

引用格式:魏阿龙,刘春生,孙景亮.不确定非线性多智能体系统的最优协同控制[J].电光与控制,2018,25(9):12-16,48. WEI A L, LIU C S, SUN J L. Optimal coordination control for uncertain nonlinear multi-agent systems[J]. Electronics Optics & Control, 2018, 25(9):12-16, 48.

不确定非线性多智能体系统的最优协同控制

魏阿龙, 刘春生, 孙景亮
(南京航空航天大学自动化学院,南京 211106)

摘要: 针对不确定非线性多智能体系统的最优协同控制问题,提出了基于辨识-评价结构的在线自适应最优控制方案。由于存在不确定性,针对每一个智能体,分别使用一个辨识神经网络用于估计多智能体系统的动态,以及一个评价神经网络用于近似耦合哈密顿-雅可比方程的解,继而导出最优协同控制律。基于辨识-评价结构,辨识神经网络和评价神经网络可以同步调整权值。通过李雅普诺夫(Lyapunov)直接法同时保证辨识神经网络和评价神经网络的权值误差一致最终有界,及闭环系统的稳定。最后通过仿真实验验证了所提方案的有效性。

关键词: 最优控制; 不确定非线性系统; 多智能体系统; 协同控制

中图分类号: V271.4; O213.2 **文献标志码:** A **doi:**10.3969/j.issn.1671-637X.2018.09.003

Optimal Coordination Control for Uncertain Nonlinear Multi-agent Systems

WEI A-long, LIU Chun-sheng, SUN Jing-liang

(College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: An online adaptive optimal control scheme based on identification-evaluation architecture is presented for the optimal coordination control of uncertain, non-linear, multi-agent systems. In light of systematic uncertainty, an identification Neural Network (NN) is utilized to each agent to approximate the uncertain system dynamics, and an evaluation NN is used to estimate the solution of the coupled Hamilton-Jacobi (HJ) equations. Then, the optimal control law can be derived. Based on the developed architecture, the weights of the identification NN and the evaluation NN can be updated synchronously. By using Lyapunov's direct method, it is guaranteed that the weight errors of the identification NN and the evaluation NN are uniformly and ultimately bounded, and that the closed-loop system is stable. Finally, a numerical example is provided to demonstrate the effectiveness of the proposed control scheme.

Key words: optimal control; uncertain nonlinear system; multi-agent system; coordination control

0 引言

最近几十年,多智能体的一致性(如编队、集群、机器人位置同步等)吸引了越来越多专家学者的关注。文献[1]被看作研究多智能体系统一致性问题的典范,考虑了网络信息流的方向、由于通信失败造成的网络拓扑时变的鲁棒性和时间延迟等实际问题,提出了解决此类问题的理论框架。近年来,一些学者在求

解一致性问题时,考虑了最优控制问题^[2-4]。对于多智能体协同控制问题,实际应用中总是期望每个智能体能够消耗尽量少的能量达到一致性。此外,最优性还能带来系统的其他优良特性,比如幅值裕度和相角裕度增强了系统的鲁棒性。

从数学的观点来看,最优控制问题最终转化为解哈密顿-雅可比-贝尔曼(HJB)方程,这是最优解存在的充分条件。对于非线性系统来说,HJB方程实际上是一个非线性偏微分方程,数学上一般难以求解。尽管动态规划为求解最优控制问题提供了一种方法,但是其缺点也是明显的,随着系统维数的增加,该算法的计算量急剧增加,被称为“维数灾”^[5],此外,其逆序求解思想也阻碍了它在实时控制中的应用。为了克服以上缺点,WERBOS提出了一种自适应动态规划(ADP)算法。

收稿日期:2017-09-04

修回日期:2018-05-22

基金项目:国家自然科学基金(61473147);南京航空航天大学研究生创新基地(实验室)开放基金(kfj20160322);中央高校基本科研业务费专项资金

作者简介:魏阿龙(1992—),男,安徽阜阳人,硕士生,研究方向为自适应动态规划及其在飞行器中的应用。

ADP 算法理论融合了强化学习、动态规划以及函数近似等方法,逼近系统的最优解,从而有效地解决非线性系统的最优控制问题^[5],然而大多数的 ADP 算法需要系统动态的精确信息^[2]。鉴于非线性系统的精确信息一般难以获得,且实际系统中又存在多种内部或外部扰动,一般的 ADP 算法难以应用于实际的控制。文献[6]提出 ADP 的在线迭代算法解决线性多智能体系统的一致性问题,基于此方案,系统的状态矩阵和输入矩阵可以是未知的;文献[7]针对不确定非线性系统提出基于 NN 的在线最优控制。受前面工作的启发,本文研究系统动态模型部分未知的一类不确定多智能体一致性问题,可以看作是对文献[2]中工作的延伸。首先使用一个神经网络逼近系统的不确定项,然后基于辨识的系统设计最优协同控制律,通过李雅普诺夫稳定性理论证明了神经网络权值的收敛性及闭环系统的稳定性,最后通过仿真验证了方案的有效性。

1 图论基础知识

用 $G = (v, \mathcal{E})$ 表示有 N 个节点的多智能体系统通信拓扑图,其中, $v = \{v_1, v_2, \dots, v_N\}$ 为 N 个节点的集合, \mathcal{E} 表示任意两个可达节点构成的边的集合(即,如果存在一条有向路径从节点 v_j 到 v_i ,那么这条有向路径 (v_j, v_i) 属于 \mathcal{E})。第 i 个节点的邻居的集合为 $N_i = \{j | (v_j, v_i) \in \mathcal{E}\}$ 。加权邻接矩阵 $A = [a_{ij}]$,若 $(v_j, v_i) \in \mathcal{E}$,则 $a_{ij} > 0$,否则 $a_{ij} = 0$ 。本文只考虑没有自环和多重边的图,因此 $a_{ij} = 0$ 。节点索引号 i 属于索引集合 $\mathcal{I} = \{1, 2, \dots, N\}$ 。定义图的入度矩阵为 $D = \text{diag}\{d_i\} \in \mathbf{R}^{N \times N}$, $i = 1, 2, \dots, N$,其中, $d_i = \sum_{j=1}^N a_{ij}$ 为 A 的第 i 行行和。定义图的加权拉普拉斯矩阵为 $L = D - A$ ^[8]。注意到,拉普拉斯矩阵所有行和都是零,即 $L\mathbf{1}_N = 0$,其中, $\mathbf{1}_N = [1, 1, \dots, 1]^T \in \mathbf{R}^N$ 为长度为 N 、元素全为 1 的向量。本文提出的方案只要求系统通信拓扑图 G 中包含以领航机为根节点的有向生成树,而不必是强连通的,这意味着,只要求领航机到每架僚机之间至少存在一条有向路径。

