

基于策略迭代算法的连续时间线性Markov跳变系统 非零和微分反馈Nash控制

朱国政, 张茂光, 何舒平[†]

(1. 安徽大学 电气工程与自动化学院, 安徽 合肥 230601; 2. 安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230601)

摘要: 针对一类连续时间线性Markov跳变系统, 本文提出了一种新的策略迭代算法用于求解系统的非零和微分反馈Nash控制问题. 通过求解耦合的数值迭代解, 以获得具有线性动力学特性和无限时域二次成本的双层非零和微分策略的Nash均衡解. 在每一个策略层, 采用策略迭代算法来计算与每一组给定的反馈控制策略相关联的最小无限时域值函数. 然后, 通过子系统分解将Markov跳变系统分解为 N 个并行的子系统, 并将该算法应用于跳变系统. 本文提出的策略迭代算法可以很容易求解非零和微分策略所对应的耦合代数Riccati方程, 且对高维系统有效. 最后通过仿真示例证明了本文设计方法的有效性和可行性.

关键词: 策略迭代; Markov跳变线性系统; 非零和; 微分反馈Nash策略

引用格式: 朱国政, 张茂光, 何舒平. 基于策略迭代算法的连续时间线性Markov跳变系统非零和微分反馈Nash控制. 控制理论与应用, 2020, 37(8): 1749 – 1756

DOI: 10.7641/CTA.2020.90603

Policy iteration-based non-zero sum differential feedback Nash control for continuous-time Markov jump linear systems

ZHU Guo-zheng, ZHANG Mao-guang, HE Shu-ping[†]

(1. School of Electrical Engineering and Automation, Anhui University, Hefei Anhui 230601, China;
2. Key Laboratory of Intelligent Computing & Signal Processing of Ministry of Education, Anhui University,
Hefei Anhui 230601, China)

Abstract: In this paper, a new policy iterative algorithm is proposed to solve the non-zero sum differential feedback Nash control problems for a class of continuous-time Markov jump linear systems. The Nash equilibrium solution of a double-layer non-zero sum differential policy with linear dynamics and infinite time-domain secondary cost is found by solving the coupled numerical iteration solutions. At each policy layer, a policy iterative algorithm is used to calculate the minimum infinite time-domain value function associated with the set of given feedback control strategies. Then, Markov jump linear systems is decomposed into N parallel subsystems by subsystems transformation. And the algorithm is applied to jump systems. The policy iteration algorithm proposed in this paper can easily solve the coupled algebraic Riccati equations corresponding to the non-zero and differential policy. It is effective for high-dimensional systems. Finally, a simulation example is given to prove the effectiveness and feasibility of the design method.

Key words: policy iteration; Markov jump linear systems; non-zero sum; differential feedback Nash strategy

Citation: ZHU Guozheng, ZHANG Maoguang, HE Shuping. Policy iteration-based non-zero sum differential feedback Nash control for continuous-time Markov jump linear systems. *Control Theory & Applications*, 2020, 37(8): 1749 – 1756

1 引言

Markov 跳变系统是 1961 年由 Krasovskii 和 Lidskii^[1]提出的一类多模态随机系统. 在Markov跳变系统中, 它包含模态和状态两个动态特性. 作为一种多模态的系统, 该系统可用来模拟具有突变特性的动

态系统, 如制造系统、网络控制系统以及容错控制系统等. 在实际过程中, 由于零部件失效、环境干扰以及子系统的连接方式改变等因素, Markov跳变系统也能用于模拟很多结构和参数具有突变效应的动态过程.

近几十年来, Markov跳变系统引起了学者们的研

收稿日期: 2019-07-23; 录用日期: 2020-03-28.

[†]通信作者. E-mail: shuping.he@ahu.edu.cn.

本文责任编辑: 方海涛.

国家自然科学基金项目(61673001), 安徽省杰出青年基金项目(1608085J05), 安徽省高校优秀青年人才支持重点项目(gxydZD2017001)资助.

Supported by the National Natural Science Foundation of China (61673001), the Foundation for Distinguished Young Scholars of Anhui Province (1608085J05) and the Key Support Program of University Outstanding Youth Talent of Anhui Province (gxydZD2017001).

究兴趣并取得了很多研究成果. Sworder^[2]研究了连续时间线性随机Markov跳变系统的二次型优化控制问题. 之后, Blair等人^[3]在连续时间Markov跳变系统的基础上提出了离散时间线性Markov跳变系统的最优控制方法. 基于动态规划理论, Wonham^[4]提出了另一个解决连续时间线性随机Markov跳变系统二次型优化控制问题的方法. Zhang等人^[5]研究了一类具有部分未知转移概率的连续时间和离散时间线性Markov跳变系统的稳定性问题. Park等人^[6-7]研究了离散时间Markov跳变系统的预测控制问题. 当然还有很多其他关于Markov跳变系统的研究成果, 如H₂控制器设计^[8-10]、H_∞控制器设计^[11-13]、鲁棒控制和滤波问题^[14-16]、故障诊断问题^[17-18]等.

在控制领域, 优化控制问题一直是研究热点. 要求解系统的优化控制器, 最常用的解决方法是求解对应的Riccati方程. Kleinman等介绍了一种离线迭代算法求解Riccati方程^[19]. 文献[20-21]给出了相关方法的收敛性证明.

随后, 相关的Nash微分博弈的零和、非零和的优化控制问题也被应用到求解对应的Riccati方程中^[22-27]. Nash微分博弈是由文献[25]首次提出的. Nash微分博弈通常出现在网络、无线通信^[28]等较大的耦合系统研究中^[29]. 当被控系统受到多个控制输入的影响时, 就可以采用Nash博弈方法. 当系统具有线性动态特性和无限时域二次成本函数指标时, Nash均衡的求解等价于求解一组耦合的代数Riccati方程, 如文献[22, 25, 30-31]. 文献[23, 26, 32-34]等采用了策略迭代的方法求解耦合代数Riccati方程. 这些算法可以得到收敛的代价函数系列. 但是文献[26, 32-34]中的收敛结果还没有确定. 虽然在线性系统中关于Nash微分博弈的零和、非零和的优化控制问题取得了初步的研究成果, 但是如何求解跳变系统的非零和微分反馈Nash控制问题, 尚未取得较好的成果.

