Recurrence conditions and the existence of average optimal strategies for inventory problems on a countable state space

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics

PROBABILITY THEORY, STATISTICS AND OPERATIONS RESEARCH GROUP

Recurrence conditions and the existence of average optimal strategies for inventory problems on a countable state space.

by

J. Wijngaard

Memorandum COSOR 77-03

Eindhoven, February 1977

The Netherlands

Introduction

The existence of average optimal strategies in Markovian decision processes has been investigated frequently. See, for instance, Blackwell [2] (finite state space), Ross [5] , Hordijk [3] (countable state space), Tijms [7], Wijngaard [8] (arbitrary state space). Sufficient conditions for the existence of an avery optimal strategy consist in general of some recurrence conditions and some continuity- and compactness conditions. The conditions derived in [8] include for instance the existence of the expected time and costs until the first visit to some subset A of the state space, the continuity of this recurrencetime and -costs on the space of strategies and the compactness of this space.
If the recurrence conditions are weak or strong depends on the structure of the problem. In inventory problems for instance the one-period costs are high if the inventorylevel is far from zero. Therefore the good strategies have to bring back the inventorylevel near to zero. That means that in this sort of problems one may require rather strong recurrence conditions, without loss of generality.
The main point of this paper is the investigation of this last statement. In section 3 two sets of conditions are given sufficient for the existence of an average optimal strategy for Markovian decision problems with a counteble state space. In section 4 it is shown that these conditions are satisfied for inventory processes where one orders at least a certain quantity R if the inventorylevel is below a certain level m.
In section 5 the problem is considered if the recurrence conditions stated in section 3 are satisfied for alle "good" strategies in inventory problems. The one-period costs are assumed to be unbounded on all infinite intervals and the existence of at least one strategy α_0 is required such that the avery costs g_{α_0} exist. A good strategy can be defined then as a strategy for which the average costs are smaller than g_{α_0}.

2. Preliminaries

Let V be a countable set. A stationary Markovian decision problem (SMD) is defined as a set of pairs $\{(P_\alpha, c_\alpha)\}$, $\alpha \in A$, where P_α for $\alpha \in A$ is a Markov process on V and c_α a nonnegative function on V (the costfunction). An SMD can be interpreted as a Markovian decision process where only stationary strategies are allowed, but the product property is not necesserily satisfied in an SMD.

The sum $\sum_{v\in B} P_\alpha(u,v)f(v)$, for f some function on V and B a subset of V, is denoted by $(P_{\alpha B}f)(u)$. If B = V we will write $(P_\alpha f)(u)$.

The average costs of α, starting in u, are equal to

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\ell=0}^{n-1} (P_\alpha^\ell c_\alpha)(u)$$

if this limit exists, and are denoted by $g_\alpha(u)$. If P_α has only one ergodic set the function g_α is constant on V.

A strategy $\alpha_0 \in A$ is called averege optimal if $g_{\alpha_0}(u) \le g_\alpha(u)$ for $u \in V$, $\alpha \in A$

The concept embedded Markov process is frequently used.

Let $A \subset V$ and $A' := V \setminus A$. If $\lim_{n\to\infty} (P^n_{\alpha A'}1)(u) = 0$ for all $u \in V$, the embedded Markov process of P on A exists and the transition probability, $Q_{\alpha A}$, is given by $Q_{\alpha A}(u,v) = \sum_{n=0}^{\infty} (P^n_{\alpha A'} P_{\alpha A} 1_v)(u)$, where 1_v is the characteristic function of $\{v\}$. The total expected costs and time until the first visit to A, starting in u, are equal to $\sum_{n=0}^{\infty} (P^n_{\alpha A'} c_\alpha)(u)$ and $\sum_{n=0}^{\infty} (P^n_{\alpha A'} 1)(u)$. These sums are sometimes denoted by $T_{\alpha A} c_\alpha$ and $T_{\alpha A} 1$.

For ω a positive function on V the function space B_ω is defined as the space of all complex valued functions f such that $\frac{|f(u)|}{\omega(u)}$ is bounded in u. With the norm $\|f\|_\omega := \sup_u \frac{|f(u)|}{\omega(u)}$ this space is a Banach space.

