Differential Equations and Dynamical Systems, Vol.9,Nos.1 & 2, January & April 2001, pp.105-112.

A Note on Stochastic Minimax Principle*

Mrinal K. Ghosh[†] K. Suresh Kumar,

A.K. Nandakumaran and K.S. Mallikarjuna Rao[‡]

(Submitted by: V. Sree Hari Rao)

Abstract

We study zero-sum stochastic differential games of fixed duration where the state is given by controlled (possibly) degenerate diffusions. Using the framework of relaxed strategies, we derive a stochastic minimax principle.

1 Introduction

We study a zero-sum stochastic differential game on the finite horizon where the state $X(\cdot)$ is an $\mathbb{R}^d$-valued controlled degenerate diffusion given

[0]Received on April 15, 2001.

[0]AMS (MOS) 2000 Subject classifications: 90D25, 90D26.

*This work is partially supported by a grant from Department of Science and Technology, Government of India, Project No. DST/MS/111/99.

[†]Corresponding author: Email: mkg@math.iisc.ernet.in, Fax: 91-80-3600146.

[‡]The author is a CSIR research fellow during this work.

[0]

by

$$\begin{aligned} dX(t) &= b(t, X(t), u_1(t), u_2(t))\, dt + \sigma(X(t))\, dW(t), \quad t \in (0, T] \\ X(0) &= x, \end{aligned} \tag{1.1}$$

where $b : [0, T] \times \mathrm{R}^d \times U_1 \times U_2 \to \mathrm{R}^d$, U_1 and U_2 are action sets of players 1 and 2 respectively; $W(\cdot)$ is a standard d-dimensional Brownian motion; $\sigma : \mathrm{R}^d \to \mathrm{R}^{d\times d}$; $u_i : [0, T] \to U_i$ is progressively measurable with respect to the σ-field generated by $W(\cdot)$ which is the strategy of player $i, i = 1, 2$. Player 1 tries to maximize his expected payoff

$$R(x, u_1(\cdot), u_2(\cdot)) := E\Big[\int_0^T r(t, X(t), u_1(t), u_2(t))\, dt + g(X(T))\Big] \tag{1.2}$$

over his strategies $u_1(\cdot)$, whereas player 2 tries to minimize the same over his strategies $u_2(\cdot)$. Here r is the running payoff function and g is the terminal payoff function. Precise conditions on b, σ, g, u_1, u_2 will be given in next section. The stochastic differential game (SDG for short) has a value if

$$\inf_{u_2(\cdot)} \sup_{u_1(\cdot)} R(x, u_1(\cdot), u_2(\cdot)) = \sup_{u_1(\cdot)} \inf_{u_2(\cdot)} R(x, u_1(\cdot), u_2(\cdot)).$$

A strategy $u_1^*(\cdot)$ is said to be optimal for player 1 if

$$R(x, u_1^*(\cdot), \tilde{u}_2(\cdot)) \geq \inf_{u_2(\cdot)} \sup_{u_1(\cdot)} R(x, u_1(\cdot), u_2(\cdot))$$

for any strategy $\tilde{u}_2(\cdot)$ of player 2. Similarly, a strategy $u_2^*(\cdot)$ is said to be optimal for player 2 if

$$R(x, \tilde{u}_1(\cdot), u_2^*(\cdot)) \leq \sup_{u_1(\cdot)} \inf_{u_2(\cdot)} R(x, u_1(\cdot), u_2(\cdot))$$

for any strategy $\tilde{u}_1(\cdot)$ of player 1. A pair of optimal strategies for both players satisfies

$$R(x, u_1(\cdot), u_2^*(\cdot)) \leq R(x, u_1^*(\cdot), u_2^*(\cdot)) \leq R(x, u_1^*(\cdot), u_2(\cdot))$$

for any pair of strategies $(u_1(\cdot), u_2(\cdot))$ of the players. Thus $(u_1^*(\cdot), u_2^*(\cdot))$ constitutes a saddle point equilibrium. Conversely, a pair of saddle point strategies $(u_1^*(\cdot), u_2^*(\cdot))$ is clearly a pair of optimal strategies for both players.