为了获得Markov跳变系统的非零和微分反馈Nash控制策略, 本文将相关问题转化为求解对应的耦合代数Riccati方程, 并通过不断迭代耦合代数Riccati方程来计算可行解. 本文提出的策略迭代算法过程包括: 首先对跳变系统进行子系统的转换, 在此基础上对模态之间的耦合关系解耦, 并构成了新的策略迭代算法, 最后通过数值仿真和对应的迭代更新结果证实算法的有效性和可行性.

2 系统描述

在概率空间 (Ω, F, P) 中, 考虑如下连续时间Markov跳变线性系统:

$$\begin{cases} \dot{x}(t) = A(r(t))x(t) + B_1(r(t))u_1(t) + \\ B_2(r(t))u_2(t), \\ x(0) = x_0, r(0) = r_0. \end{cases} \quad (1)$$

该系统遵循以下概率分布:

$$P\{r(t + \Delta t) = j | r(t) = i\} = \begin{cases} \pi_{ij}\Delta t + o(\Delta t), & j \neq i, \\ 1 + \pi_{ii}\Delta t + o(\Delta t), & j = i. \end{cases} \quad (2)$$

式(1)-(2)中, Ω , F 和 P 分别表示样本空间、事件场和定义在事件上的概率; $x(t) \in \mathbb{R}^n$ 是跳变系统的状态; x_0 是初始状态; $u_1(t), u_2(t) \in \mathbb{R}^m$ 是系统的控制输入; $A(r(t)), B_1(r(t)), B_2(r(t))$ 是系统的系数矩阵; $r(t)$ 是一个连续时间和离散状态的Markov随机过程, 它表示系统在某一时刻的模态, 并在一个离散集中取值 $M = \{1, 2, \dots, N\}$; r_0 表示初始模态. 本文假设式(1)是随机稳定的.

式(2)表示模态从时间 $t \rightarrow t + \Delta t$, 系统从模态 i 跳到 j 的概率, 其中 $i, j \in M$, $\Delta t > 0$, $\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0$, $\pi_{ij} \geq 0 (i \neq j)$, 并满足 $\pi_{ii} = -\sum_{i \neq j} \pi_{ij}$. 为了方便, 令 $r(t) = i$, 那么 $A(r(t)), B_1(r(t)), B_2(r(t))$ 3个系数矩阵可表示为 A_i, B_{i1}, B_{i2} .

3 新的策略迭代算法

求解跳变系统(1)的非零和微分的反馈Nash策略前, 先考虑如下连续时间线性系统:

$$\begin{cases} \dot{x}(t) = Ax(t) + B_1u_1(t) + B_2u_2(t), \\ x(0) = x_0. \end{cases} \quad (3)$$

给出如下的与控制量相关的二次成本函数:

$$J_i = \int_{t_0}^{\infty} (u_1^T R_{i1} u_1 + u_2^T R_{i2} u_2 + x^T Q_i x) d\tau, \quad (4)$$

其中 $R_{ij} > 0$, $i, j \in \{1, 2\}$ 和 Q_i 均为实对称正定的权值矩阵. 本文设定矩阵对 $(A, B_i, Q_i^{1/2})$, $i \in \{1, 2\}$ 是镇定可测的.

若每个控制器都要制定最佳策略, 那么在非零和反馈控制策略下, 可以使用具有闭环信息结构的Nash均衡, 即一对反馈控制策略来获得. 参考文献[35], 系统(3)对应的闭环系统的非零和反馈Nash均衡的最优控制策略可以通过以下方式得到:

$$\begin{cases} u_1^*(x) = -R_{11}^{-1} B_1^T P_1 x, \\ u_2^*(x) = -R_{22}^{-1} B_2^T P_2 x. \end{cases} \quad (5)$$

上式两个控制输入的优化值, 可通过正定矩阵 P_i , $i \in \{1, 2\}$ 获得. 对于式(5)中的唯一正定对称矩阵 P_i , $i \in \{1, 2\}$, 可通过以下具有耦合关系的Riccati方程求解:

$$\begin{cases} A^T P_1 + P_1 A + Q_1 - P_1 S_{11} P_1 - \\ P_2 S_{22} P_1 - P_1 S_{22} P_2 + P_2 S_{12} P_2 = 0, \\ A^T P_2 + P_2 A + Q_2 - P_2 S_{22} P_2 - \\ P_2 S_{11} P_1 - P_1 S_{11} P_2 + P_1 S_{21} P_1 = 0. \end{cases} \quad (6)$$

上式中的 S_{ij} , $i, j \in \{1, 2\}$, 可表示为如下形式:

$$\begin{cases} S_{ii} = B_i R_{ii}^{-1} B_i^T, \\ S_{ij} = B_j R_{jj}^{-1} R_{ij}^{-1} R_{jj}^{-1} B_j^T. \end{cases} \quad (7)$$

那么, 对于线性系统(3)的两层非零和反馈Nash策略的相关问题可表述为耦合Riccati方程(6)的求解. 为了求解相应的Riccati方程, 给出如下引理.

引理 1^[26] 满足耦合代数Riccati方程(6)的唯一正定实对称矩阵对 (P_1, P_2) 是一组稳定解, 如果闭环矩阵 $(A - S_{11}P_1 - S_{22}P_2)$ 是Hurwitz的, 即它的所有特征值均具有负实部.

虽然, 在 $(A, B_i, Q_i^{1/2})$, $i \in \{1, 2\}$ 镇定可测的条件下, 对于求解耦合代数Riccati方程(6)和(7)的解取得了部分成果^[22, 35-36], 如在文献[35-36]中, 分别给出了两种数值算法求解耦合代数Riccati方程(6)和(7)的可能解. 但是以上算法并没有给出方程稳定解存在性的证明, 而且算法的收敛性也没有完全建立起来.

为了获得新的算法求解耦合代数Riccati方程(6)的一对稳定解 (P_1, P_2) , 本文给出如下假设.

