Optimal Stopping with Random Lag under General Markov Processes

Yingda Song  Pengzhan Chen
University of Science and Technology of China

May, 2017

Abstract
This paper studies the optimal stopping problem with infinite horizon and random lag under general Markov process, where the random lag can have arbitrary distribution. We solve the problem by first constructing a continues-time Markov chain (CTMC) to approximate the general Markov process model, and then deriving the solution of the optimal stopping problem under the CTMC as an approximation. The convergence of our approximate solution is proved. In addition, we show that the approximation formula can be efficiently computed under the diffusion models. Based on our method, we compare the optimal solutions when the time lag follows different distributions, and find that the impact of misidentifying the distribution could be crucial.

Keywords: optimal stopping problem; random lag; infinite horizon; continuous-time Markov chain

1 Introduction
Along with the development of the theory of probability and stochastic processes, one of the most important problem is the optimal stopping problem, which is trying to find the best stopping strategy to obtain the maximum reward. The standard optimal stopping problem with infinite horizon has many real-world applications and has been thoroughly studied in the literature. For example, the seminal paper of Mcdonald and Siegel (1987) studies the problem to find the optimal timing of investment in an irreversible project. This kind of problem also has been applied to study Franchise Monopoly, optimal scrapping of a project, and so on. Recently, Lee and Choi (2015) studies the problem to analyze the effect of implementation lag on development (R&D) investment of pharmaceutical companies.

However in the real world, the decision time and the implementation time may be different, which leads to an implementation lag. For example, a pharmaceutical company may decide to start new drug research and development, or set up a new factory to expand production, but it takes time to implement the decision. There are also many study on different type of investment delays. The lag considered has a relation to the asset’s value at stopping time in Alvarez and Keppo (2002). In Øksendal (2011) the authors solve the optimal stopping problem with a constant lags under strong Markov process. Costeniuc and Taschini (2008) study the optimal entry and exit decision with implementation lag under diffusion model, and the delay they add in their model is constant, too. In addition, Lempa (2012) considering the time delay of the exponential distribution and solve this type problem by resolvent operator method. They transform the non-standard optimal stopping problem into an adjusted standard problem where original payoff function is mapped with the resolvent operator. As far as we known in the literature, they only solved the problem under diffusion processes and the lag is constant and exponential. But our method can solve it under general Markov process and the random lag can have any distribution.

The contribution of this paper is mainly in the following three aspects: (1) We derive the explicit solution of optimal stopping problem with any random lag under general Markov processes by using a CTMC to approximates the Markov process. Also, the convergence of the solution has been proved in appendix B. (2) In
the diffusion processes case, the detail calculation of optimal state is provided, which contains any distributed random lag. (3) We compare our result to Lempa (2012) where the random lag is exponential distributed. And we solve an application of optimal stopping problem (see Mcdonald and Siegel (1987)) with implementation delay. Also, we designed a experiment to show the loss if the assumption of lag was wrong.

To solve this kind of non-standard stopping problem under general Markov processes $X_t$, our ideas are mainly divided into three parts: (1) By discretizing the state space we generate a CTMC $M_t$ to approximates the Markov process $X_t$; (2) We solve the optimal stopping problem with random lag under CTMC $M_t$ in explicit from; (3) Prove the solution of the problem under CTMC is convergence to the solution under Markov process $X_t$. The algorithm of generating CTMC to approximate the targeted Markov process is in Mijatovic and Pistorius (2010), and the details of the part (2) is introduced in section 3. Then, the proof of the convergence of the solution is placed in Appendix B.

The remainder of this paper is organized as follows. In section 3 the main result of the non-standard optimal stopping problem with general distributed random lag under Markov processes has been provided. In section 4 we use the method in section 3 to analysis the problem under diffusion processes as an application. In section 5 We compared the accuracy of our method with Lempa (2012) and we solve an application of optimal stopping problem with implementation lag, also an experiment is designed to show the influence of wrong assumption of the distribution of the lag. Section 5 concludes the paper. Appendix contains the algorithm of generating CTMC and some proofs.

2 Preliminary Results

First, we shall introduce the notations used throughout this paper. Suppose $A$ is an $N \times N$ matrix and $a$ is an $N \times 1$ vector. We use $A(x, y)$ to denote the entry of $A$ at the $x$-th row and $y$-th column, and $a(x)$ the $x$-th element of $a$. In addition, for a function $g$ defined on a finite set $\{m_1, m_2, \ldots, m_N\}$, we use $g$ to denote the $N \times 1$ vector with the $x$-th element $g(x) = g(m_x)$, with $x = 1, 2, \ldots, N$. We also use the conventional notation $I_n$ to denote a $n \times n$ identity matrix, and $0_{m \times n}$ the $m \times n$ matrix with all entries equal to 0, while the subscripts may be omitted if the size of the matrices can be determined trivially from the context.

