Front. Electr. Electron. Eng. China 2009, 4(1): 55–59
DOI 10.1007/s11460-008-0076-5

RESEARCH ARTICLE

Yankai XU, Xi CHEN

Optimization of Markov jump linear system with controlled modes jump probabilities

Abstract The optimal control of a Markov jump linear quadratic model with controlled jump probabilities of modes is investigated. Two kinds of mode control policies, i.e., open-loop control policy and closed-loop control policy, are considered. Using the concepts of policy iteration and performance potential, the sufficient condition needed for the optimal closed-loop control policy to perform better than the optimal open-loop control policy is proposed. The condition is helpful for the design of an optimal controller. Furthermore, an efficient algorithm to construct a closed-loop control policy, which is better than the optimal open-loop control policy, is given with policy iteration.

Keywords Markov jump system, optimal control, policy iteration

1 Introduction

In recent years, switching systems have received great attention because of their potential applications in engineering systems 11. The Markov jump linear system (MJLS) is a class of switching systems that has been fully studied. MJLS is widely applied in systems with abrupt changes in operating points or disturbances [2,3], including flexible manufacturing systems, power systems, economic systems, fault-tolerant systems and inventory systems [4–10].

This paper considers the discrete-time jump linear quadratic Gaussian (JLQG) model. The mode jump of a standard Markov jump system is governed by a Markov chain, whose transition probability matrix is given a priori. In practice, the jump of modes is random, but the jump probabilities can often be controlled (or selected from a finite set). For example, the probability of the jump from normal mode to fault mode of a machine depends on daily maintenance frequency; and the switch between received mode and lost mode of data packages in networked control systems depends on the strength of the communication signal. However, problems with controlled jump probabilities of modes are rarely studied [11–13]. This paper discusses two classes of control policies for modes jump probabilities: open-loop mode control and closed-loop mode control. The relation between the two mode control policies and the standard JLQG model is analyzed. Generally, optimization of the open-loop mode control is an easier problem in a smaller policy space compared to optimization of the closed-loop mode control. By using the performance potential concept and policy iteration approach, this paper presents the sufficient condition under which the optimal closed-loop mode control is better than the optimal open-loop mode control. Based on this condition, we can easily construct a closed-loop mode control policy, which has better performance than the optimal open-loop mode control.

2 Formulation

The Markov jump system is a class of hybrid systems with two kinds of dynamics, i.e., the mode described by a discrete Markov chain, and the state described by the state-space equation under a mode. Consider the following discrete-time linear system

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}(\theta_k)\boldsymbol{x}_k + \boldsymbol{B}(\theta_k)\boldsymbol{u}_k + \boldsymbol{\sigma}(\theta_k)w_k, \tag{1}$$

where k is the discrete time epoch, $\boldsymbol{x}_k \in R^n$ is the state, $\theta_k \in \mathcal{S} = \{1,2,\ldots,S\}$ is the mode, $\boldsymbol{u}_k \in R^m$ is the control variable, and w_k is an i.i.d. random variable with mean zero and covariance matrix $\boldsymbol{I}$ (identity matrix). $\boldsymbol{A}(\theta_k)$, $\boldsymbol{B}(\theta_k)$, $\boldsymbol{\sigma}(\theta_k)$ are suitably-dimensional matrices depending on the mode. The

Translated from *Control and Decision*, 2008, 23(3): 246–250 [译自: 控制与决策]

Yankai XU (✉), Xi CHEN
Tsinghua National Laboratory for Information Science and Technology, Center for Intelligent and Networked Systems, Department of Automation, Tsinghua University, Beijing 100084, China
E-mail: xuyankai99@mails.thu.edu.cn

mode θ_k satisfies the Markov chain whose jump probability matrix is $\boldsymbol{p} = \{p_{ij}\}_{i,j\in\mathcal{S}}$. Assume that the Markov chain is ergodic, and it has a steady-state probability row vector $\boldsymbol{\pi} = [\pi_1,\pi_2,...,\pi_S]$. The criterion to be optimized is

$$J(\boldsymbol{x}_0,\theta_0) = \lim_{K\to\infty}\frac{1}{K}E\left\{\sum_{k=0}^{K-1}[f^u(\theta_k,\boldsymbol{x}_k)]\middle|\boldsymbol{x}_0,\theta_0\right\}, \quad (2)$$

where the cost function is

$$f^u(\theta_k,\boldsymbol{x}_k) = \boldsymbol{x}_k{}^{\mathrm{T}}\boldsymbol{M}(\theta_k)\boldsymbol{x}_k + \boldsymbol{u}_k{}^{\mathrm{T}}\boldsymbol{N}(\theta_k)\boldsymbol{u}_k.$$

