MATHEMATICS OF OPERATIONS RESEARCH
Vol. 17, No. 2, May 1992
*Printed in U.S.A.*

# A UNIFORM TAUBERIAN THEOREM IN DYNAMIC PROGRAMMING*

EHUD LEHRER AND SYLVAIN SORIN

We prove that, in dynamic programming framework, uniform convergence of $v_\lambda$ implies uniform convergence of $v_n$ and vice versa. Moreover, both have the same limit.

**1. Introduction.** A deterministic dynamic programming problem is defined by a set of states $S$, a (nonvoid) correspondence $\Gamma$ from $S$ to itself and a bounded real function on $S$, say with values in $[0, 1]$.

Given the state $s$, one chooses $t$ in $\Gamma(s)$ and gets a payoff of $f(s)$. A *strategy* is such a sequence of (history dependent) choices at each stage $n = 0, 1, \dots$. Any strategy induces a *play* at $s$, i.e., a sequence $h = (s = s_0, s_1, \dots, s_n, \dots)$ with $s_{n+1} \in \Gamma(s_n)$. Each play $h$ induces an $n$-average payoff $f_n(h) = (1/n)\sum_{m=0}^{n-1} f(s_m)$ and a $\lambda$-discounted payoff $f_\lambda(h) = (1-\lambda)\sum_{m=0}^{\infty}\lambda^m f(s^m)$. Taking the supremum on all strategies of the above functions defines the $n$-stage value $v_n(s)$ and the $\lambda$-discounted value $v_\lambda(s)$.

We will consider here the asymptotic behavior of these two families of functions (as $n \to \infty$ or $\lambda \to 1$) and prove that the uniform convergence of one implies the uniform convergence of the other, and both to the same limit. Note that without the uniformity condition $\lim v_n$ and $\lim v_\lambda$ may exist and differ (see Example (§2)). Moreover, this condition does not imply the equality with $v_\infty$ (defined through $f_\infty(h) = \liminf f_n(h)$) (Lehrer and Monderer 1989).

The proofs and the result extend easily to the general (stochastic) case (see §6).

Formally, let

(A) $v_\lambda \to_{\lambda \to 1} v$, uniformly on $S$ and

(B) $v_n \to_{n \to \infty} w$, uniformly on $S$.

The purpose of this paper is to prove the following.

THEOREM. (a) *If* (A) *then* (B) *and* $v = w$;
(b) *If* (B) *then* (A) *and* $v = w$.

**2. Example.** Take $S = \mathbb{N}^* \times \mathbb{N}$, where $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$, $\Gamma(n, 0) = \{(n+1, 0), (n, 1)\}$ and $\Gamma(n, m) = \{(n, m+1)\}$ for $m > 0$. $f(n, m) = 1$ iff $1 \le m \le n$ and 0 otherwise. In words, at each state $(n, 0)$ either you choose to get 1 for the next $n$ stages and then always 0 or you proceed to state $(n+1, 0)$ and get 0 at that stage.

Let $s = (1, 0)$. The feasible sequences of payoffs are of the form: $n$ times 0, $n$ times 1 and then only 0 (say, on a play $h_n$ at $s$); or always 0. Obviously, $\lim v_n(s) = \frac{1}{2}$. $f_\lambda(h_n) = \lambda^n - \lambda^{2n}$; hence $\lim v_\lambda(s) = \frac{1}{4}$ and, finally, $v_\infty(s) = 0$.

*Received September 1, 1989; revised July 27, 1990.
*AMS 1980 subject classification.* Primary: 90C39.
*IAOR 1973 subject classification.* Main: Programming: Dynamic.
*OR/MS Index 1978 subject classification.* Primary: 111 Dynamic Programming.
*Key words.* Deterministic dynamic programming, discounted value function, $n$-stage value function, asymptotic behaviors of value functions.



0364-765X/92/1702/0303/$01.25

**3. Preliminary results.** Let us begin by proving properties that hold in a general framework. The first one shows that lim sup is decreasing on plays.

PROPOSITION 1. *For any play $h = (s_m)$ at s one has*

$$\limsup v_n(s_m) \le \limsup v_n(s) \quad \textit{and}$$

$$\limsup v_\lambda(s_m) \le \limsup v_\lambda(s) \quad \textit{for all } m.$$

PROOF. Given $m$, choose $N > 2m/\epsilon$ and a play at $s_m$, $h'$, satisfying $f_n(h') \ge \limsup v_n(s_m) - \epsilon/2$ with $n \ge N$. Then

$$f_{n+N}(s_0, s_1, \dots, s_{m-1}, h') \ge \limsup v_n(s_m) - \epsilon.$$

Similarly, in the discounted case, let $\lambda_0$ with $\lambda_0^m > (1 - \epsilon/2)$ and $h''$ at $s_m$ satisfying $f_\lambda(h'') \ge \limsup v_\lambda(s_m) - \epsilon/2$ with $\lambda \ge \lambda_0$. Then

$$f_\lambda(s_0, \dots, s_{m-1}, h'') \ge \limsup v_\lambda(s_m) - \epsilon. \quad //$$

The second property is that, given an $n$ average on a play, there exists a state (on this play) from which all averages of small length (compared to $n$) give at least this amount. Note that this result may be useful for related problems (cf. Lehrer and Monderer 1989).