假设 1 为了保证求解的控制律是使得系统稳定的, 本文假设 $(A, B_1, Q_1^{1/2})$ 和 $(A, B_2, Q_2^{1/2})$ 是镇定可测的.

基于此, 本文给出计算耦合Riccati方程(6)所采用新的策略迭代算法. 通过算式变换将(6)改写为如下线性方程:

$$\begin{cases} (A - S_{11}P_1^{l-1})^T P_2^l + P_2^l (A - S_{11}P_1^{l-1}) + \\ Q_2 - P_2^l S_{22} P_2^l + P_1^{l-1} S_{21} P_1^{l-1} = 0, \\ (A - S_{22}P_2^l)^T P_1^l + P_1^l (A - S_{22}P_2^l) + \\ Q_1 - P_1^l S_{11} P_1^l + P_2^l S_{12} P_2^l = 0, \end{cases} \quad (8)$$

其中 l 表示迭代步数, 且为大于0的整数.

本文所设计的用于计算耦合Riccati方程(6)的新的策略迭代算法具体步骤如下:

Step 1 给出初始矩阵 $P_1^0 \geq 0$, 令 $l = 1$;

Step 2 定义 (P_1^l, P_2^l) . 在 $l \geq 1$ 下, 通过迭代方程(8)求解正定解 (P_1^l, P_2^l) ;

Step 3 设定精确度阈值 ε_1 和 ε_2 . 对于任意 $l \geq 1$, 通过迭代算法(8)计算 P_1^l 和 P_2^l , 直到满足 $\|P_1^{l+1} - P_1^l\| \leq \varepsilon_1$, $\|P_2^{l+1} - P_2^l\| \leq \varepsilon_2$; 否则取 $l = l + 1$, 继续执行Step 2;

Step 4 获取非零和微分反馈Nash策略的正定解 P_1^l 和 P_2^l .

引理 2^[37] 考虑如下代数Lyapunov方程:

$$A^T X + X A + Q = 0, \quad (9)$$

其中: A 为一般矩阵, $Q \in \mathbb{R}^{n \times n}$ 是正定对称矩阵, 即 $Q = Q^T \geq 0$. 如果 A 是Hurwitz矩阵且 $Q \geq 0$, 则方程

(9)的唯一解 X 是正定矩阵.

引理 3^[26] 如果 $(A, B_1, Q_1^{1/2})$ 和 $(A, B_2, Q_2^{1/2})$ 是镇定可测的, 那么由方程(6)得到的矩阵对 (P_1, P_2) 是唯一存在且正定的对称矩阵对.

如下定理(1)将给出在该算法中解得的矩阵序列 (P_1, P_2) 收敛的充分条件.

定理 1 对于耦合代数Riccati方程(6), 在任意 $l \geq 1$ 时, 存在如下正定矩阵:

$$\begin{cases} \Gamma_l^{yP_1} = (P_2^{l+1} - P_2^l)^T S_{22} P_1^l + P_1^l S_{22} (P_2^{l+1} - \\ P_2^l) - (P_1^l - P_1^{l+1})^T S_{11} (P_1^l - P_1^{l+1}) \geq 0, \\ \Gamma_l^{yP_2} = (P_1^l - P_1^{l-1})^T S_{11} P_2^l + P_2^l S_{11} (P_1^l - \\ P_1^{l-1}) - (P_2^l - P_2^{l+1})^T S_{22} (P_2^l - P_2^{l+1}) \geq 0, \end{cases} \quad (10)$$

那么由迭代方程(8)解得的迭代序列 (P_1, P_2) 是收敛的.

证 由引理3可知, 如果 (A, B_1) 和 (A, B_2) 都是稳定的, 则通过耦合Riccati方程(6)求出的解 (P_1, P_2) 是正定对称的. 考虑如下方程:

$$(A - S_{11}P_1^l - S_{22}P_2^l)^T P_1^l + P_1^l (A - S_{11}P_1^l - S_{22}P_2^l) + Q_1 + P_1^l S_{11} P_1^l + P_2^l S_{12} P_2^l = 0. \quad (11)$$

根据方程(10), 可对方程(11)做出如下变换:

$$(A - S_{11}P_1^{l+1} - S_{22}P_2^{l+1})^T P_1^l + P_1^l (A - S_{11}P_1^{l+1} - S_{22}P_2^{l+1}) + Q_1 + P_1^{l+1} S_{11} P_1^{l+1} + P_2^{l+1} S_{12} P_2^{l+1} + \Gamma_l^{yP_1} = 0. \quad (12)$$

在迭代步数为 $l + 1$ 时, 方程(11)可表达为

$$(A - S_{11}P_1^{l+1} - S_{22}P_2^{l+1})^T P_1^{l+1} + P_1^{l+1} (A - S_{11}P_1^{l+1} - S_{22}P_2^{l+1}) + Q_1 + P_1^{l+1} S_{11} P_1^{l+1} + P_2^{l+1} S_{12} P_2^{l+1} = 0. \quad (13)$$

将方程(12)减去方程(13)可得

$$(A - S_{11}P_1^{l+1} - S_{22}P_2^{l+1})^T (P_1^l - P_1^{l+1}) + (P_1^l - P_1^{l+1}) (A - S_{11}P_1^{l+1} - S_{22}P_2^{l+1}) + \Gamma_l^{yP_1} = 0. \quad (14)$$

考虑到 $\Gamma_l^{yP_1} \geq 0$, 并对所有 $l \geq 0$, $(A - S_{11}P_1 - S_{22}P_2)$ 是Hurwitz矩阵, 通过引理1和引理2可知对于所有 $l \geq 1$ 都有 $P_1^l > P_1^{l+1}$.