2 问题描述

考虑包含 N 个节点的多智能体系统,第 i 个节点的动力学方程为

$$\dot{x}_i = f_i(x_i) + g_i(x_i)u_i \quad (1)$$

式中: $x_i \in \mathbf{R}^n$, $u_i \in \mathbf{R}^m$ 分别是节点 v_i 的状态和控制输入; $f_i(x_i) \in \mathbf{R}^n$ 和 $g_i(x_i) \in \mathbf{R}^{n \times m}$ 在一个包含原点的紧集 Ω_i 上利普希茨连续, $f_i(x_i)$ 未知但满足 $f_i(0) = 0$ 。

全局网络动态方程为

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

式中:全局状态向量 $x = [x_1^T, x_2^T, \dots, x_N^T]^T \in \mathbf{R}^M$;全局动态 $f(x) = [f_1^T(x_1), f_2^T(x_2), \dots, f_N^T(x_N)]^T \in \mathbf{R}^M$; $g(x) = \text{diag}\{g_i(x_i)\} \in \mathbf{R}^{M \times M}$, $i \in \mathcal{I}$;全局控制输入向量 $u = [u_1^T, u_2^T, \dots, u_N^T]^T \in \mathbf{R}^M$, $M = m_1 + \dots + m_N$ 。

领航机的动力学模型为

$$\dot{x}_0 = f_0(x_0) \quad (3)$$

式中: $x_0(t) \in \mathbf{R}^n$; $f_0(x_0)$ 为未知的连续可微函数。

定义节点 v_i 的局部邻域一致性误差为

$$\delta_i = \sum_{j \in N_i} a_{ij}(x_i - x_j) + b_i(x_i - x_0) \quad (4)$$

式中, $\delta_i \in \mathbf{R}^n$ 。牵制增益 $b_i \geq 0$,可以直接获取领航机信息的节点 $b_i > 0$,否则 $b_i = 0$ 。

局部邻域一致性误差动态为

$$\begin{aligned} \dot{\delta}_i &= ((L_i + B_i) \otimes I_n)(\dot{x} - \dot{x}_0) = ((L_i + B_i) \otimes I_n)(f_e(t) + \\ &g(x)u) = \sum_{j \in \bar{N}_i} (l_{ij} + b_{ij})(f_{e_j}(t) + g_j(x_j)u_j) \end{aligned} \quad (5)$$

式中: $B = \text{diag}\{b_i\} \in \mathbf{R}^{N \times N}$; L_i, B_i 分别表示 L 和 B 的第 i 行; $\bar{x}_0 = \underline{I}x_0$; $\underline{I} = I_N \otimes I_n \in \mathbf{R}^{M \times n}$; $f_e(t) = f(x) - f(x_0)$; $f(x_0) = \underline{I}f(x_0)$; $\bar{N}_i = \{i, N_i\}$; $f_{e_j}(t) = f_j(x_j) - f_0(x_0)$ 。

给定节点 v_j 的性能指标为

$$\begin{aligned} J_i(\delta_i(0), u_i, u_{(j)}) &= \int_0^\infty r_i(\delta_i, u_i, u_{(j)}) d\tau = \\ &\int_0^\infty [\delta_i^T Q_i \delta_i + \sum_{j \in \bar{N}_i} u_j^T R_j^i u_j] d\tau \end{aligned} \quad (6)$$

式中: $u_{(j)} = \{u_j | j \in \bar{N}_i\}$; Q_i, R_j^i 皆为常值矩阵且 $Q_i > 0, R_j^i > 0, j \in \bar{N}_i$,否则 $R_j^i = 0$ 。这意味着,节点 v_i 的性能指标只取决于其本身及其邻节点的信息。

假设 1 控制矩阵 $g_i(x_i)$ ($i \in \mathcal{I}$) 已知且有界,即,存在正常数 g_{M_i} ,对于 $\forall x_i \in \Omega_i$ 使得 $\|g_i(x_i)\| \leq g_{M_i}$ 。

定义 1 如果控制律 u_i 不但使式(5)误差系统在紧集 Ω_i 上稳定,还能使式(6)性能指标有界,那么 u_i 称作容许的协同控制策略。

2.1 辨识器设计

基于 Weierstrass 高阶近似定理^[9],选用单隐层神经网络来近似非线性函数 $\mathcal{F}_i(x_i)$,

$$\mathcal{F}_i(x_i) = W_i^T \psi_i(V_i^T x_i) + e_i(x_i) \quad (7)$$

式中: $\psi_i(\cdot) \in \mathbf{R}^l$ 为神经网络激活函数; $e_i(x_i) \in \mathbf{R}^n$ 为重构误差; $V_i \in \mathbf{R}^{n \times l}$ 和 $W_i \in \mathbf{R}^{l \times n}$ 分别对应输入层到隐层和隐层到输出层的权值,隐层神经元个数为 l_i 。激活函数应选取单调非减且有界函数($\|\psi_i(\cdot)\| \leq \psi_{M_i}, \psi_{M_i} > 0$ 为常数)。

由式(1)可得

$$\dot{x}_i = A_i x_i + \mathcal{F}_i(x_i) + g_i(x_i)u_i \quad (8)$$

式中, $\mathcal{F}_i(x_i) = f_i(x_i) - A_i x_i$, $A_i \in \mathbf{R}^{n \times n}$ 为已知常值矩

阵。将式(7)代入式(8)可得

$$\dot{\hat{x}}_i = A_i \hat{x}_i + \hat{W}_i^T \psi_i(V_i^T \hat{x}_i) + g_i(\hat{x}_i) u_i + e_i(x_i) \quad (9)$$

