

PHASE-TYPE DISTRIBUTIONS AND OPTIMAL STOPPING FOR AUTOREGRESSIVE PROCESSES

SÖREN CHRISTENSEN,* *Christian-Albrechts-Universität zu Kiel*

Abstract

Autoregressive processes are intensively studied in statistics and other fields of applied stochastics. For many applications, the overshoot and the threshold time are of special interest. When the upward innovations are in the class of phase-type distributions, we determine the joint distribution of these two quantities and apply this result to problems of optimal stopping. Using a principle of continuous fit, this leads to explicit solutions.

Keywords: Autoregressive process; threshold time; phase-type innovation; optimal stopping

2010 Mathematics Subject Classification: Primary 60G99

Secondary 60G40; 62L15

1. Introduction

Autoregressive processes play an important role in many areas of applied probability and statistics. They can be seen as one of the building blocks for many models in time series analysis and estimation, and testing techniques are well developed. In this article we study the following setting. Let $0 < \lambda \leq 1$, let $(Z_n)_{n \in \mathbb{N}}$ be a sequence of independent and identically distributed random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and let $(\mathcal{F}_n)_{n \in \mathbb{N}}$ be the filtration generated by $(Z_n)_{n \in \mathbb{N}}$. Define the autoregressive process of order 1 (AR(1) process) $(X_n)_{n \in \mathbb{N}_0}$ by

$$X_n = \lambda X_{n-1} + Z_n \quad \text{for all } n \in \mathbb{N},$$

i.e.

$$X_n = \lambda^n X_0 + \sum_{k=0}^{n-1} \lambda^k Z_{n-k}.$$

The random variables $(Z_n)_{n \in \mathbb{N}}$ are called the innovations of $(X_n)_{n \in \mathbb{N}_0}$. Using the difference notation, the identity $X_n = \lambda X_{n-1} + Z_n$ can be written as

$$\Delta X_n = -(1 - \lambda)X_{n-1} \Delta n + \Delta L_n,$$

where $\Delta X_n = X_n - X_{n-1}$, $\Delta n = n - (n - 1) = 1$, and $\Delta L_n = \sum_{k=1}^n Z_k - \sum_{k=1}^{n-1} Z_k = Z_n$. This shows that AR(1) processes are the discrete-time analog to (Lévy-driven) Ornstein–Uhlenbeck processes. Many arguments in the following can be carried over to Ornstein–Uhlenbeck processes as well.

Autoregressive processes were studied in detail in the last decades. The joint distribution of the threshold time

$$\tau_b = \inf\{n \in \mathbb{N}_0 : X_n \geq b\}$$

Received 21 October 2010; revision received 12 September 2011.

* Postal address: Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel, Ludewig-Meyn-Str. 4, 24098 Kiel, Germany. Email address: christensen@math.uni-kiel.de

and, under the assumption that $\tau_b < \infty$ \mathbb{P}_x -almost surely (\mathbb{P}_x -a.s.) for all x , the overshoot

$$X_{\tau_b} - b \quad \text{over a fixed level } b$$

was of special interest. If $\lambda = 1$, the process $(X_n)_{n \in \mathbb{N}_0}$ is a random walk and many results about this distribution are well known. Most of them are based on techniques using the Wiener–Hopf factorization; see [4, Chapter VII] for an overview. Unfortunately, no analog to the Wiener–Hopf factorization is known for AR(1) processes, so other ideas are necessary. To get rid of well-studied cases, we assume that $\lambda < 1$ in the following.

This first passage problem for AR(1) processes was considered in different applications, such as signal detection and surveillance analysis; cf. [6]. In applications the distribution of the first passage time is approximated using Monte Carlo simulations or Markov chain approximations; cf., e.g. [17]. But, e.g. for questions of optimization, analytic solutions are necessary.

Using martingale techniques, exponential bounds for the expectation of τ_b can be found. Most of these results are based on martingales defined using integrals of the form

$$\int_0^\infty e^{uy - \phi(u)} u^{v-1} du, \quad (1.1)$$

where ϕ is the logarithm of the Laplace transform of the stationary distribution discussed in Section 3. For the integral to be well defined, it is necessary that $E(e^{uZ_1}) < \infty$ for all $u \in [0, \infty)$; cf. [12] and the references therein. If however we want to obtain explicit results for the joint distribution of τ_b and the overshoot, it is useful to assume that Z_1 is exponentially distributed. In this case explicit results are given in [3, Section 3] by setting up and solving differential equations. Unfortunately, in this case not all exponential moments of Z_1 exist and the integral described above cannot be used.

The contribution of this article is twofold.

1. We find the joint distribution of τ_b and the overshoot for a wide class of innovations. We assume that $Z_1 = S_1 - T_1$, where S_1 and T_1 are independent, S_1 has a phase-type distribution and $T_1 \geq 0$ is arbitrary. This generalizes the assumption of exponentially distributed innovations to a much wider class. In Section 2 we establish that τ_b and the overshoot are, conditioned on certain events, independent, and we find the distribution of the overshoot. In Section 3 we use a series inspired by integral (1.1) to construct martingales with the objective of finding the distribution of τ_b . This leads to explicit expressions for expectations of the form $E_x(\rho^{\tau_b} g(X_{\tau_b}))$ for general functions g and $\rho \in (0, 1)$. For corresponding results in a continuous-time setting for $g = \exp$, using complex contour integrals, we refer the reader to [9].
2. As an application, we consider the (Markovian) problem of optimal stopping for $(X_n)_{n \in \mathbb{N}_0}$ with discounted nonnegative continuous gain function g , i.e. we study the optimization problem

$$v(x) = \sup_{\tau \in \mathcal{T}} E_x(\rho^\tau g(X_\tau)) = \sup_{\tau \in \mathcal{T}} E(\rho^\tau g(\lambda^\tau x + X_\tau)), \quad x \in \mathbb{R}, \quad 0 < \rho < 1,$$

where \mathcal{T} denotes the set of stopping times with respect to $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$; to simplify the notation here and in the following, we set the payoff equal to 0 on $\{\tau = \infty\}$. Only very few results are known for this problem. In [3] and [11] the innovations are assumed to be exponentially distributed, and in [5] asymptotic results were given for $g(x) = x$.

Following the approach described in [3], the problem can be reduced to determining an optimal threshold. This is summarized in Section 4. In a second step we use the joint distribution of τ_b and the overshoot to find the optimal threshold. To this end, we use the principle of continuous fit, which is established and illustrated with an example in Section 5.

2. Innovations of phase type

In this section we recall some basic properties of phase-type distributions and identify the connection to AR(1) processes. In Subsection 2.1 we establish the terminology and state some well-known results that are of interest for our purpose. All results can be found in [1], discussed from the perspective of queueing theory.

In Subsection 2.2 we concentrate on the threshold time distribution for autoregressive processes when the positive part of the innovations is of phase type. The key result for the next sections is that, conditioned to certain events, the threshold time is independent of the overshoot, and the overshoot is phase-type distributed as well.

