UNIFORM CONVERGENCE OF VALUE ITERATION POLICIES FOR DISCOUNTED MARKOV DECISION PROCESSES

DANIEL CRUZ-SUÁREZ AND RAÚL MONTES-DE-OCA

ABSTRACT. This paper deals with infinite horizon Markov Decision Processes (MDPs) on Borel spaces. The objective function considered, induced by a nonnegative and (possibly) unbounded cost, is the expected total discounted cost. For each of the MDPs analyzed, the existence of a unique optimal policy is assumed. Conditions that guarantee both pointwise and uniform convergence on compact sets of the minimizers of the value iteration algorithm to the optimal policy are provided. The theory developed in this paper is illustrated with three examples: an inventory/production system, the linear regulator problem, and a nonlinear additive-noise system with unbounded cost function.

1. Introduction

This paper is related to infinite horizon Markov decision processes (MDPs) on Borel spaces (see [5], [12], [13] and [16]).

A nonnegative and possibly unbounded cost function is considered. The expected total discounted cost is supposed to be the objective function. For such Markov decision process (MDP), the existence of a unique stationary optimal policy $f^*$ is assumed (see [8] for conditions to guarantee the uniqueness of optimal policies in discounted MDPs).

Consider an MDP with such description. Denote the state and the action spaces by $X$ and $A$, respectively. Let $A(x), x \in X$ be the admissible action sets, and let $Q$ represent the transition probability law. Take $V^*$ as the optimal value function, and for each $n = 1, 2, \cdots$, let $V_n$ and $f_n$ denote the minimum and the minimizer corresponding to the step $n$ of the value iteration algorithm (see [12]), respectively.

This paper deals mainly with establishing conditions that ensure the uniform convergence on compact sets of $\{f_n\}$ to $f^*$ (see Theorem (3.41) below).

Additionally, for the pointwise convergence of $\{f_n\}$ to $f^*$, conditions which are fewer and weaker than those presented in the paper for the uniform convergence on compact sets of $\{f_n\}$ to $f^*$ are provided (see Section 5 below).

The uniform convergence on compact sets of $\{f_n\}$ to $f^*$ guarantees, given a compact set $\varsigma \subset X$ and $\varepsilon > 0$, the existence of a positive integer $N(\varsigma, \varepsilon)$ such that $f_{N(\varsigma, \varepsilon)}(x)$ is in the $\varepsilon$-neighborhood of $f^*(x)$ for all $x \in \varsigma$. Therefore this integer $N(\varsigma, \varepsilon)$ can be interpreted as an extended version of the standard concept known in the literature of MDPs as the Forecast Horizon (FH) (see [6], [7], [10], [13], [19], [20], [21], and [22]). The FH is a positive integer $N^*$ such that $f_n = f^*$, for all $n \geq N^*$. Note that the FH allows us to obtain a
strong class of convergence of \( \{f_n\} \) to \( f^\ast \), i.e. the limit \( f^\ast \) is attained in \( N^\ast \) steps. This excludes important control problems in which \( f_n \neq f^\ast \) for all \( n \), for instance, the linear-quadratic model (see [5] and [12]). In this paper the linear-quadratic model satisfies practically all the conditions proposed.

The results presented here are based mainly on the continuity of the cost function, of \( V^\ast \) and \( V_n \), \( n = 1, 2, \cdots \), as well as on the continuity on \( \mathbb{K} = \{(x, a) \mid x \in X, a \in A(x)\} \) of

\[
\int V_n(y)Q(dy \mid \cdot, \cdot),
\]

\( n = 1, 2, \cdots \), and also of

\[
\int V^\ast(y)Q(dy \mid \cdot, \cdot)
\]

(these integrals are associated with the right-hand side of the Optimality Equation, see [5], [11], [12], [13] and [16]).

Due to the fact that the cost function is assumed nonnegative, \( V_n \uparrow V^\ast \) (see [12]) and \( \int V_n(y)Q(dy \mid \cdot, \cdot) \uparrow \int V^\ast(y)Q(dy \mid \cdot, \cdot) \). In this situation, the assumptions of Dini’s Theorem (see [14]), are satisfied. This important fact will be used in the proof of the results (see Lemma (3.28) below).

Several examples are presented to illustrate the theory developed in this paper. These examples are: an inventory/production system (see [11]), the linear-quadratic model (see [5]), and a nonlinear additive-noise system with unbounded cost function.

Now some comments about previous results in the literature will be provided.

The study of the pointwise convergence has been dealt with in Stokey and Lucas [23] for MDPs on Euclidean spaces, compact and convex actions sets and assuming strict concavity with respect to the actions in the right-hand side of the Optimality Equation (Stokey and Lucas [23] deal with rewards instead of costs). Furthermore, if \( X \) is compact, the uniform convergence of \( \{f_n\} \) to \( f^\ast \) has been obtained in [23].

It is important to mention the result obtained by Schäl [17] (see also [15]) which permits, in the case of compact admissible actions sets, the conclusion that for each \( x \in X \), \( f^\ast (x) \) is an accumulation point of the sequence \( \{f_n(x)\} \).

Finally, for general optimization problems with applications to deterministic MDPs, the convergence of optimal solutions of finite horizon problems to the optimal solutions of the infinite horizon problems is analyzed in [19].

The paper is organized as follows. Section 2 provides the preliminaries, i.e. the basic theory that will be used in the paper. In Section 3, the main assumptions and the theorem of the uniform convergence are provided. In Section 4 examples are presented. In Section 5, some remarks about pointwise convergence are given.

