A weak dynamic programming principle for zero-sum stochastic differential games

Pedro Vitória, Diogo A. Gomes (Adviser)

June 29, 2010

Abstract

We extend a weak version of the dynamic programming principle, first proven in [1] for stochastic control problems, to the context of zero-sum stochastic differential games. By doing so we are able to derive the Hamilton-Jacobi-Bellman-Isaacs equation when one of the players is allowed to use strategies taking values in an unbounded set.

Keywords: stochastic differential games, value function, dynamic programming principle, viscosity solutions.

Contents

1 Introduction

In [1], Bouchard and Touzi propose a weak version of the dynamic programming principle in the context of stochastic optimal control. Their objective was to avoid technical difficulties related to the measurable selection argument. By doing this they were able to derive the dynamic programming equation without requiring the value function to be measurable. One question that arises naturally is how to extend this approach to stochastic differential games.

Zero-sum stochastic differential games were studied rigorously for the first time by Fleming and Souganidis in [2]. These problems are usually studied in a setting where the strong assumptions imply that the value function is continuous, hence measurable. Considering a weak version of the dynamic programming principle gives us the opportunity of studying these problems in a more general setting where the value function does not have *a priori* much regularity.

This paper tackles the problem of extending the weak dynamic programming principle of [1] to the context of stochastic differential games.

2 Preliminaries

In this paper we consider the *classical Wiener space* $(\Omega, \mathcal{F}, \mathbb{P})$, on a finite horizon T. More precisely, Ω is the space of continuous functions $\omega : [0,T] \to \mathbb{R}^N$ such that $\omega(0) = 0$, and $\mathcal{F}$ is the Borel $\sigma-$algebra in Ω completed with respect to the *Wiener measure* $\mathbb{P}$. We denote by W the *standard $N-$dimensional Brownian motion* corresponding to the coordinate process, $W_s(\omega) = \omega_s$.

We consider in $(\Omega, \mathcal{F}, \mathbb{P})$ the natural *filtration induced* by W augmented with the $\mathbb{P}-$null sets, $\mathbb{F} = \{\mathcal{F}_s, 0 \leq s \leq T\}$, where

$$\mathcal{F}_s = \sigma\{W_r : r \leq s\} \vee \mathcal{N}_{\mathbb{P}}.$$

We denote by $\mathcal{T}$ the collection of all stopping times in $\mathbb{F}$. Given τ_1, τ_2 such that $\tau_1 \leq \tau_2$, $\mathcal{T}_{[\tau_1,\tau_2]}$ denotes the collection of all $\tau \in \mathcal{T}$ such that $\tau_1 \leq \tau \leq \tau_2$. When $\tau_1 = 0$ we simply write $\mathcal{T}_{\tau_2}$.

The essential extrema of a random variable is defined as

Definition 1. *Let X be a random variable. Then $M \in \mathbb{R} \cup \{+\infty\}$ is said to be the* essential supremum *of X, $M = \operatorname{esssup} X$, if:*

- $X \leq M \;\; \mathbb{P}-\text{a.s.}$*;*
- *If there exists $\tilde{M} \in \mathbb{R}$ such that $X \leq \tilde{M}$, $\mathbb{P}-\text{a.s.}$, then $M \leq \tilde{M}$.*

The essential infimum *of X is defined as*

$$\operatorname{essinf} X := -\operatorname{esssup} - X.$$

We consider processes in the following space

$$\mathbb{H}^\infty(t,T;A) := \left\{ \psi \in \mathbb{H}^{p,\infty}(t,T;A) : \operatorname{esssup} \sup_{s\in[t,T]} |\psi_s| < \infty \right\}.$$

3 The Markovian scenario

We consider *zero-sum stochastic differential games* between two players. There is a controlled state process that determines the reward of each of them. This process is a mapping, taking values in $\mathbb{R}^d$, satisfying the following stochastic differential equation (SDE):

$$\left\{ \begin{array}{lcl} dX^{a,b}_{t,x}(s) & = & \mu\left(s, X^{a,b}_{t,x}(s); a_s, b_s\right) ds + \sigma\left(s, X^{a,b}_{t,x}(s); a_s, b_s\right) dW_s \\ X^{a,b}_{t,x}(t) & = & x, \end{array} \right. \tag{1}$$

where

$$\mu : \mathbb{S} \times A \times B \to \mathbb{R}^d,$$
$$\sigma : \mathbb{S} \times A \times B \to \mathbb{R}^{d\times N},$$

are continuous functions satisfying the typical global Lipschitz and linear growth conditions that assure existence and uniqueness of a strong solution for (1):

$$\left\{ \begin{array}{l} |\mu(t,x;a,b) - \mu(t,y;a,b)| + |\sigma(t,x;a,b) - \sigma(t,y;a,b)| \leq K|x-y| \\ |\mu(t,x;a,b)| + |\sigma(t,x;a,b)| \leq K(1 + |x| + |a| + |b|). \end{array} \right. \tag{2}$$

