

引用格式:蒲俊,马清亮,顾凡.基于自适应动态规划的未知模型非线性系统 H_2/H_∞ 控制[J].电光与控制,2018,25(9):17-21. PU J, MA Q L, GU F. H_2/H_∞ control of an unknown model nonlinear system based on adaptive dynamic programming[J]. Electronics Optics & Control, 2018, 25(9): 17-21.

基于自适应动态规划的未知模型非线性系统 H_2/H_∞ 控制

蒲俊, 马清亮, 顾凡
(火箭军工程大学, 西安 710025)

摘要: 提出了一种在线的自适应动态规划算法, 近似求解耦合的哈密尔顿雅可比 (Hamilton-Jacobi-Isaacs, HJI) 方程, 获得非线性系统混合 H_2/H_∞ 控制的纳什均衡策略。通过在控制策略和干扰策略中加入已知噪声, 从而不依赖系统的模型信息, 得到一个求解混合 H_2/H_∞ 控制问题的未知模型的近似动态规划算法。分别使用 2 个评价神经网络和 2 个执行神经网络, 同步在线更新 2 个值函数、控制策略和干扰策略, 神经网络未知参数通过最小二乘法进行估计。仿真结果验证了算法的可行性。

关键词: 自适应动态规划; H_2/H_∞ 控制; 耦合 HJIE; 最优控制; 神经网络

中图分类号: V271.4; TP13 **文献标志码:** A **doi:**10.3969/j.issn.1671-637X.2018.09.004

H_2/H_∞ Control of an Unknown Model Nonlinear System Based on Adaptive Dynamic Programming

PU Jun, MA Qing-liang, GU Fan
(Rocket Force University of Engineering, Xi'an 710025, China)

Abstract: An online adaptive dynamic programming algorithm is proposed for getting the approximate solution of the coupled Hamilton-Jacobi-Isaacs Equations (HJIE), and obtaining the Nash equilibrium strategy of mixed H_2/H_∞ control of nonlinear system. By adding the detection signal to the control strategy and the interference strategy, an approximate dynamic programming algorithm is acquired for solving mixed H_2/H_∞ control problems with unknown model without depending on model information of the system. Two critic neural networks and two executive neural networks are used to synchronously update two value functions, control strategies and interference strategies online. The unknown parameters of the neural network are estimated by generalized least squares. The simulation results verify the feasibility of the algorithm.

Key words: adaptive dynamic programming; H_2/H_∞ control; coupled Hamilton-Jacobi-Isaacs equations; optimum control; neural network

0 引言

H_2 范数和 H_∞ 范数是最优控制理论中的两个非常重要的系统性能度量指标。 H_2 性能指标反映了系统调节性能的好坏程度, 在设计 H_2 控制器时, 就会损失系统鲁棒稳定能力; H_∞ 性能指标反映了系统鲁棒稳定性的好坏程度, 在设计 H_∞ 控制器时, 就会损失系统调节能力。但是在对非线性系统研究时, 往往要求系统在具有一定调节能力的同时拥有很好的鲁棒性, 在这

样的背景下, BERSTEIN^[1]等首次提出混合 H_2/H_∞ 控制的设想。此后, 很多学者对混合 H_2/H_∞ 控制作了进一步研究, 潘伟^[2]等提出了基于遗传算法设计线性系统 H_2/H_∞ 状态反馈控制器; 姜海波等通过在满足给定 H_∞ 性能下达到 H_2 性能来设计 H_2/H_∞ 控制器; 叶思隽^[3]等提出直接迭代的 LMI 方法设计不确定系统的混合 H_2/H_∞ 鲁棒控制器; 马清亮^[4]等基于多项式 Lyapunov 函数与平方和分解技术, 提出了混合 H_2/H_∞ 模糊控制器的设计方法。但是上述方法由于在求解控制器时, 引入了约束条件, 使所求问题具有一定的保守性, 通过求解混合 H_2/H_∞ 控制问题所产生的耦合哈密尔顿雅可比 (Hamilton-Jacobi-Isaacs, HJI) 方程, 在不引入任何约束的条件下设计控制器还很少见, 也是此类问题的难点。

