Operations Research Letters 44 (2016) 307–312

Contents lists available at ScienceDirect

Operations Research Letters

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

Finite horizon linear quadratic dynamic games for discrete-time stochastic systems with N-players[☆]

Huai-Nian Zhu [*], Cheng-Ke Zhang

School of Economics and Commence, Guangdong University of Technology, Guangzhou, Guangdong 510520, China

ARTICLE INFO

Article history:
Received 16 September 2015
Received in revised form
17 February 2016
Accepted 17 February 2016
Available online 26 February 2016

Keywords:
Pareto optimal strategy
Nash strategy
Discrete-time stochastic systems
Matrix-valued equations

ABSTRACT

In this paper, dynamic games for a class of finite horizon linear stochastic system governed by Itô's difference equation are investigated. Particularly, both Pareto and Nash strategies are discussed. After defining the equilibrium condition, sufficient conditions for the existence of the strategy sets are obtained, which are associated with the solvability of the corresponding generalized difference Riccati equations (GDREs). Furthermore, an iterative algorithm is proposed to solve the related GDREs and a simple numerical example is given to show the reliability and usefulness of the considerable results.

1. Introduction

Linear quadratic (for short, LQ hereafter) dynamic games and their applications have been investigated extensively in several reports, see [1,4,15,2]. The concepts of Pareto optimization and a Nash game have their roots in decision making, and they have been applied to various control fields. Although dynamic LQ games provide a very general framework for optimal control, the same concept should also be considered for a wider class of various systems.

Over the last decade, stochastic control problems governed by Itô's equation have attracted considerable research interest. Rami et al. [14] and Huang et al. [6] studied the indefinite stochastic LQ control for discrete-time systems with state and control-dependent noises in finite horizon and infinite horizon, respectively. Zhang et al. [16] made a contribution to H_2/H_∞ control for discrete-time stochastic linear systems with state and disturbance-dependent noise. Recently, stochastic Nash game and Pareto optimal strategy for continuous-time and discrete-time stochastic Itô systems have been widely studied, see [3,7–9,11,12,10] and the references therein. Although many results are available on stochastic dynamic games of Nash and Pareto optimal strategies, they are limited to infinite horizon, corresponding results for finite horizon discrete-time stochastic Itô systems have not reported. Therefore, a general dynamic game with state and control-dependent noise for finite horizon discrete-time stochastic linear systems deserves further study.

In this paper, we concentrate our attentions on the Pareto and Nash strategies for finite horizon discrete-time linear stochastic systems with state and control-dependent noise. It will be shown that for our general systems, the existence conditions of equilibrium strategies are associated with the solvability of the corresponding GDREs.

The rest of the paper is organized as follows. In Section 2, some preliminaries are made; Section 3 contains our main theorems of Pareto and Nash strategies; Section 4 presents a numerical algorithm to solve the corresponding GDREs.

For convenience, throughout this paper we adopt the following traditional notations.

A': the transpose of a matrix A; A^{-1}: the inverse of a matrix A. $A > 0$: the positive definite symmetric matrix A. $\mathbf{R}^n$: the n-dimensional real vector space with the corresponding 2-norm $\|\cdot\|$. $\mathbf{R}^{m\times n}$: the vector space of all $m\times n$ matrices with entries in $\mathbf{R}$. $\mathbf{S}_n(\mathbf{R})$: the set of all real $n\times n$ symmetric matrices; I: the identity matrix; $\mathbf{E}(x)$: the mathematical expectation of x; $\mathbf{N}$: the set of positive integers; $N_t := \{0, 1, 2, \ldots, t\}$.

2. Preliminaries

In this section, we shall give the dynamic systems to be discussed and present some preliminary results that are needed in our later development.

☆ Fully documented templates are available in the elsarticle package on CTAN.

