Dynamic Programming via Measurable Selection

Roy Cerqueti

Department of Economic and Financial Institutions

University of Macerata, Italy

Via Crescimbeni, 20 - 62100 - Macerata, Italy

E-mail: roy.cerqueti@unimc.it

Abstract

The aim of this paper is to provide the proof of a Dynamic Programming Principle for a certain class of stochastic control problems with exit time. To this end, a Measurable Selection Theorem is also proved.

Keywords: Dynamic Programming, Stochastic Optimal Control, Exit Time, Measurable Selection.

Mathematices Subject Classification: 90C39, 93E20.

Introduction

The name *mathematical control theory* has been introduced about half a century ago. Although this fact, the nature of the optimal control problem has been the focus of research in optimization since the fifteen century. The precursor of the techniques involved in optimal control is commonly seen in calculus of variations. For a very interesting survey of the early optimization problems, we suggest (Yong and Zhou, 1999, Historical Remarks, pp. 92). In the 1940s and at the beginning of the 1950s, the theory of differential games has been developed in the U.S. and in the former Soviet Union for military purposes. The statements of the Bellman Dynamic Programming Method (Bellman, 1952, 1957) and the Pontryagin Theory (announced in 1956, see Pontryagin, 1959, 1986) are grounded in this scientific environment, and rely on a deterministic framework. Bellman was among the first that pointed out the necessity to introduce randomness in the optimal control theory, and mentioned the *stochastic* optimal control theory (Bellman, 1958). Nevertheless, stochastic differential equations and Ito's Lemma were not involved in (Bellman, 1958), and the first paper dealing with the diffusion systems, Markov processes and differential equations was (Florentin, 1961). Nowadays, the literature on this field grows continuously, with applications in economics, biology, finance, engineering and so on.

Several monographs give a complete survey on the mathematical control theory. For the deterministic case, we remind the reader to (Bardi and Capuzzo Dolcetta, 1997). The stochastic control theory is described in (Borkar, 1989; Fleming and Soner, 1993; Krylov, 1980; Yong and Zhou, 1999).

The keypoint of the optimal control theory is represented by an optimization problem, where the constraints are associated to some functions' properties (called controls α), that are elements of a certain functional space (called admissible region $\mathcal{A}$). Thus, the objective function J is a functional depending on the controls. The optimum with respect to the controls of such objective functional is called value function V.

The stochastic framework is related to the analysis of cases with admissible region given by stochastic processes spaces.

Starting from the objective functional and the definition of the admissible region, there are basically two methods to proceed: the Stochastic Maximum Principle (strongly related to the martingale theory) and the Dynamic Programming (that let intervene the theory of differential equations). In the first case, a set of necessary conditions for stochastic optimal controls are provided through forward-backward stochastic differential equations for adjoint variables and related stochastic Hamiltonian systems. In the latter case, one has to prove an optimality principle, named Dynamic Programming Principle, and rely the value function to the (classical) solution (if it exists, if it is unique) of a differential equation, named Hamilton Jacobi Bellman (HJB) equation. The HJB equation states *formally*, in the sense that we derive it by assuming the right regularity of the value function. Since the value function is generally not regular enough, a weak solution definition is needed: the viscosity solution. For the concept of viscosity solution, we remind to the seminal works (Crandall and Lions, 1981, 1983, 1987; Crandall et al., 1984; Lions, 1981, 1983). For a complete survey, we remind the reader to (Lions, 1982; Barles, 1994; Fleming and Soner, 2006, Chapter 2) and the celebrated User's Guide (Crandall et al, 1992).

Several papers establish existence and uniqueness results for both the value function and the optimal control. Among the others, we recall (Fleming, 1968; Ahmed and Teo, 1974, 1975; Davis, 1975; Fleming and Pardoux, 1982). Furthermore, the prove of the optimality principle has been the focus of some important research works. In (Davis and Varaiya, 1973), dynamic programming conditions for a certain class of stochastic optimal control problems have been obtained using the martingale method. The Girsanov measure transformation method has been applied to the solutions of the dynamical equations, in order to allow weaker requirements for the optimality principle. In (Haussmann, 1975) a different approach is used, and the optimality principle has been proved by applying a result on extremals due to Neustadt (Neustadt, 1969). In (Elliott, 1977) a semimartingale approach is adopted. Thus, by invoking the unique decomposition of special semi-martingales, some strong hypotheses required by Davis and Varaiya and Haussmann are avoided.

Different from the quoted papers, we show the validity of the optimality principle by using analytical arguments. More precisely, our contribution on the literature on this topic is an original step-by-step proof via Measurable Selection of a Dynamic Programming Principle, for a certain class of stochastic control problems with exit time.