自适应动态规划 (Adaptive Dynamic Programming,

收稿日期: 2017-12-13 修回日期: 2018-08-17

基金项目: 国家自然科学基金 (61773387)

作者简介: 蒲俊 (1994—), 男, 四川绵阳人, 硕士生, 研究方向为先导控制理论、非线性系统控制。

ADP)理论是结合最优控制、自适应控制、强化学习、智能控制等先进控制理论发展而来的一种新型近似最优控制理论^[5-6],利用函数近似结构估计代价函数,这就为近似求解 HJI 方程提供了可能^[7-8]。VAMVOUDAKIS^[9]等提出策略迭代算法通过逐步求解类非线性 Lyapunov 方程来求解耦合的 HJI 方程;ALIYU^[10]等提出迭代的方法近似求解耦合 HJI 方程,得到纳什均衡策略。但是大多研究都是建立在已知系统确切的模型信息基础上,而实际系统不可避免地会受到多种因素的影响,很大程度上造成其模型未知或部分未知。

针对上述问题,基于 ADP 理论提出一种用于解决混合 H_2/H_∞ 控制问题的近似动态规划算法,通过分别在控制和干扰两个策略中加入已知噪声 e ,在线采集系统实时数据,用最小二乘(LS)法估计神经网络的参数,从而在不需系统确切的模型信息的情况下,求解混合 H_2/H_∞ 控制问题的非零和博弈的纳什均衡策略。采用4个神经网络,同时在线更新2个值函数、控制策略和干扰策略,最后仿真结果验证了算法的可行性。

1 问题描述

考虑如下连续非线性系统

$$\begin{cases} \dot{x} = f(x) + g(x)u + k(x)w \\ z = h(x) \end{cases} \quad (1)$$

式中: $x \in \mathbf{R}^n$ 为系统状态; $u \in \mathbf{R}^m$ 为控制输入; $w \in \mathbf{R}^q$ 为外部干扰输入; $z \in \mathbf{R}^p$ 为虚拟输出; $f(x) \in \mathbf{R}^n$, $g(x) \in \mathbf{R}^{n \times m}$ 和 $k(x) \in \mathbf{R}^{n \times q}$ 为系统内部模型。假设系统模型都是未知的, $x=0$ 是系统(1)的一个平衡点,即系统是零状态可观测。

设计混合 H_2/H_∞ 控制器,就是找到控制输入 $u = u^*$ 满足如下条件。

1) 存在 $x=0$ 的邻域 N ,对可积函数 $w: [0, \infty] \rightarrow \mathbf{R}^q$,从 $x_0=0$ 开始,闭环系统 $\dot{x} = f(x) + g(x)u^* + k(x)w$ 轨迹始终落在 N 里,且满足 L_2 增益 (L_2 -gain) 在集合 Ω 上小于等于 γ ,即

$$\int_0^\infty (h^T(x)h(x) + u^T R u) dt \leq \gamma^2 \int_0^\infty \|w\|^2 dt \quad (2)$$

式中: u^* 为最优控制策略; γ 为大于零的常数; R 为适当维数的对角矩阵。

2) 在最坏干扰情况下依然可以实现非线性二次最优控制(H_2 控制),即

$$\begin{aligned} \min_{u \in \Omega} J(u, w^*) &= J(u^*, w^*) = \\ \min_{u \in \Omega, w = w^*} \int_0^\infty (h^T(x)h(x) + u^T R u) dt \end{aligned} \quad (3)$$

式中: Ω 为容许控制策略集合; $J(\cdot)$ 为性能指标函数; w^* 是最坏干扰。

3) 系统 $\dot{x} = f(x) + g(x)u^*(x)$ 在平衡点 $x=0$ 是局部渐进稳定的。

定义如下两个性能指标

$$J_1(x_0, u, w) = \int_0^\infty (h^T(x)h(x) + u^T R_1 u) dt \triangleq \int_0^\infty r_1(x, u, w) dt \quad (4)$$

$$J_2(x_0, u, w) = \int_0^\infty (h^T(x)h(x) + u^T R_1 u - \gamma^2 w^T R_2 w) dt \triangleq \int_0^\infty r_2(x, u, w) dt \quad (5)$$

