AIMS *Mathematics*

Research article

(s, S) **Inventory policies for stochastic controlled system of Lindley-type with lost-sales**

Rubén Blancas-Rivera*, **Hugo Cruz-Suárez, Gustavo Portillo-Ramírez and Ruy López-Ríos**

Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, Av. San Claudio, San Manuel, Ciudad Universitaria, 72570, Puebla, México

* **Correspondence:** Email: ruben.blancasr@correo.buap.mx.

Abstract: This paper presents a characterization of (s, S)-inventory policies for Lindley systems with possibly unbounded costs, where the objective is to minimize the expected discounted total cost by ordering (production) strategies. Moreover, the existence of a subsequence of minimizers of the value iteration functions that converge to a (s, S) optimal inventory system policy is shown. A numerical example is given to illustrate the theory.

Keywords: Markov control processes; Lindley random walk; inventory control system; (s, S) policies
Mathematics Subject Classification: 49J53, 49K99

1. Introduction

This paper considers a discrete-time, single-item, infinite-stock inventory system in which excess demand is not backlogged. The dynamics of the system is determined by a random walk with a lower barrier; in particular, the state zero is considered as a barrier. Also, it is assumed that the control variable is affected by a random noise in the dynamic of the system; see (2.1). Furthermore, the following costs are considered: ordering/production, holding, and shortage. Then, the objective function is the infinite-horizon expected total discounted cost. The goal of the paper is to guarantee the existence of stationary policies and characterize the optimal stationary policies as (s, S) policies (see [5, 18, 23, 28]). The (s, S)-policies operate in the following manner: If the inventory level falls below a minimum $s \geq 0$, the controller will request a replenishment demand to restore the inventory stock to a maximum number $S \geq s$.

As mentioned earlier, a crucial component of the inventory model is the dynamics of the system; in this paper is considered a controlled version of a Lindley random walk. This random walk has been used to model waiting times in a queuing system with a server (see [20]). A controlled version of this dynamic was introduced in [14] to illustrate the convergence of the value iteration algorithm

for Markov control processes with average costs. Furthermore, this dynamic has been used in several contexts of Markov control processes; see, e.g., [16, 27]. Despite this, in these works, the optimal policy is not characterized, except by [11], where it is considered as a compact action set and bounded cost. The characterization of (s, S) policies is an interesting problem in inventory theory, this problem can be traced back to the papers of [13, 22, 28]. One advantage of guaranteeing the existence of (s, S) policies is their ease of implementation. However, it is not always possible to guarantee the existence of a (s, S) inventory policy; see, for instance, [2]. Therefore, it is necessary to establish conditions in the stochastic control model that guarantee the existence of (s, S) inventory policies. In addition, there is recent work in applied areas where strategies (s, S) are implemented in the real world, such as in healthcare [1, 29] and in machine learning [8], to name a few. A survey work presenting applications of inventory theory in the real world can be found in [22]. On the other hand, recent research from the theoretical point of view of the model is presented in [6], where the optimality of strategies (s, S) is proved as a particular case of their setup concave linear piecewise ordering costs. The perishable inventory with backlogging demand was studied in [30]. In contrast to the work described in this manuscript, there is an extensive literature on continuous-time inventories, e.g., the recent work of [4], which studies perishable inventories.

Based on the review conducted, the main objectives of this work can be described as follows:

- To provide conditions that guarantee the optimality of inventory strategies for a system with lost sales (see Theorem 4.10).
- To propose approximation procedures for the strategies described in the previous point (see Theorem 5.3).

It is important to note that the goal of the first point is a more complex challenge, as opposed to a backlog inventory system as described in [22].

The methodology of the paper is as follows. First, the dynamic programming approach is validated, and the existence of optimal stationary policies is verified. Second, results of convex analysis are applied to guarantee that the minimizers of the value iteration functions and the optimal policies of the inventory system are (s, S)-policies. This characterization is achieved under assumptions of continuity and monotonicity in the components of the inventory control model; these conditions guarantee that the cost function is convex. The convexity property has been used previously in the inventory systems literature, see for example [9, 10, 13, 18, 23, 25, 26, 28, 31].

The paper is organized as follows. Section 2 introduces the inventory control model and the assumptions about the components of the inventory control model. Section 3 introduces the dynamic programming approach. Section 4 presents the characterization of (s, S) policies. Section 5 contains an analysis of the convergence of minimizers of value iteration. Finally, a numerical example is given in Section 6.

2. An inventory control model with a Lindley dynamic system

Consider a discrete-time inventory control system. If X_t denotes the inventory at time $t = 0, 1, \ldots$, the evolution of the system is modeled by the following Lindley-type dynamical system:

$$X_{t+1} = (X_t + \eta_t a_t - D_{t+1})^+, \tag{2.1}$$

with $X_0 = x \in X := [0, \infty)$ known, where

i) $\{\eta_t\}$ is a sequence of independent and identically distributed (*i.i.d.*) Bernoulli random variables with parameter p, $0 < p < 1$, in this case the event $\{\eta_t = 1\}$ means that an order placed at instant t has been supplied.

ii) a_t denotes the control or decision applied at time t and represents the quantity ordered by the inventory manager (or decision maker).

iii) The sequence $\{D_t\}$ is conformed by *i.i.d.* non-negative random variables with a common distribution F. D_t denotes the demand within period t. It is assumed that sequence $\{D_t\}$ is independent of the sequence $\{\eta_t\}$.

Note that the difference equation given in (2.1) induces a stochastic kernel that can be expressed on X by

$$\mathbb{K} := \{(x, a) : x \in X, a \in [0, \infty)\},$$

is defined as follows

$$Q(X_{t+1} \in (-\infty, y] | X_t = x, a_t = a)$$
$$= p(F(x + a) - F(x + a - y)) + p(1 - F(x + a)) + (1 - p)(F(x) - F(x - y)) + (1 - p)(1 - F(x)))$$
$$\tag{2.2}$$

with $y, x, a \in [0, \infty)$ and

$$Q(X_{t+1} \in (-\infty, y] | X_t = x, a_t = a) = 0,$$

if $x, a \in [0, \infty)$ and $y < 0$.

Suppose further that it is associated with a one-step cost function $C : \mathbb{K} \longrightarrow [0, \infty)$, defined as follows:

$$C(x, a) = KI_{\{a:a>0\}}(a) + ca + \mathbb{E}[h((x + \eta a - D)^+)] + \mathbb{E}[l((D - (x + \eta a))^+)], \tag{2.3}$$

where $K \geq 0$ is a fixed order price, $c > 0$ is the order price per unit, $h : [0, \infty) \longrightarrow [0, \infty)$ denotes the holding cost per period, $l : [0, \infty) \longrightarrow [0, \infty)$ indicates the shortage cost for unfilled demand and $\mathbb{E}$ denotes the expectation with respect to the joint distribution of the random vector (η, D), where (η, D) is a generic element of the sequence $\{(\eta_t, D_t)\}$. In all the following sections, the cost function is determined by the sum of two cost functions:

$$C(x, a) = g(a) + H(x, a),$$

where $g(a) = KI_{\{a:a>0\}}(a) + ca$ and

$$H(x, a) = \mathbb{E}[h((x + \eta a - D)^+)] + \mathbb{E}[l((D - (x + \eta a))^+)], \tag{2.4}$$

with $(x, a) \in \mathbb{K}$.