For $\omega(u)=1$, $u \in V$ this space is the space of bounded functions with the sup-norm. See [4], for the use of this sort of function spaces in dynamic **programming**.

3. Sufficient conditions for the existence of average optimal strategies

Let $\{(P_\alpha, c_\alpha)\}$, $\alpha \in A$ be an SMD on a countable state space V. A set of rather weak conditions, sufficient for the existence an average optimal strategy is the following.

<u>Ia</u> There is a finite subset A of V such that the expected time and costs until the first visit to A, $\sum_{n=0}^{\infty} (P^n_{\alpha A'} 1)(u)$ and $\sum_{n=0}^{\infty} (P^n_{\alpha A'} c_\alpha)(u)$, exist for all starting states $u \in V$ and are bounded in α for each $u \in A$.

<u>Ib</u> There is a topology on $\mathcal{A}$ such that the transition probability $Q_{\alpha A}(u,v)$, the recurrence time $(T_{\alpha A} 1)(u)$ and the recurrence costs $(T_{\alpha A} c_\alpha)(u)$ are continuous in α for all $u,v \in A$.

<u>Ic</u> $\mathcal{A}$ is compact.

<u>Id</u> $Q_{\alpha A}$ has only one ergodic set for all $\alpha \in \mathcal{A}$.

The proof of the sufficiency of these conditions will not be given here (see [8]). It is rather straightforward and based on the fact that one can write the average costs as the quotient of the average recurrence costs and the average recurrence time,

$$g_\alpha = \frac{\sum_{u \in A} \pi_\alpha(u)(T_{\alpha A} c_\alpha)(u)}{\sum_{u \in A} \pi_\alpha(u)(T_{\alpha A} 1)(u)}$$

where π_α is the unique invariant probability of Q_α.
The condition Id may be replaced by a communicatingness condition, see [1], [3], [8]. The communicatingness of a Markovian decision problem implies that the set $\mathcal{A}$ is dominated by the subset $\mathcal{A}_1$ of $\mathcal{A}$ with all α such that $Q_{\alpha A}$ has only one ergodic set.
The conditions Ib, c are always satisfied if the number of possible actions in each state is finite.

The difficulty with the conditions I is that they are not easy to check. Especially the continuity conditions Ib are hard to verify, since they are expressed in infinite sums of $P^n_{\alpha A'} f$. We prefer continuity conditions directly on P_α, c_α. In the following set of conditions this is realised.

<u>IIa</u> There is a finite set $A \subset V$ and a positive function ω on V such that $1, c_\alpha \in \mathcal{B}_\omega$, $\|c_\alpha\|_\omega$ is bounded on $\mathcal{A}$, $P_{\alpha A'}$ is a bounded linear operator in $\mathcal{B}_\omega$, $\|P_{\alpha A'}\|_\omega$ is bounded on $\mathcal{A}$ and $\|P^n_{\alpha A'}\|_\omega \le \rho < 1$ for some integer n and $\rho < 1$, uniform in α.

IIa' For each $u \in V$, $\varepsilon > 0$, infinite set $E \subset V$, there is a finite set $E_{u\varepsilon} \subset V$ such that $|(P_{\alpha E}\,\omega)(u) - (P_{\alpha E_{u\varepsilon}}\,\omega)(u)| < \varepsilon$ for all $\alpha \in A$

IIb $P_\alpha(u,v)$, $c_\alpha(u)$ are continuous in α for all $u,v \in V$

IIc See Ic

IId See Id

In the next lemma it is shown that the conditions II are stronger than the conditions I.

Lemma 1 The conditions IIa,a',b imply the conditions Ia,b.

Proof The condition Ia follows directly from the condition IIa. From the last part of IIa it follows also that, to prove Ib, it is sufficient to show the continuity of $(P^n_{\alpha A'}\omega)(u)$ for all $u \in A$ and for all n.

Using IIb it is possible to prove for each $\varepsilon > 0$, $n=0,1,2,\ldots$, the existence of finite intervals $B_1, B_2, \ldots, B_n$ such that

$$|(P^n_{\alpha A'}\omega)(u) - (P_{\alpha B_1} P_{\alpha B_2} \ldots P_{\alpha B_n}\,\omega)(u)| < \varepsilon \text{ for all } \alpha \in A,\ u \in A$$

The rest of the proof is straightforward. □

The recurrence conditions IIa look rather strong, but if one considers problems with somewhat more structure it turns out that they are not too bad. In the next section it is shown that they are satisfied for a rather large class of inventory problems.

