Optimality of an Affine Intensity Policy for Maximizing the Probability of an Arrival Count in Point-Process Intensity Control

Boris Defourny[1]

Department of Industrial and Systems Engineering, Lehigh University, 200 W Packer Ave, Bethlehem, PA, 18015, USA

Abstract

This paper considers the problem of maximizing the probability of attaining a prescribed count of arrivals generated by a point process, by controlling its intensity. Our analysis shows the existence of optimal intensity switching times that are affine in the arrival count, thereby contributing to the literature on the optimality of affine policies. The optimal intensity control law is established, along with closed-form expressions for its numerical parameters. Several properties of the value function are listed as well.

Keywords: Point processes, Optimal stochastic control, Intensity control, Affine decision rules

1. Introduction

This paper considers the problem of controlling the intensity of a point process in order to maximize the probability that a target number of arrivals is met exactly by a deadline, assuming the intensity is allowed to vary within a finite range [1]. The point process is assumed to be a simple point process, that is, arrivals happen one at a time. Mathematically, the problem can be formulated as an optimal point-process intensity control problem,

$$\begin{aligned} V &= \max_{\pi} \mathbb{P}^{\pi} \left[S_T = C | S_0 = 0 \right] \\ &= \max_{\pi} \mathbb{E}^{\pi} \left[R_T(S_T) | S_0 = 0 \right], \end{aligned} \tag{1}$$

where the state S_t represents the count of arrivals during the time period $(0, t]$, following a point process with controlled intensity

$$\lambda_t = \lambda_t^{\pi}(S_t) \in [\lambda_a, \lambda_b], \qquad 0 < \lambda_a < \lambda_b < \infty. \tag{2}$$

The terminal reward function is defined as

$$R_T(S_T) = 1_C(S_T) \tag{3}$$

which is equal to 1 if $S_T = C$ and 0 otherwise. The control law λ_t^{π} to be optimized is a function of the state $S_t \in \mathbb{N}$ and of the time $t \in [0, T]$.

The number of arrivals during a small time interval $(t, t + dt]$ follows a Poisson distribution of mean $\lambda_t dt$. At the first-order, $\mathbb{P}(S_{t+dt} - S_t = 1) = \lambda_t dt + o(dt)$, $\mathbb{P}(S_{t+dt} - S_t \geq 2) = o(dt)$, and $\mathbb{P}(S_{t+dt} - S_t = 0) = (1 - \lambda_t dt) + o(dt)$ [2]. Thus, while λ_t is controlled, there is no direct control over the arrival times (since $\lambda_b < \infty$).

Our interest in (1) stems from the fact that it represents one of the simplest possible point-process control problems where the decision is the intensity, and yet its optimal solution has not been described satisfactorily. What is known is the existence of an optimal bang-bang intensity policy [1, 3]. The Hamilton-Jacobi-Bellman equations characterizing the optimal solution to the continuous-time, infinite-dimensional control problem are known [4]. As a result, solutions to discretized versions of the problem can be obtained via numerical solution algorithms, developed for instance in [5]. An infinite-dimensional, nonconvex formulation for policy optimization is known as well [1].

While numerical approaches are applicable, there is value in pursuing the analysis of the optimal control problem further. Additional insights facilitate sensitivity and robustness studies and suggest approaches to tackle higher-dimensional problems. Point-process control has applications in areas such as queueing systems, inventory control, and revenue management. There is often a direct relationship between intensities and prices – for instance, [6] hypothesizes a relationship between prices and rate of arrival of customers to optimize a price-setting policy for selling an inventory of perishable products. In [7], the authors hypothesize that productivity is optimized by receiving the right workload, and let the manager control the rate of arrival of tasks. Interest in intensity-control problems is also seen in [8], motivated by online social network applications where the target state is a desired user behavior. In our case, as the probability of meeting a count drops to zero when the count exceeds the target, (1) captures situations where exceeding a capacity can have catastrophic consequences. These situations are found in the airline industry, insurance industry, and electricity industry.

The present paper describes the structure of the optimal value function, and provides the exact optimal intensity

[1]Corresponding author. Email address: defourny@lehigh.edu

control policy in closed-form, thereby furnishing a definitive answer to (1). Namely, given a feasible intensity range $[\lambda_a, \lambda_b]$, we establish the existence of a *characteristic duration*

$$\theta = \frac{\log(\lambda_b/\lambda_a)}{\lambda_b - \lambda_a} \quad (4)$$

that fully describes the optimal policy, in the sense that it is optimal to have the intensity at time t set to λ_a if t is below the critical time $T - \theta(C - S_t)$, and to λ_b if t exceeds the critical time $T - \theta(C - S_t)$. The intensity switches back to λ_a if there is an arrival at time t but t does not exceed the critical time of the new state.