2.1. Definition and some properties

Let $m \in \mathbb{N}$, $E = \{1, \dots, m\}$, $\Delta = m + 1$, and $E_\Delta = E \cup \{\Delta\}$. In this subsection we consider a Markov chain $(J_t)_{t \geq 0}$ in continuous time with state space E_Δ . The states $1, \dots, m$ are assumed to be transient and Δ is absorbing. Denote the generator of $(J_t)_{t \geq 0}$ by $\hat{Q} = (q_{ij})_{i,j \in E_\Delta}$, i.e.

$$\hat{q}_{ij}(h) := P(J_{t+h} = j \mid J_t = i) = q_{ij}h + o(h) \quad \text{for all } i \neq j \in E_\Delta$$

and

$$\hat{q}_{ii}(h) := P(J_{t+h} = i \mid J_t = i) = 1 + q_{ii}h + o(h) \quad \text{for all } i \in E_\Delta, h \rightarrow 0, \text{ and } t \geq 0.$$

If we write $\hat{Q}(h) = (\hat{q}_{ij}(h))_{i,j \in E_\Delta}$ for all $h \geq 0$ then $(\hat{Q}(h))_{h \geq 0}$ is a semigroup and the general theory yields

$$\hat{Q}(h) = e^{\hat{Q}h} \quad \text{for all } h \geq 0.$$

Since Δ is assumed to be absorbing, \hat{Q} has the form

$$\hat{Q} = \begin{pmatrix} Q & -Q\mathbf{1} \\ 0, \dots, 0 & 0 \end{pmatrix}$$

for an $m \times m$ matrix Q , where $\mathbf{1}$ denotes the column vector with entries 1.

We consider the survival time of $(J_t)_{t \geq 0}$, i.e. the random variable

$$\eta = \inf\{t \geq 0: J_t = \Delta\}.$$

Let $\hat{\alpha} = (\alpha, 0)$ be an initial distribution of $(J_t)_{t \geq 0}$. Here and in the following, $\alpha = (\alpha_1, \dots, \alpha_m)$ is assumed to be a row vector.

Definition 2.1. We call $P_{\hat{\alpha}}^\eta$ a distribution of phase type (PH) with parameters (Q, α) and write $P_{\hat{\alpha}}^\eta = \text{PH}(Q, \alpha)$ for short.

Let $m = 1$ and $Q = (-\beta)$ for a parameter $\beta > 0$. In this case it is well known that η is exponentially distributed with parameter β . This special case will be the key example we

often refer to. Furthermore, note that the class of phase-type distributions is stable under convolutions and mixtures. This shows that the important classes of Erlang and hyperexponential distributions are of phase type.

Exponential distributions have a very special structure, but phase-type distributions are flexible.

Proposition 2.1. *The distributions of phase type are dense in the space of all probability measures on $(0, \infty)$ with respect to convergence in distribution.*

Proof. See [1, Section III, Theorem 4.2].

The definition of phase-type distributions makes calculations with these distributions difficult, but the theory of semigroups leads to simple formulae for the density and the Laplace transform, as the next lemma shows. All the formulae contain matrix exponentials. The explicit calculation of such exponentials can be complex in higher dimensions, but many algorithms are available for a numerical approximation.

Proposition 2.2. (a) *The eigenvalues of Q have negative real part.*

(b) *The distribution function of $\text{PH}(Q, \alpha)$ is given by*

$$H_\alpha(s) := \mathbb{P}_\alpha(\eta \leq s) = 1 - \alpha e^{Qs} \mathbf{1}, \quad s \geq 0.$$

(c) *The density is given by*

$$h_\alpha(s) = \alpha e^{Qs} q, \quad s \geq 0,$$

where $q = -Q\mathbf{1}$.

(d) *For all $s \in \mathbb{C}$ with $\mathbb{E}_{\hat{\alpha}}(e^{\text{Re}(s)\eta}) < \infty$, it holds that*

$$\hat{H}_\alpha(s) := \mathbb{E}_{\hat{\alpha}}(e^{s\eta}) = \alpha(-sI - Q)^{-1}q,$$

where I is the $m \times m$ identity matrix. In particular, \hat{H}_α is a rational function.

Proof. See [1, Corollary 4.9, Section II, and Theorem 4.1, Section III].

An essential property for the applicability of the exponential distribution in modeling and examples is the memoryless property, which even characterizes the exponential distribution. A generalization of this property to distributions of phase type is well known and is the basis of the following lemma.

Lemma 2.1. *Let $S, T \geq 0$ be stochastically independent random variables, where S is $\text{PH}(Q, \alpha)$ -distributed. Furthermore, let $r \geq 0$ and $Z = S - T$. Then*

$$\mathbb{P}_{\hat{\alpha}}(r \leq Z \leq r + s) = \sum_{i \in E} \lambda_i(r) H_{e_i}(s), \quad s \geq 0,$$

where $\lambda_i(r) = \int \mathbb{P}_{\hat{\alpha}}(J_{r+t} = i) \mathbb{P}_{\hat{\alpha}}^T(dt)$.

Proof. The well-known generalization of the memoryless property to phase-type distributions in our case reads as

$$\mathbb{P}_{\hat{\alpha}}(\eta \leq s + t \mid \eta \geq t) = \sum_{i \in E} \mathbb{P}_{\hat{\alpha}}(J_t = i \mid \eta \geq t) \mathbb{P}_{e_i}(\eta \leq s).$$

Using Fubini’s theorem, this leads to

$$\begin{aligned} P_{\hat{\alpha}}(r \leq Z \leq r + s) &= \int P_{\hat{\alpha}}(r + t \leq S \leq r + s + t) P_{\hat{\alpha}}^T(dt) \\ &= \sum_{i \in E} \int H_{e_i}(s) P_{\hat{\alpha}}(J_{t+r} = i) P_{\hat{\alpha}}^T(dt). \end{aligned}$$

2.2. Phase-type distributions and overshoot of AR(1) processes

We again consider the situation of Section 1. In addition, we assume that the innovations have the structure

$$Z_n = S_n - T_n \quad \text{for all } n \in \mathbb{N},$$

where S_n and T_n are nonnegative and independent, and S_n is PH(Q, α)-distributed. In this context we remark that each probability measure Q on \mathbb{R} with $Q(\{0\}) = 0$ can be written as $Q = Q_+ * Q_-$, where Q_+ and Q_- are probability measures with $Q_+((-\infty, 0)) = Q_-((0, \infty)) = 0$, and ‘*’ denotes the convolution operator (cf. [4, p. 383]).

As a motivation, we consider the case of exponentially distributed innovations. If Z_n is exponentially distributed then it holds that, for all $\rho \in (0, 1]$ and measurable $g : \mathbb{R} \rightarrow [0, \infty)$,

$$E_x(\rho^\tau g(X_\tau)) = E_x(\rho^\tau) E(g(R + b)), \tag{2.1}$$

where $\tau = \tau_b$ is a threshold time, $x < b$, and R is exponentially distributed with the same parameter as the innovations (cf. [3, Theorem 3.1]). This fact is well known for random walks; cf. [4, Chapter XII]. The representation of the joint distribution of the overshoot and τ reduces to finding an explicit expression for the Laplace transform of τ . In this subsection we prove that a generalization of this phenomenon holds in our more general situation.