Then, the inventory control model is identified with a Markov control process [17], in short, MCP. The components of the associate Markov control model are as follows: $X := [0, \infty)$ is the state space, $A := [0, \infty)$ is the action space, the dynamic and the cost function are given by (2.2) and (2.3), respectively. Consequently, the inventory system evolves as follows: If the stock of the inventory systems occupies state $X_t = x$ at time t and a controller (inventory manager) requests the quantity of product $a_t = a$. Then, a cost $C(x, a)$ is incurred and the system jumps into a state X_{t+1} according to

the transition law $Q(\cdot|x, a)$. Once the transition into the new state has occurred, a new order (control) is requested and the process is repeated. Thus, for each $t \geq 1$ an admissible history h_t of the inventory system up to the transition t given by, $h_t = (X_0, a_0, \ldots, X_{t-1}, a_{t-1}, X_t)$. Let $\mathbb{H}_t$, $t \geq 1$, be the set of all admissible histories of the system up to the transition t. The following definition will be used to characterize optimal strategies in the inventory control model.

Definition: A *policy* $\pi = \{\pi_t\}$ is a sequence of stochastic kernels π_t on A given $\mathbb{H}_t$, satisfying the constraint: $\pi_t(A|h_t) = 1$, for each $h_t \in \mathbb{H}_t$, $t \geq 1$. The collection of all policies is denoted by Π. Define $\mathbb{F}$ as the set of all measurable functions $f : X \longrightarrow A$. Thus, a *Markov policy* is a sequence f_t such that $f_t \in \mathbb{F}$, for $t \geq 1$. In particular, a Markov policy $\pi = \{f_t\}$ is said to be *stationary* if f_t is independent of t, i.e., $f_t = f \in \mathbb{F}$, for all $t \geq 1$, in this case, f_t is denoted by f and refers to $\mathbb{F}$ as the set of stationary policies.

In the subsequent, for each $x \in X$ and $\pi \in \Pi$ will be denoted by $\mathbb{P}_x^\pi$ the measure defined on the measurable space $\Omega := ((X \times A)^\infty, \mathcal{F})$, where $\mathcal{F}$ is the corresponding product σ-algebra. The measure $\mathbb{P}_x^\pi$ is induced by the theorem of Ionescu Tulcea [3]. The expectation operator with respect to $\mathbb{P}_x^\pi$ is denoted by $\mathbb{E}_x^\pi$.

Inventory control problem

The goal of this subsection is to introduce the inventory control problem, then consider $\pi \in \Pi$, $x \in X$ and define the following objective function:

$$v(\pi, x) := \mathbb{E}_x^\pi \left[\sum_{t=0}^\infty \alpha^t C(X_t, a_t) \right], \tag{2.5}$$

where $\alpha \in (0, 1)$ denotes a discount factor. The performance criterion defined in (2.5) is called *expected total discounted cost* (see [17]). Hence, the optimal inventory control problem consists in determining a policy $\pi^* \in \Pi$ such that,

$$v(\pi^*, x) = \inf_{\pi \in \Pi} v(\pi, x), \tag{2.6}$$

for each $x \in X$, in this case π^* is denominated *optimal policy*. The function V defined by

$$V(x) := \inf_{\pi \in \Pi} v(\pi, x), \tag{2.7}$$

for each $x \in X$, it will be called the *optimal value function*.

In subsequent sections, the following assumption will be considered.

Basic Assumption (BA):

i) h is a non-decreasing convex function such that $h(0_+) = h(0)$, where $h(0_+) := \inf\{h(y) : y > 0\}$.

ii) l is a convex function such that $l(0) = 0$ and $l'(u)$ exists and is non-negative for each $u \geq 0$.

iii) $\mathbb{E}[h((u - D)^+)] < \infty$ and $\mathbb{E}[l((D - u)^+)] < \infty$, for each $u \geq 0$ and D is a generic element of the sequence $\{D_t\}$.

iv) If D is a generic element of the sequence $\{D_t\}$, it is assumed that D has a continuous density denoted by Δ.

BA will not be mentioned in each lemma or theorem throughout the paper, and we will assume that it holds.

Remark 2.1. *i) Observe that the cost function (2.4) is convex, due to the convexity and monotonicity of functions h and l (see BA i) and ii)). Therefore, the cost function (2.3) is a convex function, this property will be used in the proof of Lemma 4.3.*

ii) Moreover, under BA i) and ii), h and l are continuous functions. This characterization will help us apply to demonstrate the validity of Lemma 3.1.

3. Dynamic programming approach

Assuming BA, the inventory control model in the MCP literature is identified as a semi-continuous model [17]. Moreover, the following lemmas guarantee the existence of optimal stationary policies and the dynamic programming approach; see Theorem 3.7 below.

Lemma 3.1. *The cost function C, is a lower semi-continuous (l.s.c.) function, i.e., the set $\{(x, a) \in \mathbb{K} : C(x, a) \leq \lambda\}$ is closed for all $\lambda \in \mathbb{R}$.*

Proof. First, we can see directly that the cost g is a continuous function for $a > 0$ and *l.s.c.* at $a = 0$. We now want to prove that the function H is continuous for any $(x, a) \in \mathbb{K}$. For this purpose, we fix $x \in X$ and $a \in A$. Take $x_n \in X$, $a_n \in A$, $n \geq 1$, such that $x_n \longrightarrow x$, $a_n \longrightarrow a$, when n goes to infinity. Consequently, Eq (2.4) gives $H(x_n, a_n) \longrightarrow H(x, a)$ as $n \longrightarrow \infty$, by applying the dominated convergence theorem [3] and continuity of the functions h and l; the latter by Remark 2.1. Therefore, the cost function C is a *l.s.c.*. $\square$

Corollary 3.2. *The cost C is an inf-compact function, namely the set $\{a \in A : C(x, a) \leq \lambda\}$ is compact for each $x \in X$ and $\lambda \in \mathbb{R}$.*

Proof. Note that $\{a \in A : ca \leq \lambda\} = [0, \lambda/c]$, for any $\lambda \in \mathbb{R}$. On the other hand, by Lemma 3.1, for all $\lambda \in \mathbb{R}$ and $x \in X$ the set defined as $M := \{a \in A : C(x, a) \leq \lambda\}$ is closed set. Consequently, since M is contained in the compact set $\{a \in A : ca \leq \lambda\}$, it is concluded that the cost function C is inf-compact. $\square$

Lemma 3.3. *Transition law Q (2.2) induced by the difference equation (2.1) is strongly continuous, i.e., $w(x, a) := \int_X u(y)Q(dy|x, a)$ is continuous and bounded on $\mathbb{K}$ for every measurable bounded function u on X.*

Proof. Let u be a measurable and bounded function defined on X and let $\{(x_n, a_n)\}$ be a sequence on $\mathbb{K}$ convergent to $(x, a) \in \mathbb{K}$. Then, from (2.2), we have that

$$
\begin{aligned}
w(x_n, a_n) &= \int_X u(y)Q(dy|x_n, a_n) \\
&= p \int_0^\infty u((x_n + a_n - s)^+)\Delta(s)ds + (1-p)\int_0^\infty u((x_n - s)^+)\Delta(s)ds \\
&= p \int_0^{x_n+a_n} u(x_n + a_n - s)\Delta(s)ds + (1-p)\int_0^{x_n} u(x_n - s)\Delta(s)d(s) \\
&\quad + (1-p)u(0)\int_{x_n}^\infty \Delta(s)ds \\
&\quad + pu(0)\int_{x_n+a_n}^\infty \Delta(s)ds.
\end{aligned}
$$