PROPOSITION 2. *Let $\epsilon > 0$. For all s, n there exist a play $h = (s_l)$ at s and a stage L such that*:

$$(1/T) \sum_{m=0}^{T-1} f(s_{L+m}) \ge v_n(s) - \epsilon \quad \textit{for all } 1 \le T \le [n\epsilon/2].$$

PROOF. Otherwise there exist $s$ and $n$ such that on each play $h$ at $s$ and each stage $L$, there is some $T = T(h, L)$ such that:

$$1 \le T \le [n\epsilon/2] \quad \text{and} \quad (1/T) \sum_{m=0}^{T-1} f(s_{L+m}) < v_n(s) - \epsilon,$$

in particular, if $h$ is a play at $s$ satisfying $f_n(h) \ge v_n(s) - \epsilon/2$. But then we can divide this play into segments of length at most $[n\epsilon/2]$ on each of which, except at most the last one, the average payoff is less than $v_n(s) - \epsilon$. (Taking $L_0 = 0$, then $L_1 = T(h, L_0)$, $L_2 = T(h, L_1)$, and so on.) It follows that $f_n(h) \le v_n - \epsilon + \epsilon/2$, a contradiction. //

We now compare $v_n$ and $v_\lambda$. Recall that any normalized power series with parameter in [0, 1) can be written also as a convex combination of the finite averages. Since we will need it later, we provide here the explicit formula.

If $\{a_m\}$ is a bounded sequence and $0 \le \lambda < 1$, then for all $n \in \mathbb{N} \cup \{+\infty\}$

$$(1) \qquad (1 - \lambda) \sum_{m=0}^{n} a_m \lambda^m = (1 - \lambda)^2 \sum_{m=0}^{n-1} \lambda^m (m + 1) \left( \sum_{l=0}^{m} a_l/(m + 1) \right)$$
$$+ (1 - \lambda) \lambda^n (n + 1) \left( \sum_{l=0}^{n} a_l/(n + 1) \right).$$

This relation will allow us to examine the asymptotic behavior of families of geometric distributions. We start with the following simple observation.

LEMMA 3. *Let* $M(\alpha, \beta; \lambda) = (1 - \lambda)^2 \Sigma_\alpha^\beta \lambda^m (m + 1)$:
(i) *There exist* $N_0$ *and* $\epsilon_0$ *such that* $\forall n \geq N_0$, $\forall \epsilon \leq \epsilon_0$,

$$M([(1 - \epsilon)n], n; 1 - 1/n) \geq \epsilon/2e. \tag{2}$$

(ii) $\forall \delta > 0$, *there exist* $N_0$ *and* $\epsilon_0$ *such that* $\forall n \geq N_0$, $\forall \epsilon \leq \epsilon_0$,

$$M([\epsilon n], [(1 - \epsilon)n]; 1 - 1/n\sqrt{\epsilon}) \geq 1 - \delta. \tag{3}$$

PROOF. Use (1) with $n = +\infty$ to get

$$M(\alpha, \beta; \lambda) = ((\alpha + 1)\lambda^\alpha - \alpha\lambda^{\alpha+1}) - ((\beta + 2)\lambda^{\beta+1} - (\beta + 1)\lambda^{\beta+2}).$$

For $\alpha = [(1 - \epsilon)n]$, $\beta = n$ and $\lambda = 1 - 1/n$ the first term is of the order of $(2 - \epsilon)\exp(-1 + \epsilon)$ and the second of $2/e$. For $\alpha = [\epsilon n]$, $\beta = [(1 - \epsilon)n]$ and $\lambda = 1 - 1/n\sqrt{\epsilon}$ we obtain, respectively, $(1 + \sqrt{\epsilon})\exp(-\sqrt{\epsilon})$ and $1/\sqrt{\epsilon}\exp(-1/\sqrt{\epsilon})$, providing that $n$ is much larger than $1/\epsilon$. //

PROPOSITION 4. $\forall \epsilon > 0$, $\forall N$, *there is* $\lambda_0$ *such that for all* $\lambda \geq \lambda_0$ *and all* $s$ *in* $S$ *there exists* $n \geq N$ *with* $v_n(s) \geq v_\lambda(s) - \epsilon$.

