CONF-9006183--1
FG02-88ER13939

STOCHASTIC DIFFERENTIAL GAMES WITH WEAK SPATIAL AND STRONG INFORMATIONAL COUPLING[1]

Tamer Başar[2] and R. Srikant
Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Avenue
Urbana, IL 61801/USA
Tel: (217)-333-3607
Fax: (217) 244-1764

CONF-9006183--1

DE90 011925

Abstract

We formulate a parameterized family of linear quadratic two-person nonzero-sum stochastic differential games where the players are weakly coupled through the state equation and strongly coupled through the measurements. A positive parameter ϵ characterizes this family, in terms of which the subsystems are coupled (weakly). With $\epsilon = 0$ the problem admits a unique Nash equilibrium solution, while for $\epsilon > 0$, no matter how small, no general method is available to obtain the Nash equilibrium solution and even to prove existence and uniqueness. In this paper, we develop an iterative technique whereby Nash solutions of all orders (in terms of ϵ) are obtained by starting the iteration with the unique (strong team) solution determined for $\epsilon = 0$. The Nash solutions turn out to be linear, requiring only finite-dimensional controllers, in spite of the fact that a separation (of estimation and control) result does not hold.

Keywords: Stochastic differential games, Nash equilibria, Weak coupling, stochastic measurements.

[1] Paper submitted for presentation at the *9th Internation Conf. Analysis and Optimization of Systems*, June 12-15, 1990, Antibes, France.

[2] All correspondence on this paper should be sent to this author.

1 Introduction

We formulate a class of stochastic nonzero-sum differential games where the players are weakly coupled through the state equation while sharing the same source of informational (measurement). This latter feature brings in a strong "information coupling" the presence of which makes the derivation of Nash equilibria quite a challenging task unless the weak coupling parameter (ϵ) is set equal to zero. The approach developed in the paper involves an iterative scheme which starts with the solution of the zero'th order stochastic game (obtained by setting $\epsilon = 0$), and requires at each step the solution of individual stochastic control problems each one of which is of the linear-quadratic type. We study the structural properties of these noncertainty-equivalence controllers, as well as the convergence of the proposed iterative scheme.

A precise formulation of the problem is given in the next section, followed (in Section 3) by a summary of the main results to be presented at the Conference.

2 Problem Statement

The linear-quadratic, nonzero-sum, stochastic differential game under consideration, with weak spatial and strong informational coupling between the players, can be defined in precise mathematical terms as follows: The evolution of the composite state ($x := (x^{1'}, x^{2'})'$) of the game is described by the linear Itö stochastic differential equation:

$$\begin{array}{c} dx_t = A(t;\epsilon)x_t dt + \tilde{B}^1(t)u_t^1 dt + \tilde{B}^2(t)u_t^2 dt + F(t)dw_t \\ t_0 \le t \le t_f, \qquad x_{t_0} = x_0 \end{array} \tag{1}$$

where the initial state x_0 is taken to be a Gaussian distributed random vector with mean zero and covariance $\Sigma_0 > 0$, $dim(x^i) = n_i$, $dim(u^i) = r_i$, $i = 1, 2$,

$$A(t;\epsilon) := \begin{pmatrix} A_1 & \epsilon A_{12} \\ \epsilon A_{21} & A_2 \end{pmatrix}(t); \quad \tilde{B}_1(t) := \begin{bmatrix} B_1(t) \\ \cdots \\ 0 \end{bmatrix}; \tilde{B}^2(t) := \begin{bmatrix} 0 \\ \cdots \\ B_2(t) \end{bmatrix} \tag{2}$$

$$F(t) = \text{ block diag } (F_1(t), F_2(t)); \quad F_1F_1' > 0, \quad F_2F_2' > 0,$$

$\epsilon > 0$ is a small (coupling) parameter, and the partitions of A, $\tilde{B}^1$, $\tilde{B}^2$ and F are compatible with the subsystem structure, so that with $\epsilon = 0$ the system decomposes into two completely decoupled and stochastically independent subsystems, each one controlled by a different player. The functions u_t^1 and u_t^2, $t \ge t_0$, represent the controls of Players 1 and 2, respectively, and are vector stochastic processes with continuous sample paths, as to be further explained in the sequel. The driving term $w_t := (w_t^{1'}, w_t^{2'})'$, $t \geq t_0$ is a standard vector Wiener process, that is independent of the initial state x_0.