Observe by applying an adequate change of variable, the following identity holds:

$$\int_0^{x_n+a_n} u(x_n + a_n - s)\Delta(s)ds = \int I_{(-\infty,x_n+a_n]}(s)u(s)\Delta(x_n + a_n - s)ds.$$

On the other hand, it is easy to prove that $\{I_{(-\infty,x_n+a_n]}\}$ converges to $\{I_{(-\infty,x+a]}\}$ almost everywhere (a.e.) with respect to the Lebesgue measure m on $\mathbb{R}$. Moreover, let $\theta \in (0, \infty)$ such that $|u(x)| \leq \theta$, for all $x \in X$ and consider the following functions:

$$r_n(s) = I_{(-\infty,x_n+a_n]}(s)u(s)\Delta(x_n + a_n - s),$$
$$r(s) = I_{(-\infty,x+a]}(s)u(s)\Delta(x + a - s),$$
$$g_n(s) = \theta I_{(-\infty,x_n+a_n]}(s)\Delta(x_n + a_n - s),$$
$$g(s) = \theta I_{(-\infty,x+a]}(s)\Delta(x + a - s),$$

with $s \in [0, \infty)$. Then, notice that the following statements hold:

i) $\{r_n\}$ converges to r a.e. with respect to m.
ii) $\{g_n\}$ converges to g a.e. with respect to m.
iii) For each $n = 1, 2, ...$, $|r_n(\cdot)| \leq g_n(\cdot)$.
iv) $\int g_n(s)ds = \int g(s) = \theta$.

Statements i)–iv) guarantee the hypothesis of the Dominated Convergence Theorem (see [3]), in consequence, it is obtained that $\lim_{n\to+\infty} \int r_n(s)ds = \int r(s)ds$, i.e.,

$$\int_0^{x_n+a_n} u(x_n + a_n - s)\Delta(s)ds \longrightarrow \int_0^{x+a} u(x + a - s)\Delta(s)ds. \tag{3.1}$$

Similarly, it is possible to prove that:

$$\int_0^{x_n} u(x_n - s)\Delta(s)d(s) = \int I_{(-\infty,x_n]}(s)u(s)\Delta(x_n - s)ds$$
$$\longrightarrow \int I_{(-\infty,x]}(s)u(s)\Delta(x - s). \tag{3.2}$$

Besides, due to the continuity of the distribution function F, it yields that

$$\int_{x_n}^\infty \Delta(s)ds$$
$$= 1 - F(x_n)$$
$$\longrightarrow 1 - F(x) = \int_x^\infty \Delta(s)ds,$$
$$\int_{x_n+a_n}^\infty \Delta(s)ds = 1 - F(x_n + a_n)$$
$$\longrightarrow 1 - F(x + a) = \int_{x+a}^\infty \Delta(s)ds. \tag{3.3}$$

In consequence, from (3.1)–(3.3), it is obtained that

$$
\lim_{n \to +\infty} w(x_n, a_n) = p \int_0^{x+a} u(s)\Delta(x + a - s)ds
$$
$$
+ (1 - p) \int_0^x u(s)\Delta(x - s)ds
$$
$$
+ (1 - p)u(0) \int_x^\infty \Delta(s)ds
$$
$$
+ pu(0) \int_{x+a}^\infty \Delta(s)ds
$$
$$
= w(x, a).
$$

Therefore, w is a continuous function on $\mathbb{K}$. □

The demonstration of Lemma 3.4 is based on a result of renewal processes used in classical theory of inventories (see, e.g., [4, 12]). For this purpose, define for each $x \in X$, the *renewal process*

$$
N(x) := sup\{t \geq 0 | W_t \leq x\}, \tag{3.4}
$$

where $W_0 := 0$ and $W_t := \sum_{j=1}^t D_j$. Observe that $\mathbb{E}[N(x)] < \infty$ for each $x \geq 0$, a proof of this fact can be consulted in [15]. Furthermore, consider the *residual lifetime* defined as

$$
R(x) := W_{N(x)+1} - x, x \in X. \tag{3.5}
$$

Lemma 3.4. *For all $x \in X$, $\mathbb{E}[l(R(x))] < \infty$.*

Proof. Let $x \in X$ be fixed. Observe that the tail of the distribution of the residual lifetime is given for $z \geq 0$ as follows

$$
\mathbb{P}(R(x) > z) = \sum_{n=1}^\infty \mathbb{P}(W_{n-1} \leq x, W_n > x + z)
$$
$$
= 1 - F(x + z) + \int_0^x (1 - F(x + z - s))dU(s),
$$

where $U := \mathbb{E}[N(x)]$ is the renewal function. On the other hand, from assumption BA i), it is obtained that

$$
\mathbb{E}[l(R(x))] = \int_0^{l(\infty)} \mathbb{P}(l(R(x) > z))dz
$$
$$
= \int_0^\infty \mathbb{P}(l(R(x)) > l(z)))dl(z)
$$
$$
= \int_0^\infty \mathbb{P}(R(x) > z)l'(z)dz,
$$

where $l(\infty) := lim_{u \to \infty} l(u)$. In consequence, by substituting the tail distribution of function R in the last equation, it yields that

$$
\mathbb{E}[l(R(x))] = \int_0^\infty l'(z)(1 - F(x + z))dz + \int_0^\infty \left(\int_0^x l'(z)(1 - F(x + z - s))dU(s) \right) dz.
$$

Now, observe that

$$\int_0^\infty l'(z)(1 - F(x + z))dz \le \int_0^\infty l'(z)P(D > z)dz = \mathbb{E}[l(D)],$$

and from Fubini's theorem [3], it is obtained that

$$\int_0^\infty \int_0^x l'(z)(1 - F(x + z - s))dU(s)dz$$
$$\le \int_0^x \left(\int_0^\infty P(D > z)l'(z)dz \right) dU(s)$$
$$= \int_0^x \mathbb{E}[l(D)]dU(s)$$
$$= \mathbb{E}[l(D)]\mathbb{E}[N(x)].$$

Therefore,

$$\mathbb{E}[l(R(x))] \le \mathbb{E}[l(D)](1 + \mathbb{E}[N(x)]) < \infty.$$

Since state x is arbitrary, the result follows. $\square$

Lemma 3.5. *There exists $\tilde{\pi} \in \Pi$ such that $v(\tilde{\pi}, x) < \infty$, for all $x \in X$.*

Proof. Consider $x \in X$ fixed and the stationary policy $\tilde{\pi} = \{\tilde{g}, \tilde{g}, \ldots\} \subseteq \mathbb{F}$ with $\tilde{g}(y) = 0$ for all $y \in X$. Hence, the stochastic path $\{X_t^{\tilde{g}}\}$ generated by $\tilde{\pi}$ is given by

$$X_{t+1}^{\tilde{g}} = (X_t^{\tilde{g}} - D_{t+1})^+, \tag{3.6}$$

with $X_0^{\tilde{g}} = x$. Consequently, observe that

$$X_t^{\tilde{g}} = \begin{cases} x - W_t & if \quad t = 0, \ldots, N(x), \\ 0 & if \quad t \ge N(x) + 1. \end{cases} \tag{3.7}$$