Suppose that Eq. (1) is stochastically stabilizable. For a stable system, performance criterion Eq. (2) exists and is independent of the initial values [5]. For simplicity, when $\theta_k = i$, the matrices $\boldsymbol{A}(\theta_k)$, $\boldsymbol{B}(\theta_k)$, $\boldsymbol{\sigma}(\theta_k)$, $\boldsymbol{M}(\theta_k)$ and $\boldsymbol{N}(\theta_k)$ are written as $\boldsymbol{A}_i$, $\boldsymbol{B}_i$, $\boldsymbol{\sigma}_i$, $\boldsymbol{M}_i$ and $\boldsymbol{N}_i$, respectively. Let $\boldsymbol{a}_i = \boldsymbol{\sigma}_i\boldsymbol{\sigma}_i^{\mathrm{T}}$. Assume that $\boldsymbol{M}_i$ and $\boldsymbol{N}_i$ are positive semi-definite matrices.

If the jump probabilities $\{p_{ij},\forall j\in\mathcal{S}\}$ from mode i to other modes are not given a priori, but can be chosen from a finite set Y_i, it can be concluded that the jump probabilities $\boldsymbol{p} = \{p_{ij}\}_{i,j\in\mathcal{S}}$ of the modes are controllable. This paper studies the optimization problem of JLQG with controlled mode jump probabilities. At this point, the admissible control policy $\mathcal{L}$ is a combination of the feedback control law u and the mode control policy, denoted as $\mathcal{L} = \{u,\boldsymbol{p}\}$. Consider two classes of mode control polices: For the first class, the jump probability is independent of the current system state $\boldsymbol{x}$, i.e., the jump probability takes the same value for any state $\boldsymbol{x}$, thus we call it an open-loop mode control; For the second class, the jump probability depends on the current system state $\boldsymbol{x}$, i.e., the jump probability takes a different value for a different state $\boldsymbol{x}$, thus we call it a closed-loop mode control, and the corresponding jump probability is denoted as $p_{ij}(\boldsymbol{x})$. Consider the following optimization problems:

Problem 1 Find a state feedback control law $u(i,\boldsymbol{x})$ and a closed-loop mode control $p_{ij}(\boldsymbol{x})$ to minimize performance Eq. (2). Denote the admissible control policy of Problem 1 as $\mathcal{L}_1=\{u(i,\boldsymbol{x}),p_{ij}(\boldsymbol{x})\}$, where the optimal policy is denoted as $\mathcal{L}_1^*$.

Problem 2 Find a state feedback control law $u(i,\boldsymbol{x})$ and an open-loop mode control p_{ij} to minimize performance Eq. (2). Denote the admissible control policy of Problem 2 as $\mathcal{L}_2=\{u(i,\boldsymbol{x}),p_{ij}\}$, where the optimal policy is denoted as $\mathcal{L}_2^*$.

Problem 3 Given a mode jump probability matrix $\boldsymbol{p}$, find a state feedback control law $u(i,\boldsymbol{x})$ to minimize performance Eq. (2). Denote the admissible control policy of Problem 3 as $\mathcal{L}_3(\boldsymbol{p}) = \{u(i,\boldsymbol{x})\}$, where the optimal policy is denoted as $\mathcal{L}_3^*(\boldsymbol{p})$. Problem 3 is a typical JLQG problem. The following lemma gives its optimal solution.

Lemma 1 [5] Given a mode's jump probability $\boldsymbol{p} = \{p_{ij}\}_{i,j\in\mathcal{S}}$, the optimal feedback control law is

$$u^*(i,\boldsymbol{x}) = -\boldsymbol{L}_i\boldsymbol{x},$$

$$\boldsymbol{L}_i = \left[\boldsymbol{N}_i + \boldsymbol{B}_i^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{B}_i\right]^{-1}\boldsymbol{B}_i^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{A}_i,$$

where $\boldsymbol{F}_i = \sum_{j\in\mathcal{S}}p_{ij}\boldsymbol{K}_j$, $\boldsymbol{K}_i$ is the unique solution to the coupled Riccati equation:

$$\boldsymbol{K}_i = \boldsymbol{A}_i^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{A}_i + \boldsymbol{M}_i - \boldsymbol{A}_i^{\mathrm{T}}\boldsymbol{F}_i\boldsymbol{B}_i\boldsymbol{L}_i.$$