The common observation y of the players is described by

$$\begin{aligned} dy_t &= C(t)x_t dt + G(t)dv_t, \\ C &:= (C_1, C_2) \end{aligned} \tag{3}$$

where $dim(y) = m$, C_i is $m \times r_i$, $i = 1, 2$, $GG' > 0$, and v_t, $t \geq t_0$ is another standard vector Wiener process, independent of $\{w_t\}$ and x_0. This common observation constitutes the only strong coupling between the players.

All matrices in the above formulation are taken to be continuous on the time interval $[t_0, t_f]$. Let $C_m = C_m[t_0, t_f]$ denote the space of the continuous functions on $[t_0, t_f]$, with values in $I\!R^m$. Further let $\mathcal{Y}_t$ be the sigma-field in C_m generated by the cylinder sets $\{y \in C_m, y_s \in B\}$ where B is a Borel set in $I\!R^m$, and $t_0 \leq s \leq t$. Then, the information gained by each player during the course of the game is completely determined by the information field $\mathcal{Y}_t$, $t \geq t_0$. A permissible strategy for Player i is a mapping $\gamma_i(\cdot, \cdot)$ of $[t_0, t_f] \times C_m$ into $I\!R^{r_i}$ with the following properties:

(i) $\gamma_i(t, \eta)$ is continuous in t for each $\eta \in C_m$;

(ii) $\gamma_i(t, \eta)$ is uniformly Lipschitz in η, i.e., $|\gamma_i(t, \eta) - \gamma_i(t, \xi)| \leq k\|\eta - \xi\|$, $t \in [t_0, t_f]$, η, $\xi \in C_m$, where $\|\cdot\|$ is the sup norm on C_m.

(iii) $u_t^i = \gamma_i(t, \eta)$ is adapted to the information field $\mathcal{Y}_t$.

Let us denote the collection of all strategies described above, for Player i, by Γ_i. It is known that, corresponding to any pair of strategies $\{\gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2\}$, the stochastic differential equation (1) admits a unique solution that is a sample-path-continuous second-order process. As a result, the observation process y_t, $t \geq t_0$ will also have continuous sample paths.

For any pair of strategies $\{\gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2\}$, we introduce the cost function for Player i, $i = 1, 2$, as

$$J_i(\gamma_1, \gamma_2) = E\left\{x_{t_f}' \tilde{Q}_{if}(\epsilon) x_{t_f} + \int_{t_0}^{t_f} (x_t' \tilde{Q}_i(t; \epsilon) x_t + u_t^{i'} u_t^i) dt\right\}$$

where all the matrices are nonnegative definite and

$$\tilde{Q}_{1f}(\epsilon) := \text{block diag } (Q_{1f}, \epsilon Q_{12f})$$

$$\tilde{Q}_{2f}(\epsilon) := \text{block diag } (\epsilon Q_{21f}, Q_{2f})$$

$$\tilde{Q}_1(t;\epsilon) := \text{block diag } (Q_1(t), \epsilon Q_{12}(t))$$

$$\tilde{Q}_2(t;\epsilon) := \text{block diag } (\epsilon Q_{21}(t), Q_2(t)).$$

Furthermore, $u_t^i = \gamma_i(t, y_0^t)$, with y_0^t denoting the stochastic process restricted to the time interval $[t_0, t]$. Note that the players' costs are also coupled weakly, so that if $\epsilon = 0$ each cost function involves only that player's state vector and control function. Of course, even with $\epsilon = 0$ there is still an "informational coupling" through the common observation, which implicitly couples the cost functions under any equilibrium solution concept.

Adopting the Nash equilibrium solution concept, we seek a pair of strategies $(\gamma_1^* \in \Gamma_1, \gamma_2^* \in \Gamma_2)$ satisfying the pair of inequalities

$$J_1(\gamma_1^*, \gamma_2^*) \leq J_1(\gamma_1, \gamma_2^*); \quad J_2(\gamma_1^*, \gamma_2^*) \leq J_2(\gamma_1^*, \gamma_2) \tag{4}$$

for all $\gamma_1 \in \Gamma_1$, $\gamma_2 \in \Gamma_2$. To show the explicit dependence of the Nash policies on the available information and the coupling parameter ϵ, we will sometimes use the notation $\gamma_i^*(t, y_0^t; \epsilon)$.

