Stochastic Processes and their Applications 8 (1978) 59–76.

MARKOV DECISION PROCESSES AND STRONGLY EXCESSIVE FUNCTIONS*

K.M. van HEE and J. WESSELS
Department of Mathematics, Eindhoven University of Technology, The Netherlands

Received 10 March 1976
Revised 12 April 1978

Strongly excessive functions play an important role in the theory of Markov decision processes and Markov games. In this paper the following question is investigated: What are the properties of Markov decision processes which possess a strongly excessive function? A probabilistic characterization is presented in the form of a random drift through a partitioned state space. For strongly excessive functions which have a positive lower bound a characterization is given in terms of the lifetime distribution of the process.

Finally we give a characterization in terms of the spectral radius.

Markov decision process	excessive function
transient behaviour	exponentially bounded stopping time
spectral radius	

1. Introduction

When analyzing (semi-) Markov decision processes and Markov games one often applies contraction properties of certain operators in a Banach space. This technique has been introduced by Blackwell [1], using 1) the boundedness of the immediate return in the supremum norm and 2) discounting, which is equivalent to a positive probability β of leaving the system in each state (for all strategies). The idea has been generalized by Denardo [2] who weakened the discounting condition by assuming a positive probability of leaving the system in N stages (uniform in the starting state and the strategy).

In order to obtain weaker conditions other norms might be used. Norms which appear to be useful are of the weighted supremum norm type. First attempts in this direction have been made by Veinott [12] in case of a finite state space and by Lippman [7] using a polynomial and a special condition on the transition probabilities. More general approaches have been presented by Lippman [8] and by one of the present authors [14]. Hinderer [5] uses a similar technique as [14] for finite stage programs. Wijngaard [15, Chapter 5] uses weighted supremum

* Memorandum COSOR 77-11; revised and extended version of Memorandum COSOR 75-22.

norms (with an exponential weight function) for analyzing average costs inventory problems.

In this paper we will investigate some properties of the decision processes when the transition probabilities satisfy the conditions imposed by the weighted supremum norm approach. These conditions may be formulated (see below) as the existence of a function on the state space which is excessive in a somewhat stronger sense than usual (compare Hordijk [6]).

In Section 3 we give a characterization of the existence of a strongly excessive function in the form of a random drift through a partition of the state space and in Section 4 these properties are related to the lifetime distribution of the process.

Further we give in Section 5, an analytic equivalent for the existence of a strongly excessive function in terms of the spectral radius of the decision process.

A *Markov decision process* is determined by a pair $(E, \mathcal{P})$ where E is called the *state space* (supposed to be countable in this paper), $\mathcal{P}$ is a set of sub-Markov matrices ($P \in \mathcal{P}$ is a nonnegative function on $E \times E$ with $\sum_{j \in E} P(i,j) \leq 1$ for all $i \in E$). It is usual to define a Markov decision process as a triple $(E, \mathcal{P}, r)$ where r is a real function on $E \times \mathcal{P}$ with the interpretation of a reward function. However in this paper we are only dealing with the state space E and the transition probabilities $\mathcal{P}$.

At this moment we do not require any structure on $\mathcal{P}$ but in Section 2 we make an assumption which is fulfilled if we are dealing with the usual law of motion of a Markov decision process. Consider a positive function μ on E and introduce the Banach space V_μ of all real valued functions v on E which satisfy

$$\|v\|_\mu := \sup_{i \in E} \frac{|v(i)|}{\mu(i)} < \infty.$$

We call $\|\cdot\|_\mu$ the *weighted supremum norm* and μ a *bounding function*. The norm concept in V_μ induces a norm for the matrices $P \in \mathcal{P}$, viz. the operator norm

$$\|P\|_\mu := \sup_{i \in E} \mu(i)^{-1} \sum_{j \in E} P(i,j)\mu(j).$$

Definition 1. A Markov decision process $(E, \mathcal{P})$ is called *contracting* if there exists a bounding function μ on E and a number $0 < \rho < 1$ such that

$$\|P\|_\mu \leq \rho \quad \text{for all } P \in \mathcal{P}.$$

In [14] the *contracting dynamic programming* model is analysed extensively. Note that $\|P\|_\mu \leq \rho < 1$ for all $P \in \mathcal{P}$ is equivalent to

$$P\mu(i) := \sum_{j \in E} P(i,j)\mu(j) \leq \rho\mu(i) \quad \text{for all } P \in \mathcal{P}.$$

For $\rho = 1$ this condition becomes the usual requirement for excessivity of the function μ (see Hordijk [6]). So our condition is stronger.

Definition 2. A bounding function v on E is called *strongly excessive* with respect to $(E, \mathscr{P})$ if there is a number ρ $(0<\rho<1)$ with $Pv \leqslant \rho v$ for all $P \in \mathscr{P}$. The number ρ is called an *excessivity factor.*

Remark. The contracting condition may be used in the total expected reward case and in the total expected discounted reward case, viz. if in the discounted case Q is a transition matrix we define a matrix P as βQ where β is the discount factor $(0 \leqslant \beta < 1)$.

In the same way discounted semi-Markov decision processes may be handled by defining

$$\beta_Q(i,j) := \int_{0^-}^{\infty} e^{-\alpha t}\, dF_Q(t; i,j), \qquad P(i,j) := \beta_Q(i,j)Q(i,j).$$

2. Prerequisites and notations

A *Markov strategy* R is a sequence $(P_0, P_1, \ldots)$ of elements of $\mathscr{P}$. The set of all Markov strategies is denoted by $\mathscr{M}$.

Since the matrices $P \in \mathscr{P}$ are supposed to be sub-Markov, it might be desirable to extend the state space E to $\bar{E}$ by adding a new state x in the following way: $P(x,x) := 1$, $P(i,x) := 1 - \sum_{j\in E} P(i,j)$ for all $i \in E$, for all $P \in \mathscr{P}$.

All functions on E are extended to functions on $\bar{E}$ by defining them 0 in x. (Note that a strongly excessive function v on $(E, \mathscr{P})$ with excessivity factor ρ satisfies $\rho v \geqslant Pv$ on $\bar{E}$ for all $P \in \mathscr{P}$).

Any starting state $i \in \bar{E}$ and any $R \in \mathscr{M}$ determine a (nonhomogeneous)-Markov chain on $\bar{E}$ and so a probability $\mathbf{P}_{i,R}$ on $(\bar{E})^\infty$ with transition matrix P_n at time n. The random variable X_n denotes the state at time n. Let $\mathbf{E}_{i,R}$ be the expectation with respect to $\mathbf{P}_{i,R}$.

The following lemmas will be used in Section 3.