Then the optimal performance is

$$J^{\mathcal{L}_3^*(\boldsymbol{p})} = \sum_{i\in\mathcal{S}}\pi_i\mathrm{tr}(\boldsymbol{a}_i\boldsymbol{F}_i).$$

For Problem 3, by applying the results in Ref. [5] and the approach in Ref. [13], it is easy to verify that the value function (also called performance potential [14]) of the optimal feedback control law $u^*(i,\boldsymbol{x})$ is

$$g^{\mathcal{L}_3^*(\boldsymbol{p})}(i,\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{K}_i\boldsymbol{x} + q_i, \quad (3)$$

where q_i is the solution to the following equation:

$$(\boldsymbol{I}-\boldsymbol{p})\boldsymbol{q} + J^{\mathcal{L}_3^*(\boldsymbol{p})}\boldsymbol{e} - \tilde{\boldsymbol{f}} = 0, \quad (4)$$

where $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{q}=[q_1,q_2,...,q_S]^{\mathrm{T}}$, $\boldsymbol{e} = [1,1,...,1]^{\mathrm{T}}$, $\tilde{\boldsymbol{f}} = \left[\tilde{f}_1,\tilde{f}_2,\dots,\tilde{f}_S\right]^{\mathrm{T}}$ and $\tilde{f}_i = \mathrm{tr}(\boldsymbol{a}_i\boldsymbol{F}_i)$. The solution of Eq. (4) is not unique, i.e., if $q_i,\forall i\in\mathcal{S}$ is a solution, then for any constant c, $q_i+c,\forall i\in\mathcal{S}$ is also a solution to Eq. (4). Reference [14] gives a special solution in the following form:

$$\boldsymbol{q} = (\boldsymbol{I}-\boldsymbol{p}+\boldsymbol{e\pi})^{-1}\tilde{\boldsymbol{f}}. \quad (5)$$

If p_{ij} is controlled, $\{\mathcal{L}_3^*(\boldsymbol{p}),p_{ij}\}$ is an admissible control policy of Problem 2. Let

$$\boldsymbol{p}^* = \underset{\boldsymbol{p}}{\mathrm{argmin}}\left\{J^{\{\mathcal{L}_3^*(\boldsymbol{p}),p_{ij}\}}\right\},$$

then $\mathcal{L}_2^* = \{\mathcal{L}_3^*(\boldsymbol{p}^*),p_{ij}^*\}$ is the optimal admissible control policy of Problem 2, and the performance of $\mathcal{L}_3^*(\boldsymbol{p}^*)$ is the same as that of $\mathcal{L}_2^*$, i.e., $J^{\mathcal{L}_2^*} = J^{\mathcal{L}_3^*(\boldsymbol{p}^*)}$. By solving Problem 3, we can find the optimal open-loop mode control in the set $\{Y_i,\forall i\in\mathcal{S}\}$, and thus obtain the optimal policy $\mathcal{L}_2^*$ of Problem 2. Another way to address Problem 2 is to apply the gradient-based method introduced in Ref. [13] to find the optimal policy.

Solving Problem 1 is rather difficult. An admissible control policy of Problem 2 can be viewed as a special admissible control policy of Problem 1, thus the policy space of Problem 2 is a subset of the policy space of Problem 1. Therefore, the optimal performance of Problem 1 is no worse than the optimal performance of Problem 2.

Dynamic systems can be formulated with Markov systems. We define transition function $P_i(B|\boldsymbol{x})$ as the probability of the transition from mode i, state $\boldsymbol{x}$ to Borel set B, and $P(j,B|i,\boldsymbol{x})$ as the probability of the transition from mode i, state $\boldsymbol{x}$ to mode j, Borel set B. Then we have $P(j,B|i,\boldsymbol{x}) = p_{ij}P_i(B|\boldsymbol{x})$. For any cost function $f(i,\boldsymbol{x})$, define a transition operation

$$\mathcal{P}f(i,\boldsymbol{x}) = \sum_{j\in\mathcal{S}} p_{ij} \int_{\boldsymbol{y}\in R^n} P_i(\mathrm{d}\boldsymbol{y}|\boldsymbol{x})f(j,\boldsymbol{y}). \tag{6}$$

In the above definition, $\mathcal{P}f(i,\boldsymbol{x})$ represents the expected cost function at the next time epoch after one step transition when the current mode is i and the state is $\boldsymbol{x}$. Let $\mathcal{L}$ be a control policy which can be any admissible control policy of Problems 1, 2 and 3. Under this policy, the transition function, cost function and performance are denoted as $P^{\mathcal{L}}$, $f^{\mathcal{L}}$ and $J^{\mathcal{L}}$ respectively. Reference [15] gives the policy iteration formula to optimize performance.