PROOF. Given $\epsilon > 0$ and $N$ let $\lambda_0$ be such that:

$$(1 - \lambda)^2 \sum_{m=0}^{N-1} \lambda^m (m + 1) < \epsilon/2 \quad \text{for } \lambda \geq \lambda_0.$$

By (1) this implies that on an $\epsilon/2$ optimal play at $s$ for $v_\lambda$, say $h$, there exists some $n \geq N$ with $f_n(h) \geq v_\lambda(s) - \epsilon$. //

COROLLARY 5. $\limsup v_n \geq \limsup v_\lambda$.

**4. Proof of part (a).** We assume (A).

LEMMA 6. $\forall \epsilon > 0$, *there is an* $N$ *such that* $n \geq N$ *implies* $v_n \leq v + \epsilon$.

PROOF. Otherwise, let $\epsilon > 0$ such that for all $N$ there exist $n \geq N$ and $s$ in $S$ with $v_n(s) > v(s) + \epsilon$. Let $\lambda$ such that $\|v_\lambda - v\| \leq \epsilon/8$ (by (A)) and $N$ such that

$$(1 - \lambda)^2 \sum_{m=0}^{[n\epsilon/4]-1} \lambda^m (m + 1) \geq 1 - \epsilon/8 \quad \text{for } n \geq N.$$

We now use Proposition 2 with $\epsilon/2$ to get a play $h$ at $s$ and a stage $L$ with:

$$(1/T) \sum_{m=0}^{T-1} f(s_{L+m}) \geq v_n(s) - \epsilon/2 > v(s) + \epsilon/2 \quad \text{for all } 1 \leq T \leq [n\epsilon/4].$$

(1) then implies that

$$v_\lambda(s_L) \geq v(s) + \epsilon/2 - \epsilon/8.$$

Hence,

$$v(s_L) \geq v(s) + \epsilon/2 - \epsilon/8 - \epsilon/8,$$

a contradiction to Proposition 1. //

LEMMA 7. *$\forall \epsilon > 0$, there is an $N$ such that $n \geq N$ implies $v_n \geq v - \epsilon$.*

PROOF. Otherwise, let $\epsilon > 0$ such that for all $N$ there exist $n \geq N$ and $s$ with $v_n(s) < v(s) - \epsilon$. In particular, for $N$ large enough, on any play $h$ at $s$ one has:

$$(1/T) \sum_{m=0}^{T-1} f(s_m) \leq v(s) - \epsilon/2 \quad \text{for } [(1-\epsilon/2)n] \leq T \leq n.$$

Choose $N$ large enough and $\epsilon$ small enough to get that the weight of these stages is at least $\delta = \epsilon/4e$ for $\lambda = 1 - 1/n$; i.e., $M([(1-\epsilon/2)n], n; 1 - 1/n) \geq \epsilon/4e$ (by (2) in Lemma 3). Finally, let $K$ be such that $v_K(s) \leq v(s) + \delta\epsilon/8$ for $n \geq K$, by Lemma 6. Choose $N$ large enough to guarantee for $n \geq N$ and $\lambda = 1 - 1/n$: $\|v_\lambda - v\| < \delta\epsilon/5$ (by (A)), and furthermore, $(1-\lambda)^2 \Sigma_{m=0}^{K-1} \lambda^m (m+1) < \epsilon\delta/8$. This implies by (1) that:

$$\begin{aligned} f_\lambda(h) &\leq \epsilon\delta/8 + \delta(v(s) - \epsilon/2) + (1 - \delta - \epsilon\delta/8)(v(s) + \delta\epsilon/8) \\ &\leq v(s) - \delta\epsilon/4, \end{aligned}$$

a contradiction. //

Lemmas 6 and 7 give (a).

REMARK. Notice that the previous proof shows also that uniform convergence of a sequence $V_{\lambda_i}$, where $\lambda_i \to 1$, implies uniform convergence of the sequence $V_{n_i}$, where $n_i = [1/(1-\lambda_i)]$. Moreover, both converge to the same limit.