式中, x_0 为零时刻系统状态。

事实上,式(4)为 H_2 最优控制性能指标,式(5)为 H_∞ 最优控制性能指标^[10]。

定义控制输入 u 和干扰输入 w 的值函数分别为

$$V_1(x_t, u, w) = \int_t^\infty (h^T(x)h(x) + u^T R_1 u) dt \quad (6)$$

$$V_2(x_t, u, w) = \int_t^\infty (h^T(x)h(x) + u^T R_1 u - \gamma^2 w^T R_2 w) dt \quad (7)$$

式中, x_t 为 t 时刻系统状态。

根据博弈论相关理论可知,求解混合 H_2/H_∞ 最优控制策略 u^* ,可以看成是求解一个非零和博弈问题,其目标是同时最小化值函数式(6)、式(7),找到满足下列不等式的最优控制组(u^*, w^*),即纳什均衡策略为

$$\begin{cases} V_1(x_t, u^*, w^*) \leq V_1(x_t, u, w^*) \\ V_2(x_t, u^*, w^*) \leq V_2(x_t, u^*, w) \end{cases} \quad (8)$$

对值函数微分得到非线性 Lyapunov 方程为

$$r_i(x, u, w) + (\nabla V_i)^T (f(x) + g(x)u + k(x)w) = 0 \quad i=1,2 \quad (9)$$

式中, $\nabla V_i = \partial V_i / \partial x$ 。

定义哈密尔顿函数为

$$H_i(x, \nabla V_i, u, w) = r_i(x, u, w) + (\nabla V_i)^T (f(x) + g(x)u + k(x)w) \quad i=1,2 \quad (10)$$

由 $\frac{\partial V_i}{\partial u} = 0, \frac{\partial V_i}{\partial w} = 0$ 得到

$$\begin{cases} u = -\frac{1}{2} R_1^{-1} g^T(x) \nabla V_1 \\ w = \frac{1}{2\gamma^2} R_2^{-1} k^T(x) \nabla V_2 \end{cases} \quad (11)$$

将式(11)代入式(10)得到如下两个耦合的 HJI 方程

$$h^T(x)h(x) - \frac{1}{4} (\nabla V_1)^T g(x) R_1^{-1} g^T(x) \nabla V_1 + (\nabla V_1)^T \cdot$$

$$f(x) + \frac{1}{2\gamma^2} (\nabla V_1)^T k(x) R_2^{-1} k^T(x) \nabla V_2 = 0 \quad (12)$$

$$h^T(x)h(x) + \frac{1}{4\gamma^2} (\nabla V_2)^T k(x) R_2^{-1} k^T(x) \nabla V_2 +$$

$$\begin{aligned} & (\nabla V_2)^T f(x) + \frac{1}{4} (\nabla V_1)^T g(x) R_1^{-1} g^T(x) \nabla V_1 - \\ & \frac{1}{2} (\nabla V_2)^T g(x) R_1^{-1} g^T(x) \nabla V_1 = 0 \quad (13) \end{aligned}$$

式(12)、式(13)是两个耦合的非线性偏微分方程,得到其解析解是非常困难的,但是基于自适应动态规划思想,利用神经网络强大的逼近能力,近似求解耦合的HJI方程,可以证明^[10],方程的解(V_1, V_2)一致收敛到(V_1^*, V_2^*)。

假设1 选择 $\gamma > 0$ 且系统(1)是零状态可观的。在集合 $\Omega \in \mathbf{R}^n$ 上存在一个控制输入 u 使系统渐进稳定且有 L_2 -gain $\leq \gamma$ 。HJI 方程式(13)在 Ω 上存在一个光滑解^[11]。