Let us first list a few known facts on the Nash solution of this stochastic differential game, when ϵ is not necessarily a small parameter.

1. Conditions under which a Nash equilibrium exists are not known. What is known, however, is that the solution (whenever it exists) will not satisfy any separation principle (between estimation and control), which is in some sense true even for the zero-sum version of the problem [1].

2. The discrete-time version of the problem, but with private measurements for the players that are shared with a delay of one time unit, has been considered before in [2] where it has been shown that the Nash equilibrium solution is unique and linear in the available information for each player. The procedure developed there can readily be used to derive a similar result for the common information case (in discrete-time), when even though a separation result does not apply, the Nash controllers for the players have finite dimensional representations (i.e., the controller dimension does not grow with the number of stages in the problem) [3].

3. For the continuous-time problem, however, the procedure of [2] does not apply, and consequently a proof of existence and uniqueness of linear Nash equilibria has been quite elusive for the past decade. For the zero-sum version (of the continuous-time problem), it is possible to prove existence of a unique linear saddle-point equilibrium, though using an indirect approach that employs explicitly the interchangeability property saddle-points [1].

In view of this past experience, we adopt in this paper a different approach for the nonzero-sum stochastic differential game, that exploits the weakness of the coupling between the two subsystems: Suppose that there exists a pair of Nash equilibrium policies that are analytic in ϵ, that is, for every $n \geq 1$,

$$\gamma_i^*(t, y_0^t; \epsilon) = \gamma_i^{(0)}(t, y_0^t) + \sum_{i=1}^{n} \gamma_i^{(n)}(t, y_0^t)\epsilon^n + 0(\epsilon^n). \tag{6}$$

Then the question we raise is whether it is possible to obtain the different terms in this expansion by solving simpler game or stochastic control problems, and whether it is possible to prove unicity of equilibrium in this class (of analytic-in-ϵ policies). Our findings are summarized in the next section.

3 Summary of Main Results

Our first result is that the zero'th order term in the expansion (6) can be obtained by solving the original stochastic differential game after setting $\epsilon = 0$. Note that with $\epsilon = 0$ the game dynamics are completely decoupled, but the players' policy choices are still coupled through the common measurement.

To obtain the Nash equilibrium of the "zero'th order game", we first fix $u_t^2 = \gamma_2^{(0)}(t, y_0^t)$ and minimize J_1 over $\gamma_1 \in \Gamma_1$ subject to (1) with $\epsilon = 0$. This is a standard stochastic control problem, admitting the unique solution

$$\gamma_1^{(0)}(t, y_0^t) = -B_1' S_1 \hat{x}_t^1 \tag{7}$$

where $S_1 \geq 0$ satisfies the Riccati equation

$$\dot{S}_1 + A_1' S_1 + S_1 A_1 - S_1 B_1 B_1' S_1 + Q_1 = 0; \quad S_1(t_f) = Q_{1f} \tag{8}$$

and

$$d\hat{x}_t^1 = (A_1 - B_1 B_1' S_1)\hat{x}_t^1 dt + K_1(dy_t - (C_1 \hat{x}_t^1 + C_2 \hat{x}_t^2)dt); \hat{x}_{t_0}^1 = 0 \tag{9a}$$

$$d\hat{x}_t^2 = A_2\hat{x}_t^2 dt + B_2\gamma_2^{(0)}(t, y_0^t)dt + K_2(dy_t - (C_1\hat{x}_t^1 + C_2\hat{x}_t^2)dt); \hat{x}_{t_0}^2 = 0 \tag{9b}$$

$$K := \begin{pmatrix} K_1 \\ K_2 \end{pmatrix} = \Sigma C'(GG')^{-1} \tag{9c}$$

$$\dot{\Sigma} - A_0\Sigma - \Sigma A_0' - FF' + KGG'K = 0; \quad \Sigma(t_0) = \Sigma_0 \tag{9d}$$

$$A_0 := \text{ block diag } (A_1, A_2).$$

Note that $\gamma_2^{(0)}$ enters the solution only through the estimator of the second block component of x.