**5. Proof of part (b).** We assume (B).

LEMMA 8. *For any $\epsilon > 0$ small enough, there exists $N$ such that for all $n \geq N$ and all $s$, there is a play $h = (s_l)$ at $s$ satisfying:*

$$(1/T) \sum_{m=0}^{T-1} f(s_m) \geqslant w(s) - \epsilon \quad \textit{for all } [\epsilon n] \leq T \leq [(1-\epsilon)n].$$

PROOF. Use (B) to get $N$ such that $\|v_n - w\| \leq \delta$ for $n \geq [\epsilon N]$ with $\delta = \epsilon^2/3$. Given $n \geq N$ let $h = (s_l)$ at $s$ with $f_n(h) \geq v_n(s) - \delta$. For $T \leq [(1-\epsilon)n]$ we obtain on $h$:

$$v_{n-T}(s_T) \leq w(s_T) + \delta \leq w(s) + \delta$$

by Proposition 1. Thus,

$$n(v_n(s) - \delta) \leq \sum_{m=0}^{T-1} f(s_m) + (n-T)(w(s) + \delta).$$

Hence,

$$\begin{aligned} (1/T) \sum_{m=0}^{T-1} f(s_m) &\geq w(s) - n/T \cdot 3\delta \\ &\geq w(s) - \epsilon \quad \text{for } T \geq [\epsilon n]. \quad // \end{aligned}$$

LEMMA 9. *$\forall\delta > 0$, there exists $\lambda_0$ such that $\lambda \geq \lambda_0$ implies $v_\lambda \geq w - \delta$.*

PROOF. Choose $\epsilon_0$ and $N_0$ as in (3) (Lemma 3) with $\delta/3$. Then use Lemma 8 to get with $\epsilon \leq \delta/3$ for any $n$ large enough and any $s$, the existence of a play $h$ at $s$ with:

$$f_\lambda(h) \geq (1-\lambda)^2 \sum_{T=[\epsilon n]}^{[(1-\epsilon)n]} \lambda^T (T+1)\left((1/T)\sum_{m=0}^{T-1} f(s_m)\right)$$

$$\geq (1-\delta/3)(w-\delta/3) \quad \text{for } \lambda = \lambda_n = 1 - 1/n\sqrt{\epsilon}.$$

Note that $\lambda_n \leq \lambda \leq \lambda_{n+1}$ implies

$$f_{\lambda_n}(h)/(1-\lambda_n) \leq f_\lambda(h)/(1-\lambda) \leq f_{\lambda_{n+1}}(h)/(1-\lambda_{n+1}).$$

Hence, the result for $\lambda$ large enough. //

LEMMA 10. *$\forall\delta > 0$, there exists $\lambda_0$ such that $\lambda \geq \lambda_0$ implies $v_\lambda \leq w + \delta$.*

PROOF. Follows from Proposition 4, using (B). //

**6. Comments.** The result extends to the stochastic case as follows. Consider first a countable-Borel framework where $S$ is a countable set of states, $A$ is a Borel set of actions, $q$ is a transition probability from $S \times A$ to $S$ and $f$ is a measurable bounded payoff function from $S \times A$ to $\mathbb{R}$. (Recall that if $S$ and $A$ are finite, $v$, $w$ and $v_\infty$ exist, are equal and realized with a pure stationary strategy (Blackwell 1962).)

Define, for any Markov strategy $\sigma$ (Blackwell 1965), a play starting from $s$ by a sequence $\{w_n\}$ of probabilities on $S$ with $w_0 = \delta_s$, $w_{n+1}(S') = \int_S q(S' \mid t, \sigma_n(t)) w_n(dt)$, for all Borel sets $S' \subset S$. The corresponding sequence of payoffs is $\{x_n\}$ with

$$x_n = \int_S f(t, \sigma_n(t)) w_n(dt).$$

The proof goes then word-for-word.

If one leaves the countable state set up, a selection theorem is needed in Proposition 1. Hence, one can use an analytic framework (Blackwell, Freedman and Orkin 1974), where plays are defined by strategies.

## References

Blackwell, D. (1962). Discrete Dynamic Programming. *Ann. Math. Statist.* **33** 719–726.

______ (1965). Discounted Dynamic Programming. *Ann. Math. Statist.* **36** 226–235.

______, Freedman, D. and Orkin, M. (1974). The Optimal Reward Operator in Dynamic Programming. *Ann. Probab.* **2** 926–941.

Lehrer, E. and Monderer, D. (1989). Discounting Versus Averaging in Dynamic Programming. Northwestern University Discussion Paper No. 837; *Games and Economic Behavior* (to appear).

LEHRER: DEPARTMENT OF MATHEMATICS AND DEPARTMENT OF MANAGERIAL ECONOMICS AND DECISION SCIENCE, J. L. KELLOGG GRADUATE SCHOOL OF MANAGEMENT, NORTHWESTERN UNIVERSITY, 2001 SHERIDAN ROAD, EVANSTON, ILLINOIS 60208

SORIN: DÉPARTEMENT DE MATHÉMATIQUE, UNIVERSITÉ LOUIS PASTEUR, 7 RUE RENÉ DESCARTES, 67084 STRASBOURG, FRANCE