Automatica 57 (2015) 65–77

Contents lists available at ScienceDirect

Automatica

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

Discrete-time mean-field Stochastic linear–quadratic optimal control problems, II: Infinite horizon case[☆]

Yuan-Hua Ni[a,b], Robert Elliott[c,d,1], Xun Li[e]

[a] Department of Mathematics, School of Science, Tianjin Polytechnic University, Tianjin 300160, China
[b] Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[c] School of Mathematical Sciences, University of Adelaide, SA, 5005, Australia
[d] Haskayne School of Business, University of Calgary, Calgary, Alberta, Canada
[e] Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China

ARTICLE INFO

Article history:
Received 5 August 2013
Received in revised form
1 November 2014
Accepted 25 March 2015
Available online 16 May 2015

Keywords:
Stochastic linear–quadratic optimal control
Mean-field theory
Generalized algebraic Riccati equation

ABSTRACT

This paper first presents results on the equivalence of several notions of L^2-stability for linear mean-field stochastic difference equations with random initial value. Then, it is shown that the optimal control of a mean-field linear–quadratic optimal control with an infinite time horizon uniquely exists, and the optimal control can be expressed as a linear state feedback involving the state and its mean, via the minimal nonnegative definite solution of two coupled algebraic Riccati equations. As a byproduct, the open-loop L^2-stabilizability is proved to be equivalent to the closed-loop L^2-stabilizability. Moreover, the minimal nonnegative definite solution, the maximal solution, the stabilizing solution of the algebraic Riccati equations and their relations are carefully investigated. Specifically, it is shown that the maximal solution is employed to construct the optimal control and value function to another infinite time horizon mean-field linear–quadratic optimal control. In addition, the maximal solution being the stabilizing solution, is completely characterized by properties of the coefficients of the controlled system. This enriches the existing theory about stochastic algebraic Riccati equations. Finally, the notion of exact detectability is introduced with its equivalent characterization of stochastic versions of the Popov–Belevitch–Hautus criteria. It is then shown that the minimal nonnegative definite solution is the stabilizing solution if and only if the uncontrolled system is exactly detectable.

1. Introduction

This paper discusses the mean-field stochastic linear–quadratic (LQ) optimal control problem with an infinite horizon. Specifically, we wish to minimize the cost functional

$$J(\xi;u)=\sum_{k=0}^{\infty}\mathbb{E}\left[x_k^TQx_k+(\mathbb{E}x_k)^T\bar{Q}\mathbb{E}x_k+u_k^TRu_k+(\mathbb{E}u_k)^T\bar{R}\mathbb{E}u_k\right] \quad (1.1)$$

subject to the dynamics

$$\begin{cases} x_{k+1} = (Ax_k+\bar{A}\mathbb{E}x_k+Bu_k+\bar{B}\mathbb{E}u_k) \\ \qquad\quad +(Cx_k+\bar{C}\mathbb{E}x_k+Du_k+\bar{D}\mathbb{E}u_k)w_k, \\ x_0=\xi, \quad k\in\{0,1,2,\ldots\}\triangleq\mathbb{N}. \end{cases} \quad (1.2)$$

System (1.2) is a discrete-time stochastic difference equation (SDE) of McKean–Vlasov type, which is also referred to as a mean-field SDE (MF-SDE). Here, the term "mean-field" comes from the mean-field theory, which is developed to study the collective behavior resulting from individuals' mutual interactions in various physical and sociological dynamical systems. According to the mean-field theory, the interactions among agents can be modeled by a mean-field term. Letting the number of individuals approach infinity, the mean-field term can approximate the expected value. As a feature of MF-SDEs, the dynamics depend on the statistical distribution of the solution. This provides effective techniques for studying large systems by reducing the dimension and the complexity. In addition to the mean-field theory above, the study of mean-field stochastic LQ optimal control is motivated by optimal control theory: People might like to have the optimal control, as well as the optimal

☆ Dr. Yuan-Hua Ni wishes to thank the supports from the National Natural Science Foundation of China under grant 11101303 and 11471242, the China Postdoctoral Science Foundation (2014M560128) and the China Scholarship Council (201308120011). Professor Robert Elliott wishes to thank the Australian Research Council (DP0877639) and Natural Sciences and Engineering Research Council of Canada for support. Dr. Xun Li wishes to thank the Hong Kong RGC for support with grants GRF520412 and GRF15209614. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Valery Ugrinovskii under the direction of Editor Ian R. Petersen..