In this note, we establish a (stochastic) minimax principle in the framework of relaxed strategies which gives necessary conditions for optimality. This framework is explained in the next section. Our paper is organized as follows. In Section 2, we introduce the basic notation and assumptions. The minimax principle is derived in Section 3.

2 Preliminaries

Let $U_i, i = 1, 2$, be given compact metric spaces and $\mathcal{M}_i = \mathcal{P}(U_i)$, the space of probability measures on U_i. Let $T > 0$ be fixed. Let

$$\bar{b} : [0, T] \times \mathrm{R}^d \times U_1 \times U_2 \to \mathrm{R}^d$$

and

$$\sigma : [0, T] \times \mathrm{R}^d \to \mathrm{R}^{d \times d}.$$

We assume that
(A1) the functions $\bar{b}$ and σ are continuous and there exists a constant $C_1 > 0$ such that

$$|\bar{b}(t, x, u_1, u_2) - \bar{b}(s, y, u_1, u_2)| + |\sigma(t, x) - \sigma(s, y)| \leq C_1(|t - s| + |x - y|)$$

for all $(u_1, u_2) \in U_1 \times U_2$ and $(t, x), (s, y) \in [0, T] \times \mathrm{R}^d$. Define

$$b : [0, T] \times \mathrm{R}^d \times \mathcal{M}_1 \times \mathcal{M}_2 \to \mathrm{R}^d$$

by

$$b(t, x, \mu_1, \mu_2) = \int_{U_2} \int_{U_1} \bar{b}(t, x, u_1, u_2)\ \mu_1(du_1)\ \mu_2(du_2).$$

The state of the system $X(\cdot)$ evolves according to the controlled stochastic differential equation of Ito type

$$\begin{aligned} dX(t) &= b(t, X(t), \mu_1(t), \mu_2(t))\ dt + \sigma(t, X(t))\ dW(t), \quad t \in (0, T], \\ X(0) &= x \in \mathrm{R}^d. \end{aligned} \tag{2.1}$$

Here $W(\cdot)$ is a standard d-dimensional Wiener process; $\mu_i(\cdot)$ is an $\mathcal{M}_i$-valued process which is progressively measurable with respect to σ-fields

$\mathcal{F}_t^W := \sigma(W(t') : 0 \le t' \le t)$. The process $\mu_i(\cdot)$ is called an (admissible) relaxed strategy for player i. Let $\mathcal{A}_i$ denote the set of all admissible strategies for player i. A relaxed strategy $\mu_i(\cdot)$ of player i is called a pure strategy if $\mu_i(\cdot)$ is a Dirac measure, i.e., $\mu_i(\cdot) = \delta_{u_i(\cdot)}$, where $u_i(\cdot)$ is a U_i-valued process.

Let $\bar{r} : [0,T] \times \mathrm{R}^d \times U_1 \times U_2 \to \mathrm{R}$ be the running payoff function. If x is the state at time t and the players choose actions $(u_1, u_2) \in U_1 \times U_2$, then player 1 receives a payoff $\bar{r}(t, x, u_1, u_2)$ from player 2 at time t. Let $g : \mathrm{R}^d \to \mathrm{R}$ denote the terminal payoff function.

We assume that:

(A2) (i) the functions $\bar{r}$ and g are bounded and continuous;

(ii) there exist constants $C_2 > 0$, $C_3 > 0$, such that

$$|\bar{r}(t,x,u_1,u_2) - \bar{r}(s,y,u_1,u_2)| \le C_2(|t-s| + |x-y|),$$

$$|g(x) - g(y)| \le C_3|x-y|,$$

for all $t, s \in [0,T]$, $x, y \in \mathrm{R}^d$, $(u_1, u_2) \in U_1 \times U_2$.

