JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 64, No. 3, MARCH 1990

An Optimal Control Problem with a Random Stopping Time[1]

E. K. BOUKAS,[2] A. HAURIE,[3] AND P. MICHEL[4]

Communicated by G. Leitmann

Abstract. This paper deals with a stochastic optimal control problem where the randomness is essentially concentrated in the stopping time terminating the process. If the stopping time is characterized by an intensity depending on the state and control variables, one can reformulate the problem equivalently as an infinite-horizon optimal control problem. Applying dynamic programming and minimum principle techniques to this associated deterministic control problem yields specific optimality conditions for the original stochastic control problem. It is also possible to characterize extremal steady states. The model is illustrated by an example related to the economics of technological innovation.

Key Words. Stochastic control, infinite-horizon optimal control, minimum principle.

1. Introduction

The aim of this paper is to establish optimality conditions in the form of a modified minimum principle for a class of stochastic control problems where the randomness is essentially concentrated in the stopping time. This class of control problems and its relationship with a production and maintenance planning problem for an FMS have already been considered by Boukas and Haurie (Ref. 1). These authors have also shown that these problems are closely related to the so-called piecewise deterministic control problems studied by Davis (Ref. 2) and Vermes (Ref. 3).

[1] This research has been supported by NSERC-Canada, Grants 36444 and A4952; by FCAR-Québec, Grant 88EQ3528, Actions Structurantes; and by MESS-Québec, Grant 6.1/7.4(28).
[2] Professeur, GERAD, École Polytechnique, Montréal, Québec, Canada.
[3] Professeur, Université de Genève, Geneva, Switzerland, and GERAD, École des Hautes Études Commerciales, Montréal, Québec, Canada.
[4] Professeur, Université de Paris 1, Paris, France.

Actually this stochastic control problem can be reformulated equivalently as an infinite-horizon deterministic optimal control problem in a larger state space. One can obtain readily the Hamilton-Jacobi-Bellman equation if one assumes that the sufficient regularity conditions for the dynamic programming approach work. Under more general conditions, one can invoke the infinite-horizon version of the minimum principle (Ref. 4), with asymptotic transversality conditions (Ref. 5). The optimality conditions can then be expressed in terms of the original state variable only, and an interesting economic interpretation can be given to them. This completes and clarifies the preliminary investigation of optimality conditions for this class of systems presented in Boukas and Haurie (Ref. 1).

As the system dynamics is stationary, except for the discounting term, there is a possibility to define an extremal steady state for which the intensity of the stopping time remains constant as well as the state and control variables. It is worth noticing that this extremal steady state is not the stationary solution of the associated infinite-horizon deterministic control problem.

2. Optimal Control under a Random Stopping Time

Consider a system described by the state equation

$$dx(t)/dt = f(x(t), u(t)), \qquad t \geq 0, \tag{1}$$

$$x(0) = x^0, \qquad \text{given initial state,} \tag{2}$$

where $x \in \mathbb{R}^p$ is the state variable and $u \in U \subset \mathbb{R}^q$ is the control variable. The function $f: \mathbb{R}^p \times \mathbb{R}^q \mapsto \mathbb{R}^p$, as well as the gradient w.r.t. x, f'_x, are assumed to be continuous with respect to x and u. An admissible pair $(u(\cdot), x(\cdot))$ is defined as a measurable control $u(\cdot): [0, \infty) \mapsto U$ and the associated absolutely continuous trajectory $x(\cdot): [0, \infty) \mapsto \mathbb{R}^p$, unique solution of (1)-(2).

Let (Ω, B, P_u) be a probability space; and let T be a random variable, called the stopping time for the system (1)-(2). The probability measure P_u depends on the control $u(\cdot)$ in the following way: for an admissible pair $(u(\cdot), x(\cdot))$ at initial state x^0, we assume that, for any $t \geq 0$ and dt, the following holds:

$$P_u[T \in (t, t+dt) \mid T \geq t] = q(x(t), u(t))\, dt + o(dt), \tag{3}$$

where $q: \mathbb{R}^p \times \mathbb{R}^q \mapsto \mathbb{R}^+$(called the intensity of the jump time T) as well as its gradient w.r.t. x, q'_x, are continuous functions, and $o(dt)/dt \to 0$, uniformly in x and u, when $dt \to 0$.

