Systems & Control Letters 93 (2016) 69–76

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Mean-field stochastic linear–quadratic optimal control with Markov jump parameters

Yuan-Hua Ni [a], Xun Li [b,*], Ji-Feng Zhang [c]

[a] Department of Mathematics, School of Science, Tianjin Polytechnic University, Tianjin 300387, PR China

[b] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hunghom, Kowloon, Hong Kong

[c] Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China

ARTICLE INFO

Article history:
Received 23 January 2015
Received in revised form
28 October 2015
Accepted 8 April 2016
Available online 13 May 2016

Keywords:
Mean-field
Markov jump
Stochastic control

ABSTRACT

This paper considers a class of mean-field stochastic linear–quadratic optimal control problems with Markov jump parameters. The new feature of these problems is that means of state and control are incorporated into the systems and the cost functional. Based on the modes of Markov chain, the corresponding decomposition technique of augmented state and control is introduced. It is shown that, under some appropriate conditions, there exists a unique optimal control, which can be explicitly given via solutions of two generalized difference Riccati equations. A numerical example sheds light on the theoretical results established.

1. Introduction

During the past three decades, Markov jump systems have gained a great deal of attention. Such systems often arise in reality with component failures or repairs, changing subsystem interconnections, and abrupting environmental disturbances. It can be found in robotic manipulator systems, aircraft control systems, large scale flexible structures for space stations (such as antenna, solar arrays, among others), and flexible manufacturing systems, on which an actuator or a sensor failure is a quite common occurrence. Without any intention of being exhaustive here, we mention [1–9] and the monographs [10–12] to see different aspects of control problems corresponding to Markov jump systems.

In this paper, a kind of mean-field stochastic linear–quadratic (LQ) optimal control problem with Markov jump parameters is investigated. Compared with the standard stochastic LQ optimal control problems with Markov jump parameters, an important feature of the problem in this paper is that the cost functional involves nonlinearly the states and the controls as well as their expected values. Such a feature roots itself in the category of mean-field theory, which is developed to study the collective behaviors resulting from individuals' mutual interactions in various physical and sociological dynamical systems. There exist many successful applications of the mean-field formulation in various field of engineering, games, finance and economics in the past few years. Recently, stochastic maximum principles of mean-field type are extensively studied in several works [13–15], which specify the necessary conditions for the optimality. As applications, [13,14] studied the Markowitz mean–variance portfolio selection and a class of mean-field LQ problems using stochastic maximum principle. [15] considered mean-field control problems with partial information. [16] investigated the definite mean-field LQ control over a finite time horizon using a variational method and a decoupling technique. It is shown that the optimal control is of linear feedback form and that the gains are represented by solutions of two coupled differential Riccati equations. [17] formulated the discrete-time definite mean-field LQ problem as an operator stochastic LQ optimal control problem. By the kernel-range decomposition representation of the expectation operator and its pseudo-inverse, an optimal control is obtained based on the solutions of two Riccati difference equations. Furthermore, the closed-loop formulation is also investigated. Later, [18,19] generalized results obtained in [16,17] to the case of infinite time horizon.

It is worth noting that the recent research on controlled mean-field stochastic differential and difference equations is partially relighted by a surge of interest in mean-field games [20–27]. Particularly, [21–23] investigated large population stochastic dynamical games with mean-field terms. [24] considered similar

* Corresponding author.
E-mail addresses: yhni@amss.ac.cn (Y.-H. Ni), malixun@polyu.edu.hk (X. Li), jif@iss.ac.cn (J.-F. Zhang).

http://dx.doi.org/10.1016/j.sysconle.2016.04.002
0167-6911/

problems from the viewpoint of mean-field theory. [25,27] dealt with the asymptotically optimal decentralized control problem for the large population multi-agent systems with Markov jump parameters. [26] considered risk-sensitive mean-field games with some interesting aspects, and [28] considered LQ mean-field games via the adjoint equation approach. It is worth pointing out that mean-field games can be reduced to a standard control problem, but the mean-field type control is a non-standard control problem (see [20]). [29,27] studied the mean-field games involving random coefficients. [29] established a stochastic maximum principle for general nonlinear system, which provides necessary conditions for the existence of Nash equilibria in a certain form of N-agent mean-field stochastic differential game. [27] investigated an infinite horizon mean-field LQ games with Markov jump coefficients. Specifically, the distributed strategies were given by solving a Markov jump tracking problem. It is shown that the closed-loop system is uniformly stable, and the distributed strategies are asymptotically optimal in the sense of Nash equilibrium, as the number of agents grows to infinity.