To this end, we use an embedding of $(X_n)_{n \in \mathbb{N}_0}$ into a stochastic process in continuous time as follows. For all $n \in \mathbb{N}$, denote the Markov chain which generates the phase-type distribution of S_n by $(J_t^{(n)})_{t \geq 0}$, and write

$$J_t = J_{t - \sum_{k=1}^{n_t} S_k}^{(n_t+1)}, \quad \text{where } n_t = \max \left\{ n \in \mathbb{N}_0 : \sum_{k=1}^n S_k \leq t \right\} \quad \text{for all } t \geq 0.$$

Hence, the process $(J_t)_{t \geq 0}$ is constructed by compounding the processes $J^{(n)}$ restricted to their lifetime. Obviously, $(J_t)_{t \geq 0}$ is a continuous-time Markov chain with state space E ; cf. [1, Section III, Proposition 5.1]. Furthermore, we define a process $(Y_t)_{t \geq 0}$ by

$$Y_t = \lambda X_{n_t} - T_{n_t+1} + t - \sum_{k=1}^{n_t} S_k.$$

See Figure 1 for an illustration. It holds that

$$X_n = Y_{(S_1 + \dots + S_n)^-} \quad \text{for all } n \in \mathbb{N},$$

so we can find $(X_n)_{n \in \mathbb{N}_0}$ in $(Y_t)_{t \geq 0}$. Now let $\hat{\tau}$ be the threshold time of the process $(Y_t)_{t \geq 0}$ over the threshold b , i.e.

$$\hat{\tau} = \inf\{t \geq 0 : Y_t \geq b\}.$$

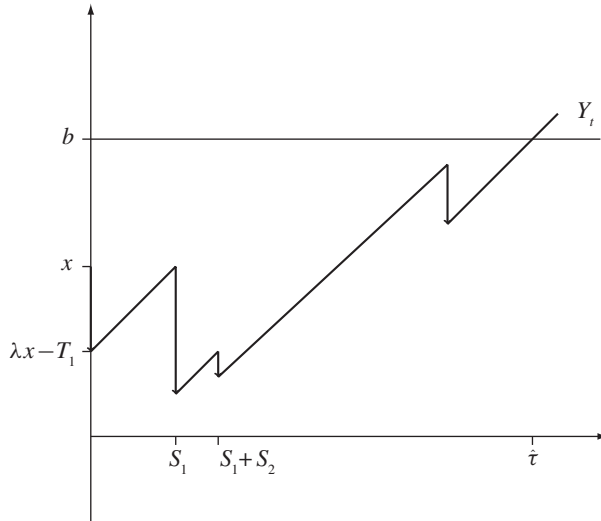


FIGURE 1: A path of $(Y_t)_{t \geq 0}$.

By the definition of $(Y_t)_{t \geq 0}$, it holds that

$$Y_t = b \iff t = -\lambda X_{n_t} + T_{n_t+1} + \sum_{k=1}^{n_t} S_k + b \quad \text{for all } t \geq 0. \tag{2.2}$$

For the following result, we need the event that the associated Markov chain is in state i when $(Y_t)_{t \geq 0}$ crosses $b \geq 0$, i.e. the event $G_i = \{J_{\hat{\tau}} = i\}$ for $i \in E$. For the following considerations, we fix the threshold $b \geq 0$.

In generalization of the result for exponentially distributed innovations the following theorem states that, conditioned on G_i , the threshold time and the overshoot are independent and the overshoot is phase-type distributed as well.

Theorem 2.1. *Let $x < b$, $n \in \mathbb{N}$, and $y \geq 0$, and write*

$$\tau = \tau_b = \inf\{n \in \mathbb{N}_0 : X_n \geq b\}.$$

Then

$$P_x(X_\tau - b \leq y, \tau = n) = \sum_{i \in E} H_{e_i}(y) P_x(\tau = n, G_i).$$

Proof. Using Lemma 2.1 and identity (2.2), we obtain

$$\begin{aligned} P_x(X_\tau - b \leq y, \tau = n) &= E_x(\mathbf{1}_{\{\tau \geq n\}} P_x(X_n \geq b, X_n - b \leq y \mid \mathcal{F}_{n-1})) \\ &= E_x(\mathbf{1}_{\{\tau \geq n\}} P_x(b \leq X_n \leq b + y \mid \mathcal{F}_{n-1})) \\ &= E_x(\mathbf{1}_{\{\tau \geq n\}} P_x(b - \lambda X_{n-1} \leq Z_n \leq b + y - \lambda X_{n-1} \mid \mathcal{F}_{n-1})) \\ &= E_x\left(\mathbf{1}_{\{\tau \geq n\}} \sum_{i \in E} H_{e_i}(y) P_x(J_{b - \lambda X_{n-1} + T_n}^{(n)} = i \mid \mathcal{F}_{n-1})\right) \\ &= \sum_{i \in E} H_{e_i}(y) P_x(\tau = n, J_{\hat{\tau}} = i), \end{aligned}$$

where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function.

This immediately implies a generalization of (2.1) to the case of general phase-type distributions.

Corollary 2.1. *It holds that*

$$E_x(\rho^\tau g(X_\tau)) = \sum_{i \in E} E_x(\rho^\tau \mathbf{1}_{G_i}) E(g(b + R^i)),$$

where R^i is a $\text{PH}(Q, e_i)$ -distributed random variable (under \mathbb{P}).

3. Explicit representations of the joint distribution of the threshold time and overshoot

Corollary 2.1 reduces the problem of finding expectations of the form $E_x(\rho^\tau g(X_\tau))$ to finding $\Phi_i^b(x) = \Phi_i(x) = E_x(\rho^\tau \mathbf{1}_{G_i})$ for $\tau = \tau_b$ and $b > x$. The aim of this section is to construct martingales of the form $(\rho^{n \wedge \tau} h(X_{n \wedge \tau}))_{n \in \mathbb{N}}$ as a tool for the explicit representation of $\Phi_i^b(x)$. To this end, some definitions are necessary.

We assume the setting of the previous section, i.e. we assume that the innovations can be written in the form

$$Z_n = S_n - T_n \quad \text{for all } n \in \mathbb{N},$$

where S_n and T_n are nonnegative and independent, and S_n is $\text{PH}(Q, \alpha)$ -distributed.

Let $\exp(\psi)$ be the Laplace transform of Z_1 , i.e. $\psi(u) = \log E(e^{uZ_1})$ for all $u \in \mathbb{C}_+ := \{z \in \mathbb{C} : \text{Re}(z) \geq 0\}$ with real part $\text{Re}(u)$ so small that the expectation exists. Since $E(e^{uZ_1}) = E(e^{uS_1}) E(e^{-uT_1})$ and $T_1 \geq 0$, Proposition 2.2 yields the existence of $\psi(u)$ for all u with $\text{Re}(u)$ smaller than the smallest eigenvalue of $-Q$. On this strip ψ is analytic and, because of independence, it holds that

$$\psi(u) = \psi_1(u) + \psi_2(u),$$

where $\exp(\psi_1)$ denotes the Laplace transform of S_1 and $\exp(\psi_2)$ is the Laplace transform of $-T_1$. On \mathbb{C}_+ , ψ_2 is analytic, and ψ_1 can be analytically extended to $\mathbb{C}_+ \setminus \text{Sp}(-Q)$ by Proposition 2.2. Here $\text{Sp}(\cdot)$ denotes the spectrum, i.e. the set of all eigenvalues. Hence, ψ can be extended to $\mathbb{C}_+ \setminus \text{Sp}(-Q)$ as well and this extension is again denoted by ψ . Note that this extension cannot be interpreted from a probabilistic point of view because $E(e^{uZ_1})$ does not exist for $u \in \mathbb{C}_+$ when $\text{Re}(z)$ is large.

