DISCOUNTED DYNAMIC PROGRAMMING ON COMPACT METRIC SPACES

By ASHOK MAITRA

Indian Statistical Institute

SUMMARY. It is proved that under certain assumptions there exists a stationary optimal plan for the discounted dynamic programming problem with continuous state and action spaces.

1. Introduction

A dynamic programming problem is specified by four objects : S, A, q, r, where S is a non-empty Borel subset of a Polish (i.e. complete, separable, metric) space, the set of states of some system, A is a non-empty Borel subset of a Polish space, the set of actions available to you, q is the law of motion of the system—it associates (Borel measurably) with each pair (s, a) a probability measure $q(.|s, a)$ on the Borel subsets of S : when the system is in state s and action a is chosen, the system moves to the state s' according to the distribution $q(.|s, a)$; and r is a bounded Borel measurable function on $S \times A$, the immediate return—when the system is in state s, choose action a, and receive an income $r(s, a)$. A *plan* π is a sequence $\pi_1, \pi_2, \ldots$, where π_n tells you how to select an action on the n-th day, as a function of the previous history $h = (s_1, a_1, \ldots, a_{n-1}, s_n)$ of the system, by associating with each h (Borel measurably) a probability distribution $\pi_n(.|h)$ on the Borel subsets of A.

A Borel function f from S into A defines a plan. When in state s, choose action $f(s)$ (independently of when and how you have arrived at state s). Denote the corresponding plan by $f^{(\infty)}$. Such plans will be called *stationary*.

A plan π associates with each initial state s a corresponding n-th day expected return $r_n(\pi)(s)$ and an expected discounted total return

$$I(\pi)(s) = \sum_{n=1}^{\infty} \beta^{n-1} r_n(\pi)(s),$$

where β is a fixed discount factor, $0 \leqslant \beta < 1$.

A plan π^* will be called *optimal* if $I(\pi^*)(s) \geqslant I(\pi)(s)$ for all plans π and $s \epsilon S$. The problem, then, is to find a n optimal plan.

Blackwell (1965) and Strauch (1966) have studied this problem extensively. Blackwell (1965) has given an example in which not even ε-optimal plans exist. In this paper, we are going to make additional assumptions about A, q and r. Throughout the paper, the following assumptions will remain operative : (a) A is a compact metric space; (b) r is a bounded upper semi-continuous function on $S \times A$; and (c) if $s_n \rightarrow s$, $a_n \rightarrow a$, then $q(\cdot|s_n, a_n)$ converges weakly to $q(.|s, a)$. We shall show that under these restrictions there will always exist a stationary optimal plan.

The proof of the existence of an optimal plan rests on a Selection theorem due to Dubins and Savage (1965, see Chapter 2.16). In Section 3 of this paper, we give a proof of the Dubins-Savage theorem, which is rather technical in nature (Section 3 is, therefore, expository. It seems worthwhile to include a detailed proof here). Section 2 is devoted to an exposition of certain topological notions needed for the Selection theorem. In Section 4, we establish the existence of an optimal stationary plan.

2. Topological prerequisites

Let Δ be a compact metric space with metric ρ. Denote by 2^{Δ} the collection of all non-empty closed subsets of Δ. We introduce a metric d on 2^{Δ}—the Hausdorff metric—as follows : for any $A, B \epsilon 2^{\Delta}$,

$$d(A, B) = \max \left(\sup_{x \epsilon A} \rho(x, B), \sup_{y \epsilon B} \rho(y, A) \right)$$

where, for $z \epsilon \Delta$ and $D \subset \Delta$, $\rho(z, D) = \inf_{z' \epsilon D} \rho(z, z')$.

Proposition 2.1 : *$(2^{\Delta}, d)$ is a compact metric space.*

The proof may be found in Kuratowski (1950, page 21).

Let us now study convergence in the Hausdorff metric d. For any sequence $\{A_n, n = 1, 2, ...\} \subset 2^{\Delta}$, define $\overline{\lim}\, A_n = \{p \epsilon \Delta$: there exists an increasing sequence $k_1 < k_2 < ... <$ of natural numbers such that $p_{k_n} \epsilon A_{k_n}$ and $p_{k_n} \rightarrow p\}$; define $\underline{\lim}\, A_n = \{p \epsilon \Delta$: there exists a sequence p_n such such that $p_n \epsilon A_n$ and $p_n \rightarrow p\}$. It is clear that $\overline{\lim}\, A_n$, $\underline{\lim}\, A_n$ are closed. In case $\overline{\lim}\, A_n = \underline{\lim}\, A_n$ we say that the limit exists and denote it by $\lim A_n$. As Δ is compact, it is clear that $\lim A_n \epsilon 2^{\Delta}$.