Lemma 2 [15] Policy $\mathcal{L}'$ is better than policy $\mathcal{L}$, if and only if

1) For any $i \in \mathcal{S}$, $\boldsymbol{x} \in R^n$, we have

$$\mathcal{P}^{\mathcal{L}'}g^{\mathcal{L}}(i,\boldsymbol{x}) + f^{\mathcal{L}'}(i,\boldsymbol{x}) \leqslant \mathcal{P}^{\mathcal{L}}g^{\mathcal{L}}(i,\boldsymbol{x}) + f^{\mathcal{L}}(i,\boldsymbol{x});$$

2) There exists $l \in \mathcal{S}$ and a set $X \subset R^n$ with non-zero measure, such that for all $\boldsymbol{x} \in X$, we have

$$\mathcal{P}^{\mathcal{L}'}g^{\mathcal{L}}(l,\boldsymbol{x}) + f^{\mathcal{L}'}(l,\boldsymbol{x}) < \mathcal{P}^{\mathcal{L}}g^{\mathcal{L}}(l,\boldsymbol{x}) + f^{\mathcal{L}}(l,\boldsymbol{x}),$$

$g^{\mathcal{L}}$ is the performance potential under policy $\mathcal{L}$, which satisfies the Poisson equation:

$$J^{\mathcal{L}} + g^{\mathcal{L}} = f^{\mathcal{L}} + \mathcal{P}^{\mathcal{L}}g^{\mathcal{L}}.$$

3 Analysis of two classes of mode control policies

We shall discuss the condition under which the optimal admissible control policy of Problem 1 is better than that of Problem 2.

Lemma 3 If $\exists\, l\in\mathcal{S}, X\subset R^n$ and $\bar{p}_{lj}\in Y_l$, such that for l and $\boldsymbol{x}\in X$, we have

$$\sum_{j\in\mathcal{S}} \bar{p}_{lj} \int_{\boldsymbol{y}\in R^n} P_l^{\mathcal{L}_3^*(\boldsymbol{p}^*)}(\mathrm{d}\boldsymbol{y}|\boldsymbol{x})g^{\mathcal{L}_2^*}(j,\boldsymbol{y})$$
$$< \sum_{j\in\mathcal{S}} p_{lj}^* \int_{\boldsymbol{y}\in R^n} P_l^{\mathcal{L}_3^*(\boldsymbol{p}^*)}(\mathrm{d}\boldsymbol{y}|\boldsymbol{x})g^{\mathcal{L}_2^*}(j,\boldsymbol{y}), \tag{7}$$

then $\mathcal{L}_1^*$ is better than $\mathcal{L}_2^*$.

Proof $\mathcal{L}_2^*$ can be considered as a special admissible control policy of Problem 1:

$$\tilde{\mathcal{L}}_1 = \left\{\mathcal{L}_3^*(\boldsymbol{p}^*), \tilde{p}_{ij}(\boldsymbol{x}) = p_{ij}^*, \forall\boldsymbol{x}\right\}.$$

Policies $\mathcal{L}_2^*$ and $\tilde{\mathcal{L}}_1$ have the same performance. We can construct an admissible control policy for Problem 1 as follows:

$$\mathcal{L}_1' = \left\{\mathcal{L}_3^*(\boldsymbol{p}^*), p_{ij}'(\boldsymbol{x})\right\},$$

where

$$p_{ij}'(\boldsymbol{x}) = \begin{cases} \bar{p}_{ij}, & \text{if } i = l \text{ and } \boldsymbol{x} \in X, \\ p_{ij}^*, & \text{otherwise.} \end{cases} \tag{8}$$