The fact that the critical times triggering the intensity switches from λ_a to λ_b are affine in the arrival count is a remarkable feature of the problem. To the best of our knowledge, this is the first time an intensity policy affine in the switching times is shown to be optimal, thereby contributing to the literature on the optimality of affine policies [9]. An additional pleasant characteristic of the optimal intensity policy is the availability of a closed-form expression for its numerical parameters.

The presentation is organized as follows. Section 2 recalls the optimality conditions for the intensity control problem. Section 3 describes the structure of the optimal value function, which represents the probability of reaching the target count under the optimal intensity policy, given the current information. Section 4 establishes the optimality of the affine intensity policy. Section 5 illustrates the results, and Section 6 concludes.

2. Optimality conditions

Optimality conditions for the stated problem are well known, see e.g. Chapter VII §2 in [4], and lead to the conclusion that there exists an optimal policy such that $\lambda_t^\pi(S_t) \in \{\lambda_a, \lambda_b\}$. Our statement slightly differs from [4] in that we use directional derivatives, while [4] assumes the differentiability of the value function.

Let $V_{T-t}(s) = \text{Prob}(S_T = C \mid S_{T-t} = s)$ under an optimal policy. From the properties of simple point processes, for a small $dt > 0$, it holds that $\text{Prob}(S_{t+dt} - S_t = 1) = \lambda dt + o(dt)$, $\text{Prob}(S_{t+dt} - S_t \geq 2) = o(dt)$, and $\text{Prob}(S_{t+dt} - S_t = 0) = 1 - \lambda dt + o(dt)$. Therefore, by backward induction,

$$\begin{aligned}
&V_{T-(t+dt)}(s)\\
&= \sup_{\lambda \in [\lambda_a, \lambda_b]} \left[\textstyle\sum_{n=0}^{\infty} \text{Prob}(S_{T-t} - S_{T-(t+dt)} = n) V_{T-t}(s+n)\right]\\
&= \sup_{\lambda \in [\lambda_a, \lambda_b]} [(1 - \lambda dt) V_{T-t}(s) + \lambda dt V_{T-t}(s+1) + o(dt)],\\
&V_{T-(t+dt)}(s) - V_{T-t}(s)\\
&= \sup_{\lambda \in [\lambda_a, \lambda_b]} [\lambda dt (V_{T-t}(s+1) - V_{T-t}(s)) + o(dt)]. \qquad (5)
\end{aligned}$$

The $o(dt)$ term can be considered to be independent of λ, as $|V_{T-t}(s)| \leq 1$ for all s and the supremum is over a compact interval. Dividing both sides by $dt = 1/m$, letting $m \to \infty$, and invoking the uniform convergence of $f_m(\lambda) = \lambda(V_{T-t}(s+1) - V_{T-t}(s)) + o(1/m)$ to $f(\lambda) = \lambda(V_{T-t}(s+1) - V_{T-t}(s))$ at least locally over $[\lambda_a, \lambda_b]$, this relation becomes $\partial V_{T-t}(s)/\partial t = \sup_{\lambda \in [\lambda_a, \lambda_b]} [\lambda(V_{T-t}(s+1) - V_{T-t}(s))]$, where $\partial V_{T-t}(s)/\partial t = \lim_{dt \to 0^+} [V_{T-(t+dt)}(s) - V_{T-t}(s)]/dt$ denotes the derivative in the direction of an increasing time-to-go. As $f(\lambda)$ is linear in λ, the optimum is attained at the boundary of the feasibility set and we have

$$\begin{aligned}
\partial V_{T-t}(s)/\partial t &= \lambda_a (V_{T-t}(s+1) - V_{T-t}(s))\\
&\quad \text{if} \quad V_{T-t}(s+1) - V_{T-t}(s) < 0,\\
\partial V_{T-t}(s)/\partial t &= \lambda_b (V_{T-t}(s+1) - V_{T-t}(s))\\
&\quad \text{if} \quad V_{T-t}(s+1) - V_{T-t}(s) > 0. \qquad (6)
\end{aligned}$$

We have $V_{T-t}(s) = 0$ for $s \geq C + 1$, since the target is missed forever if exceeded. At $t = 0$ we have $V_{T-0}(C) = 1$ and $V_{T-0}(s) = 0$ for $s \neq C$. The choice of λ where $V_{T-t}(s+1) - V_{T-t}(s) = 0$ is inconsequential as far as optimality is concerned.

3. Unimodality properties

We first show by induction that for each fixed time $T - t$, the sequence $\{V_{T-t}(s)\}_{s=0,1,\ldots}$ over the states s is unimodal.