In presence of exit time, the objective functional is not easy to treat due to the difficulty to prove the optimality principle. The main problems are due to the measurability questions associated to the control processes in the stochastic intervals. Therefore, in order to prove the dynamic programming principle, we prove a measurable selection result, which has its roots in an important result in functional analysis due to Jankov and von Neumann. The Jankov-von Neumann's Lemma implies in our case, as we shall see, a regularity condition for a certain class of admissible controls. We report the statement of the result of Jankov and Von Neumann, and we remind the reader to (Bertsekas and Shreve, 1978, pp. 182) for further details.

This paper is organized as follows. In the next section, the optimal control problem is formalized. In the third section the Measurable Selection Theorem is proved. The fourth section presents the Dynamic Programming Principle and its proof. Last section concludes the paper.

The optimal control problem

Consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ on which we define a standard Brownian motion W with respect to $\{\mathcal{F}_t\}_{t\geq 0}$ under P. Here $\{\mathcal{F}_t\}_{t\geq 0}$ represents the P-augmentation of the natural filtration generated by W; that is, $\mathcal{F}_t = \sigma\{W(u)|u \in [0,t]\} \wedge \mathcal{N}$, where $\mathcal{N}$ is the collection of all P-null sets or sets of measure zero under P.

Let us denote with $\mathcal{T}$ the set of the (optional) stopping times in $[0, +\infty]$, i.e.

$$\mathcal{T} := \{\tau : \Omega \to [0, +\infty] \,|\, \{\tau \leq t\} \in \mathcal{F}_t, \ \forall\, t \geq 0\}. \tag{1}$$

The controlled system is described by the following stochastic differential equation with initial data

$$\begin{cases} \mathrm{d}X(t) = \mu(X(t), \alpha(t))\mathrm{d}t + \sigma(X(t), \alpha(t))\mathrm{d}W(t), \\ \\ X(\eta) = \zeta \end{cases} \tag{2}$$

where X is a Markovian process and

- $\eta \in \mathcal{T}$,
- $t \in [[\eta, +\infty)$, where "[[" represents the lower bound of a stochastic interval
- it results

$$X : [0, +\infty) \times \Omega \to B \subseteq \mathbf{R}^n,$$

  where B is the solvency region, and it is open and bounded;

- fixed $t \in [0, +\infty)$, $X(t)$ is an $\mathcal{F}_t$-measurable and square integrable random variable with respect to P;
- fixed $\omega \in \Omega$ and $\eta \in \mathcal{T}$, define $X(\eta)(\omega) := (X \circ \eta)(\omega)$, where $\circ$ is the usual composition operator.

Denote the usual euclidean norm as $||\cdot||$. Then

- fixed $\eta \in \mathcal{T}$, $\alpha \in \mathcal{A}(\eta, \zeta)$, that is the set of admissible Markov controls, and it is defined as

$$\mathcal{A}(\eta,\zeta) := \Big\{\alpha : [[\eta, +\infty)\times\Omega \to A \subseteq \mathbf{R}^n \quad \{\mathcal{F}_t\}_{t\in[[\eta,+\infty)} - \text{progressively measurable}, \\ \text{such that } \mathbf{E}\Big[\int_\eta^{+\infty} e^{-\delta s}||\alpha(s)||\mathrm{d}s\Big] < +\infty\Big\}, \tag{3}$$

- ζ is an integrable random variable measurable with respect to $\mathcal{F}_\eta$,

- it results

$$\mu : B \times A \to \mathbf{R}^n,$$

$$\sigma : B \times A \to \mathbf{R}^{n \times n};$$

- $\exists\, L_1 > 0 \,|\, \forall\, x, y \in B,\ \alpha \in A,$

$$||\mu(x,\alpha) - \mu(y,\alpha)|| \leq L_1 ||x - y||,$$

$$||\sigma(x,\alpha) - \sigma(y,\alpha)|| \leq L_1 ||x - y||.$$

- $\exists\, L_2 > 0 \,|\, \forall\, x \in B,\ \alpha \in A,$

$$||\mu(x,\alpha)|| \leq L_2(1 + ||x||),$$

$$||\sigma(x,\alpha)|| \leq L_2(1 + ||x||).$$

Remark 1 *By the regularity hypothesis on the drift and diffusion coefficients and standard stochastic theory, we ensure the existence and uniqueness for the strong solution of the stochastic differential equation (2).*

Consider $B \subseteq \mathbf{R}^n$ and define the exit time τ as

$$\tau := \inf \Big\{ t \geq 0 \,|\, X(t) \notin B \Big\}. \tag{4}$$

Remark 2 $\{\tau \leq t\}$ *is measurable with respect to the* σ*-field* $\mathcal{F}_t$*, for each* $t \geq 0$.

We define the objective functional of the control process α.