针对式(1)系统设计神经网络辨识器如下

$$\dot{\hat{x}}_i = A_i \hat{x}_i + \hat{W}_i^T \psi_i(\hat{V}_i^T \hat{x}_i) + g_i(\hat{x}_i) u_i + S_i \tilde{x}_i \quad (10)$$

式中: $\hat{x}_i \in \mathbf{R}^n$ 为系统辨识器状态; $\hat{V}_i \in \mathbf{R}^{n \times l_i}$ 和 $\hat{W}_i \in \mathbf{R}^{l_i \times n}$ 分别对应理想权值 V_i 和 W_i 的估计值; $\tilde{x}_i(t) \triangleq x_i(t) - \hat{x}_i(t)$ 为辨识误差; 参数矩阵 $S_i \in \mathbf{R}^{n \times n}$ 应使得 $A_i - S_i$ 为赫尔维茨矩阵。

由式(9)、式(10)可得辨识误差动力学方程为

$$\dot{\tilde{x}}_i = (A_i - S_i) \tilde{x}_i + \tilde{W}_i^T \psi_i(\hat{V}_i^T \hat{x}_i) + \mathfrak{S}_i(x_i) \quad (11)$$

式中: $\tilde{W}_i = W_i - \hat{W}_i$; $\mathfrak{S}_i(x_i) = W_i^T [\psi_i(V_i^T x_i) - \psi_i(\hat{V}_i^T \hat{x}_i)] + [g_i(x_i) - g_i(\hat{x}_i)] u_i + e_i(x_i)$ 。

假设2 辨识神经网络理想权值 V_i 和 W_i 存在上界 V_{M_i} 和 W_{M_i} , 使得 $\|V_i\| \leq V_{M_i}$, $\|W_i\| < W_{M_i}$ 。辨识神经网络重构误差 $e_i(x_i)$ 存在上界 e_{M_i} , 使得 $\|e_i(x_i)\| \leq e_{M_i}$ 。

引理1 若 $A_i - S_i$ 为赫尔维茨矩阵, 则存在唯一的对称正定矩阵 P_i 满足李雅普诺夫方程^[7] $(A_i - S_i)^T P_i + P_i (A_i - S_i) = -\alpha_i I_n$, 其中, $\alpha_i > 0$ 为设计参数。

定理1 对于式(1)系统, 设计式(10)辨识器, 若权值 \hat{V}_i 和 \hat{W}_i 更新为

$$\begin{aligned} \dot{\hat{W}}_i &= -c_{i1} \psi_i(\hat{V}_i^T \hat{x}_i) \tilde{x}_i^T (A_i - S_i)^{-1} - k_{i1} \|\tilde{x}_i\| \hat{W}_i \quad (12) \\ \dot{\hat{V}}_i &= -c_{i2} \text{sgn}(\hat{x}_i) \tilde{x}_i^T (A_i - S_i)^{-1} \hat{W}_i^T \times \\ &\quad (I_{l_i} - \mathfrak{B}(\hat{V}_i^T \hat{x}_i)) - k_{i2} \|\tilde{x}_i\| \hat{V}_i \quad (13) \end{aligned}$$

式中: 设计参数 $c_{ij} > 0$, $k_{ij} > 0$, $j = 1, 2$; $\mathfrak{B}(\hat{V}_i^T \hat{x}_i) = \text{diag}\{\psi_{ij}^2 \cdot (\hat{V}_{ij}^T \hat{x}_i)\}$, $j = 1, \dots, l_i$; $\text{sgn}(\cdot)$ 为符号函数。则可使神经网络辨识误差收敛于一个小的紧集, 即

$$\Omega_{\tilde{x}_i} = \left\{ \tilde{x}_i : \|\tilde{x}_i\| \leq \frac{2\mathfrak{D}_i}{\alpha_i} \right\} \quad (14)$$

此外, 神经网络权值估计误差 $\tilde{V}_i = V_i - \hat{V}_i$ 和 $\tilde{W}_i = W_i - \hat{W}_i$ 都一致最终有界。

证明 考虑如下李雅普诺夫函数

$$L_i(t) = L_{i1}(t) + L_{i2}(t) \quad (15)$$

式中: $L_{i1}(t) = \frac{1}{2} \text{tr}(\tilde{W}_i^T c_{i1}^{-1} \tilde{W}_i) + \frac{1}{2} \text{tr}(\tilde{V}_i^T c_{i2}^{-1} \tilde{V}_i)$;
 $L_{i2}(t) = \frac{1}{2} \tilde{x}_i^T P_i \tilde{x}_i$ 。

对 $L_{i1}(t)$ 求时间导数并将式(12)、式(13)代入可得

$$\begin{aligned} \dot{L}_{i1}(t) &= \text{tr}\{\tilde{W}_i^T \psi_i(\hat{V}_i^T \hat{x}_i) \tilde{x}_i^T (A_i - S_i)^{-1} + \frac{k_{i1}}{c_{i1}} \|\tilde{x}_i\| \times \\ &\quad \tilde{W}_i^T (W_i - \tilde{W}_i)\} + \text{tr}\{\tilde{V}_i^T \text{sgn}(\hat{x}_i) \tilde{x}_i^T \times (A_i - S_i)^{-1} \times \\ &\quad (W_i - \tilde{W}_i)^T (I_{l_i} - \mathfrak{B}(\hat{V}_i^T \hat{x}_i)) - \\ &\quad \frac{k_{i2}}{c_{i2}} \|\tilde{x}_i\| \tilde{V}_i^T (V_i - \tilde{V}_i)\} \quad (16) \end{aligned}$$