The next proposition connects convergence in the metric d with convergence defined above.

Proposition 2.2 : *Let A_n, $n = 1, 2, ...$ be a sequence* of elements of 2^{Δ}. *Then $d(A_n, A) \rightarrow 0$ if and only if $\lim A_n = A$.*

See Kuratowski (1950), page 21, for a proof.

Let X be a metric space and let F be a map from X into 2^{Δ}. We shall say that F is *upper semi-continuous in the sense of Kuratowski* (abbreviated, hereafter, by u.s.c. (K)) if $x_n \rightarrow x$ implies $\overline{\lim}\, F(x_n) \subset F(x)$.

We shall need the following fact about u.s.c. (K) maps.

Proposition 2.3 : If F is u.s.c. (K) from a metric space X into 2^{Δ}, then F is Borel measurable.

See Kuratowski (1950), page 38, for a proof.

3. Selection theorem

In this section, we prove the Selection theorem of Dubins and Savage. Throughout this section, A will be a compact metric space, S a Borel subset of a Polish space and v a bounded, upper semi-continuous (abbreviated, hereafter, by u.s.c.) function on A (that is, $a_n \to a$ implies $\limsup v(a_n) \leqslant v(a)$). Assume that $|v(a)| \leqslant M$ for all $a \epsilon A$.

We need some preliminary lemmas for the proof of the Selection theorem.

Lemma 3.1 : *Define* v^* *on* 2^A *into* R *by* $v^*(K) = \max_{a \epsilon K} v(a)$. *Then* v^* *is u.s.c.*

Proof: As v is u.s.c., and K compact, it follows from a well-known result that there exists $a_0 \epsilon K$ such that $v^*(K) = v(a_0)$.

Now suppose $K_n \to K$ and assume that for some $a_n \epsilon K_n$, $v^*(K_n) = v(a_n)$. Choose a subsequence $\{v(a_{n'})\}$ such that $v(a_{n'}) \to \limsup v^*(K_n)$. As A is compact, there exists a subsequence $\{a_{n''}\}$ of $\{a_{n'}\}$ such that $a_{n''} \to a$. It follows that $a \epsilon K$ and, since v is u.s.c., $\limsup v^*(K_n) = \lim_{n''} v(a_{n''}) \leqslant v(a) \leqslant v^*(K)$, which proves that v^* is u.s.c.

Lemma 3.2 : *For each* $K \epsilon 2^A$ *and* $x \epsilon [-M, M]$, *define* $\tilde{V}(K, x) = \{a \epsilon K : v(a) \geqslant x\}$. *Denote by dom* $\tilde{V}$ *the set* $\{(K, x) \epsilon 2^A \times [-M, M] : V(K, x) \neq \phi\}$. *Then dom* $\tilde{V}$ *is closed in* $2^A \times [-M, M]$ *and so a compact metric space. Furthermore,* $\tilde{V}$ *is u.s.c. (K) from dom* $\tilde{V}$ *into* 2^A.

Proof: v being u.s.c., for any real c, $\{v \geqslant c\}$ is closed in A. Hence $\tilde{V}(K, x)$, if non-empty, is an element of 2^A. Next, let us show that dom $\tilde{V}$ is closed. Let $(K_n, x_n) \epsilon$ dom $\tilde{V}$, $n = 1, 2, \ldots$ and suppose $(K_n, x_n) \to (K, x)$. Let $a_n \epsilon \tilde{V}(K_n, x_n)$, $n = 1, 2, \ldots,$. Since A is compact, there exists a subsequence $\{a_{n_k}\}$ of $\{a_n\}$ such that $a_{n_k} \to a$. Consequently, $a \epsilon K$ and $x = \lim x_{n_k} \leqslant \limsup v(a_{n_k}) \leqslant v(a)$, so that $a \epsilon \tilde{V}(K, x)$. Hence $(K, x) \epsilon$ dom $\tilde{V}$, which is, therefore, closed.

Finally, in order to prove that $\tilde{V}$ is u.s.c. (K), we have to show that $(K_n, x_n) \to (K, x)$, $\tilde{a}_n \epsilon \tilde{V}(K_n, x_n)$, $n = 1, 2, \ldots, a_n \to a$ imply that $a \epsilon \tilde{V}(K, x)$. Since $K_n \to K$, $a \epsilon K$. Consequently, since $x = \lim x_n \leqslant \limsup v(a_n) \leqslant v(a)$, $a \epsilon \tilde{V}(K, x)$. This completes the proof of Lemma 3.2.

