Ann. sc. math. Québec, 1986, Vol. 10, No 2, pp.181-187

# OPTIMAL STOCHASTIC CONTROL OF A CLASS OF PROCESSES WITH AN EXPONENTIAL COST FUNCTION[1]

Mario Lefebvre

Summary

In this note, the optimal control of a class of processes with non-linear/Gaussian dynamics and negative exponential cost criterion is given in terms of the first-passage distribution over the terminal set for the uncontrolled processes. The result is applied to the optimal control of a process with lognormal transitions.

Résumé

Dans cette note, la commande optimale d'une classe de processus avec dynamique non linéaire/gaussienne et coûts exponentiels est donnée en termes de la distribution de premier passage à l'intérieur de la région d'arrêt pour les processus non contrôlés. Le résultat est appliqué à la commande optimale d'un processus avec transitions lognormales.

## 1. Introduction and theoretical results

In [6] Whittle and Gait considered processes with linear/Gaussian dynamics and quadratic control costs, but with general terminal costs. They showed that, for a class of cases, the optimal control of the processes can be obtained from an expectation over the coordinate of first entry of the uncontrolled processes into the

[1] Research supported by the Natural Sciences and Engineering Council of Canada. Grant no. A7989.

termination region. Whittle has generalized this result to the case of processes with non-linear dynamics. (See also Lefebvre [3] and [4].) In this note, a theorem is established which relates the optimal control of processes with non-linear/Gaussian dynamics and *exponential* cost criterion to an expectation for the uncontrolled processes.

So, consider the continuous time dynamic system with process equation

$$dx/dt = a + Bu + \varepsilon, \tag{1}$$

where the state variable $x$ is in $R^n$, the control variable $u$ is in $R^m$, $a$ is an arbitrary $n$-vector function of $\xi = (x,t)$ and $\varepsilon$ is Gaussian white noise of zero mean and covariance rate $N$. The $n$-square matrix $N$ and the $n \times m$ matrix $B$ may be $\xi$-dependent. The aim is to minimize the expected value of the negative exponential cost function (see Jacobson [1])

$$J(\xi) = -\exp[-\int_t^\tau (u'Qu)/2 \, ds - K(x(\tau),\tau)], \tag{2}$$

where the positive definite $m$-square matrix $Q$ may possibly be $\xi$-dependent, and $\tau$ is the first moment at which $x(t)$ enters a termination region $D$, having started from $\xi$.

THEOREM. *Suppose that ultimate entry of the uncontrolled process*

$$dx/dt = a + \varepsilon \tag{3}$$

*into a prescribed termination region* $D$ *is certain, and that* $\xi = (x,t)$ *and* $u$ *are unrestricted. Then if* $Q$ *is chosen so that*

$$N = BQ^{-1}B' \tag{4}$$

*the optimal control is given by*

$$u = Q^{-1}B'\Phi_x/2\Phi, \tag{5}$$

*where*

(6) $$\Phi(\xi) = E_\xi[\exp(-2K(x(\tau),\tau)],$$

*the expectation being over the time* $\tau$ *and coordinate* $x(t)$ *of first passage into* $D$ *for the uncontrolled process* (3), *conditional on starting from* $\xi$.

PROOF. Let $F(\xi)$ be the minimal expected cost incurred from position $\xi$; that is,

(7) $$F(\xi) = \inf_u E_\xi[-\exp(-\int_t^\tau (u'Qu)/2\, ds)] \qquad (C),$$

where the (C) indicates that the equation holds in the continuation region $C$, the complement of $D$ in the domain of definition of $x(t)$. Correspondingly, we have the boundary condition

(8) $$F(\xi) = -\exp[-K(\xi)] \qquad (D).$$

Then we can write for $\xi \in C$

$$F(\xi) = \inf_u E_\xi[-\exp\{-\Delta(u'Qu)/2 - \int_{t+\Delta}^\tau (u'Qu)/2\, ds + o(\Delta)\}].$$

Using the principle of optimality and Taylor's formula in $n$ variables we may write

$$F(\xi) = \inf_u\{[1 - \Delta(u'Qu)/2 + o(\Delta)][F(\xi) + \Delta(a+Bu)'F_x + \Delta F_t + \Delta tr(NF_{xx})/2 + o(\Delta)][1 + o(\Delta)]\},$$

where $F_x$ and $F_{xx}$ are the column vector of first derivatives and the matrix of second derivatives of $F$, respectively. Hence we have

$$0 = \inf_u[(a+Bu)'F_x + tr(NF_{xx})/2 + F_t - (u'Qu)F/2].$$

One finds that

(9) $$u = Q^{-1}B'F_x/F$$

and it follows that

$$F_t + a'F_x + tr(NF_{xx})/2 + (F_x'BQ^{-1}B'F_x)/2F = 0.$$

That is, since $N = BQ^{-1}B'$,

(10) $$F_t + a'F_x + tr(NF_{xx})/2 + (F_x'NF_x)/2F = 0.$$

Next, let

$$\Phi(\xi) = F^2(\xi). \tag{11}$$

Then, remembering that if $A$ is $n \times p$ and $B$ is $p \times n$

$$\mathrm{tr}(AB) = \mathrm{tr}(BA),$$

we find that (10) is linearized to

$$\Phi_t + a'\Phi_x + \mathrm{tr}(N\Phi_{xx})/2 = 0 \tag{12}$$

and, from (8), with boundary condition

$$\Phi(\xi) = \exp[-2K(\xi)] \qquad (\xi \in D). \tag{13}$$

Now equation (12) is in fact the Kolmogorov backward equation that the expectation (6) satisfies and (13) is the appropriate boundary condition.

