# An explicit solution to a discrete-time stochastic optimal control problem

MARIO LEFEBVRE
Department of Mathematics and Industrial Engineering
Polytechnique Montréal
2500, chemin de Polytechnique, Montréal (Québec) H3T 1J4
CANADA
https://www.polymtl.ca/expertises/en/lefebvre-mario

*Abstract:* The problem of controlling a one-dimensional Markov chain until is leaves a given set $C$ is considered. The optimizer tries to minimize the time spent by the Markov chain inside $C$. The control variable can take two different values. An exact formula is obtained for the value function, from which the optimal control is deduced.

*Key-Words:* Markov chain, homing problem, value function, difference equations, absorption problems.

Received: July 12, 2022. Revised: March 7, 2023. Accepted: March 19, 2023. Published: April 26, 2023.

## 1 Introduction

Assume that the controlled discrete-time stochastic process $\{X_n, n = 0, 1, \ldots\}$ is such that $X(0) = x$ and

$$X_{n+1} = X_n + u_n + \epsilon_n, \tag{1}$$

where $u_n$ is the control variable and $\epsilon_n$ is a random variable. We assume that both $u_n$ and $\epsilon_n$ can take a finite number of integer values. Thus, $\{X_n, n = 0, 1, \ldots\}$ is a controlled one-dimensional Markov chain.

In [4], the authors considered the following problem: find the value of the control variable that minimizes the expected value of the cost function

$$J(x) = \sum_{n=0}^{T(x)-1} (u_n^2 + \lambda), \tag{2}$$

where

$$T(x) := \inf\{n > 0 : X_n \in D \mid X_0 = x \in C\}, \quad (3$$

in which $D$ is a subset of $\mathbb{Z}$ and $\lambda$ is a non-zero constant. The random variable $T(x)$ is a called a *first-passage time* in probability theory. The set $D$ is the stopping region, which is the complement of the continuation region $C$.

This type of problem, in which the optimizer tries to either minimize (if $\lambda > 0$)or maximize (if $\lambda < 0$) the time spent by a controlled stochastic process in a given region, while taking the control costs into account, is known as a *homing problem*. [6], considered this problem for $n$-dimensional diffusion processes. In [7], the author treated the case when the cost criterion is risk-sensitive; see also [1], [5]. Recent papers on homing problems include the following ones by the author: [2], [3].

In [4], authors found the exact solutions to two particular problems. They also considered the following case: assume that the set $C$ is $\{-k + 1, \ldots, k - 1\}$, where $k \geq 2$, $u_n \in \{-2, -1, 1, 2\}$ and $\epsilon_n = \pm 1$ with probability 1/2. Moreover, the parameter $\lambda$ in the cost function is strictly positive and

$$T(x) := \min\{n > 0 : |X_n| \geq k \mid X_0 = x \in C\}. (4)$$

Next, we define the *value function*

$$F(x) = \min_{u_n,\, n=0,\ldots,T(x)-1} E[J(x)]. \tag{5}$$

Using dynamic programming, we can state the following lemma (see, [4]).

**Lemma 1.1.** *The function $F(x)$ satisfies the dynamic programming equation*

$$\begin{aligned} F(x) &= \min_{u_0} \left\{ u_0^2 + \lambda + \frac{1}{2}[F(x + u_0 - 1) \right. \\ &\qquad \left. + F(x + u_0 + 1)] \right\}. \end{aligned} \tag{6}$$

*Moreover, we have the boundary condition*

$$F(x) = 0 \quad \textit{if } |x| \geq k. \tag{7}$$

By symmetry, we only have to consider the case when $x \in \{0, 1, \ldots, k - 1\}$. Moreover, because $\lambda$ is positive, it is clear that the optimal control is equal to either $+1$ or $+2$ for any $x \geq 0$. It follows that Equation (6) becomes

$$\begin{aligned} F(x) &= \min \left\{ 1 + \lambda + \frac{1}{2}[F(x) + F(x+2)], \right. \\ &\qquad \left. 4 + \lambda + \frac{1}{2}[F(x+1) + F(x+3)] \right\}. \end{aligned} \tag{8}$$

In [4], the authors showed how we can compute the optimal control $u_0^*(x)$ first for $x = k-1$, then for $x = k-2$ and $x = k-3$. However, as they mentioned in their paper, it is not easy to give a general formula for any $x \in \{0, 1, \ldots, k-1\}$ in terms of the parameter $\lambda$. In the current paper, we will show that we can give a general expression for the value function $F(x)$ for any fixed value of $\lambda$. We can also then deduce the value of the optimal control $u_0^*(x)$.