Lemma 1. *Suppose $A \subset E$, $i \in E$, $R_0 \in \mathscr{M}$. Then*

$$\sum_{n=0}^{\infty} \mathbf{P}_{i,R_0}[X_n \in A] \leqslant \sup_R \sup_{j\in A} \sum_{m=0}^{\infty} \mathbf{P}_{j,R}[X_m \in A].$$

Proof. Suppose $i \notin A$, since for $i \in A$ the assertion is trivial. Let $R_0 = (P_0, P_1, \ldots)$ and define $R_k := (P_k, P_{k+1}, \ldots)$. It is easy to verify that:

$$\mathbf{P}_{i,R_0}[X_n \in A \mid X_k = j] = \mathbf{P}_{j,R_k}[X_{n-k} \in A] \quad \text{for } k = 1, \ldots, n \text{ and } j \in A.$$

Note that

$$\sum_{n=0}^{\infty} \mathbf{P}_{i,R_0}[X_n \in A]$$

$$= \sum_{n=0}^{\infty} \sum_{k=1}^{n} \sum_{j\in A} \mathbf{P}_{i,R_0}[X_n \in A \mid X_k = j]\mathbf{P}_{i,R_0}[X_k = j, X_l \notin A (1 \leqslant l < k)].$$

Hence

$$\sum_{n=0}^{\infty} \mathbf{P}_{i,R_0}[X_n \in A]$$

$$\leq \sum_{k=1}^{\infty} \sum_{j \in A} \mathbf{P}_{i,R_0}[X_k = j, X_l \notin A (1 \leq l < k)] \sup_{j \in A} \sup_{R} \sum_{n=k}^{\infty} \mathbf{P}_{j,R}[X_{n-k} \in A]$$

$$\leq \sup_{R} \sup_{j \in A} \sum_{n=0}^{\infty} \mathbf{P}_{j,R}[X_n \in A].$$

Lemma 2. *Suppose $(E, \mathcal{P})$ is contracting with excessivity factor $\rho < 1$. Suppose $1 < \alpha < \rho^{-1}$. Then there exists a strongly excessive function c for $(E, \mathcal{P})$ with excessivity factor $\alpha\rho$, such that c maps E into the set $\{\alpha^l \mid l \in \mathbf{Z}\}$.*[1]

Proof. Let b be strongly excessive for $(E, \mathcal{P})$ with excessivity factor ρ. Define on E the function c: $c(i) := \alpha^l$ if $\alpha^{l-1} < b(i) \leq \alpha^l$, $i \in E$. Then $\alpha^{-1}c < b \leq c$. Hence $Pc \leq \alpha Pb \leq \alpha\rho b \leq \alpha\rho c$.

Lemma 2 allows us to consider only exponentials as candidates for strongly excessive functions. If there is a strongly excessive function that is bounded away from zero the same holds for equidistant grids:

Lemma 3. *Suppose b is a strongly excessive function for $(E, \mathcal{P})$ satisfying $0 < \delta \leq b(i)$ for $i \in E$. Suppose $\rho + \alpha\delta^{-1} < 1$, where $\alpha > 0$, and ρ is an excessivity factor for b. Then there exists a strongly excessive function c for $(E, \mathcal{P})$ with excessivity factor $\rho + \alpha\delta^{-1}$, such that c maps E into the set $\{\delta + l\alpha \mid l = 1, 2, \ldots\}$.*

Proof. Define $c(i) = \delta + l\alpha$ if $\delta + (l-1)\alpha \leq b(i) < \delta + l\alpha$. Hence $c(i) - \alpha \leq b(i) < c(i)$, which implies

$$Pc \leq P(b + \alpha e) \leq \rho b + \alpha e \leq \rho c + \alpha e \leq \rho c + \alpha\delta^{-1}c = (\rho + \alpha\delta^{-1})c$$

where $e(i) = 1$ for all $i \in E$ and where we used $c \geq \delta e$.

We now introduce an assumption for $\mathcal{P}$ which will be supposed to hold throughout the rest of this paper.

Assumption. Let $P_1, P_2, \cdots \in \mathcal{P}$ and let $A_1, A_2, \ldots$ be a partition of E. Then P defined by

$$P(i, \cdot) := P_j(i, \cdot), \quad \text{if } i \in A_j$$

is also an element of $\mathcal{P}$.

[1] $\mathbf{Z}$ is the set of integers.

This assumption is fulfilled if and only if there is for each $i \in E$ a collection Q_i of (defective) probabilities on E, i.e. $Q_i \subset \{q \mid q : E \to [0, 1], \sum_{j \in E} q(j) \leq 1\}$ for all $i \in E$, such that $P \in \mathscr{P}$ if and only if $P(i, \cdot) \in Q_i$ for all $i \in E$.

The set Q_i can be considered as the set of available actions in state i.

In the sequel we will frequently consider a function v on E defined by

$$v(i) := \sup_R \mathbf{E}_{i,R}\left[\sum_{n=0}^{\infty} \lambda^n r(X_n)\right]$$

for some positive number λ and some nonnegative function r on E. If $v < \infty$, then it may be shown that v satisfies Bellman's optimality principle:

$$v \geq r + \lambda P v \quad \text{for all } P \in \mathscr{P}.$$

The proof of this property is completely analogous to the proof of Theorem 6.1 in Ross [10]. For this proof one only needs, that v is not only the supremum over all Markov strategies, but also over a more general class of strategies. The equality of both suprema has been proved in [4], under more general conditions.

Note that $\lambda > 1$ is allowed here.

The following lemma gives the two forms in which we will use the property $v \geq r + \lambda P v$ for the construction of strongly excessive functions.

Lemma 4. *Consider a Markov decision process $(E, \mathscr{P})$. If there is a function $r \geq 0$ on E such that one of the following conditions holds for all $i \in E$*

(i) $v(i) := \sup_R \mathbf{E}_{i,R}[\sum_{n=0}^{\infty} r(X_n)] < \infty$, $r(i) \geq (1-\rho) v(i)$ *and* $v(i) > 0$ *for some* $0 < \rho < 1$,

(ii) $z(i) := \sup_R \mathbf{E}_{i,R}[\sum_{n=0}^{\infty} \lambda^n r(X_n)] < \infty$ *and* $z(i) > 0$, *for some* $\lambda > 1$, *then there is a strongly excessive function for* $(E, \mathscr{P})$.

To prove this note that, by Bellman's optimality principle $v \geq r + Pv$. Hence $v \geq (1-\rho)v + Pv$ and therefore $\rho v \geq Pv$ if (i) is true. Further $z \geq r + \lambda P z$ and so $\lambda^{-1} z \geq Pz$, if (ii) holds.