Proposition 1. *For each time $T - t$ there exists a critical state C^*_{T-t} such that $V_{T-t}(s-1) \leq V_{T-t}(s)$ for $s \leq C^*_{T-t}$, and $V_{T-t}(s) \geq V_{T-t}(s+1)$ for $s \geq C^*_{T-t}$.*

Proof. The property is true for $t = 0$ with $C^*_{T-0} = C$, since for any $k \geq 1$ we have $0 = V_T(C-k) \leq V_T(C) = 1 \geq V_T(C+k) = 0$. Suppose this is true at time $T - t$. Then at time $T - (t + dt)$, from (5) it is optimal to choose $\lambda = \lambda_b$ at states $s \leq C^*_{T-t} - 1$, and $\lambda = \lambda_a$ at states $s \geq C^*_{T-t}$. Furthermore, (i) At states $s \geq C^*_{T-t}$ we have

$$\begin{aligned}
&V_{T-(t+dt)}(s)\\
&= (1 - \lambda_a dt) V_{T-t}(s) + (\lambda_a dt) V_{T-t}(s+1) + o(dt)\\
&\geq (1 - \lambda_a dt) V_{T-t}(s+1) + (\lambda_a dt) V_{T-t}(s+2) + o(dt)\\
&= V_{T-(t+dt)}(s+1),
\end{aligned}$$

using $V_{T-t}(C^*_{T-t}) \geq V_{T-t}(s) \geq V_{T-t}(s+1) \geq V_{T-t}(s+2)$ from the induction hypothesis. This establishes $V_{T-(t+dt)}(s) \geq V_{T-(t+dt)}(s+1)$ for states $s \geq C^*_{T-t}$.

(ii) At states $s \leq C^*_{T-t} - 1$, we have

$$\begin{aligned}
&V_{T-(t+dt)}(s-1)\\
&= (1 - \lambda_b dt) V_{T-t}(s-1) + (\lambda_b dt) V_{T-t}(s) + o(dt)\\
&\leq (1 - \lambda_b dt) V_{T-t}(s) + (\lambda_b dt) V_{T-t}(s+1) + o(dt)\\
&= V_{T-(t+dt)}(s),
\end{aligned}$$

using $V_{T-t}(s-1) \leq V_{T-t}(s) \leq V_{T-t}(s+1) \leq V_{T-t}(C^*_{T-t})$ from the induction hypothesis. This establishes $V_{T-(t+dt)}(s-1) \leq V_{T-(t+dt)}(s)$ for states $s \leq C^*_{T-t} - 1$.

(iii) At state $s = C^*_{T-t}$, we can have either $V_{T-(t+dt)}(s-1) \leq V_{T-(t+dt)}(s)$ (and thus we set $C^*_{T-(t+dt)} = C^*_{T-t}$) or

$V_{T-(t+dt)}(s-1) > V_{T-(t+dt)}(s)$ (and thus we set $C^*_{T-(t+dt)} = C^*_{T-t} - 1$).

Together, (i-ii-iii) prove the claim. $\square$

Regarding the evolution of the critical state C^*_{T-t} over the time-to-go t, the property $C^*_{T-(t+dt)} \in \{C^*_{T-t} - 1, C^*_{T-t}\}$ shows that C^*_{T-t} is nonincreasing in t. This leads to the conclusion that $V_{T-t}(s)$ is unimodal in t for each state s:

Proposition 2. *For each fixed state $s \geq 0$, there exists a critical time-to-go t^*_s such that $V_{T-t}(s)$ is nondecreasing over $t \in [0, t^*_s)$ and nonincreasing over $t \in [t^*_s, T]$.*

Proof. For any fixed state s, having $s \geq C^*_{T-t}$ implies $s \geq C^*_{T-(t+dt)}$ since $C^*_{T-t} \geq C^*_{T-(t+dt)}$. This implies that we can define for each state s a critical time-to-go

$$t^*_s = \inf\{t \in [0, T] : s \geq C^*_{T-t}\}$$

such that $s \leq C^*_{T-t} - 1$ over $t \in [0, t^*_s)$ and $s \geq C^*_{T-t}$ over $t \in [t^*_s, T]$. In the time-to-go interval $[0, t^*_s)$ we have $V_{T-t}(s) \leq V_{T-t}(s+1)$ by definition of C^*_{T-t}, and thus by (6), λ_b is optimal. In the interval $[t^*_s, T]$ we have $V_{T-t}(s) \geq V_{T-t}(s+1)$ by definition of C^*_{T-t}, and thus λ_a is optimal.

Referring back to (6) again, for each s we have $\partial V_{T-t}(s)/\partial t \geq 0$ over $[0, t^*_s)$ and then $\partial V_{T-t}(s)/\partial t \leq 0$ over $[t^*_s, T]$, showing that $V_{T-t}(s)$ as a function of t is unimodal, and attains its maximum at t^*_s. $\square$

A key property of the critical time-to-go is that it is nonincreasing in the state:

Proposition 3. *It holds that $s \leq s'$ implies $t^*_s \geq t^*_{s'}$.*

Proof. By definition of t^*_s, we have $s' \geq s \geq C^*_{T-t^*_s}$. Then by definition of $t^*_{s'}$ we have $s' \geq C^*_{T-t^*_{s'}}$ for some $C^*_{T-t^*_{s'}} \geq C^*_{T-t^*_s}$. As C_{T-t} is nonincreasing in t, we have $t^*_{s'} \leq t^*_s$. $\square$