Lemma 3.3 : *Define* V *on* 2^A *by* $V(K) = \{a \epsilon K : v(a) = v^*(K)\}$. *Then* V *is a Borel measurable map from* 2^A *into* 2^A.

Proof: As v is u.s.c., $V(K)$ is non-empty. Let us show that it is closed. Let $a_n \epsilon V(K)$, $n = 1, 2$, and suppose that $a_n \to a$. Then, since v is u.s.c. $v^*(K) = \limsup v(a_n) \leqslant v(a)$ and as K is closed, $a \epsilon K$. Consequently, $a \epsilon V(K)$. Hence V maps 2^A into 2^A.

Let $\mathcal{K}$ be a Borel subset of 2^A. Note that $V(K) = \tilde{V}(K, v^*(K))$. Consequently

$$\{K \epsilon 2^A : V(K) \epsilon \mathcal{K}\} = \text{proj}_{2^A} [\{(K, x) \epsilon \text{ dom } \tilde{V} : \tilde{V}(K, x) \epsilon \mathcal{K}\} \cap \{(K, x) : v^*(K) = x\}] \quad \ldots (1)$$

As v^* is u.s.c. (by Lemma 3.1), it is a Borel function from 2^A into $[-M, M]$ and hence its graph is a Borel set in $2^A \times [-M, M]$ (cf. Kuratowski, 1952, page 291). Also $\tilde{V}$ is u.s.c. (K) (by Lemma 3.2) and so it is a Borel map from dom $\tilde{V}$ into 2^A (see Proposition 2.3). Consequently, $\{(K, x) \epsilon \text{ dom } \tilde{V} : \tilde{V}(K, x) \epsilon K\}$ is a Borel subset of dom $\tilde{V}$, which being closed in $2^A \times [-M, M]$, the former set is a Borel subset of $2^A \times [-M, M]$ as well. Thus, the set within square brackets on the right-hand side of (1) is a Borel subset of $2^A \times [-M, M]$. Finally, projection being a continuous map and, moreover, 1—1 in this case, it follows by a well-known theorem of Lusin (cf. Kuratowski, 1952, page 396) that $\{K \epsilon 2^A : V(K) \epsilon K\}$ is a Borel subset of 2^A. Hence V is a Borel map. This completes the proof of Lemma 3.3.

Lemma 3.4 : *Let u be a bounded u.s.c. function on $S \times A$. Define $u^* : S \rightarrow R$ by $u^*(s) = \max_{a \epsilon A} u(s, a)$. Then u^* is u.s.c.*

Proof : As u is u.s.c., for fixed a, $u(s, .)$ is u.s.c. in a, so that $u^*(s)$ is well defined. Let $s_n \rightarrow s$ and suppose $u^*(s_n) = u(s_n, a_n)$, $n = 1, 2, \ldots$. Choose a subsequence $\{u^*(s_{n'})\}$ such that $u^*(s_{n'}) \rightarrow \lim\sup u^*(s_n)$. Moreover, as A is compact, there is a subsequence $\{a_{n''}\}$ of $\{a_{n'}\}$ such that $a_{n''} \rightarrow a$. Since u is u.s.c., it follows that $\lim\sup u^*(s_n) = \lim_{n''} u(s_{n''}, a_{n''}) \leqslant u(s, a) \leqslant u^*(s)$. Hence u^* is u.s.c.

Lemma 3.5 : *Let u be a bounded u.s.c. function on $S \times A$. Define $U : S \rightarrow 2^A$ by $U(s) = \{a \epsilon A : u(s, a) = \max_{a' \epsilon A} u(s, a')\}$. Then U is a Borel map.*

The proof of Lemma 3.5 is omitted as it is similar to that of Lemma 3.3.

Selection Theorem : *Let u be a bounded u.s.c. function on $S \times A$. Then there exists a Borel measurable map f from S into A such that $u(s, f(s)) = \max_{a \epsilon A} u(s, a)$ for all $s \epsilon S$.*