Theorem 1 (A Limiting Property of the Matrix Exponential) Let $\Lambda = (\lambda_{ij})_{N \times N}$ be a transition rate matrix, such that $\lambda_{ij} \geq 0$ for $i \neq j$, and $\sum_{j=1}^{N} \lambda_{ij} = 0$ for $i = 1, \ldots, N$. $r > 0$. For $n \in \{2, \ldots, N\}$, define the $N \times N$ matrix $\tilde{\Lambda}^{(n)}_r$ by

$$
\tilde{\Lambda}^{(n)}_r(x, y) := \begin{cases} 
\Lambda(x, y) - r & \text{if } x < n \text{ and } x = y \\
\Lambda(x, y) & \text{if } x < n \text{ and } x \neq y \\
0 & \text{if } x \geq n
\end{cases}
$$

In addition, define the $(n - 1) \times (n - 1)$ matrix $A_n$ and $(n - 1) \times (N - n + 1)$ matrix $B_n$ by

$$
A_n(x, y) := \begin{cases} 
\Lambda(x, y) - r & \text{if } x = y \\
\Lambda(x, y) & \text{if } x \neq y
\end{cases}, \quad 1 \leq x, y \leq n - 1,
$$

$$
B_n(x, y) := \Lambda(x, y + n - 1), \quad 1 \leq x \leq n - 1, 1 \leq y \leq N - n + 1,
$$

Then $A_n$ is invertible and

$$
\lim_{T \to \infty} e^{rT\tilde{\Lambda}^{(n)}_r} = \begin{pmatrix} 0_{(n-1)\times(n-1)} & -A_n^{-1}B_n \\ 0_{(N-n+1)\times(n-1)} & I_{N-n+1} \end{pmatrix}.
$$
Proof. First, we show \( A_n \) is invertible. Actually, \( A_n \) is strictly diagonally dominant, since for any \( i \in \{1, \ldots, n-1\} \),

\[
|A_n(i, i)| = |\Lambda(i, i) - r| = r - \Lambda(i, i) > -\Lambda(i, i) = \sum_{j=1, j \neq i}^{N} \Lambda(i, j) \geq \sum_{j=1, j \neq i}^{n-1} \Lambda(i, j) = \sum_{j=1, j \neq i}^{n-1} |A(i, j)|.  \tag{1}
\]

By Lévy-Desplanques theorem (see, e.g., Horn and Johnson 1985, Corollary 5.6.17), we know that a strictly diagonally dominant matrix is non-singular, and hence \( A_n \) is invertible.

Next we prove the existence of \( \lim_{T \to \infty} e^{T\Lambda(n)} \). According to Gershgorin circle theorem, (see, e.g., Horn and Johnson 1985, Theorem 6.1.1), all the eigenvalues of \( \Lambda(n) \) are located in the union of \( N \) discs

\[
\bigcup_{i=1}^{N} \left\{ z \in \mathbb{C} : |z - \Lambda_{i,n}(i, i)| \leq \sum_{j=1, j \neq i}^{N} |\Lambda_{i,n}(i, j)| \right\}
\]

By the definition of \( \Lambda_{i,n} \) and a similar argument as in (1), we know for all \( i \in \{1, \cdots, N\} \),

\[
-\Lambda_{i,n}(i, i) \geq \sum_{j=1, j \neq i}^{N} |\Lambda_{i,n}(i, j)|.
\]

Therefore all the eigenvalues of \( \Lambda_{i,n} \) have non-positive real parts. Denote the distinguished non-zero eigenvalues of \( \Lambda_{i,n} \) as \( \lambda_1, \cdots, \lambda_k \). According to Jordan decomposition, there exists invertible matrix \( H \) such that \( \Lambda_{i,n} = H^{-1}JH \), where

\[
J = \begin{pmatrix}
J_1 & & \\
& \ddots & \\
& & J_k
\end{pmatrix}, \text{ with each Jordan block } J_q = \begin{pmatrix}
\lambda_q & 1 & & \\
& \lambda_q & \ddots & \\
& & \ddots & 1 \\
& & & \lambda_q
\end{pmatrix}.
\]

It follows that

\[
e^{T\Lambda_{i,n}} = H^{-1}e^{TJ}H = H^{-1} \begin{pmatrix}e^{TJ_1} & & \\
& \ddots & \\
& & e^{TJ_k}
\end{pmatrix}H.
\]

For a given \( q \),

\[
e^{TJ_q} = e^{T(\lambda_q I+N)} = e^{\lambda_q T}e^{TN} = e^{\lambda_q T} \sum_{l=0}^{\infty} \frac{T^l N^l}{l!}, \tag{2}
\]

where \( I \) is an identity matrix and \( N \) is a nilpotent matrix, both with the same size of \( J_q \). Notice that the series in (2) only has a finite number of non-zero terms. Since \( \text{Re}(\lambda_q) < 0 \), we know \( \lim_{T \to \infty} e^{TJ_q} = 0 \), and hence \( \lim_{T \to \infty} e^{T\Lambda_{i,n}} \) exits.

Last, we compute the limit. For ease of exposition, we define \( L_n := \lim_{T \to \infty} e^{T\Lambda_{i,n}} \). By L’Hôpital’s rule,

\[
L_n = \lim_{T \to \infty} \frac{e^{T} \cdot e^{T\Lambda(n)}}{e^{T}} = \lim_{T \to \infty} \left( e^{T\Lambda(n)} + \frac{d}{dT} e^{T\Lambda(n)} \right) = L_n + \Lambda_{i,n}L_n \equiv L_n + L_n\Lambda_{i,n}
\]

Therefore

\[
\Lambda_{i,n}L_n = L_n\Lambda_{i,n} = 0. \tag{3}
\]
By definition, we can write \( \tilde{\Lambda}^{(n)} = \begin{pmatrix} A_n & B_n \\ 0 & 0 \end{pmatrix} \). Using method of induction, it is easy to see that \( \left( \tilde{\Lambda}^{(n)} \right)^j \) is always of the form \( \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix} \) for positive integer \( j \). Therefore, \( L_n = \lim_{T \to \infty} e^{T \tilde{\Lambda}^{(n)}} = \lim_{T \to \infty} I + \sum_{j=1}^{\infty} \frac{T^j (\tilde{\Lambda}^{(n)})^j}{j!} \) should be of the form \( \begin{pmatrix} * & * \\ 0 & I \end{pmatrix} \). Let \( L_n = \begin{pmatrix} X & Y \\ 0 & I \end{pmatrix} \), where the block sizes correspond with \( \tilde{\Lambda}^{(n)} \). Then by (3) we get

\[
XA_n = 0, \quad A_n Y + B_n = 0
\]