Finally, since we have assumed ultimate absorption to be certain, the solution of (12) and (13) is unique and the theorem follows. □

Suppose now that the vector function $a$ is independent of $t$, and that the matrices $B$, $N$ and $Q$ and the terminal cost function $K$ are also time-invariant. Let $\lambda$ be a real parameter and consider the cost function

$$J(x) = -\exp\{-\int_0^\tau [(u'Qu)/2 + \lambda]\, dt - K(x(\tau))\}. \tag{14}$$

We have the following result.

COROLLARY. *Under the same hypotheses as in the theorem above, the optimal control is given by*

$$u = Q^{-1}B'\Phi_x/2\Phi, \tag{15}$$

*where*

$$\Phi(x) = E_x\{\exp[-2\lambda\tau - 2K(x(\tau))]\}. \tag{16}$$

When the parameter $\lambda$ is greater than zero, the aim is to leave the continuation region as soon as possible, whereas when $\lambda$ is negative, one is trying to maximize the time spent in the continuation region, taking into account the control costs $(u'Qu)/2$ and terminal cost $K$ thus incurred in both cases. When $\lambda$ is equal to zero, one obviously takes $u \equiv 0$.

## 2. An example

Consider the one-dimensional process equation

$$dx/dt = x + bu + \varepsilon, \tag{17}$$

where $b$ is a constant and $\varepsilon$ is Gaussian white noise of zero mean and covariance rate $N = 2x^2$. Suppose that we wish to minimize the expected value of the cost function (or maximize the expected reward)

$$J(x) = -\exp\{-\int_0^\tau [(u'Qu)/2 + \lambda] \, dt\}, \tag{18}$$

where $\tau$ is the first moment at which the process $x(t)$ enters the termination region $D = [d,\infty)$, having started from $x(0)$ in the interval $(0,d)$. We take $\lambda$ positive; so we encourage early departure from the region $C = (0,d)$. Ultimate entry of the uncontrolled process into $D$ being certain, we deduce from the Corollary that if we choose

$$Q = b^2/2x^2, \tag{19}$$

then the optimal control is

$$u = (b/Q)\Phi_x/2\Phi$$

i.e.,

$$u = (x^2/b)\Phi_x/\Phi \qquad (b \neq 0). \tag{20}$$

The function $\Phi$ can be obtained either by solving the equation

$$-2\lambda\Phi + x\Phi_x + x^2\Phi_{xx} = 0 \qquad x \in (0,d), \tag{21}$$

subject to

$$\Phi(d) = 1, \tag{22}$$

or by evaluating the mathematical expectation

$$E[\exp(-2\lambda\tau) \mid x(0) = x].$$

Indeed, in this case the process $x(t)$ has lognormal transitions and its first-passage time density to the stopping region $D$ is given by (see Kannan [2])

$$g(d,\tau;x,0) = [\log(d/x)/2\pi^{1/2}\pi^{3/2}]\exp\{[-\log^2(d/x)]/4\tau\} \tag{23}$$

and $\Phi$ is just the Laplace transform of $g$. One easily finds that

$$\Phi(x) = (x/d)^{\theta} \qquad x \in (0,d], \tag{24}$$

where

$$\theta = (2\lambda)^{\frac{1}{2}}. \tag{25}$$

Hence the optimal control is

$$u = (\theta/b)x. \tag{26}$$

The optimal control is thus linear in $x$.

If the continuation region is the interval $(d,\infty)$ and the stopping region is $(0,d]$, we find that

$$\Phi(x) = (d/x)^{\theta} \qquad x \in [d,\infty) \tag{27}$$

and

$$u = (-\theta/b)x. \tag{28}$$

## 3. Acknowledgment

The author would like to thank a referee for his judicious remarks.

## References

[1] JACOBSON, D.H., *Extensions of Linear-Quadratic Control, Optimization and Matrix Theory*, Academic Press, London, 1977.

[2] KANNAN, D., *An Introduction to Stochastic Processes*, North Holland, New York, 1979.

[3] LEFEBVRE, M., *Models of Hazard Survival*, Ph.D. Thesis, University of Cambridge, 1983.

[4] LEFEBVRE, M. and WHITTLE, P., *Survival optimization for a dynamic system* (submitted for publication).

[5] WHITTLE, P., *Optimization over Time*, Vol. I, Wiley, Chichester, 1982.

[6] WHITTLE, P. and GAIT, P.A., *Reduction of a class of stochastic control problems*, J. Inst. Maths. Applcs., 6(1970), 131-140.

*Département de mathématiques appliquées*
*Ecole polytechnique*
*Campus de l'Université de Montréal*
*Case postale 6079, Succ. "A"*
*Montréal, Québec H3C 3A7*

*Manuscrit reçu le 3 février 1986.*
*Revision le 15 avril 1986.*