2. Preliminaries

Let \( (X, A, \{A(x) : x \in X\}, Q, c) \) be a discrete-time, stationary Markov decision model (see [12] for notation and terminology) which consists of the state space \( X \), the control or action set \( A \), the admissible action sets \( A(x), x \in X \), the transition law \( Q \), and the one-stage cost \( c \).
The sets $X$ and $A$ are assumed to be Borel spaces, with Borel σ-algebras $\mathcal{B}(X)$ and $\mathcal{B}(A)$, respectively. Moreover, for every $x \in X$ there is a nonempty Borel set $A(x) \subset A$ whose elements are the feasible control actions when the state of the system is $x$. Define $\mathcal{K} := \{(x, a) \mid x \in X, a \in A(x)\}$. The transition law $Q(B \mid x, a), B \in \mathcal{B}(X), x \in X, a \in A(x)$ is a stochastic kernel on $X$, given $\mathcal{K}$ (that is, $Q(\cdot \mid x, a)$ is a probability measure on $X$ for every $(x, a) \in \mathcal{K}$; and $Q(B \mid \cdot)$ is a measurable function on $\mathcal{K}$ for every $B \in \mathcal{B}(X)$). Finally, $c : \mathcal{K} \to \mathbb{R}$ is a measurable function which represents the cost-per-stage.

A policy $\pi$ is a (measurable, possibly randomized) rule for choosing actions, and at each time $t = 0, 1, \cdots$, the control prescribed by $\pi$ may depend on the current state as well as on the history of previous states and actions. The set of all policies is denoted by $\Pi$. Given the initial state $x_0 = x$, any policy $\pi$ defines the unique probability distribution of the state-action processes $\{(x_t, a_t)\}$. For details see, for instance, [12]. This probability distribution is denoted by $P^\pi_x$, whereas $E^\pi_x$ stands for the corresponding expectation operator. Let $\mathcal{F}$ be the set of all measurable functions $f : X \to A$, such that $f(x) \in A(x)$ for every $x \in X$. A policy $\pi \in \Pi$ is stationary if there exists $f \in \mathcal{F}$ such that, under $\pi$, the control $f(x_t)$ is applied at each time $t = 0, 1, \cdots$. The set of all stationary policies is identified with $\mathcal{F}$.

The focus here is on the expected total discounted cost defined as

\begin{equation}
V(\pi, x) = E^\pi_x \left[ \sum_{t=0}^{\infty} \alpha^t c(x_t, a_t) \right],
\end{equation}

when the policy $\pi \in \Pi$ is used, and $x \in X$ is the initial state. In (2.1), $\alpha \in (0, 1)$ is a given discount factor.

A policy $\pi^*$ is said to be optimal if

\begin{equation}
V(\pi^*, x) = V^*(x),
\end{equation}

$x \in X$, where

\begin{equation}
V^*(x) := \inf_{\pi \in \Pi} V(\pi, x),
\end{equation}

$x \in X$, is the so-called optimal value function.

Now some assumptions and results to be used in the next sections will be listed.

**Assumption (2.3).**

a. The one-stage cost $c$ is nonnegative, lower semicontinuous (l.s.c.) and inf-compact on $\mathcal{K}$. (Recall that $c$ is inf-compact on $\mathcal{K}$ if the set

$\{a \in A(x) \mid c(x, a) \leq s\}$

is compact for every $x \in X$ and $s \in \mathbb{R}$.)

b. The transition law $Q$ is strongly continuous, i.e.,

$\mu'(x, a) := \int \mu(y) Q(dy \mid x, a)$

is continuous and bounded on $\mathcal{K}$ for every measurable bounded function $\mu : X \to \mathbb{R}$.

c. There is a policy $\pi$ such that $V(\pi, x) < \infty$ for each $x \in X$. 

Definition (2.4). The value iteration functions are defined as

\[ V_n(x) = \min_{a \in A(x)} \left[ c(x, a) + \alpha \int V_{n-1}(y) Q(dy \mid x, a) \right], \]

for all \( x \in X \) and \( n = 1, 2, \ldots \), with \( V_0(\cdot) = 0 \).

Remark (2.6). Using Assumption (2.3) it is possible to demonstrate (see [8]) that for each \( n = 1, 2, \ldots \), \( V_n \) is a measurable function and there exists a stationary policy \( f_n \in \mathcal{F} \) such that the minimum in (2.5) is attained, i.e.,

\[ \forall x \in X. \]

Lemma (2.8). ([12], Theorem 4.2.3) Suppose that Assumption (2.3) holds. Then

a. The optimal value function \( V^* \) defined in (2.2) is the (pointwise) minimal solution of the Optimality Equation (OE), i.e., for all \( x \in X \):

\[ V^*(x) = \min_{a \in A(x)} \left[ c(x, a) + \alpha \int V^*(y) Q(dy \mid x, a) \right], \]

and if \( u \) is another solution to the OE, then \( u(\cdot) \geq V^*(\cdot) \).

There is also \( f^* \in \mathcal{F} \) such that

\[ V^*(x) = c(x, f^*(x)) + \alpha \int V^*(y) Q(dy \mid x, f^*(x)), \]

\( x \in X \), and \( f^* \) is optimal.

b. For every \( x \in X \), \( V_n(x) \uparrow V^*(x) \) as \( n \to +\infty \).

Assumption (2.11). Suppose that \( f^* \) given in (2.10) is unique.

Remark (2.12). See [8] for conditions to ensure the uniqueness of optimal policies of discounted MDPs.

Throughout the paper, MDPs that satisfy Assumptions (2.3) and (2.11) are considered. Assumptions (2.3) and (2.11) will not be mentioned in each Lemma or Theorem in this paper, but they are supposed to hold.

Lemma (2.13). Let \( \{g_n\} \) be a sequence of continuous real-valued functions on a metric space \((Y, d_Y)\). Let \( g \) be a continuous real-valued function on \( Y \). Then \( g_n(z_n) \to g(z), n \to +\infty \) for every \( z \in Y \) and every sequence \( \{z_n\} \) in \( Y \) which converges to \( z \) if and only if \( \{g_n\} \) converges uniformly on compact sets to \( g \).