Let $r : [0,T] \times \mathrm{R}^d \times \mathcal{M}_1 \times \mathcal{M}_2 \to \mathrm{R}$ be defined by

$$r(t,x,\mu_1,\mu_2) = \int_{U_2}\int_{U_1} \bar{r}(t,x,u_1,u_2)\ \mu_1(du_1)\ \mu_2(du_2).$$

When the initial state of the system at time t is x and the players use relaxed strategies $(\mu_1(\cdot), \mu_2(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2$, then the payoff of player 1 is given by

$$R(x,\mu_1(\cdot),\mu_2(\cdot)) = E\left[\int_0^T r(s,X(s),\mu_1(s),\mu_2(s))\ ds + g(X(T))\right], \tag{2.2}$$

where $X(\cdot)$ is given by (2.1). A relaxed strategy $\mu_1^*(\cdot) \in \mathcal{A}_1$ is said to be optimal for player 1, if

$$R(x,\mu_1^*(\cdot),\mu_2(\cdot)) \ge \inf_{\mu_2(\cdot)\in\mathcal{A}_2}\ \sup_{\mu_1(\cdot)\in\mathcal{A}_1} R(x,\mu_1(\cdot),\mu_2(\cdot)),$$

for any $\mu_2(\cdot) \in \mathcal{A}_2$. Similarly, a relaxed strategy $\mu_2^*(\cdot) \in \mathcal{A}_2$ is said to be optimal for player 2, if

$$R(t,x,\mu_1(\cdot),\mu_2^*(\cdot)) \le \sup_{\mu_1(\cdot)\in\mathcal{A}_1}\ \inf_{\mu_2(\cdot)\in\mathcal{A}_2} R(x,\mu_1(\cdot),\mu_2(\cdot)),$$

for any $\mu_1(\cdot) \in \mathcal{A}_1$. Thus a pair of optimal strategies constitute a saddle point equilibrium.

Remark 2.1. Since $\mu_i(\cdot)$ is progressively measurable with respect to the σ-field generated by $W(\cdot)$, there exists a progressively measurable function f_i such that $\mu_i(t) = f_i(t, W(\cdot)), 0 \leq t \leq T$. Thus the player 1 chooses the function f_1, whereas the player 2 chooses f_2. This way the noncooperative nature of the game is preserved.

3 Stochastic Minimax Principle

In this section, we derive a stochastic minimax principle. We make the following additional assumptions.

(A3) (i) For each $(t, u_1, u_2) \in [0, T] \times U_1 \times U_2$, the functions $\bar{b}(t, \cdot, u_1, u_2)$, $\bar{r}(t, \cdot, u_1, u_2)$ and $\sigma(\cdot)$ are continuously differentiable.

(ii) The function g is continuously differentiable.

Let $(\mu_1^*(\cdot), \mu_2^*(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2$ be a pair of optimal strategies and $X^*(\cdot)$ the process corresponding to $(\mu_1^*(\cdot), \mu_2^*(\cdot))$ with initial condition $X^*(0) = x$. We define the following "adjoint" process $p(\cdot)$ given by

$$\begin{aligned} dp(t) &= -b_x'(t, X^*(t), \mu_1^*(t), \mu_2^*(t))\ p(t)\ dt \\ &\quad -D_x r(t, X^*(t), \mu_1^*(t), \mu_2^*(t))\ dt \\ p(T) &= D_x g(X^*(T)), \end{aligned} \tag{3.1}$$

where b_x denotes the Jacobian matrix and b_x', its adjoint. Let $\varepsilon > 0$ and $\mu_1(\cdot) \in \mathcal{A}_1$. Let $X^\varepsilon(\cdot)$ denote the process

$$dX^\varepsilon(t) = b(t, X^\varepsilon(t), \mu_1^\varepsilon(t), \mu_2^*(t))\ dt + \sigma(t, X^\varepsilon(t))\ dW(t), \quad X^\varepsilon(0) = x, \tag{3.2}$$

where $\mu_1^\varepsilon(\cdot) = \mu_1^*(\cdot) + \varepsilon(\mu_1(\cdot) - \mu_1^*(\cdot))$. Then we can prove the following lemma by invoking Gronwall's lemma. We omit the details.