In the next sections we shall search for functions $r \geq 0$ on E satisfying one of the conditions (i) and (ii) of Lemma 4 to prove the existence of strongly excessive functions.

We conclude this section with some conventions. Recall that the unit function on E is denoted by e, so $e(i) = 1$ for all $i \in E$. The characteristic function of some set A is denoted by χ_A, so $\chi_A(i) = 1$ if $i \in A$ and 0 otherwise.

3. Probabilistic equivalent for strong excessivity

In this section it is shown that strong excessivity is related to certain drifting properties of the Markov chain involved.

Theorem 1. *$(E, \mathcal{P})$ is contracting, if and only if there exist a partition $\{E_k \mid k \in \mathbf{Z}\}$ of E and numbers $\alpha > 1$, $\beta > 1$, such that for all $R \in \mathcal{M}$*

$$\sum_{n=0}^{\infty} \mathbf{P}_{i,R}[X_n \in E_k] \leq \beta \min\{1, \alpha^{l-k}\} \quad \text{if } i \in E_l.$$

Proof. We first prove the "if"-part. Suppose we have a partition of E with the properties mentioned in the assertion. Choose ε with $0 < \varepsilon < 1$, $\varepsilon\alpha > 1$. The positive functions r and v on E are defined by

$$r(i) := (\varepsilon\alpha)^k \quad \text{if } i \in E_k \quad \text{and} \quad v(i) := \sup_R \sum_{n=0}^{\infty} \mathbf{E}_{i,R}[r(X_n)].$$

Then for $i \in E_l$

$$v(i) = \sup_R \sum_{k \in Z} (\varepsilon\alpha)^k \sum_{n=0}^{\infty} \mathbf{P}_{i,R}[X_n \in E_k]$$

$$\leq \beta \sum_{k<l} (\varepsilon\alpha)^k + \beta \sum_{k \geq l} (\varepsilon\alpha)^k \alpha^{l-k} = \beta(\varepsilon\alpha)^l\{(\varepsilon\alpha - 1)^{-1} + (1-\varepsilon)^{-1}\}$$

$$= \beta r(i)\{(1-\varepsilon)^{-1} + (\varepsilon\alpha - 1)^{-1}\}$$

$$= (1-\rho)^{-1} r(i) \quad \text{with } \varepsilon < \rho < 1$$

Hence, according to Lemma 4(i), we have $\rho v \geq Pv$ for all $P \in \mathcal{P}$.

Now the "only if"-part will be proved, hence it is assumed that $(E, \mathcal{P})$ is contracting with excessivity factor $\rho < 1$. Without loss of generality, we may assume that the strongly excessive function b is equal to α^l for $i \in E_l$, where $\{E_l \mid l \in \mathbf{Z}\}$ is some partition on E and $1 < \alpha < \rho^{-1}$ (Lemma 2). Note that $\rho b \geq Pb$ for all $P \in \mathcal{P}$ implies for any Markov strategy $R = (P_0, P_1, \ldots)$

$$\rho^n b \geq P_0 \cdots P_{n-1} b, \quad \text{or} \quad (1-\rho)^{-1} b \geq \sum_{n=0}^{\infty} \left(\prod_{m=0}^{n-1} P_m \right) b,$$

(where an empty product is equal to the unit matrix). For $i \in E_l$ this means

$$(1-\rho)^{-1} \alpha^l \geq \sum_{k \in Z} \alpha^k \sum_{n=0}^{\infty} \mathbf{P}_{i,R}[X_n \in E_k],$$

hence

$$(*) \qquad (1-\rho)^{-1} \alpha^l \geq \alpha^k \sum_{n=0}^{\infty} \mathbf{P}_{i,R}[X_n \in E_k].$$

With $\beta := (1-\rho)^{-1}$ this settles the assertion for $k \geq l$. For $k < l$ we apply Lemma 1 with $A = E_k$:

$$\sum_{n=0}^{\infty} \mathbf{P}_{i,R}[X_n \in E_k] \leq \sup_{R_0} \sup_{j \in E_k} \sum_{m=0}^{\infty} \mathbf{P}_{j,R_0}[X_m \in E_k] \leq (1-\rho)^{-1}.$$

where the inequality $(*)$ has been used with $l = k$.

Remarks. (i) In the construction of a strongly excessive function in the first part of the proof, the choice of ε determines the excessivity factor ρ. For $\varepsilon = \alpha^{-1/2}$ the constructed ρ is minimal, viz. $1-(\alpha^{1/2}-1)\beta^{-1}(\alpha^{1/2}+1)^{-1}$.

(ii) If there is a strongly excessive function b such that $b(i) \geq \delta > 0$ for all $i \in E$ then there is a partition $\{E_k, k = 1, 2, 3, \dots\}$ of E, which has the drifting property described in Theorem 1.

4. Strong excessivity and exponentially bounded life times

Since we supposed $\sum_j P(i,j) \leq 1$, there may be positive probabilities for certain i that the process does not exist any more after one step. Hence we can speak of the life time T of the process. We say that $T \geq n$ iff $X_n \in E$, hence for $R = (P_0, P_1, \dots)$

$$\mathbf{P}_{i,R}[X_n \in E] = \mathbf{P}_{i,R}[T \geq n] = (P_0 P_1 \dots P_{n-1} e)(i).$$

In this section we will investigate the relation between strong excessivity and exponential boundedness of the life time distributions. Exponentially bounded lifetimes are well known in statistical sequential analysis (cf. Ferguson [3]).

Definition 2. (1) The life time T of $(E, \mathscr{P})$ is said to be *exponentially bounded* iff there exist a real number γ $(0 < \gamma < 1)$ and a positive function a on E with for all $R \in \mathcal{M}$ and $i \in E$

$$\mathbf{P}_{i,R}[T \geq n] \leq a(i)\gamma^n$$

(2) If the function a in Definition 2.1 does not depend on i, the life time T of $(E, \mathscr{P})$ is said to be *uniformly exponentially bounded.*

Remarks. (1) In the case of discounting we have $P(i,j) := \beta Q(i,j)$ with $0 < \beta < 1$, $\sum_j Q(i,j) \leq 1$. Hence the life time is uniformly exponentially bounded with $a(i) = 1$, $\gamma = 1 - \beta$.

(2) In the case of discounted semi-Markov decision processes we have

$$\beta_Q(i,j) := \int_{0^-}^{\infty} e^{-\alpha t}\, dF_Q(t; i, j), \qquad P(i,j) := \beta_Q(i,j) Q(i,j), \qquad \sum_j Q(i,j) \leq 1.$$

Hence the life time is uniformly exponentially bounded if

$$\beta_Q(i,j) \leq \beta < 1.$$

(3) In the nondiscounted case we have: if there exist a natural number M and a real number $\varepsilon > 0$, such that for all R, i

$$\mathbf{P}_{i,R}[X_M \in E] \leq 1 - \varepsilon,$$

then $(E, \mathscr{P})$ is uniformly exponentially bounded.