* Corresponding author.
E-mail address: huainian258@163.com (H.-N. Zhu).
URL: http://www.elsevier.com (H.-N. Zhu).

http://dx.doi.org/10.1016/j.orl.2016.02.010
0167-6377/

Consider the following discrete-time stochastic system with N-players involving state and control-dependent noise

$$\begin{cases} x(t+1) = A(t)x(t) + \sum_{i=1}^{N} B_i(t)u_i(t) \\ \qquad + \left[C(t)x(t) + \sum_{i=1}^{N} D_i(t)u_i(t) \right] w(t), \\ x(0) = x_0 \in \mathbf{R}^n, \quad t \in N_T, \end{cases} \tag{1}$$

where $x(\cdot) \in \mathbf{R}^n$, $u_i(\cdot) \in \mathbf{R}^{m_i}$, $i = 1, 2, \ldots, N$, are the system state, and ith control strategy of player i, respectively. $A(\cdot)$ and $C(\cdot)$ are $n \times n$ matrix-valued functions, $B_i(\cdot)$ and $D_i(\cdot)$ are $n \times m_i$ matrix-valued functions. $\{w(t) \in \mathbf{R}, t \in N_T\}$ is a sequence of real random variables defined on a complete probability space $\{\Omega, \mathcal{F}, \mu\}$, which is a wide sense stationary, second-order process with $\mathbf{E}(w(t)) = 0$ and $\mathbf{E}(w(t)w(s)) = \delta_{st}$ with δ_{st} being a Kronecker function. Similar to [13, Definition 3.1.2], we denote the σ-algebra generated by $w(s)$, i.e., $\mathcal{F}_t = \sigma(w(s) : s \in N_t)$.

We assume that the control trajectory u_i belongs to an admissible control space U_i, and denoting the joint control $(u_1(\cdot), \ldots, u_N(\cdot)) \in U_1 \times \cdots \times U_N = \mathcal{U}$, the cost function for each player is defined by

$$J_i(x_0, u_1(\cdot), \ldots, u_N(\cdot)) = \sum_{t=0}^{T} \mathbf{E}\left[x'(t)Q_i(t)x(t) + u_i'(t)R_i(t)u_i(t) \right], \tag{2}$$

where $Q_i(\cdot)$ and $R_i(\cdot)$, $i = 1, 2, \ldots, N$, are $n \times n$ and $m_i \times m_i$ symmetric matrix-valued functions, respectively. The joint admissible control space is taken to be the space of square integrable functions as

$$\mathcal{U} \triangleq L^2_{\mathcal{F}}(N_T, \mathbf{R}^m) = \left\{ u : N_T \times \Omega \to \mathbf{R}^m, \sum_{t=0}^{T} \mathbf{E}\,\|u(t)\|^2 < \infty \right\}.$$

Obviously, for any $T \in \mathbf{N}$ and $(x_0, u_1(\cdot), \ldots, u_N(\cdot)) \in \mathbf{R}^n \times \mathcal{U}$, there exists a unique solution $x(\cdot) \equiv x(\cdot; x_0, u_1(\cdot), \ldots, u_N(\cdot)) \in L^2_{\mathcal{F}}(N_T, \mathbf{R}^m)$ to (1) under some mild conditions on the coefficients. It should be noted that cost function (2) includes $u_i(\cdot)$ only, it is possible to consider more general forms of cost function (for example, $J_i(\cdot)$ also explicitly contains $u_j(\cdot), j \neq i$, etc.). We take the above form for simplicity of the presentation in this paper.

We concentrate on considering a situation in which player i designs his control strategy on the basis of the state information. The design specifications of the player i can be expressed in terms of a cost function $J_i(\cdot)$. First, it is assumed that all the players decide their strategies through mutual cooperation under various constraints. The solution to such a problem is found in the class of Pareto optimal strategies. This means that no deviation from the Pareto optimal strategy can decrease the costs of all players. Second, a major step toward an understanding of non-cooperative games with several players is provided by the newly introduced concept of non-cooperative equilibrium. In this case, a set of strategies is formulated as a Nash equilibrium if, whenever a single player modifies his strategy, his own payoff will not increase [5]. These two strategies are investigated in next section.

