mode fault-tolerant control for affine nonlinear systems with state constraints[J]. *Neurocomputing*, 2016, 216: 78
- [18] XIE C H, YANG G H. Model-free fault-tolerant control approach for uncertain state-constrained linear systems with actuator faults[J]. *International Journal of Adaptive Control and Signal Processing*, 2017, 31(2): 223
- [19] YANG Y L, DING D W, XIONG H Y, et al. Online barrier-actor-critic learning for H_∞ control with full-state constraints and input saturation[J]. *Journal of the Franklin Institute*, 2020, 357(6): 3316
- [20] SONG R Z, LIU L, XIA L N, et al. Online optimal event-triggered H_∞ control for nonlinear systems with constrained state and input[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, 53(1): 131
- [21] 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
- [22] ZHAO B, WANG D, SHI G A, et al. Decentralized control for large-scale nonlinear systems with unknown mismatched interconnections via policy iteration[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018, 48(10): 1725

Decentralized stabilization of state constrained interconnected nonlinear systems based on adaptive dynamic programming

ZHAO Bo DU Wenqian YUAN Guoling KONG Jie

(School of Systems Science, Beijing Normal University, 100875, Beijing, China)

Abstract A decentralized stabilization method based on adaptive dynamic programming (ADP) is proposed for a class of interconnected nonlinear systems with constant-value state constraints. A barrier function is introduced so that the original system is converted into an unconstrained system by coordinate transformation. Auxiliary subsystems and improved cost functions enabled transformation of robust decentralized stabilization problem into an optimal regulation problem. The Hamilton-Jacobi-Bellman (HJB) equation is solved by policy iteration after constructing a local critic neural network for each auxiliary subsystem so that an approximate optimal stabilization control law is obtained. According to the Lyapunov stability theory, the proposed method can drive estimation errors of closed-loop interconnected system and local critic neural networks to be ultimately uniformly bounded dynamically. Numerical simulations validate the effectiveness of proposed decentralized stabilization method.

Keywords adaptive dynamic programming; reinforcement learning; state constraints; interconnected nonlinear systems; decentralized stabilization; optimum control

【责任编辑: 武 佳】