Theorem 2. *The life time of $(E, \mathscr{P})$ is exponentially bounded if and only if there is a strongly excessive function b on $(E, \mathscr{P})$ satisfying $0<\delta \leq b(i)$ for all $i \in E$ and certain δ.*

Proof. We first prove the "if"-part assuming $\rho b \geq Pb$ for $P \in \mathscr{P}$ with $0<\rho<1$, $b(i) \geq \delta > 0$. The assumption implies $\rho^n b \geq P_0 \cdots P_{n-1} b \geq \delta P_0 \cdots P_{n-1} e$. Hence for any $i \in E$, $R \in \mathscr{M}$

$$\mathbf{P}_{i,R}[X_n \in E] \leq \rho^n \frac{b(i)}{\delta}.$$

So the choices $\gamma := \rho$ and $a(i) := b(i)\delta^{-1}$ prove the assertion.

We now prove the "only if"-part assuming the exponential boundedness of the life time T. Define

$$b(i) := \sup_R \mathbf{E}_{i,R}\left[\sum_{n=0}^{\infty} \lambda^n e(X_n)\right], \quad i \in E$$

with $1<\lambda<\gamma^{-1}$. It is easy to verify that $\mathbf{E}_{i,R}[e(X_n)] = \mathbf{P}_{i,R}[T \geq n]$. Hence $1 \leq b(i) \leq a(i)(1-\gamma\lambda)^{-1} < \infty$ and therefore by Lemma 4(ii) b has the desired properties.

Corollary. *The life time of $(E, \mathscr{P})$ is uniformly exponentially bounded if and only if $(E, \mathscr{P})$ is contracting with a strongly excessive function b satisfying $0<\delta \leq b(i) \leq \Delta$ for all $i \in E$ and certain δ and Δ.*

Proof. The "if"-part follows from the first part of the proof of Theorem 2:

$$b(i)\delta^{-1} \leq \Delta\delta^{-1}, \quad \text{hence } a(i) := \Delta\delta^{-1} \text{ suffices.}$$

For the "only if"-part we use the construction of the second part in the proof of Theorem 2. Then we obtain b with

$$1 \leq b(i) \leq a(1-\gamma\lambda)^{-1} \quad \text{if } a(i) = a \text{ for all } i \in E.$$

Remark. In the case of a Markov decision process with a uniformly exponentially bounded life time, another strongly excessive function may be constructed in the following way

$$b(i) := \sup_R \sum_{n=0}^{\infty} \mathbf{P}_{i,R}[X_n \in E], \quad i \in E.$$

Namely we have $1 \leq b(i) \leq a(1-\gamma)^{-1}$, $(a \geq 1)$ and

$$b(i) = \sup_R \sum_{n=0}^{\infty} \mathbf{E}_{i,R}[e(X_n)].$$

So by Lemma 4(i) with $0 \leq 1-\rho \leq a^{-1}(1-\gamma)$ we have that b is strongly excessive, with excessivity factor ρ.

5. Strong excessivity and the spectral radius

In this section we will present an analytical characterization of contracting Markov decision processes. If $(E, \mathscr{P})$ is contracting with respect to the bounding function μ, then $\|P\|_\mu \leq \rho < 1$ and consequently $\|P^n\|_\mu^{1/n} \leq \rho$. So we have for a contracting Markov decision process that the spectral radii of all $P \in \mathscr{P}$ are at most ρ:

$$\sup_{P\in\mathscr{P}} \limsup_{n\to\infty} \|P^n\|_\mu^{1/n} \leq \rho < 1.$$

The topic of this section will be the investigation of the reverse proposition.

Definition 4. The *spectral radius of a Markov decision process* $(E, \mathscr{P})$ *with respect to a bounding function* μ is defined as

$$\sup_{P\in\mathscr{P}} \limsup_{n\to\infty} \|P^n\|_\mu^{1/n}.$$

The main result of this section will be: A Markov decision process $(E, \mathscr{P})$ is contracting (with respect to some bounding function) if and only if the spectral radius of $(E, \mathscr{P})$ with respect to some bounding function μ is less than one and

$$\sup_{P\in\mathscr{P}} \|P\|_\mu < \infty.$$

The "only if"-part of this statement has been proved in the introduction of this section. The "if"-part is a trivial consequence of the following theorem.

Theorem 3. *If for a Markov decision process* $(E, \mathscr{P})$ *there is a bounding function* μ *such that*

(i) $\rho^* := \sup_{P\in\mathscr{P}} \limsup_{n\to\infty} \|P^n\|_\mu^{1/n} < 1,$

(ii) $M := \max\{1, \sup_{P\in\mathscr{P}} \|P\|_\mu\} < \infty,$

then there is for each $\varepsilon > 0$ *a bounding function* $\tilde{\mu}$ *such that*

(i) $(\rho^* + \varepsilon)\tilde{\mu} \geq P\tilde{\mu}$ *for all* $P \in \mathscr{P}$.

(ii) $\mu \leq \tilde{\mu} \leq L\mu$ *for some constant* L.

We shall postpone the proof of this theorem to the end of this section and we consider some useful lemmas first.

In Lemmas 5 and 6 we transform the assumptions (i) and (ii) of Theorem 3 into equivalent conditions, which allow us to consider only the bounding function $\mu = e$.

Lemma 5. *If assumption* (i) *of Theorem* 3 *holds, then there is for each* $1 < \lambda < \rho^{*-1}$ *a number* ρ, $0 < \rho < 1$ *and for each* $P \in \mathscr{P}$ *there is a number* $b_P > 0$ *such that*

$$\lambda^n P^n \mu \leq b_P \rho^n \mu \quad \text{for all } P \in \mathscr{P} \text{ and } n = 0, 1, 2, \ldots.$$

Proof. Choose $\lambda > 1$ with $\lambda\rho^* < 1$. Then

$$\sup_{P\in\mathcal{P}} \limsup_{n\to\infty} \|\lambda^n P^n\|_\mu^{1/n} = \lambda\rho^*$$

Choose ρ with $\lambda\rho^* < \rho < 1$. Then there is a n_P for any P with

$$\|\lambda^n P^n\|_\mu \leq \rho^n \quad \text{for } n \geq n_P.$$

Hence b_P may be chosen such that

$$\|\lambda^n P^n\|_\mu \leq b_P\rho^n \quad \text{for } n = 0, 1, 2, \ldots.$$

Lemma 6. *Let the assumptions* (i) *and* (ii) *of Theorem* 3 *hold. Define* $\mathcal{P}^*$ *by* $\mathcal{P}^* := \{P^* \mid P^*(i,j) := M^{-1}\mu^{-1}(i)P(i,j)\mu(j),\ P\in\mathcal{P}\}$ *and* λ^* *by* $\lambda^* := \lambda M$. *Then*

(i) $\sum_j P^*(i,j) \leq 1$.