注意到, 对于 $\forall X \in \mathbf{R}^n, Y \in \mathbf{R}^{1 \times n}, Z, \tilde{Z} \in \mathbf{R}^{m \times n}$, 都有 $\text{tr}(XY) = \text{tr}(YX) = YX, \text{tr}[\tilde{Z}^T(Z - \tilde{Z})] \leq \|\tilde{Z}\| \|Z\| - \|\tilde{Z}\|^2$, 进而可得

$$\begin{aligned} \dot{L}_{i1}(t) &\leq \beta_i \|\tilde{x}_i\| \|\tilde{W}_i\| + \frac{k_{i1}}{c_{i1}} \|\tilde{x}_i\| (W_{M_i} \|\tilde{W}_i\| - \|\tilde{W}_i\|^2) + \\ &\quad \beta_i \|I_{l_i} - \mathfrak{B}(\hat{V}_i^T \hat{x}_i)\| \|\tilde{x}_i\| (W_{M_i} + \|\tilde{W}_i\|) \|\tilde{V}_i\| + \\ &\quad \frac{k_{i2}}{c_{i2}} \|\tilde{x}_i\| (V_{M_i} \|\tilde{V}_i\| - \|\tilde{V}_i\|^2) \quad (17) \end{aligned}$$

式中, $\beta_i = \|(A_i - S_i)\|^{-1}$ 。

另一方面, 对 $L_{i2}(t)$ 求时间导数得

$$\begin{aligned} \dot{L}_{i2}(t) &= -\frac{\alpha_i}{2} \tilde{x}_i^T \tilde{x}_i + \tilde{x}_i^T P_i [\tilde{W}_i^T \psi_i(\hat{V}_i^T \hat{x}_i) + \mathfrak{S}_i(x_i)] \leq \\ &\quad -\frac{\alpha_i}{2} \|\tilde{x}_i\|^2 + \|\tilde{x}_i\| \|P_i\| (\|\tilde{W}_i\| + \mathfrak{S}_{M_i}) \quad (18) \end{aligned}$$

式中, $\|\mathfrak{S}_i(x_i)\| \leq \mathfrak{S}_{M_i}$ 。

结合式(16)、式(18), 且 $\|I_{l_i} - \mathfrak{B}(\hat{V}_i^T \hat{x}_i)\| \leq 1$, 可得

$$\begin{aligned} \dot{L}_i(t) &\leq -\frac{\alpha_i}{2} \|\tilde{x}_i\|^2 + [\mathfrak{S}_{M_i} \|P_i\| + \left(\frac{k_{i1}}{c_{i1}} - \frac{\beta_i^2}{4}\right) \gamma_{i1}^2 + \\ &\quad \left(\frac{k_{i2}}{c_{i2}} - 1\right) \gamma_{i2}^2 - \left(\frac{k_{i1}}{c_{i1}} - \frac{\beta_i^2}{4}\right) \|\tilde{W}_i + \gamma_{i1}\|^2 - \left(\frac{k_{i2}}{c_{i2}} - 1\right) \cdot \\ &\quad \|\tilde{V}_i + \gamma_{i2}\|^2 - \left(\frac{\beta_i}{2} \|\tilde{W}_i\| - \|\tilde{V}_i\|\right)^2] \|\tilde{x}_i\| \quad (19) \end{aligned}$$

式中: $\gamma_{i1} = \frac{2c_{i1}(\beta_i + \|P_i\|) + 2k_{i1}W_{M_i}}{\beta_i^2 c_{i1} - 4k_{i1}}$; $\gamma_{i2} = \frac{\beta_i c_{i2} W_{M_i} + k_{i2} V_{M_i}}{2(c_{i2} - k_{i2})}$ 。

选择参数使得 $k_{i1} > \frac{\beta_i^2 c_{i1}}{4}$, $k_{i2} > c_{i2}$, 则有

$$\begin{aligned} \dot{L}_i(t) &\leq -\frac{\alpha_i}{2} \|\tilde{x}_i\|^2 + \left[\mathfrak{S}_{M_i} \|P_i\| + \left(\frac{k_{i1}}{c_{i1}} - \frac{\beta_i^2}{4}\right) \gamma_{i1}^2 + \right. \\ &\quad \left. \left(\frac{k_{i2}}{c_{i2}} - 1\right) \gamma_{i2}^2\right] \|\tilde{x}_i\| = -\left(\frac{\alpha_i}{2} \|\tilde{x}_i\| - \mathfrak{D}_i\right) \|\tilde{x}_i\| \quad (20) \end{aligned}$$

式中, $\mathfrak{D}_i = \mathfrak{S}_{M_i} \|P_i\| + \left(\frac{k_{i1}}{c_{i1}} - \frac{\beta_i^2}{4}\right) \gamma_{i1}^2 + \left(\frac{k_{i2}}{c_{i2}} - 1\right) \gamma_{i2}^2$ 。

2.2 最优分布式协同控制器设计

由上节的讨论可知, 式(1)系统可描述为

$$\dot{\hat{x}}_i = \mathfrak{h}_i(\hat{x}_i) + g_i(\hat{x}_i) u_i \quad (21)$$

式中, $\mathfrak{h}_i(\hat{x}_i) = A_i \hat{x}_i + \hat{W}_i^T \psi_i(\hat{V}_i^T \hat{x}_i) + S_i \tilde{x}_i$ 。

节点 v_i 的代价函数为

$$\begin{aligned} V_i(\delta_i(t)) &= \int_t^\infty r_i(\delta_i, u_i, u_{(j)}) d\tau = \\ &\quad \int_t^\infty [\delta_i^T Q_i \delta_i + \sum_{j \in N_i} u_j^T R_j^T u_j] d\tau \quad (22) \end{aligned}$$

非线性李雅普诺夫方程为

$$0 = r_i(\delta_i, u_i, u_{(j)}) + V_{\delta_i}^T \mathfrak{L}_i(\mathfrak{h}_i(t) + g_i(\hat{x}_i) u) \quad (23)$$

式中, $\mathfrak{h}_i(t) = [\mathfrak{h}_1^T(t), \dots, \mathfrak{h}_N^T(t)]^T \in \mathbf{R}^{Nn}$, $\mathfrak{h}_i(t) = \mathfrak{h}_i^T(\hat{x}_i) - \mathfrak{h}_0^T(\hat{x}_0)$, $i \in \mathcal{L}$ 。