As \( A_n \) is invertible, we have

\[
L_n = \begin{pmatrix} 0 & -A_n^{-1}B_n \\ 0 & I \end{pmatrix}.
\]

\[\square\]

3 General Distributed Random Lag Under Markov Processes

We consider the following optimal stopping problem:

\[
V(x) = \sup_{\tau \in \Pi(X)} E_x \left[ e^{-r (\tau + U)} g(X_{\tau + U}) \right].
\]  

(4)

Here \( \{X_t\}_{t \geq 0} \) is a one-dimensional Markov process. In this paper we assume the reward function \( g(x) \) takes the form of \( (x - K)^+ \) where \( K \) is constant, however, similar discussions can be made under other types of reward functions. \( \Pi(X) \) contains all the stopping strategies. \( U \) is a nonnegative random variable independent of \( \{X_t\}_{t \geq 0} \), representing the implementation lag. \( E_x[\cdot] \) denotes the expectation given that the initial value of the underlying process is \( x \). What we trying to do is to find the optimal stopping time \( \tau^* \in \Pi(X) \) that maximizes the present value of the reward, where the discount rate \( r \geq 0 \) is a constant.

We propose to solve this problem in three steps:

**Step1:** Construct a sequence of continue-time Markov chains \( \{M^{(n)}_t\}_{t \geq 0} (n = 1, 2, \ldots) \) that weakly converges to \( X_t \).

**Step2:** Find the solution to the new optimal stopping problem:

\[
v_n(x) = \sup_{\tau \in \Pi(M^{(n)})} E_x \left[ e^{-r (\tau + U)} g(M^{(n)}_{\tau + U}) \right]
\]  

(5)

**Step3:** Prove the solutions of problem (5) converge to the solution of problem (4).

The algorithm of generating CTMC and the proof of convergence are introduced in the Appendix A and B. Theorem 2 gives the solution to the optimal stopping problem (5). For brevity of notation, we suppress the superscript \( (n) \).

**Theorem 2 (Optimal Stopping with Random Lag under CTMC)** Let \( M_1 \) be a CTMC with finite state space \( S = \{m_1, m_2, \ldots, m_N\} \) \((m_1 < m_2 < \cdots < m_N)\) and transition rate matrix \( \Lambda \), and \( M_0 = m_{i_0}, 1 \leq i_0 < N \). Let \( U \) be a nonnegative random variable independent of \( \{M_t\}_{t \geq 0} \). For \( i_0 < i < N \), define \( \tau_i = \inf \{t \geq 0 : M_t \geq m_i\} \).
Then according to Theorem 1, \( \lim_{t \to \infty} \frac{1}{\tau} \) exists. We prove this theorem in three progressive steps.

Then, according to Theorem 1, \( \lim_{t \to \infty} \frac{1}{\tau} \) exists. We prove this theorem in three progressive steps.

In addition, if \( U \) follows a hyper-exponential distribution with density function \( f_U(x) = \sum_{j=1}^{n} p_j e^{-\alpha_j x} \) with \( \alpha_j, p_j > 0 \), \( \sum_{j=1}^{n} p_j = 1 \), then (6) can be further simplified as

\[
v(m) = \max_{m \in S} \left( \sum_{j=1}^{n} p_j (I - \alpha_j^{-1}(A - rI))^{-1} g \right)(m), \quad m \in S.
\]

Proof. We prove this theorem in three progressive steps.

(i) Consider the case without lag:

\[
\bar{v}(m) = \sup_{\tau \in \Pi(M)} E_m \left[ e^{-r \tau} g(M_{\tau}) \right]
\]

Since \( \{M_t\} \) has a finite state space, we know that \( \Pi(M) = \{\tau_1, \cdots, \tau_N\} \) has only finite elements, where \( \tau_i := \inf \{ t \geq 0 : M_t \geq m_i \} \) for \( m_i \in S \). Notice that for \( i \leq i_0 \), we have \( \tau_i = \tau_{i_0} = 0 \). For \( i > i_0 \), given \( T > 0 \), we have \( |e^{-r \tau_i T} g(M_{\tau_i T})| \leq g(m_N) < \infty \), and hence by bounded convergence theorem, we have

\[
E_m \left[ e^{-r \tau_i} g(M_{\tau_i}) \right] = \lim_{T \to \infty} E_m \left[ e^{-r \tau_i T} g(M_{\tau_i T}) \right].
\]

While according to Theorem 3.1 in Mijatovic and Pistorius (2010), \( E_m \left[ e^{-r \tau_i T} g(M_{\tau_i T}) \right] = \left( e^{T \bar{A}^{(i)} g} \right)(m) \), where

\[
\bar{A}^{(i)}(x, y) := \begin{cases} 
\Lambda(x, y) - r & \text{if } x < i, x = y \\
\Lambda(x, y) & \text{if } x < i, x \neq y \\
0 & \text{if } x \geq i
\end{cases}
\]

Notice that by definition,

\[
\bar{A}^{(i)} = \left( \begin{array}{cc}
A_i & B_i \\
0 & 0
\end{array} \right).
\]