Proof. Suppose that \( g_n(z_n) \to g(z), n \to +\infty \) for every \( z \in Y \) and every sequence \( \{z_n\} \) in \( Y \) which converges to \( z \in Y \), but there is a compact set \( \varsigma \subset Y \) such that \( \{g_n\} \) does not converge uniformly to \( g \) on \( \varsigma \). Consequently, there exist \( \varepsilon > 0 \), a sequence of positive integers \( \{n_k\}, n_1 < n_2 < \cdots \), and a sequence \( \{y_k\} \subset \varsigma \), such that

\[ |g_{n_k}(y_k) - g(y_k)| \geq \varepsilon, \]

\( k = 1, 2, \ldots \). It is possible to assume without losing generality that there exists \( y \in \varsigma \) such that \( y_k \to y, k \to +\infty \) (recall that \( \varsigma \) is a compact set). Now take \( z_{n_k} := y_k, k = 1, 2, \ldots \), and observe that \( z_{n_k} \to y, k \to +\infty \). By the hypothesis,
it follows that $g_{n_k}(z_{n_k}) \to g(y)$, $k \to +\infty$. Therefore, by the continuity of $g$, it results that

$$\tag{2.15} \left| g_{n_k}(z_{n_k}) - g(z_{n_k}) \right| \leq \left| g_{n_k}(z_{n_k}) - g(y) \right| + \left| g(y) - g(z_{n_k}) \right| \to 0, \quad k \to +\infty.$$ 

On the other hand, putting $y_k = z_{n_k}$, $k = 1, 2 \cdots$, in (2.14), it follows that

$$\tag{2.16} \left| g_{n_k}(z_{n_k}) - g(z_{n_k}) \right| \geq \varepsilon > 0,$$

for all $k = 1, 2, \cdots$, which is a contradiction to (2.15).

Conversely, assume that $\{g_n\}$ converges uniformly on compact sets to $g$. Take $z \in Y$ and let $\{z_n\}$ be a sequence in $Y$ which converges to $z$. Put $\varsigma := \{z_n\} \cup \{z\}$. Hence $\varsigma$ is a compact set and, by hypothesis, $\{g_n\}$ converges uniformly on $\varsigma$ to $g$. Thus, this fact, the continuity of $g$, and the inequality

$$\tag{2.17} \left| g_n(z_n) - g(z) \right| \leq \left| g_n(z_n) - g(z_n) \right| + \left| g(z_n) - g(z) \right|,$$

$n = 1, 2, \cdots$, imply that $g_n(z_n) \to g(z)$, $n \to +\infty$. This completes the proof of Lemma (2.13). \qed

Remark (2.18). Lemma (2.13) is a particular case of result 7.5 of Chapter XII in [9]. In fact, Lemma (2.13) holds for $g_n$ and $g$ which are defined in a topological space $Y$ (that is 1st countable) and take values in a metric space $Z$ (see [9]). The proof of this Lemma is presented here for completeness of the paper.

Definition (2.19). Let $X$ and $Y$ be (nonempty) Borel spaces. A multifunction $\Psi$ from $X$ to $Y$ is said to be

a. \textit{upper semicontinuous} (u.s.c.) if $\{x \in X \mid \Psi(x) \cap F \neq \emptyset\}$ is closed in $X$ for every closed $F \subset Y$;

b. \textit{lower semicontinuous} (l.s.c.) if $\{x \in X \mid \Psi(x) \cap G \neq \emptyset\}$ is open in $X$ for every open set $G \subset Y$;

c. \textit{continuous} if it is both u.s.c. and l.s.c.

A convenient characterization of a compact-valued u.s.c. multifunction will be given in the following result (for the proof of this result see Theorem 16.20 and Theorem 16.21, p. 534 in [1]). It will be used in Sections 3 and 4 below.

Lemma (2.20). Let $X$ and $Y$ be (nonempty) Borel spaces. Let $\Psi$ be a multifunction from $X$ to $Y$. Suppose that $\Psi(x) \neq \emptyset$ for every $x \in X$ and, moreover, $\Psi$ is compact-valued. Then the following statements are equivalent:

a. $\Psi$ is u.s.c.

b. If $x_n \to x$, $n \to +\infty$ in $X$ and $y_n \in \Psi(x_n)$, $n = 1, 2, \cdots$, there exist a subsequence $\{y_{n_k}\}$ of $\{y_n\}$ and $y \in \Psi(x)$ such that $y_{n_k} \to y$, $k \to +\infty$.

3. Uniform Convergence on Compact Sets

In this section the assumptions and the theorem which ensure the uniform on compact sets convergence of minimizers of the value iteration algorithm of discounted MDPs to the optimal policies are presented.

The following notation will be used in the rest of the paper.
**Notation (3.1).** a. For \( x \in X \), denote
\[
\widehat{A}(x) := \{ a \in A(x) \mid c(x, a) \leq V^*(x) \},
\]
where \( V^* \) is the function given in (2.2).

b. For each \( n = 1, 2, \cdots \),
\[
(3.2) \quad G_n(x, a) := c(x, a) + \alpha \int V_{n-1}(y) \, Q(dy \mid x, a),
\]
\((x, a) \in \mathbb{K}\), where \( V_{n-1} \) is the function given in (2.5),
c. \( G(x, a) := c(x, a) + \alpha \int V^*(y) \, Q(dy \mid x, a), (x, a) \in \mathbb{K}. \)
d. Denote
\[
(3.3) \quad \widehat{\mathbb{K}}_s := \{(x, a) \in \mathbb{K} \mid x \in s, a \in \widehat{A}(x)\},
\]
where \( s \subset X \) is a nonempty compact set.

**Assumption (3.4).** a. The multifunction \( x \mapsto A(x) \) is u.s.c. and closed-valued.
b. \( c(\cdot, \cdot) \) is a continuous function on \( \mathbb{K} \).
c. \( V_n(\cdot), n = 1, 2, \cdots \), and \( V^*(\cdot) \) are continuous functions on \( X \).
d. The integrals
\[
(3.5) \quad \int V_n(y) \, Q(dy \mid \cdot, \cdot), n = 1, 2, \cdots ,
\]
and
\[
(3.6) \quad \int V^*(y) \, Q(dy \mid \cdot, \cdot)
\]
are finite and continuous functions on \( \mathbb{K} \).

**Assumption (3.7).** a. \( A \), the control space, is a compact set.
b. The multifunction \( x \mapsto A(x) \) is compact-valued.
c. The one-stage cost \( c \) is strictly unbounded, that is, there exist nondecreasing sequences of compact sets \( X_n \uparrow X, n \to +\infty \), and \( A_n \uparrow A, n \to +\infty \) such that \( \Lambda_n := X_n \times A_n \) is a subset of \( \mathbb{K} \), and
\[
(3.8) \quad \lim_{n \to +\infty} \inf_{(x,a) \in \Lambda_n^c} c(x,a) = +\infty,
\]
where \( \Lambda_n^c \) denotes the complement of \( \Lambda_n \).