定义节点 v_i 的耦合哈密顿函数为

$$H_i(\boldsymbol{\delta}_i, \mathbf{V}_{\delta_i}, u_i, u_{(j)}) = r_i(\boldsymbol{\delta}_i, u_i, u_{(j)}) + \mathbf{V}_{\delta_i}^T \mathfrak{L}_i(\boldsymbol{h}_e(t) + g(\hat{x})\mathbf{u}) \quad (24)$$

式中, $\mathfrak{L}_i = ((L_i + B_i) \otimes I_n)$ 。

由最优控制原理的必要条件得

$$u_i^* = -\frac{1}{2} \mathbf{R}_i^{i-1} \mathbf{g}_i^T(\hat{x}_i) (l_{ii} + b_i) \mathbf{V}_{\delta_i}^* \quad (25)$$

式中, l_{ii} 为拉普拉斯矩阵 L 的元素。

将式(25)最优控制律代入式(23)可得

$$\begin{aligned} 0 = & \boldsymbol{\delta}_i^T \mathbf{Q}_i \boldsymbol{\delta}_i - \frac{1}{4} \mathbf{V}_{\delta_i}^{*T} (l_{ii} + b_i)^2 \mathbf{g}_i(\hat{x}_i) \mathbf{R}_i^{i-1} \mathbf{g}_i^T(\hat{x}_i) \mathbf{V}_{\delta_i}^* + \\ & \frac{1}{4} \sum_{j \in \bar{N}_i} \mathbf{V}_{\delta_j}^{*T} (l_{jj} + b_j)^2 \mathbf{g}_j(\hat{x}_j) \mathbf{R}_j^{j-1} \mathbf{R}_j^j \mathbf{R}_j^{j-1} \mathbf{g}_j^T(\hat{x}_j) \mathbf{V}_{\delta_j}^* - \\ & \frac{1}{2} \mathbf{V}_{\delta_i}^{*T} \sum_{j \in \bar{N}_i} (l_{jj} + b_j)^2 \mathbf{g}_j(\hat{x}_j) \mathbf{R}_j^{j-1} \mathbf{R}_j^j \mathbf{R}_j^{j-1} \mathbf{g}_j^T(\hat{x}_j) \mathbf{V}_{\delta_j}^* + \\ & \mathbf{V}_{\delta_i}^{*T} \sum_{j \in \bar{N}_i} (l_{jj} + b_j) \boldsymbol{h}_j(t) \quad (26) \end{aligned}$$

注意到,一般如式(26)的耦合非线性偏微分方程在数学上难以得到解析解。本文使用单隐层神经网络模型(评价网络)近似理想值函数 $V_i^*(\boldsymbol{\delta}_i)$ 及其关于 $\boldsymbol{\delta}_i$ 的偏导数,形如

$$V_i^*(\boldsymbol{\delta}_i) = \boldsymbol{\theta}_i^T \boldsymbol{\sigma}(\boldsymbol{\delta}_i) + \varepsilon_i(\boldsymbol{\delta}_i) \quad (27)$$

$$\mathbf{V}_{\delta_i}^* = \nabla \boldsymbol{\sigma}^T(\boldsymbol{\delta}_i) \boldsymbol{\theta}_i + \nabla \varepsilon_i \quad (28)$$

式中: $\boldsymbol{\theta}_i = [\theta_{i1}, \theta_{i2}, \dots, \theta_{i\bar{h}}]^T \in \mathbf{R}^{\bar{h}}$ 为理想权值向量; $\boldsymbol{\sigma}(\boldsymbol{\delta}_i): \mathbf{R}^n \rightarrow \mathbf{R}^{\bar{h}}$ 为激活函数; $\varepsilon_i(\boldsymbol{\delta}_i)$ 为神经网络近似误差; $\nabla \boldsymbol{\sigma}(\boldsymbol{\delta}_i) = \frac{\partial \boldsymbol{\sigma}(\boldsymbol{\delta}_i)}{\partial \boldsymbol{\delta}_i}$, $\nabla \varepsilon_i = \frac{\partial \varepsilon_i}{\partial \boldsymbol{\delta}_i}$ 。

由于理想权值未知,用其估计值 $\hat{\boldsymbol{\theta}}_i$ 表示为

$$\hat{V}_i(\boldsymbol{\delta}_i) = \hat{\boldsymbol{\theta}}_i^T \boldsymbol{\sigma}(\boldsymbol{\delta}_i) \quad (29)$$

$$\hat{\mathbf{V}}_{\delta_i}(\boldsymbol{\delta}_i) = \nabla \boldsymbol{\sigma}^T(\boldsymbol{\delta}_i) \hat{\boldsymbol{\theta}}_i \quad (30)$$

进而,式(24)哈密顿函数可表示为

$$\begin{aligned} e_i = & H_i(\boldsymbol{\delta}_i, \hat{\boldsymbol{\theta}}_i, \hat{u}_i, \hat{u}_{(j)}) = r_i(\boldsymbol{\delta}_i, u_i, u_{(j)}) + \\ & \hat{\mathbf{V}}_{\delta_i}^T \mathfrak{L}_i(\boldsymbol{h}_e(t) + g(\hat{x})\mathbf{u}) \quad (31) \end{aligned}$$

给定容许的协同控制策略 $u_i, u_{(j)}$, 理想的 $\hat{\boldsymbol{\theta}}_i$ 应该使残差平方 $E_i(\hat{\boldsymbol{\theta}}_i)$ 最小

$$E_i(\hat{\boldsymbol{\theta}}_i) = \frac{1}{2} \mathbf{e}_i^T \mathbf{e}_i \quad (32)$$

由梯度下降法得,权值更新律为

$$\dot{\hat{\boldsymbol{\theta}}}_i = -\lambda_i \boldsymbol{\phi}_i(\boldsymbol{\phi}_i^T \hat{\boldsymbol{\theta}}_i + r_i(\boldsymbol{\delta}_i, u_i, u_{(j)})) \quad (33)$$

式中: $\lambda_i > 0$ 为学习率; $\boldsymbol{\phi}_i = \nabla \boldsymbol{\sigma} \mathfrak{L}_i(\boldsymbol{h}_e(t) + g(\hat{x})\mathbf{u})$ 。