First of all, we introduce the random variable J_0 as

$$J_0 : [0, +\infty) \times B \times A \times \Omega \to \mathbf{R}, \tag{5}$$

such that

$$J_0(t, x, \alpha(\cdot))(\omega) := \Big[\int_t^\tau f(X(s), \alpha(s)) e^{-\delta s} \mathrm{d}s + h(X(\tau)) e^{-\delta \tau} \,\Big|\, \mathcal{F}_t \Big](\omega), \tag{6}$$

where $X(t) = x$,

$$f : \bar{B} \times A \to \mathbf{R}$$

and

$$h : B \to \mathbf{R}.$$

are the running and the terminal reward, respectively, $\bar{B}$ is the closure of the set B, and δ is the discount factor. Furthermore, suppose that f satisfies a growth condition with respect to both the state and the control. More precisely, there exists $C > 0$ and $p \geq 1$ such that

$$|f(x, \alpha)| \leq C(1 + ||x|| + ||\alpha||^p), \quad \forall\ (x, \alpha) \in \bar{B} \times A.$$

Now we have the instruments to introduce the objective functional as conditional expectation of the random variable J_0 under the measure P.

We assume

$$J(t, x, \alpha(\cdot)) := \mathbf{E}\Big[J_0(t, x, \alpha(\cdot)) \,\Big|\, \mathcal{F}_t \Big] =$$

$$= \mathbf{E}\Big[\int_t^\tau f(X(s),\alpha(s))e^{-\delta s}\mathrm{d}s + h(X(\tau))e^{-\delta\tau}\,\Big|\,\mathcal{F}_t\Big]. \tag{7}$$

Compounding the stochastic elements, we are able to define the objective functional of the control problem analyzed in the case of random boundary data using the dynamics, i.e. for the (2).

In consistent with the definition of the functional J in (7), we define the objective functional of our control problem in the case of stochastic boundary data as

$$\bar{J}(\eta,\zeta,\alpha(\cdot)) := \mathbf{E}\Big[\bar{J}_0(\eta,\zeta,\alpha(\cdot))|\mathcal{F}_\eta\Big]. \tag{8}$$

Here we consider a maximization problem. The value function of the problem is

$$\bar{V}(\eta,\zeta) = \sup_{\alpha\in\mathcal{A}(\eta,\zeta)} \bar{J}(\eta,\zeta,\alpha(\cdot)). \tag{9}$$

Assume that $\bar{V}(\eta,\zeta) < +\infty$. Now we want to provide the definition of a particular class of admissible controls, which is useful for later development.

Definition 3 *Let us consider $\epsilon > 0$.*

Consider the state equation (2) with initial condition $X(\eta) = \zeta$, where $\eta \in \mathcal{T}$ and ζ is an integrable random variable measurable with respect to $\mathcal{F}_\eta$.

An admissible control $\alpha \in \mathcal{A}(\eta,\zeta)$ is said to be ϵ-optimal for the initial condition (η,ζ) if

$$\bar{J}(\eta,\zeta,\alpha(\cdot)) > \bar{V}(\eta,\zeta) - \epsilon, \quad P - a.s.$$

Remark 4 *Let us fix* (η, ζ) *the random initial condition for (2).*

By definition of sup, there exists an ϵ*-optimal control for* (η, ζ)*.*

A Measurable Selection Theorem

To prove the Dynamic Programming Principle, we need a measurable selection theorem. The aim of this section is to develop the measurable selection for our class of optimal control problems. To this end, we first need the description of the admissible region, with the main features of the admissible controls, and the analysis of some interesting properties of the solution of the state equation (2).

Let us fix (η, ζ) the initial data of the (2).

Define the space of functions

$$\Xi(\eta, \zeta) := \Big\{ u : [[\eta, +\infty) \times \Omega \to \mathbf{R}^n \,|\, u(.) \text{ is } \{\mathcal{F}_t\}_{t \geq \eta} \text{ progressively measurable} \Big\}.$$

We formalize the main properties of the admissible region. The following lemma is based on a result due to (Soner and Touzi, 2003), that can be adapted to our setting.

Lemma 5 *The set of admissible controls* $\mathcal{A}(\eta, \zeta)$ *is a Borel subset of* $\Xi(\eta, \zeta)$ *which satisfies the following conditions.*

- *(A0) Define a weighted norm on* $\mathcal{A}(\eta,\zeta)$

$$||\cdot||_w : \mathcal{A}(\eta,\zeta) \to \mathbf{R}^+, \tag{10}$$

such that

$$||\alpha||_w := \mathbf{E}\Big[\int_\eta^{+\infty} e^{-\delta s}||\alpha(s)||\mathrm{d}s, \Big| X(\eta)=\zeta\Big], \tag{11}$$

and consider $\Gamma_{(\eta,\zeta)}$ *the topology induced by the weighted norm (11) on* $\mathcal{A}(\eta,\zeta)$*. Then* $(\mathcal{A}(\eta,\zeta),\Gamma_{(\eta,\zeta)})$ *is a topological separable metric space.*

- *(A1) Closure under stopping time concatenation:*

$\forall\, \tau_1 \in \mathcal{T}$ *it results*

$$\nu := \alpha_1 \mathbf{1}_{[[\eta,\tau_1)} + \alpha_2 \mathbf{1}_{[[\tau_1,+\infty)} \in \mathcal{A}(\eta,\zeta),\ \forall\, \alpha_1,\alpha_2 \in \mathcal{A}(\eta,\zeta).$$