Proof : Choose a sequence $\{v_n, n = 1, 2, \ldots\}$ of continuous real-valued functions on A, which separate points in A (for instance, one may choose a sequence of functions dense in $C(A)$). For each v_i, define, for $K \epsilon 2^A$, $V_i(K) = \{a \epsilon K : v_i(a) = \max_{a' \epsilon K} v_i(a')\}$. Then by Lemma 3.3, each V_i is a Borel map from 2^A into 2^A. Let U be as in Lemma 3.5. Define $U_1(s) = V_1(U(s))$ and $U_n(s) = V_n(U_{n-1}(s))$, $n \geqslant 2$. By virtue of Lemma 3.5 it follows that each V_n is a Borel map from S into 2^A. Moreover, for each s, $U(s) \supset U_1(s) \supset U_2(s) \supset \ldots$. Consequently, the family $\{U_n(s) : n = 1, 2, \ldots\}$ of closed subsets of A has the finite intersection property, so that, as A is compact, $\bigcap_{n=1}^{\infty} U_n(s) \neq \phi$ for every $s \epsilon S$. Suppose now that for some $s \epsilon S$, $a, a' \epsilon \bigcap_{n=1}^{\infty} U_n(s)$ and $a \neq a'$. Then, for every n, as $a, a' \epsilon U_n(s)$, it follows that $v_n(a) = v_n(a')$, which contradicts the separating property of the sequence $\{v_n\}$. Hence $a = a'$ and for each s, $\bigcap_{n=1}^{\infty} U_n(s)$ is a singleton say, $\{f(s)\}$.

Next, let us show that for every $s\epsilon S$, $\{f(s)\} = \lim_{n\to\infty} U_n(s)$ in the Hausdorff metric of 2^A (see Section 2). Fix s and suppose that $a\epsilon \overline{\lim}\ U_n(s)$. Then there exists a sequence $\{a_{n_k}\}$ with $a_{n_k}\epsilon U_{n_k}(s)$ and $a_{n_k}\to a$. As each $U_m(s)$ is closed and $a_{n_k}\epsilon U_m(s)$ for all $n_k \geqslant m$, it follows that $a\epsilon U_m(s)$ and, consequently, $a\,\epsilon \bigcap_{n=1}^{\infty} U_n(s)$. Hence $\overline{\lim}\ U_n(s) \subset \{f(s)\}$. Also it is clear that $\{f(s)\} \subset \underline{\lim}\ U_n(s)$. Hence, by Proposition 2.2, $\{f(s)\} = \lim U_n(s)$ for each $s\epsilon S$. As each U_n is a Borel map from S into 2^A, it now follows that $\phi : s \to \{f(s)\}$ is a Borel measurable map from S into 2^A.

Finally, it is easy to check that the class of all singletons belonging to 2^A is isometric to A. It follows that f is a Borel measurable map from S into A. Moreover, as $f(s)\epsilon U(s)$ for each $s\epsilon S$, we get: $u(s, f(s)) = \max_{a\epsilon A} u(s, a)$ for every $s\epsilon S$. This completes the proof of the Selection theorem.

4. Existence of optimal plans

Let us return to the dynamic programming problem posed in Section 1. We shall assume that S is a Borel subset of a Polish space, A a compact metric space, r a bounded u.s.c. function on $S\times A$ and q is continuous, that is, $(s_n, a_n)\to(s, a)$ implies $q(.|s_n, a_n)$ converges weakly to $q(.|s, a)$.

Lemma 4.1 : *Let $w : S\to R$ be a bounded u.s.c. function. Then $g : S\times A \to R$ defined by $g(s, a) = \int w(.)dq(.|s.a)$ is u.s.c.*

Proof: If w is continuous, then clearly g is continuous. Now if w is bounded and u.s.c. there exists a sequence of bounded continuous functions $w_n \downarrow w$. Let $g_n(s, a) = \int w_n(.)dq(.|s, a)$, $(s, a)\,\epsilon\, S\times A$. Each g_n is continuous and, moreover, by the dominated convergence theorem, $g_n \downarrow g$. Hence g is u.s.c. This terminates the proof of Lemma 4.1.

Denote by $C_1(S)$ the class of all bounded u.s.c. functions on S. For $u, v\epsilon C_1(S)$, define $d_1(u, v) = \|u-v)\| = \sup_{s\epsilon S} |u(s)-v(s)|$. d_1 is a metric on $C_1(S)$.

Lemma 4.2 : *The metric space $(C_1(S), d_1)$ is complete.*

Proof: It suffices to show that $C_1(S)$ is closed under uniform convergence. Let $v_n\epsilon C_1(S)$, $n = 1, 2, \ldots$ and suppose v_n converges uniformly to v on S. Let $s_n\to s_0$. Given $\varepsilon > 0$, choose N_ε such such that $n \geqslant N_\varepsilon$ implies $|v_n(s)-v(s)| < \varepsilon$ for all $s\epsilon S$. Hence we have $v(s_n) < v_{N_\varepsilon}(s_n)+\varepsilon$ for all n and $v_{N_\varepsilon}(s_0) < v(s_0)+\varepsilon$. Consequently, $\limsup v(s_n) \leqslant \limsup v_{N}(s_n)+\varepsilon \leqslant v_{N_\varepsilon}(s_0)+\varepsilon < v(s_0)+2\varepsilon$. As ε is arbitrary, this proves that $v\,\epsilon\, C_1(S)$.