**Remark (3.9).** a. Notice that the inf-compactness on \( \mathbb{K} \) of the cost function \( c \) implies that \( \widehat{A}(x) \) is compact for every \( x \in X \); moreover, it is direct to obtain that \( f_n(x) \) and \( f^*(x) \) belong to \( A(x) \), for every \( x \in X \) and every \( n = 1, 2, \cdots \).
b. Observe that \( \mathbb{K} \) is closed as an immediate consequence of Assumption (3.4a) and Proposition 7, p. 110 [2] (see also Proposition D3 in Appendix D, p. 182 [12]). In addition, it is easily seen that Assumptions (3.4a), (3.4b) and (3.4c) imply that \( \widehat{\mathbb{K}}_s \) is closed in \( X \times A \) for every nonempty compact set \( s \subset X \).

**Lemma (3.10).** Suppose that Assumptions (3.4a), (3.4b), and (3.4c) hold. Then each Assumption (3.7a), (3.7b), or (3.7c) implies that \( \widehat{\mathbb{K}}_s \) is a compact set for every nonempty compact set \( s \subset X \).
**Proof.** Suppose that Assumptions (3.4a), (3.4b), and (3.4c) hold. Let $\varsigma$ be an arbitrary, fixed, nonempty compact subset of $X$. Observe that $\hat{K}_\varsigma$ is closed in $X \times A$ (see Remark (3.9(b))). Therefore, a suitable compact set $\hat{J} \subset X \times A$ such that $\hat{K}_\varsigma \subset \hat{J}$ will be shown for each Assumption (3.7a), (3.7b), or (3.7c). This allows to conclude the compactness of $\hat{K}_\varsigma$.

Firstly, suppose Assumption (3.7a) holds. In this case, taking $\hat{K}_\varsigma \subset \hat{J} := \varsigma \times \Lambda_{n_0}$, it results that $\hat{K}_\varsigma$ is compact.

Secondly, suppose Assumption (3.7b) holds. Now observe that $\cup_{x \in \varsigma} A(x)$ is a compact set (see [4] p. 72). As

$$\hat{K}_\varsigma \subset \hat{J} := \varsigma \times \cup_{x \in \varsigma} A(x),$$

then $\hat{K}_\varsigma$ is a compact set.

Thirdly, suppose Assumption (3.7c) holds. It will be shown that there exists a positive integer $n_0$ such that $\hat{K}_\varsigma \subset \Lambda_{n_0}$. By contradiction, assume that for each $n = 1, 2, \ldots$, there exists $(x_n, a_n) \in \hat{K}_\varsigma$, such that $(x_n, a_n) \notin \Lambda_n$. Observe that

$$c(x_n, a_n) \leq V^*(x_n) \leq L := \sup_{x \in \varsigma} V^*(x) < +\infty,$$

for all $n = 1, 2, \ldots$ (recall that Assumption (3.4c) holds, hence $V^*$ is a continuous function).

Therefore

$$\limsup_{n \to \infty} c(x_n, a_n) \leq L.$$

Finally, as

$$\inf_{(x, a) \in \Lambda_{n_0}^c} c(x, a) \leq c(x_n, a_n),$$

$n = 1, 2, \ldots$, then

$$+\infty = \lim_{n \to \infty} \inf_{(x, a) \in \Lambda_{n_0}^c} c(x, a) \leq \limsup_{n \to \infty} c(x_n, a_n),$$

which is a contradiction to (3.13). Thus, there exists $n_0$ such that $\hat{K}_\varsigma \subset \hat{J} := \Lambda_{n_0}$, which implies that $\hat{K}_\varsigma$ is a compact set.

Since $\varsigma$ is arbitrary, the desired result follows.

**Corollary (3.16).** Suppose that Assumptions (3.4a), (3.4b) and (3.4c) hold. Then each of Assumptions (3.7a), (3.7b) or (3.7c) implies that the multifunction $x \to \hat{A}(x)$ is u.s.c.

**Proof.** Observe that $\hat{A}(x)$ is compact for every $x \in X$ (see Remark (3.9a)). Now suppose $x_n \to x$ in $X$ and take $a_n \in \hat{A}(x_n)$, $n = 1, 2, \ldots$. Let $\varsigma = \{x_n\} \cup \{x\}$. Notice that $\varsigma$ is a compact set, so from Lemma (3.10), $\hat{K}_\varsigma$ is also compact. Therefore, since $(x_n, a_n) \in \hat{K}_\varsigma$, $n = 1, 2, \ldots$, then there exist a subsequence $\{(x_{n_k}, a_{n_k})\}$ of $\{(x_n, a_n)\}$ and $(x, a) \in \hat{K}_\varsigma$ such that

$$c(x_{n_k}, a_{n_k}) \to c(x, a),$$

$k \to \infty$. In particular, notice that $a_{n_k} \to a$, $k \to +\infty$ and $a \in \hat{A}(x)$. Hence, from Lemma (2.20), the result follows.
**Lemma (3.18).** Suppose that Assumption (3.4) and one of Assumptions (3.7a), (3.7b), or (3.7c) hold. Then the stationary optimal policy $f^*$ is a continuous function.

**Proof.** Suppose that $f^*$ is not continuous. Then there exist $x \in X$ and a sequence $\{x_n\}$ in $X$ such that
\begin{equation}
 x_n \to x,
\end{equation}
but
\begin{equation}
 f^*(x_n) \not\to f^*(x).
\end{equation}

Then there exist $\varepsilon > 0$ and a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that
\begin{equation}
 d(f^*(x_{n_k}), f^*(x)) \geq \varepsilon,
\end{equation}
for all $k = 1, 2, \ldots$. (\(d\) denotes the metric in the control space $A$.)