- *(A2) Stability under measurable selection:*

Denote as $\mathcal{B}_{\mathcal{A}(\eta,\zeta)}$ *the Borel* σ*-field of* $\mathcal{A}(\eta,\zeta)$*.* $\forall\,\theta_1 \in \mathcal{T}$ *and any measurable map*

$$\phi : (\Omega, \mathcal{F}_{\theta_1}) \to (\mathcal{A}(\eta,\zeta), \mathcal{B}_{\mathcal{A}(\eta,\zeta)}),$$

there exists $\nu \in \mathcal{A}(\eta,\zeta)$ *such that*

$$\phi(\omega)(t,\omega) = \nu(t,\omega) \text{ on } [[\theta_1,+\infty)\times\Omega,\ \mathcal{L}\times P - \text{a.e.}$$

where $\mathcal{B}_{\mathcal{A}(\eta,\zeta)}$ *is the set of the Borel subsets of* $\mathcal{A}(\eta,\zeta)$*;* $\mathcal{L}$ *is the Lebesgue measure on* $[[\theta_1,+\infty)$*.*

Proof. The proof is due to (Soner and Touzi, 2003) and standard stochastic calculus. ■

The following result summarizes the main properties of the solution of the state equation. As in Lemma 5, the following result is also grounded on (Soner and Touzi, 2003).

Lemma 6 *Let us denote the solution of (2) as $X^{\alpha}_{\eta,\zeta}(t)$, to indicate the initial data $X(\eta) = \zeta$, $\eta \in \mathcal{T}$ and ζ is an integrable random variable, and the control α.*

- *(SP1) Consistency in law with deterministic initial data:*

$$\mathbf{E}[f(X^{\alpha}_{\eta,\zeta}(s))|(\eta,\zeta) = (t,z)] = \mathbf{E}[f(X^{\alpha}_{t,z}(s))],$$

  where f is a Borel-measurable bounded function and $s \geq t$.

- *(SP2) Pathwise uniqueness:*

  $\forall\, \tau, \theta \in \mathcal{T}$ with $\theta \leq \tau$, P-a.s., it results

$$X^{\alpha}_{\tau,\zeta} = X^{\alpha}_{\theta,\gamma}, \text{ on } [[\tau, +\infty) \times \Omega, \text{ where } \zeta = X^{\alpha}_{\theta,\gamma}(\tau).$$

- *(SP3) Causality:*

  $\forall\, \alpha_1, \alpha_2 \in \mathcal{A}(\eta, \zeta)$ such that $\alpha_1 = \alpha_2$ on $[[\theta, \tau]]$, where $\tau, \theta \in \mathcal{T}$ such that $P(\tau \geq \theta) = 1$,

$$X^{\alpha_1}_{\theta,\gamma} = X^{\alpha_2}_{\theta,\gamma} \text{ on } [[\theta, \tau]] \times \Omega.$$

- *(SP4) Measurability:*

  $X^{\alpha}_{t,z}$ *is Borel measurable with respect to the variables* t, z *and* α.

Proof. In order to prove this result, we remind the reader to (Soner and Touzi, 2003). ■

The next result provides an useful generalization of the definitions of the functional $\bar{J}$ and of the value function $\bar{V}$ given in (8) and (9).

Lemma 7 *Let us consider* $\alpha \in \mathcal{A}$ *and* $\tilde{\alpha} \in \mathcal{A}$ *such that, for each fixed* $\eta \in \mathcal{T}$,

$$\alpha(s+\eta) = \tilde{\alpha}(s),$$

for each $s \geq 0$.

Then

$$\bar{J}(\eta, x, \alpha) = e^{-\delta\eta}\bar{J}(0, x, \tilde{\alpha}) := e^{-\delta\eta}J(x, \tilde{\alpha}), \tag{12}$$

and

$$\bar{V}(\eta, x) = e^{-\delta\eta}\bar{V}(0, x) := e^{-\delta\eta}V(x), \tag{13}$$

for each $\eta \in \mathcal{T}$, $x \in B$ *and* α *control process.*

The proof is omitted.

Remark 8 *The formulas (12) and (13) allow to treat the case in which the starting time is deterministic. In fact, a constant time is a special case of a stopping case, and, so* $t \in \mathcal{T}$, *for each* $t \in [0, +\infty)$.

For sake of completeness, we recall a result useful to prove the Measurable Selection Theorem.

Lemma 9 (Jankov-von Neumann) *Let X and Y Borel sets and A an analytic subset of $X \times Y$.*

Let us define

$$\mathrm{proj}_X(A) := \Big\{ x \in X \,|\, \exists\, y \in Y \text{ such that } (x, y) \in A \Big\} \subseteq X.$$

Then there exists an analytically measurable function

$$\phi : \mathrm{proj}_X(A) \to Y$$

such that

$$\mathrm{Gr}(\phi) := \Big\{ (x, \phi(x)) \,|\, x \in \mathrm{proj}_X(A) \Big\} \subseteq A.$$

We need a remarkable property of the Borel sets.