2 基于 ADP 的无模型策略迭代算法

为了分析未知模型的近似动态规划算法,首先推导出模型已知的策略迭代(Policy Iteration, PI)算法。

算法1 步骤如下所述。

Step 1 分别给定初始稳定的控制、干扰输入 u_1 和 w_1 , 此时 $i = 1$ 。

Step 2(策略评价) 依据已知 u_i 和 w_i , 利用如下 Lyapunov 方程求解 V_1^i 和 V_2^i

$$\begin{cases} r_1(x, u_i, w_i) + (\nabla V_1^i)^T(x) (f(x) + g(x)u_i(x) + k(x)w_i(x)) = 0 \\ r_2(x, u_i, w_i) + (\nabla V_2^i)^T(x) (f(x) + g(x)u_i(x) + k(x)w_i(x)) = 0 \end{cases} \quad (14)$$

Step 3(策略提高) $i = i + 1$, 更新 u_i 和 w_i 如下

$$\begin{cases} u_{i+1}(x) = -\frac{1}{2} R_1^{-1} g^T(x) \nabla V_1^{i+1} \\ w_{i+1}(x) = \frac{1}{2\gamma^2} R_2^{-1} k^T(x) \nabla V_2^{i+1} \end{cases} \quad (15)$$

Step 4 如果 $\|V_j^i - V_j^{i-1}\| \leq \varepsilon$, ($j = 1, 2, \varepsilon$ 为足够小且大于零的常数) 停止计算, 输出 V_1^i, V_2^i , 否则 $i = i + 1$, 并转至 Step 2。

定理1 假设1成立且满足 Kantorovitch's 收敛条件^[12], 在式(14)和式(15)迭代相当于牛顿方法求解HJI方程式(13), 当 $i \rightarrow \infty$ 时, 有

$$\begin{cases} \lim_{i \rightarrow \infty} V_j^i = V_j^* \\ \lim_{i \rightarrow \infty} u_i = u^* \\ \lim_{i \rightarrow \infty} w_i = w^* \end{cases} \quad j = 1, 2 \quad (16)$$

式中, V_j^* 为耦合 HJI 方程的最优解。

证明 文献[11]已作详细证明, 不赘述。

算法1计算过程需要系统模型 $f(x), g(x)$ 和

$k(x)$ 信息, 为了解难以得到系统模型的困难, 消除算法对系统模型信息的依赖, 采取将探测信号 e 依次加入到控制输入 u_i 和干扰输入 w_i 中, 这里的探测信号是已知且不为零的有界可测信号, 即 $\|e\| \leq e_M, e_M$ 为有界信号的界限。

系统(1)改写为

$$\begin{cases} \dot{x} = f(x) + g(x)(u_i + e) + k(x)w_i \\ \dot{x} = f(x) + g(x)u_i + k(x)(w_i + e) \\ z = h(x) \end{cases} \quad (17)$$

值函数 V_1^i, V_2^i 对时间求导得到

$$\begin{cases} \dot{V}_1^i = (\nabla V_1^i)^T (f(x) + g(x)(u_i + e) + k(x)w_i) \\ \dot{V}_2^i = (\nabla V_2^i)^T (f(x) + g(x)u_i + k(x)(w_i + e)) \end{cases} \quad (18)$$

在区间 $[t, t + T]$ 上积分式(18), 并结合式(14)和式(15)经一系列数学变换得

$$\begin{aligned} V_1^i(x_{t+T}) - V_1^i(x_t) &= - \int_t^{t+T} r_1(x, u_i, w_i) dt - \\ & 2 \int_t^{t+T} u_{i+1}^T R_1 e dt \end{aligned} \quad (19)$$

$$\begin{aligned} V_2^i(x_{t+T}) - V_2^i(x_t) &= - \int_t^{t+T} r_2(x, u_i, w_i) dt + \\ & 2\gamma^2 \int_t^{t+T} u_{i+1}^T R_2 e dt \end{aligned} \quad (20)$$

式中, x_t, x_{t+T} 分别为 $x(t), x(t + T)$ 的简记, T 为采样时间。

式(19)、式(20)中不再含有 $f(x), g(x)$ 和 $k(x)$, 继而可以推导出如下求解混合 H_2/H_∞ 控制律的未知模型近似自适应动态(ADP)算法。