Let $y_{n_k} := f^*(x_{n_k})$, $k = 1, 2, \ldots$. Observe that $y_{n_k} \in \hat{A}(x_{n_k})$, for all $k = 1, 2, \ldots$, and that the multifunction $x \mapsto \hat{A}(x)$ is compact-valued, and it is also u.s.c. as a consequence of Corollary (3.16). Now, from Lemma (2.20) there exist a subsequence $\{y_{n_{k_l}}\}$ of $\{y_{n_k}\}$ and $y \in \hat{A}(x)$ such that
\begin{equation}
 y_{n_{k_l}} \to y.
\end{equation}

Then, from (3.21) it results that
\begin{equation}
 d(y_{n_{k_l}}, f^*(x)) \geq \varepsilon,
\end{equation}
for all $l = 1, 2, \ldots$. Letting $l \to +\infty$ in (3.23), it follows that
\begin{equation}
 d(y, f^*(x)) \geq \varepsilon.
\end{equation}

On the other hand, (2.10) implies that
\begin{equation}
 V^*(x_{n_{k_l}}) = c(x_{n_{k_l}}, y_{n_{k_l}}) + \alpha \int V^*(y) Q(dy | x_{n_{k_l}}, y_{n_{k_l}}),
\end{equation}
for all $l = 1, 2, \ldots$. Hence, letting $l \to +\infty$ and using Assumption (3.4) in (3.25), it results that
\begin{equation}
 V^*(x) = c(x, y) + \alpha \int V^*(z) Q(dz | x, y),
\end{equation}
but
\begin{equation}
 V^*(x) = c(x, f^*(x)) + \alpha \int V^*(z) Q(dz | x, f^*(x)).
\end{equation}

Therefore, from (3.26), (3.27) and the uniqueness of $f^*$, it follows that $y = f^*(x)$, which is a contradiction to (3.24). This completes the proof of Lemma (3.18).

**Lemma (3.28).** Suppose that Assumption (3.4) holds. Then $\{G_n\}$ converges uniformly to $G$ on every nonempty compact subset of $\mathbb{K}$.

**Proof.** Let $\Theta$ be an arbitrary, fixed, nonempty compact subset of $\mathbb{K}$. Note that from Lemma (2.8) it results that $\{G_n\}$ converges pointwise on $\Theta$ to $G$, in fact, $G_n \uparrow G$, $n \to +\infty$. In addition, from Assumption (3.4), $G$ and $G_n$, $n = 1, 2, \ldots$, are continuous functions. Then, by Dini’s Theorem ([14] p. 239), $G_n \to G$ uniformly on $\Theta$. Since $\Theta$ is arbitrary, the result follows.
Lemma (3.29). Suppose that Assumption (3.4) and one of Assumptions (3.7a), (3.7b) or (3.7c) hold. Then, for each nonempty compact set \( \varsigma \subset X \), the following holds: for every \( \varepsilon > 0 \), there exists \( \delta > 0 \) such that, for all \( (x, a) \in \hat{K}_\varsigma \), if
\[
\left| G(x, f^*(x)) - G(x, a) \right| < \delta,
\]
then
\[
d(f^*(x), a) < \varepsilon.
\]

Proof. Let \( \varsigma \) be an arbitrary, fixed, nonempty compact subset of \( X \). The proof is by contradiction. Suppose that there exists \( \varepsilon > 0 \), such that for every \( \delta > 0 \), there is \( (x_\delta, a_\delta) \in \hat{K}_\varsigma \), which satisfies
\[
\left| G(x_\delta, f^*(x_\delta)) - G(x_\delta, a_\delta) \right| < \delta,
\]
and
\[
d(f^*(x_\delta), a_\delta) \geq \varepsilon.
\]

Let \( \delta = \frac{1}{n}, n = 1, 2, \ldots \). Then there exist \( (x_n, a_n) \in \hat{K}_\varsigma \), such that
\[
\left| G(x_n, f^*(x_n)) - G(x_n, a_n) \right| < \frac{1}{n},
\]
and
\[
d(f^*(x_n), a_n) \geq \varepsilon,
\]
for all \( n = 1, 2, \ldots \). Now, since \( \hat{K}_\varsigma \) is a compact set, there exist a subsequence \( \{(x_n, a_n)\} \) of \( \{(x_n, a_n)\} \) and \( (x, a) \in \hat{K}_\varsigma \), such that \( (x_n, a_n) \to (x, a), k \to \infty \). Notice that
\[
\left| G(x_n, f^*(x_n)) - G(x_n, a_n) \right| < \frac{1}{n_k}
\]
for all \( k = 1, 2, \ldots \). Therefore, letting \( k \to \infty \) and using Assumption (3.4) and Lemma (3.18) in (3.36), it follows that
\[
G(x, a) = G(x, f^*(x)).
\]

This implies that
\[
a = f^*(x),
\]
(recall the uniqueness of \( f^* \)).

Nevertheless, from (3.35)
\[
d(f^*(x_n), a_n) \geq \varepsilon,
\]
for all \( k = 1, 2, \ldots \).

So, if \( k \to \infty \) in the last inequality and using Lemma (3.18),
\[
d(f^*(x), a) \geq \varepsilon,
\]
which is a contradiction to (3.38). Since \( \varsigma \) is arbitrary, the proof of Lemma (3.29) is finished.

Now, the uniform convergence on compact sets of minimizers of the value iteration algorithm of Discounted MDPs to optimal policies will be proved.
Theorem (3.41). Suppose that Assumption (3.4) and one of Assumptions (3.7a), (3.7b), or (3.7c) hold. Then \( \{ f_n \} \) converges uniformly on compact sets to \( f^* \).

Proof. Let \( \varsigma \) be an arbitrary, fixed, non-empty compact of \( X \). Observe that \( f_n(x) \) minimizes to \( G_n(x, \cdot) \) on \( A(x) \), and \( f^*(x) \) is the (unique) minimizer to \( G(x, \cdot) \) on \( A(x) \), for all \( x \in \varsigma \), and for all \( n \). Therefore,

\[
0 \leq G(x, f_n(x)) - G(x, f^*(x)) \leq G(x, f_n(x)) - G_n(x, f_n(x)) + G_n(x, f^*(x)) - G(x, f^*(x)) \leq 2 \sup_{(x,a) \in \hat{K}_{\varsigma}} |G(x, a) - G_n(x, a)|,
\]

(3.42)

for all \( x \in \varsigma \), and for all \( n \).