Lemma 10 *Let X be a Borel set. Then every Borel subset of X is analytic.*

Proof. We remind the reader to (Bertsekas and Shreve, 1978). ■

Now we prove the main result of this section. We have the following.

Theorem 11 (Measurable Selection Theorem) *Let us consider a stopping time $\eta \in \mathcal{T}$.*

For any product measure π on the space $[0, +\infty) \times B$ given the product of a

Lebesgue measure on $[0,+\infty)$ *and a probability measure on* B *and for each* $\epsilon > 0$*, there exists a Borel-measurable function*

$$\phi_\pi^\epsilon : \Big([0,+\infty) \times B, \mathcal{B}_{[0,+\infty)\times B}\Big) \to (\mathcal{A}(\cdot,\cdot), \mathcal{B}_{\mathcal{A}(\cdot,\cdot)})$$

such that $\phi_\pi^\epsilon(t,x)$ *is an* ϵ*-optimal control for starting point* $X(t) = x$*, for each* $(t,x) \in [0,+\infty) \times B$ π*-a.e.*

Proof. The proof consists of three steps.

- *First step*

  Given $\epsilon > 0$, let us define the space

  $$G_\epsilon := \Big\{(t,x,\alpha) \in [0,+\infty) \times B \times \mathcal{A}(t,x) \,|\, V(t,x) - J(t,x,\alpha) < \epsilon\Big\}.$$

  The space G_ϵ can be interpreted as follows: $\forall\,(t,x,\alpha) \in G_\epsilon$, $\alpha(t,x)$ is an ϵ-optimal control, for each $(t,x) \in [0,+\infty) \times B$. We want to prove that G_ϵ is a Borel-measurable set.

  In order to prove this claim we need to give the proof that J and V are measurable functions.

  Let us fix the initial condition (t,x).

  - J is a measurable function of (t,x,α). We get this property by the measurability of the state process $X(t)$ (by the property (SP4)), by the measurability of α, by the measurability of τ as stopping time and by the measurability of the functions f and g. So J is composed by measurable functions, hence J is measurable.

- $V(t,x)$ is measurable. By definition of $V(t,x)$ as supremum of $J(t,x,\alpha)$ with respect to the controls $\alpha \in \mathcal{A}(t,x)$, we get that $V(t,x)$ is measurable if and only if $\mathcal{A}(t,x)$ is countable. By the separability property of $\mathcal{A}(t,x)$, proved in Lemma 5, we get that there exists a set $\mathcal{D}(t,x) \subseteq \mathcal{A}(t,x)$ that is countable and dense in $\mathcal{A}(t,x)$. So $V(t,x)$ is measurable.

So G_ϵ is Borel-measurable.

- *Second step*

  By Lemma 10, we get that G_ϵ is an analytic subset of $[0,+\infty) \times B \times \mathcal{A}(\cdot,\cdot)$ (since it is a Borel set). By the Jankov-von Neumann Lemma, we obtain directly the existence of an analytically measurable function $\phi_\epsilon : [0,+\infty) \times B \to \mathcal{A}(\cdot,\cdot)$ such that $\mathrm{Gr}(\phi_\epsilon) \subseteq G_\epsilon$, i.e. $\phi_\epsilon(t,x)$ is an ϵ-optimal control, $\forall\,(t,x) \in [0,+\infty) \times B$.

- *Third step*

  It remains to construct a Borel measurable map ϕ_π^ϵ such that $\phi_\pi^\epsilon = \phi_\epsilon$ π-a.e.

  Let us define $\Pi([0,+\infty) \times B)$ as the set of all product measures on

$[0, +\infty) \times B$ and, given $\pi \in \Pi([0, +\infty) \times B)$, let us define $\mathcal{B}_{[0,+\infty)\times B}(\pi)$ as the completion of the Borel σ-algebra $\mathcal{B}_{[0,+\infty)\times B}$ with all π-null sets or sets of measure zero under π. Moreover, let us define σ-algebra

$$\Theta_{[0,+\infty)\times B} := \bigcap_{\pi \in \Pi([0,+\infty)\times B)} \mathcal{B}_{[0,+\infty)\times B}(\pi).$$

One can prove that every analytic subset of a Borel set X is measurable with respect to $\Theta_{[0,+\infty)\times B}$ (see, for example, (Bertsekas and Shreve, 1978)). As a particular case, we have that every analytic map ϕ_ϵ is measurable with respect to $\Theta_{[0,+\infty)\times B}$.

By definition, we get

$$\Theta_{[0,+\infty)\times B} \subseteq \mathcal{B}_{[0,+\infty)\times B}(\pi), \quad \forall\, \pi \in \Pi([0, +\infty) \times B),$$

and so ϕ_ϵ is measurable with respect $\mathcal{B}_{[0,+\infty)\times B}(\pi)$.