以上是对于控制输入 u_1 所对应的正定解 P_1 收敛性的证明, 同理也可以对控制输入 u_2 所对应的正定解 P_2 做出相应处理:

$$(A - S_{11}P_1^{l+1} - S_{22}P_2^{l+1})^T (P_2^l - P_2^{l+1}) + (P_2^l - P_2^{l+1}) (A - S_{11}P_1^{l+1} - S_{22}P_2^{l+1}) + \Gamma_l^{yP_2} = 0. \quad (15)$$

同样, 考虑到 $\Gamma_l^{yP_2} \geq 0$, 并对所有 $l \geq 0$, $(A - S_{11}P_1 - S_{22}P_2)$ 是 Hurwitz 矩阵, 通过引理1和引理2可知对于所有 $l \geq 1$ 都有 $P_2^l > P_2^{l+1}$. 证毕.

对于式(10), 由于 $S_{11} \geq 0, S_{22} \geq 0, P_i^l$ 在迭代中不断收敛, 所以只要满足初始条件 $\Gamma_l^{yP_1} \geq 0, \Gamma_l^{yP_2} \geq 0$, 那么在迭代过程中 $\Gamma_l^{yP_1} \geq 0, \Gamma_l^{yP_2} \geq 0$ 将会一直满足.

推论 1 如果引理1和定理1的条件满足, 则由耦合方程(6)和(7)所定义的迭代序列 (P_1, P_2) 最终是收敛的.

注 1 以上是将一类连续时间线性系统的非零和微分反馈 Nash 控制问题转化为求解相应的 Riccati 方程, 给出了求解 Riccati 方程的策略迭代算法. 但是该求解方法只适用于线性系统, 不能直接应用于跳变系统. 为了将该方法应用于跳变系统需要考虑到: 首先要根据跳变系统的特性将系统进行子系统分解, 然后对每个子系统进行解耦. 进而, 利用相应的迭代算法对跳变系统进行求解. 接下去将对跳变系统的子系统分解和跳变解耦进行描述, 并给出具体的算法步骤.

4 子系统分解和跳变解耦

对于 Markov 跳变系统(1), 给出如下的与控制量相关的二次成本函数:

$$J_i = \int_{t_0}^{\infty} (u_1^T R_{i1}(r(t))u_1 + u_2^T R_{i2}(r(t))u_2 + x^T Q_i(r(t))x) dt, \quad (16)$$

其中 $R_{ij} > 0 (i, j \in \{1, 2\})$ 和 Q_i 均为实对称正定的权值矩阵. 设定矩阵对 $(A, B_i, Q_i^{l/2}), i \in \{1, 2\}$ 是镇定可测的.

若每层都需要制定最佳策略, 参考对线性系统的分析, 系统(1)对应的闭环系统的非零和反馈 Nash 均衡的最优控制策略可以通过以下方式得到:

$$\begin{cases} u_1^*(x) = -R_{11}^{-1}B_1^T P_1 x, \\ u_2^*(x) = -R_{22}^{-1}B_2^T P_2 x. \end{cases} \quad (17)$$

上式两个控制输入的优化值, 可通过正定矩阵 $P_i, i \in \{1, 2\}$ 获得. 与线性系统最优控制策略(5)不同的是上式中 R_{11}, R_{22}, P_1, P_2 都和跳变的模态相关, 它们之间存在着耦合关系. 所以应用于线性系统的求解方法不能直接应用到跳变系统.

考虑到上述原因, 本文对跳变系统进行子系统分解和跳变模态解耦^[38].

定义 1 对于跳变系统(1), 在某一时刻 t , 可将其分解成多个子系统, 那么第 i 个系统可写成如下的形式:

$$\begin{cases} \dot{x}_i(t) = (A_i + \frac{\pi_{ii}}{2}I)x_i(t) + B_{i1}u_{i1}(t) + B_{i2}u_{i2}(t), \\ x_i(0) = x_0. \end{cases} \quad (18)$$

显然, 式(18)中的各个子系统存在耦合关系. 针对第 i 个模态, 其对应的耦合代数 Riccati 方程可表示如下:

$$\begin{cases} A_i^T P_{i1} + P_{i1} A_i + Q_{i1} - P_{i1} S_{11} P_{i1} - P_{i2} S_{22} P_{i1} - P_{i1} S_{22} P_{i2} + P_{i2} S_{12} P_{i2} = 0, \\ A_i^T P_{i2} + P_{i2} A_i + Q_{i2} - P_{i2} S_{22} P_{i2} - P_{i2} S_{11} P_{i1} - P_{i1} S_{11} P_{i2} + P_{i1} S_{21} P_{i1} = 0. \end{cases} \quad (19)$$

方程(19)可改写成如下形式:

$$\begin{cases} (A_i - S_{11} P_{i1}^{l-1})^T P_{i2}^l + P_{i2}^l (A_i - S_{11} P_{i1}^{l-1}) + Q_{i2}^l - P_{i2}^l S_{22} P_{i2}^l + P_{i1}^{l-1} S_{21} P_{i1}^{l-1} = 0, \\ (A_i - S_{22} P_{i2}^l)^T P_{i1}^l + P_{i1}^l (A_i - S_{22} P_{i2}^l) + Q_{i1}^l - P_{i1}^l S_{11} P_{i1}^l + P_{i2}^l S_{12} P_{i2}^l = 0, \end{cases} \quad (20)$$

其中 i 表示对应的模态, 且在一个离散集中取值 $M = \{1, 2, \dots, N\}$, 并有 $A_i = A_i + \frac{\pi_{ii}}{2}I$. 跳变系统的各个模态之间的解耦关系可表述为

$$\begin{cases} Q_{i2}^l = Q_{i2} + \sum_{j=1, j \neq i}^N \pi_{ij} P_{j2}^l, \\ Q_{i1}^l = Q_{i1} + \sum_{j=1, j \neq i}^N \pi_{ij} P_{j1}^l. \end{cases} \quad (21)$$

注 2 对于 Markov 跳变系统的每一个模态, 需要获得对应的稳定解 (P_{i1}, P_{i2}) . 在求解相应的 Riccati 方程时不仅存在同一模态下双控制器之间的耦合关系, 也会存在跳变系统各模态间的耦合关系. 对于这种双耦合关系, 在处理过程中为了使得算法简单, 本文假定每一模态下具有相同的 S_{ij} , 其中 $i, j \in M = \{1, 2, \dots, N\}$.