Now, let \( \varepsilon > 0 \). By Lemma (3.29) applied to \( \varsigma \), there exists \( \delta > 0 \) such that for all \( (x, a) \in \hat{K}_{\varsigma} \), if

\[
|G(x, f^*(x)) - G(x, a)| < \delta,
\]

(3.43)

then

\[
d(f^*(x), a) < \varepsilon.
\]

Furthermore, Lemma (3.28) guarantees the existence of a positive integer \( R \) such that

\[
2 \sup_{(x,a) \in \hat{K}_{\varsigma}} |G(x, a) - G_n(x, a)| < \delta,
\]

(3.45)

if \( n \geq R \). Now, from (3.42) and (3.45), if \( n \geq R \),

\[
|G(x, f^*(x)) - G(x, f_n(x))| < \delta,
\]

(3.46)

for all \( x \in \varsigma \).

Combining (3.46), (3.43) and (3.44) it follows that

\[
d(f^*(x), f_n(x)) < \varepsilon,
\]

(3.47)

for all \( n \geq R \), and for all \( x \in \varsigma \). So, \( f_n \to f^* \) uniformly on \( \varsigma \).

Since \( \varsigma \) is arbitrary, the result follows. \( \square \)

4. Examples

Remak (4.1). Consider MDPs that satisfy Assumptions (2.3), (2.11), (3.4a) and (3.4b) and that both integrals (3.5) and (3.6) are finite.

Concerning the continuity required in Assumptions (3.4c) and (3.4d), for \( V_n, n = 1, 2, \cdots, V^* \), and the integrals (3.5) and (3.6), observe the following.

a. The continuity mentioned trivially holds for discrete models (i.e., MDPs for which both \( X \) and \( A \) are finite or denumerable sets).

b. For bounded models (i.e., MDPs with bounded cost functions and compact admissible action sets), the continuity of the integrals (3.5) and (3.6) follows directly from the strong continuity of the transition law \( Q \). If moreover, the multifunction \( x \mapsto A(x) \) is continuous (see Definition (2.19c)), then the continuity of \( V_n, n = 1, 2, \cdots \), and \( V^* \) is an immediate consequence of Proposition
D.3(c) p. 130 in [11], using the continuity of the cost function \( c \) and of the integrals (3.5) and (3.6), and equations (2.5) and (2.9).

For such a bounded model, see Example (4.2) below.

c. In [8] two convexity conditions have been presented (see Conditions C1 and C2 in [8]), each of which guarantees (in particular) that \( V_n, n = 1, 2, \cdots, V^* \), and the integrals (3.5) and (3.6) are convex functions (see Lemma 6.2 p. 433 and its proof in [8]). If, in addition, the spaces \( X, A \) and \( \mathbb{K} \) are open sets, then the continuity required in Assumptions (3.4c) and (3.4d) is obtained (see Theorem 3, p. 113 [3]). This is illustrated in Examples (4.9) and (4.15) below.

An inventory/production model is presented below (see [11] p. 10 for the precise description of this example in the context of inventory/production area).

**Example (4.2).** Let \( M \) be a fixed positive constant. Let \( X = A = [0, M] \), \( A(x) = [0, M - x] \), \( x \in X \), and consider

\[
x_{t+1} = [x_t + a_t - \xi_t]^+, \quad t = 0, 1, \cdots,
\]

where \( z^+ := \max\{0, z\} \). Here \( \xi_0, \xi_1, \cdots \) are i.i.d. random variables taking values in \( S = [0, \infty) \), and with common density \( \Delta \).

**Assumption (4.4).**

a. \( \Delta \) is a bounded continuous function. (Notice that the distribution function \( \hat{G} \) of \( \xi \) is a continuous function, where \( \xi \) is a generic element of the sequence \( \{\xi_t\} \).)

b. \( c \) is non-negative, continuous and strictly convex on \( \mathbb{K} \) (observe that \( \mathbb{K} \) in this example is a compact set); also \( c \) is an increasing function in the first variable.

**Lemma (4.5).**

a. Example (4.2) satisfies Assumptions (2.3), (2.11), (3.7a), and (3.7b).

b. The Assumption (3.4) holds.

**Proof.**

a. Example (4.2) clearly satisfies Assumptions (2.3a) and (2.3c). In reference to Assumption (2.3b), notice that if \( \mu : X \to \mathbb{R} \) is a measurable and bounded function, then a simple computation shows that

\[
\int \mu(y) Q(dy | x, a) = \mu(0) \left[ 1 - \hat{G}(x + a) \right] + \int I_{[0,x+a]}(u) \mu(u) \Delta(x + a - u) du,
\]

(4.6)

\((x, a) \in \mathbb{K}, \) where \( I_{[\cdot, \cdot]} \) denotes the indicator function of the subset \([\cdot, \cdot]\).

Since \( \hat{G} \) is a continuous function, it follows that \( \mu(0) \left[ 1 - \hat{G}(x + a) \right] \) is a continuous function on \( \mathbb{K} \). Now, as \( \mu \) is a bounded function and \( \Delta \) is a bounded continuous function, it follows directly, using the Dominated Convergence Theorem, that

\[
\int I_{[0,x+a]}(u) \mu(u) \Delta(x + a - u) du
\]

(4.7)

is a continuous function on \( \mathbb{K} \). Hence

\[
\int \mu(y) Q(dy | \cdot, \cdot)
\]

(4.8)
is a continuous function on $K$.

Observe that by means of straightforward computations it is possible to verify that this Example satisfies Condition C1 in [8], therefore, the uniqueness of the optimal policy holds.