The definition of $\mathcal{B}_{[0,+\infty)\times B}(\pi)$ implies that there exists a Borel measurable map ϕ_π^ϵ such that $\phi_\pi^\epsilon = \phi_\epsilon$ π-a.e.

The theorem is completely proved. ■

Dynamic Programming Principle

We have proved in a general case a Measurable Selection Theorem. Now we are able to prove the Principle of Optimality.

First of all, we need a preliminary technical result.

Lemma 12 *Let us consider $\gamma \in \mathcal{T}$ such that $\gamma \in [[0, \tau]]$ and $\alpha \in \mathcal{A}$.*

Then

$$J(\gamma, X(\gamma), \alpha(\cdot)) = \mathbf{E}\Big[\int_{\gamma}^{\tau} f(X^{\alpha}_{\gamma,X(\gamma)}(s), \alpha(s))e^{-\delta s}\mathrm{d}s + h(X^{\alpha}_{\gamma,X(\gamma)}(\tau))e^{-\delta\tau} \,\Big|\, \mathcal{F}_{\gamma}\Big]. \tag{14}$$

Proof. The proof comes from (Yong and Zhou, 1999), Lemma 3.2, pg. 179, the definitions provided by (7) and (8) and the Markovian property of the state process X. ■

Theorem 13 (Dynamic Programming Principle) *Let us consider $\eta \in \mathcal{T}$. Then*

$$V(t, x) = \sup_{\alpha \in \mathcal{A}(t,x)} \mathbf{E}\Big[\int_{t}^{\eta\wedge\tau} \bar{f}(s, \alpha(s), X(s))\mathrm{d}s + V(\eta\wedge\tau, X(\eta \wedge \tau))|X(t) = x\Big], \tag{15}$$

where the sup is taken over all α admissible controls over the stochastic interval $[t, \eta \wedge \tau]]$ and, in the setting proposed for our general model,

$$\bar{f}(s, \alpha(s), X(s)) := e^{-\delta s} f(\alpha(s), X(s)).$$

Proof. We prove the double inequality so as to prove the validity of (15).

First step

Let us consider a stopping time $\eta \in \mathcal{T}$. We can write

$$V(t, x) = \sup_{\alpha \in \mathcal{A}} \mathbf{E}\Big[\int_{t}^{\eta\wedge\tau} \bar{f}(s, \alpha(s), X(s))\mathrm{d}s +$$

$$+\int_{\eta\wedge\tau}^{\tau} e^{-\delta t} f(\alpha(t), X(t))\mathrm{d}t + e^{-\delta\tau} h(X(\tau)) \Big| \mathcal{F}_t\Big].$$

Fix $\alpha_1 \in \mathcal{A}$. Accordingly with the notation introduced in Lemma 6, let us consider the controlled dynamic with starting point $X_{x,t}^{\alpha_1}(\eta\wedge\tau)$ and let π be a product measure on $[0,+\infty)\times B$ induced by $X_{x,t}^{\alpha_1}(\eta\wedge\tau)$. For each $\epsilon > 0$, by Theorem 11, there exists a Borel-measurable function

$$\phi_\pi^\epsilon : ([0,+\infty)\times B, \mathcal{B}_{[0,+\infty)\times B} \to (\mathcal{A}(\cdot,\cdot), \mathcal{B}_{\mathcal{A}(\cdot,\cdot)})$$

such that, $\forall\,(t,x) \in [0,+\infty)\times B$, $\phi_\pi^\epsilon(t,x)$ is an ϵ-optimal control at (t,x) π-a.e.. Now, let us consider the function

$$\xi : \Omega \to \mathcal{A}$$

such that:

$$\omega \to^{\xi} \phi_\pi^\epsilon(\eta\wedge\tau(\omega), X(\eta\wedge\tau)(\omega)).$$

For each ω, $\phi_\pi^\epsilon(\eta\wedge\tau(\omega), X(\eta\wedge\tau)(\omega))$ is ϵ-optimal at $(\eta\wedge\tau(\omega), X(\eta\wedge\tau)(\omega))$. Thanks to (A2), we have that there exists $\alpha_2 \in \mathcal{A}$ such that

$$\xi(\omega) = \phi_\pi^\epsilon(\eta\wedge\tau(\omega), X(\eta\wedge\tau)(\omega)) = \alpha_2(\omega) \qquad \forall\,\omega\in\Omega,$$

and α_2 is an ϵ-optimal admissible control at $(\eta\wedge\tau, X(\eta\wedge\tau))$.

Furthermore, by (A1) we get that for each $\alpha_1 \in \mathcal{A}$, there exists α defined as

$$\alpha := \alpha_1 \mathbf{1}_{[[t,\eta\wedge\tau)} + \alpha_2 \mathbf{1}_{[[\eta\wedge\tau,+\infty)} \tag{16}$$

that is an admissible control at (t,x).