算法2 步骤如下所述。

Step 1 分别给定初始稳定的控制、干扰输入 u_1 和 w_1 , 此时 $i = 1$ 。

Step 2 依据已知 u_i, w_i 和 e 求解如下方程

$$\begin{aligned} V_1^i(x_t) &= V_1^i(x_{t+T}) + \int_t^{t+T} r_1(x, u_i, w_i) dt + \\ & 2 \int_t^{t+T} u_{i+1}^T R_1 e dt \end{aligned} \quad (21)$$

$$\begin{aligned} V_2^i(x_t) &= V_2^i(x_{t+T}) + \int_t^{t+T} r_2(x, u_i, w_i) dt - \\ & 2\gamma^2 \int_t^{t+T} u_{i+1}^T R_2 e dt \end{aligned} \quad (22)$$

Step 3 如果 $\|V_j^i - V_j^{i-1}\| \leq \varepsilon, j = 1, 2, \varepsilon > 0$, 停止计算, 输出 V_1^i, V_2^i , 否则 $i = i + 1$, 并转至 Step 2。

3 算法实现

采用如图1所示的4个3层前馈神经网络(评价-执行-评价干扰)。

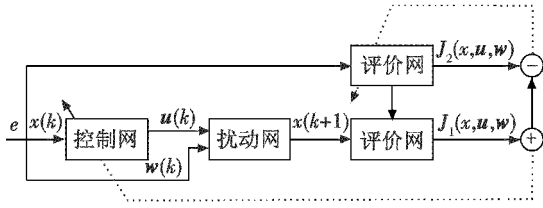


图1 神经网络结构图

Fig.1 Neural network structure

两个评价网络近似值函数 V_1, V_2 , 执行网络近似控制输入 u 和扰动网络近似扰动输入 w 。为了便于描述, 记 $m = q = 1$, 对式 (21) 和式 (22), $V_j^i(x), j = 1, 2, u_{i+1}(x), w_{i+1}(x)$ 用神经网络表示为

$$V_1^i(x) = (\mathbf{W}_i^1)^T \boldsymbol{\phi}_1(x) + \varepsilon_i^1(x) \quad (23)$$

$$V_2^i(x) = (\mathbf{W}_i^2)^T \boldsymbol{\phi}_2(x) + \varepsilon_i^2(x) \quad (24)$$

$$u_{i+1}(x) = (\mathbf{W}_{i+1}^3)^T \boldsymbol{\phi}_3(x) + \varepsilon_{i+1}^3(x) \quad (25)$$

$$w_{i+1}(x) = (\mathbf{W}_{i+1}^4)^T \boldsymbol{\phi}_4(x) + \varepsilon_{i+1}^4(x) \quad (26)$$

式中: $\mathbf{W}_i^j (j=1, 2, 3, 4)$ 是神经网络适当维数的未知理想权值矩阵; $\boldsymbol{\phi}_j (j=1, 2, 3, 4)$ 是激活函数; $\varepsilon_i^j (j=1, 2, 3, 4)$ 是神经网络的近似误差。当神经网络隐含层神经元数 $n \rightarrow \infty$ 时, 误差 $\varepsilon \rightarrow 0$ 。由于理想权值未知, 假设神经网络输出为

$$\hat{V}_1^i(x) = (\hat{\mathbf{W}}_i^1)^T \boldsymbol{\phi}_1(x) \quad (27)$$

$$\hat{V}_2^i(x) = (\hat{\mathbf{W}}_i^2)^T \boldsymbol{\phi}_2(x) \quad (28)$$

$$\hat{u}_{i+1}(x) = (\hat{\mathbf{W}}_{i+1}^3)^T \boldsymbol{\phi}_3(x) \quad (29)$$

$$\hat{w}_{i+1}(x) = (\hat{\mathbf{W}}_{i+1}^4)^T \boldsymbol{\phi}_4(x) \quad (30)$$