Finally, notice that Assumptions (3.7a) and (3.7b) trivially hold.

b. Firstly, it is evident that $A(x)$ is a closed set for each $x \in X$. Secondly, since $c$ is bounded, it follows that for each $n = 1, 2, \cdots$, $V_n$ and also $V^*$ are bounded and they are also measurable from Assumption (2.3) (which holds from item a). Therefore, from the fact that $V_n$ and $V^*$ are bounded, it follows that the integrals (3.5) and (3.6) are finite; and the strong continuity of $Q$ allows to conclude that (3.5) and (3.6) are continuous on $K$. Finally, direct computations show that the multifunction $\mathbf{x} \mapsto \mathbf{A} (\mathbf{x})$ is continuous (i.e. $x \mapsto A(x)$ is both u.s.c. and l.s.c. (see Definition (2.19) and Lemma (2.20)). Hence, the continuity of the integrals (3.5) and (3.6), the continuity of the cost $c$, (2.5), (2.9) and Proposition D.3 (c) p. 130 in [11] imply that $V_n$ and $V^*$ are continuous. \hfill $\blacksquare$

**Example (4.9).** Consider a simple linear system

\begin{equation*}
\mathbf{x}_{t+1} = \widehat{\gamma} \mathbf{x}_t + \delta \mathbf{a}_t + \xi_t, \quad t = 0, 1, \cdots,
\end{equation*}

with quadratic cost

\begin{equation*}
c(\mathbf{x}, \mathbf{a}) = q \mathbf{x}^2 + r \mathbf{a}^2,
\end{equation*}

$x, a \in \mathbb{R}$. Here $X = A = A(x) = \mathbb{R}$, for all $x \in X$.

**Assumption (4.12).**

a. $\widehat{\gamma} \delta \neq 0$, both $q$ and $r$ are positive.

b. The disturbances $\xi_t$, $t = 0, 1, \cdots$ are i.i.d. random variables with values in $S = \mathbb{R}$. Moreover, suppose that $\xi_0$ has a continuous density $\Delta$, zero mean value and a finite variance $\sigma^2 > 0$.

**Remark (4.13).** Assumptions (2.3) and (2.11) have been proved in Example 4.8 in [8] (in particular, Condition C2 in [8] holds for this model). This Example satisfies trivially Assumptions (3.4a), (3.4b), and (3.7c). On the other hand, it is easy to verify that $W(x) = l \left( x^2 + \frac{\alpha}{1 - \alpha} \sigma^2 \right)$, with $l := q + r \left( \widehat{\gamma}^2 / \delta^2 \right)$, is an upper bound for $V_n(x)$, $n = 1, 2, \cdots$, and $V^*(x), x \in X$. In fact, $W(x) = V(f, x), x \in X$, for $f(x) = - \left( \widehat{\gamma} / \delta \right) x, x \in X$ (see Lemma 4.9(c) in [8]). Since

\begin{equation*}
\int W(y) Q(dy | x, a) = l \left[ (\widehat{\gamma} x + \delta a)^2 + \sigma^2 \right] + \frac{\alpha \sigma^2}{1 - \alpha} < \infty,
\end{equation*}

$(x, a) \in \mathbb{R}^2$, it follows that the integrals (3.5) and (3.6) are finite. Finally, since for this Example the Condition C2 in [8] holds, hence from Lemma 6.2 (and its proof) in [8] it follows that $V_n$ and $V^*$ are convex on $\mathbb{R}$, and the integrals (3.5) and (3.6) are convex on $\mathbb{R}^2$. Consequently, Assumptions (3.4c) and (3.4d) hold (see Remark (4.1)).

Now an example of a nonlinear additive-noise system with unbounded cost function will be presented.

**Example (4.15).** Consider

\begin{equation*}
\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{a}_t^2 + \xi_t, \quad t = 0, 1, \cdots,
\end{equation*}
with cost
\begin{equation}
(4.17) \quad c(x, a) = h(x) + g(a),
\end{equation}
where \( h(x) = e^x \), \( g(a) = 2a^4 + a + 1 \), \( x, a \in \mathbb{R} \), and \( X = A = A(x) = \mathbb{R} \), for all \( x \in X \).

**Assumption (4.18).** The disturbances \( \xi_t, t = 0, 1, \cdots \) are i.i.d. random variables with values in \( S = \mathbb{R} \). Moreover, suppose that \( \xi_0 \) has a continuous density \( \Delta \), and
\begin{equation}
(4.19) \quad k := \int e^s \Delta(s) \, ds < \infty,
\end{equation}
with \( 0 < \alpha k < 1 \).

**Remark (4.20).** Clearly, \( c \) is non-negative (in fact, \( h(x) > 0, x \in X \) and \( g(a) \geq 5/8 \), \( a \in A \), therefore \( c(x, a) > 5/8 \) for all \( (x, a) \in \mathbb{R} \)), and continuous, and Assumptions (3.4a) and (3.4b) hold. The inf-compactness of \( c \) in \( \mathbb{R} \) and Assumption (3.7c) follow directly from the fact that
\begin{equation}
(4.21) \quad \lim_{a \to +\infty} g(a) = \lim_{a \to -\infty} g(a) = +\infty.
\end{equation}
Also this Example trivially satisfies Condition C1 in [8] (notice that this Example does not satisfy Condition C2 in [8]). Hence Assumption (2.11) follows.

**Lemma (4.22).**

\begin{enumerate}[(a)]
\item \( Q \) defined by (4.16) is strongly continuous.
\item There is \( \pi \in \Pi \) such that \( V(\pi, x) < \infty \).
\end{enumerate}

**Proof.**

(a) (4.16) and the (well-known) Change of Variable Theorem permit to get that
\begin{equation}
(4.23) \quad Q(B \mid x, a) = \int I_B [x + a^2 + s] \Delta(s) \, ds = \int_B \Delta(s - (x + a^2)) \, ds
\end{equation}
for \( x \in X, a \in A \) and \( B \in \mathcal{B}(X) \). Hence from Assumption (4.18) and from Example C.6 in Appendix C in [12], it results that \( Q \) is strongly continuous.