We consider the following spaces of *admissible controls* for players 1 and 2:

$$\mathcal{A}_t := \mathbb{H}^\infty(t,T;A), \;\; \mathcal{B}_t := \mathbb{H}^\infty(t,T;B).$$

The following definition is useful to compare controls:

Definition 2. *Given controls $a, \tilde{a}$, stopping times τ_1, τ_2, we write*

$$a \equiv \tilde{a} \text{ on } [\tau_1, \tau_2] \tag{3}$$

to mean

$$\mathbb{P}(\{X_s = Y_s; s \text{ a.e. on } [\tau_1, \tau_2]\}) = 1.$$

If $a \equiv \tilde{a}$ on $[t, T]$ we say that a and $\tilde{a}$ are equivalent.

The *terminal reward* of player 1 is the random variable

$$J(t, x; a, b) := \mathbb{E}\left[f\left(X_{t,x}^{a,b}(T) \right) \Big| \mathcal{F}_t \right],$$

where f is a measurable function with polynomial growth.

We now introduce the concept of non-anticipative strategy.

Definition 3. *A* strategy *for player 2 is a function $\beta : \mathcal{A}_t \to \mathcal{B}_t$ that maps equivalent controls to equivalent controls.*

A strategy, β, is said to verify the non-anticipativity property *if for every $a_1, a_2 \in \mathcal{A}_t$, $\tau \in \mathcal{T}_{[t,T]}$ we have:*

$$a_1 \equiv a_2 \text{ on } [t, \tau] \Rightarrow \beta[a_1] \equiv \beta[a_2] \text{ on } [t, \tau].$$

The space of non-anticipative strategies for player 2 is denoted by $\Delta(t)$.

The definition of strategy for player 1 is analogous. The space of non-anticipative strategies for player 1 is denoted by $\Gamma(t)$.

We can now define the upper and lower values of a stochastic differential game. In the lower value we allow player 2 to use strategies. This gives him an advantage over player 1. In the upper value we have the opposite situation. More precisely, we have the following:

Definition 4. *The* lower value *of a stochastic differential game is defined as*

$$V(t, x) = \operatorname*{essinf}_{\beta \in \Delta(t)} \operatorname*{esssup}_{a \in \mathcal{A}_t} J(t, x; a, \beta[a]).$$

Similarly, the upper value *of a stochastic differential game is*

$$U(t, x) = \operatorname*{esssup}_{\alpha \in \Gamma(t)} \operatorname*{essinf}_{b \in \mathcal{B}_t} J(t, x; \alpha[b], b).$$

The name upper and lower values is justified by the following inequality, which is intuitive but needs to be proved,

$$V(t, x) \leq U(t, x).$$

This result is established in an indirect way through the uniqueness of solution of the Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation (see [3]).

These value functions, which are *a priori* random variables, are in fact deterministic. More precisely:

Theorem 5. *The random variables $V(t, x)$ and $U(t, x)$ are constant, i.e.,*

$$V(t, x) = \mathbb{E}[V(t, x)],$$
$$U(t, x) = \mathbb{E}[U(t, x)].$$

Proof. See [3]. □

Our objective is to establish that V, U are solutions of the associated HJBI equations, using a weak version of the dynamic programming principle.

4 Weak dynamic programming principle

The dynamic programming principle states that

$$\begin{aligned}
V(t,x) &= \operatorname*{essinf}_{\beta\in\Delta(t)} \operatorname*{esssup}_{a\in\mathcal{A}_t} \mathbb{E}\left[V\left(\theta, X_{t,x}^{a,\beta[a]}(\theta)\right)\middle|\mathcal{F}_t\right], \\
U(t,x) &= \operatorname*{esssup}_{\alpha\in\Gamma(t)} \operatorname*{essinf}_{b\in\mathcal{B}_t} \mathbb{E}\left[U\left(\theta, X_{t,x}^{\alpha[b],b}(\theta)\right)\middle|\mathcal{F}_t\right].
\end{aligned}$$

Obviously, for this statement to have mathematical sense, V, U need to be measurable functions. Proving this requires in general deep measurable selection arguments.