式中, $\hat{\mathbf{W}}_i^1, \hat{\mathbf{W}}_i^2, \hat{\mathbf{W}}_{i+1}^3, \hat{\mathbf{W}}_{i+1}^4$ 是当前的估计值。

神经网络的参数用 LS 法进行估计, 由式 (27) ~ 式 (30) 可将式 (21)、式 (22) 分别写为

$$(\psi_k^1)^T \begin{bmatrix} \hat{\mathbf{W}}_i^1 \\ \hat{\mathbf{W}}_{i+1}^3 \end{bmatrix} = \theta_k^1 \quad (31)$$

$$(\psi_k^2)^T \begin{bmatrix} \hat{\mathbf{W}}_i^2 \\ \hat{\mathbf{W}}_{i+1}^4 \end{bmatrix} = \theta_k^2 \quad (32)$$

式中, $\theta_k^j = \int_{t+(k-1)T}^{t+kT} r_j(x, u_i, w_i) dt, \psi_k^1 = [(\boldsymbol{\phi}_1(x_{t+(k-1)T}) - \boldsymbol{\phi}_1(x_{t+kT}))^T, -2 \int_{t+(k-1)T}^{t+kT} \mathbf{R}_1 e \boldsymbol{\phi}_3^T(x) dt], \psi_k^2 = [(\boldsymbol{\phi}_2(x_{t+(k-1)T}) - \boldsymbol{\phi}_2(x_{t+kT}))^T, 2\gamma^2 \int_{t+(k-1)T}^{t+kT} \mathbf{R}_2 e \boldsymbol{\phi}_4^T(x) dt], j = 1, 2$ 。

注意到, 式 (31) 和式 (32) 为一维方程, 其解的唯一性是不能保证的, 用 LS 法来计算参数向量, 方法如下

$$\begin{cases} (\boldsymbol{\Phi}_i^1)^T \begin{bmatrix} \hat{\mathbf{W}}_i^1 \\ \hat{\mathbf{W}}_{i+1}^3 \end{bmatrix} = \boldsymbol{\Theta}_i^1 \\ (\boldsymbol{\Phi}_i^2)^T \begin{bmatrix} \hat{\mathbf{W}}_i^2 \\ \hat{\mathbf{W}}_{i+1}^4 \end{bmatrix} = \boldsymbol{\Theta}_i^2 \end{cases} \quad (33)$$

式中: $\boldsymbol{\Phi}_i^j = [\psi_k^j, \dots, \psi_k^j], j = 1, 2; \boldsymbol{\Theta}_i^j = [\theta_k^j, \dots, \theta_k^j], j = 1, 2$ 。

假设 $\boldsymbol{\Phi}_i^j$ 列满秩, 式 (33) 变为

$$\begin{cases} \begin{bmatrix} \hat{\mathbf{W}}_i^1 \\ \hat{\mathbf{W}}_{i+1}^3 \end{bmatrix} = ((\boldsymbol{\Phi}_i^1)^T \boldsymbol{\Phi}_i^1)^{-1} (\boldsymbol{\Phi}_i^1)^T \boldsymbol{\Theta}_i^1 \\ \begin{bmatrix} \hat{\mathbf{W}}_i^2 \\ \hat{\mathbf{W}}_{i+1}^4 \end{bmatrix} = ((\boldsymbol{\Phi}_i^2)^T \boldsymbol{\Phi}_i^2)^{-1} (\boldsymbol{\Phi}_i^2)^T \boldsymbol{\Theta}_i^2 \end{cases} \quad (34)$$

LS 问题可以采集系统的实时数据求解, N_1 和 N_2 分别为权值矩阵 $\hat{\mathbf{W}}_i^1$ 和 $\hat{\mathbf{W}}_{i+1}^3$ 的维数, 为了保证解存在, 采样数据 k 要取得足够大, $k_{\min} = \text{rank}(\boldsymbol{\Phi}_i^j) = N_1 + N_2$ 。