Then according to Theorem 1, \( \lim_{T \to \infty} e^{T \bar{A}^{(i)}} = L_i \). Therefore

\[
\bar{v}(m) = \max_{i \in \{1, 2, \cdots, N\}} E_m \left[ e^{-r \tau_i} g(M_{\tau_i}) \right] = \max_{m \in S} \lim_{T \to \infty} \left( e^{T \bar{A}^{(i)} g} \right)(m) = \max_{m \in S} (L_i g)(m).
\]

(ii) Consider the case with a constant lag \( t \) as follows:

\[
\tilde{v}(m, t) = \sup_{\tau \in \Pi(M)} E_m \left[ e^{-r(\tau+t)} g(M_{\tau+t}) \right]
\]
According to the property of conditional expectation and the Markov property of CTMC \( M_t \), we have

\[
E_m \left[ e^{-r(\tau_1+t)}g(M_{\tau_1+t}) \right] = E_m \left[ E[e^{-r(\tau_1+t)}g(M_{\tau_1+t})|F_{\tau_1}] \right]
\]

\[
= E_m \left[ e^{-r\tau_1} E_{M_{\tau_1}} [e^{-rt}g(M_t)] \right]
\]

\[
= E_m \left[ e^{-r\tau_1} \tilde{g}_t(M_{\tau_1}) \right]
\]

for each \( i \in \{1, 2, \cdots, N\} \), where \( \tilde{g}_t(x) = E_x [e^{-rt}g(M_t)] = (e^{-rt}e^{At}g)(x) \). Using the result from the first step, we have

\[
\tilde{v}(m, t) = \max_{i_0 \leq t \leq N} (L_i \tilde{g}_t)(m) = \max_{t \leq i \leq N} \left(L_i e^{(A-r)^t}g\right)(m).
\]

(iii) Consider the case with random lag \( U \). Since \( U \) is independent of \( \{M_t\}_{t \geq 0} \), using the result from the second step, we have

\[
E_m \left[ e^{-r(\tau_1+U)}g(M_{\tau_1+U}) \right] = E \left[ E_m[e^{-r\tau_1}\tilde{g}_U(M_{\tau_1})|U] \right] = E \left[ (L_i e^{(A-r)U}g)(m) \right].
\]

Therefore,

\[
v(m) = \max_{0 \leq i \leq N} \left( L_i e^{(A-r)U}g \right)(m).
\]

\[\square\]

In addition, suppose \( U \) follows the hyper-geometric distribution, and we have \( E[U^k] = \sum_{j=1}^n p_j k^j \alpha_j^k \). Therefore,

\[
E \left( L_i e^{(A-r)U}g \right)(m) = \left( L_i \sum_{k=0}^\infty \frac{E(U^k)}{k!} (A-r)^k g \right)(m)
\]

\[
= \left( L_i \sum_{k=0}^\infty \sum_{j=1}^n \frac{p_j}{\alpha_j^k} (A-r)^k g \right)(m)
\]

\[
= \sum_{j=1}^n p_j L_i \left( I - \alpha_j^{-1} (A-rI) \right)^{-1} g \)(m)
\]

So

\[
v(m) = \sum_{j=1}^n p_j L_i \left( I - \alpha_j^{-1} (A-rI) \right)^{-1} g \)(m)
\]

\[\square\]

4 General Distributed Random Lag Under Diffusion Model

In this section, we first derive an more efficient solution for diffusion model by theorem 3, then we give some formulas and needed parameters such as transition rate matrix \( A \) for calculation. The statement is hereby given that the \( M_t \) on behalf of the CTMC approximates the diffusion model \( X_t \) by algorithm in appendix A.

4.1 Theoretical derivation

According to the diffusion process have almost surely continuous sample path, so the transition rate matrix \( A \) of the CTMC \( M_t \) we constructed is a tridiagonal matrix. That means the process can only transfer to the adjacent states by each step, and this property simplified our result in section 3. We show it in the following theorem.
Theorem 3 Consider a diffusion process \( \{X_t\}_{t \geq 0} \) and a sequence of CTMC \( \{M_t^{(n)}\}_{t \geq 0} \) weakly converges to it. Then we have

\[
v(m) = \max_{m_0 \leq i \leq N} E \left( e^{(A-r)U} g \right) (m_i) \times (L, 1) (m)
\]

In addition, if \( U \) follows a hyper-exponential distribution, then we have

\[
v(m) = \max_{m_0 \leq i \leq N} \left( \sum_{j=1}^{n} p_j (I - \alpha_j^{-1}(A - rI))^{-1} g \right) (m_i) \times (L, 1) (m), \quad m \in S.
\]

Where \( m \) is the initial position and \( m_i \) is the \( i \)th state of CTMC.