将式(28)代入式(25)得

$$\hat{u}_i = -\frac{1}{2} \mathbf{R}_i^{i-1} \mathbf{g}_i^T(\hat{x}_i) ((l_{ii} + b_i) \otimes I_n)^T \nabla \boldsymbol{\sigma}^T(\boldsymbol{\delta}_i) \hat{\boldsymbol{\theta}}_i \quad (34)$$

将式(28)代入式(23)可得

$$r_i(\boldsymbol{\delta}_i, u_i, u_{(j)}) + \boldsymbol{\theta}_i^T \boldsymbol{\phi}_i = \varepsilon_{H_i} \quad (35)$$

式中, $\varepsilon_{H_i} = -\nabla \boldsymbol{\varepsilon}_i^T \mathfrak{L}_i(\boldsymbol{h}_e(t) + g(\hat{x})\mathbf{u})$ 。

定义权值估计误差 $\tilde{\boldsymbol{\theta}}_i = \hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i$, 由式(33)和式(35)得

$$\dot{\tilde{\boldsymbol{\theta}}}_i = \dot{\hat{\boldsymbol{\theta}}}_i - \dot{\boldsymbol{\theta}}_i = \dot{\hat{\boldsymbol{\theta}}}_i = -\lambda_i \boldsymbol{\phi}_i(\boldsymbol{\phi}_i^T \tilde{\boldsymbol{\theta}}_i + \varepsilon_{H_i}) \quad (36)$$

定理2 对于部分未知的多智能体系统(1),使用式(25)协同控制策略,其中,各权值分别按照式(12)、式(13)、式(33)调整,则局部一致性误差 $\boldsymbol{\delta}_i$ 、辨识神经网络权值误差 $\tilde{\mathbf{W}}_i$ 、 $\tilde{\mathbf{V}}_i$ 和评价网络权值估计误差 $\tilde{\boldsymbol{\theta}}_i$ 都是一致最终有界的。控制节点轨迹 $x_0(t)$ 为协同一致最终有界的,即,所有节点与领航状态一致。此外,近似协同控制律 \hat{u}_i 与理想协同控制律 u_i 的误差在一个小的领域内,即 $\|\hat{u}_i - u_i\| \leq \varepsilon_{u_i}(t \rightarrow \infty)$ 。

证明 考虑如下李雅普诺夫函数

$$J_i(t) = J_{i1}(t) + J_{i2}(t) + J_{i3}(t) \quad (37)$$

式中: $J_{i1}(t) = L_i(t)$, $L_i(t)$ 在定理1中定义, $J_{i2}(t) = \frac{\text{tr}(\tilde{\boldsymbol{\theta}}_i^T \tilde{\boldsymbol{\theta}}_i)}{2\lambda_i}$, $J_{i3}(t) = \boldsymbol{\delta}_i^T \boldsymbol{\delta}_i + 2\Gamma_i V_i(\boldsymbol{\delta}_i)$, $\Gamma_i > 0$ 。

$$\begin{aligned} \dot{J}_{i2}(t) = & -\frac{\text{tr}(\tilde{\boldsymbol{\theta}}_i^T \dot{\tilde{\boldsymbol{\theta}}}_i)}{\lambda_i} = -\tilde{\boldsymbol{\theta}}_i^T \boldsymbol{\phi}_i \boldsymbol{\phi}_i^T \tilde{\boldsymbol{\theta}}_i - \\ & 2 \frac{\lambda_i}{\sqrt{2\lambda_i}} \tilde{\boldsymbol{\theta}}_i^T \boldsymbol{\phi}_i \frac{1}{\sqrt{2\lambda_i}} \varepsilon_{H_i} \quad (38) \end{aligned}$$

由于 $\tilde{\boldsymbol{\theta}}_i^T \boldsymbol{\phi}_i \boldsymbol{\phi}_i^T \tilde{\boldsymbol{\theta}}_i > 0$, 存在 $q_i > 0$ 使得 $q_i \|\tilde{\boldsymbol{\theta}}_i\|^2 \leq \tilde{\boldsymbol{\theta}}_i^T \boldsymbol{\phi}_i \boldsymbol{\phi}_i^T \tilde{\boldsymbol{\theta}}_i$ ($q_i \leq \|\boldsymbol{\phi}_i\|^2$), 可得

$$\dot{J}_{i2}(t) \leq (-q_i + \frac{\lambda_i}{2} \|\boldsymbol{\phi}_i\|^2) \|\tilde{\boldsymbol{\theta}}_i\|^2 + \frac{1}{2\lambda_i} b_{\varepsilon_i} \quad (39)$$

$$\dot{J}_{i3}(t) = 2\boldsymbol{\delta}_i^T \dot{\boldsymbol{\delta}}_i + 2\Gamma_i \dot{V}_i(\boldsymbol{\delta}_i) \leq \sum_{j \in \bar{N}_i} (\|((l_{ij} + b_{ij}) \otimes I_n) \mathbf{g}_{M_i}\|^2 -$$

$$2\Gamma_i \lambda_{\min}(\mathbf{R}_i^j) \|\mathbf{u}_j\|^2 + (2(\Lambda_i + 1) - 2\Gamma_i \lambda_{\min}(\mathbf{Q}_i)) \|\boldsymbol{\delta}_i\|^2 + \sum_{j \in \bar{N}_i} \|((l_{ij} + b_{ij}) \otimes I_n) \boldsymbol{h}_e(t)\|^2) \quad (40)$$

式中, Λ_i 为节点 v_i 的邻节点的个数。进而可得

$$\begin{aligned} \dot{J}_i(t) \leq & -\frac{1}{2} \alpha_i \left(\|\tilde{\mathbf{x}}_i\| - \frac{\mathfrak{D}_i}{\alpha_i} \right)^2 + \frac{\mathfrak{D}_i^2}{2\alpha_i} + \left(-q_i + \frac{\lambda_i}{2} \|\boldsymbol{\phi}_i\|^2 \right) \cdot \\ & \|\tilde{\boldsymbol{\theta}}_i\|^2 + \frac{1}{2\lambda_i} b_{\varepsilon_i} + (2(\Lambda_i + 1) - 2\Gamma_i \lambda_{\min}(\mathbf{Q}_i)) \|\boldsymbol{\delta}_i\|^2 + \\ & \sum_{j \in \bar{N}_i} (\|((l_{ij} + b_{ij}) \otimes I_n) \mathbf{g}_{M_i}\|^2 - 2\Gamma_i \lambda_{\min}(\mathbf{R}_i^j)) \|\mathbf{u}_j\|^2 + \\ & \sum_{j \in \bar{N}_i} \|((l_{ij} + b_{ij}) \otimes I_n) \boldsymbol{h}_e(t)\|^2 \quad (41) \end{aligned}$$