To guarantee the convergence of $(X_n)_{n \in \mathbb{N}_0}$, we assume a weak integrability condition—the well-known Vervaat condition:

$$E(\log(1 + |Z_1|)) < \infty;$$

see [8, Theorem 2.1] for a characterization of such conditions in the theory of perpetuities. We do not go into details here, but we do use the fact that $(X_n)_{n \in \mathbb{N}_0}$ converges to a (finite) random variable θ in distribution that fulfills the stochastic fixed point equation

$$\mathbb{P}^\theta = \mathbb{P}^{\lambda\theta} * \mathbb{P}^{Z_1}.$$

Since the AR(1) process has the representation

$$X_n = \lambda^n X_0 + \sum_{k=0}^{n-1} \lambda^k Z_{n-k}$$

and convergence in distribution is equivalent to the pointwise convergence of the Laplace transforms, the Laplace transform $\exp(\phi)$ of θ fulfills

$$\phi(u) = \sum_{k=0}^{\infty} \psi(\lambda^k u) \tag{3.1}$$

for all $u \in \mathbb{C}_+$ such that the Laplace transform of S_1 exists. The right-hand side defines a holomorphic function on $\mathbb{C}_+ \setminus \hat{P}$ that is also denoted by ϕ , where we write $\hat{P} = \bigcup_{n \in \mathbb{N}_0} \text{Sp}(-\lambda^{-n} Q)$. For the convergence of the series, note that, as described above, it converges for all $u \in \mathbb{C}_+$ such that $E(e^{uZ_1}) < \infty$. For all other $u \in \mathbb{C}_+$, the series also converges since there exists a k_0 such that $E(e^{\lambda^{k_0} u Z_1}) < \infty$ for all $k \geq k_0$.

Furthermore, the identity

$$\phi(u) = \phi(\lambda u) + \psi(u) \tag{3.2}$$

holds whenever u and λu are in the domain of ϕ . To avoid problems concerning the applicability of (3.2), we assume that

$$\text{Sp}(\lambda^n Q) \cap \text{Sp}(Q) = \emptyset \quad \text{for all } n \in \mathbb{N}. \tag{3.3}$$

Note that the function ϕ was also used and studied in [12].

The next two lemmas are helpful in the construction of the martingales.

Lemma 3.1. *Let $\delta \in \mathbb{C}_+$ such that $E(e^{\delta S_1})$ exists. Then, for all $x < b$, it holds that*

$$\rho E(e^{\delta(\lambda x + Z_1)} \mathbf{1}_{\{\lambda x + Z_1 \geq b\}}) = \alpha_\delta e^{-\lambda x} Q q,$$

where $\alpha_\delta = \rho \alpha (-\delta I - Q)^{-1} e^{(\delta I + Q)b + \psi_2(-Q)}$.

Proof. In the following calculation we use the facts that all matrices are commutative and all eigenvalues of $\delta I + Q$ have negative real part. It holds that

$$\begin{aligned} E(e^{\delta(\lambda x + Z_1)} \mathbf{1}_{\{\lambda x + S_1 \geq b\}}) &= e^{\delta \lambda x} \int_0^\infty E(e^{\delta(S-t)} \mathbf{1}_{\{\lambda x + S_1 - t \geq b\}}) P^T(dt) \\ &= e^{\delta \lambda x} \int_0^\infty \int_{b+t-\lambda x}^\infty e^{\delta(s-t)} \alpha e^{Qs} q ds P^T(dt) \\ &= e^{\delta \lambda x} \int_0^\infty e^{-\delta t} \alpha \int_{b+t-\lambda x}^\infty e^{(\delta I + Q)s} ds q P^T(dt) \\ &= -e^{\delta \lambda x} \int_0^\infty e^{-\delta t} \alpha (\delta I + Q)^{-1} e^{(\delta I + Q)(b+t-\lambda x)} q P^T(dt). \end{aligned}$$

We obtain

$$\begin{aligned} \rho E(e^{\delta(\lambda x + Z_1)} \mathbf{1}_{\{\lambda x + Z_1 \geq b\}}) &= -\rho e^{\delta \lambda x} \alpha \int_0^\infty e^{Q t} P^T(dt) (\delta I + Q)^{-1} e^{(\delta I + Q)(b-\lambda x)} q \\ &= -\rho \alpha \int_0^\infty e^{-Qs} P^{-T}(ds) (\delta I + Q)^{-1} e^{-Q\lambda x} e^{(\delta I + Q)b} q \\ &= \alpha_\delta e^{-Q\lambda x} q. \end{aligned}$$

We write $Q_\gamma = -\gamma Q$ for short. For all $\gamma \in \mathbb{C}_+$ fulfilling

$$\text{Sp}(\lambda^n Q_\gamma) \cap \hat{P} = \emptyset \quad \text{for all } n \in \mathbb{N}, \tag{3.4}$$

we define the function

$$f_\gamma : \mathbb{R} \rightarrow \mathbb{C}^{m \times m}, \quad x \mapsto \sum_{n \in \mathbb{N}} e^{x\lambda^n Q_\gamma - \phi(\lambda^n Q_\gamma)} \rho^{n-1}.$$

This series converges because $e^{x\lambda^n Q_\gamma - \phi(\lambda^n Q_\gamma)}$ is bounded in n . Note that the summand of this series is similar to the integrand in (1.1).

Lemma 3.2. *There exists a $\delta > 0$ such that, for all $x \in \mathbb{R}$ and $\gamma \in \mathbb{C}_+$ with $|\gamma| < \delta$, it holds that*

$$\rho E(f_\gamma(\lambda x + Z_1)) = f_\gamma(x) - e^{\lambda x Q_\gamma - \phi(\lambda Q_\gamma)}.$$

Proof. For all $\gamma \in \mathbb{C}_+$ with sufficiently small $|\gamma|$, the expected value $E(e^{Q_\gamma \lambda^n Z_1})$ exists for all $n \in \mathbb{N}$ since Q has (finitely many) negative eigenvalues. This leads to

$$\begin{aligned} E(f_\gamma(\lambda x + Z_1)) &= E\left(\sum_{n \in \mathbb{N}} e^{(\lambda x + Z_1)\lambda^n Q_\gamma - \phi(\lambda^n Q_\gamma)} \rho^{n-1}\right) \\ &= \sum_{n \in \mathbb{N}} e^{(\lambda x)\lambda^n Q_\gamma - \phi(\lambda^n Q_\gamma)} E(e^{\lambda^n Q_\gamma Z_1}) \rho^{n-1} \\ &= \sum_{n \in \mathbb{N}} e^{x\lambda^{n+1} Q_\gamma - \phi(\lambda^n Q_\gamma) + \psi(\lambda^n Q_\gamma)} \rho^{n-1} \\ &\stackrel{(3.2)}{=} \frac{1}{\rho} \sum_{n \in \mathbb{N}} e^{x\lambda^{n+1} Q_\gamma - \phi(\lambda^{n+1} Q_\gamma)} \rho^n \\ &= \frac{1}{\rho} (f_\gamma(x) - e^{\lambda Q_\gamma x - \phi(\lambda Q_\gamma)}). \end{aligned}$$

The next step is to find a family of equations characterizing

$$\Phi(x) = (\Phi_1(x), \dots, \Phi_m(x)) = (E_x(\rho^\tau \mathbf{1}_{G_1}), \dots, E_x(\rho^\tau \mathbf{1}_{G_m}))$$

using martingale techniques, where $\tau = \tau_b$ and $x < b$. To this end, we consider

$$h_{\gamma,\delta} : \mathbb{R} \rightarrow \mathbb{C}, \quad x \mapsto e^{\delta x} \mathbf{1}_{\{x \geq b\}} + \beta_{\gamma,\delta} f_\gamma(x)q,$$

for all $\delta \in \mathbb{C}_+ \setminus \text{Sp}(-Q)$ and γ fulfilling (3.4), where $\beta_{\gamma,\delta} = \alpha_\delta e^{\phi(\lambda Q_\gamma)}$. For the special value $\gamma = 1$, we write $h_\delta = h_{\gamma,\delta}$ and this function is well defined by (3.3).