Then, due to (2.3) and (2.4), it yields that

$$v(\tilde{\pi}, x) = \mathbb{E}_x^{\tilde{\pi}}\left[\sum_{t=0}^\infty \alpha^t H(X_t^{\tilde{g}}, 0) \right]$$
$$= \mathbb{E}_x^{\tilde{\pi}}\left[\sum_{t=0}^\infty \alpha^t \left(\mathbb{E}[h((X_t^{\tilde{g}} - D_{t+1})^+)] + \mathbb{E}[l((D_{t+1} - X_t^{\tilde{g}})^+)] \right) \right].$$

Now, from (3.6) and the renewal process (3.4), it may be found that

$$v(\tilde{\pi}, x) \le \frac{\mathbb{E}[l(D)] + h(0)}{1 - \alpha} + \mathbb{E}_x^{\tilde{\pi}}\left[\sum_{t=0}^{N(x)} \alpha^t (\mathbb{E}[h((X_t^{\tilde{g}} - D_{t+1})^+)] + \mathbb{E}[l((D_{t+1} - X_t^{\tilde{g}})^+)]) \right].$$

On the other hand, observe that

$$\sum_{t=0}^{N(x)} \alpha^t (\mathbb{E}[h((X_t^{\tilde{g}} - D_{t+1})^+)] + \mathbb{E}[l((D_{t+1} - X_t^{\tilde{g}})^+)])$$

$$= \sum_{t=0}^{N(x)} \alpha^t (\mathbb{E}[h((x - W_{t+1})^+)] + \mathbb{E}[l((W_{t+1} - x)^+)])$$

$$\leq \frac{1}{\alpha} \sum_{t=1}^{N(x)} \alpha^t (\mathbb{E}[h((x - W_t)^+)] + \mathbb{E}[l((W_t - x)^+)]) + h(0) + \mathbb{E}[l(W_{N(x)+1} - x)]$$

$$\leq \frac{1}{1 - \alpha}(h(x) + l(0)) + h(0) + \mathbb{E}[l(W_{N(x)+1} - x)], \quad \mathbb{P}_x^{\tilde{\pi}} - a.s,$$

where the second inequality was obtained since h is a non-decreasing function (see BA i)). Finally, both inequalities follow from (3.4). In consequence, since $l(0) = 0$, it is obtained that

$$v(\tilde{\pi}, x) \leq \frac{\mathbb{E}[l(D)] + h(x) + h(0)}{1 - \alpha} + \mathbb{E}[l(W_{N(x)+1} - x)].$$

By applying Lemma 3.4 to the last relation and due to $\mathbb{E}[l(D)] < \infty$ as a consequence of BA iii), it is obtained that $v(\tilde{\pi}, x) < \infty$. Since $x \in X$ is arbitrary, the result holds. $\qquad\square$

Remark 3.6. *Note that Lemma 3.5 guarantees that $V(x) < \infty$, for all $x \in X$, due to $V(x) \leq v(\tilde{\pi}, x)$, for all $x \in X$.*

As a consequence of the previous results, the following theorem is valid (see Theorem 4.2.3 and Lemma 4.2.8 in [17]).

Theorem 3.7. *The following statements hold:*

i) For each $x \in X$, the optimal value function satisfies the dynamic programming equation:

$$V(x) = \min_{a \in A} \{C(x, a) + \alpha \mathbb{E}[V((x + \eta a - D)^+)]\}.$$

ii) There exists an optimal stationary policy $f \in \mathbb{F}$ such that for each $x \in X$ the following equation holds

$$V(x) = C(x, f(x)) + \alpha \mathbb{E}[V((x + \eta f(x) - D)^+)].$$

iii) The value iteration functions, defined as $V_0(x) := 0$ and

$$V_n(x) := \min_{a \in A} \{C(x, a) + \alpha \mathbb{E}[V_{n-1}((x + \eta a - D)^+)]\},$$

$n \geq 1, x \in X$, converge monotonically increasing to the optimal value function V.

Remark 3.8. *Minimizers of the value iteration functions will be denoted by f_n, $n \geq 0$. These minimizers satisfy the following, $f_0(x) = 0$ and for each $n \geq 1$: $V_n(x) = C(x, f_n(x)) + \alpha \mathbb{E}[V_{n-1}((x + \eta f_n(x) - D)^+)]$, for all $x \in X$.*

4. Characterization of (s, S) policies

This section deals with the characterization of the minimizers of the value iteration functions and the optimal policies by (s, S) policies.

Definition 4.1. *Let $f \in \mathbb{F}$ be a stationary policy, if there exists $(s, S) \in \mathbb{R}^2$ such that $0 \le s \le S$ and*

$$f(x) = \begin{cases} S - x & if \quad x \le s, \\ \\ 0 & if \quad x > s, \end{cases} \tag{4.1}$$

f is called a (s, S) stationary policy.

Define the following functions

$$\hat{V}_n(u) := \mathbb{E}[V_n((u - D)^+)], \quad n = 0, 1, \ldots, \tag{4.2}$$

and

$$G_n(u) := cu + p\hat{H}(u) + \alpha p \hat{V}_{n-1}(u), \quad n = 1, 2, \ldots, \tag{4.3}$$

for $u \in X$, where $\hat{H}(u) = H(u, 0)$. In consequence, value iteration functions $\{V_n\}$ can be expressed for each $n \ge 1$ and $x \in X$ as follows,

$$\begin{aligned} V_n(x) &= \min\{cx + p\hat{H}(x) + \alpha p \hat{V}_{n-1}(x), \inf_{a>0}\{K + c(x + a) + p\hat{H}(x + a) + \alpha p \hat{V}_{n-1}(x + a)\}\} \\ &\quad - cx + (1 - p)\hat{H}(x) + \alpha(1 - p)\hat{V}_{n-1}(x) \\ &= \min\{G_n(x), \inf_{a>0}\{K + G_n(x + a)\}\} - cx + (1 - p)\hat{H}(x) + \alpha(1 - p)\hat{V}_{n-1}(x). \end{aligned}$$

Making the change of variable $y := x + a$, the previous equation is equivalent to

$$V_n(x) = \min\{G_n(x), \inf_{y \ge x}\{K + G_n(y)\}\} - cx + (1 - p)\hat{H}(x) + \alpha(1 - p)\hat{V}_{n-1}(x), \tag{4.4}$$

for $n \ge 1$ and $V_0(x) = 0$, $x \in X$.

The following definition will be applied to characterize the optimal policy [21].

Definition 4.2. *A function $\vartheta : [0, \infty) \longrightarrow [0, \infty)$ is called* norm-like *if $\vartheta(x) \to \infty$ as $x \to \infty$; this means that the sub-level sets $\{x : \vartheta(x) \le r\}$ are precompact for each $r > 0$.*

Lemma 4.3. *For each $n = 1, 2, \ldots$, the following statements hold*

  i) V_n is a non-decreasing convex function,
 ii) G_n is a convex function on X,
iii) G_n is norm-like.