(ii) $\lambda^{*n} P^{*n} e \leq b_P\rho^n e$.

(iii) $P^*\iota \leq \lambda^{*-1}\nu$ *for some bounding function* ν *implies* $P(\mu\otimes\nu) \leq \lambda^{-1}\mu\otimes\nu$, *where* $(\mu\otimes\nu)(i) := \mu(i)\nu(i)$ *(and reversely)*.

Proof. (i) is trivial, (ii) is a consequence of Lemma 5 and the property

$$(P^{*n})(i,j) = M^{-n}\mu^{-1}(i)(P^n)(i,j)\mu(j)$$

and (iii) is proved by inspection.

As a consequence of these two lemmas it is sufficient for the proof of Theorem 3 to show that: if for some Markov decision process $(E, \mathcal{P})$

(*) $Pe \leq e$,

(**) $\lambda^n P^n e \leq b_P\rho^n e$ for $n = 0, 1, 2, \ldots$ and all $P\in\mathcal{P}$, where $0 < \rho < 1 < \lambda$,

then there is some bounding function ν such that

$$\lambda^{-1}\nu \geq P\nu \quad \text{for all } P\in\mathcal{P}.$$

In Lemma 7 we construct such a bounding function ν under an additional assumption. The technique we use is familiar. We show that Howard's policy iteration method converges for this Markov decision process $(E, \mathcal{P})$ with unit rewards and discount factor λ (cf. Ross [10]). However, we need the extra assumption that the value function of this Markov decision process is finite. Afterwards, we show that this assumption is implied by (*) and (**).

This appears to be the most tedious step in the proof of Theorem 3. From now on till the end of this section we assume that (*) and (**) hold. Hence $\lambda > 1$ fixed.

Lemma 7. *Define* v_P *by* $v_P := \sum_{n=0}^{\infty} \lambda^n P^n e$ *and* v *by* $v := \sup_{P\in\mathcal{P}} v_P$. *If* $v < \infty$, *then* *there is a bounding function ν such that*

(i) $\lambda P\nu \leqslant \nu$ *for all* $P \in \mathscr{P}$,

(ii) $e \leqslant \nu \leqslant v$.

Proof. We start by formulating the policy iteration method for $(E, \mathscr{P})$ with reward one in each state: Choose $P_0 \in \mathscr{P}$ and define P_{n+1} recursively for $n = 0, 1, 2, \ldots$ in the following way. First of all fix a function $\varepsilon_n : E \to [0, 1)$ such that

$$\varepsilon_n(i) = 0 \quad \text{if } v_{P_n}(i) = \sup_P \{1 + \lambda(Pv_{P_n})(i)\},$$

or

$$0 < \varepsilon_n(i) \leqslant \sup_P \{1 + (\lambda P v_{P_n})(i)\} - v_{P_n}(i) \quad \text{otherwise.}$$

Then choose P_{n+1} such that

$$1 + \lambda(P_{n+1}v_{P_n})(i) \geqslant \sup_P \{1 + (\lambda P v_{P_n})(i)\} - \varepsilon_n(i).$$

The sequence $\{P_n\}$ satisfies $v_{P_n} \leqslant e + \lambda P_{n+1} v_{P_n}$. Iterating this equation yields

$$v_{P_n} \leqslant \sum_{k=0}^{N-1} \lambda^k P_{n+1}^k e + \lambda^N P_{n+1}^N v_{P_n} \quad \text{and hence} \quad v_{P_n} \leqslant v_{P_{n+1}}.$$

Namely,

$$\lambda^N P_{n+1}^N v_{P_n} \leqslant \lambda^N P_{n+1}^N b_{P_n}(1-\rho)^{-1} e \leqslant b_{P_{n+1}} \rho^N b_{P_n}(1-\rho)^{-1} e,$$

which tends to zero for $N \to \infty$. Since $v_{P_n} \leqslant v$ and $v_{P_n} \leqslant v_{P_{n-1}}$ we obtain: $\nu := \lim_{n\to\infty} v_{P_n} < v$. For the proof of $\nu \geqslant \lambda P \nu$ for all $P \in \mathscr{P}$, note that

$$v_{P_{n+1}} = e + \lambda P_{n+1} v_{P_{n+1}} \geqslant e + \lambda P_{n+1} v_{P_n} \geqslant e + \lambda P v_{P_n} - \varepsilon_n \quad \text{for all } P \in \mathscr{P}.$$

This implies $v_{P_{n+1}} \geqslant \lambda P v_{P_n}$, which gives $\nu \geqslant \lambda P \nu$.

However, in order to prove $v < \infty$ we need four more lemmas. By P_A we will denote the sub-Markov matrix of the Markov chain with matrix P restricted to A, i.e.

$$P_A(i,j) := P(i,j) \quad \text{if } i, j \in A,$$
$$:= 0 \qquad \text{otherwise.}$$

Lemma 8. *Suppose for some* $P \in \mathscr{P}$, $K \in \mathbf{R}$, $i \in E$

$$\sum_{n=0}^{\infty} \lambda^n (P^n e)(i) \geqslant K.$$

Then there is for each $\varepsilon > 0$ *a finite subset* A *of* E *with*

$$\sum_{n=0}^{\infty} \lambda^n (P_A^n e)(i) \geqslant K - \varepsilon.$$

Proof (see Ornstein [9] for a similar construction). Choose N with

$$\sum_{n=0}^{N} \lambda^n (P^n e)(i) \geq v_P(i) - \tfrac{1}{2}\varepsilon.$$

It is easy to verify that for each n there exists a finite set $A_n \subset E$ with

$$(P^n_{A_n} e)(i) \geq (P^n e)(i) - \frac{\varepsilon}{2N} \lambda^{-n}.$$

Hence $A := \bigcup_{n=0}^{N} A_n$ has the required properties.

The foregoing lemma says that in some sense restriction to finite Markov chains is allowed for fixed P. The next lemma shows that the expected number of visits to state i (discounted with factor λ) is bounded as a function of P.

