https://doi.org/10.52326/ic-ecco.2022/CS.09

An Inverse Stochastic Optimal Control Problem

Mario Lefebvre [1], ORCID: 0000-0002-9451-543X
[1] Polytechnique Montréal, 2900 boul. Édouard-Montpetit, Montréal, Canada, mlefebvre@polymtl.ca,
https://www.polymtl.ca/expertises/en/lefebvre-mario

***Abstract*—The problem of controlling a compound Poisson process until it leaves an interval is considered. In this paper, instead of choosing the density function of the jumps and trying to find the corresponding value function, from which the optimal control follows at once, we consider the inverse problem: we fix the value of the value function and we look for admissible density functions for the jumps.**

Keywords—homing problem; Poisson random jumps; first-passage time; integro-differential equation; dynamic programming

I. Introduction

In [1], the author considered the controlled jump-diffusion process $\{X_u(t), t \geq 0\}$ defined by

$$X_u(t) = X_u(0) + \mu t + \int_0^t b[X_u(s)]\, u[X_u(s)]\, \mathrm{d}s + \sigma B(t) + \sum_{i=1}^{N(t)} Y_i, \qquad (1)$$

where $\mu \in |$ and $\sigma > 0$ are constants, $b(\cdot) \neq 0$, $u(\cdot)$ is the control variable, $\{B(t), t \geq 0\}$ is a standard Brownian motion and $\{N(t), t \geq 0\}$ is a Poisson process with rate λ. Moreover, $\{B(t), t \geq 0\}$ and $\{N(t), t \geq 0\}$ are independent, and the random variables $Y_1, Y_2, \ldots$ are independent and identically distributed. The aim was to find the control that minimizes the expected value of the cost criterion

$$J(x) := \int_0^{T(x)} \left\{ \frac{1}{2} q[X_u(t)]\, u^2[X_u(t)] + \theta \right\} \mathrm{d}t + K[X_u(T(x))], \qquad (2)$$

where θ is a real constant, $q(\cdot) > 0$, $K(\cdot)$ is a general termination cost function and the final time $T(x)$ is a random variable (known as a *first-passage time* in probability theory) defined by

$$T(x) = \inf\{t \geq 0 : X_u(t) \int (a, b) \mid X_u(0) = x\}, \qquad (3)$$

where $x \in [a, b]$. Explicit solutions to particular problems were obtained when $Y_1, Y_2, \ldots$ are exponentially distributed, which implies that the random jumps are always positive, and we assume that the ratio $b^2(x)/q(x)$ is a (positive) constant:

$$\kappa := \frac{b^2(x)}{2q(x)} \qquad (4)$$

This type of stochastic control problem, in which the final time is a first-passage time, is called a *homing problem*; see Whittle [2] and/or Whittle [3]. In the case when the parameter θ is positive (respectively, negative), the optimizer must try to minimize (respectively, maximize) the time spent by the controlled process in the continuation region (a, b), taking the quadratic control costs $q[X_u(t)]\, u^2[X_u(t)]/2$ and the termination cost $K[X_u(T(x))]$ into account.

In this paper, we set $\sigma = 0$. It follows that $\{X_u(t), t \geq 0\}$ becomes a controlled compound Poisson process (with drift μ); see, for example, Ross [4]. Moreover, we define

$$T(x) = \inf\{t \geq 0 : X_u(t) \geq 1 \mid X_u(0) = x \leq 1\}. \qquad (5)$$

That is, the continuation region is the interval $(-\infty, 1)$.

II. Optimal Control

Let $f_Y(y)$ be the common density function of the random variables $Y_1, Y_2, \ldots$ The following result is an extension of Proposition 2.1 in [1].

Remark. The function $f_Y(y)$ can involve the Dirac delta function. That is, Y can be a discrete or a mixed type random variable.