Proof. i) The proof is by induction. It will be proved that V_1 is a non-decreasing function,

$$\min_{a \in [0,\infty)} C(x, a) \le \min_{a \in [0,\infty)} C(y, a),$$

if $x, y \in X$ and $x \le y$, due to $C(\cdot, a)$ is a non-decreasing function for each $a \in [0, \infty)]$ (see (2.3) and AIM). Then, V_1 is a non-decreasing function. On the other hand, using Lemma 1 in [19] together with

the fact that C is a convex function, it is obtained that V_1 is a convex function. Now, suppose that V_n is a non-decreasing convex function. To prove that V_{n+1} is a non-decreasing function, observe that, as a consequence of V_n is non-decreasing, the following inequality holds

$$V_n((x + \eta a - s)^+) \le V_n((y + \eta a - s)^+),$$

for all $x, y \in X$, with $x \le y$ and $a, s \in [0, \infty)$. In consequence,

$$C(x, a) + \alpha \mathbb{E}[V_n((x + \eta a - D)^+)] \le C(y, a) + \alpha \mathbb{E}[V_n((y + \eta a - D)^+)], \tag{4.5}$$

for all $x, y \in X$, with $x \le y$ and $a, s \in [0, \infty)$. Consequently, taking the minimum with respect to a, in each side of inequality (4.5), it results that $V_{n+1}(x) \le V_{n+1}(y)$, $x, y \in X$ with $x \le y$. Now, it will be proved that V_{n+1} is a convex function. To this end, note that $C(x, a) + \alpha \mathbb{E}[V_n((x + \eta a - D)^+)]$ is a convex function for each $(x, a) \in \mathbb{K}$, this statement is a consequence of the induction hypothesis and the convexity of the cost function. Then, from Lemma 1 of [19], it may be concluded that V_{n+1} is a convex function.

ii) The previous statement implies that $\hat{V}_n$ is a convex function for each $n \ge 1$, as the following relations evidence it:

$$\begin{aligned}
\hat{V}_n(\lambda u_1 + (1 - \lambda)u_2) &= \mathbb{E}\left[V_n((\lambda u_1 + (1 - \lambda)u_2 - D)^+)\right] \\
&\le \mathbb{E}\left[\lambda V_n((u_1 - D)^+) + (1 - \lambda)V_n((u_2 - D)^+)\right] \\
&= \lambda \hat{V}_n(u_1) + (1 - \lambda)\hat{V}_n(u_2),
\end{aligned}$$

for each $u_1, u_2 \in X$, $\lambda \in [0, 1]$. Consequently, since $\hat{H}$ is convex, (4.3) implies that G_n is a convex function for each $n \ge 1$.

iii) Observe that $G_n(u) \ge cu$, for $n \ge 1$ with $c > 0$. Hence, for each $n \ge 0$, $G_n(u) \to \infty$, when $u \to \infty$. This implies that G_n is a norm-like function for each $n \ge 1$. $\qquad \square$

Remark 4.4. *Observe that for each $n \ge 1$, we have that*
i) G_n (4.3) is a continuous function on $(0, \infty)$, as a consequence of convexity of function G_n.
ii) Given that G_n is a norm-like function it follows that

$$\arg \min_{y \in X} G_n(y) := \{z \ge 0 : G_n(z) = \min_{y \in X} G_n(y)\},$$

is a non-empty set.

Lemma 4.5 is a modified version of Lemma 2.1 of [7], the proof of this lemma is presented here for the completeness of the paper.

Lemma 4.5. *For each $n \ge 1$, let $S_n \in \arg \min_{y \ge 0} G_n(y)$ and $s_n := \inf\{0 \le x \le S_n : G_n(x) \le K + G_n(S_n)\}$ then the following statements hold:*
i) $G_n(s_n) = K + G(S_n)$, if $s_n > 0$.
ii) $G_n(x) \le K + G_n(S_n)$, $0 \le s_n \le x \le S_n$.
iii) $K + G_n(S_n) \le G_n(x)$, $0 \le x < s_n$.
iv) $G_n(x)$ is a decreasing function on $[0, s_n)$.

Proof. Let $n \geq 1$. *i)* Suppose that $s_n > 0$ then $G_n(0) > K + G_n(S_n) \geq G_n(s_n)$. Hence, there exists $u \in (0, s_n)$ such that $G_n(u) = G_n(S_n) + K$, because of the continuity of function G_n (see Remark 4.4). Now, observe that $s_n \leq u$, due to the definition of s_n, then $s_n = u$ and the result holds.

ii) This statement follows directly from the definition of s_n.

iii) Consider $0 \leq x < s_n$, by Lemma 4.3 *ii)*, it follows that

$$G_n(s_n) \leq \frac{s_n - x}{S_n - x} G_n(S_n) + \frac{S_n - s_n}{S_n - x} G_n(x),$$

then

$$G_n(s_n) + \frac{S_n - s_n}{s_n - x}(G_n(s_n) - G_n(x))$$
$$= \frac{S_n - s_n + s_n - x}{s_n - x} G_n(s_n) - \frac{S_n - s_n}{s_n - x} G_n(x)$$
$$\leq G_n(S_n)$$
$$\leq K + G_n(S_n).$$

Thus,

$$G_n(s_n) + \frac{S_n - s_n}{s_n - x}(G_n(s_n) - G_n(x)) \leq K + G_n(S_n). \tag{4.6}$$

Now, due to $K + G_n(S_n) - G_n(s_n) = 0$ given that and $x \leq s_n \leq S_n$, (4.6) implies that $G_n(s_n) \leq G_n(x)$. Therefore, $K + G_n(S_n) = G_n(s_n) \leq G_n(x)$ with $0 \leq x < s_n$.

iv) Let $x_1, x_2 \in [0, s_n)$ with $x_1 \leq x_2$. Hence, by Lemma 4.3 *ii)*,

$$K + G_n(S_n) \geq G_n(x_2) + \frac{S_n - x_2}{x_2 - x_1}(G_n(x_2) - G_n(x_1)). \tag{4.7}$$

Now, from statement *iii)*, it is obtained that $G_n(x_2) \geq K + G_n(S_n)$, this relation and (4.7) together lead to $0 \geq G_n(x_2) - G_n(x_1)$ and the result follows. $\qquad\square$

A consequence of Lemma 4.5 is the following result.

Theorem 4.6. *Consider $\{(s_n, S_n) : n = 1, 2, \ldots\}$ in Lemma 4.5. Then, the minimizers of the value iteration functions are given by*

$$f_n(x) = \begin{cases} S_n - x & if \quad x \leq s_n \\ 0 & if \quad x > s_n, \end{cases} \tag{4.8}$$

with $0 < s_n \leq S_n$ and $n = 1, 2, \ldots$

Proof. Let $n \geq 1$ and $x \in X$. The proof will proceed by considering three cases depending on the order relation between state x, s_n, and S_n. The following simple but important claim is established.

Claim 1: $u \in \arg\min_{y \geq x} \{K + G_n(y)\}$ if and only if $u - x \in \arg\min_{a \geq 0} \{K + G_n(x + a)\}$.

Case 1. Suppose that $x < s_n$, Lemma 4.5 *iii)* yields that $K + G_n(S_n) \leq G_n(x)$. Consequently, by (4.4) and Claim, it follows that $f_n(x) = S_n - x$.

Case 2. Consider $s_n \leq x \leq S_n$, then by Lemma 4.5 *ii*), it is obtained that $G_n(x) \leq K + G_n(S_n)$. In consequence, (4.4) implies that $f_n(x) = 0$.

Case 3. Finally, assume that $S_n < x$, since G_n is a convex function, we have that G_n is a non-decreasing function on (S_n, ∞), this fact implies that $x \in \arg\min_{y \geq x} \{K + G_n(y)\}$. Now, from Claim 1, it follows that $0 \in \arg\min_{a \geq 0} \{K + G_n(x + a)\}$. Therefore, $f_n(x) = 0$.