Let us consider now $\epsilon > 0$ and $X(\cdot)$ a stochastic process with starting point a B-random variable and controlled by α defined as in (16), where α_1 is an arbitrary admissible control and α_2 is an ϵ-optimal control for starting point $(\eta \wedge \tau, X(\eta \wedge \tau))$.

By (16) it results $\alpha = \alpha_1$ in $[[0, \eta \wedge \tau)$ and $\alpha = \alpha_2$ in $[[\eta \wedge \tau, \tau]]$. So, by definition of V, by (SP3) and by (14), we get the following inequalities:

$$V(t,x) \geq J(t,x,\alpha) = \mathbf{E}\Big[\int_t^{\eta\wedge\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s \,\Big|\, \mathcal{F}_t\Big] +$$

$$\mathbf{E}\Big[\int_{\eta\wedge\tau}^{\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s + e^{-\delta\tau}h(X(\tau)) \,\Big|\, \mathcal{F}_{\eta\wedge\tau}\Big] = \mathbf{E}\Big[\int_t^{\eta\wedge\tau} \bar{f}(s,\alpha_1(s),X(s))\mathrm{d}s \,\Big|\, \mathcal{F}_t\Big] +$$

$$+\mathbf{E}\Big[\int_{\eta\wedge\tau}^{\tau} \bar{f}(s,\alpha_2(s),X(s))\mathrm{d}s + e^{-\delta\tau}h(X(\tau)) \,\Big|\, \mathcal{F}_{\eta\wedge\tau}\Big]$$

$$= \mathbf{E}\Big[\int_t^{\eta\wedge\tau} \bar{f}(s,\alpha_1(s),X(s))\mathrm{d}s \,\Big|\, \mathcal{F}_t\Big] + J(\eta\wedge\tau, X(\eta\wedge\tau), \alpha_2)$$

$$\geq \mathbf{E}\Big[\int_t^{\eta\wedge\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s \,\Big|\, \mathcal{F}_t\Big] + V(\eta\wedge\tau, X(\eta\wedge\tau)) - \epsilon.$$

Then

$$V(t,x) \geq \sup_{\alpha\in\mathcal{A}} \mathbf{E}\Big[\int_t^{\eta\wedge\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s + V(\eta\wedge\tau, X(\eta\wedge\tau)) | X(t)) = x\Big] - \epsilon.$$

Second step: let us consider $\epsilon > 0$ and $\alpha(\cdot)$ an ϵ-optimal control for (t,x).

In order to proceed, we need to remark that, given $\eta \in \mathcal{T}$, it results:

$$\alpha \in \mathcal{A} \;\Rightarrow\; \alpha \mathbf{1}_{[[\gamma,+\infty)} \in \mathcal{A}, \;\; \forall \gamma > \eta. \tag{17}$$

Moreover, by (13) and (14), we get

$$\sup_{\alpha\in\mathcal{A}} \mathbf{E}\Big[\int_{\eta\wedge\tau}^{\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s + e^{-\delta\tau}h(X(\tau))\,\Big|\,\mathcal{F}_{\eta\wedge\tau}\Big] = V(\eta\wedge\tau, X(\eta\wedge\tau)). \tag{18}$$

So, by (17) and (18), we have:

$$V(t,x)-\epsilon \leq \mathbf{E}\Big[\int_{t}^{\eta\wedge\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s\,\Big|\,\mathcal{F}_t\Big]+$$

$$+\mathbf{E}\Big[\int_{\eta\wedge\tau}^{\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s+e^{-\delta\tau}h(X(\tau))\,\Big|\,\mathcal{F}_{\eta\wedge\tau}\Big] \leq \mathbf{E}\Big[\int_{t}^{\eta\wedge\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s\,\Big|\,\mathcal{F}_t\Big]+$$

$$\begin{aligned} &+\sup_{\alpha\in\mathcal{A}} \mathbf{E}\Big[\int_{\eta\wedge\tau}^{\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s + e^{-\delta\tau}h(X(\tau))\,\Big|\,\mathcal{F}_{\eta\wedge\tau}\Big] \\ &= \mathbf{E}\Big[\int_{t}^{\eta\wedge\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s\,\Big|\,\mathcal{F}_t\Big] + V(\eta\wedge\tau, X(\eta\wedge\tau)). \end{aligned}$$

Then

$$V(t,x)-\epsilon \leq \sup_{\alpha\in\mathcal{A}} \mathbf{E}\Big[\int_{t}^{\eta\wedge\tau} \bar{f}(s,\alpha(s),X(s))\mathrm{d}s + V(\eta\wedge\tau, X(\eta\wedge\tau))|X(t)=x\Big].$$

The theorem is completely proved. ■

Conclusions

In this paper a Dynamic Programming Principle for a certain class of optimal control problems with exit time is proved. To this end, a Measurable Selection result is firstly showed. The optimality principle can be used to treat several dynamic optimization problems, involving economic, financial, engineering or physical applications.