Proposition 2.1. *The value function*

$$V(x) := \inf_{\substack{u[X_u(t)] \\ 0 \le t \le T(x)}} E[J(x)] \tag{6}$$

satisfies the first-order, non-linear integro-differential equation

$$0 = \theta + \mu V'(x) - \frac{1}{2}\frac{b^2(x)}{q(x)}\left[V'(x)\right]^2 + \lambda \int_0^\infty [V(x+y) - V(x)]\, f_Y(y)\,\mathrm{d}y \tag{7}$$

for x < 1. *Moreover, this equation is subject to the boundary condition*

$$V(x) = K(x) \quad \text{if } x \ge 1. \tag{8}$$

$$\begin{aligned} V(x) &= \inf_{\substack{u[X_u(t)] \\ 0\le t\le \Delta t}} E\Bigg[\int_0^{\Delta t}\left\{\frac{1}{2}q[X_u(t)]\,u^2[X_u(t)] + \theta\right\} \\ &\quad + V\left(x + [\mu + b(x)\,u(x)]\,\Delta t + \sum_{i=1}^{N(\Delta t)} Y_i\right)\Bigg] \\ &= \inf_{\substack{u[X_u(t)] \\ 0\le t\le \Delta t}} E\Bigg[\left\{\frac{1}{2}q(x)\,u^2(x) + \theta\right\}\Delta t \\ &\quad + V\left(x + [\mu + b(x)\,u(x)]\,\Delta t + \sum_{i=1}^{N(\Delta t)} Y_i\right) \\ &\quad + o(\Delta)\Bigg]. \end{aligned}$$

Proof. Bellman's principle of optimality implies that

(9)

Moreover, we deduce from the properties of the Poisson distribution that

$$P[N(\Delta t) = 0] = e^{-\lambda\Delta t} = 1 - \lambda\Delta t + o(\Delta t) \tag{10}$$

and

$$P[N(\Delta t) = 1] = \lambda\Delta t\; e^{-\lambda\Delta t} = \lambda\Delta t + o(\Delta t). \tag{11}$$

Hence, we can write that

$$\begin{aligned} &E\left[V\left(x + [\mu + b(x)\,u(x)]\,\Delta t + \sum_{i=1}^{N(\Delta t)} Y_i\right)\right] = \\ &E\left[V\Big(x + [\mu + b(x)\,u(x)]\,\Delta t + Y_1\Big)\right]\lambda\Delta t \\ &+ E\left[V\Big(x + [\mu + b(x)\,u(x)]\,\Delta t\Big)\right](1 - \lambda\Delta t) \\ &+ o(\Delta t). \end{aligned} \tag{12}$$

Next, assuming that the function $V(x)$ is differentiable with respect to x, Taylor's formula yields that

$$\begin{aligned} &E\left[V\Big(x + [\mu + b(x)\,u(x)]\,\Delta t\Big)\right](1 - \lambda\Delta t) = \\ &V(x) + [\mu + b(x)\,u(x)]\,\Delta t\, V'(x) - \lambda\Delta t\, V(x) \\ &\quad + o(\Delta t) \end{aligned} \tag{13}$$

and

$$\begin{aligned} &E\left[V\Big(x + [\mu + b(x)\,u(x)]\,\Delta t + Y_1\Big)\right]\lambda\Delta t = \\ &E[V(x + Y_1)]\,\lambda\Delta t + o(\Delta t) = \\ &\lambda\Delta t \int_0^\infty V(x+y)\, f_Y(y)\,\mathrm{d}y + o(\Delta t). \end{aligned} \tag{14}$$

It follows, substituting (13) and (14) into Eq. (9), that

$$\begin{aligned} 0 = \inf_{\substack{u[X_u(t)] \\ 0\le t\le \Delta t}} \Bigg\{ &\left(\frac{1}{2}q(x)\,u^2(x) + \theta\right)\Delta t \\ &+ [\mu + b(x)\,u(x)]\,\Delta t\, V'(x) - \lambda\Delta t\, V(x) \\ &+ \lambda\Delta t \int_{-\infty}^{\infty} V(x+y)\, f_Y(y)\,\mathrm{d}y + o(\Delta t)\Bigg\} \end{aligned} \tag{15}$$