E-mail addresses: yhni@amss.ac.cn (Y.-H. Ni), relliott@ucalgary.ca (R. Elliott), malixun@polyu.edu.hk (X. Li).

[1] Tel.: +61 8 83135077; fax: +61 8 83133696.

http://dx.doi.org/10.1016/j.automatica.2015.04.002
0005-1098/© 2015 Elsevier Ltd. All rights reserved.

state, to be not too "random", as pointed out in Yong (2013). To achieve that, one could include variation of the state process and/or variation of the control process in the cost functional. In fact, a particular example of this is the Markowitz mean–variance portfolio selection problem in financial investment, where the risk to be minimized is quantified by using the variance of the wealth. This Noble-Prize-winning approach became the foundation of modern finance theory and inspired hundreds of extensions and applications.

The investigation of continuous-time mean-field stochastic differential equations may be traced back to 1960s (McKean, 1966); see also Sznitman (1989) for early developments. Recently, there is an increasing interest in the mean-field control theory in mathematics and the control communities. In Ahmed and Ding (2001), to cope with the possible time-inconsistency of optimal control, an extended version of the dynamic programming principle was derived by using the Nisio nonlinear operator semigroup. Subsequently, stochastic maximum principles were studied in several works (Ahdersson & Djehiche, 2011; Buckdahn, Djehiche, & Li, 2011; Li, 2012), which specify necessary conditions for optimality. The results range from the case of a convex action space to a general action space. It is valuable to mention that the adjoint equations are mean-field backward stochastic differential equations (Buckdahn, Djehiche, Li, & Peng, 2009). As applications, the Markowitz mean–variance portfolio selection and a class of mean-field LQ problems are studied in Ahdersson and Djehiche (2011) and Li (2012) using the stochastic maximum principle. In Yong (2013), mean-field LQ control with a finite time horizon is studied 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 using solutions of the two coupled differential Riccati equations. Later, Huang, Li, and Yong (2015) generalizes results in Yong (2013) to the case with an infinite time horizon. For other interesting aspects of mean-field optimal control problems, readers may refer to, for example, Ahmed (2007), Bjork and Murgoci (2010), Meyer-Brandis, Oksendal, and Zhou (2012) and Yong (2013) and related works. It is worth mentioning that the study of controlled mean-field stochastic differential equations is also partially motivated by a recent surge of interest in mean-field games (Huang, Caines, & Malhame, 2003, 2007; Lasry & Lions, 2007; Li & Zhang, 2008a,b). Compared with the topic of this paper, mean-field games use decentralized controls, that is, the controls are selected to achieve each individual's own goal by using local information.

This paper is a continuation of Elliott, Li, and Ni (2013) where a discrete-time mean-field LQ problem with a finite time horizon was studied. As pointed out in Elliott et al. (2013), a typical case for studying discrete-time systems is that the signal values are available only for measurement or manipulation at certain times, for example, a continuous time system is sampled at certain times. In Elliott et al. (2013), the mean-field LQ optimal control problem is formulated as an operator stochastic LQ optimal control problem, and is solved using the kernel-range decomposition representation of the expectation operator and its pseudo-inverse. This formulation may be viewed as an alternative to the variational method. In this paper, we shall deal with the infinite horizon case of mean-field LQ control. The content and the contribution of this paper are presented in what follows.