4. Inventory problems

The inventory problems considered in this section are one-point inventory problems with leadtime 0 and with backlogging. The state of the system can be represented by the inventory-position. For convenience we assume the existence of an upperbound M on the inventory. The state space V is therefore the set of all integers on $(-\infty, M]$.
An action is a quantity to order and a stationary strategy is a nonnegative function $\alpha(.)$ on V where $\alpha(u)$ gives the quantity to order in state u.
The boundedness of the inventorylevel from above, by M, implies $u + \alpha(u) \le M$.

Let $\varphi(.)$ be the probabilitydensity fuction of the demand per period, then $P_\alpha(u,v) = \varphi(u + \alpha(u) - v)$.
Let $r_1(x)$ be the costs of ordering a quantity x and $r_2(y)$ the costs of having an inventorylevel y (inventory- and stockoutcosts), then $c_\alpha(u) = r_1(\alpha(u)) + r_2(u + \alpha(u))$.

Now it will be shown that the conditions IIa are satisfied for these inventory problems if the quantity to order is "large enough" for small u.

Theorem 2 Let $r_1, r_2 \in \mathcal{B}_\omega$ with $\omega(u) := e^{|u|}$, $u \in (-\infty, M]$.
If there exist integers $m < 0$, $R > 0$ such that $\sum_{x=0}^{\infty} e^x \varphi(x) < e^R$
and if for all $\alpha \in A$, $\alpha(u) \geq R$ for $u \leq m$, then the inventory problem satisfies the condition IIa with $A := [m+1, M]$ and $\omega(u) = e^{|u|}$.

Proof Since $e^{|u|} \geq 1$ the function 1 is an element of $\mathcal{B}_\omega$.
From $r_1(\alpha(u)) \leq \|r_1\|_\omega e^{\alpha(u)} \leq \|r_1\|_\omega e^{M-u}$
and $r_2(u+\alpha(u)) \leq \|r_2\|_\omega e^{|u+\alpha(u)|} \leq \|r_2\|_\omega e^M . e^{|u|}$
it follows that $c_\alpha \in \mathcal{B}_\omega$ and $\|c_\alpha\|_\omega$ bounded in α.
Now we have to consider $P_{\alpha A'} f$ for $f \in \mathcal{B}_\omega$.

$$(P_{\alpha A'} f)(u) = \sum_{v=-\infty}^{m} f(v)\varphi(u+\alpha(u)-v) = \sum_{x=u+\alpha(u)-m}^{\infty} f(u+\alpha(u)-x)\varphi(x)$$

Hence, for $u \leq m$,

$$(1) \qquad \left|\frac{(P_{A'} f)(u)}{e^{|u|}}\right| = \frac{1}{e^{-u}}\left|\sum_{x=u+\alpha(u)-m}^{\infty} f(u+\alpha(u)-x)\varphi(x)\right| \leq$$

$$\leq {}_{-\infty}^{\ m}\|f\|_\omega \cdot \sum_{x=u+\alpha(u)-m}^{\infty} e^{x-\alpha(u)}\varphi(x) \leq {}_{-\infty}^{\ m}\|f\|_\omega \cdot e^{-R} \cdot \left(\sum_{x=0}^{\infty} e^x \varphi(x)\right)$$

where $\quad {}_{-\infty}^{\ m}\|f\|_\omega = \sup_{u \in (-\infty, m]} \frac{|f(u)|}{\omega(u)}$

For $m < u \leq M$ we get,

$$(2) \qquad \left|\frac{(P_{\alpha A'} f)(u)}{e^{|u|}}\right| \leq {}_{-\infty}^{\ m}\|f\|_\omega \cdot \frac{1}{e^{-u}} \sum_{x=u+\alpha(u)-m}^{\infty} e^{-u-\alpha(u)+x}\varphi(x) \leq$$

$$\leq {}_{-\infty}^{\ m}\|f\|_\omega \cdot \sum_{x=0}^{\infty} e^{x-\alpha(u)}\varphi(x) \leq {}_{-\infty}^{\ m}\|f\|_\omega \cdot \sum_{x=0}^{\infty} e^x \varphi(x)$$

From (1) and (2) it follows that $P_{\alpha A'}$ is indeed a bounded linear operator in

B_ω and that $\| P_{\alpha A'} \|_\omega$ is bounded in α.