4. Optimal intensity policy

Consider the reachable space $E = \{(t, k) \in \mathbb{R} \times \mathbb{Z} : t \in [0, T], k \leq C\}$ of pairs (t, k) where t is the time-to-go (which physically decreases from T to 0) and k is the count-to-go (which physically starts at $k = C$ and then is decremented of one unit at each arrival). Let E_a, E_b denote the subsets of E where it is optimal to choose λ_a and λ_b respectively. The key result of the paper is that the optimal switching times from λ_a to λ_b are affine in the count-to-go:

Proposition 4. *The boundary between the regions E_a, E_b is described by the indifference line*

$$t = k\theta, \qquad \theta = \frac{\log(\lambda_b/\lambda_a)}{\lambda_b - \lambda_a}.$$

In terms of the physical time $T-t$ and arrival count $s = C-k$ at time $T-t$, an optimal intensity policy is

$$\lambda^\pi_{T-t}(s) = \begin{cases} \lambda_a & \text{if } T-t < T-(C-s)\theta, \\ \lambda_b & \text{if } T-t \geq T-(C-s)\theta. \end{cases} \tag{7}$$

The remainder of the section is concerned with establishing the proposition. The originality of the approach is that we proceed by induction over the states, rather than by induction backward in time, as usually done in dynamic programming. This is technically possible because along sample paths, with probability 1 the count of arrivals is nondecreasing.

It will be convenient to set

$$\rho_a = \frac{\lambda_a}{\lambda_b - \lambda_a}, \qquad \rho_b = \frac{\lambda_b}{\lambda_b - \lambda_a}. \tag{8}$$

We start from state $s = C$. At this state, the optimal intensity is $\lambda = \lambda_a$. From $\frac{\partial V_{T-t}}{\partial t}(C) = \lambda_a(0 - V_{T-t}(C))$ and the initial condition $V_{T-0}(C) = 1$, we find

$$V_{T-t}(C) = e^{-\lambda_a t} \quad \text{for } 0 \leq t \leq T.$$

At state $s = C-1$, we have $\lambda^\pi_{T-t}(s) = \lambda_b$ for $0 \leq t < t_1$ and $\lambda^\pi_{T-t}(s) = \lambda_a$ for $t_1 \leq t \leq T$, for some t_1 to be determined. This structure holds thanks to Propositions 2 and 3: $V_{T-t}(C-1)$ starts from 0 at $t = 0$ and increases over t until it meets $V_{T-t}(C)$ at $t = t_1$, reaching its maximum and triggering the switch to λ_b while $V_{T-t}(C)$ keeps decreasing. The solution over $0 \leq t \leq t_1$ must satisfy $\frac{\partial V_{T-t}}{\partial t}(C-1) = \lambda_b(V_{T-t}(C) - V_{T-t}(C-1))$, with initial condition $V_{T-0}(C-1) = 0$. We find

$$\begin{aligned} V_{T-t}(C-1) &= \rho_b[e^{-\lambda_a t} - e^{-\lambda_b t}] \\ &= \rho_b e^{-\lambda_a t}(1 - e^{-(\lambda_b - \lambda_a)t}) \quad \text{for } 0 \leq t \leq t_1. \end{aligned}$$

The time t_1 is precisely when $V_{T-t}(C-1) = V_{T-t}(C)$, which corresponds to $\partial V_{T-t}(C-1)/\partial t = 0$. This condition translates to $e^{-\lambda_a t}\rho_b[1 - e^{-(\lambda_b-\lambda_a)t}] = e^{-\lambda_a t}$, that is, $\rho_b[1 - e^{-(\lambda_b-\lambda_a)t_1}] = 1$, that is, $e^{-(\lambda_b-\lambda_a)t_1} = \lambda_a/\lambda_b$, that is,

$$t_1 = \frac{\log(\lambda_b) - \log(\lambda_a)}{\lambda_b - \lambda_a} := \theta.$$

Note the maximum $\bar{v}_{C-1} := V_{T-t_1}(C-1) = (\lambda_a/\lambda_b)^{\rho_a}$.