# 2 An explicit expression for the value function

Let us denote the function $F(x)$ by $F_i(x)$ if the optimizer chooses $u_0 = i$, for $i = 1, 2$. We deduce from Eq. (8) that we have

$$F_1(x) = 1 + \lambda + \frac{1}{2}\left[F_1(x) + F_1(x+2)\right] \quad (9)$$

and

$$F_2(x) = 4 + \lambda + \frac{1}{2}\left[F_2(x+1) + F_2(x+3)\right]. \quad (10)$$

We can find the general solution of both difference equations. First, Eq. (9) is a second-order linear difference equation with constant coefficients:

$$F_1(x+2) - F_1(x) + 2(1+\lambda) = 0. \quad (11)$$

Its general solution can be written as follows:

$$F_1(x) = c_1(-1)^x + c_2 - (1+\lambda)x, \quad (12)$$

where $c_1$ and $c_2$ are arbitrary constants. To determine the values of $c_1$ and $c_2$, we impose the conditions $F_1(k) = F_1(k+1) = 0$. Moreover, we set $F_1(k+2) = 0$.

Next, Eq. (10) is a third-order linear difference equation with constant coefficients:

$$F_2(x+3) + F_2(x+1) - 2F_2(x) + 2(4+\lambda) = 0. \quad (13)$$

We find that

$$\begin{aligned} F_2(x) \;=\; & d_1 + d_2\left(-\frac{1}{2} + \frac{\sqrt{7}i}{2}\right)^x \\ & + d_3\left(-\frac{1}{2} - \frac{\sqrt{7}i}{2}\right)^x - \frac{4+\lambda}{2}x, \quad (14) \end{aligned}$$

where $d_1$, $d_2$ and $d_3$ are constants that are determined from the boundary conditions $F_2(k) = F_2(k+1) = F_2(k+2) = 0$.

*Remarks.* (i) Even though the expression for the function $F_2(x)$ contains complex terms, it is actually real for any integer $x \in \{0, 1, \ldots, k-1\}$.
(ii) The function $F_i(x)$ corresponds to the expected cost if we choose $u_0(x) \equiv i$, for $i = 1, 2$.

We can state the following proposition.

**Proposition 2.1.** *The value function $F(x)$ can be expressed as follows:*

$$\begin{aligned} F(x) \;=\; \min\Big\{ & 1 + \lambda + \frac{1}{2}[\min\{F_1(x), F_2(x)\} \\ & + \min\{F_1(x+2), F_2(x+2)\}], 4 + \lambda \\ & + \frac{1}{2}[\min\{F_1(x+1), F_2(x+1)\} \\ & + \min\{F_1(x+3), F_2(x+3)\}]\Big\} \quad (15) \end{aligned}$$

*for $x = 0, 1, \ldots, k-1$.*

Let

$$\begin{aligned} G(x) \;:=\; & 1 + \lambda + \frac{1}{2}[\min\{F_1(x), F_2(x)\} \\ & + \min\{F_1(x+2), F_2(x+2)\}] \quad (16) \end{aligned}$$

and

$$\begin{aligned} H(x) \;:=\; & 4 + \lambda + \frac{1}{2}[\min\{F_1(x+1), F_2(x+1)\} \\ & + \min\{F_1(x+3), F_2(x+3)\}], \quad (17) \end{aligned}$$

so that

$$F(x) = \min\{G(x), H(x)\}. \quad (18)$$

To determine the optimal control $u_0^*(x)$ for any $x$ in $\{0, 1, \ldots, k-1\}$, we can compare the value of $G(x)$ with that of $H(x)$.

In [4], the authors proved that the value function satisfies a non-linear difference equation.