在线性系统迭代算法的基础上, 本文所设计的 Markov 跳变系统策略迭代算法和解耦过程的具体步骤如下:

Step 1 给出初始矩阵对 (P_{i1}^0, P_{i2}^0) , 令 $l = 1$;

Step 2 定义 (P_{i1}^l, P_{i2}^l) . 在任意 $l \geq 1$ 下, 通过策略迭代方程(20)和策略更新方程(21)求解对应的正定解 (P_{i1}^l, P_{i2}^l) ;

Step 3 设定精确度阈值 ε_1 和 ε_2 . 对于任意 $l \geq 1$, 通过迭代算法(20)和(21)计算 P_{i1}^l 和 P_{i2}^l , 直到满足 $\|P_{i1}^{l+1} - P_{i1}^l\| \leq \varepsilon_1, \|P_{i2}^{l+1} - P_{i2}^l\| \leq \varepsilon_2$; 否则取 $l = l + 1$, 继续执行 Step 2;

Step 4 获取 Markov 跳变系统非零和微分反馈 Nash 策略正定解 P_{i1}^l 和 P_{i2}^l .

上文所设计算法解得的矩阵序列 (P_{i1}^l, P_{i2}^l) 收敛的充分条件可由如下定理2给出.

定理 2 对于耦合代数 Riccati 方程(19), 在任意 $l \geq 1$ 时, 如果存在如下正定矩阵:

$$\begin{cases} \Gamma_l^{yP_{i1}} = \\ (P_{i2}^{l+1} - P_{i2}^l)^T S_{22} P_{i1}^l + P_{i1}^l S_{22} (P_{i2}^{l+1} - P_{i2}^l) - \\ (P_{i1}^l - P_{i1}^{l+1})^T S_{11} (P_{i1}^l - P_{i1}^{l+1}) \geq 0, \\ \Gamma_l^{yP_{i2}} = \\ (P_{i1}^l - P_{i1}^{l-1})^T S_{11} P_{i2}^l + P_{i2}^l S_{11} (P_{i1}^l - P_{i1}^{l-1}) - \\ (P_{i2}^l - P_{i2}^{l+1})^T S_{22} (P_{i2}^l - P_{i2}^{l+1}) \geq 0, \end{cases} \quad (22)$$

那么由迭代方程(20)和(21)解得的迭代序列 (P_{i1}^l, P_{i2}^l) 是收敛的.

由于定理2和定理1的证明相似, 在此不再详细证明.

推论 2 如果引理1和定理2的条件满足, 则由耦合方程(19)所定义的迭代序列 (P_{i1}, P_{i2}) 最终是收敛的.

对于Markov跳变系统的非零和微分反馈Nash策略的问题, 新的策略迭代算法在对跳变系统子系统分解和解耦合后, 将该问题转化为求解耦合代数Riccati方程(19). 在跳变系统的每一个模态都能得到一对稳定解 (P_{i1}, P_{i2}) , 即可以通过求解 N 个并行的方程(19)实现对该系统的优化控制.

在文献[35]所提出的算法中, 代数Riccati方程的解依赖于系统的系数矩阵和终值, 所以无法保证算法的收敛性. 文献[36]所提出的算法中, 给出了求解代数Riccati方程的方法, 却没有给出保证迭代解收敛的充分条件. 本文所提出的算法中, 给出了算法解得的矩阵序列的充分条件, 并且将其应用于Markov跳变系统.

接下来将给出相应的数值仿真示例来验证本文所设计算法的正确性和迭代收敛的效果.

5 算法仿真实例

5.1 实验 1

本文将通过给出如下示例来验证上述算法的有效性和收敛性. 在求解之前先给出初始数据.

对于跳变系统(1)所对应的耦合Riccati方程(19), 给出模态数 $N = 2$ 的如下四维系数矩阵:

$$A_1 = \begin{bmatrix} -2.2361 & -1.1358 & 1 & 0.6324 \\ -0.1024 & -3 & 0.3835 & 0.85 \\ 0.7112 & 11.2346 & -36.8199 & 4 \\ 1.0692 & 13.4230 & 20.1185 & -12 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -1.5326 & -1.2436 & 0.5458 & 0.7136 \\ -0.8 & -2.9346 & 0.0920 & 0.42 \\ 11.1634 & 23 & -26.4655 & -1.8347 \\ 25 & 8.3132 & -3.8714 & -31.4631 \end{bmatrix},$$

$$B_{11} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, B_{12} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

为了求解方便, 令

$$B_{21} = B_{11}, B_{22} = B_{12},$$

$$Q_{i1} = Q_{i2} = \text{eye}(4), i \in \{1, 2\}.$$

选取跳变系统的转移矩阵如下:

$$\Pi = \begin{bmatrix} -3 & 3 \\ 2.5 & -2.5 \end{bmatrix}, I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

那么子系统分解后的系数矩阵可通过如下公式获得:

$$A_1 = A_1 + \left(\frac{\Pi(1, 1)}{2}\right) \times I,$$

$$A_2 = A_2 + \left(\frac{\Pi(2, 2)}{2}\right) \times I.$$

两个模态下的 R_{ij} 的值如下:

$$R_{11} = 1, R_{12} = 0.25, R_{21} = 0.6, R_{22} = 2.$$

其他相关的数据可以通过方程(18)解得.

设定两个模态下初始的稳定值分别为 $P_{12}^0 = I_4, P_{22}^0 = I_4$, 精确度阈值分别为 $\varepsilon_1 = 10^{-9}, \varepsilon_2 = 10^{-9}$. 通过策略迭代和策略更新算法后, 获得耦合Riccati方程的解表述如下.