The solution over $t_1 \leq t \leq T$ must satisfy $\frac{\partial V_{T-t}}{\partial t}(C-1) = \lambda_a(V_{T-t}(C) - V_{T-t}(C-1))$, with $V_{T-t_1}(C-1) = \bar{v}_{C-1}$. We find

$$\begin{aligned} &V_{T-t}(C-1) \\ &= (\lambda_a t)e^{-\lambda_a t} + (1 - \lambda_a t_1)e^{-\lambda_a t} \quad \text{for } t_1 \leq t \leq T. \end{aligned}$$

We proceed with state $s = C-2$. At this state we have $\lambda^\pi_{T-t}(s) = \lambda_b$ for $0 \leq t < t_2$ and $\lambda^\pi_{T-t}(s) = \lambda_a$ for $t_2 \leq t \leq T$, for some $t_2 \geq t_1$. The solution over $0 \leq t \leq t_2$ must satisfy $\frac{\partial V_{T-t}}{\partial t}(C-2) = \lambda_b(V_{T-t}(C-1) - V_{T-t}(C-2))$, with initial condition $V_{T-0}(C-1) = 0$. The solution restricted to $0 \leq t \leq t_1$ is

$$\begin{aligned} &V_{T-t}(C-2) \\ &= \rho_b^2[e^{-\lambda_a t} - e^{-\lambda_b t}(1 + t(\lambda_b - \lambda_a))] \\ &= \rho_b^2 e^{-\lambda_a t}[1 - e^{-(\lambda_b-\lambda_a)t}(1 + (\lambda_b - \lambda_a)t)] \quad \text{for } 0 \leq t \leq t_1. \end{aligned}$$

At t_1 we have $V_{T-t_1}(C-2) = (\lambda_a/\lambda_b)^{\rho_a}\rho_b(1-\rho_a\log(\lambda_b/\lambda_a))$. The solution restricted to $t_1 \le t \le t_2$, using the value of $V_{T-t_1}(C-2)$ as the boundary condition, is

$$V_{T-t}(C-2) = \rho_b[\lambda_a t e^{-\lambda_a t} + e^{-\lambda_a t}(1-\rho_a-\lambda_a t_1) + e^{-\lambda_b t}(\rho_b - \lambda_b t_1)] \quad \text{for } t_1 \le t \le t_2.$$

The time t_2 corresponds to the maximum of $V_{T-t}(C-2)$ over $t \ge t_1$. The condition $\partial V_{T-t}(C-2)/\partial t = 0$ corresponds to finding the solution $t = t_2$ satisfying

$$\frac{1}{\lambda_b-\lambda_a}\Big[e^{-\lambda_a t}\lambda_a^2(1-t(\lambda_b-\lambda_a)+\log(\lambda_b/\lambda_a)) - e^{-\lambda_b t}\lambda_b^2(1-\log(\lambda_b/\lambda_a))\Big] = 0.$$

With $t = 2t_1$, we have $e^{-\lambda_a t}\lambda_a^2 = (\lambda_a e^{-\lambda_a t_1})^2 = (\lambda_b e^{-\lambda_b t_1})^2 = e^{-\lambda_b t}\lambda_b^2$, and also $1-t(\lambda_b-\lambda_a)+\log(\lambda_b/\lambda_a) = 1-2\log(\lambda_b/\lambda_a)+\log(\lambda_b/\lambda_a) = 1-\log(\lambda_b/\lambda_a)$, showing that the solution is

$$t_2 = 2t_1 = 2\theta.$$

The value of the maximum is $\bar{v}_{C-2} := V_{T-t_2}(C-2) = (\lambda_a/\lambda_b)^{2\rho_a}[1-\rho_a\log(\lambda_a/\lambda_b)]$.

Finally, the solution restricted to $t_2 \le t \le T$ must satisfy $\frac{\partial V_{T-t}}{\partial t}(C-2) = \lambda_a(V_{T-t}(C-1) - V_{T-t}(C-2))$, with $V_{T-t_1}(C-2) = \bar{v}_{C-2}$. We find

$$V_{T-t}(C-2) = \frac{(\lambda_a t)^2}{2}e^{-\lambda_a t} + (1-\lambda_a t_1)e^{-\lambda_a t}(1+\lambda_a t) \quad \text{for } t_2 \le t \le T.$$

At this point, we have established that $V_{T-t}(C-i)$ attains its maximum over t at $t_i = i\theta$ for $i = 0,1,2$. To generalize the reasoning to all states and all times, we proceed inductively. Suppose that $V_{T-t}(C-i)$ attains its maximum over t at $t_i = i\theta$ for $i = 0,\dots,k$. Under this assumption, the value function can be expressed as follows. For convenience, we index vectors and matrices starting from the index 0. Let $x_t \in \mathbb{R}^{C+1}$ be the vector with elements

$$x_{tk} = V_{T-t}(C-k) \quad \text{for } k = 0,\dots,C. \tag{9}$$

The initial conditions at $t=0$ impose $x_{00} = 1$ and $x_{0k} = 0$ for $k = 1,\dots,C$. Let $A_\ell \in \mathbb{R}^{(C+1)\times(C+1)}$ for $\ell = 0,1,\dots,C$ be the matrix with nonzero elements $(A_\ell)_{ii} = -\lambda_a$ for $i = 0,\dots,\ell$, $(A_\ell)_{i,i-1} = \lambda_a$ for $i = 1,\dots,\ell$, and $(A_\ell)_{ii} = -\lambda_b$ and $(A_\ell)_{i,i-1} = \lambda_b$ for $i = \ell+1,\dots,C$.