Let $r := e^{-R} . \sum_{x=0}^{\infty} e^x \varphi(x)$, then by (1)

$$\frac{|(P^n_{\alpha A'} f)(u)|}{e^{|u|}} \le r^n \, {}_{-\infty}^{m}\|f\|_\omega \text{ for } u \le m$$

and by (2)

$$\frac{|(P^n_{\alpha A'} f)(u)|}{e^{|u|}} \le r^n . e^R . \, {}_{-\infty}^{m}\|f\|_\omega \text{ for } m < u \le M$$

These two relations imply the existence of an integer n and a $\rho < 1$ such that $\| P^n_{\alpha A'} \|_\omega \le \rho < 1$ for all $\alpha \in A$, which completes the proof of condition IIa. □

Since the set of possible actions in each state is finite, the conditions IIb,c are always satisfied and the condition IIa' is satisfied as soon as IIa is satisfied. The condition IId is satisfied for instance if $\varphi(x) > 0$ for all $x = 0,1,2,\ldots$

5. Exclusion of bad strategies

In inventory problems the one-period costs are usually assumed to be high for low inventory levels. That gives the idea that the recurrence conditions in I, II and theorem 2 are not so strong. Strategies under which the inventorylevel stays too low can not be good ones.
This will be formalized in this section.

First we state a new set of conditions.

IIIa There is a positive function h on V ($=(-\infty,M]$) such that
- i $c_\alpha(u) \ge h(u)$ for all $u \in V$
- ii h has a positive lower bound but is unbounded from above on each infinite set.

IIIb There is a strategy α_0 with average costs g_{α_0}

IIIc
- i $\varphi(x) > 0$ for all $x = 0,1,2,\ldots$
- ii there is an integer N such that $\varphi(x)$ is decreasing in x for $x > N$.

The conditions III imply that under strategies with average costs less than g_{α_0} there has to be a certain recurrency to finite sets.

<u>Theorem 3</u> Let the conditions III be satisfied and define $A' := \{\alpha \in A \mid g_\alpha \le g_{\alpha_0}\}$, then the condition IIa are satisfied for the inventory problem with A' as set of strategies.

<u>Proof</u> Choose the real number b such that $h(u) > 2g_{\alpha_0}$ for all $u \in (-\infty, b-1]$.

Let $B := (-\infty, b-1]$ and $A := [b, M]$. Then for all $\alpha \in A'$, $u \in V$, $\lim_{n\to\infty} (P^n_{\alpha B} 1)(u) < \frac{1}{2}$.

Let $\ell_{\alpha B}(u) := \lim_{n\to\infty} (P^n_{\alpha B} 1)(u)$.

Since $\ell_{\alpha B}(u) = (P_{\alpha B}\ell_{\alpha B})(u) = \lim_{n\to\infty} (P^n_{\alpha B}\ell_{\alpha B})(u)$ this implies that $\ell_{\alpha B}(u) = 0$ for all $u \in V$, (the inventorylevel returns to A almost surely).

By condition IIIc i the embedded Markov process $Q_{\alpha A}$ has a unique invariant probability (π_α).

Now the following modification of the process P_α is considered:

As soon as the process is N periods outside of A the transitionprobability $Q_{\alpha A}$ is applied instead of the transitionprobability P_α. That means that the state of the system jumps back to A without changing the embedded Markov process on A. The one-period costs are also changed, outside of A the costs are assumed to be equal to $2g_{\alpha_0}$ and on A equal to zero.