We avoid this technical problems by considering a weak version of this principle:

Theorem 6 (Weak dynamic programming principle)**.** *Suppose that A is compact and f is locally Lipschitz.*

Let $\phi : \mathbb{S} \to \mathbb{R}$ be a continuous function and $\theta^{a,b}$ be a non-anticipative controlled stopping time[1] *such that*

$$W(t,x;a,b) := \mathbb{E}\left[\phi\left(\theta^{a,b}, X_{t,x}^{a,b}(\theta^{a,b})\right)\middle|\mathcal{F}_t\right]$$

makes sense for every $a \in \mathcal{A}_t, b \in \mathcal{B}_t$.

Under some technical assumptions[2]*, it follows that:*

1. *If $\phi \geq V$ then*

$$V(t,x) \leq \operatorname*{essinf}_{\beta\in\Delta(t)} \operatorname*{esssup}_{a\in\mathcal{A}_t} W(t,x;a,\beta[a]). \tag{4}$$

2. *If $\phi \leq V$ then*

$$V(t,x) \geq \operatorname*{essinf}_{\beta\in\Delta(t)} \operatorname*{esssup}_{a\in\mathcal{A}_t} W(t,x;a,\beta[a]). \tag{5}$$

Proof. See [3]. □

5 Hamilton-Jacobi-Bellman-Isaacs equation

We know from Theorem 5 that $V(t,x)$ is deterministic, that is, we can think of it as a function $V : \mathbb{S} \to \mathbb{R}$. In this section we use the weak dynamic programming principle to give a PDE characterization for V by means of a HJBI equation.

We introduce the Hamiltonians $H^{\pm} : \mathbb{S} \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \to \mathbb{R}$ by

$$\begin{aligned}
H^-(t,x,p,X) &= \inf_{a\in A} \sup_{b\in B} H^{a,b}(t,x,p,X), \\
H^+(t,x,p,X) &= \sup_{b\in B} \inf_{a\in A} H^{a,b}(t,x,p,X),
\end{aligned}$$

where

$$H^{a,b}(t,x,p,X) = -\langle \mu(t,x;a,b), p\rangle - \frac{1}{2}\mathrm{Tr}\left(\left(\sigma\sigma^T\right)(t,x;a,b)X\right).$$

Using Theorem 6 we can prove the next important Theorem. An analogous result holds for the upper value, U.

[1]For the definition of non-anticipative controlled stopping time see [3].
[2]For further details see [3].

Theorem 7. *Suppose A is compact. Then:*

1. *V^* is a viscosity subsolution of*

$$\begin{cases} -\partial_t V^* + H^-(., DV^*, D^2V^*) \leq 0, & on \quad (0,T)\times\mathbb{R}^d \\ -V^* \leq -f & on \quad \{T\}\times\mathbb{R}^d; \end{cases}$$

2. *V_* is a viscosity supersolution of*

$$\begin{cases} -\partial_t V_* + (H^-)^*(., DV_*, D^2V_*) \geq 0, & on \quad (0,T)\times\mathbb{R}^d \\ -V_* \geq -f & on \quad \{T\}\times\mathbb{R}^d. \end{cases}$$

Proof. See [3]. □

Uniqueness for the above equation is proved in the space of functions

$$\Theta \;=\; \Big\{ w : \mathbb{S} \to \mathbb{R} : \exists \tilde{C} > 0 \text{ such that } \lim_{|x|\to+\infty} w(t,x) \exp\Big(-\tilde{C}\Psi(x)\Big) = 0, \text{ uniformly in } t \Big\},$$

where

$$\Psi(x) := \big(\log(|x|^2+1)+1\big)^2 .$$

However, only the case where A, B are both compact is studied. This result follows from the following comparison principle:

Theorem 8 (Comparison principle for the HJBI equation)**.** *Suppose A, B are compact. If $v_1, v_2 \in \Theta$ are continuous functions such that v_1 is a viscosity subsolution and v_2 is a viscosity supersolution of*

$$\begin{cases} -\partial_t V + H^-(., DV, D^2V) = 0, & on \quad (0,T)\times\mathbb{R}^d \\ -V = -f & on \quad \{T\}\times\mathbb{R}^d. \end{cases} \tag{6}$$

then $v_1 \leq v_2$.

Proof. See [3]. □

6 Conclusions and further research

We have extended the weak dynamic programming principle of [1] to the context of stochastic two-person zero-sum differential games. We used this principle to derive, under weaker assumptions than the ones existing in the literature, that both value functions are viscosity solutions of the associated Hamilton-Jacobi-Bellman-Isaacs equations.

There are two points where the setting considered in this paper is more general than the one in [2, 4]: we consider a larger control space for one of the players and a larger set of payoff functions. Indeed, to derive that the lower value function is a viscosity solution of the HJBI equation, the set B is allowed to be unbounded and the payoff function f needs only to be locally Lipschitz with polynomial growth (as opposed to globally Lipschitz).