Let $L:\mathbb{R}^p \times \mathbb{R}^q \mapsto \mathbb{R}^+$ and $\Phi:\mathbb{R}^p \mapsto \mathbb{R}^+$ be two given continuous functions with continuous gradients L'_x and Φ'_x. We define the cost $g(x^0, u(\cdot))$ associated with x^0 and $u(\cdot)$ as the following conditional expectation:

$$g(x^0, u(\cdot)) = E_u\left[\int_0^T \exp(-\rho t) L(x(t), u(t))\, dt + \exp(-\rho T)\Phi(x(T)) \mid x^0\right]. \tag{4}$$

We want to find a control $u^*(\cdot)$ which minimizes (4) subject to (1)-(3); i.e., we want to solve

$$V(x^0) = \min\{g(x^0, u(\cdot)) \mid (u(\cdot), x(\cdot)) \text{ admissible}\}. \tag{5}$$

3. Reformulation as a Deterministic Optimal Control Problem

The stochastic control problem defined above in (1)-(5) can be reformulated as an equivalent infinite-horizon deterministic optimal control problem.

Let us first consider the elementary probability of the time interval $(t, t+dt)$ for the stopping time T, given an admissible pair $(u(\cdot), x(\cdot))$, at initial state x^0. From (3), this probability is given by

$$q(x(t), u(t)) \exp\left[-\int_0^t q(x(s), u(s))\, ds\right] dt.$$

Therefore, the expression of the cost $g(x^0, u(\cdot))$ can be rewritten as follows:

$$g(x^0, u(\cdot)) = \int_0^\infty \left\{\int_0^t \exp(-\rho s) L(x(s), u(s))\, ds + \exp(-\rho t)\Phi(x(t))\right\} \times q(x(t), u(t)) \exp\left[-\int_0^t q(x(s), u(s))\, ds\right] dt. \tag{6}$$

Let us integrate by parts the term $\int_0^\infty \mathcal{U}\, d\mathcal{V}$, where

$$\mathcal{U}(t) = \int_0^t \exp(-\rho s) L(x(s), u(s))\, ds,$$

$$\mathcal{V}(t) = -\exp\left[-\int_0^t q(x(s), u(s))\, ds\right].$$

Assume that

$$\lim_{t\to\infty} \mathcal{U}(t)\mathcal{V}(t) = 0. \tag{7}$$

Then, we can rewrite

$$g(x^0, u(\cdot)) = \int_0^\infty \exp(-\rho t)\{L(x(t), u(t)) + \Phi(x(t))q(x(t), u(t))\} \times \exp\left[-\int_0^t q(x(s), u(s))\, ds\right] dt. \tag{8}$$

Remark 3.1. Condition (7) is implied by the following:

$$\int_0^\infty \exp(-\rho s) L(x(s), u(s))\, ds < \infty, \tag{7a}$$

$$\int_0^\infty q(x(s), u(s))\, ds = \infty. \tag{7b}$$

Condition (7a) is standard when $\rho > 0$, whereas (7b) means that the stopping time occurs almost surely before infinity. Assumption (7) may hold even if $\rho = 0$.

Introducing the following auxiliary state equation:

$$dz(t)/dt = q(x(t), u(t)), \qquad z(0) = 0, \tag{9}$$

the original stochastic optimization problem can be recast in the form of a deterministic infinite-horizon optimal control problem, with the extended state variable (x, z) and with cost

$$g(x^0, u(\cdot)) = \int_0^\infty \exp(-\rho t - z(t)) \times \{L(x(t), u(t)) + \Phi(x(t))q(x(t), u(t))\}\, dt. \tag{8'}$$

Remark 3.2. This reformulation as a deterministic control problem with infinite-time horizon would permit us to use known results on existence of optimal controls [e.g., Baum (Ref. 6) or Toman (Ref. 7)] to give conditions under which a solution to problem (1)-(5) exists.

4. Hamilton–Jacobi–Bellman Equation

For the associated deterministic problem, the Bellman value functional is defined as

$$W(t, x, z) = \inf_{u(\cdot)} \left\{ \int_t^\infty \exp(-\rho s - z(s))\{L(x(s), u(s)) + \Phi(x(s))q(x(s), u(s))\}\, ds \,\middle|\, x(t) = x, z(t) = z \right\}. \tag{10}$$

It is easy to check the following link with the value function defined in (5) for the original problem:

$$W(t, x, z) = \exp(-\rho t - z)V(x). \tag{11}$$

The Hamilton-Jacobi-Bellman equation for the deterministic problem is

$$-\partial W/\partial t = \inf_{u \in U} \{\exp(-\rho t - z)\{L(x, u) + \Phi(x)q(x, u)\} + (\partial W/\partial x)f(x, u) + (\partial W/\partial z)q(x, u)\}. \tag{12}$$