Lemma 3.1. Assume (A1) - (A3). Let $Z(\cdot)$ denote the process given by

$$\begin{aligned} dZ(t) &= b_x(t, X^*(t), \mu_1^*(t), \mu_2^*(t))\ Z(t)\ dt + \sigma_x(X^*(t))\ Z(t)\ dW(t) \\ &\quad +(b(t, X^*(t), \mu_1(t), \mu_2^*(t)) - b(t, X^*(t), \mu_1^*(t), \mu_2^*(t))\ dt, \\ Z(0) &= 0. \end{aligned} \tag{3.3}$$

Then

$$E\left|\frac{1}{\varepsilon}\int_0^T (r(s, X^\varepsilon(s), \mu_1^\varepsilon(s), \mu_2^*(s)) - r(s, X^*(s), \mu_1^*(s), \mu_2^*(s)))\, ds - \int_0^T [D_x r(s, X^*(s), \mu_1^*(s), \mu_2^*(s)) \cdot Z(s) + r(s, X^*(s), \mu_1(s), \mu_2^*(s)) - r(s, X^*(s), \mu_1^*(s), \mu_2^*(s))] ds\right| \to 0 \quad \text{as} \quad \varepsilon \to 0.$$

Let $G : [0,T] \times \mathrm{R}^d \times \mathrm{R}^d \times \mathcal{M}_1 \times \mathcal{M}_2 \to \mathrm{R}$ be defined by

$$G(t, x, p, \mu_1, \mu_2) = b(t, x, \mu_1, \mu_2) \cdot p + r(t, x, \mu_1, \mu_2).$$

We, now, prove the following minimax principle.

Theorem 3.2. Assume (A1) - (A3). Let $(\mu_1^*(\cdot), \mu_2^*(\cdot), X^*(\cdot))$ be as above. Then for a.e. $t \in [0,T]$,

$$\begin{aligned}\min_{\mu_2 \in \mathcal{M}_2} \max_{\mu_1 \in \mathcal{M}_1} \quad & E[G(t, X^*(t), p(t), \mu_1, \mu_2)] \\ &= \max_{\mu_1 \in \mathcal{M}_1} E[G(t, X^*(t), p(t), \mu_1, \mu_2^*(t))] \\ &= \max_{\mu_1 \in \mathcal{M}_1} \min_{\mu_2 \in \mathcal{M}_2} E[G(t, X^*(t), p(t), \mu_1, \mu_2)] \\ &= \min_{\mu_2 \in \mathcal{M}_2} E[G(t, X^*(t), p(t), \mu_1^*(t), \mu_2)].\end{aligned} \tag{3.4}$$

Proof. Using Lemma 3.1, we have

$$\begin{aligned}&\frac{d}{d\varepsilon} R(x, \mu_1^\varepsilon(\cdot), \mu_2^*(\cdot))\Big|_{\varepsilon=0} \\ &\quad = E\Big[\int_0^T \Big[D_x r(t, X^*(t), \mu_1^*(t), \mu_2^*(t)) \cdot Z(t) + r(t, X^*(t), \mu_1(t), \mu_2^*(t)) - \\ &\quad -r(t, X^*(t), \mu_1^*(t), \mu_2^*(t))\Big] dt + D_x g(X^*(T)) \cdot Z(T)\Big].\end{aligned} \tag{3.5}$$