Dividing each side of Eq. (15) by Δt and letting Δt decrease to zero, we obtain the *dynamic programming equation*

$$\begin{aligned} 0 = \inf_{u(x)} \Bigg\{ &\frac{1}{2}q(x)\,u^2(x) + \theta + [\mu + b(x)\,u(x)]\,V'(x) \\ &- \lambda V(x) + \lambda \int_{-\infty}^{\infty} V(x+y)\, f_Y(y)\,\mathrm{d}y\Bigg\}. \end{aligned} \tag{16}$$

From the above equation, we deduce that the optimal control $u^*(x)$ can be expressed in terms of the value function $V(x)$:

$$u^*(x) = -\frac{b(x)}{q(x)}\,V'(x) \tag{17}$$

for $x < 1$. Equation (7) is obtained by replacing this expression for $u^*(x)$ into Eq. (16) and by noticing that we can write

$$V(x) = \int_0^\infty V(x)\,\alpha\, e^{-\alpha y}\mathrm{d}y. \tag{18}$$

Finally, because the jumps are strictly positive, the controlled process $\{X_u(t),\ t \geq 0\}$ cannot be smaller than the endpoint a of the interval $[a, b]$. However, it can cross the boundary $x = b$. Hence, we obtain the boundary condition in Eq. (8). □

Now, in this paper, instead of choosing the density function $f_Y(y)$ and trying to find the corresponding value function $V(x)$ (and hence the optimal control from (17)), we consider the inverse problem: we fix the value of $V(x)$ and we look for admissible density functions $f_Y(y)$.

First, assume that the value function $V(x)$ is a constant V_0 in the interval $(-\infty,1)$ (so that $u^*(x) \equiv 0$). Then, we deduce from (7) that we must have

$$\begin{aligned} 0 &= \theta - \lambda V_0 + \lambda \left\{ \int_{-\infty}^{1-x} V_0 f_Y(y)\,\mathrm{d}y \right. \\ &\left. + \int_{1-x}^{\infty} K(x+y) f_Y(y)\,\mathrm{d}y \right\}. \end{aligned} \tag{19}$$

If the function $K(x)$ is also equal to the constant V_0, the above equation reduces to

$$0 = \theta - \lambda V_0 + \lambda V_0 \int_{-\infty}^{\infty} f_Y(y)\,\mathrm{d}y. \tag{20}$$

Thus, this solution is valid for any (non-defective) density function $f_Y(y)$ if and only if $\theta = 0$.

Remark. When $\theta = 0$ and $K(x)$ is also equal to V_0, the optimal control is trivially identical to zero. However, when $K(x) \equiv K_0 \neq V_0$, Eq. (19) becomes

$$0 = \theta - \lambda V_0 + \lambda\,\{V_0 F_Y(1-x) + K_0[1 - F_Y(1-x)]\}, \tag{21}$$

where $F_Y(\cdot)$ denotes the common cumulative distribution function of the random variables Y_1, Y_2, ... Then, the function $F_Y(1-x)$ must not depend on x. This is true if the jumps are always negative, so that $F_Y(y) = 1$ for any $y \geq 0$. It follows that the solution is valid again if and only if $\theta = 0$.

Next, we look for value functions $V(x)$ that are affine functions of x:

$$V(x) = c_1 x + c_0, \tag{22}$$

where $c_1 \neq 0$, so that

$$u*(x) = -c_1 \frac{b(x)}{q(x)}. \tag{23}$$

Remark. We assumed that the ratio $b^2(x)/2q(x)$ is a positive constant κ. However, the ratio $b(x)/q(x)$ is not necessarily a constant. Hence, the optimal control is not necessarily a constant either.

Substituting the function defined in (22) into the integro-differential equation (7), we obtain that

$$\begin{aligned} 0 &= \theta + \mu c_1 - \kappa c_1^2 - \lambda(c_1 x + c_0) \\ &\quad + \lambda \int_{-\infty}^{\infty} V(x+y) f_Y(y)\,\mathrm{d}y. \end{aligned} \tag{24}$$

The simplest case is when we choose $K(x) = c_1 x + c_0$ for any $x \geq 1$. Then, the above equation becomes

$$\begin{aligned} 0 &= \theta + \mu c_1 - \kappa c_1^2 - \lambda(c_1 x + c_0) \\ &\quad + \lambda \int_{-\infty}^{\infty} [c_1(x+y) + c_0] f_Y(y)\,\mathrm{d}y \\ &= \theta + \mu c_1 - \kappa c_1^2 + \lambda c_1 E[Y]. \end{aligned} \tag{25}$$