The average costs of the modified process are equal to

$$g^N_\alpha := \frac{2g_{\alpha_0} \cdot \sum_{u\in A} \pi_\alpha(u) t^N_\alpha(u)}{\sum_{u\in A} \pi_\alpha(u)\ (1+t^N_\alpha(u))}, \text{ where } t^N_\alpha(u) := \sum_{n=1}^{N} (P^n_{\alpha B}\ 1)(u)$$

If $g^N_\alpha > g_{\alpha_0}$ then also $g_\alpha > g_{\alpha_0}$, hence

$$\sum_{u\in A} \pi_\alpha(u)\ t^N_\alpha(u) \le 1 \text{ for all } \alpha \in A',\ N = 1,2,3,\ldots, \text{ and}$$

$$(1) \qquad (T_{\alpha A} 1)(u) = 1 + \sum_{n=1}^{\infty} (P^n_{\alpha B}\ 1)(u) \le 1 + \frac{1}{\pi_\alpha(u)} \text{ for } \alpha \in A',\ u \in A.$$

To get an upperbound for the recurrence costs we consider the same modified process, but with the costs changed in another way. The one-period costs, $c_\alpha^K(u)$, are assumed to be equal to $c_\alpha(u)$ if $c_\alpha(u) \le K$ and equal to K if $c_\alpha(u) > K$ (for some $K > 0$). The average costs are then equal to

$$g_\alpha^{NK} := \frac{\sum_{u\in A} \pi_\alpha(u) c_\alpha^{NK}(u)}{\sum_{u\in A} \pi_\alpha(u)(1+t_\alpha^N(u))}, \text{ where } c_\alpha^{NK}(u) := \sum_{n=0}^{N} (P_{\alpha B}^n c_\alpha^K)(u).$$

If $g_\alpha^{NK} > g_{\alpha_0}$ then also $g_\alpha > g_{\alpha_0}$, hence

$$\sum_{u\in A} \pi_\alpha(u)\, c_\alpha^{NK}(u) \le 2g_{\alpha_0} \quad \text{for all } \alpha \in A',\ K > 0,\ N=1,2,3,\ldots, \text{ and}$$

(2) $(T_{\alpha A} c_\alpha)(u) \le \dfrac{2g_{\alpha_0}}{\pi_\alpha(u)}$ for $\alpha \in A'$, $u \in A$

Using the conditions IIIc it is straightforward now to complete the proof. □

Notice that in the derivation of the relations (1) and (2) of the proof only the conditions III a,b are used and the inventory structure is not essential in these conditions.

In the rest of this section the conditions III a,b,c are assumed to be satisfied.
The set A is chosen as in the proof of theorem 3. The conditions Ia are satisfied for this set A with A' as set of strategies. Now an extra condition (IIIb') is considered which implies that the stronger recurrence conditions IIa are also satisfied.

<u>IIIb'</u> There is an $L > 0$ such that $(T_{\alpha_0 A} c_{\alpha_0})(u) < L.h(u)$ for all $u \in V$.

This condition implies that the recurrence costs (to A) are of the same shape as the one-period costs. This condition is only satisfied in general if it is possible to reach A in a finite number of steps, independent of the starting state. It is related to the Doeblin condition.

Before continuing with this condition it has to be remarked that the conditions Ia and Id imply the existence and uniqueness of the relative values of α, $v_\alpha(u)$, and that

$$v_\alpha(u) = c_\alpha(u) - g_\alpha + (P_\alpha v_\alpha)(u) = T_{\alpha A}(c_\alpha - g_\alpha)(u) + (Q_{\alpha A} v_\alpha)(u).$$

Furthermore, if $T_{\alpha A}(c_\alpha - g_{\alpha_0})(u) + (Q_{\alpha A} v_{\alpha_0})(u) \le v_{\alpha_0}$ for all $u \in V$ then $g_\alpha \le g_{\alpha_0}$ (a sort of policy improvement). This is easily seen by substituting the inequality $v_{\alpha_0} \ge T_{\alpha A}(c_\alpha - g_{\alpha_0}) + Q_{\alpha A} v_{\alpha_0}$ in its righthand side,

$$v_{\alpha_0} \ge T_{\alpha A}(c_\alpha - g_{\alpha_0}) + Q_{\alpha A} v_{\alpha_0} \ge \ldots\ldots \ge \sum_{\ell=0}^{n-1} Q_{\alpha A}^{\ell} T_{\alpha A}(c_\alpha - g_{\alpha_0}) + Q_\alpha^n v_{\alpha_0},$$