Proof. Because the continuity property of diffusion model, the stopping time \( \tau_i = \inf \{ t \geq 0 : M_t \geq m_i \} = \inf \{ t \geq 0 : M_t = m_i \} \). Then we have \( M_{\tau_i} = m_i \). According to the result in theorem 2,

\[
E_m \left[ e^{-r(\tau_i+U)} g(M_{\tau_i+U}) \right] = E \left[ E_m [e^{-r\tau_i} \tilde{g}_U(M_{\tau_i})] | U \right]
\]

\[
= E \left[ \tilde{g}_U(m_i) | U \right] \times E_m \left( e^{-r\tau_i} \right)
\]

\[
= E \left( e^{(A-r)U} g \right) (m_i) \times (L, 1) (m)
\]

(8)

where \( \tilde{g}_U(x) = E_x \left[ e^{-rU} g(M_U) \right] = \left( e^{(A-r)U} g \right) (x) \), Then

\[
v(m) = \max_{m_0 \leq i \leq N} E \left( e^{(A-r)U} g \right) (m_i) \times (L, 1) (m)
\]

\[\square\]

4.2 The details of the calculation

Now we calculate optimal stopping time with random lag of a diffusion model. Assume that the diffusion model \( \{X_t\}_{t \geq 0} \) subject to the following SDE. Let \( \gamma \) be a constant and \( \sigma(x) \) be a measurable function. \( \{B_t\}_{t \geq 0} \) is the standard Brownian motion.

\[
\frac{dX_t}{X_t} = \gamma dt + \sigma(X_t) dB_t
\]

First we construct a CTMC \( M_t \) approximates \( X_t \) by algorithm in appendix A. Well-known facts in the theory of diffusion processes and continues-time Markov chains imply that the coefficients of the transition rate matrix \( A \) of \( M_t \) must satisfy the following equations.

\[
\sum_{j=1}^{N} A_{i,j} = 0
\]

\[
\sum_{j=1}^{N} A_{i,j} (m_j - m_i) = \gamma m_i
\]

\[
\sum_{j=1}^{N} A_{i,j} (m_j - m_i)^2 = \sigma^2 (m_i) m_i^2
\]
Through a simple matrix operations, the $i^{th}$ row of $\Lambda$ can be expressed as

$$\begin{pmatrix}
\Lambda_{i,1} \\
\vdots \\
\Lambda_{i,i-2} \\
\Lambda_{i,i-1} \\
\Lambda_{i,i} \\
\Lambda_{i,i+1} \\
\vdots \\
\Lambda_{i,N}
\end{pmatrix}^T =
\begin{pmatrix}
0 \\
\cdots \\
0 \\
\frac{\gamma m_i - \gamma^2 (m_i) m_i^2 + \gamma m_i (m_i, m_i-1)}{m_i+1} \\
\frac{\gamma m_i - \gamma^2 (m_i) m_i^2 + \gamma m_i (m_i, m_i-1)}{m_i+1} \\
\frac{\sigma^2 m_i^2 + \gamma m_i (m_i, m_i-1)}{(m_i+1)(m_i-1)} \\
0 \\
\cdots \\
0
\end{pmatrix}^T$$

$i = 2, 3, ..., N$

If $i = 1$,

$$\begin{pmatrix}
\Lambda_{1,1} \\
\Lambda_{1,2} \\
\Lambda_{1,3} \\
\vdots \\
\Lambda_{1,N}
\end{pmatrix}^T =
\begin{pmatrix}
\frac{-\gamma m_1}{m_2-m_1} \\
\frac{-\gamma m_1}{m_2-m_1} \\
0 \\
\cdots \\
0
\end{pmatrix}^T$$

If $i = N$,

$$\begin{pmatrix}
\Lambda_{N,1} \\
\vdots \\
\Lambda_{N,N-2} \\
\Lambda_{N,N-1} \\
\Lambda_{N,N}
\end{pmatrix}^T =
\begin{pmatrix}
0 \\
\cdots \\
0 \\
\frac{-\gamma m_N}{m_N-m_{N-1}} \\
\frac{-\gamma m_N}{m_N-m_{N-1}}
\end{pmatrix}^T$$

After we have the transition rate matrix, we can calculate $A_i, B_i$ and $L_i$ defined in theorem 1. Since the algorithm to discretize the state space is introduced in Appendix A, we can solve the optimal stopping with random lag now by bring the above parameters in formula (8).

5 Numerical Illustrations

In this section we illustrate three main numerical results. First, we test the accuracy of our result with Lempa (2012) where the delay is exponential distributed. Then we show the effect of random lag on investment decision by compare with Mcdonald and Siegel (1987). Finally, we give an experiment to show the impact of misidentifying the distribution of random lag.

5.1 Test of accuracy

In this paper we provided a general framework to solve non-standard optimal stopping problem with a random lag. According to the main method used in this paper is using CTMC to approximates any Markov process, so we need to compare our result with others to see the accuracy of our method.

Lempa (2012) solve this kind problem by using resolvent semigroup method, and their method can deal with the optimal stopping problem with exponential lag. Table 5.1 provides the optimal state under exponential
distributed where $x_{N=500}$ shows our result when states of CTMC are 500 and $x_{Lempa}$ is the result in Lempa (2012). The parameter settings are same as Lempa (2012), which $\lambda$ is the rate parameter of exponential distribution, and the parameter configuration are $\mu = 0.025$, $r = 0.05$, $\sigma = 0.15$, and $K = 2$. Also, we consider the value function has the form $g(x) = (x - K)^+$. 

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>0.01</th>
<th>0.1</th>
<th>1</th>
<th>10</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{N=100}$</td>
<td>2.808</td>
<td>4.615</td>
<td>5.269</td>
<td>5.269</td>
<td>5.613</td>
</tr>
<tr>
<td>$x_{N=500}$</td>
<td>2.835</td>
<td>4.499</td>
<td>5.269</td>
<td>5.416</td>
<td>5.450</td>
</tr>
<tr>
<td>$x_{N=1000}$</td>
<td>2.872</td>
<td>4.478</td>
<td>5.298</td>
<td>5.416</td>
<td>5.420</td>
</tr>
<tr>
<td>$x_{Lempa}$</td>
<td>2.879</td>
<td>4.480</td>
<td>5.296</td>
<td>5.412</td>
<td>5.424</td>
</tr>
</tbody>
</table>

Table 1: Optimal states under GBM model with a exponential distributed lag

From table 5.1, we can see that the difference of results between our method and Lempa (2012) is quite small as the state number increasing.