Substituting (11) in (12) and collecting terms yields

$$\rho V(x) = \inf_{u \in U} \{L(x, u) + (\partial V/\partial x)f(x, u) + q(x, u)[\Phi(x) - V(x)]\}. \tag{13}$$

This equation has basically the same structure as the dynamic programming equation established by Rishel (Ref. 8) for control systems with jump Markov disturbances. In order to use (13) as necessary and sufficient optimality conditions, one has to assume rather stringent regularity conditions [e.g., see Boltyanskii (Ref. 9), Mirica (Ref. 10), or Rishel (Ref. 11)]. Under much milder conditions, one can use the minimum principle approach, as shown in the next section.

5. Minimum Principle

According to Halkin (Ref. 4), if $u^*(\cdot)$ is an optimal control generating the extended state trajectory $(x^*(\cdot), z^*(\cdot))$, then there exist absolutely continuous functions $\mu(\cdot): [0, \infty) \mapsto \mathbb{R}^p$ and $\nu(\cdot): [0, \infty) \mapsto \mathbb{R})$ and a constant α such that

(i) $H(u^*(t), x^*(t), z^*(t), \alpha, \mu(t), \nu(t), t)$
$= \min_{u \in U} H(u, x^*(t), z^*(t), \alpha, \mu(t), \nu(t), t)$,

(ii) $d\mu(t)/dt = -H'_x(u^*(t), x^*(t), z^*(t), \alpha, \mu(t), \nu(t), t)$, a.e.,

(iii) $d\nu(t)/dt = -H'_z(u^*(t), x^*(t), z^*(t), \alpha, \mu(t), \nu(t), t)$, a.e.,

(iv) $(\alpha, \mu(0), \nu(0)) \neq 0$,

where

$$H(u, x, z, \alpha, \mu, \nu, t) = \alpha \exp(-\rho t - z)\{L(x, u) + \Phi(x)q(x, u)\} + \mu f(x, u) + \nu q(x, u).$$

As shown by Halkin, the infinite-horizon minimum principle does not necessarily include the asymptotic transversality conditions for the adjoint variables,

$$\lim_{t \to \infty} (\mu(t), \nu(t)) = 0. \tag{14}$$

However, Michel (Ref. 5) has shown that the following asymptotic property must hold:

$$\lim_{t\to\infty} \min_{u\in U} H(u, x^*(t), z^*(t), \alpha, \mu(t), \nu(t), t) = 0. \tag{15}$$

Using (15), one can establish the transversality condition for the one-dimensional ν variable only under the following assumption.

Assumption A1. There exists a control $\tilde{u}(\cdot)$ satisfying $f(x^*(t), \tilde{u}(t)) = 0$ for t large enough, and such that

$$\lim_{t\to\infty} q(x^*(t), \tilde{u}(t)) > 0,$$

$$\lim_{t\to\infty} \exp[-\rho t - z^*(t)]\{L(x^*(t), \tilde{u}(t)) + \Phi(x^*(t))q(x^*(t), \tilde{u}(t))\} = 0.$$

A direct consequence of (15) and Assumption A1 is that $\lim_{t\to\infty} \nu(t) \geq 0$. It is also easy to see that $\nu(t)$ is always nonpositive, and therefore the limit is 0. Indeed, consider the same control problem, except with Eq. (9) replaced by

$$dz(t)/dt = q(x(t), u(t)) - v(t), \qquad z(0) = 0, \tag{9'}$$

where the additional control variable v takes its value in $\mathbb{R}^+$. A nonidentically zero control $v(t)$ increases the value of the cost functional. Therefore, the optimal trajectory for this new problem is the same as before with $v^*(t) \equiv 0$, and the minimum of the Hamiltonian w.r.t. v implies that $\nu(t) \leq 0$.

Remark 5.1. The corollary in Michel (Ref. 5), giving (14), does not apply directly to this model, since the intensity function $q(x, u)$ is nonnegative. However, the sign property of $\nu(t)$ allows one to obtain (14) under the condition that, for some compact subset $K \subset U$, the set $\{f(x^*(t), u), q(x^*(t), u): u \in K\}$ contains a neighborhood of 0 in $\mathbb{R}^p \times \mathbb{R}^+$ for t large enough.

We are now in a position to prove the following theorem.