Putting together the results of Lemma 3.1 and Lemma 3.2 we obtain the equation

$$\rho E_x(h_{\gamma,\delta}(X_1)) - h_{\gamma,\delta}(x) = \alpha_\delta e^{-\lambda x Q} q - \alpha_\delta e^{-\lambda \gamma x Q} q \quad \text{for all } x < b \tag{3.5}$$

for all $\gamma, \delta \in \mathbb{C}_+$ with sufficiently small modulus.

Before stating the equations we need one more technical result.

Lemma 3.3. *Let R^i be a $\text{PH}(Q, e_i)$ -distributed random variable, and denote by $\psi^i(\cdot) = \log(e_i(-\cdot I - Q)^{-1}q)$ the holomorphic extension of the logarithmized Laplace transform*

of R_i . Here e_i denotes the i th unit vector. Let $|\gamma|$ and $|\delta|$ be so small that $E(h_{\gamma,\delta}(b + R^i))$ exists. Then it holds that

$$E(h_{\gamma,\delta}(b + R^i)) = e^{\delta b} \alpha_{\gamma,i} (-\delta I - Q)^{-1} q =: \eta_{\gamma,\delta,i},$$

where

$$\alpha_{\gamma,i} = e_i + \alpha e^{Qb + \psi_2(-Q) + \phi(\lambda Q_\gamma)} \sum_{n \in \mathbb{N}} e^{b\lambda^n Q_\gamma - \phi(\lambda^n Q_\gamma) + \psi^i(\lambda^n Q_\gamma)} \rho^n.$$

Proof. Simple calculus similar to the above yields the result.

Theorem 3.1. For all $x < b$ and $\delta \in \mathbb{C}_+ \setminus \text{Sp}(-Q)$, it holds that

$$\sum_{i=1}^m \eta_{\delta,i} \Phi_i(x) = h_\delta(x),$$

where $\eta_{\delta,i} = \eta_{1,\delta,i}$ is given in Lemma 3.3.

Proof. Write $h := h_{\gamma,\delta}$ for $\delta, \gamma \in \mathbb{C}_+$ with $|\delta|$ and $|\gamma|$ so small that $E(h(Z_1)) < \infty$. The discrete version of Itô's formula yields

$$\begin{aligned} \rho^n h(X_n) - \sum_{i=0}^{n-1} \rho^i (\rho E_x(h(X_{i+1}) \mid X_i) - h(X_i)) \\ = h(X_0) + \sum_{i=0}^{n-1} \rho^{i+1} (h(X_{i+1}) - E_x(h(X_{i+1}) \mid X_i)) \\ =: M_n, \end{aligned}$$

and $(M_n)_{n \in \mathbb{N}}$ is a martingale. The optional sampling theorem applied to $\tau = \tau_b$ yields

$$\begin{aligned} h(x) &= E_x(M_{\tau \wedge n}) \\ &= E_x(\rho^{n \wedge \tau} h(X_{n \wedge \tau})) - E_x\left(\sum_{i=0}^{n \wedge \tau - 1} \rho^i (\rho E_x(h(X_{i+1}) \mid X_i) - h(X_i))\right) \\ &= E_x(\rho^{n \wedge \tau} h(X_{n \wedge \tau})) - E_x\left(\sum_{i=0}^{n \wedge \tau - 1} \rho^i (\alpha_\delta e^{-\lambda X_i} Q q - \alpha_\delta e^{-\lambda \gamma X_i} Q q)\right) \end{aligned}$$

using equality (3.5). The dominated convergence theorem shows that

$$h_{\delta,\gamma}(x) = E_x(\rho^\tau h_{\delta,\gamma}(X_\tau)) - E_x\left(\sum_{i=0}^{\tau-1} \rho^i (\alpha_\delta e^{-\lambda X_i} Q q - \alpha_\delta e^{-\lambda \gamma X_i} Q q)\right);$$

note that the dominated convergence theorem is applicable to both summands since Q has negative eigenvalues and so e^{-sQ} is bounded in s for s with $\text{Re}(s)$ being bounded above. Corollary 2.1 leads to

$$h_{\gamma,\delta}(x) = \sum_{i=1}^m E_x(\rho^\tau \mathbf{1}_{G_i}) E(h_{\gamma,\delta}(R_i + b)) - E_x\left(\sum_{i=0}^{\tau-1} \rho^i (\alpha_\delta e^{-\lambda X_i} Q q - \alpha_\delta e^{-\lambda \gamma X_i} Q q)\right),$$

where R_i is $\text{PH}(Q, e_i)$ -distributed and Lemma 3.3 implies that

$$E(h_{\gamma,\delta}(R_i + b)) = e^{\delta b} \alpha_{\gamma,i} (-\delta I - Q)^{-1} q.$$

Since $\sum_{i=0}^{\tau-1} \rho^i (\alpha_\delta e^{-\lambda X_i} Q q - \alpha_\delta e^{-\lambda \gamma X_i} Q q)$ is bounded, both sides of the equation

$$h_{\delta,\gamma}(x) = \sum_{i=1}^m \eta_{\gamma,\delta,i} \Phi_i(x) - E_x \left(\sum_{i=0}^{\tau-1} \rho^i (\alpha_\delta e^{-\lambda X_i} Q q - \alpha_\delta e^{-\lambda \gamma X_i} Q q) \right)$$

are holomorphic in $\{\gamma \in \mathbb{C}_+ : \hat{P} \cap \text{Sp}(-\gamma \lambda^n Q) = \emptyset \text{ for all } n \in \mathbb{N}\}$ and it follows from the identity theorem for holomorphic functions that these extensions agree on their domains. Keeping (3.3) in mind we obtain, for $\gamma = 1$,

$$h_\delta(x) = \sum_{i=1}^m \eta_{\delta,i} \Phi_i(x).$$

Furthermore, both sides of the equations are again holomorphic functions in δ on $\mathbb{C}_+ \setminus \text{Sp}(-Q)$. Another application of the identity theorem proves the assertion.

The equation in Theorem 3.1 appears useful and flexible enough for the explicit solution, as shown in the next subsections.

3.1. The case of exponential positive innovations

As described above, the case of $\text{Exp}(\mu)$ -distributed positive innovations is of special interest. In this case we obtain the solution directly from the results above. Hence, let $m = 1$, $\alpha = 1$, $Q = -\mu$, and $q = \mu$. It is not relevant which δ we take, but we choose $\delta = 0$ as this choice leads to simpler expressions. Then we obtain

$$\begin{aligned} \eta_{0,1} &= 1 + e^{\psi_2(\mu) - \mu b + \phi(\lambda\mu)} \sum_{n \in \mathbb{N}} e^{\lambda^n \mu b - \phi(\lambda^n \mu) + \psi_1(\lambda^n \mu)} \rho^n \\ &\stackrel{(3.2)}{=} 1 + e^{\psi_2(\mu) - \mu b + \phi(\lambda\mu)} \sum_{n \in \mathbb{N}} e^{\lambda^n \mu b - \phi(\lambda^{n+1} \mu) - \psi_2(\lambda^n \mu)} \rho^n \\ &= e^{\psi_2(\mu) - \mu b + \phi(\lambda\mu)} \sum_{n \in \mathbb{N}_0} e^{\lambda^n \mu b - \phi(\lambda^{n+1} \mu) - \psi_2(\lambda^n \mu)} \rho^n \end{aligned}$$

and

$$h_0(x) = e^{\psi_2(\mu) - \mu b + \phi(\lambda\mu)} \sum_{n \in \mathbb{N}} e^{\lambda^n \mu x - \phi(\lambda^n \mu)} \rho^n$$

for $x < b$. Theorem 3.1 yields the following result.