In matrix form, the differential equations for $V_{T-t}(C-k)$ are expressed as $\partial x_t/\partial t = A_\ell x_t$ over $\ell\theta \le t \le (\ell+1)\theta$. Thus, the value function backwards in time has the dynamics of a linear switched system [10] with switching times $i\theta$, and we have, using matrix exponentials,

$$x_t = \exp(A_\ell(t-\ell\theta))\exp(A_{\ell-1}\theta)\dots\exp(A_0\theta)x_0, \quad \ell = \min(C, \lfloor t/\theta\rfloor). \tag{10}$$

Now, we show that $V_{T-t}(C-(k+1))$ attains its maximum over t at $t_{k+1} = (k+1)\theta$. From the differential equation valid for $t_k \le t \le t_{k+1}$, where t_{k+1} is to be determined, we have $x_t = \exp(A_k(t-t_k))x_{t_k}$ and consequently

$$\frac{\partial x_t}{\partial t} = A_k[\exp(A_k(t-t_k))]x_{t_k}, \quad t_k \le t \le t_{k+1}. \tag{11}$$

We have to show that $(\partial x_t/\partial t)_{k+1} = (A_k(\exp(A_k(t-k\theta))\exp(A_{k-1}\theta)\dots\exp(A_0\theta)x_0))_{k+1} = 0$ at $t = t_{k+1}$. By inspection of A_k, this means that the elements k and $k+1$ of $x_t = [\exp(A_k(t-k\theta))\,\exp(A_{k-1}\theta)\,\dots\,\exp(A_0\theta)]x_0$ should be equal at $t = t_{k+1}$.

We proceed by verification. Based on the conjecture that $t_{k+1} = (k+1)\theta$, we calculate the expression of x_t at $t = (k+1)\theta$ for the relevant states. Our result, valid with $\theta = (\log(\lambda_b/\lambda_a))/(\lambda_b-\lambda_a)$, is the following. Define

$$c_{k+1,j} := \frac{j(-(k+1))^{j-1} + (-(k+1))^j}{j!}\ .$$

For a given $k+1$ and $0 \le i \le k+1$, one has

$$\begin{aligned} V_{T-t_{k+1}}(C-i) &= x_{t_{k+1},i} \\ &= [\exp(A_k\theta)\exp(A_{k-1}\theta)\dots\exp(A_0\theta)x_0]_i \\ &= (\lambda_a/\lambda_b)^{(k+1)\rho_a}\sum_{j=0}^{i} c_{k+1,j}\,(-\lambda_a\theta)^j \\ &= (\lambda_a/\lambda_b)^{(k+1)\rho_a}\sum_{j=0}^{i}\left(1-\frac{j}{k+1}\right)\frac{((k+1)\lambda_a\theta)^j}{j!}\ . \end{aligned} \tag{12}$$

To find this, we focused on the element $x_{t_{k+1},k+1}$. We evaluated the matrix exponential products using symbolic computations for a few k's, and then tried to identify the pattern behind the numerical coefficients $c_{k+1,j}$ of the terms of the sum over j. Namely, we find that the coefficients $c_{k+1,j}$ for $(x_{t_{k+1}})_{k+1}$ are produced by the generating function

$$g(z) = (1+z)e^{-(k+1)z},$$

and from there we obtain the analytical form of coefficients via $c_{k+1,j} = \frac{\partial^j g(z)/\partial z^j}{j!}\Big|_{z=0}$.

Observe now that setting $i = k$ or $i = k+1$ leads to the same value of the sum over $j = 0,\dots,i$, since the factor $(1-j/(k+1))$ is zero at $j = k+1$. This completes the proof that $(x_{t_{k+1},k}) = (x_{t_{k+1},k+1})$. Therefore, we have established the optimality of the policy (7), completing the proof of Proposition 4.

As a side note, while the expression of $V_{T-t_k}(C-i))$ for $i = 0,\dots,k$ has been established in the course of the proof (by reading (12) with k replacing $k+1$), with a similar approach we can establish the expression of $V_{T-t_k}(C-i)$ for $i = k+1$:

$$x_{t_k,i}|_{i=k+1} = \left(\frac{\lambda_a}{\lambda_b}\right)^{k\rho_a}\sum_{j=0}^{i-1}\Big[\frac{\lambda_b}{\lambda_a}(i-j)i^{j-1}(-\rho_a)^{i-j} + (k-j)k^{j-1}\left(1-(-\rho_a)^{i-j}\right)\Big]\frac{(\lambda_a\theta)^j}{j!}\ . \tag{13}$$