Theorem 5.1. If $(u^*(\cdot), x^*(\cdot))$ is an optimal solution for the original stochastic control problem (1)–(5) and if Assumption A1 holds, then there exist an absolutely continuous function $\lambda(\cdot)$: $[0, \infty) \mapsto \mathbb{R}^p$ and a constant α such that

(i) $$\mathcal{H}(u^*(t), x^*(t), \alpha, \lambda(t), \mathcal{L}^*(t)) = \min_{u\in U} \mathcal{H}(u, x^*(t), \alpha, \lambda(t), \mathcal{L}^*(t)),$$

(ii) $$d\lambda(t)/dt = -\mathcal{H}'_x(u^*(t), x^*(t), \alpha, \lambda(t), \mathcal{L}^*(t)) + (\rho + q(x^*(t), u^*(t))\lambda(t), \qquad \text{a.e.},$$

(iii) $(\alpha, \lambda(0)) \neq 0$,

where

$$\mathcal{H}(u, x, \alpha, \lambda, \mathcal{L}) = \alpha\{L(x, u) + (\Phi(x) - \mathcal{L})q(x, u)\} + \lambda f(x, u),$$

$$\mathcal{L}^*(t) = \int_t^\infty \exp(-\rho(s-t)) \times \{L(x^*(s), u^*(s)) + \Phi(x^*(s))q(x^*(s), u^*(s))\} \times \exp\left[\int_t^s q(x^*(\tau), u^*(\tau))\, d\tau\right] ds. \tag{16}$$

Proof. Halkin's necessary condition (iii) is identical to

$$(d/dt)\nu(t) = -\alpha(d/dt)W(t, x^*(t), z^*(t)). \tag{17}$$

equation (17), with transversality condition $\lim_{t\to\infty}\nu(t) = 0$ resulting from Assumption A1, is equivalent to

$$\begin{aligned}\nu(t) &= -\alpha W(t, x^*(t), z^*(t))\\ &= -\alpha \exp(-\rho t - z^*(t))V(x^*(t))\\ &= -\alpha \exp(-\rho t - z^*(t))\mathcal{L}^*(t).\end{aligned} \tag{18}$$

Substituting (18) in Halkin's necessary conditions (i), (ii), and (iv) yields the desired result. □

Remark 5.2. The term $\mathcal{L}^*(t) = V(x^*(t))$ is also equal to

$$\min_{u(\cdot)} E_u\left[\int_t^T \exp[-\rho(s-t)]L(x(s), u(s))\, ds + \exp[(-\rho(T-t)]\Phi(x(T) \mid T > t, x(t) = x^*(t)\right],$$

i.e., the minimum expected cost-to-go. This expression, which is related to the adjoint variable $\nu(t)$, anticipates the future behavior of the system cost along the optimal trajectory.

Thereafter, we exclude the degenerate case $\alpha = 0$, and α will be taken equal to 1. The minimum principle stated above has the following economic interpretation. The Hamiltonian expresses the usual trade-off between an instantaneous cost and the marginal effect of state modification along the optimal trajectory. The instantaneous cost takes a peculiar form here, since in addition to the elementary cost $L(x, u)\, dt$, it takes into account the conditional expectation of the net cost due to a jump $(\Phi(x) - \mathcal{L}^*(t))q(x, u)\, dt$ over the elementary time interval $[t, t+dt)$.

6. Extremal Steady State

In this section, we characterize an extremal steady state, viz., a time-invariant admissible pair $(\hat{u}, \hat{x})$, associated with a constant adjoint variable $\hat{\lambda}$ which satisfies the necessary optimality conditions given in Theorem 5.1. Such a steady state is determined as a solution of the following relations:

$$f(\hat{x}, \hat{u}) = 0, \tag{19a}$$

$$\mathcal{H}(\hat{u}, \hat{x}, \hat{\lambda}, \hat{\mathcal{L}}) = \min_{u \in U} \mathcal{H}(u, \hat{x}, \hat{\lambda}, \hat{\mathcal{L}}), \tag{19b}$$

$$\mathcal{H}'_x((\hat{u}, \hat{x}, \hat{\lambda}, \hat{\mathcal{L}}) = (\rho + q(\hat{x}, \hat{u}))\hat{\lambda}, \tag{19c}$$

with

$$\hat{\mathcal{L}} = \{1/[\rho + q(\hat{x}, \hat{u})]\}[L(\hat{x}, \hat{u}) + \Phi(\hat{x})q(\hat{x}, \hat{u})]. \tag{20}$$

Remark 6.1. Notice that this solution is not, in general, a stationary solution of the associated deterministic control problem introduced in Section 3. This would require the additional condition $\dot{z} = q(\hat{x}, \hat{u}) = 0$ which is not a natural steady-state condition for the original stochastic control problem.