Note that the cost function $f^{\mathcal{L}}(i,\boldsymbol{x})$ only depends on the feedback control law, and is independent of mode control policy, thus

$$\mathcal{P}^{\mathcal{L}_1'}g^{\tilde{\mathcal{L}}_1}(i,\boldsymbol{x}) + f^{\mathcal{L}_1'}(i,\boldsymbol{x})$$
$$= \sum_{j\in\mathcal{S}} p_{ij}'(\boldsymbol{x}) \int_{\boldsymbol{y}\in R^n} P_i^{\mathcal{L}_3^*(\boldsymbol{p}^*)}(\mathrm{d}\boldsymbol{y}|\boldsymbol{x})g^{\mathcal{L}_2^*}(j,\boldsymbol{y}) + f^{\mathcal{L}_3^*(\boldsymbol{p}^*)}(i,\boldsymbol{x}),$$

and

$$\mathcal{P}^{\tilde{\mathcal{L}}_1}g^{\tilde{\mathcal{L}}_1}(i,\boldsymbol{x}) + f^{\tilde{\mathcal{L}}_1}(i,\boldsymbol{x})$$
$$= \sum_{j\in\mathcal{S}} p_{ij}^* \int_{\boldsymbol{y}\in R^n} P_i^{\mathcal{L}_3^*(\boldsymbol{p}^*)}(\mathrm{d}\boldsymbol{y}|\boldsymbol{x})g^{\mathcal{L}_2^*}(j,\boldsymbol{y}) + f^{\mathcal{L}_3^*(\boldsymbol{p}^*)}(i,\boldsymbol{x}).$$

According to Eqs. (7) and (8), for l and $\boldsymbol{x} \in X$, we have

$$\mathcal{P}^{\mathcal{L}_1'}g^{\tilde{\mathcal{L}}_1}(l,\boldsymbol{x}) + f^{\mathcal{L}_1'}(l,\boldsymbol{x}) < \mathcal{P}^{\tilde{\mathcal{L}}_1}g^{\tilde{\mathcal{L}}_1}(l,\boldsymbol{x}) + f^{\tilde{\mathcal{L}}_1}(l,\boldsymbol{x}).$$

For any other mode i and state $\boldsymbol{x}$, we have

$$\mathcal{P}^{\mathcal{L}_1'}g^{\tilde{\mathcal{L}}_1}(i,\boldsymbol{x}) + f^{\mathcal{L}_1'}(i,\boldsymbol{x}) = \mathcal{P}^{\tilde{\mathcal{L}}_1}g^{\tilde{\mathcal{L}}_1}(i,\boldsymbol{x}) + f^{\tilde{\mathcal{L}}_1}(i,\boldsymbol{x}).$$

According to Lemma 2, $\mathcal{L}_1'$ is better than $\tilde{\mathcal{L}}_1$. And $\mathcal{L}_1^*$ is no worse than $\mathcal{L}_1'$, thus $\mathcal{L}_1^*$ is better than $\tilde{\mathcal{L}}_1$. Since $\mathcal{L}_2^* = \tilde{\mathcal{L}}_1$, $\mathcal{L}_1^*$ is better than $\mathcal{L}_2^*$.

Lemma 4 If $\exists\, l\in\mathcal{S}, X\subset R^n$ and $\bar{p}_{lj}\in Y_l$, such that for l and $\boldsymbol{x} \in X$, we have

$$\sum_{j\in\mathcal{S}} \bar{p}_{lj}\left[\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}_l - \boldsymbol{B}_l\boldsymbol{L}_l)^{\mathrm{T}}\boldsymbol{K}_j(\boldsymbol{A}_l - \boldsymbol{B}_l\boldsymbol{L}_l)\boldsymbol{x} + \mathrm{tr}\left(\boldsymbol{a}_l\boldsymbol{K}_j\right) + q_j\right]$$
$$< \sum_{j\in\mathcal{S}} p_{lj}^*\left[\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}_l - \boldsymbol{B}_l\boldsymbol{L}_l)^{\mathrm{T}}\boldsymbol{K}_j(\boldsymbol{A}_l - \boldsymbol{B}_l\boldsymbol{L}_l)\boldsymbol{x} + \mathrm{tr}\left(\boldsymbol{a}_l\boldsymbol{K}_j\right) + q_j\right]. \tag{9}$$

Thus $\mathcal{L}_1^*$ is better than $\mathcal{L}_2^*$. $\boldsymbol{K}_i, q_i, \boldsymbol{L}_i, \forall\, i\in\mathcal{S}$ are solutions to Problem 3 with a given $\boldsymbol{p}^*$.

Proof Put Eq. (3) into Eq. (7), and the proof can be obtained with Lemmas 3 and 4.