For every $w\,\epsilon\, C_1(S)$, let $Tw : S\to R$ be the function defined by:

$$(Tw)(s) = \max_{a\epsilon A} \left[\int (r(s, a)+\beta w(.))\, dq(.|s, a)\right]. \qquad \ldots \quad (2)$$

Note that, by virtue of Lemma 4.1, the expression within square brackets on the right-hand side of (2) is u.s.c. in s and a, and, consequently, the maximum is assumed for every s. Moreover, Lemma 3.4 implies that Tw is u.s.c. and, since it is obviously bounded, $Tw \in C_1(S)$. Thus T maps $C_1(S)$ into $C_1(S)$.

Lemma 4.3 : *T is a contraction mapping on $C_1(S)$ and, consequently, has a unique fixed point.*

Proof: Let $w_1, w_2 \in C_1(S)$. Clearly $w_1 \leqslant w_2+\|w_1-w_2\|$. Since, as is easy to check, T is monotone, $Tw_1 \leqslant T(w_2+\|w_1-w_2\|) = Tw_2+\beta\|w_1-w_2\|$. Consequently, $Tw_1-Tw_2 \leqslant \beta\|w_1-w_2\|$. Interchanging w_1 and w_2, we get $Tw_2-Tw_1 \leqslant \beta\|w_1-w_2\|$. Hence $\|Tw_1-Tw_2\| \leqslant \beta\|w_1-w_2\|$, which proves that T is a contraction mapping, as $\beta < 1$. Since $C_1(S)$ is a complete metric space (Lemma 4.2), it follows from the Banach Fixed Point theorem that T has a unique fixed point in $C_1(S)$. This completes the proof.

Theorem : *There exists a stationary optimal plan and the optimal return $(= \sup\limits_{\pi} I(\pi))$ is u.s.c. on S.*

Proof: With each Borel map $g : S \to A$, associate the operator $L(g)$ on $M(S)$ (= the collection of all bounded Borel measurable functions on S) which sends $u \in M(S)$ into $L(g)u \in M(S)$, where $L(g)u$ is defined by :

$$(L(g)u)(s) = \int [r(s, g(s))+\beta u(.)]dq(.\,|\,s, g(s)), s \in S. \qquad \ldots \quad (3)$$

It is known that $I(g^{(\infty)})$ is the unique fixed point of the operator $L(g)$ (see Theorem 5.1 (b) in Strauch (1965).

Now let T be as above and let $w^* \in C_1(S)$ be its unique fixed point (Lemma 4.3), i.e., $Tw^* = w^*$. It now follows from the Selection Theorem (of Section 3) that there exists a Borel map f from S into A such that $Tw^* = L(f)w^*$. Consequently $L(f)w^* = w^*$ so that by the remark made in the preceding paragraph, $w^* = I(f^{(\infty)})$. Hence $Tw^* = w^*$ can be rewritten as :

$$I(f^{(\infty)})(s) = \max_{a \in A} \left[\int (r(s, a)+\beta I(f^{(\infty)})(.))dq(.\,|\,s, a)\right] s \in S.$$

Thus $I(f^{(\infty)})$ satisfies the optimality equation, so that by a theorem of Blackwell (1965, Theorem 6(f)) $f^{(\infty)}$ is an optimal plan. Moreover, as $w^* = I(f^{(\infty)})$ and $w^* \in C_1(S)$, it follows that the optimal return is u.s.c. This completes the proof of the theorem.

ACKNOWLEDGMENTS

I am grateful to Messers N. D. Prabhakar and A. B. Raha for useful discussions.

REFERENCES

BLACKWELL, D. (1965) : Discounted dynamic programming. *Ann. Math. Stat.*, **36**, 226-235.

DUBINS, L. E. and SAVAGE, L. J. (1965) : *How to Gamble If You Must*, McGraw-Hill, New York.

KURATOWSKI, C. (1952) : *Topologie*, **1**, Warsaw, Poland.

——— (1950) : *Topologie*, **2**, Warsaw, Poland.

STRAUCH, R. E. (1966) : Negative dynamic programming. *Ann. Math. Stat.*, **37**, 871-890.

Paper received: December, 1967.