Knowing the optimal policy, we can compute $V_{T-t}(C-k)$ over $t_k \le t \le T$ recursively for $k = 0, \ldots, C$, by finding functions $V_{T-t}(C-k)$ such that $\dfrac{\partial V_{T-t}}{\partial t}(C-k) = \lambda_a(V_{T-t}(C-(k-1)) - V_{T-t}(C-k))$, subject to the boundary condition $V_{T-t_k}(C-k) = V_{T-t_k}(C-(k-1))$. The general solution of this problem, restricted to $T-t \in [0, T-k\theta]$, is

$$V_{T-t}(C-k) = \frac{(\lambda_a t)^k}{k!} e^{-\lambda_a t} + (1-\lambda_a\theta) e^{-\lambda_a t} \sum_{j=0}^{k-1} \frac{(\lambda_a t)^j}{j!}. \tag{14}$$

The result can be verified by checking that $V_{T-t}(C-k)$ satisfies the differential equations and that the boundary conditions hold, for $k = 0, 1, \ldots, C$. The expression $e^{-\lambda_a t} \sum_{j=0}^{k-1} (\lambda_a t)^j / j!$ in (14) coincides with the cumulative distribution function of a Poisson random variable of mean $\lambda_a t$ evaluated at $k-1$. It can thus also be evaluated as $Q(k, \lambda_a t)$ where $Q(k, x) = \int_x^\infty u^{k-1} e^{-u} du \Big/ \int_0^\infty u^{k-1} e^{-u} du$ is the upper regularized gamma function.

If $T \ge \theta C$, we obtain the optimal value (1) by setting $k = C$ and $t = T$ in (14):

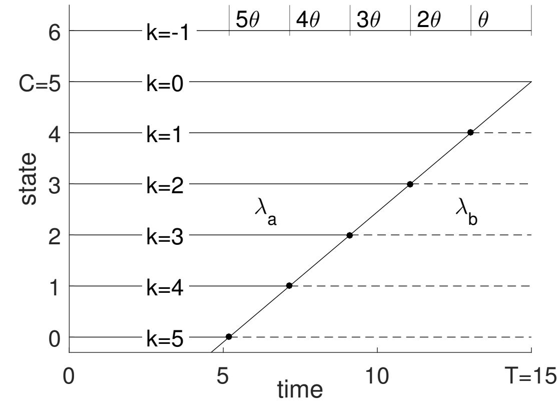
$$V = \frac{(\lambda_a T)^C}{C!} e^{-\lambda_a T} + (1-\lambda_a\theta)\, Q(C, \lambda_a T) \;\text{ if } T \ge \theta C.$$

If $T < \theta C$, the value function can be obtained via (10).

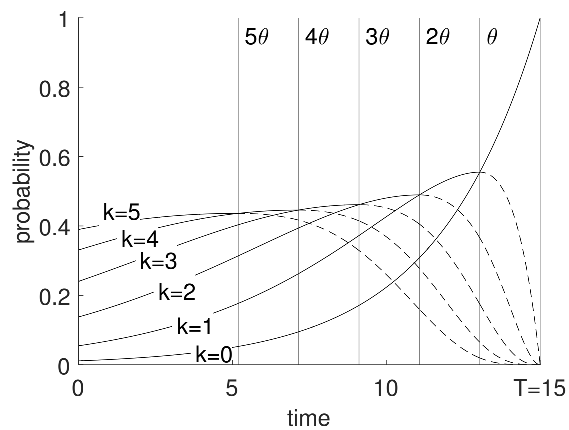
5. Numerical Illustration

The solution for the case $\lambda_a = 0.3$, $\lambda_b = 0.8$, $C = 5$, $T = 15$ is illustrated on Figure 1. Figure 1a depicts the optimal policy and can be read as follows. Being at time t in state s corresponds to the point (t, s). The passage of time increases t, while the random arrivals increment s (or equivalently decrement k). Having (t, s) on a continuous line indicates that the intensity λ_a is optimal. Having (t, s) on a dashed line indicates that the intensity λ_b is optimal. The critical times are indicated with dots. The dot markers are connected by a line to emphasize the affine dependence in s.

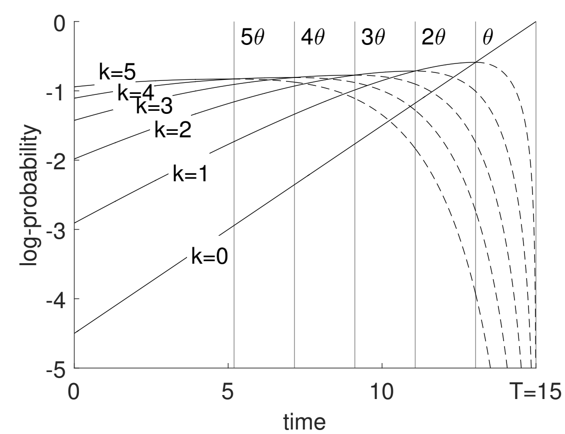
Figure 1b depicts the value function $V_t(C-k)$, which represents the optimal probability of meeting the count C at time T, given that the count is $s = C-k$ at time t. The optimal intensities can be read from the continuous or dashed line convention. Initially $t = 0$ and $k = C = 5$. If the curve relative to a count-to-go k is above the curve relative to $k-1$, the intensity λ_a is used to maximize the probability of staying on the curve k. If the curve relative to k is below the curve relative to $k-1$, the intensity λ_b is used to maximize the probability of jumping to the curve $k-1$. Figure 1c is Figure 1b but with a logarithmic scale for the value function.