Theorem 3.2. *It holds that*

$$E_x(\rho^\tau) = \frac{h_0(x)}{\eta_{0,1}} = \frac{\sum_{n \in \mathbb{N}} e^{\lambda^n \mu x - \phi(\lambda^n \mu)} \rho^n}{\sum_{n \in \mathbb{N}_0} e^{\lambda^n \mu b - \phi(\lambda^{n+1} \mu) - \psi_2(\lambda^n \mu)} \rho^n} \quad \text{for all } x < b.$$

In [3] the special case of positive, exponentially distributed innovations was treated by finding and solving ordinary differential equations for $E_x(\rho^\tau)$. For this case, i.e. $T_1 = 0$, we obtain

$$E_x(\rho^\tau) = \frac{\sum_{n \in \mathbb{N}} e^{\lambda^n \mu x - \phi(\lambda^n \mu)} \rho^n}{\sum_{n \in \mathbb{N}_0} e^{\lambda^n \mu b - \phi(\lambda^{n+1} \mu)} \rho^n}.$$

To obtain more explicit results, we need a simple expression for $\phi(\lambda^n \mu)$. Using identity (3.1), we find such an expression as

$$\begin{aligned} e^{\phi(\lambda^n \mu)} &= \prod_{k=0}^{\infty} e^{\psi_1(\lambda^{n+k} \mu)} \\ &= \prod_{k=0}^{\infty} \frac{\mu}{\mu - \lambda^{n+k} \mu} \\ &= \prod_{k=0}^{\infty} \frac{1}{1 - \lambda^{n+k}} \\ &= \frac{\prod_{k=1}^{n-1} (1 - \lambda^k)}{\prod_{k=1}^{\infty} (1 - \lambda^k)} \\ &= \frac{(\lambda, \lambda)_{n-1}}{\phi_E(\lambda)}, \end{aligned}$$

where $(a, q)_n = \prod_{k=1}^{n-1} (1 - aq^{k-1})$ denotes the q -Pochhammer symbol and $\phi_E(q) = (q, q)_{\infty}$ denotes the Euler function. This leads to

$$E_x(\rho^\tau) = \frac{\sum_{n \in \mathbb{N}} \rho^n e^{\lambda^n \mu x} / (\lambda, \lambda)_{n-1}}{\sum_{n \in \mathbb{N}_0} \rho^n e^{\lambda^n \mu b} / (\lambda, \lambda)_n}$$

and the numerator is given by

$$\begin{aligned} \sum_{n \in \mathbb{N}} \frac{\rho^n}{(\lambda, \lambda)_{n-1}} e^{\lambda^n \mu x} &= \sum_{k \in \mathbb{N}_0} \frac{(\mu x)^k}{k!} \sum_{n \in \mathbb{N}} \frac{(\rho \lambda^k)^n}{(\lambda, \lambda)_{n-1}} \\ &= \rho \sum_{k \in \mathbb{N}_0} \frac{(\mu x \lambda)^k}{k!} \sum_{n \in \mathbb{N}} \frac{(\rho \lambda^k)^{n-1}}{(\lambda, \lambda)_{n-1}} \\ &= \rho \sum_{k \in \mathbb{N}_0} \frac{(\mu x \lambda)^k}{k!} \frac{1}{(\rho \lambda^k, \lambda)_{\infty}} \\ &= \frac{\rho}{(\rho, \lambda)_{\infty}} \sum_{k \in \mathbb{N}_0} \frac{(\rho, \lambda)_k (\mu \lambda x)^k}{k!}. \end{aligned}$$

Note that we used the q -binomial theorem in the third step (see [7, Equation (1.3.15)] for a proof). An analogous calculation for the denominator yields

$$\sum_{n \in \mathbb{N}_0} \frac{\rho^n}{(\lambda, \lambda)_n} e^{\lambda^n \mu b} = \frac{1}{(\rho, \lambda)_{\infty}} \sum_{k \in \mathbb{N}_0} \frac{(\rho, \lambda)_k (\mu b)^k}{k!},$$

leading to the following result.

Theorem 3.3. *If S_1 is $\text{Exp}(\mu)$ -distributed and $T_1 = 0$, it holds that*

$$E_x(\rho^{\tau_b}) = \rho \frac{\sum_{k \in \mathbb{N}_0} (\rho, \lambda)_k (\mu x \lambda)^k / k!}{\sum_{k \in \mathbb{N}_0} (\rho, \lambda)_k (\mu b)^k / k!} \text{ for all } x < b.$$

For the special case $\mu = 1$, this formula was obtained using differential equations based on the generator in [11, Theorem 3].

Noting that

$$E_b(\rho^{\tau_{b+}}) = \rho \frac{\sum_{k \in \mathbb{N}_0} (\rho, \lambda)_k (\mu b \lambda)^k / k!}{\sum_{k \in \mathbb{N}_0} (\rho, \lambda)_k (\mu b)^k / k!},$$

by direct calculation we find that

$$\frac{d}{db} E_x(\rho^{\tau_b}) = E_x(\rho^{\tau_b}) \mu (E_b(\rho^{\tau_{b+}}) - 1).$$

This reproduces Theorem 3.3 of [3]. Note that in that article the stopping time $\tilde{\tau}_b = \inf\{n \in \mathbb{N}_0 : X_n > b\}$ was considered. But this leads to analogous results since

$$\tau_{b+} = \tilde{\tau}_b \quad \text{under } P_b$$

and

$$\tau_b = \tilde{\tau}_b \quad P_x\text{-a.s. for all } x \neq b.$$

3.2. The general case

In this subsection we show that Theorem 3.1 is a powerful tool for the explicit calculation of Φ in many cases of interest. By Lemma 3.3 we see that $\eta_{\delta,i}$ is a rational function of δ with poles in $\text{Sp}(-Q)$ for all $i = 1, \dots, m$. We assume for simplicity that all eigenvalues are pairwise different (for the general case, see Remark 3.1 below). Then partial fraction decomposition yields the representation

$$\eta_{\delta,i} = \sum_{j=1}^m \frac{a_{i,j}}{\mu_j - \delta} \quad \text{for some } a_{i,1}, \dots, a_{i,m},$$

and since $h_\delta(x)$ is rational in δ with the same poles, we may write

$$h_\delta(x) = \sum_{j=1}^m \frac{c_j(x)}{\mu_j - \delta} \quad \text{for some } c_1(x), \dots, c_m(x).$$

Theorem 3.1 reads

$$\sum_{j=1}^m \frac{\sum_{i=1}^m a_{ij} \Phi_i(x)}{\mu_j - \delta} = \sum_{j=1}^m \frac{c_j(x)}{\mu_j - \delta}$$

and the uniqueness of the partial fraction decomposition yields

$$\sum_{i=1}^m a_{ij} \Phi_i(x) = c_j(x),$$

i.e.

$$A \Phi(x) = c(x),$$

where $A = (a_{ij})_{i,j=1}^m$ and $c(x) = (c_j(x))_{j=1}^m$. This leads to the following result.

Theorem 3.4. *If A is invertible then $\Phi(x)$ is given by*

$$\Phi(x) = A^{-1}c(x) \quad \text{for all } x < b.$$

Remark 3.1. Note that the assumption of distinct eigenvalues was made for simplicity only. When it is not fulfilled, we can use the general partial fraction decomposition formula and obtain the analogous result.