对于双控制器输入模型, 第1个模态下的稳定解为

$$P_{11} = \begin{bmatrix} 0.3759 & -0.0850 & 0.0156 & 0.0150 \\ -0.0850 & 0.5047 & 0.0471 & 0.0706 \\ 0.0156 & 0.0471 & 0.0257 & 0.0216 \\ 0.0150 & 0.0706 & 0.0216 & 0.0481 \end{bmatrix},$$

$$P_{12} = \begin{bmatrix} 0.3198 & -0.0880 & 0.0117 & 0.0105 \\ -0.0880 & 0.44847 & 0.04241 & 0.0640 \\ 0.0117 & 0.04241 & 0.0249 & 0.0206 \\ 0.0105 & 0.0640 & 0.0206 & 0.0466 \end{bmatrix}.$$

应用同样的迭代算法, 计算出第2个模态下的稳定解为

$$P_{21} = \begin{bmatrix} 0.6408 & -0.1305 & 0.0135 & 0.0221 \\ -0.1305 & 0.3116 & 0.0118 & 0.0023 \\ 0.0135 & 0.0118 & 0.0193 & -0.0013 \\ 0.0221 & 0.0023 & -0.0013 & 0.0164 \end{bmatrix},$$

$$P_{22} = \begin{bmatrix} 0.6114 & -0.1358 & 0.0127 & 0.0212 \\ -0.1358 & 0.3098 & 0.0117 & 0.0021 \\ 0.0127 & 0.0117 & 0.0193 & -0.0013 \\ 0.0212 & 0.0021 & -0.0013 & 0.0164 \end{bmatrix}.$$

在以下的仿真图中, 图1—4分别绘制出了Markov跳变系统非零和微分的反馈Nash策略问题的求解更新过程. 图1绘制了模态 $i = 1$ 时求解的正定解 P_{i1} 的

更新过程, 图中的“*”体现出 P_{i1} 迭代更新的过程; 通过模值的范数可以看出 P_{i1} 最终是收敛的, 且在迭代步数 $l = 9$ 时收敛到精确值. 图2绘制了模态 $i = 1$ 时求解的正定解 P_{i2} 的更新过程, 从图中可看出在迭代步数 $l = 8$ 时 P_{i2} 收敛到精确值. 图3绘制了模态 $i = 2$ 时求解的正定解 P_{i1} 的更新过程, 从图中可看出在迭代步数 $l = 9$ 时 P_{i1} 收敛到精确值. 图4绘制了模态 $i = 2$ 时求解的正定解 P_{i2} 的更新过程, 从图中可看出在迭代步数 $l = 8$ 时 P_{i2} 收敛到精确值. 该仿真示例更进一步体现了算法的正确性和可行性.

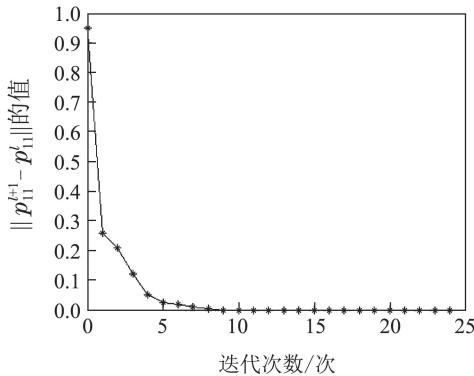


图1 $\|P_{11}^{l+1} - P_{11}^l\|$ 运行轨迹图
Fig. 1 The trajectory diagram of $\|P_{11}^{l+1} - P_{11}^l\|$

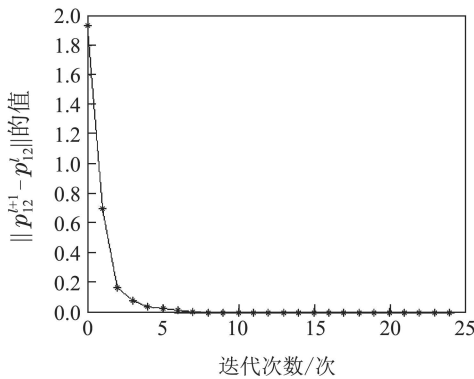


图2 $\|P_{12}^{l+1} - P_{12}^l\|$ 运行轨迹图
Fig. 2 The trajectory diagram of $\|P_{12}^{l+1} - P_{12}^l\|$

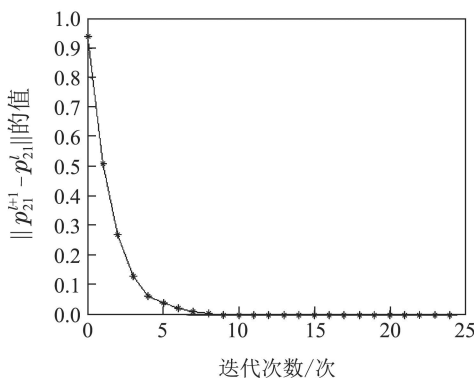


图3 $\|P_{21}^{l+1} - P_{21}^l\|$ 运行轨迹图
Fig. 3 The trajectory diagram of $\|P_{21}^{l+1} - P_{21}^l\|$

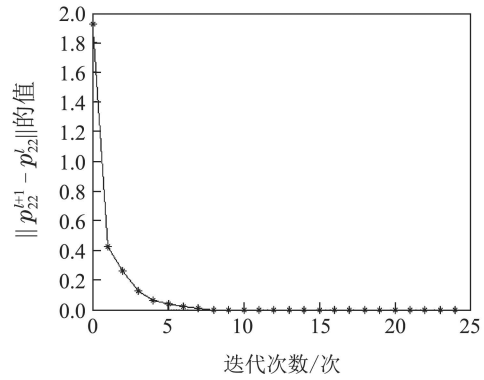


图4 $\|P_{22}^{l+1} - P_{22}^l\|$ 运行轨迹图
Fig. 4 The trajectory diagram of $\|P_{22}^{l+1} - P_{22}^l\|$

5.2 实验2

为了验证本文中所提出的算法对于非零和耦合代数Riccati方程的优越性, 笔者将其和文献[35-36]中算法1(Algorithm 1, A.1)及算法2(Algorithm 2, A.2)的两种数值算法进行了对比, 并做了如下实验:

对于耦合代数Riccati方程(5)和(6), 给出如下矩阵:

$$A = \begin{bmatrix} 0.4199 & 0.7939 \\ 0.7537 & 0.9200 \end{bmatrix},$$

$$S_{11} = \begin{bmatrix} 8.3052 & 3.9682 \\ 3.9682 & 3.7217 \end{bmatrix}, S_{22} = \begin{bmatrix} 7.4092 & 6.2590 \\ 6.2590 & 5.7807 \end{bmatrix},$$

$$Q_1 = \begin{bmatrix} 9.8905 & 6.8816 \\ 6.8816 & 6.0301 \end{bmatrix}, Q_2 = \begin{bmatrix} 3.2543 & 4.8155 \\ 4.8155 & 10.9606 \end{bmatrix},$$

$$S_{12} = \begin{bmatrix} 5.2835 & 6.4178 \\ 6.4178 & 7.8537 \end{bmatrix}, S_{21} = \begin{bmatrix} 4.4255 & 4.4501 \\ 4.4501 & 4.5081 \end{bmatrix}.$$

设初始值为 $P_1^0 = \begin{bmatrix} 8.4335 & 1.2583 \\ 1.2583 & 0.8392 \end{bmatrix}$, 精确度阈值为 $\varepsilon = 10^{-9}$. 通过策略迭代和策略更新算法后, 获得耦合Riccati方程的解如下所示:

$$P_1 = \begin{bmatrix} 0.9520 & -0.1448 \\ -0.1448 & 1.2026 \end{bmatrix},$$

$$P_2 = \begin{bmatrix} 0.2782 & 0.1805 \\ 0.1805 & 1.0388 \end{bmatrix}.$$

图5和图6分别绘制了3种算法对于求解非零和耦合代数Riccati方程的解 P_1 和 P_2 的更新过程. 其中: 菱形线表示本文中所设计算法的求解更新过程, 实线代表文献[36]中算法2的求解更新过程, 点线代表文献[35]中算法1的求解更新过程. 从图中可以看出本文所设计的算法在求解非零和耦合代数Riccati方程时可以更快地得到稳定解. 图5表示在求解 P_1 时, 本文的算法需要3次迭代就可以求得稳定解, 而A.2和A.1分别需要7次和8次迭代才能求得稳定解. 图6表示在求解 P_2 时, 本文的算法需要2次迭代就可以求得稳定解, 而

A.2和A.1分别需要4次和6次迭代才能求得稳定解. 此外, 文献[35–36]中没有对两种算法的收敛性进行证明, 而本文给出了对所设计算法中解得的矩阵序列 (P_1, P_2) 收敛的充分条件, 并且给出了证明. 从以上分析可以看出本文的算法更有优越性.

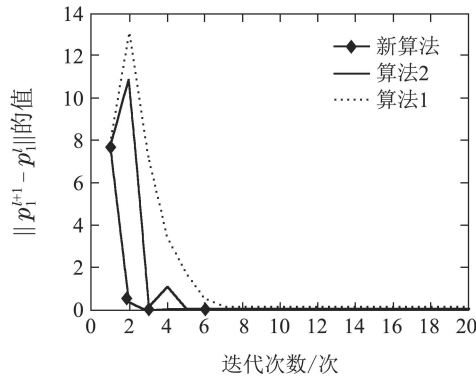


图 5 $\|P_1^{l+1} - P_1^l\|$ 运行轨迹图

Fig. 5 The trajectory diagram of $\|P_1^{l+1} - P_1^l\|$

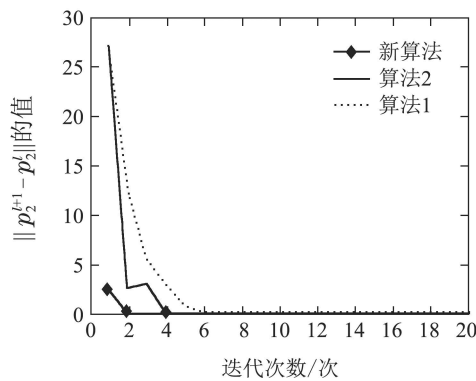


图 6 $\|P_2^{l+1} - P_2^l\|$ 运行轨迹图

Fig. 6 The trajectory diagram of $\|P_2^{l+1} - P_2^l\|$

6 结语

本文研究了一类连续时间线性Markov跳变系统的非零和微分反馈Nash控制策略, 并将相关问题转化为求解对应的耦合代数Riccati方程. 通过子系统的转换和模态之间的解耦等, 设计了一类新的策略迭代算法, 并通过数值仿真和对应的迭代更新结果证实了所设计算法的有效性和可行性.

参考文献:

- [1] KRASOVSKII N N, LIDSKII E A. Analytical design of controllers in systems with random attributes. *Automation and Remote Control*, 1961, 22(3): 1021 – 1025.
- [2] SWORDER D. Feedback control of a class of linear systems with jump parameters. *IEEE Transactions on Automatic Control*, 1969, 14(1): 9 – 14.
- [3] BLAIR J, SWORDER D. Feedback control of a class of linear discrete systems with jump parameters and quadratic cost criteria. *International Journal of Control*, 1975, 21(5): 833 – 841.
- [4] WONHAM W M. Random differential equations in control theory. *Matematika*, 1973, 17(4): 129 – 167.