(a) Optimal intensity policy $\lambda_t^\pi(s)$, with $k = C - s$



(b) Optimal Probability $\mathbb{P}^\pi(S_T = C \mid S_t = C-k)$



(c) Log-Probability $\log \mathbb{P}^\pi(S_T = C \mid S_t = C-k)$

Figure 1: Solution for $\lambda_a = 0.3$, $\lambda_b = 0.8$, target $C = 5$, horizon $T = 15$. $k = C - s$ is the count-to-go. Continuous lines indicate that $\lambda = \lambda_a$ under the optimal policy, dashed lines indicate $\lambda = \lambda_b$.

6. Discussion

As a function of k, the value function $V_{T-t}(C-k)$ coincides with the probability of having k arrivals during the remaining period of duration t. Given the form of the optimal policy, we can decompose t into $[t_0, t_1), [t_1, t_2), \ldots, [t_m, t)$ where $t_\ell = \ell\theta$ and $m = \min(C, \lfloor t/\theta \rfloor)$. Equation (10) can then be interpreted as follows: the elements $i, j \ge 0$ of the matrix $M_\ell := \exp(A_\ell \theta)$ is the probability of reaching state $C-j$ at time $T-\theta\ell$ while starting from state $C-i$ at time $T-\theta(\ell+1)$. $M_{\ell,ij}$ also represents the probability of $N_\ell = j-i$ arrivals during the period $(T-t_{\ell+1}, T-t_\ell]$, while starting from state $C-i$ at time $T-\theta(\ell+1)$.

We conclude the paper with two conjectures.

Conjecture 1. *$V_{T-t}(C-k)$ is log-concave in t.*

Conjecture 1 is illustrated in Figure 1c, which suggests that the functions $t \mapsto \log V_t(C-k)$ are concave. The conjecture is true when $\lambda_a = \lambda_b = \lambda$, since in this case $x_{tk} = V_{T-t}(C-k) = e^{-\lambda t}(\lambda t)^k/k!$. One verifies that $x_{t,k+1}/x_{tk} \leq x_{tk}/x_{t,k-1}$, which shows that for each fixed t, the sequence x_{tk} in k is log-concave. From there one verifies that $\partial^2 \log V_{T-t}(s)/\partial t^2 \leq 0$, and finally, $V_{T-t}(s)$ log-concave in t implies $V_t(s)$ log-concave in t.

Conjecture 2. *An optimal intensity policy with switching times affine in the positive count-to-go exists when the terminal reward $1_C(S_T)$ is extended to the 0-1 indicator for $S_T \in \{C, C+1, \ldots, C+r\}$.*

Conjecture 2 originates from tests on discrete-time numerical approximations. We do not know if a new characteristic duration θ can be found in closed-form.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1610825.

References

[1] P. Brémaud, Bang-bang controls of point processes, Advances in Applied Probability 8 (2) (1976) 385–394.

[2] D. Daley, D. Vere-Jones, An Introduction to the Theory of Point Processes, Vol. I, 2nd Edition, Springer, New York, 2003.

[3] R. Boel, P. Varaiya, Optimal control of jump processes, SIAM J. Control and Optimization 15 (1) (1977) 92–119.

[4] P. Brémaud, Point Processes and Queues: Martingale Dynamics, Springer, New York, NY, 1981.

[5] H. Kushner, P. Dupuis, Numerical Methods for Stochastic Control Problems in Continuous Time, Applications of Mathematics, Springer-Verlag, 1992.

[6] G. Gallego, G. van Ryzin, Optimal dynamic pricing of inventories with stochastic demand over finite horizons, Management Sci. (1994) 999–1020.

[7] H. Van Ooijen, J. Bertrand, The effects of a simple arrival rate control policy on throughput and work-in-process in production systems with workload dependent processing rates, International Journal of Production Economics 85 (1) (2003) 61–68.

[8] Y. Wang, G. Williams, E. Theodorou, L. Song, Variational policy for guiding point processes, in: Proc. 34th Internat. Conf. Machine Learning, Vol. 70 of Proc. Machine Learning Research, Sydney, Australia, 2017, pp. 3684–3693.

[9] D. A. Iancu, M. Sharma, Supermodularity and affine policies in dynamic robust optimization, Operations Research 61 (4) (2013) 941–956.

[10] H. Lin, P. Antsaklis, Stability and stabilizability of switched linear systems: A survey of recent results, IEEE Trans. Automatic Control 54 (2) (2009) 308–322.