### 5.2 Effect of random lag on investment decision

We consider a franchise monopolist who has an investment opportunity of building a new plant to produce special products, and he is protected from competition. Also, the profit and cost of this project are $X_t$ and $F$. From Mcdonald and Siegel (1987), the $X_t$ is the following process

$$
\frac{dX_t}{X_t} = \alpha_X dt + \sigma_X dz_X
$$

Also, the value function of this opportunity is like this

$$
V(X_0) = \sup_\tau E_X[e^{-r\tau}(X_\tau - F)]
$$

For simplicity, we consider the cost of building is constant. Then the opportunity of investment can be seen as a real option. When $X_0 = F = 1$, we calculate the value of this investment option by using our result in section 4 and compared the results with Mcdonald and Siegel (1987) in table 5.2.

<table>
<thead>
<tr>
<th>$\sigma_X^2$</th>
<th>M&amp;F</th>
<th>S&amp;C</th>
<th>EU = 0.1</th>
<th>EU = 0.2</th>
<th>EU = 0.5</th>
<th>EU = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.194</td>
<td>0.194</td>
<td>0.191</td>
<td>0.187</td>
<td>0.179</td>
<td>0.166</td>
</tr>
<tr>
<td>0.3</td>
<td>0.365</td>
<td>0.365</td>
<td>0.360</td>
<td>0.356</td>
<td>0.343</td>
<td>0.323</td>
</tr>
<tr>
<td>0.5</td>
<td>0.463</td>
<td>0.462</td>
<td>0.456</td>
<td>0.450</td>
<td>0.435</td>
<td>0.413</td>
</tr>
<tr>
<td>1</td>
<td>0.599</td>
<td>0.595</td>
<td>0.589</td>
<td>0.582</td>
<td>0.566</td>
<td>0.551</td>
</tr>
</tbody>
</table>

Table 2: Value of Investment Opportunity with uniform distributed lag

From 5.2 we have the following conclusions. First, the value are almost same when there are no lag, which means our method are accuracy under no lag. Second, the value of investment opportunity are decreasing as the expectation of random lag increasing. Third, the value are increasing as the volatility increasing.

### 5.3 Impact of misidentifying the distribution of random lag

Sometimes, we may do not even know the actual probability distribution of the random delay when we are faced with decision making. Also, to find out the distribution of the random lag is still not an easy thing. So when we deal with the optimal stopping problem with random lag, if we have a wrong assumption of the random lag, how much loss will we face?
Now we give an experiment to show if we don’t know the distribution of random lag, how different assumption of the lag will have different impact to our profit. Assume we are working on a research and development project, and the real distribution of lag about this project is exponential distribution, and we do not know the real distribution of random lag when we calculate the optimal stopping state. Therefore, we analyzed the result if we assume the lag is constant or exponential distributed but the real distribution of lag is exponential distribution. The reason why we have the above assumption is that as far as we know in the past literature can only solve the problem with exponential lag (see Lempa (2012)) and constant lag (see Øksendal (2011)). The loss of different wrong assumption has been illustrated in figure 1. And the parameter configuration are as follow: CTMC state number $N=200$, discount rate $r=0.15$, $\alpha_V = -0.05$, $\sigma^2_V = 0.3$ and the variance of uniform lag is $\frac{1}{12}$.

![Figure 1: The loss of profit when we have wrong assumption of lag’s distribution](image)

From Figure 1 we can see, an exponential assumption seems inferior to a constant assumption of random lag when the real lag is uniform distributed. The sawtooth in the figure 1 will be small when the states of CTMC we generated increase. We can still find that use exponential assumption is not good when real distribution of delay is uniform. But constant assumption is better, from the figure 1 we can see the loss is so small in constant assumption.

What we can see from this experiment is the distribution of time delay is influential to the decision-making, therefore, a reasonable estimate of implementation delay’s distribution is very important in practice.
6 Conclusion

From above analyses, we can conclude that the optimal stopping problem with general random lag under general Markov processes has been solved in explicit form by using CTMC approximation. Also, We simplify the calculation in the special case of diffusion process basing on the continuity property of diffusion model. Also we have a more efficient solution when the lag is hyper-exponential distributed.

To illustrate the result, we first compare our solution to Lempa (2012) where random lag is exponential distributed, then we show the effect of random lag on investment decision. Finally, we design an experiment to show what level of loss we will face if we have wrong assumption of random lag.

Additional research should focus on the following field: (i) Consider the case that random lag is related to some parameters in stochastic processes; (ii) To solve a variety of applications with more different value function by exploring the form of stopping boundary.

A Algorithm of Construction of the CTMC

This algorithm first raised by Mijatovic and Pistorius (2010), and we use it under assumption of Cai et al. (2015). Assume the initial position of the stochastic process is $x$. To generate the state space of CTMC approximates a Markov process $X_t$, we choose $S_1 = 0.01 \times x$ and $S_N = 40 \times x$. and the other states $\{S_2, \ldots, S_{N-1}\}$ generated by the procedure.

$$S_k = \begin{cases} 
 x + \sinh[(1 - \frac{k-1}{N} - 1) \times \text{arsinh}(S_1 - x)] & \text{if } 2 \leq k \leq \frac{N}{2} \\
 x + \sinh[\frac{k-2N}{N} \times \text{arsinh}(S_N - x)] & \text{if } \frac{N}{2} < k \leq N - 1
\end{cases}$$

Intuitively, these $N$ states are non-uniformly distributed over $[S_0, S_N]$ and are placed more densely around the initial position $x$.