The previous cases guarantee the truth of Theorem 4.6. □

Remark 4.7. *Observe that Theorem 3.7 and (4.4) imply that the optimal value function satisfies the following equation*

$$V(x) = \min\{G(x), \inf_{y \geq x} G(y)\} - cx + (1 - p)\hat{H}(x) + \alpha(1 - p)V^*(x), \tag{4.9}$$

with

$$G(u) := cu + p\hat{H}(u) + \alpha p V^*(u), \quad u \geq 0, \tag{4.10}$$

and

$$V^*(u) := E[V((u - \xi)^+)].$$

Since $V_n \uparrow V$ because of Theorem 3.7 and by the Dominated Convergence Theorem, it yields that

$$G_n \uparrow G. \tag{4.11}$$

Lemma 4.8. *G is a norm-like and convex function.*

Proof. Observe that $G(x) \geq cx$ for each $x \geq 0$. Hence, $G(x) \to \infty$ as x goes to infinity. Thus, by Definition 4.2, G is a norm-like function. Furthermore, (4.11) implies that G is a convex function since G_n is a convex function for each $n \geq 1$ (see Lemma 4.3). □

The proofs of Lemma 4.9 and Theorem 4.10 are in similar lines with the proofs of Lemma 4.5 and Theorem 4.6, respectively. Thus, the proofs will be omitted.

Lemma 4.9. *Let $S \in \arg\min_{y \geq 0} G(y)$ and $s := \inf\{0 \leq x \leq S : G(x) \leq K + G(S)\}$ then the following statements are valid.*
i) $G(s) = G(S) + K$, if $s > 0$.
ii) $G(x) \leq K + G(S)$, $0 \leq s \leq x \leq S$.
iii) $G(S) + K \leq G(x)$, $0 \leq x < s$.
iv) $G(x)$ is a decreasing function on $[0, s)$.

Theorem 4.10. *Consider (s, S) as in Lemma 4.9. Hence, the optimal policy for the inventory system is an (s, S) policy given by*

$$f(x) = \begin{cases} S - x & si \quad x \leq s, \\ \\ 0 & si \quad x > s, \end{cases}$$

with $0 < s \leq S$.

Remark 4.11. *Observe that in Theorem 4.10, if $s = 0$ then $G(x) \leq G(S) + K$, for all $x \geq 0$, which implies that $f(x) = 0$, for all $x \geq 0$. A similar argumentation guarantees the following assertion: $f_n(x) = 0$, $x \in X$, if $s_n = 0$ for $n \geq 1$.*

5. Convergence of minimizers of the value iteration functions

In this section, the convergence of the minimizers of the value iteration function will be analyzed. First, the next auxiliary results are exposed.

Lemma 5.1. *Let $\hat{C} \subset (0, \infty)$ be a compact set. Then, the following statements hold.*
i) $\{G_n\}$ converges uniformly to G on $\hat{C}$.
ii) For each $\{u_n\} \subset \hat{C}$ such that $u_n \to u \in \hat{C}$, $\lim_{n\to\infty} G_n(u_n) = G(u)$.

Proof. i) This statement is a direct consequence of (4.11) and Theorem 7.13 in [24].

ii) Applying *i)* of Lemma 5.1, it follows that for all $\epsilon > 0$ there exists $N_1 \geq 1$ such that for each $x \in \hat{C}$,

$$|G_n(x) - G(x)| < \epsilon/2, \tag{5.1}$$

if $n \geq N_1$. Furthermore, by continuity of function G on $\hat{C}$, there exists $\delta > 0$ such that

$$|G(y) - G(x)| < \epsilon/2, \tag{5.2}$$

if $|y - x| < \delta$. Now, since u_n converges to u when n goes to infinity, there exists $N_2 \geq 1$ such that

$$|u_n - u| < \delta, \tag{5.3}$$

if $n \geq N_2$. Let $N = \max\{N_1, N_2\}$. Thus, (5.2) and (5.3) together leads to

$$|G(u_n) - G(u)| < \epsilon/2, \tag{5.4}$$

if $n \geq N$. Then, taking $x = u_n$ in (5.1), it yields that

$$|G_n(u_n) - G(u_n)| < \epsilon/2. \tag{5.5}$$

Therefore, by (5.4) and (5.5), $|G_n(u_n) - G(u)| < \epsilon$. This last relation implies that $G_n(u_n)$ converges to $G(u)$, when n goes to infinity. $\square$

Lemma 5.2. *Let $B := \{x > 0 | G_1(x) \leq \inf_{y\geq 0} G(y) + K\}$ and $\{(s_n, S_n) : n \geq 1\}$ as in Lemma 4.5. Then, the following statements hold*
i) B is a compact set in $\mathbb{R}$.
ii) $s_n, S_n \in B$, for each $n = 1, 2, \ldots$.
iii) There exists a subsequence $\{(s_{n_k}, S_{n_k})\}$ of $\{(s_n, S_n)\}$, which converges to (s^*, S^*) such that $S^* \in \arg\min_{y\geq 0} G(y)$ and $G(s^*) = K + G(S^*)$ with $0 < s^* \leq S^*$.

Proof. i) First, observe that B is a closed set, since G_1 is a continuous function on $(0, \infty)$. Now suppose that B is not bounded, then there exists a sequence $\{b_n\} \subset B$ such that b_n converges to infinity. Thus,

$$\infty = \liminf_{n\to\infty} G_1(b_n) \leq \inf_{y\geq 0} G(y) + K.$$

The previous relation is a contradiction since $\inf_{y\geq 0} G(y) + K < \infty$. Therefore B is a compact set.

ii) Observe that for each $n = 1, 2, \ldots$ the following inequalities are valid

$$G_1(S_n) \leq G_n(S_n) = \inf_{y \geq 0} G_n(y) \leq \inf_{y \geq 0} G(y) + K,$$

$$G_1(s_n) \leq G_n(s_n) = \inf_{y \geq 0} G_n(y) + K \leq \inf_{y \geq 0} G(y) + K,$$

since $\{G_n\}$ is a non-decreasing sequence whose limit is G and the equalities are valid due to Lemma 4.5. Then, for each $n = 1, 2, \ldots, s_n, S_n \in B$.

iii) The previous statements *i*) and *ii*) imply that there exists a subsequence $\{(s_{n_k}, S_{n_k})\}$ convergent to $(s^*, S^*) \in B^2$, with $0 < s^* \leq S^*$. Furthermore, due to Lemma 5.1 *ii*), it yields that

$$\lim_{k \to \infty} G_{n_k}(S_{n_k}) = G(S^*), \tag{5.6}$$

$$\lim_{k \to \infty} G_{n_k}(s_{n_k}) = G(s^*). \tag{5.7}$$

Now, Lemma 4.5 implies that $G_{n_k}(S_{n_k}) \leq G_{n_k}(x)$, $x \geq 0$. Then, when k goes to infinity in the last inequality, it is obtained that $G(S^*) \leq G(x)$, $x \geq 0$. In consequence, $S^* \in \arg \min_{y \geq 0} G(y)$. On the other hand, as a consequence of Lemma 4.5 *ii*), (5.6) and (5.7), it follows that $G(s^*) = K + G(S^*)$. $\square$

Theorem 5.3. *Let $\{f_n\}$ be the sequence of minimizers of value iteration functions. Then, there exists a subsequence $\{f_{n_k}\}$ such that converges uniformly on X to an (s^*, S^*) optimal policy.*