**Proposition 2.2.** *The value function $F(x)$ satisfies the non-linear third-order difference equation*

$$\begin{aligned} 0 \;=\; & 2F^2(x) - F(x)[F(x+1) + 2F(x+2) \\ & + F(x+3) + 12 + 6\lambda] \\ & + 2(4+\lambda)F(x+2) + 2(1+\lambda)[F(x+1) \\ & + F(x+3)] + F(x+1)F(x+2) \\ & + F(x+2)F(x+3) + (5+2\lambda)^2 - 9 \quad (19) \end{aligned}$$

*for $x = 0, 1, \ldots, k-1$. The boundary conditions are $F(x) = 0$ if $x = k, k+1, k+2$.*

*Remark.* There are a few misprints in [4].

## 2.1 A particular problem

Assume that $k = 4$. We find that

$$F_1(x) = -\frac{1}{2}(1+\lambda)(-1)^x + \frac{9}{2}(1+\lambda) - (1+\lambda)x \quad (20)$$

and that the constants $d_1$, $d_2$ and $d_3$ in Eq. (14) are given by

$$d_1 = \frac{19}{8}(4+\lambda), \quad d2 = (4+\lambda)\frac{(3i - \sqrt{7})\sqrt{7}}{56i\sqrt{7} + 168} \quad (21)$$

Table 1: Functions $F(x)$, $F_1(x)$, $F_2(x)$, $G(x)$ and $H(x)$, and optimal control $u_0^*(x)$ for $x = 0, 1, 2, 3$ when $\lambda = 1$

| $x$ | $F(x)$ | $F_1(x)$ | $F_2(x)$ | $G(x)$ | $H(x)$ | $u_0^*(x)$ |
|---|---|---|---|---|---|---|
| 0 | 8 | 8 | 11.875 | 8 | 11 | 1 |
| 1 | 7 | 8 | 8.75 | 8 | 7 | 2 |
| 2 | 4 | 4 | 7.5 | 4 | 7 | 1 |
| 3 | 4 | 4 | 5 | 4 | 5 | 1 |

Table 2: Functions $F(x)$, $F_1(x)$, $F_2(x)$, $G(x)$ and $H(x)$, and optimal control $u_0^*(x)$ for $x = 0, 1, 2, 3$ when $\lambda = 2$

| $x$ | $F(x)$ | $F_1(x)$ | $F_2(x)$ | $G(x)$ | $H(x)$ | $u_0^*(x)$ |
|---|---|---|---|---|---|---|
| 0 | 12 | 12 | 14.25 | 12 | 14.25 | 1 |
| 1 | 9 | 12 | 10.5 | 11.25 | 9 | 2 |
| 2 | 6 | 6 | 9 | 6 | 9 | 1 |
| 3 | 6 | 6 | 6 | 6 | 6 | 1 or 2 |

and

$$d_3 = -(4+\lambda)\,\frac{i\sqrt{7}}{56}. \tag{22}$$

Table 1, Table 2, Table 3, Table 4 give the value function $F(x), F_1(x), F_2(x), G(x), H(x)$ and the optimal control $u_0^*(x)$ for $x = 0, 1, 2, 3$ for various values of the parameter $\lambda$. Notice that, as expected, when $\lambda$ is large, the optimal control is most often $u_0^*(x) = 2$.

To conclude this section, we will check that the values of the function $F(x)$ given in Table 1 (and using the fact that $F(x) = 0$ for $x \geq 4$) are such that Eq. (19) with $\lambda = 1$ is indeed satisfied for $x = 0, 1, 2, 3$. First, when $x = 0$, we have

$$\begin{aligned} 0 &= 2\times 8^2 - 8(7+2\times 4+4+12+6)+10\times 4 \\ &\quad +4(7+4)+7\times 4+4\times 4+40. \end{aligned} \tag{23}$$

Similarly, for $x = 1$, $x = 2$ and $x = 3$ we have

Table 3: Functions $F(x)$, $F_1(x)$, $F_2(x)$, $G(x)$ and $H(x)$, and optimal control $u_0^*(x)$ for $x = 0, 1, 2, 3$ when $\lambda = 5$

| $x$ | $F(x)$ | $F_1(x)$ | $F_2(x)$ | $G(x)$ | $H(x)$ | $u_0^*(x)$ |
|---|---|---|---|---|---|---|
| 0 | 21.375 | 24 | 21.375 | 22.6875 | 21.375 | 2 |
| 1 | 15 | 24 | 15.75 | 18.375 | 15 | 2 |
| 2 | 12 | 12 | 13.5 | 12 | 13.5 | 1 |
| 3 | 9 | 12 | 9 | 10.5 | 9 | 2 |