Let

$$\boldsymbol{\Pi} \triangleq (\boldsymbol{A}_l - \boldsymbol{B}_l\boldsymbol{L}_l)^{\mathrm{T}}\sum_{j\in\mathcal{S}} \boldsymbol{K}_j\left(\bar{p}_{lj} - p_{lj}^*\right)(\boldsymbol{A}_l - \boldsymbol{B}_l\boldsymbol{L}_l),$$

and

$$\Gamma \triangleq \sum_{j\in\mathcal{S}} \left(\mathrm{tr}(\boldsymbol{a}_l\boldsymbol{K}_j) + q_j\right)\left(p_{lj}^* - \bar{p}_{lj}\right),$$

then Eq. (9) is equivalent to

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Pi}\boldsymbol{x} < \Gamma. \tag{10}$$

With the above lemmas, we obtain the sufficient condition under which the optimal closed-loop mode control is better than the optimal open-loop mode control.

Theorem 1 If $\exists\, l \in \mathcal{S}, X \subset R^n$ and $\bar{p}_{lj} \in Y_l$, such that one of the following three conditions is satisfied, then $\mathcal{L}_1^*$ is better than $\mathcal{L}_2^*$.

1) $\boldsymbol{\Pi}$ is an indefinite matrix;

2) $\boldsymbol{\Pi}$ is a non-zero positive semi-definite matrix, and $\Gamma > 0$;

3) $\boldsymbol{\Pi}$ is a non-zero negative semi-definite matrix, and $\Gamma < 0$.

$\boldsymbol{K}_i, q_i, \boldsymbol{L}_i, \forall i \in \mathcal{S}$ are solutions to Problem 3 when $\boldsymbol{p}^*$ is given.

Proof First, Eq. (10) does not constantly hold. Otherwise, from Lemma 2 we obtain that $\bar{p}_{lj}$ is better than p_{lj}^*. This conclusion conflicts with the definition that $\boldsymbol{p}^*$ is the jump probability matrix corresponding to the optimal admissible control policy of Problem 2. Thus the solution to Eq. (10) is not R^n space. Therefore, the Lemma 4 condition is equivalent to that for the optimal admissible control policy of Problem 2, $\mathcal{L}_2^*$, $\exists\, l \in \mathcal{S}$ and $\bar{p}_{lj} \in Y_l$, such that Eq. (10) has a solution.

The matrix $\boldsymbol{\Pi}$ may be found in four cases: non-zero positive semi-definite, non-zero negative semi-definite, zero and indefinite. We will discuss these four cases respectively.

When condition 1) holds, i.e., $\boldsymbol{\Pi}$ is indefinite, then Eq. (10) always has a solution no matter what Γ is.

When condition 2) holds, i.e., $\boldsymbol{\Pi}$ is non-zero positive semi-definite, then we have $0 \leqslant \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Pi}\boldsymbol{x} < +\infty$. Thus, Eq. (10) has a solution and does not hold constantly, which means $\Gamma > 0$.

When condition 3) holds, i.e., $\boldsymbol{\Pi}$ is non-zero negative semi-definite, then we have $-\infty < \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Pi}\boldsymbol{x} \leqslant 0$. If $\Gamma \geqslant 0$, then from Lemma 2 we obtain that $\bar{p}_{lj}$ is better than p_{lj}^*, which conflicts with the definition of $\boldsymbol{p}^*$. Therefore, $\Gamma < 0$.

If $\boldsymbol{\Pi}$ is zero, when $\Gamma > 0$, the solution of Eq. (10) is R^n. And when $\Gamma \leqslant 0$, Eq. (10) has no solution. Therefore, $\boldsymbol{\Pi}$ cannot be zero. Theorem 1 is proved.

Theorem 1 gives the sufficient condition under which the optimal closed-loop mode control is better than the optimal open-loop mode control. If this condition is satisfied, the closed-loop mode control achieves better performance than the open-loop mode control. This result is helpful for designing controllers. For the optimization problem of JLQG with controlled jump probabilities, seeking the optimal open-loop mode control is relatively simpler with a smaller policy space. However, seeking the optimal closed-loop mode control is rather difficult, since if the jump of modes depends on system state, the state feedback control law then cannot be solved by the standard JLQG model any more. The only existing methods are dynamic programming or simulation-based policy iteration approaches [15] which are time consuming. If the optimal open-loop mode control policy is obtained, we can apply the sufficient condition provided by Theorem 1 to check if there is any better closed-loop mode control policy. Thereafter, we can improve the jump probability of modes.