4. Applications to optimal stopping

To tackle the optimal stopping problem

$$v(x) = \sup_{\tau \in \mathcal{T}} E_x(\rho^\tau g(X_\tau)),$$

it is useful to reduce the (infinite-dimensional) set of stopping times to a finite-dimensional subclass. A class often used for this kind of problem is the class of threshold times. This reduction can be carried out in two different ways.

- (i) We use elementary arguments to reduce the set of potential optimal stopping times to the subclass of threshold times, i.e. to stopping times of the form

$$\tau_b = \inf\{t \geq 0: X_t \geq b\}$$

for some $b \in \mathbb{R}$. Then we find the optimal threshold. A summary of examples where this approach can be applied is given below.

- (ii) We make the ansatz that the optimal stopping time is of threshold type, identify the optimal threshold, and use a verification theorem to prove that this stopping time is indeed optimal.

(i) In [3, Section 2] the idea of an elementary reduction of optimal stopping problems to threshold problems was studied in detail. For arbitrary innovations, this approach can be applied to the power-gain function, i.e. $g(x) = x^n$, $n \in \mathbb{N}$. This problem is known as the Novikov–Shiryaev problem and was completely solved for random walks in [13] and [14]. Furthermore, the approach applies to gain functions of call type, i.e. $g(x) = (x - K)^+$. For the special case of nonnegative innovations, a much wider class of gain functions can be handled, such as exponential functions.

(ii) To use this approach, the following easy verification theorem is useful.

Lemma 4.1. *Let $b^* \in \mathbb{R}$, write $v^*(x) = E_x(\rho^{\tau_{b^*}} g(X_{\tau_{b^*}}))$, and assume that*

- (a) $v^*(x) \geq g(x)$ for all $x < b^*$,
 (b) $E(\rho v^*(\lambda x + Z_1)) \leq v^*(x)$ for all $x \in \mathbb{R}$.

Then $v = v^*$ and τ_{b^*} is optimal.

Proof. By the independence of $(Z_n)_{n \in \mathbb{N}}$, property (b) implies that $(\rho^n v^*(X_n))_{n \in \mathbb{N}}$ is a supermartingale under each measure P_x . Since it is positive, the optional sampling theorem leads to

$$v^*(x) \geq \sup_{\tau \in \mathcal{T}} E_x(\rho^\tau v^*(X_\tau)) \geq \sup_{\tau \in \mathcal{T}} E_x(\rho^\tau g(X_\tau)) \quad \text{for all } x \in \mathbb{R},$$

where the second inequality holds by (a) since $v^*(x) = g(x)$ for all $x \geq b$. On the other hand, $v^*(x) \leq v(x)$, i.e. $v^*(x) = v(x)$ and τ_{b^*} is optimal.

Now we are prepared to solve the optimal stopping problem

$$v(x) = \sup_{\tau} E_x(\rho^\tau g(X_\tau)), \quad x \in \mathbb{R}.$$

Section 4 gives conditions for the optimality of threshold times. In this case we can simplify the problem to

$$v(x) = \sup_b E_x(\rho^{\tau_b} g(X_{\tau_b})), \quad x \in \mathbb{R}.$$

Now take an arbitrary starting point x . Then we have to maximize the real function

$$\Psi_x : (x, \infty) \rightarrow \mathbb{R}, \quad b \mapsto E_x(\rho^{\tau_b} g(X_{\tau_b})) = \sum_{i=1}^m \Phi_i^b(x) E(g(b + R^i)),$$

where R^i is PH(e_i, Q)-distributed and $b \geq 0$. The results of the previous section give rise to explicit calculations of $\Phi_i^b(x)$ and Ψ_x . Hence, we are faced with the well-studied maximization problem for real functions that can, e.g. be solved using the standard tools from differential calculus. If we have a maximum point b^* of Ψ_x and $\Psi_x(b^*) > g(x)$ then

$$\tau^* = \inf\{n \in \mathbb{N}_0 : X_n \geq b^*\}$$

is an optimal stopping time when $(X_n)_{n \in \mathbb{N}_0}$ is started in x . A more elegant approach for finding the optimal threshold b^* is the principle of continuous fit.

5. The principle of continuous fit

The principles of smooth and continuous fit play an important role in the study of many optimal stopping problems. The principle of smooth fit was introduced in [10] and has been applied to a variety of problems, ranging from sequential analysis to mathematical finance. The principle of continuous pasting is more recent and was introduced in [15] as a variational principle to solve sequential testing and disorder problems for the Poisson process. For a discussion in the case of Lévy processes and further references, we refer the reader to [2]. Another overview is given in [16, Chapter IV.9], from where we draw the following summary (see page 49):

If X enters the interior of the stopping region S immediately after starting on ∂S , then the optimal stopping point x^* is selected so that the value function v is smooth in x^* . If X does not enter the interior of the stopping region immediately, then x^* is selected so that v is continuous in x^* .

Most applications of this principle involve processes in continuous time. In discrete time an immediate entrance is of course not possible, so we cannot expect the smooth-fit principle to hold. In this section we prove that the continuous-fit principle holds in our setting and we illustrate how it can be used for an easy determination of the optimal threshold.

We keep the notation and assumptions of the previous sections, and, as before, we assume that the optimal stopping set is an interval of the form $[b^*, \infty)$ and consider the optimal stopping time $\tau_{b^*} = \tau = \inf\{n \in \mathbb{N}_0 : X_n \geq b^*\}$.

Furthermore, we assume that

$$\lim_{\varepsilon \searrow 0} \Phi_i^{b^*}(b^* - \varepsilon) = \lim_{\varepsilon \searrow 0} \Phi_i^{b^* + \varepsilon}(b^*) \quad \text{for all } i = 1, \dots, m. \tag{5.1}$$

Note that this condition is obviously fulfilled in the cases discussed above. If g is continuous under an appropriate integrability condition, it furthermore holds that

$$E(g(R_i + \varepsilon + b^*)) \rightarrow E(g(R_i + b^*)) \quad \text{as } \varepsilon \rightarrow 0, i = 1, \dots, m. \tag{5.2}$$

Proposition 5.1. *Assume that (5.1) and (5.2) hold. Then*

$$\lim_{b \nearrow b^*} v(b) = g(b^*).$$

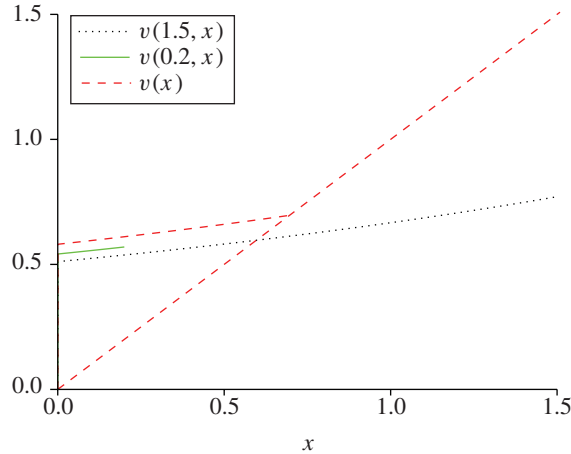


FIGURE 2: Some candidate solutions for different thresholds in the case $g(x) = x$.