Now

$$d(p(t) \cdot Z(t)) = Z(t) \cdot dp(t) + p(t).dZ(t).$$

Integrating the above from 0 to T, taking expectation and then rearranging the terms using (3.1) and (3.3), it follows that

$$\begin{aligned}&E\Big[\int_0^T D_x r(t, X^*(t), \mu_1^*(t), \mu_2^*(t)) \cdot Z(t)\, dt + D_x g(X^*(T)) \cdot Z(T)\Big] \\ &\quad = E\Big[\int_0^T (b(t, X^*(t), \mu_1(t), \mu_2^*(t)) - b(t, X^*(t), \mu_1^*(t), \mu_2^*(t))) \cdot p(t)\, dt\Big].\end{aligned}$$

Hence by (3.5), we obtain

$$\frac{d}{d\varepsilon}R(x,\mu_1^\varepsilon(\cdot),\mu_2^\star(\cdot))\Big|_{\varepsilon=0} = E\Big[\int_0^T\Big\{G(t,X^\star(t),p(t),\mu_1(t),\mu_2^\star(t)) - -G(t,X^\star(t),p(t),\mu_1^\star(t),\mu_2^\star(t))\Big\}\,dt\Big]. \quad (3.6)$$

Since $(\mu_1^\star(\cdot),\mu_2^\star(\cdot))$ is an optimal pair, we have

$$E\Big[\int_0^T\Big\{G((t,X^\star(t),p(t),\mu_1(t),\mu_2^\star(t)) -G(t,X^\star(t),p(t),\mu_1^\star(t),\mu_2^\star(t))\Big\}\,dt\Big] \le 0. \quad (3.7)$$

Let $\mu_1(\cdot)\in\mathcal{A}_1$ be given by

$$\mu_1(s) = \begin{cases}\mu_1, & s\in(t,t+\varepsilon)\\ \mu_1^\star(s), & s\in[0,t]\cup[t+\varepsilon,T]\end{cases}$$

where $\mu_1\in\mathcal{M}_1$ is arbitrary. Since $\varepsilon>0$ is arbitrary, by (3.7), it follows that for any $\mu_1\in\mathcal{M}_1$, a.e. t

$$E[G(t,X^\star(t),p(t),\mu_1,\mu_2^\star(t))] \le E[G(t,X^\star(t),p(t),\mu_1^\star(t),\mu_2^\star(t))]. \quad (3.8)$$

Similarly, we can show that for any $\mu_2\in\mathcal{M}_2$, a.e. t

$$E[G(t,X^\star(t),p(t),\mu_1^\star(t),\mu_2^\star(t))] \le E[G(t,X^\star(t),p(t),\mu_1^\star(t),\mu_2)]. \quad (3.9)$$

From (3.8) and (3.9), the desired result follows.

Remark 3.3. We refer to [1], [2] for stochastic maximum principle for stochastic optimal control problem. Here we have adapted the arguments in [1]. There is, however, a crucial technical difference. Note that in [1], the maximum principle involves the derivative with respect to the control variable, whereas in the framework of relaxed strategies the derivative term with respect to the control variable does not appear. See [3] for a detailed discussion on stochastic maximum principle.

References

[1] A. Bensoussan,"Lectures on Stochastic Control", Part I, Lecture Notes in Math., # 972, Springer-Verlag, 1983, pp. 1–39.

[2] U.G. Haussman," A Stochastic Maximum Principle for Optimal Control of Diffusions, Pitman, 1986.

[3] X.Y. Zhou, A unified treatment of maximum principle and dynamic programming in stochastic controls, Stoch. and Stoch. Reports, 36(1991), 137–161.

Mrinal K. Ghosh, A.K. Nandakumaran and K.S. Mallikarjuna Rao: Department of Mathematics, Indian Institute of Science, Bangalore - 560 012, India.

K. Suresh Kumar: Department of Mathematics, Indian Institute of Technology, Powai, Mumbai - 400 076, India.