To our knowledge, most of the existing results about mean-field LQ optimal control problems mainly focus on deterministic coefficients. In the real problem, however, one often encounters systems with random coefficients. In the case of deterministic coefficients, it is shown that the optimal controls are linear feedback forms of the state x_k and its expectation $\mathbb{E}x_k$. For a deterministic matrix M_k, $\mathbb{E}(M_kx_k) = M_k\mathbb{E}x_k$, which is an essential property to obtain the optimal control to mean-field LQ problems. But, when M_k becomes random, the property of $\mathbb{E}(M_kx_k) = M_k\mathbb{E}x_k$ no longer holds. This may result in fundamental difficulty in tackling such stochastic control problems with random coefficients (see [16]).

In this paper, we introduce a decomposition technique of the state and the control based on the modes of Markov chain, which is shown to be efficient to attack Markov jump mean-field LQ problem. By completing the square for two different parts of the augmented state and control, the optimal control is constructed via solutions to two generalized difference Riccati equations. The optimal control is shown to be a linear feedback of the current state and its expectation of decomposition of the state.

The decomposition technique adopted in this paper is motivated by [10,3], where a decomposition technique of the state was introduced corresponding to the modes of Markov chain, and the stability of the control-free systems was investigated. In this paper, based on the modes of Markov chain, not only the state and the control are decomposed, but also the mean-field LQ optimal control problem with Markov jump parameters is decomposed to a solvable formulation. By the augmented state and control, we can successfully construct the optimal control of the original mean-field LQ optimal control problem with Markov jumps. A numerical example in Section 4 illustrates that our results are significantly different from those results corresponding to standard Markov jump stochastic LQ problems.

The rest of this paper is organized as follows. Section 2 gives some preliminaries. Section 3 presents the main results of this paper. Section 4 introduces a numerical example. Concluding remarks are given in Section 5.

2. Problem formulation

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space which is assumed to be abundant enough such that two processes $\theta \equiv \{\theta_k\}$, $w \equiv \{w_k\}$ and a random ζ live on it.

(a) θ is a homogeneous Markov chain taking values in a finite set $\{1, .., m\} \equiv \mathcal{M}$ with a stationary one-step transition probability matrix $\Lambda = (p_{ij})$. The (i, j)th entry of Λ is

$$p_{ij} = P(\theta_{k+1} = j|\theta_k = i), \quad i, j \in \mathcal{M},\ k = 0, 1, \dots. \tag{2.1}$$

The initial distribution of θ_0 is denoted by $\nu = (\nu_1, \dots, \nu_m)^T$, where the superscript T denotes the transposition of a matrix or a vector.

(b) w is a martingale difference sequence in the sense that $\mathbb{E}[w_{k+1}|\mathcal{F}_k] = 0$ with $\mathcal{F}_k$ being the σ-algebra generated by $\{\zeta, w_l, \theta_l, l = 0, 1, \dots, k\}$. It is assumed that w has the property

$$\mathbb{E}[(w_{k+1})^2|\mathcal{F}_k] = 1, \tag{2.2}$$

and that θ and w are independent of each other.

(c) ζ is square integrable with a known distribution.