First, we introduce a lemma which will be used in our subsequent analysis [16].

Lemma 1. *Consider the following discrete-time stochastic system*

$$\begin{cases} x(t+1) = A(t)x(t) + B(t)u(t) + [C(t)x(t) + D(t)u(t)]\, w(t), \\ x(0) = x_0 \in \mathbf{R}^n, \quad t \in N_T. \end{cases} \tag{3}$$

Suppose $T \in \mathbf{N}$ is given and $P(0), P(1), \ldots, P(T+1)$ is an arbitrary family of matrices in $\mathbf{S}_n(\mathbf{R})$, then for any $x_0 \in \mathbf{R}^n$, we have

$$\sum_{t=0}^{T} \mathbf{E} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}' Q(P(t)) \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} = \mathbf{E}\left[x'(T+1)P(T+1)x(T+1) \right] - x_0'P(0)x_0, \tag{4}$$

where $Q(P(t))$ is given in Box I.

3. Main results

3.1. Pareto optimal strategy

In this subsection, Pareto optimal strategies as one of the cooperative game theory are considered. It is assumed that each player wants to minimize his own cost function described in (2). A more mathematical formulation is given below.

A Pareto solution is a set $(u_1(\cdot), \ldots, u_N(\cdot))$, which minimizes

$$J(x_0, u_1(\cdot), \ldots, u_N(\cdot)) = \sum_{i=1}^{N} \gamma_i J_i(x_0, u_1(\cdot), \ldots, u_N(\cdot)),$$

$$0 < \gamma_i < 1, \quad \sum_{i=1}^{N} \gamma_i = 1 \tag{5}$$

for some γ_i, $i = 1, 2, \ldots, N$ [9].

From the above description, we can see that the stochastic LQ regulator problem is a special case of this problem when the players agree on a choice of γ_i, $i = 1, 2, \ldots, N$, as weight factors.

Now we introduce a type of generalized difference Riccati equation associated with the Pareto optimal strategies.

Definition 1. The following constrained matrix-valued difference equation

$$\begin{cases} -P(t) + \mathcal{N}(P(t+1)) - \mathcal{L}'(P(t+1)) \\ \qquad \times \mathcal{R}^{-1}(P(t+1))\mathcal{L}(P(t+1)) = 0, \\ P(T+1) = 0, \\ \gamma_i R_i(t) + B_i'(t)P(t+1)B_i(t) + D_i'(t)P(t+1)D_i(t) > 0, \\ \quad t \in N_t, \end{cases} \tag{6}$$

where $i = 1, \ldots, N$, $\mathcal{N}(P(t+1))$, $\mathcal{R}(P(t+1))$, and $\mathcal{L}(P(t+1))$ are given in Box II is called a generalized difference Riccati equation (GDRE).

The Pareto optimal strategies are given below.

Theorem 1. *Consider the stochastic system (1) and the cost function (5). If GDRE (6) admits a solution $P(t) > 0$, $t \in N_T$, then the state feedback Pareto strategy set is given below.*

$$\mathbf{u}^*(t) = \begin{pmatrix} u_1^*(t) \\ \vdots \\ u_N^*(t) \end{pmatrix} = K(t)x(t) = -\mathcal{R}^{-1}(P(t+1))\mathcal{L}(P(t+1))x(t). \tag{7}$$

Furthermore,

$$J(x_0, u_1^*(\cdot), \ldots, u_N^*(\cdot)) = x_0'P(0)x_0. \tag{8}$$

Proof. From Lemma 1 and Eq. (6), by using the square completion technique, for any $P(t) \in \mathbf{S}_n(\mathbf{R})$, and $(x_0, u_1(\cdot), \ldots, u_N(\cdot)) \in \mathbf{R}^n \times \mathcal{U}$, we obtain

$J(x_0, u_1(\cdot), \ldots, u_N(\cdot))$