(b) Let \( f \in \mathbb{F} \) given \( f(x) = 0 \), for all \( x \in X \). Then it is possible to prove by straightforward induction argument that
\begin{equation}
(4.24) \quad E_x^t \left[ c(x_t, f) \right] = k^t e^x + 1,
\end{equation}
for each \( t = 0, 1, \cdots \), and \( x \in X \), where \( k \) was defined in Assumption (4.18). Now, from (4.24) and Assumption (4.18), it results that, for each \( x \in X \):
\begin{equation}
(4.25) \quad V(f, x) = \sum_{t=0}^{\infty} \alpha^t E_x^t \left[ c(x_t, f) \right] = \frac{e^x}{1 - \alpha k} + \frac{1}{1 - \alpha} < \infty.
\end{equation}
\[ \Box \]

**Remark (4.26).** Finally, similar to the previous Example (see Remark (4.13)), Assumptions (3.4c), and (3.4d) follow from Lemma 6.2 (and its proof) in [8].
5. Remarks on the Pointwise Convergence of Minimizers

Remark (5.1). As a consequence of Theorem (3.41), it is easily obtained that \( \{f_n\} \) converges pointwise to \( f^* \) on \( X \). Actually, for the pointwise convergence of \( \{f_n\} \) to \( f^* \), fewer and weaker assumptions are needed than those proposed for the uniform convergence on compact sets of \( \{f_n\} \) to \( f^* \) (see also Remark (5.14) below); in fact, it is sufficient with a weaker version of Assumption (3.4) (see Assumption (5.2) below), and Assumption (3.7) is not necessary.

Assumption (5.2). a. For each \( x \in X \), \( c(x, \cdot) \) is a continuous function on \( A(x) \).

b. For each \( x \in X \), \( V_n(y)Q(dy|x, \cdot), n = 1, 2, \cdots \),

are finite and continuous functions on \( A(x) \).

Lemma (5.5). Suppose that Assumption (5.2) holds. Then for each \( x \in X \), \( G_n(x, \cdot) \to G(x, \cdot) \) uniformly on every nonempty compact subset of \( A(x) \).

Proof. It is similar to the proof of Lemma (3.28).

Theorem (5.6). Suppose that Assumption (5.2) holds. Then \( f_n(x) \to f^*(x) \), \( n \to +\infty \) for each \( x \in X \).

Proof. The proof is by contradiction. Suppose that there exists \( x \in X \) such that \( f_n(x) \not\to f^*(x) \). Let \( \{f_{n_k}(x)\} \) be a subsequence of \( \{f_n(x)\} \), and \( \epsilon > 0 \) such that

\[
d(f_{n_k}(x), f^*(x)) \geq \epsilon,
\]

for all \( k = 1, 2, \cdots \).

Since \( \hat{A}(x) \) is a compact set (see Remark (3.9a)), there exist \( a_x \in \hat{A}(x) \), and a subsequence \( \{f_{n_{k_l}}(x)\} \) of \( \{f_{n_k}(x)\} \), such that

\[
f_{n_{k_l}}(x) \to a_x,
\]

\( l \to +\infty \). Observe that from (5.7) it results that

\[
d(a_x, f^*(x)) \geq \epsilon.
\]

Now note that by Assumption (5.2), Lemma (2.13), and Lemma (5.5) specialized to the compact \( \hat{A}(x) \), it follows that

\[
\lim_{l \to +\infty} G_{n_{k_l}}(x, f_{n_{k_l}}(x)) = G(x, a_x).
\]

On the other hand, for each \( l \geq 1 \),

\[
G_{n_{k_l}}(x, f_{n_{k_l}}(x)) = V_{n_{k_l}}(x).
\]
Hence from (2.10), (5.10), (5.11) and Lemma (2.8b), it follows that
\[
\lim_{l \to +\infty} G_{n_{k_l}}(x, f_{n_{k_l}}(x)) = V^*(x) = G(x, f^*(x)) = G(x, a_x).
\]
(5.12)

Then by uniqueness of \( f^* \) it results that \( f^*(x) = a_x \), which is a contradiction to (5.9). This completes the proof of Theorem (5.6) \( \Box \)

Now a corollary with two results which represent special cases of Theorem (5.6) will be shown. They are related to the concept known in the literature of MDPs as the Forecast Horizon (see [6], [7], [10], [13], [19], [20], [21] and [22]).

**Corollary (5.13).** If \( A \) is a finite set or a denumerable set, then for each \( x \in X \) there exists a positive integer \( N^*(x) \) such that \( f_n(x) = f^*(x) \), for all \( n \geq N^*(x) \), supposing that integrals (5.3) and (5.4) are finite.

**Proof.** Fix \( x \in X \). Suppose that \( A \) is a finite set or a denumerable set with the discrete metric. Since in this case Assumption (5.2) trivially holds, it follows that \( f_n(x) \to f^*(x) \). Then there exists a positive integer \( N^*(x) \) such that \( f_n(x) \in \{f^*(x)\} \) for all \( n \geq N^*(x) \) (recall that in the discrete metric, \( \{f^*(x)\} \) is an open set). Therefore \( f_n(x) = f^*(x) \) for all \( n \geq N^*(x) \). Since \( x \) is arbitrary, the result follows. \( \Box \)

**Remark (5.14).** Observe that under Assumptions (2.3) and (2.11), the pointwise convergence of \( \{f_n\} \) to \( f^* \) is a direct consequence of Theorem 4.6.5, p. 67 in [12] provided the control space \( A \) is locally compact. On the other hand, Proposition 12.2 in [18] yields the pointwise convergence for the case of admissible compact actions sets and under Assumption (2.11).

**Acknowledgement**

The authors wish to thank two anonymous referees and the associate editor for helpful comments and for suggesting us Remark (5.14).

*Received September 29, 2004*

*Final version received December 02, 2005*

DANIEL CRUZ SUÁREZ
DIVISIÓN ACADÉMICA DE CIENCIAS BÁSICAS
UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO
APDO. POSTAL 5
86690 CUNDUACÁN, TAB.
MÉXICO
daniel.cruz@basicas.ujat.mx

RAÚL MONTES-DE-OCA
DEPARTAMENTO DE MATEMÁTICAS UNIVERSIDAD AUTÓNOMA METROPOLITANA-IZTAPALAPA
AV. SAN RAFAEL ATLIXCO 186, COL. VICENTINA
09340 MÉXICO D.F.
MÉXICO
momr@xanum.uam.mx
REFERENCES