- [5] ZHANG L, BOUKAS E K. Stability and stabilization of Markovian jump linear systems with partly unknown transition probabilities. *Automatica*, 2009, 45(2): 463 – 468.
- [6] PARK B G, LEE J W, KWON W H. Receding horizon control for linear discrete systems with jump parameters. *Proceedings of the 36th IEEE Conference on Decision and Control*. San Diego: IEEE, 1997, 4: 3956 – 3957.
- [7] PARK B G, KWON W H. Robust one-step receding horizon control of discrete-time Markovian jump uncertain systems. *Automatica*, 2002, 38(7): 1229 – 1235.
- [8] COSTA O L V, DO VAL J B R, GEROMEL J C. Continuous-time state-feedback H_2 -control of Markovian jump linear systems via convex analysis. *Automatica*, 1999, 35(2): 259 – 268.
- [9] CHIZECK H J, WILLSKY A S, CASTANON D. Discrete-time Markovian-jump linear quadratic optimal control. *International Journal of Control*, 1986, 43(1): 213 – 231.
- [10] JI Y, CHIZACK H J. Controllability, observability and discrete-time Markovian jump linear quadratic control. *International Journal of Control*, 1988, 48(2): 481 – 498.
- [11] BOUKAS E K. *Stochastic Switching Systems: Analysis and Design*. Boston: Springer Science & Business Media, 2007.
- [12] DE FARIAS D P, GEROMEL J C, DO VAL J B R, et al. Output feedback control of Markov jump linear systems in continuous-time. *IEEE Transactions on Automatic Control*, 2000, 45(5): 944 – 949.
- [13] KARIMI H R. A sliding mode approach to H_∞ synchronization of master-slave time-delay systems with Markovian jumping parameters and nonlinear uncertainties. *Journal of the Franklin Institute*, 2012, 349(4): 1480 – 1496.
- [14] SHI P, LI F, WU L, et al. Neural network-based passive filtering for delayed neutral-type semi-markovian jump systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(9): 2101 – 2114.
- [15] MAO Z, JIANG B, SHI P. H_∞ fault detection filter design for networked control systems modelled by discrete Markovian jump systems. *IET Control Theory & Applications*, 2007, 1(5): 1336 – 1343.
- [16] COSTA O, MARQUES R. Robust H_2 -control for discrete-time Markovian jump linear systems. *International Journal of Control*, 2000, 73(1): 11 – 21.
- [17] LIU M, SHI P, ZHANG L, et al. Fault-tolerant control for nonlinear Markovian jump systems via proportional and derivative sliding mode observer technique. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2011, 58(11): 2755 – 2764.
- [18] NGUANG S K, SHI P. Robust H_∞ output feedback control design for Takagi-Sugeno systems with Markovian jumps: a linear matrix inequality approach. *Journal of Dynamic Systems, Measurement, and Control*, 2006, 128(3): 617 – 625.
- [19] KLEINMAN D. On an iterative technique for Riccati equation computations. *IEEE Transactions on Automatic Control*, 1968, 13(1): 114 – 115.
- [20] RALL L B. A note on the convergence of Newton's method. *Society for Industrial and Applied Mathematics Journal on Numerical Analysis*, 1974, 11(1): 34 – 36.
- [21] DAMM T, HINRICHSSEN D. Newton's method for a rational matrix equation occurring in stochastic control. *Linear Algebra and Its Applications*, 2001, 332: 81 – 109.
- [22] ABOU-KANDIL H, FREILING G, JANK G. Necessary conditions for constant solutions of coupled Riccati equations in Nash games. *Systems & Control Letters*, 1993, 21(4): 295 – 306.
- [23] LI T, GAJIC Z. *Lyapunov iterations for solving coupled algebraic Riccati equations of Nash differential games and algebraic Riccati equations of zero-sum games*. Boston: Birkhauser, 1995: 333 – 351.

- [24] MUKAIDANI H. Optimal numerical strategy for Nash games of weakly coupled large scale systems. *Dynamics of Continuous Discrete and Impulsive Systems Series B*, 2006, 13(2): 249 – 268.
- [25] STARR A, HO Y. Nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 1969, 3(3): 184 – 206.
- [26] CHERFI L, CHITOUR Y, ABOU-KANDIL H. A new algorithm for solving coupled algebraic Riccati equations. *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce*. Vienna: IEEE, 2005, 1: 83 – 88.
- [27] VRABIE D, LEWIS F. Integral reinforcement learning for online computation of feedback nash strategies of nonzero-sum differential games. *Proceedings of the 49th IEEE Conference on Decision and Control*. Atlanta: IEEE, 2010: 3066 – 3071.
- [28] SHAH V M. *Power control for wireless data services based on utility and pricing*. New Brunswick, America: Rutgers University, 1998.
- [29] MUKAIDANI H. Newton's method for solving cross-coupled signindefinite algebraic Riccati equations for weakly coupled large-scale systems. *Applied Mathematics and Computation*, 2007, 188(1): 103 – 115.
- [30] BASAR T, OLSDER G. *Dynamic noncooperative game theory*. New York, America: Society for Industrial and Applied Mathematics, 1999.
- [31] ENGWERDA J. *LQ dynamic optimization and differential games*. Tilburg, Netherland: John Wiley & Sons, 2005.
- [32] CHERFI L, ABOU-KANDIL H, BOURLES H. Iterative method for general algebraic Riccati equation. *Proceedings of International Conference on Automatic Control and System Engineering*. Cairo, Egypt: 2005: 85 – 88.
- [33] FREILING G, JANK G, ABOU-KANDIL H. On global existence of solutions to coupled matrix Riccati equations in closed-loop Nash games. *IEEE Transactions on Automatic Control*, 1996, 41(2): 264 – 269.
- [34] JUNGERS M, DE PIERI E, ABOU-KANDIL H. Solving coupled algebraic Riccati equations from closed-loop Nash strategy, by lack of trust approach. *International Journal of Tomography & Statistics*, 2007, 7: 49 – 55.
- [35] GAJIC Z, SHEN X. *Parallel Algorithms for Optimal Control of Large Scale Linear Systems*. Berlin, Germany: Springer Science & Business Media, 1993.
- [36] ABOU-KANDIL H, FREILING G, IONESCU V, et al. *Matrix Riccati Equations in Control and Systems Theory*. Berlin, Germany: Birkhauser Verlag, 2003.
- [37] ABOU-KANDIL H, FREILING G, IONESCU V. Matrix Riccati equations in control and systems theory. *IEEE Transactions on Automatic Control*, 2003, 49(11): 2094 – 2095.
- [38] HE S, SONG J, LIU F, et al. Data-driven policy iteration algorithm for optimal control of continuous-time Ito stochastic system with Markovian jumps. *IET Control Theory and Applications*, 2016, 10(12): 1431 – 1439.

作者简介:

朱国政 硕士研究生, 主要研究方向为跳变系统的自适应优化控制, E-mail: 2572165091@qq.com;

张茂光 硕士, 主要研究方向为跳变系统的自适应优化控制, E-mail: 1057200891@qq.com;

何舒平 教授, 博士, 主要研究方向为随机系统控制、优化控制、强化学习与智能信息处理, E-mail: shuping.he@ahu.edu.cn.