Firstly, several notions of L^2-stability for the control-free MF-SDE (1.2) ($B = \bar{B} = D = \bar{D} = 0$) are introduced, and are shown to be equivalent by using the discrete-time semigroup theory. In addition, two notions about the stabilizability of system (1.2) are introduced: the closed-loop L^2-stabilizability and the open-loop L^2-stabilizability. Clearly, the closed-loop L^2-stabilizability implies the open-loop L^2-stabilizability. *Secondly*, under appropriate conditions, the optimal control in $\mathcal{U}_{ad}$ (Problem (MF-LQ)) uniquely exists, and is represented using the minimal nonnegative definition solution to two coupled discrete-time algebraic Riccati equations (AREs) (4.2). As a byproduct, it is shown that for (1.2) the closed-loop L^2-stabilizability is equivalent to the open-loop L^2-stabilizability. *Thirdly*, from the AREs (4.2), we construct another ARE (5.8), from which the existence of the maximal solution of the AREs (4.2) is easily achieved. In addition, it is shown that the optimal control in $\overline{\mathcal{V}}_{ad}$ (Problem (MF-LQ*)) uniquely exists and is represented using the maximal solution of the AREs (4.2). This result is more integrated than that in Huang et al. (2015), where the maximal solution is only shown to be used to represent the corresponding value function. Moreover, the maximal solution being the stabilizing solution is completely characterized by the properties of the coefficients of system (1.2); for this, see Theorems 5.4 and 5.5. To our best knowledge, this is the first such result for stochastic AREs. Therefore, this result enriches the existing theory of stochastic AREs, which is viewed as one of the main contributions of this paper. It is valuable to mention that such result for deterministic ARE (Corollary 5.3) is also new; for this, see Remark 5.4. Furthermore, exact detectability is introduced. It is shown that the minimal nonnegative definite solution is the stabilizing solution if and only if the uncontrolled system $[A, \bar{A}; C, \bar{C}|Q, \bar{Q}]$ is exactly detectable. Under this detectability condition, the AREs (4.2) admit one nonnegative definite solution.

The paper is organized as follows. The next section gives the problem formulation. In Section 3, the stability and stabilizability of MF-SDEs are studied. Section 4 gives a complete solution to the infinite horizon mean-field LQ problem. In Section 5, the nonnegative definite solutions are thoroughly investigated. Some concluding remarks are given in Section 6.

2. Problem formulation

In (1.1) and (1.2), $Q, \bar{Q}, A, \bar{A}, C, \bar{C} \in \mathbb{R}^{n\times n}$, $R, \bar{R} \in \mathbb{R}^{m\times m}$ and $B, \bar{B}, D, \bar{D} \in \mathbb{R}^{n\times m}$ are given constant deterministic matrices, and $\mathbb{E}$ is the expectation operator. $\{x_k, k \in \mathbb{N}\}$ and $\{u_k, k \in \mathbb{N}\}$ are the state process and control process, respectively, and ξ is a square-integrable random variable. $\{w_k, k \in \mathbb{N}\}$ represents the stochastic disturbance, which is assumed to be a martingale difference sequence in the sense that

$$\mathbb{E}[w_{k+1}|w_l, l = 0, 1, \ldots, k] = 0,$$

with bounded second-order conditional moments

$$\mathbb{E}[(w_{k+1})^2|w_l, l = 0, 1, \ldots, k] = 1. \tag{2.1}$$

In this paper, we assume ξ and $\{w_k, k \in \mathbb{N}\}$ are independent of each other. Let $\mathcal{F}_k$ be the σ-algebra generated by $\{\xi, w_l, l = 0, 1, \ldots, k-1\}$. Before proceeding, introduce a space of L^2-summable functions:

$$L^2_{\mathcal{F}}(\mathbb{R}^l) = \left\{ \phi \left| \begin{array}{l} \phi_k \in \mathbb{R}^l \text{ and is } \mathcal{F}_k\text{-measurable}, k \in \mathbb{N}, \\ \sum_{k=0}^{\infty} \mathbb{E}|\phi_k|^2 < \infty \end{array} \right. \right\},$$

where $\phi = (\phi_0, \phi_1, \ldots)$. Clearly, $L^2_{\mathcal{F}}(\mathbb{R}^l)$ is a Hilbert space endowed with norm $\|\phi\| = \left(\sum_{k=0}^{\infty} \mathbb{E}|\phi_k|^2\right)^{\frac{1}{2}}$, where $|\phi_k| = (\phi_k^T \phi_k)^{\frac{1}{2}}$. To make the cost functional meaningful, we introduce the following set of admissible controls $u = (u_0, u_1, \ldots)$:

$$\mathcal{U}_{ad} = \left\{ u \in L^2_{\mathcal{F}}(\mathbb{R}^m) \middle| J(\xi; u) < \infty \right\}.$$

The optimal control problem studied in this paper is as follows.

Problem (MF-LQ). For any given initial value ξ, find a $u^\circ \in \mathcal{U}_{ad}$ such that

$$J(\xi; u^\circ) = \inf_{u \in \mathcal{U}_{ad}} J(\xi; u). \tag{2.2}$$

Problem (MF-LQ) is a free end-point problem, as the end point of the trajectory is free when $k \to \infty$.