B Proof of the Convergence of Solution

According to the main idea in this paper is using CTMC approximate any Markov processes, and we solve the problem under CTMC framework. The convergence of the solution must be proved. Our proof is similar to Eriksson et al. (2015), but the problem we faced under unlimited time and there is a random lag in the problem. Therefore some element that they abandoned in the proof we must to prove. The details of the proof summarized in the following theorem.

**Theorem 4** Assume that stochastic processes $X$ is a Feller processes with state-space $R_+$ that defined on some filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F} = \{\mathcal{F}_t\}_{t \geq 0}$. Also we assume that $\{e^{-rt}X_t\}_{t \geq 0}$ is a square-integrable martingale. And we can construct a series of CTMC $X^{(n)}$ weakly convergence in $X$ by algorithm in appendix A. Let $g$ be a bounded Lipschitz-continuous function with bound $L \in R_+$ and Lipschitz constant $K$. And let $U$ be the general distributed random lag. If the following two functions are Lipschitz-continuous on $[0, \infty)$ with respective Lipschitz constants given by $C^2(c_1 x + c_2)^2$ and $D(n)^2(d_1 x + d_2)$ for some $C, D(n), c_1, c_2, d_1, d_2 \in R_+$ and $\sup_{n \in N} D(n)$ is finite.

$$t \mapsto \mathbb{E}_x[(e^{-r}X)_t]$$
$$t \mapsto \mathbb{E}_x[(e^{-r}X^{(n)})_t]$$
then we have the main result as follow.

\[ V^{(n)}(x) \to v(x) \quad \text{if} \quad n \to \infty \]  \hspace{1cm} (B.1)

where

\[ V^{(n)}(x) = \sup_{\tau \in \Pi(X)} E_x \left[ e^{-r(T+U)} g(X^{(n)}_{\tau+U}) \right] \]
\[ v(x) = \sup_{\tau \in \Pi(X)} E_x \left[ e^{-r(T+U)} g(X_{\tau+U}) \right] \]

And \( \Pi(X) \) contain all the \( F \)-stopping times taking values in \( R_+ \).

**Proof:** Denote \( T = \{k\Delta : k \in \mathbb{N}^* \} \) with \( \Delta = 1/M, \ M \in \mathbb{N}^* \), and

\[ V^{(n),M}(x) = \sup_{\tau \in \Pi^M(X)} E_x \left[ e^{-r(T+U)} g(X^{(n)}_{\tau+U}) \right] \]
\[ v^M(x) = \sup_{\tau \in \Pi^M(X)} E_x \left[ e^{-r(T+U)} g(X_{\tau+U}) \right] \]

where \( \Pi^M(X) \) denotes the collection of \( F \)-stopping times taking values in the grid \( T \).

First, we prove the following two inequalities.

\[ |v^M(x) - v(x)| \leq \frac{\tilde{C}}{\sqrt{M}}, \quad |V^{(n),M}(x) - V^{(n)}(x)| \leq \frac{\tilde{D}}{\sqrt{M}} \]

It’s easy to find that the stopping time of the form \( \tau_M = \inf\{s \geq \tau : s \in T\} \) for \( \tau \in \Pi(X) \) is equal to the set \( \Pi^M(X) \). Then

\[ |v(x) - v^M(x)| = \sup_{\tau \in \Pi(X)} E_x \left[ e^{-r(T+U)} g(X_{\tau+U}) - e^{-r(T+U)} g(X_{\tau_M+U}) \right] \]
\[ \leq \sup_{\tau \in \Pi(X)} E_x \left[ e^{-r(T+U)} g(X_{\tau+U}) - e^{-r(T_M+U)} g(X_{\tau_M+U}) \right] \]
\[ \leq \sup_{\tau \in \Pi(X)} E_x \left[ |e^{-r(T+U)} - e^{-r(T_M+U)}| g(X_{\tau+U}) \right] \]
\[ \leq \frac{rL}{M} + K \sup_{\tau \in \Pi(X)} E_x \left[ e^{-r(T_M+U)} |X_{\tau+U} - X_{\tau_M+U}| \right] \]

where \( K \) is the Lipschitz constant and \( L \) is the bound of \( g(x) \). Then we have

\[ E_x \left[ e^{-r(T_M+U)} |X_{\tau+U} - X_{\tau_M+U}| \right] \leq E_x \left[ e^{-r(T_M+U)} |X_{\tau_M+U} - X_{\tau_M+U}| \right] \]

Since

\[ E_x[|X_{\tau_M+U} - X_{\tau_M+U}|] \leq E_x \left[ |X_0 - e^{-r(T_M+U)} X_{\tau_M+U}| \right] + E_x \left[ (e^{-r(T_M+U)} - 1) X_{\tau_M+U} \right] \]
\[ := e_1(s) + e_2(s) \]  \hspace{1cm} (B.3)