算法3 收敛性分析。

定理2 假设1成立, 在探测信号 e 作用下, 对任意 $\varepsilon > 0$, 存在 $i^* > 0, N_1^* > 0, N_2^* > 0$, 那么如果 $i^* > i, N_2 > N_2^*$ 则

$$\begin{cases} |\hat{V}_i^j(x) - V_i^j(x)| \leq \varepsilon, \|\hat{u}_{i+1} - u_{i+1}\| \leq \varepsilon, \\ \|\hat{w}_{i+1} - w_{i+1}\| \leq \varepsilon \\ |\hat{V}_i^j(x) - V_i^{j*}(x)| \leq \varepsilon, \|\hat{u}_{i+1} - u^*\| \leq \varepsilon, \\ \|\hat{w}_{i+1} - w^*\| \leq \varepsilon \end{cases} \quad (35)$$

对所有 $x \in \Omega$ 成立, 类似的结论证明详见文献 [13]。

4 算例仿真

本节通过 Matlab 软件仿真结果来说明算法2的有效性。

考虑连续非线性系统 $\begin{cases} \dot{x} = f(x) + g(x)u + k(x)w \\ z = h(x) \end{cases}$,

其中, $f(x) = \begin{bmatrix} -0.25x_1 \\ 0.5x_2^3 - 0.25\gamma^{-2}x_2x_1^2 - 0.5x_2 \end{bmatrix}, g(x) = \begin{bmatrix} 0 \\ x_2 \end{bmatrix}, k(x) = \begin{bmatrix} 0 \\ x_1 \end{bmatrix}, h(x) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 。

定义性能指标式 (4), 式 (5) 中 $\mathbf{R}_1 = 2\mathbf{I}, \mathbf{R}_2 = \mathbf{I}, \gamma = 1$, 这里 \mathbf{I} 是单位矩阵。

初始状态设置为 $x_0 = [1 \ -1]^T$, 初始容许控制输入 $u_1 = 0, w_1 = 0$, 初始权值 \mathbf{W} 参数均设为 0, 激活函数为 $\boldsymbol{\phi}_1(x) = \boldsymbol{\phi}_2(x) = \boldsymbol{\phi}_3(x) = \boldsymbol{\phi}_4(x) = [x_1^2 \ x_1x_2 \ x_2^2]^T$, 选用 $e = \sin t$ 为已知噪声。算法在线运行, 采集系统响应数据的周期为 $T = 0.05$ s, LS 法采集数据为 20 组, 则激活函数参数每 1 s 更新一次。图 2 为值函数 V_1, V_2 估计参数对更新次数的变化; 图 3 为控制输入 u , 干扰输入 w 估计参数随更新次数的变化。可以看出, 参数在 5 次更新以后基本保持稳定。图 4 为在所得的控制器和干扰器作用下的系统响应图。

由图 2 ~ 图 3 可以看出, 值函数和控制器的权值 \mathbf{W} 在参数 5 次更新后基本收敛到稳定值, 在 LS 估计的

理想权值参数下,最优值函数为 $V_1^* = 5.5396x_1^2 - 2.3891x_1x_2 - 0.1796x_2^2, V_2^* = -6.9093x_1^2 + 9.7127x_1x_2 + 2.4373x_2^2$ 。纳什均衡策略(鞍点)为 $u^* = -0.4460x_1^2 + 0.3035x_1x_2 + 0.2018x_2^2, w^* = 1.4854x_1^2 - 1.8993x_1x_2 - 1.5787x_2^2$ 。

由图4可以看出,在所设计的控制器和干扰器的共同作用下,系统状态收敛到零或稳定在极小区间内,表明所得到的控制器可以使系统一致最终有界,且具有混合 H_2/H_∞ 性能。