Proof. First, in view of Theorem 4.6, for each $n \geq 1$, $f_n(x) = (S_n - x)I_{\{x : x \leq s_n\}}(x)$ with $0 < s_n \leq S_n$. Then, using *iii*) of Lemma 5.2, it is obtained a subsequence (s_{n_n}, S_{n_k}) such that $(s_{n_n}, S_{n_k}) \longrightarrow (s^*, S^*)$ when k goes to infinity and (s^*, S^*) satisfies the statements *i*) $-$ *iv*) of Lemma 4.9. Thus, consider

$$f^*(x) = \begin{cases} S^* - x & si \quad x \leq s^* \\ \\ 0 & si \quad x > s^*, \end{cases}$$

and observe that,

$$\sup_{x \in X} \left| f_{n_k}(x) - f^*(x) \right| \leq \left| S_{n_k} - S^* \right| \longrightarrow 0,$$

when $k \longrightarrow \infty$. Furthermore, f^* is an optimal policy due to Theorem 4.10. This concludes the proof of the theorem. $\square$

6. A numerical example

Consider an inventory control system with an exponential distribution on the demand with parameter $\lambda = 0.1$ and suppose that the parameter of the variable η is $p = 0.5$. On the other hand, for each $u \geq 0$, suppose that $h(u) = \gamma_1 u$, $l(u) = \gamma_2 u$, where γ_1, γ_2 are non negative constants. In this case, γ_1 represents the cost per unit shortage and γ_2 represents the cost per unit demand not supplied.

Furthermore, observe that for each $u \geq 0$,

$$\mathbb{E}[h((u - D)^+)] = \frac{\gamma_1}{\lambda} \left(\lambda u + e^{-\lambda u} - 1 \right),$$

$$\mathbb{E}[l((D-u)^+)] = \gamma_2\left(\frac{e^{-\lambda u}}{\lambda}\right).$$

Then by (4.3) for $n = 1$, it yields that

$$G_1(u) = (p\gamma_1 + c)u + p(\gamma_1 + \gamma_2)\frac{e^{-\lambda u}}{\lambda} - \frac{\gamma_1}{\lambda}, \quad u \geq 0.$$

Lemma 6.1. *For each $n = 1, 2, ..., G_n$ is a strictly convex function.*

Proof. First, observe that the derivative of G_1 is given by $G_1'(u) = p\gamma_1 + c - p(\gamma_1 + \gamma_2)e^{-\lambda u}$ and the second derivative $G_1''(u) = \lambda p(\gamma_1 + \gamma_2)e^{-\lambda u} > 0$, $u \geq 0$. Consequently, the function G_1 is strictly convex. Then, since $G_n(u) = G_1(u) + \alpha p\hat{V}_{n-1}(u)$, $u \geq 0$ (see 4.3), it is concluded that for each $n = 1, 2, ..., G_n$ is a strictly convex function, due to $\hat{V}_n$ is a convex function for each $n \geq 1$ (see proof of Lemma 4.3) and G_1 is a strictly convex function. Then, the result follows. □

Remark 6.2. *A consequence of Lemma 6.1 is the uniqueness of minimizers of the value iteration functions (see (4.4)). Furthermore, observe that (4.10) can be rewritten for each $u \geq 0$ as follows $G(u) = G_1(u) + \alpha pV^*(u)$, which is a strictly convex function, as a consequence of Lemma 6.1 and Remark 4.7. Hence, the optimal policy is unique due to the Eq (4.9).*

Lemma 6.3. *The unique minimizer of function G_1 is $y^* = -\lambda^{-1}(ln(\gamma_1 + c/p) - ln(\gamma_1 + \gamma_2) > 0$, if $c/p < \gamma_2$.*

Proof. Observe that, due to Lemma 6.1, the minimizer y^* is characterized by the first order condition:

$$p\gamma_1 + c - p(\gamma_1 + \gamma_2)e^{-\lambda y^*} = 0. \tag{6.1}$$

Thus, $y^* = -\lambda^{-1}ln((\gamma_1 + c/p)(\gamma_1 + \gamma_2)^{-1})$. In particular, since $c/p < \gamma_2$, it is obtained that $(\gamma_1 + c/p)(\gamma_1 + \gamma_2)^{-1} < 1$, in consequence, $y^* > 0$. □

The following algorithm is applied to approximate the optimal value function V and the optimal policy f, for each state x given as initial condition. The validity of the next algorithm is sustained in the following results: Lemma 4.9, Theorem 4.10, Lemmas 5.2 and 6.3.

Algorithm 1 Approximation of the optimal policy f and the optimal value function V

Require: $K > 0, c > 0, \lambda > 0, \gamma_1 > 0, \gamma_2 > c/p > 0, x > 0$ and $\epsilon, \alpha \in (0, 1)$
1: $S_1 \leftarrow y^*$
2: $G_1(u) \leftarrow (p\gamma_1 + c)u + p(\gamma_1 + \gamma_2)e^{-\lambda u}/\lambda$
3: $s_1 \leftarrow x^*$ such that $G_1(x^*) = G_1(S_1) + K$
4: **Repeat**
5:    **for** n **do**
6:      $\hat{V}_n(u) \leftarrow \mathbb{E}[V_n((u - D)^+)]$
7:      **if** $u \leq s_n$
8:        $V_n(u) \leftarrow G_n(S_n) + K - cu + (1 - p)\hat{H}(u) + \alpha(1 - p)\hat{V}_{n-1}(u)$
9:      **else**
10:       $V_n(u) \leftarrow G_n(u) - cu + (1 - p)\hat{H}(u) + \alpha(1 - p)\hat{V}_{n-1}(u)$

11:     $G_n(u) \leftarrow G_1(u) + \hat{V}_n(u)$

12:     $S_n \leftarrow \arg\min_{u \geq 0} G_n(u)$

13:     $s_n \leftarrow x^*$ such that $G_n(x^*) = G_n(S_n) + K$

14:     until $|S_{n-1} - S_n| < \epsilon$ and $|s_{n-1} - s_n| < \epsilon$

15: $S^* \leftarrow S_n$ and $s^* \leftarrow s_n$

16: $f(x) \leftarrow S^* - x$ **if** $x \geq s^*$ **else** $f(x) \leftarrow 0$

17: $V(x) \leftarrow V_n(x)$

18: **return** $f(x)$ and $V(x)$

Consider the following parameters: $\gamma_1 = \gamma_2 = 30$, $K = 1.5$, $c = 2.5$, $\epsilon = 0.01$, $\alpha = 0.2$ and $x = 40$, observe that $\gamma_2 > c/p$. Table 1 illustrates the numerical results obtained by applying the previous Algorithm 1.

n	s_n	S_n	$V_n(x)$	$f_n(x)$
1	49.79	53.89	1200.22	13.89
3	52.75	55.28	1438.65	14.80
5	55.72	55.75	1496.62	15.28
7	53.82	53.89	1497.04	15.75
8	53.82	53.89	1498.92	15.75

Table 1. The table displays convergence of value iteration function and their minimizers.

Observe that the convergence of the optimal value function and (s, S) policy are illustrated, see Theorem 4.10 and Lemma 5.2, respectively. Therefore, $f(x) = 15.75$ and $V(x) = 1498.92$ for $x = 40$ with an error $\epsilon = 0.01$.