Proof. Let $\varepsilon > 0$. First note that $v(b) > g(b)$ for all $b < b^*$, so

$$\liminf_{b \nearrow b^*} v(b) \geq \liminf_{b \nearrow b^*} g(b) = g(b^*).$$

Furthermore, using Corollary 2.1,

$$\begin{aligned} v(b^* - \varepsilon) - g(b^*) &= E_{b^* - \varepsilon}(\rho^\tau g(X_\tau)) - v(b^*) \\ &\leq E_{b^* - \varepsilon}(\rho^\tau g(X_\tau)) - E_{b^*}(\rho^{\tau_{b^* + \varepsilon}} g(X_{\tau_{b^* + \varepsilon}})) \\ &= \sum_{i=1}^m (\Phi_i^{b^*} (b^* - \varepsilon) E(g(R_i + b^*)) - \Phi_i^{b^* + \varepsilon} (b^*) E(g(R_i + \varepsilon + b^*))) \\ &\rightarrow 0 \end{aligned}$$

as $\varepsilon \searrow 0$. This proves that $\limsup_{b \nearrow b^*} v(b) \geq g(b^*)$.

In Figure 2 we illustrate how the continuous-fit principle can be used. We consider the candidate solutions

$$v(b, x) = \begin{cases} \Psi_x(b), & x < b, \\ g(x), & x \geq b, \end{cases}$$

and solve the equation $\Psi_{b-}(b) = g(b)$, where Ψ is defined as in the previous section. If the equation has a unique solution, we can conclude that this solution must be the optimal threshold, as illustrated in the following example.

Example 5.1. We consider the gain function $g(x) = x$ and $\text{Exp}(\mu)$ -distributed innovations; in this setting we always assume that $(X_n)_{n \in \mathbb{N}_0}$ has values in $[0, \infty)$. The discussion in Section 4 guarantees that the optimal stopping time is of threshold type. The optimal threshold can be found by the continuous-fit principle described in the previous section.

The problem is solved if we find a unique $b^* \in [0, \infty)$ that solves the equation

$$b = \Psi_{b-}(b) = \Phi_{b-}(b) \left(b + \frac{1}{\mu} \right) = \frac{\rho \sum_{k \in \mathbb{N}_0} (\rho, \lambda)_k (\mu b \lambda)^k / k!}{\sum_{k \in \mathbb{N}_0} (\rho, \lambda)_k (\mu b)^k / k!} \left(b + \frac{1}{\mu} \right),$$

where we have used Theorem 3.3 in the last step. This equation is equivalent to

$$\sum_{k=0}^{\infty} (\rho, \lambda)_k \frac{\mu^k}{k!} b^{k+1} = \sum_{k=0}^{\infty} \rho (\rho, \lambda)_k \frac{\mu^k \lambda^k}{k!} b^{k+1} + \sum_{k=0}^{\infty} \frac{\rho}{\mu} (\rho, \lambda)_k \frac{\mu^k \lambda^k}{k!} b^k,$$

i.e.

$$\frac{\rho}{\mu} - \sum_{k=0}^{\infty} (\rho, \lambda)_k \frac{\mu^k}{k!} \left(1 - \rho \lambda^k - \frac{\rho}{\mu} (1 - \rho \lambda^k) \frac{\mu \lambda^{k+1}}{k+1} \right) b^{k+1} = 0,$$

i.e.

$$f(b) = 0,$$

where

$$f(b) = \frac{\rho}{\mu} - \sum_{k=0}^{\infty} (\rho, \lambda)_{k+1} \frac{\mu^k}{k!} \left(1 - \frac{\rho \lambda^{k+1}}{k+1} \right) b^{k+1}.$$

Note that $f(0) = \rho/\mu > 0$ and

$$f'(b) = - \sum_{k=0}^{\infty} (\rho, \lambda)_{k+1} \frac{\mu^k}{k!} \left(1 - \frac{\rho \lambda^{k+1}}{k+1} \right) (k+1) b^k < 0 \quad \text{for all } b \in [0, \infty).$$

Since $f(b) \leq \rho/\mu - (1 - \rho)(1 - \rho\lambda)b$, we furthermore obtain $f(b) \rightarrow -\infty$ for $b \rightarrow \infty$. Hence, there exists a unique solution b^* of the transcendental equation $f(b) = 0$.

The optimal stopping time is

$$\tau^* = \inf\{n \in \mathbb{N} : X_n \geq b^*\}$$

and the value function is given by

$$v(x) = \begin{cases} \left(b + \frac{1}{\mu} \right) \rho \frac{\sum_{k \in \mathbb{N}_0} (\rho, \lambda)_k (\mu x \lambda)^k / k!}{\sum_{k \in \mathbb{N}_0} (\rho, \lambda)_k (\mu b^*)^k / k!}, & x < b^*, \\ x, & x \geq b^*. \end{cases}$$

In Figure 2 v is plotted for the parameters $\mu = 1$ and $\rho = \lambda = \frac{1}{2}$.

References

- [1] ASMUSSEN, S. (2003). *Applied Probability and Queues* (Appl. Math. **51**), 2nd edn. Springer, New York.
- [2] CHRISTENSEN, S. AND IRLE, A. (2009). A note on pasting conditions for the American perpetual optimal stopping problem. *Statist. Prob. Lett.* **79**, 349–353.
- [3] CHRISTENSEN, S., IRLE, A. AND NOVIKOV, A. (2011). An elementary approach to optimal stopping problems for AR(1) sequences. *Sequent. Anal.* **30**, 79–93.
- [4] FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications*, Vol. II. John Wiley, New York.
- [5] FINSTER, M. (1982). Optimal stopping on autoregressive schemes. *Ann. Prob.* **10**, 745–753.
- [6] FRISÉN, M. AND SONESSON, C. (2006). Optimal surveillance based on exponentially weighted moving averages. *Sequent. Anal.* **25**, 379–403.
- [7] GASPER, G. AND RAHMAN, M. (2004). *Basic Hypergeometric Series* (Encyclopedia Math. Appl. **96**), 2nd edn. Cambridge University Press.
- [8] GOLDIE, C. M. AND MALLER, R. A. (2000). Stability of perpetuities. *Ann. Prob.* **28**, 1195–1218.
- [9] JACOBSEN, M. AND JENSEN, A. T. (2007). Exit times for a class of piecewise exponential Markov processes with two-sided jumps. *Stoch. Process. Appl.* **117**, 1330–1356.

- [10] MIKHALEVICH, V. S. (1958). Bayesian choice between two hypotheses for the mean value of a normal process. *Visnik Kiyv. Univ.* **1**, 101–104.
- [11] NOVIKOV, A. A. (2009). On distributions of first passage times and optimal stopping of AR(1) sequences. *Theory Prob. Appl.* **53**, 419–429.
- [12] NOVIKOV, A. AND KORDZAKHIA, N. (2008). Martingales and first passage times of AR(1) sequences. *Stochastics* **80**, 197–210.
- [13] NOVIKOV, A. A. AND SHIRYAEV, A. N. (2004). On an effective case of the solution of the optimal stopping problem for random walks. *Teor. Veroyat. Primen.* **49**, 373–382 (in Russian). English translation: *Theory Prob. Appl.* **49** (2005), 344–354.
- [14] NOVIKOV, A. AND SHIRYAEV, A. (2007). On solution of the optimal stopping problem for processes with independent increments. *Stochastics* **79**, 393–406.
- [15] PEŠKIR, G. AND SHIRYAEV, A. N. (2000). Sequential testing problems for Poisson processes. *Ann. Statist.* **28**, 837–859.
- [16] PEŠKIR, G. AND SHIRYAEV, A. (2006). *Optimal Stopping and Free-Boundary Problems*. Birkhäuser, Basel.
- [17] RUNGER, G. C. AND PRABHU, S. S. (1996). A Markov chain model for the multivariate exponentially weighted moving averages control chart. *J. Amer. Statist. Assoc.* **91**, 1701–1706.