Now we reverse the roles of the players,, fix $\gamma_1^{(0)}$ arbitrarily and minimize J_2 over Γ_2, with $\epsilon = 0$ to arrive again at a unique solution:

$$\gamma_2^{(0)}(t, y_0^t) = -B_2'S_2\hat{x}_t^2 \tag{10}$$

where $S_2 \geq 0$ satisfies

$$\dot{S}_2 + A_2'S_2 + S_2A_2 - S_2B_2B_2'S_2 + Q_2 = 0; \quad S_2(t_f) = Q_{2f} \tag{11}$$

and $\hat{x}_t^2$ is given by (9b) with $\gamma_2^{(0)}$ replaced by the expression in (10). Since (9b) depends on $\hat{x}_t^1$ also, we will need here (9a) with the term $-B_1B_1'S_1\hat{x}_t^1$ replaced by $B_1\gamma_1^{(0)}(t, y_0^t)$. Since (9) and (10) are unique responses, it readily follows that the Nash equilibrium policies are unique, and given by (9) and (10) with $\hat{x}_2^1$ and $\hat{x}_t^2$ given by (9a)–(9b), with $\gamma_2^{(0)}$ in (9b) replaced by the expression in (10). This completes the derivation of the zero'th order solution. It is useful to note that this zero'th order solution is also the unique solution to a team problem with objective function any convex combination of $J_1^{(0)}$ and $J_2^{(0)}$, where $J_i^{(0)}$ is J_1 with $\epsilon = 0$.

To obtain the first order terms, $\gamma_1^{(1)}$ and $\gamma_2^{(1)}$, we perform the following minimizations, which turn out to be the first step in a policy-space iteration algorithm:

$$\min J_1(\gamma_1, \gamma_2^{(0)}); \qquad \min J_2(\gamma_1^{(0)}, \gamma_2)$$

where ϵ is *not* set equal to zero. The solutions to these individual stochastic control problems are unique and analytic in $\epsilon > 0$ – – let us denote them by $\hat{\gamma}_1$ and $\hat{\gamma}_2$, respectively. Then we can show that

$$\hat{\gamma}_1(t, y_0^t; \epsilon) = \gamma_1^{(0)}(t, y_0^t) + \epsilon\gamma_1^{(1)}(t, y_0^t) + 0(\epsilon) \tag{12a}$$

$$\hat{\gamma}_2(t, y_0^t; \epsilon) = \gamma_2^{(0)}(t, y_0^t) + \epsilon\gamma_2^{(1)}(t, y_0^t) + 0(\epsilon) \tag{12b}$$

where the zero'th order terms are precisely the ones given by (7) and (10), and the first-order terms are the ones in the expansion (6). The controllers turn out to be of the same dimension as in the zeroth order case, but this time the control gain matrices depend on the Kalman gain K.

For the next step we substitute the expression for $\gamma_1 = \gamma_1^{(0)} + \epsilon\gamma_1^{(1)}$ into J_2, and that for $\gamma_2 = \gamma_2^{(0)} + \epsilon\gamma_2^{(1)}$ into J_1, and minimize them with respect to γ_2 and γ_1, respectively. The result is the unique solution (as the counterpart of (12)):

$$\hat{\gamma}_1(t, y_0^t; \epsilon) = \gamma_1^{(0)}(t, y_0^t) + \sum_{i=1}^{2} \epsilon^i \gamma_1^{(i)}(t, y_0^t) + 0(\epsilon^2)$$

$$\hat{\gamma}_2(t, y_0^t; \epsilon) = \gamma_2^{(0)}(t, y_0^t) + \sum_{i=1}^{2} \epsilon^i \gamma_2^{(i)}(t, y_0^t) + 0(\epsilon^2)$$

where all the terms (zero'th, first and second-order) yield exactly the corresponding terms in (6) (details of verification of this result will be provided in the final version of the paper at the Conference). Hence, following this procedure iteratively, we arrive at policies of all orders, to yield approximations to the Nash policies, to any degree of accuracy. In all cases the controllers are linear and of the same order as the zero'th order Nash controllers.

In addition to providing the precise expressions for these linear controllers, the final version of the paper to be presented at the Conference will also include a study of the convergence of the generated sequence of controllers, and a characterization of the limiting Nash policies.

References

[1] T. Başar, "On the saddle-point solution of a class of stochastic differential games," *Journal on Optimization Theory and Applications*, vol. 33, pp. 539–556, April 1981.

[2] T. Başar, "Two-criteria LQG decision problems with one-step delay observation sharing pattern," *Information and Control*, vol. 32, no. 1, pp. 21–50, 1978.

[3] T. Başar, *Theory of Stochastic Dynamic Noncooperative Games.* Oxford, England: Basil Blackwell, 1990.