Lemma 9. *For all $P \in \mathcal{P}$ we have*

$$P_i(\lambda) := \sum_{n=0}^{\infty} \lambda^n P^n(i, i) \leq (1-\rho)^{-1} \quad \text{for } i \in E.$$

Proof. Assume there is a number $\varepsilon > 0$ such that for some n

$$\lambda^n P^n(i, i) \geq \rho^n + \varepsilon.$$

Then, by a standard argument we have

$$\lambda^{nk} P^{nk}(i, i) \geq (\rho^n + \varepsilon)^k \quad \text{for } k = 1, 2, 3, \ldots.$$

However, for k sufficiently large: $(\rho^n + \varepsilon)^k > b_P \rho^{nk}$. This is contradictory to assumption (**). Hence

$$\lambda^n P^n(i, i) \leq \rho^n \quad \text{for } n = 0, 1, 2, \ldots.$$

Lemma 10. *Let $B \subset E$, such that for some $P \in \mathcal{P}$ and a real number K:*

$$\sum_{n=0}^{\infty} \lambda^n P^n_B e \leq K e.$$

Then we have for the set $C := B \cup \{j\}$, $j \in E \backslash B$:

$$\sum_{n=0}^{\infty} \lambda^n P^n_C e \leq K' e$$

where the constant K' is determined by λ, ρ and K. (Note that $K \geq 1$ and K' independent of P.)

Proof. Define for $i, k \in E$ and $A \subset E$:

$$_A P^n(i, k) := \sum_{l_1, \ldots, l_{n-1} \in A} P(i, l_1) \cdot P(l_1, l_2) \cdots P(l_{n-1}, k).$$

Note that $_A P^n(i, k) = P^n_A(i, k)$ if $i, k \in A$.

Consider

$$\sum_{k\in C} {}_CP^n(j,k)=\sum_{m=1}^{n} {}_CP^{n-m}(j,j)\sum_{k\in B} {}_BP^m(j,k)+{}_CP^n(j,j) \quad \text{for } n\geq 0.$$

Hence

$$(*) \qquad \sum_{n=0}^{\infty}\lambda^n\sum_{k\in C} {}_CP^n(j,k)=\sum_{n=0}^{\infty}\lambda^n {}_CP^n(j,j)\Big\{1+\sum_{n=1}^{\infty}\lambda^n\sum_{k\in B} {}_BP^n(j,k)\Big\}$$

$$\leq\frac{1}{1-\rho}\Big\{1+\lambda\sum_{i\in B}P(j,i)\sum_{n=0}^{\infty}\lambda^n\sum_{k\in B} {}_BP^n(i,k)\Big\}$$

$$\leq\frac{1}{1-\rho}\Big\{1+\lambda\sum_{i\in B}P(j,i)K\Big\}\leq\frac{1+\lambda K}{1-\rho},$$

where the first inequality is justified by Lemma 9, and the second one by the assumption of the lemma.

Further consider for $i\in B$:

$$\sum_{k\in C} {}_CP^n(i,k)=\sum_{k\in B} {}_BP^n(i,k)+\sum_{m=1}^{n} {}_BP^m(i,j)\sum_{k\in C} {}_CP^{n-m}(j,k).$$

Hence

$$(**) \qquad \sum_{n=0}^{\infty}\lambda^n\sum_{k\in C} {}_CP^n(i,k)=\sum_{n=0}^{\infty}\lambda^n\sum_{k\in B} {}_BP^n(i,k)$$

$$+\sum_{m=1}^{\infty}\lambda^m {}_BP^m(i,j)\sum_{n=0}^{\infty}\lambda^n\sum_{k\in C} {}_CP^n(j,k)\leq K+\lambda K\frac{1+\lambda K}{1-\rho},$$

where the last inequality follows from (*), the assumption of the lemma and from the inequality

$$\sum_{m=1}^{\infty}\lambda^m {}_BP^m(i,j)\leq\lambda\sum_{m=1}^{\infty}\lambda^{m-1}\sum_{k\in B} {}_BP^{m-1}(i,k)\leq\lambda K.$$

Let $K':=K+\lambda K(1+\lambda K)/(1-\rho)$. Since $\lambda K>1$, the assertion is now a consequence of (*) and (**).

Finally we prove in Lemma 11, using Lemmas 9 and 10, that the function v defined in Lemma 7 is finite. We even prove that this function is bounded. To prove this, we assume the contrary, i.e. $\sup_{i\in E} v(i)=\infty$.

Then we construct disjoint subsets of $E: A_1, A_2, A_3, \ldots$ and sub-Markov matrices $P_1, P_2, P_3, \ldots$ such that

$$\sum_{n=0}^{\infty}\lambda^n P_{A_k}^n e(i)\geq a_k \quad \text{for } i\in A_k$$

where $a_1, a_2, a_3, \ldots$ is a nondecreasing sequence tending to infinity. With these $P_1, P_2, P_3, \ldots$ we construct a $P\in\mathcal{P}$ such that the function v_P (see Lemma 7) is unbounded, which contradicts assumption (**).

Lemma 11. $\sup_{i\in E} v(i)<\infty$.

Proof. Let A be a subset of E. We define on A:

$$v_A := \sup_{P\in\mathcal{P}} \sum_{n=0}^{\infty} \lambda^n P_A^n e.$$

Note that $v(i)=v_E(i)$, $i\in E$. First note that if A is finite and $\sup v_E(i)=\infty$, then $\sup_{i\in A^c} v_{A^c}(i)=\infty$, since assume to the contrary $\sup_{i\in A^c} v_{A^c}(i)<\infty$. Then using Lemma 10 we have for $j\in A$: $\sup_{i\in A^c\cup\{j\}} v_{A^c\cup\{j\}}(i)<\infty$, and so by induction $\sup_{i\in A^c\cup A} v_{A^c\cup A}(i)<\infty$ which produces a contradiction.