References

[1] N.U. Ahmed and K.L. Teo, An existence theorem on optimal control of partially observable diffusions, SIAM Journal on Control and Optimization, 12, 351-355, (1974).

[2] N.U. Ahmed and K.L. Teo, Optimal control of stochastic Ito differential systems with fixed terminal time, Advances in Applied Probability, 7, 154-178, (1975).

[3] M. Bardi and I. Capuzzo Dolcetta, Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equation. Birkhäuser Boston, Inc.: Boston, MA, (1997).

[4] G. Barles, Viscosity solutions of Hamilton-Jacobi equations, Springer-Verlag, Paris (1994).

[5] R. Bellman, On the theory of dynamic programming, Proc. Nat. Acad. Sci. USA, 38, 716-719, (1952).

[6] R. Bellman, Dynamic programming, Princeton University Press, Princeton, N. J., (1957).

[7] R. Bellman, Dynamic programming and stochastic control process, Information & Control, 1, 228-239, (1958).

[8] D.P. Bertsekas and S.E. Shreve, Stochastic optimal control. The discrete time case. Academic Press, Inc.: New York-London, (1978).

[9] V.S. Borkar, Optimal control of diffusion processes, Pitman Research Notes in Mathematics Series, 203, Longman Scientific & Technical, Harlow, (1989).

[10] M.G. Crandall, L.C. Evans and P.-L. Lions, Some properties of viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc., 282(2), 487-502, (1984).

[11] M.G. Crandall and P.-L. Lions, Condition d'unicit pour les solutions gnralises des quations de Hamilton-Jacobi du premier ordre, C. R. Acad. Sci. Paris Sr. I Math., 292(3), 183-186, (1981).

[12] M.G. Crandall and P.-L. Lions, Remarks on the existence and uniqueness of unbounded viscosity solutions of Hamilton-Jacobi equations, Illinois J. Math., 31(4), 665-688, (1987).

[13] M.G. Crandall and P.-L. Lions, Viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc., 277(1), 1-42, (1983).

[14] M.G. Crandall, H. Ishii and P.-L. Lions, User's guide to viscosity solutions of second order partial differential equations, Bull. Amer. Math.

Soc. (N.S.), 27(1), 1-67, (1992).

[15] M.H.A. Davis, On the existence of optimal policies in stochastic control, SIAM Journal on Control and Optimization, 11, 587-594, (1973).

[16] M.H.A. Davis and P.P. Varaiya, Dynamic programming conditions for partially observable stochastic systems, SIAM Journal on Control and Optimization 11, 226261, (1973).

[17] R.J. Elliott, The optimal control of a stochastic system, SIAM Journal on Control and Optimization 15(5), 756778, (1977).

[18] R.J. Elliott, Stochastic Calculus and Applications, Springer, Berlin Heidelberg New York, (1982).

[19] W.H. Fleming, Optimal control of partial observable diffusions, SIAM Journal of Control, 6, 194-214, (1968).

[20] W.H. Fleming and E. Pardoux, Existence of optimal controls for partially observed diffusions, SIAM Journal on Control and Optimization, 20, 261- 285, (1982).

[21] W.H. Fleming and H.M. Soner, Controlled Markov processes and viscosity solutions. Applications of Mathematics (New York), Springer-Verlag: New York, (1993).

[22] J.J. Florentin, Optimal control of continuous time, Markov, stochastic systems, Journal of Electronics Control, 10, 473-488, (1961).

[23] U.G. Haussmann, General necessary conditions for the optimal control of stochastic systems, Symposium of Stochastic Systems, University of Kentucky, Lexington, (1975).

[24] N.V. Krylov, Controlled diffusion processes, Springer-Verlag, New York, (1980).

[25] P.-L. Lions, Control of diffusion processes in R^N, Comm. Pure Appl. Math., 34(1), 121-147, (1981).

[26] P.-L. Lions, Équations de Hamilton-Jacobi et solutions de viscosité, Ennio De Giorgi colloquium, Paris, 83-97, (1983).

[27] P.-L. Lions, Generalized solutions of Hamilton-Jacobi equations, Research Notes in Mathematics, 69, Pitman (Advanced Publishing Program), Boston, Mass.-London, (1982).

[28] L.W. Neustadt, A general theory of extremals, J. Comput. System Sci., 57-92, (1969).

[29] L.S. Pontryagin, Optimal processes of regulation, Uspekhi Mat. Nauk, 14, 1(85), 3-20 (1959) (in Russian); English transl. in Amer. Math. Soc. Transl. 18(2), (1961).

[30] L.S. Pontryagin, The mathematical theory of optimal control processes and differential games, Proc. Steklov Inst. Math., 4, 123-159, (1986).

[31] H.M. Soner and N. Touzi, A stochastic representation for mean curvature type geometric flows, Ann. Prob. 31, 1145-1165, (2003).

[32] J. Yong and Zhou X.Y., Stochastic controls. Hamiltonian systems and HJB equations. Applications of Mathematics, Springer-Verlag: New York, (1999).