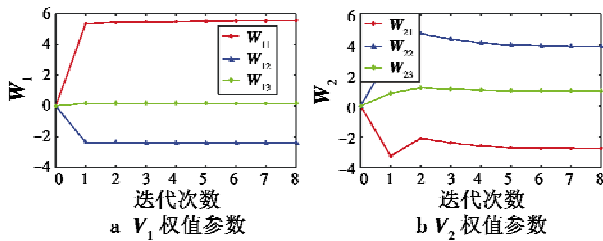


图2 V_1, V_2 权值参数更新

Fig.2 Updating of parameter V_1 and V_2

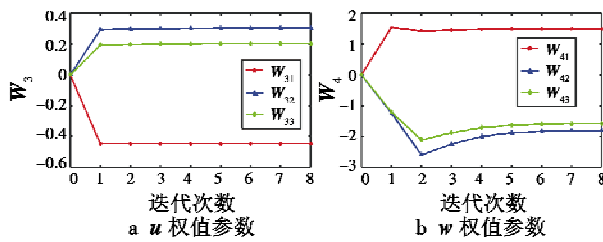


图3 u, w 权值参数更新

Fig.3 Updating of parameter u and w

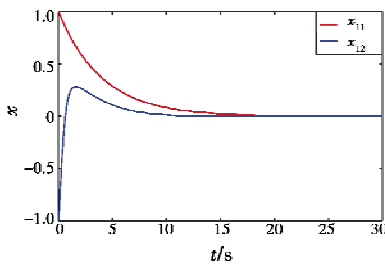


图4 系统状态响应

Fig.4 System state response

5 结束语

本文针对一类内部动力模型未知的非线性系统,基于自适应动态规划思想,提出了一种求解混合 H_2/H_∞ 最优控制问题的无模型算法,近似求解耦合的 HJI 方程,避免了非线性系统模型难以获得的难题。通过4个神经网络,同时在线更新值函数、控制策略和干扰策略,加快了算法的收敛速度。最后算例仿真验证了算法的有效性。

参考文献

- [1] BERNSTEIN D S, HADDAD W M. LQG control with an H_∞ performance bound: a riccati equation approach[C]// American Control Conference, IEEE, 2009:796-802.
- [2] 潘伟,王学勇,井元伟. 基于遗传算法的混合 H_2/H_∞ 状态反馈控制器[J]. 控制与决策, 2005, 20(2):132-136.
- [3] 叶思隽,王新民,张清江,等. 不确定系统混合 H_2/H_∞ 鲁棒控制的直接迭代 LMI 方法[J]. 控制理论与应用, 2011, 28(2):247-255.
- [4] 马清亮,杨海燕,吴旭光. 多项式模糊系统混合 H_2/H_∞ 控制[J]. 电光与控制, 2017, 24(7):1-6.
- [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] 张化光,张欣,罗艳红,等. 自适应动态规划综述[J]. 自动化学报, 2013, 39(4):303-311.
- [7] LIU D R, YANG X, WANG D, et al. Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints [J]. IEEE Transactions on Cybernetics, 2015, 45(7):1372-1385.
- [8] ZHANG H G, QIN C B, LUO Y H. Neural-network-based constrained optimal control scheme for discrete-time switched nonlinear system using dual heuristic programming [J]. IEEE Transactions on Automation Science and Engineering, 2014, 11(3):839-849.
- [9] VAMVOUDAKIS K G, LEWIS F L. Multi-player non-zero-sum games: online adaptive learning solution of coupled Hamilton-Jacobi-Equations [J]. Automatica, 2011, 47(8):1556-1569.
- [10] ALIYU M D S. An iterative computational scheme for solving the coupled Hamilton-Jacobi-Isaacs equations in nonzero-sum differential games of affine nonlinear systems [J]. Decisions in Economics & Finance, 2017(40):1-30.
- [11] ZHAO D B, XIA Z P, WANG D. Model-free optimal control for affine nonlinear systems with convergence analysis [J]. IEEE Transaction on Automation Science and Engineering, 2015, 12(4):1461-1468.
- [12] TAPIA R A. The Kantorovich theorem for Newton's method [J]. American Mathematical Monthly, 1971, 78(4):389-392.
- [13] ZHANG Q, ZHAO D, ZHU Y. Data-driven adaptive dynamic programming for continuous-time fully cooperative games with partially constrained inputs [J]. Neurocomputing, 2017(238):377-386.