Suppose $\sup_{i\in E} v(i)=\infty$. Fix a nondecreasing sequence $a_1, a_2, \ldots$ tending to infinity. Fix $\varepsilon>0$. There must be an $i_1\in E$ and a $P_1\in\mathcal{P}$ such that $v_{P_1}(i_1)\geq a_1+\varepsilon$. Hence by Lemma 8 there is a finite subset A_1 such that $i_1\in A_1$ and for $R_1=(P_1, P_1, \ldots)$

$$\sum_{n=0}^{\infty} \lambda^n \mathbf{P}_{i_1,R_1}[X_0\in A_1, \ldots, X_n\in A_1]\geq a_1.$$

Consider the process restricted to A_1^c. We have already seen that $\sup_{i\in A_1^c} v_{A_1^c}(i)=\infty$. Hence there is an $i_2\in A_1^c$ and a $P_2\in\mathcal{P}$ such that for $R_2=(P_2, P_2, \ldots)$

$$\sum_{n=0}^{\infty} \lambda^n \mathbf{P}_{i_2,R_2}[X_0\in A_1^c, \ldots, X_n\in A_1^c]\geq a_2+\varepsilon$$

and again by Lemma 8 there is a finite subset $A_2\subset A_1^c$ such that $i_2\in A_2$ and:

$$\sum_{n=0}^{\infty} \lambda^n \mathbf{P}_{i_2,R_2}[X_0\in A_2, \ldots, X_n\in A_2]\geq a_2.$$

We may apply the same argument to $(A_1\cup A_2)^c$. So we find finite sets $A_1, A_2, \ldots$ with state $i_1\in A_1$, $i_2\in A_2, \ldots$ and $P_1, P_2, \ldots\in\mathcal{P}$ such that for $R_k := (P_k, P_k, \ldots)$

$$\sum_{n=0}^{\infty} \lambda^n \mathbf{P}_{i_k,R_k}[X_0\in A_k, \ldots, X_n\in A_k]\geq a_k.$$

Consider a new element $P\in\mathcal{P}$ defined by: $P(i,j) := P_k(i,j)$ if $i\in A_k$. It is easy to verify that for all $k=1, 2, 3, \ldots$

$$v_P(i_k)\geq \sum_{n=0}^{\infty} \lambda^n \mathbf{P}_{i_k,R_k}[X_0\in A_k, \ldots, X_n\in A_k].$$

Hence

$$\sup_k v_P(i_k)=\infty.$$

On the other hand we have for all $P\in\mathcal{P}$: $v_P(i)\leq b_P/(1-\rho)$ for all $i\in E$. Therefore $\sup_i v(i)<\infty$.

The proof of Theorem 3 is a direct consequence of the foregoing lemmas.

Proof of Theorem 3. Note that we transformed the model as in Lemma 6. So we have to prove the existence of a bounding function ν assuming * and **. Fix $\varepsilon > 0$ and define $\lambda := (\rho^* + \varepsilon)^{-1}$.

From Lemmas 11 and 8 we have the existence of a bounding function ν such that

$$\lambda P\nu \leqslant \nu \quad \text{for } P \in \mathcal{P}$$

and $e \leqslant \nu \leqslant Le$, for some constant L. Hence, by Lemma 6, we have for the untransformed model

$$P(\mu \otimes \nu) \leqslant \lambda^{-1}(\mu \otimes \nu) \quad \text{for all } P \in \mathcal{P}.$$

Define $\tilde{\mu} := \mu \otimes \nu$. Then we have $P\tilde{\mu} \leqslant (\rho^* + \varepsilon)\tilde{\mu}$ and $\mu \leqslant \tilde{\mu} \leqslant \mu L$.

6. Some consequences and remarks

(1) In the proofs of Theorems 1 and 2 the assumption on $\mathcal{P}$ has only been used for the proof of the sufficiency of both conditions for strong excessivity, not for the necessity.

(2) In our definition strongly excessive functions are positive. Hinderer [5] allows the value 0 for $b(i)$. However, the strong excessivity (even if ρ is not less than 1, as in Hinderer's case) requires for the system to remain in the set of states with b-value 0 as soon as this set is entered. Hence, without restricting generality, one may assume that the state space is left when such a state is entered.

(3) Combination of Theorems 1 and 2 gives the following necessary and sufficient condition for exponential boundedness of the life time of $(E, \mathcal{P})$: there exist a partitioning $\{E_k \mid k = 0, 1, \ldots\}$ of E and numbers $\alpha > 1, \beta \geqslant 1$, such that for all $R \in \mathcal{M}$

$$\sum_{n=0}^{\infty} \mathbf{P}_{i,R}[X_n \in E_k] \leqslant \beta \min\{1, \alpha^{l-k}\} \quad \text{if } i \in E_l.$$

(4) Suppose there are a positive integer N, real numbers M and β, $0 < \rho < 1$ and a bounding function μ such that for all $P \in \mathcal{P}$:

(i) $P\mu \leqslant M\mu$ and (ii) $P^N\mu \leqslant \rho\mu$.

We say there is *N-stage contraction* in this situation. It is an immediate consequence of Theorem 3 that for each $\varepsilon > 0$ there is a bounding function $\tilde{\mu}$ such that

(iii) $P\tilde{\mu} \leqslant (\rho^{1/N} + \varepsilon)\mu$ and (iv) $\mu \leqslant \tilde{\mu} \leqslant L\mu$, for some L.

However, it can be proved more directly that N-stage contraction implies (iii) and (iv). Namely, it is easy to verify for $\lambda := (\rho^{1/N}+\varepsilon)^{-1}$ that $1<\lambda<\rho^{-N}$ and

$$v_P := \sum_{n=0}^{\infty} \lambda^n P^n \mu = \sum_{k=0}^{N-1} \lambda^k P^k \sum_{n=0}^{\infty} \lambda^{nN} P^{nN} \mu \leq \left\{ \frac{(\lambda M)^N - 1}{\lambda M - 1} \cdot \frac{1}{1-\rho\lambda^N} \right\} \mu$$

for all $P \in \mathcal{P}$

Further we can construct completely analogously to the proof of Lemma 7 a bounding function ν such that: $\lambda P \nu \leq \nu$ for all $P \in \mathcal{P}$ and

$$\mu \leq \nu \leq L\mu \quad \text{where } L := \frac{(\lambda M)^N - 1}{\lambda M - 1} \frac{1}{1-\rho\lambda^N}.$$

(5) In a decision process one considers contraction properties of a whole set $\mathcal{P}$ of operators simultaneously. If one only considers one (not necessarily linear) operator T in a complete metric space X, then N-stage contraction of T implies one-stage contraction of T in X with respect to some other distance. This has been shown by Walter [13] without using the equivalent of our condition (i).

(6) In fact we proved in Lemma 11 that

$$(*) \qquad \sup_{P \in \mathcal{P}} \limsup_{n\to\infty} \|P^n\|_e^{1/n} < 1$$

implies

$$(**) \qquad \sup_{i \in E} \sup_{P \in \mathcal{P}} \sum_{n=0}^{\infty} P^n e(i) < \infty.$$

However (**) implies, for all $0<\rho<1$, the existence of a positive integer N such that

$$(***) \qquad \sup_{P \in \mathcal{P}} \|P^N\|_e < \rho.$$

To verify this, assume to the contrary: there is a $0<\rho<1$ such that for all $N = 0, 1, 2, \ldots$ there is a $P \in \mathcal{P}$ and an $i \in E$ such that

$$P^N e(i) > \rho.$$

Hence, since $P^n e(i)$ is nonincreasing in n we have

$$\sum_{n=0}^{\infty} P^n e(i) \geq \sum_{n=0}^{N} P^n e(i) \geq (N+1)\rho.$$

And therefore

$$\sup_{i \in E} \sup_{P \in \mathcal{P}} \sum_{n=0}^{\infty} P^n e(i) = \infty$$

which contradicts (**). So we actually proved here that the assumptions (i) and (ii) of Theorem 3 imply N-stage contraction with respect to the same bounding function and 1-stage contraction with respect to another (but in some sense equivalent) bounding function.