For a special case with two modes and a one-dimensional state, Theorem 1 can be simplified.

Corollary 1 $m = n = 1$ and the mode set is $\mathcal{S} = \{1,2\}$. If $\exists\, l \in \mathcal{S}$, $\bar{p}_{lj} \in Y_l$, $\bar{p}_{lj} \neq p_{lj}^*$, such that

$$\frac{\left(\sigma_l^2 K_2 + q_2\right) - \left(\sigma_l^2 K_1 + q_1\right)}{(A_l - B_l L_l)^2 (K_1 - K_2)} > 0, \tag{11}$$

then $\mathcal{L}_1^*$ is better than $\mathcal{L}_2^*$. $K_i, q_i, L_i, \forall i \in \mathcal{S}$ are the solutions to Problem 3 when $\boldsymbol{p}^*$ is given.

Proof In this case,

$$\Pi = (A_l - B_l L_l)^2 (\bar{p}_{l1} - p_{l1}^*)(K_1 - K_2),$$

$$\Gamma = (p_{l1}^* - \bar{p}_{l1})\left[\left(\sigma_l^2 K_1 + q_1\right) - \left(\sigma_l^2 K_2 + q_2\right)\right].$$

According to Theorem 1, it is easy to obtain the proof for Corollary 1.

Furthermore, we can solve the interval X. Solve the equation $\Pi x^2 = \Gamma$, we have

$$x_+ = \sqrt{\frac{\left(\sigma_l^2 K_2 + q_2\right) - \left(\sigma_l^2 K_1 + q_1\right)}{(A_l - B_l L_l)^2 (K_1 - K_2)}},$$
$$x_- = -\sqrt{\frac{\left(\sigma_l^2 K_2 + q_2\right) - \left(\sigma_l^2 K_1 + q_1\right)}{(A_l - B_l L_l)^2 (K_1 - K_2)}}. \tag{12}$$

If Eq. (11) holds, x_+ and x_- exist. If $(\bar{p}_{l1} - p_{l1}^*)(K_1 - K_2) > 0$, we have $X = [x_-, x_+]$. If $(\bar{p}_{l1} - p_{l1}^*)(K_1 - K_2) < 0$, the interval becomes $X = (-\infty, x_-] \cup [x_+, +\infty)$.

If the condition in Theorem 1 is satisfied, we should seek an admissible control policy of Problem 1 to improve system performance. Seeking $\mathcal{L}_1^*$ is rather difficult. As a compromise, we tried to find a good admissible control policy for Problem 1, which improves system performance significantly compared with $\mathcal{L}_2^*$. All the proofs of the above lemmas and theorems are all constructive, thus we can construct a closed-loop mode control following the same idea. Combining the original feedback control law and the closed-loop mode control, we obtain a new policy, which is better than $\mathcal{L}_2^*$ and $\boldsymbol{p}^*$.

Algorithm 1 Construct an admissible control policy for Problem 1.

1) Solve the optimal admissible control policy of Problem 2, $\mathcal{L}_2^*$.

2) Check the condition contained in Theorem 1. When it is satisfied, there may exist more than one mode $l \in \mathcal{S}$, and even in one Y_l there may exist several possible $\bar{p}_{lj}$s. Then choose any mode $l \in \mathcal{S}$, which satisfies the condition and corresponding $X \subset R^n$ and $\bar{p}_{lj} \in Y_l$.

3) Implement a one-step policy iteration for $\mathcal{L}_2^*$. Construct $\mathcal{L}_1' = \left\{\mathcal{L}_3^*(\boldsymbol{p}^*), p'_{ij}(\boldsymbol{x})\right\}$ by using Eq. (8).

4 Numerical example

Example 1 Consider a system with a one-dimensional state and two modes, $\mathcal{S} = \{1,2\}$.