hence

$$0 \ge \lim_{n\to\infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q_{\alpha A}^{\ell}(T_{\alpha A} c_\alpha - g_{\alpha_0} . T_{\alpha A} 1) = \sum_{u \in A} \pi_\alpha(u)\{(T_{\alpha A} c_\alpha)(u) - g_{\alpha_0} . (T_{\alpha A} 1)(u)\}$$

and

$$g_\alpha = \frac{\sum_{u \in A} \pi_\alpha(u)(T_{\alpha A} c_\alpha)(u)}{\sum_{u \in A} \pi_\alpha(u)(T_{\alpha A} 1)(u)} \le g_{\alpha_0}$$

The conditions Ia and Id are satisfied (by theorem 3 and condition IIIc i) and this policy improvement property will be used to construct a set of strategies A^*, smaller than A', which also dominates A.

Define $A^* := \{ \alpha \in A' \mid T_{\alpha A}(c_\alpha - g_\alpha) + Q_{\alpha A} v_\alpha \le T_{\alpha_0 A}(c_{\alpha_0} - g_\alpha) + Q_{\alpha_0 A} v_\alpha \}$

<u>Lemma 4</u> A^* dominates A', for each $\alpha' \in A'$ there is an $\alpha^* \in A^*$ such that $g_{\alpha^*} \le g_{\alpha'}$.

<u>Proof</u> If $T_{\alpha_0 A}(c_{\alpha_0} - g_\alpha)(u) + (Q_{\alpha_0 A} v_\alpha)(u) < T_{\alpha A}(c_\alpha - g_\alpha)(u) + (Q_{\alpha A} v_\alpha)(u)$ for some $u \in V$ it is possible to construct a strategy α^* such that

$$T_{\alpha^* A}(c_{\alpha^*} - g_\alpha)(u) + (Q_{\alpha^* A} v_\alpha)(u) \le$$

$$\le \text{Min } \{T_{\alpha A}(c_\alpha - g_\alpha)(u) + (Q_{\alpha A} v_\alpha)(u), T_{\alpha_0 A}(c_{\alpha_0} - g_\alpha)(u) + (Q_{\alpha_0} v_\alpha)(u)\}$$

for all $u \in V$ (a result of negative dynamic programming, see Strauch [6]). The policy improvement property then implies $g_{\alpha^*} \le g_\alpha$. □

Now it will be shown that the conditions IIa are satisfied for the problem with A^* as set of strategies if the condition IIIb' is also satisfied.

Theorem 5 Let the condition IIIb' be satisfied, then the conditions IIa are satisfied for the problem with A^* as set of strategies and $\omega := h$.

Proof From condition IIIc i it follows that $|v_\alpha(u)|$ is bounded on A' for all $u \in A$. Let $K > 0$ be such that $|v_\alpha(u)| < K$ for all $u \in A$, $\alpha \in A'$.

Since $c_\alpha(u) > h(u) > 2g_{\alpha_0} \geq 2g_\alpha$ for $u \in B$, $\alpha \in A'$ we get

$$\tfrac{1}{2}(T_{\alpha A}c_\alpha)(u)-K \leq T_{\alpha A}(c_\alpha-g_\alpha)(u)+(Q_{\alpha A}v_\alpha)(u) \leq T_{\alpha_0 A}(c_{\alpha_0}-g_\alpha)(u)+(Q_{\alpha_0 A}v_\alpha) \leq L\cdot h(u)+K$$

for $u \in B$, $\alpha \in A^*$.

Together with the boundedness in α of $(T_{\alpha A}c_\alpha)(u)$for all $u \in A$, this implies the existence of a $\gamma > 0$ such that

$$(T_{\alpha A}c_\alpha)(u) \leq \gamma\cdot h(u) \text{ for all } \alpha \in A^*,\ u \in V$$

Hence (since $c_\alpha \geq h \geq (T_{\alpha A}c_\alpha)\ \frac{1}{\gamma}$)

$$P_{\alpha B}(T_{\alpha A}c_\alpha) = T_{\alpha A}c_\alpha - c_\alpha \leq (1-\frac{1}{\gamma})\ T_{\alpha A}c_\alpha$$

and $P_{\alpha B}$ is a contraction in $\mathcal{B}_\omega$ with $\omega := T_{\alpha A}c_\alpha$.