Consider the cost functional

$$\begin{aligned} J(\zeta, u; \theta_0) = &\sum_{k=0}^{N-1} \mathbb{E}\left[x_k^T Q_{\theta_k} x_k + (\mathbb{E}x_k)^T \bar{Q}_{\theta_k} \mathbb{E}x_k \right. \\ &\left. + u_k^T R_{\theta_k} u_k + (\mathbb{E}u_k)^T \bar{R}_{\theta_k} \mathbb{E}u_k\right] \\ &+ \mathbb{E}\left(x_N^T G_{\theta_N} x_N\right) + \mathbb{E}\left[(\mathbb{E}x_N)^T \bar{G}_{\theta_N} \mathbb{E}x_N\right], \end{aligned} \tag{2.3}$$

which is subject to the following dynamics

$$\begin{cases} x_{k+1} = \left[A_{\theta_k}x_k + B_{\theta_k}u_k\right] + \left[C_{\theta_k}x_k + D_{\theta_k}u_k\right]w_k, \\ x_0 = \zeta, \quad k \in \mathbb{T} \equiv \{0, 1, \dots, N-1\}. \end{cases} \tag{2.4}$$

Here, N is a positive integer; $\{x_k \in \mathbb{R}^n, k \in \bar{\mathbb{T}}\}$ and $\{u_k \in \mathbb{R}^p, k \in \mathbb{T}\}$ are the state process and the control process, respectively, with $\bar{\mathbb{T}} = \{0, 1, \dots, N\}$; θ represents the mode of system (2.4). When $\theta_k = i \in \mathcal{M}$, $A_{\theta_k}, B_{\theta_k}, C_{\theta_k}, D_{\theta_k}, Q_{\theta_k}, \bar{Q}_{\theta_k}, R_{\theta_k}, \bar{R}_{\theta_k}$ will be denoted by $A^i, B^i, C^i, D^i, Q^i, \bar{Q}^i, R^i, \bar{R}^i$, respectively, which are of compatible dimensions. Similar notations hold for G_{θ_N} and $\bar{G}_{\theta_N}$.

Throughout this paper, θ, w and ζ are assumed to be available to us. Therefore, at time point k, the information set that we have is $\mathcal{F}_{k-1}$. Let $L^2_{\mathcal{F}}(\mathbb{T}; \mathbb{R}^m)$ be the set of $\mathbb{R}^m$-valued processes $\nu = \{\nu_k, k \in \mathbb{T}\}$ such that ν_k is $\mathcal{F}_{k-1}$-measurable and $\sum_{k=0}^{N-1} \mathbb{E}|\nu_k|^2 < \infty$. The optimal control problem of this paper is as follows.

Problem (MF-JLQ). *Given ζ, find a $u^* \in \mathcal{U}_{ad}$ such that*

$$J(\zeta, u^*; \theta_0) = \inf_{u \in L^2_{\mathcal{F}}(\mathbb{T};\mathbb{R}^m)} J(\zeta, u; \theta_0). \tag{2.5}$$

We call u^* an optimal control for Problem (MF-JLQ).

3. Main results

3.1. System dynamics and cost functional

In [17], the state and the control are decomposed into two orthogonal parts, respectively. By completing the squares for these two parts, we derive the optimal control, which is a linear feedback of the state and its expectation. For a deterministic matrix M_k, we have $\mathbb{E}(M_kx_k) = M_k\mathbb{E}x_k$, which is an essential property to obtain the optimal control [17]. If M_k becomes random, the property $\mathbb{E}(M_kx_k) = M_k\mathbb{E}x_k$ no longer holds. In particular, taking expectation for both sides of (2.4), we have

$$\mathbb{E}x_{k+1} = \mathbb{E}\left[A_{\theta_k}x_k\right] + \mathbb{E}\left[B_{\theta_k}u_k\right].$$

As process θ appears, it is impossible to obtain a deterministic linear system for $\mathbb{E}x_k$. Hence, the results established in [17] cannot be directly applied to solve the case with random coefficients.

To overcome this difficulty, a decomposition technique, corresponding to the modes of Markov chain, is proposed:

$$\begin{cases} y_k^j = x_k I_{(\theta_k=j)}, \quad \forall j \in \mathcal{M}, \\ v_k^j = u_k I_{(\theta_k=j)}, \quad \forall j \in \mathcal{M}. \end{cases} \tag{3.1}$$

Based on this decomposition, the optimal control of Problem (MF-JLQ), which gets around the difficulty mentioned above, can be constructed directly.