Thus, assuming that $E[Y]$ exists (and is finite), we deduce that the solution is valid for any random variable Y if and only if the constant c_1 is given by

$$c_1 = \frac{(\mu + \lambda E[Y]) \pm \sqrt{(\mu + \lambda E[Y])^2 + 4\theta\kappa}}{2\kappa} \tag{26}$$

whereas the constant c_0 is arbitrary. The parameter θ should be such that the term in the square root is non-negative.

In the special case when $E[Y] = 0$, Eq. (26) simplifies to

$$c_1 = \frac{\mu \pm \sqrt{\mu^2 + 4\theta\kappa}}{2\kappa} \tag{27}$$

which reduces further to $c_1 = \pm\,(\theta/\kappa)^{1/2}$ when μ is equal to zero as well.

Remark. When θ is positive, the value function $V(x)$ must also be positive. Therefore, the constant c_1 must be chosen appropriately.

Finally, when $K(x) \equiv 0$, Eq. (24) becomes

$$\begin{aligned} 0 &= \theta + \mu c_1 - \kappa c_1^2 - \lambda(c_1 x + c_0) \\ &\quad + \lambda \int_{-\infty}^{1-x} [c_1(x+y) + c_0] f_Y(y)\,\mathrm{d}y \\ &= \theta + \mu c_1 - \kappa c_1^2 - \lambda(c_1 x + c_0) \\ &\quad + \lambda(c_1 x + c_0) F_Y(1-x) \\ &\quad + \lambda c_1 \int_{-\infty}^{1-x} y f_Y(y)\,\mathrm{d}y. \end{aligned} \tag{28}$$

As in the case when $V(x) \equiv V_0$, we would like the distribution function $F_Y(1-x)$ to be independent of x. We can again assume that the jumps are always negative, which implies that $F_Y(1-x) \equiv 1$ for $x < 1$. It follows that the constant c_1 must satisfy the equation

$$\begin{aligned} 0 &= \theta + \mu c_1 - \kappa c_1^2 + \lambda c_1 \int_{-\infty}^{1-x} y f_Y(y)\,\mathrm{d}y \\ &= \theta + \mu c_1 - \kappa c_1^2 + \lambda c_1 \int_{-\infty}^{0} y f_Y(y)\,\mathrm{d}y \\ &= \theta + \mu c_1 - \kappa c_1^2 + \lambda c_1 E[Y]. \end{aligned} \tag{19}$$

Hence, we retrieve the equation for c_1 in (26). However, because we assumed that $Y < 0$, we must impose the condition $\mu > 0$, otherwise $T(x)$ will be equal to infinity.

III. Conclusion

In this paper, we considered an inverse LQG homing problem for one-dimensional jump-diffusion processes. Whereas in a previous paper the jump-size distribution was fixed and the aim was to solve the dynamic programming equation (and hence obtaining the optimal control), here we tried to find admissible density functions for the jumps when the form of the value function $V(x)$ is chosen.

We were able to obtain explicit solutions to our problem in the case when $V(x)$ is a constant or an affine function of x. We could try to generalize our results to a polynomial function.

Finally, either this problem or the original one (with the constant $\sigma > 0$) could be considered in two or more dimensions.

References

[1] M. Lefebvre, “Exact solutions to optimal control problems for Wiener processes with exponential jumps,” *Journal of Stochastic Analysis*, vol. 2, no. 2, Article 1 (9 pages), 2021. DOI: 10.31390/josa.2.2.01

[2] P. Whittle, *Optimization over Time*, Vol. I, Chichester: Wiley, 1982.

[3] P. Whittle, *Risk-sensitive Optimal Control*, Chichester: Wiley, 1990.

[4] S. M. Ross, *Introduction to Probability Models*, 12th ed., Amsterdam: Elsevier/Academic Press, 2019. DOI: 10.1016/C2017-0-01324-1.