If this approach is to be applied to both the lower and upper values at the same time then stronger conditions must be considered. Indeed, for our results to apply to the lower value we need A compact while for the upper value we need B compact. Thus, to apply our results to the lower and upper values simultaneously, then both A, B need to be compact sets. However, there are situations where only one of the value functions is of interest, like for example in worst-case approaches. In such a situation our approach provides a more general framework.

Some directions of further research

We indicate four main directions where further research can be taken:

- Study in more detail the technical assumptions required to prove Theorem 6.
- Extending the result to the case where the set A is unbounded. We assume that A is compact so as to ensure that the reward function has the required continuity, which must be uniform with respect to the control of the first player. It would be interesting if we could lift this assumption so as to be able to consider both A, B to be unbounded. If this is not possible it would then be interesting to find counterexamples.
- Considering even larger control spaces. The space of admissible controls considered in this thesis is the space of essentially bounded controls, $\mathbb{H}^\infty$. In the literature on stochastic optimal control a weaker condition on the integrability of the controls is often made. For example, in the reference for stochastic optimal control [5, p. 153], the admissible controls must be such that all its moments are finite.
- Proving uniqueness of solution for the HJBI equation, when A or B are not compact. If one of these sets fails to be compact then the Hamiltonian can be discontinuous and the arguments in the proof of uniqueness need to be modified.

  If a comparison result is proved for the HJBI equation with discontinuous Hamiltonian then we would have as a consequence

  $$V(t,x) \leq V^*(t,x) \leq V_*(t,x) \leq V(t,x),$$

  which proves that V is continuous.

References

[1] B. Bouchard and N. Touzi. Weak Dynamic Programming Principle for Viscosity Solutions. *Preprint*, 2009. Available from: `http://www.ceremade.dauphine.fr/~bouchard/pdf/BT09.pdf`.

[2] W. H. Fleming and P. E. Souganidis. On the existence of value functions of two-player, zero-sum stochastic differential games. *Indiana Univ. Math. J.*, 38(2):293–314, 1989. Available from: `http://dx.doi.org/10.1512/iumj.1989.38.38015`, `doi:10.1512/iumj.1989.38.38015`.

[3] P. Vitória. *A weak dynamic programming principle for zero-sum stochastic differential games.* Instituto Superio Técnico, Technical University of Lisbon, Lisbon, 2010.

[4] Rainer Buckdahn and Juan Li. Stochastic differential games and viscosity solutions of Hamilton-Jacobi-Bellman-Isaacs equations. *SIAM J. Control Optim.*, 47(1):444–475, 2008. Available from: `http://dx.doi.org/10.1137/060671954`, http://dx.doi.org/10.1137/060671954 `doi:10.1137/060671954`.

[5] Wendell H. Fleming and H. Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability.* Springer, New York, second edition, 2006.

[6] P. Tankov and N. Touzi. *Calcul Stochastique en Finance (polycopié).* École Polytechnique, 2009. Available from: `http://www.cmap.polytechnique.fr/~touzi/Poly-MAP552.pdf`.

[7] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1991.

[8] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Applications of Mathematics.* Springer-Verlag, New York, 2008. Reprint of the 1980 edition.

[9] Byung-Il Kim. On some Gronwall type inequalities for a system integral equation. *Bull. Korean Math. Soc.*, 42(4):789–805, 2005.

[10] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia. Available from: `http://dx.doi.org/10.1007/978-0-8176-4755-1`, `doi:10.1007/978-0-8176-4755-1`.

[11] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[12] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992. Available from: `http://dx.doi.org/10.1090/S0273-0979-1992-00266-5`, http://dx.doi.org/10.1090/S0273-0979-1992-00266-5 `doi:10.1090/S0273-0979-1992-00266-5`.

[13] Avner Friedman. *Differential games*. Wiley-Interscience [A division of John Wiley & Sons, Inc.], New York-London, 1971. Pure and Applied Mathematics, Vol. XXV.

[14] Rufus Isaacs. *Differential games. A mathematical theory with applications to warfare and pursuit, control and optimization*. John Wiley & Sons Inc., New York, 1965.

[15] J. L. Doob. *Measure theory*, volume 143 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994.

[16] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[17] Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic optimal control*, volume 139 of *Mathematics in Science and Engineering*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978. The discrete time case.

[18] Nicole El Karoui, D. Huu Nguyen, and Monique Jeanblanc-Picqué. Compactification methods in the control of degenerate diffusions: existence of an optimal control. *Stochastics*, 20(3):169–219, 1987.