Table 4: Functions $F(x)$, $F_1(x)$, $F_2(x)$, $G(x)$ and $H(x)$, and optimal control $u_0^*(x)$ for $x = 0, 1, 2, 3$ when $\lambda = 10$

| $x$ | $F(x)$ | $F_1(x)$ | $F_2(x)$ | $G(x)$ | $H(x)$ | $u_0^*(x)$ |
|---|---|---|---|---|---|---|
| 0 | 33.25 | 44 | 33.25 | 38.125 | 33.25 | 2 |
| 1 | 24.5 | 44 | 24.5 | 30.25 | 24.5 | 2 |
| 2 | 21 | 22 | 21 | 21.5 | 21 | 2 |
| 3 | 14 | 22 | 14 | 18 | 14 | 2 |

respectively

$$\begin{aligned} 0 &= 2\times 7^2 - 7(4+2\times 4+0+12+6)+10\times 4 \\ &\quad +4(4+0)+4\times 4+4\times 0+40, \end{aligned} \tag{24}$$

$$\begin{aligned} 0 &= 2\times 4^2 - 4(4+2\times 0+0+12+6)+10\times 0 \\ &\quad +4(4+0)+4\times 0+0\times 0+40 \end{aligned} \tag{25}$$

and

$$\begin{aligned} 0 &= 2\times 4^2 - 4(0+2\times 0+0+12+6)+10\times 0 \\ &\quad +4(0+0)+0\times 0+0\times 0+40. \end{aligned} \tag{26}$$

# 3 Conclusion

In this note, we gave an explicit and exact expression for the value function $F(x)$ in an optimal control problem for a discrete-time and discrete-state Markov chain that was considered by Lefebvre and Kounta [4]. From this expression, it is possible to determine the optimal control $u_0^*(x)$ for any value of $x$ in the set $\{0, 1, \ldots, k-1\}$. Moreover, by symmetry, we can write that $u_0^*(-x) = -u_0^*(-x)$.

We saw that the function $F(x)$ satisfies a nonlinear third-order difference equation. Solving such equations directly is a very difficult task. However, we checked in a particular case that the values obtained for $F(x)$ are indeed such that the difference equation is satisfied.

The results presented in this note can be generalized to the case when the control variable can take more than two values. We could also consider this type of problem in two or more dimensions.

*References:*

[1] J. Kuhn, The risk-sensitive homing problem, *Journal of Applied Probability*, Vol. 22, 1985, pp. 796-803. https://doi.org/10.2307/3213947

[2] M. Lefebvre, Minimizing or maximizing the first-passage time to a time-dependent boundary, *Optimization*, Vol. 71, No. 2, 2022, pp. 387-401. https://doi.org/10.1080/02331934.2021.1914039

[3] M. Lefebvre, The homing problem for autoregressive processes, *IMA Journal of Mathematical Control and Information*, Vol. 39, No. 1, 2022, pp. 322-344. https://doi.org/10.1093/imamci/dnab047

[4] M. Lefebvre and M. Kounta, Discrete homing problems, *Archives of Control Sciences*, Vol. 23, No. 1, 2013, pp. 5-18. https://doi.org/10.2478/v10170-011-0039-6

[5] C. Makasu, Risk-sensitive control for a class of homing problems, *Automatica*, Vol. 45, No. 10, 2009, pp. 2454-2455. https://doi.org/10.1016/j.automatica.2009.06.015

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

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

## Contribution of individual authors to the creation of a scientific article (ghostwriting policy)

The author did all the research work of this study.

## Sources of funding for research presented in a scientific article or scientific article itself

This research was supported by the Natural Sciences and Engineering Research Council of Canada.

## Conflict of Interest

The author has no conflict of interest to declare that is relevant to the content of this article.

## Creative Commons Attribution License 4.0 (Attribution 4.0 International , CC BY 4.0)