To illustrate this concept on a simple example related to the economics of innovation, consider a manufacturing firm which maintains a research and development (RD) department in order to be ready to exploit a possible technological breakthrough, e.g., computerized integrated manufacturing. Let x be a state variable which measures the size of the RD department. The state equation is

$$dx(t)/dt = u(t) - \delta x(t), \qquad t \geq 0, \tag{21}$$

$$x(0) = x^0, \qquad \text{given initial state}, \tag{22}$$

where the control u represents the investment effort in the RD department and δ is the depreciation rate of the RD facility. The current cost function is given by

$$L(x, u) = a + bx + cu^2, \tag{23}$$

where a represents the current constant production cost of the firm (i.e., determined by the current technology used); b and c are the maintenance and development costs of RD, respectively. The time of occurrence of the technological breakthrough is modelled as a random stopping time with intensity

$$q(x, u) = \alpha x. \tag{24}$$

The breakthrough generates a new constant production cost. Assume that the larger the RD department is at the time of breakthrough the lower this new cost will be. This permits us to represent the ability of a prepared firm to seize the opportunity offered by a new technology. This leads us to consider the anticipated total discounted cost after the stopping time as given by the function

$$\Phi(x) = (a/\rho)\exp(-x) = \int_0^\infty \exp(-\rho t) a \exp(-x)\, dt. \tag{25}$$

If the firm minimizes the expected discounted cost over an infinite-time horizon, Eqs. (21)-(25) define a control system with a random stopping time.

Equations (19)-(20) with straightforward algebra lead to the following equation to be satisfied by the steady state $\hat{x}$:

$$a[1-\exp(-x)] + bx + c\delta^2 x^2 - (2c\delta/a)x(\rho + \alpha x) = 0. \tag{26}$$

This equation admits a unique solution $\hat{x}$ under the following conditions:

$$a\delta < 2\alpha, \qquad 2c\delta\rho < a(a+b). \tag{27}$$

If (27) holds, the left-hand-side of Eq. (26) is a concave function which is zero and increasing at the origin, and decreasing to $-\infty$ when x increases to ∞. The steady state RD investment rate is then $\hat{u} = \delta\hat{x}$.

References

1. BOUKAS, E. K., and HAURIE, A., *Optimality Conditions for Continuous-Time Systems with Controlled Jump Markov Disturbances: Application to an FMS Planning Problem*, Analysis and Optimization of Systems, Edited by A. Bensoussan and J. L. Lions, Springer-Verlag, Berlin, Germany, pp. 633-676, 1988.
2. DAVIS, M. H. A., *Control of Piecewise-Deterministic Process via Discrete-Time Dynamic Programming*, Proceedings of the 3rd Bad Honnef Symposium on Stochastic Differential Systems, Springer-Verlag, Berlin, Germany, 1985.
3. VERMES, D., *Optimal Control of Piecewise Deterministic Markov Process*, Stochastics, Vol. 14, pp. 165-208, 1985.
4. HALKIN, H., *Necessary Conditions for Optimal Control Problems with Infinite Horizons*, Econometrica, Vol. 42, pp. 267-272, 1974.
5. MICHEL, P., *On the Transversality Condition in Infinite-Horizon Optimal Control Problems*, Econometrica, Vol. 50, pp. 975-985, 1982.
6. BAUM, R. F., *Existence Theorems for Lagrange Control Problems with Unbounded Time Domain*, Journal of Optimization Theory and Applications, Vol. 19, pp. 89-116, 1976.
7. TOMAN, M. A., *Optimal Control with an Unbounded Horizon*, Journal of Economic Dynamics and Control, Vol. 9, pp. 291-316, 1985.

8. RISHEL, R., *Control of Systems with Jump Markov Disturbances*, IEEE Transactions on Automatic Control, Vol. AC-20, pp. 241-244, 1975.
9. BOLTYANSKII, V. G., *Sufficient Conditions for Optimality and the Justification of the Dynamic Programming Method*, SIAM Journal on Control, Vol. 4, pp. 326-361, 1966.
10. MIRICA, S., *On the Admissible Synthesis in Optimal Control Theory and Differential Games*, SIAM Journal on Control, Vol. 7, pp. 292-316, 1969.
11. RISHEL, R., *Dynamic Programming and Minimum Principles for Systems with Jump Markov Disturbances*, SIAM Journal on Control, Vol. 13, pp. 338-371, 1975.