(7) In Veinott [12] a *similarity* transformation for decision processes was introduced for transient models with a finite state space. This transformation for the transition probabilities has the form

$$P^*(i,j) := \frac{P(i,j)\mu(j)}{\mu(i)} P \in \mathcal{P}, \quad \mu \text{ a bounding function.}$$

A lot of properties of the decision process are invariant under this transformation. Lemma 3 in Veinott's paper (due to Hoffman) is exactly the same as the statement of Theorem 3 for a finite state space and a finite action space. Note however that in the finite case it is obvious that the function v defined in Lemma 9 is bounded. Furthermore the finiteness of E implies that $P^N\mu \leqslant \rho_0\mu$ for all P and some N, $\rho_0 < 1$ if the spectral radius of $(E, \mathcal{P})$ is less than one. This can easily be used to show that $(E, \mathcal{P})$ is contracting (compare Remark 4).

(8) At first glance one might expect that

$$r := \sup_{P\in\mathcal{P}} \sup_{i,j\in E} \limsup_{n\to\infty} \{P^{(n)}(i,j)\}^{1/n} < 1$$

is a sufficient condition for the decision process to be contracting. The quantity r may be regarded as a generalization of the concept *convergence norm* to decision processes (see Seneta [11, p. 162]). However, we produce a counterexample for this statement.

Counterexample. $E := \{-1, 0, 1, 2, 3, \cdots\}$, $\mathcal{P} = \{P_n \mid n = 1, 2, 3, \cdots\}$ where for all $n = 1, 2, 3, \cdots$

(i) $P_n(K, K-1) = 1 \quad K \geqslant 1$

(ii) $P_n(0, i) = 0 \quad i \neq 0, \quad P_n(0, 0) = \rho < 1$ and $P_n(-1, n) = 1, \quad n = 1, 2, 3, \cdots$

It is easy to verify that if $j \neq 0$: $P_n^{(N)}(i,j) = 0$ for N sufficiently large and

$$\limsup_{N\to\infty} \{P_n^{(N)}(i,0)\}^{1/N} = \rho \quad \text{for all } i \in E \text{ and } n = 1, 2, 3, \cdots.$$

Hence $r = \rho$. However if μ is strongly excessive with excessivity factor $0 < \rho^* < 1$, then

(i) $\rho^*\mu(0) \geqslant \rho\mu(0)$ hence $\rho^* \geqslant \rho$

(ii) for $K = 1, 2, 3, \ldots$ $\rho^{*K}\mu(K) \geqslant \mu(0)$ hence $\mu(K) \geq \mu(0)(\rho^*)^{-K}$ and therefore since $\mu(-1) \geqslant \mu(K)$ $K = 0, 1, 2, \ldots$ we have $\mu(-1) \geqslant \sup_K (\rho^*)^{-K}\mu(0) = \infty$. So there does not exist a strongly excessive function here.

(9) The following 6 assertions for $(E, \mathcal{P})$ are equivalent:

(A) $(E, \mathcal{P})$ is contracting with a strongly excessive function b satisfying $0 < \delta \leq b(i) \leqslant \Delta$ for all $i \in E$ and certain δ and Δ.

(B) $(E, \mathcal{P})$ is strongly excessive with a strongly excessive function which is finitely valued.

(C) For certain N and $\varepsilon > 0$

$$\mathbf{P}_{i,R}[X_N \in E] \leq 1 - \varepsilon \quad \text{for all } i \in E, R \in \mathcal{M}.$$

(D) There exist a number $\varepsilon > 0$ and a finite partition $\{E_k \mid k = 1, \ldots, N\}$ of E, such that

$$\mathbf{P}_{i,R}\left[X_1 \in \bigcup_{k=l}^{N} E_k\right] \leq 1 - \varepsilon \quad \text{for all } i \in E_l, l = 1, \ldots, N, R \in \mathcal{M}.$$

(E) The life time of $(E, \mathcal{P})$ is uniformly exponentially bounded.

(F) There exist a number ρ $(0 < \rho < 1)$, a function b on E with $0 < \delta < b(i) \in \Delta$ for certain δ, Δ, and a natural number N, such that for all $P_1, \ldots, P_N \in \mathcal{P}$ $P_1, \ldots, P_N b \leq \rho b$.

Acknowledgment

The authors are grateful to A. Hordijk, J.Th. Runnenburg and J. van der Wal for several valuable suggestions.

References

[1] D. Blackwell, Discounted dynamic programming. Ann. Math. Statist. 36 (1965) 226–235.

[2] E.V. Denardo, Contraction mappings in the theory underlying dynamic programming, SIAM-Review 9 (1967) 165–177.

[3] T.S. Ferguson, Mathematical Statistics (Academic Press, New York and London, 1967).

[4] K.M. van Hee, Markov strategies in dynamic programming, Mathematics of Operations Research, 3 (1) (1978).

[5] K. Hinderer, Bounds for stationary finite-stage dynamic programs with unbounded reward functions. Report of the "Institut für Mathematisch Stochastik der Universität Hamburg (June 1975).

[6] A. Hordijk, Dynamic programming and Markov potential theory (Amsterdam, 1974).

[7] S.A. Lippman, Semi-Markov decision processes with unbounded rewards, Management Science 19 (1973) 717–731.

[8] S.A. Lippman, On dynamic programming with unbounded rewards, Management Science 21 (1975) 1225–1233.

[9] D. Ornstein, On the existence of stationary optimal strategies, Proc. Amer. Math. Soc. 20 (1969) 563–569.

[10] S. Ross, Applied probability models with optimization applications (Holden Day, 1970).

[11] E. Seneta, Nonnegative matrices, George Allen and Unwin Ltd. (1973).

[12] A.F. Veinott, Discrete dynamic programming with sensitive discount optimality criteria, Ann. Math. Statist. 40 (1969) 1635–1660.

[13] W. Walter, A note on contraction, SIAM Review 18 (1976) 107–111.

[14] J. Wessels, Markov programming by successive approximations with respect to weighted supremum norms, J. Math. Anal. Appl. 58 (1977).

[15] J. Wijngaard, Stationary Markovian decision problems. Technological University Eindhoven (1975).