7. Conclusions

In this paper, a discrete-time inventory system was presented. In the manuscript, conditions for convexity and monotonicity in the components of the Markov control model were proposed. In this inventory system, the existence of stationary policies (s, S) was proved. To achieve this goal, the methodology of dynamic programming in discrete time was applied. Moreover, the existence of a subsequence of minimizers of the value iteration functions converging to a (s, S) optimal policy of the inventory system was proved. Finally, a numerical algorithm for approximating the optimal inventory cost and policy is presented and applied in a numerical example. Future work in this direction includes the following:

- Incorporating Markovian demand into the inventory model.
- Investigating the lost-sales inventory system under other performance measures, e.g., considering the long average or risk-sensitive criteria.
- Implementing the manuscript's proposal on real-world data by finding a suitable database for the assumptions presented in this manuscript.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors would like to thank the referees for their valuable comments, corrections and suggestions, which lead to an improvement of the original paper.

Conflict of interest

The authors declare that they have no competing interests.

References

1. E. Ahmadi, H. Mosadegh, R. Maihami, I. Ghalehkhondabi, M. Sun, G. A. Süer, Intelligent inventory management approaches for perishable pharmaceutical products in a healthcare supply chain, *Comput. Oper. Res.*, **147** (2022), 105968. https://doi.org/10.1016/j.cor.2022.105968

2. Y. Aneja, A. Noori, The optimality of (s, S) policies for a stochastic inventory problem with proportional and lump-sum penalty cost, *J. Manag. Sci.*, **33** (1987), 750–755. https://doi.org/10.1287/mnsc.33.6.750

3. R. Ash, *Real analysis and probability*, New York: Academic Press, 1972. https://doi.org/10.1016/C2013-0-06164-6

4. Y. Barron, O. Baron, QMCD Approach for perishability models: The (S, s) control policy with lead time, *IISE Trans.*, **52** (2020), 133–150. https://doi.org/10.1080/24725854.2019.1614697

5. A. Bensoussan, *Dynamic programming and inventory control*, In: Studies in Probability, Amsterdam: IOS Press, **3** (2011). https://doi.org/10.3233/978-1-60750-770-3-i

6. A. Bensoussan, M. A. Helal, V. Ramakrishna, Optimal policies for inventory systems with piecewise-linear concave ordering costs, *Available at SSRN 3601262*, 2020. https://dx.doi.org/10.2139/ssrn.3601262

7. D. Bertsekas, *Dynamic programming and optimal control*, 2 Eds., Athena scientific, 2012.

8. B. Chen, X. Chao, C. Shi, Nonparametric learning algorithms for joint pricing and inventory control with lost sales and censored demand, *Math. Oper. Res.*, **46** (2021), 726–756. https://doi.org/10.1287/moor.2020.1084

9. D. Cruz-Suárez, R. Montes-de-Oca, Uniform convergence of value iteration policies for discounted Markov decision processes, *B. Soc. Mat. Mex.*, **12** (2006), 133–148.

10. D. Cruz-Suárez, R. Montes-de-Oca, F. Salem-Silva, Conditions for the uniqueness of optimal policies of discounted Markov decision processes, *Math. Method. Oper. Res.*, **60** (2004), 415–436. https://doi.org/10.1007/s001860400372

11. H. Daduna, P. Knopov, L. Tur, Optimal strategies for an inventory system with cost functions of general form, *Cybern. Syst. Anal.*, **35** (1999), 602–618. https://doi.org/10.1007/bf02835856

12. E. Feinberg, D. Kraemer, Continuity of discounted values and the structure of optimal policies for periodic-review inventory systems with setup costs, *Nav. Res. Log.*, 2023, 1–13. https://doi.org/10.1002/nav.22108

13. E. Feinberg, Optimality conditions for inventory control, *Optim. Chall. Complex Netw. Risk. Syst. INFORMS*, 2016, 14–45. https://doi.org/10.1287/educ.2016.0145

14. E. Gordienko, O. Hernández-Lerma, Average cost Markov control processes with weighted norms: Value iteration, *Appl. Math.*, **23** (1995), 219–237. https://doi.org/10.4064/am-23-2-219-237

15. A. Gut, *Stopped random walks*, Limit Theorems and Applications, 2 Eds., New York: Springer, 2009. https://doi.org/10.1007/978-0-387-87835-5

16. X. Guo, Q. Zhu, Average optimality for Markov decision processes in Borel spaces: A new condition and approach, *J. Appl. Probab.*, **43** (2006), 318–334. https://doi.org/10.1239/jap/1152413725

17. O. Hernández-Lerma, J. Lasserre, *Discrete-time Markov control processes: Basic optimality criteria*, New York: Springer Science & Business Media, 2012.

18. D. Iglehart, Optimality of (s, S) policies in the infinite horizon dynamic inventory problem, *Manag. Sci.*, **9** (1963), 259–267. https://doi.org/10.1287/mnsc.9.2.259

19. D. Iglehart, Capital accumulation and production for the firm: Optimal dynamic policies, *Manag. Sci.*, **12** (1965), 193–205. https://doi.org/10.1287/mnsc.12.3.193

20. D. Lindley, The theory of queues with a single server, *Math. Proc. Cambridge*, **48** (1952), 277–289. https://doi.org/10.1017/S0305004100027638

21. S. Meyn, R. Tweedie, *Markov chains and stochastic stability*, New York: Springer Science & Business Media, 1993. https://doi.org/10.1007/978-1-4471-3267-7

22. S. Perera, S. Sethi, A survey of stochastic inventory models with fixed costs Optimality of (s, S) and (s, S) type policies discrete-time case, *Prod. Oper. Manag.*, **32** (2022), 131–153. https://doi.org/10.1111/poms.13820

23. E. Porteus, On the optimality of generalized (s, S) policies, *Manag. Sci.*, **17** (1971), 411–426. https://doi.org/10.1287/mnsc.17.7.411

24. W. Rudin, *Principles of mathematical analysis (Vol. 3)*, New York: McGraw-hill (1976).

25. M. Schäl, On the optimality of (s, S)-policies in dynamic inventory models with finite horizon, *SIAM J. Appl. Math.*, **30** (1976), 528–537. https://doi.org/10.1137/0130048

26. S. Sethi, F. Feng, Optimality of (s, S) policies in inventory models with Markovian demand, *Oper. Res.*, **45** (1997), 931–939. https://doi.org/10.1287/opre.45.6.931

27. O. Vega, R. Montes-de-Oca, Application of average dynamic programming to inventory systems, *Math. Method. Oper. Res.*, **47** (1998), 451–471. https://doi.org/10.1007/bf01198405

28. A. Veinott, H. Wagner, Computing optimal (s, S) inventory policies, *Manag. Sci.*, **11** (1965), 525–552. https://doi.org/10.1287/mnsc.11.5.525

29. X. Xu, S. P. Sethi, S. H. Chung, Ordering COVID-19 vaccines for social welfare with information updating: Optimal dynamic order policies and vaccine selection in the digital age, *IISE*, 2023, 1–28. https://doi.org/10.1080/24725854.2023.2204329

30. H. Zhang, J. Zhang, R. Q. Zhang, Simple policies with provable bounds for managing perishable inventory, *Prod. Oper. Manag.*, **29** (2020), 2637–2650. https://doi.org/10.2307/3214683

31. Y. Zheng, A simple proof for optimality of (s, S) policies in infinite-horizon inventory systems, *J. Appl. Probab.*, **28** (1991), 802–810. https://doi.org/10.2307/3214683

AIMS Press