$$A_1 = 0.5;\ B_1 = 2;\ M_1 = 1;\ N_1 = 10;\ \sigma_1 = 2;$$

$$A_2 = 1;\ B_2 = 1;\ M_2 = 2;\ N_2 = 10;\ \sigma_2 = 1;$$

$$Y_1 = \{[0.8\ 0.2],[0.2\ 0.8]\}, Y_2 = \{[0.5\ 0.5]\}.$$

First, find the optimal open-loop mode control. Therefore, we have

$$\boldsymbol{p}^* = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}.$$

Then $K_1 = 1.236$, $K_2 = 2.614$, $q_1 = 6.549$, $q_2 = 0.664$, $L_1 = 0.094$, $L_2 = 0.161$. When $l = 1$, we have

$$\left[\left(\sigma_1^2 K_2 + q_2\right) - \left(\sigma_1^2 K_1 + q_1\right)\right] / \left[(A_1 - B_1 L_1)^2 (K_1 - K_2)\right] = 2.772 > 0,$$

thus the optimal admissible control policy of Problem 1, $\mathcal{L}_1^*$, is better than $\mathcal{L}_2^*$. According to Eq. (12), we have

$$X = [-1.665, 1.665].$$

Using Algorithm 1, we can construct a better closed-loop mode control: When $x \in X$,

$$\boldsymbol{p}'(x) = \begin{bmatrix} 0.2 & 0.8 \\ 0.5 & 0.5 \end{bmatrix};$$

otherwise, $\boldsymbol{p}'(x) = \boldsymbol{p}^*$. Through simulation we have that the system performance under policy $\mathcal{L}_2^*$ is 4.87, while the performance under constructed policy $\mathcal{L}_1'$ is 4.06, which achieves a 17% improvement than $\mathcal{L}_2^*$.

5 Conclusions

The paper considers two classes of mode control policies, and gives the sufficient condition under which the optimal closed-loop mode control is better than the optimal open-loop mode control. When optimizing the JLQG model with controlled jump probabilities of modes, we can first find the optimal feedback control law and then the optimal open-loop mode control, after which we can check the condition that is sufficient. If it is satisfied, we can further seek better closed-loop mode control policies; otherwise, the open-loop mode control is good enough, and we can stop searching to avoid huge computations. If the optimal closed-loop mode control is difficult to obtain, we can apply Algorithm 1 to construct a closed-loop mode control, which significantly improves system performance. The numerical example verifies the efficiency of the proposed algorithm.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Nos. 60574064, 60736027).

References

1. ChengD Z, GuoY Q. Advances on switched systems. Control Theory and Applications, 2005, 22(6): 954–960 (in Chinese)
2. Abou-kandil H, De Smet O, Freiling G, et al. Flow control in a failure-prone multi-machine manufacturing system. In: Proceedings of INRIA/IEEE Symposium on Emerging Technologies and Factory Automation. 1995, 2: 575–583
3. Boukas E K, Shi P, Andijani A. Robust inventory-production control problem with stochastic demand. Optimal Control Applications and Methods, 1999, 20(1): 1–20
4. Ji Y, Chizeck H J. Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control. IEEE Transactions on Automatic Control, 1990, 35(7): 777–788
5. Costa O, Fragoso M D, Marques R P. Discrete-Time Markov Jump Linear Systems. London: Springer-Verlag, 2005
6. Xue F, Guo L. Necessary and sufficient conditions for adaptive stabilizability of jump linear systems. Communications in Information and Systems, 2001, 1(2): 205–224
7. Zhang L J, Li C W, Cheng D Z. Robust adaptive control of Markov jump systems with parameter uncertainties. Control and Decision, 2005, 20(9): 1030–1033 (in Chinese)
8. Liu F. Robust L_2-L_∞ filtering for uncertain jump systems. Control and Decision, 2005, 20(1): 32–35 (in Chinese)
9. Liu F, Zhang X H. Robust control for jump systems with L_2 gain constraints. Control Theory and Applications, 2006, 23(3): 373–377 (in Chinese)
10. Liu F, Su H Y, Chu J. Robust positive real control of Markov jump systems with parametric uncertainties. Acta Automatica Sinica, 2003, 29(5): 761–766 (in Chinese)
11. Ji Y, Chizeck H J. Optimal quadratic control of jump linear systems with separately controlled transition probabilities. International Journal of Control, 1989, 49(2): 481–491
12. Boukas E K, Liu Z K. Jump linear quadratic regulator with controlled jump rates. IEEE Transactions on Automatic Control, 2001, 46(2): 301–305
13. Xu Y K, Chen X. Discrete-time JLQG with dependently controlled jump probabilities. In: Proceedings of IEEE 22nd International Symposium on Intelligent Control, Singapore. 2007: 441–445
14. Cao X R. Stochastic Learning and Optimization: A Sensitivity-Based Approach. New York: Springer, 2007
15. Zhang K J, Xu Y K, Chen X, et al. Policy iteration based feedback control. Automatica, 2008, 44(4): 1055–1061