选择参数使 $0 < \lambda_i < \frac{2q_i}{\|\boldsymbol{\phi}_i\|^2}$, $\Gamma_i > \max \left\{ \frac{(\Lambda_i + 1)}{\lambda_{\min}(\mathbf{Q}_i)}, \left\{ \frac{\|((l_{ij} + b_{ij}) \otimes I_n) \mathbf{g}_{M_i}\|^2}{2\lambda_{\min}(\mathbf{R}_i^j)}, j \in \bar{N}_i \right\} \right\}$ 。

令 $e_i = \frac{\mathfrak{D}_i^2}{2\alpha_i} + \frac{1}{2\lambda_i} b_{e_i} + \sum_{j \in N_i} \|((l_{ij} + b_{ij}) \otimes I_n) \tilde{y}_e(t)\|^2$, 则

$$\|\tilde{x}_i\| > \frac{\mathfrak{D}_i}{\alpha_i} + \sqrt{\frac{2e_i}{\alpha_i}} \triangleq b_{\tilde{x}_i} \quad (42)$$

$$\|\tilde{\theta}_i\| > \sqrt{\frac{e_i}{q_i - \frac{\lambda_i}{2} \phi_{M_i}^2}} \triangleq b_{\tilde{\theta}_i} \quad (43)$$

$$\|\delta_i\| > \sqrt{\frac{e_i}{2\Gamma_i \lambda_{\min}(\mathbf{Q}_i) - 2(A_i + 1)}} \quad (44)$$

任一式成立时, 都有 $\dot{J}_i(t) < 0$ 。

另一方面, 由式(25)和式(34)可得

$$\|\hat{u}_i - u_i\| \leq \frac{1}{2} \|\mathbf{R}_i^{-1} \mathbf{g}_i^T(\hat{x}_i)(l_{ii} + b_{ii})\| \sqrt{2(b_{\tilde{\theta}_i}^2 + \varepsilon_{\nabla M_i}^2)} \quad (45)$$

在定理2中对一致性进行了证明, 针对所得控制律关于设计指标的最优性证明, 限于篇幅, 仅作简单说明: 当文献[10]中不考虑干扰情况, 且令其设计参数 $\varepsilon_1 = \varepsilon_2 = 1$ 时, 结论是明显的, 详细证明过程请参考文献[10]。本文是对文献[2]中工作的延伸。文献[2]中提出的控制方案需要精确的系统模型信息, 然而这在实际的工程应用中很难实现。本文通过增加一个辨识神经网络, 降低了这个苛刻条件, 即允许系统动态信息 $f(x)$ 未知, 降低了对模型的依赖程度, 增强了闭环系统的鲁棒性。

3 仿真验证

研究的多智能体系统^[2]如图1所示。

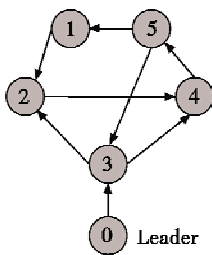


图1 有领机的多智能体系统拓扑图

Fig. 1 Topology of multiagent system with leader node

各智能体的动态方程描述为 $f(x_i) = (x_{i2} - x_{i1}^2 x_{i2} - (x_{i1} + x_{i2})(1 - x_{i1})^2)^T, i = 1, \dots, 5, \mathbf{g}_1(x_1) = (0 \ x_{12}^2)^T, \mathbf{g}_2(x_2) = (0 \ 1.5x_{22}^2)^T, \mathbf{g}_3(x_3) = (0 \ -0.2x_{32}^2)^T, \mathbf{g}_4(x_4) = (0 \ 0.3x_{42}^2)^T, \mathbf{g}_5(x_5) = (0 \ -0.9x_{52}^2)^T$ 。

辨识器相关参数分别为 $\mathbf{A}_i = [-1, 1, -0.5, -0.5], c_{i1} = c_{i2} = 40, \mathbf{S}_i = [1, 0, -0.5, 0], k_{i1} = 15, k_{i2} = 45, l_i = 8$; 令边的权值和牵制增益为1, $\mathbf{Q}_i = \mathbf{I}_2, \mathbf{R}_i^i = 10, \mathbf{R}_i^j = 1, i \neq j, \lambda_i = 0.01, i = 1, \dots, 5$ 。辨识网络初

始权值 $\hat{\mathbf{W}}_i, \hat{\mathbf{V}}_i$ 在区间 $[-2, 2]$ 随机选择, 评价网络初始权值在 $[-1, 1]$ 随机选择。不失一般性, 选用双曲正切函数 $\tanh(\cdot)$ 作为两个 NN 的激活函数。

仿真结果如图2~图6所示。图2是最优协同控制策略下的各智能体的状态响应曲线, 13 s后各节点状态与领机保持一致; 图3揭示了辨识器的性能, 13 s后辨识误差趋于零; 图4是辨识器权值的 Frobenius 范数, 13 s后辨识器权值调节结束; 图5是评价神经网络权值的学习过程; 图6是一致性误差响应曲线, 13 s后系统达到一致。由图2~图6可知, 辨识器权值和评价网络权值同时调节, 同时闭环系统稳定。