On the one hand, according to Doob’s optimal stopping theorem we have

\[ e_2(s) = E_x \left[ (e^{-r(T_M+U)} - 1) X_{\tau_M+U} \right] \]
\[ \leq \frac{r}{M} e^{\tilde{r}M} E_x \left[ e^{-r(T_M+U)} X_{\tau_M+U} \right] \]
\[ = \frac{r}{M} e^{\tilde{r}M} s \]  \hspace{1cm} (B.4)
On the other hand, by an application of Doob’s maximal inequalities, we find

\[ c_1(s) = E_x \left[ |X_0 - e^{-r(\tau_M \wedge s)}X_{\tau_M \wedge s}| \right] \]

\[ \leq E_x \left[ \sup_{t \leq \frac{s}{\sqrt{M}}} |e^{-r t}X_t - X_0| \right] \]

\[ \leq 4E_x \left[ \left( e^{-\frac{r L}{2}}X_{\frac{s}{\sqrt{M}}} - X_0 \right)^\frac{1}{2} \right]^2 \]

\[ = 4 \left( E_x \left[ |e^{-r t}X_t| \right] \right)^\frac{1}{2} \]

\[ \leq \frac{4}{\sqrt{M}} C (c_1 s + c_2) \] (B.5)

Then according to (B.3),(B.4) and (B.5) and Doob’s optimal stopping theorem, we have

\[
(B.2) \leq \frac{r L}{M} + K \sup_{\tau \in \Pi(X)} E_x \left[ e^{-r(\tau_M \wedge U)} (c_1(X_{\tau+U}) + c_2(X_{\tau+U})) \right] \\
\leq \frac{r L}{M} + K \sup_{\tau \in \Pi(X)} E_x \left[ e^{-r(\tau_M \wedge U)} \left( \frac{r}{M} e^{\frac{r L}{2}} X_{\tau+U} + \frac{4}{\sqrt{M}} C (c_1 X_{\tau+U} + c_2) \right) \right] \\
\leq \frac{r L}{M} + K \left( \frac{r}{M} e^{\frac{r L}{2}} x + \frac{4}{\sqrt{M}} C (c_1 x + c_2) \right) \leq \bar{C} \sqrt{M}
\] (B.6)

where \( \bar{C} = rL + K [re^r x + 4C (c_1 x + c_2)] \) is constant. (B.6) means \( |v^M(x) - v(x)| \leq \frac{\bar{C}}{\sqrt{M}} \). And similar procedure can derive that \( |V^{(n), M}(x) - V^{(n)}(x)| \leq \frac{\bar{C}}{\sqrt{M}} \).

Next, we will prove the main result (B.1). Denote \( \Pi \) is the filtration generated by \( \{X, X^{(n)}, n \in N^*\} \) and \( \Pi^M(X) \) is the collection of all \( \Pi \)-stopping times with grid \( T \). we have

\[
\left| v^M(x) - V^{(n), M}(x) \right| \leq \sup_{\tau \in \Pi^M(X)} E_x \left[ e^{-r(\tau_M \wedge U)} \left| g(X_{\tau+U}) - g(X^{(n)}_{\tau+U}) \right| \right] \\
\leq K \sup_{\tau \in \Pi^M(X)} E_x \left[ e^{-r(\tau_M \wedge U)} \left| X_{\tau+U} - X^{(n)}_{\tau+U} \right| \right] \\
\leq KE_x \sup_{\tau \in T} e^{-r(\tau_M \wedge U)} \left| X_{\tau+U} - X^{(n)}_{\tau+U} \right| \] (B.7)

Because \( X^{(n)}_{\tau+U} \xrightarrow{w} X_{\tau} \) in the skorokhod topology as \( n \xrightarrow{w} \infty, \forall t \in T, X^{(n)}_{\tau+U} \xrightarrow{d} X_{\tau} \) as \( n \xrightarrow{w} \infty \). According to Skorohod representation theorem, there exist another probability space \( (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}) \) and random variables \( \tilde{X}_{\tau+U}, \tilde{X}^{(n)}_{\tau+U}, n \in N^* \) in this space satisfy the following relations.

\( \tilde{X}^{(n)}_{\tau+U} \equiv X^{(n)}_{\tau+U} \), \( \tilde{X}_{\tau} \equiv X_{\tau} \), and \( \tilde{X}^{(n)}_{\tau+U} \xrightarrow{a.s.} \tilde{X}_{\tau} \)

The uniform integrability of the form collection \( \{X^{(n)}_{\tau+U}, \tilde{X}_{\tau}\} \) implies \( E_x \left\[ \left| X_{\tau+U} - X^{(n)}_{\tau+U} \right| \right\] \xrightarrow{0} 0 \) as \( n \xrightarrow{a.s.} \infty \).

The triangle inequality implies that

\[
\left| V^{(n)}(x) - v(x) \right| \leq \left| V^{(n)}(x) - V^{(n), M}(x) \right| + \left| V^{(n), M}(x) - v^M(x) \right| + \left| v^M(x) - v(x) \right|
\]

Let \( \epsilon \) be arbitrary. Since (B.6) and (B.7), there exist an \( M_\epsilon \) such that \( \forall M \geq M_\epsilon \) and \( \forall n \in N^* \)

\[
\max \left\{ \left| v^M(x) - v(x) \right|, \sup_{n \in N^*} \left| V^{(n)}(x) - V^{(n), M}(x) \right| \right\} \leq \epsilon
\] (B.8)
Fix $M$ larger than $M_\epsilon$, then there exist an $N_\epsilon$ such that

$$\left| V^{(n),M}(x) - v^M(x) \right| \leq \epsilon, \ \forall n \geq N_\epsilon \tag{B.9}$$

In combination with (B.8) and (B.9), we have

$$\left| V^{(n)}(x) - v(x) \right| \leq 3\epsilon$$

Since $\epsilon$ is arbitrary, the statement (B.1) follow. □

References