Now it is easy to prove that IIa is satisfied for $\omega := h$.

$T_{\alpha A}c_\alpha \in \mathcal{B}_h$ implies that $c_\alpha \in \mathcal{B}_h$.

It is shown already that $\| T_{\alpha A}c_\alpha \|_h \leq \gamma$, $\alpha \in A^*$

$$P_{\alpha B}h \leq P_{\alpha B}(T_{\alpha A}c_\alpha) \leq (1-\frac{1}{\gamma})T_{\alpha A}c_\alpha \leq (1-\frac{1}{\gamma})\gamma h,$$

hence $\| P_{\alpha B} \|_h$ is bounded on A^*.

Choose $\varepsilon > 0$ such that $\varepsilon\gamma < 1$ and let n_ε be such that $(1-\frac{1}{\gamma})^{n_\varepsilon} < \varepsilon$

then for all $u \in V$, $\alpha \in A^*$

$$(P_{\alpha B}^{n_\varepsilon}h)(u) \leq P_{\alpha B}^{n_\varepsilon}(T_{\alpha A}c_\alpha)(u) \leq (1-\frac{1}{\gamma})^{n_\varepsilon}(T_{\alpha A}c_\alpha)(u) \leq \varepsilon(T_{\alpha A}c_\alpha)(u) \leq \varepsilon\cdot\gamma\cdot h(u)$$

Hence $P_{\alpha B}^{n_\varepsilon}$ is a contraction in $\mathcal{B}_h$ and the contractionfactor is independent of α. □

We have proved now that the conditions IIIa,b,b',c imply condition IIa. But it is possible to proof that the conditions III imply all conditions II and hence the existence of an optimal strategy.

Corollary 6 The conditions IIIa,b,b',c imply the conditions IIa,a',b,c,d for the problem with A^* as set of strategies.

Proof Let the conditions IIIa,b,b',c be satisfied. Then the condition IIa is satisfied by theorem 5. Together with the finiteness of the numbers of possible actions in each state, this implies that the condion IIa' is also satisfied. Condition IId is satisfied by condition IIIc i. The finiteness of the set of possible actions implies the continuity of $P_\alpha(u,v)$ and $c_\alpha(u)$ on A and the compactness of A. Since $A^* \subset A$ condition IIb is satisfied. The only point to prove yet is the compactness of A^* or, since A is compact, the closedness of A^*.
Let $\alpha_1, \alpha_2, \alpha_3 \dots \in A^*$ converge to some $\alpha \in A$.
From theorem 5 we have the existence of an integer N and a $\rho < 1$ such that $P^N_{\alpha_i B} h \le \rho h$ for all $i=1,2,\dots$. Using methods as in the proof of lemma 1 it is possible to show that $P^N_{\alpha_0 B} h \le \rho h$. Together with the continuity of $P_\alpha(u,v)$ and $c_\alpha(u)$ in α this implies that α_0 is also an element of A^*. □

REFERENCES

[1] Bather, J. (1973): Optimal decision procedures for finite Markov chains, part II: communicating systems. Adv. in Appl. Prob. 5, 521-540

[2] Blackwell, D. (1962): Discrete dynamic programming, Ann.Math. Statist, 33, 719-726

[3] Hordijk, A. (1974): Dynamic programming and Markov potential theory. Math. Centre Tracts, no. 51, Amsterdam

[4] van Nunen, J.A.E.E. and Wessels J. (1975): A note on dynamic programming with unbounded rewards, Eindhoven University of Technology, Dept. of Math. (Memorandum COSOR 75-13)

[5] Ross, S.M. (1968): Nondiscounted denumerable Markovian decision models. Ann. Math. Statist. 39, 412-423

[6] Strauch, R.E. (1966): Negative dynamic programming. Ann.Math.Statist. 37, 871-889

[7] Tyms, H.C. (1975): On dynamic programming with arbitrary state space, compact action space and the average return as criterion. Report BW 55/75, Math. Centre, Amsterdam

[8] Wijngaard, J: Stationary Markovian decision problems and Perturbation theory of linear operators. Math. of Oper. Res. (to appear).