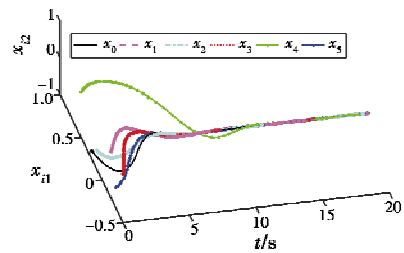


图2 各智能体的状态轨迹

Fig. 2 The evolution of each agent's state

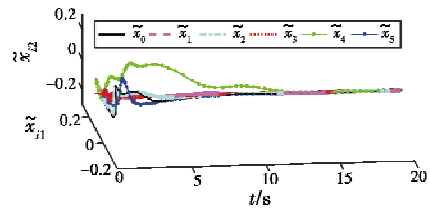


图3 系统辨识误差

Fig. 3 System identification error

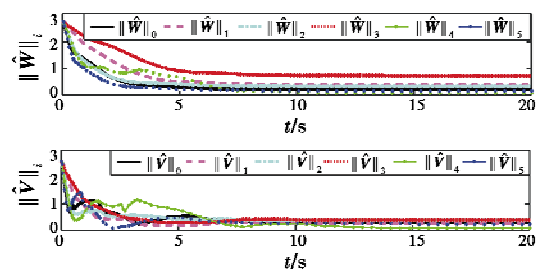


图4 辨识神经网络权值的 Frobenius 范数

Fig. 4 Frobenius-norm of identification NN weights

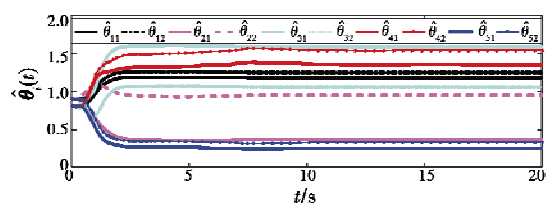


图5 评价网络权值

Fig. 5 Evaluation network weights

为双机队形不存在制导盲区。

通过对两种编队队形下的双机协同攻击区仿真结果进行分析,可以得出以下结论:1) 双机协同攻击区大于单机制导攻击区,但不是协同探测区域的并集,而是在此基础上受到的各种限制有所缩小;2) 协同攻击区同样受到平台探测角度、平台制导角度、导弹指令接收角度的限制;3) 协同攻击区的面积与编队形式有关,合适的队形能够放宽平台探测角度、制导角度、导弹指令接收角度的限制,扩大协同攻击区大小。

3 结束语

本文首先介绍协同攻击区的概念、发射/制导方式和约束条件,并阐述了协同攻击区计算模型,其次,仿真分析了两种编队队形下的协同攻击区,得到如下结论:双机协同能够充分发挥空空导弹的性能;协同攻击区同样受到平台探测角度、制导角度、导弹指令接收角度的限制;编队形式对协同攻击区的大小也有很大的影响。

(上接第16页)

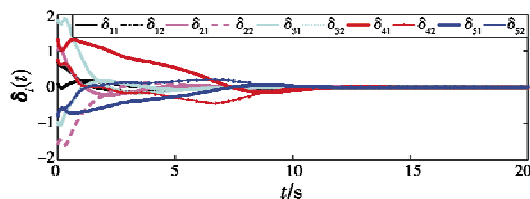


图6 局部一致性误差

Fig. 6 Evolution of the local consensus error

4 结束语

本文针对不确定非线性多智能体系统的最优控制问题,提出了在线自适应协同控制方案。使用两组神经网络:一组用于估计不确定系统;另一组用于近似最优值函数。基于提出的控制方法,两组网络同时在线学习。通过李雅普诺夫稳定性理论,证明了闭环系统的稳定性。本文所提算法的缺点之一是神经元的个数需要多次尝试后才能确定,所以接下来可以尝试设计权值个数自适应寻优的最优协同控制算法。

参考文献

[1] OLFATI-SABER R, FAX J A, MURRAY R M. Consensus and cooperation in networked multi-agent systems [J]. Proceedings of the IEEE, 2007, 95(1):215-233.

[2] ZHANG H G, ZHANG J L, 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,

参考文献

- [1] 孙隆和. 网络信息环境中的飞机(飞行作战平台)武器火力指挥控制系统[J]. 电光与控制, 2010, 17(3):1-8.
- [2] YOUNG B W. Future integrated fire control [C]//The 10th International Command and Control Research and Technology Symposium, 2005:68-72.
- [3] 高劲松,丁全心,邹庆元. 网络中心战对空战研究的影响[J]. 电光与控制, 2008, 15(9):43-47.
- [4] 寇英信,付昭旺,冯国强,等. 网络瞄准下空空导弹允许发射区模拟仿真分析[J]. 空军工程大学学报:自然科学版, 2012, 13(2):24-28.
- [5] 刁兴华,方洋旺,伍友利,等. 双机编队空空导弹协同发射区模拟仿真分析[J]. 北京航空航天大学学报, 2014, 40(3):370-376.
- [6] 周志刚. 机载火力控制系统分析[M]. 北京:国防工业出版社, 2008.
- [7] 樊会涛. 空空导弹方案设计原理[M]. 北京:航空工业出版社, 2013.
- [8] 2015, 23(1):152-163.
- [3] MOVRIC K H, LEWIS F L. Cooperative optimal control for multi-agent systems on directed graph topologies [J]. IEEE Transactions on Automatic Control, 2014, 59(3):769-774.
- [4] ZHANG H G, FENG T, YANG G H, et al. Distributed cooperative optimal control for multiagent systems on directed graphs: an inverse optimal approach [J]. IEEE Transactions on Cybernetics, 2015, 45(7):1315-1326.
- [5] 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.
- [6] 李健,沈艳军,刘允刚. 线性多智能体系统一致性的自适应动态规划求解方法[J]. 系统科学与数学, 2016, 36(7):1016-1030.
- [7] YANG X, LIU D R, HUANG Y Z. Neural-network-based on-line optimal control for uncertain non-linear continuous-time systems with control constraints [J]. IET Control Theory & Applications, 2013, 7(17):2037-2047.
- [8] 王庆领. 饱和受限的多智能体系统一致性问题研究[D]. 哈尔滨:哈尔滨工业大学, 2014.
- [9] MODARES H, LEWIS F L. Optimal tracking control of non-linear partially-unknown constrained-input systems using integral reinforcement learning [J]. Automatica, 2014, 50(7):1780-1792.
- [10] SUN J L, LIU C S, YE Q. Robust differential game guidance laws design for uncertain interceptor-target engagement via adaptive dynamic programming [J]. International Journal of Control, 2017, 90(5):990-1004.