Stationary Markovian decision problems II

Citation for published version (APA):

Document status and date:
Published: 01/01/1975

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Memorandum COSOR 75-15

Stationary Markovian decision problems II

by

J. Wijngaard

Eindhoven, September 1975

The Netherlands
Stationary Markovian decision problems II

by

J. Wijngaard

1. Introduction

A stationary Markovian decision problem (SMD) is a set of pairs \( \{(P_\alpha, r_\alpha)\} \), \( \alpha \in A \) where \( P_\alpha \) is a Markov process and \( r_\alpha \) a nonnegative function on the state space (cost function). The elements \( \alpha \in A \) are called strategies. In [5] conditions are derived for the existence of an average optimal strategy. The most important conditions were the boundedness of \( r_\alpha \) and the quasi-compactness of \( P_\alpha \), which is equivalent to the Doeblin-condition. For a countable state space the Doeblin-condition for a Markov process \( P \) is equivalent to the existence of a finite set \( A \), an integer \( n \), and an \( \varepsilon > 0 \), such that the probability of being in the set \( A \) after \( n \) transitions \( P^{(n)}(u,A) \geq \varepsilon \) for each starting state \( u \). To show how severe this condition and hence quasi-compactness is we consider the following inventory problem:

At the beginning of each period the inventory level is assumed to be \( \ldots \ldots -2, -1, 0, 1, 2, \ldots \ldots \). One may order a quantity of at most \( R \) units, the delivery is instantaneous. During the period there is a demand for \( 0, 1, 2, \ldots \) units with a probability of \( p_0, p_1, p_2, \ldots \). The transition probability under order policy \( \alpha \) is \( P_\alpha(i,j) = p_{i+\alpha(i)-j} \) (the quantity to order in state \( i \)).

If \( R \) is large enough there are policies \( \alpha \) such that for each state \( i \) one can find a finite set \( A \), an integer \( n \), and an \( \varepsilon > 0 \) such that \( P^{(n)}_\alpha(i,A) \geq \varepsilon \). However, if \( j \) is more than \( nR \) units below the lowest element of \( A \), then \( P^{(n)}_\alpha(j,A) = 0 \). Hence there is no policy such that the corresponding Markov process satisfies the Doeblin-condition. Such decision processes can be studied by introducing embedded Markov processes. In this paper we do not assume the quasi-compactness of \( P_\alpha \) nor the boundedness of \( r_\alpha \). Instead of that we state the existence of a subset \( A \) of the state space such that the embedded Markov process of \( P_\alpha \) on \( A \) exists and is quasi-compact for all \( \alpha \in A \) and the recurrence time and costs until \( A \) are bounded on \( A \).

Embedded Markov processes are introduced in section 2. We derive some properties of Markov processes with a quasi-compact embedded Markov process. Section 3 deals with the existence of the average costs for these Markov processes with unbounded cost function. The continuity of the average costs and
the existence of an optimal $\alpha$ for problems $\{(P_\alpha, r_\alpha)\}$, $\alpha \in A$ is worked out in section 4. In section 5 the results of section 4 are applied to the case with a countable state space and related to the results of Ross [3] and Hordijk [1].

2. Embedded Markov processes

We assume that $P$ is a Markov process on the measurable space $(V, \Sigma)$. For $A \in \Sigma$ the sub-markov process $I_A$ is defined by the sub-transition probability

$$I_A(u, E) := 1_{A \cap E}(u), \quad u \in V, \ E \in \Sigma.$$ 

Instead of $P|_A$ we shall write $P_A$.

The next lemma serves as an introduction to the concept of an embedded Markov process.

**Lemma 1.** Let $A \in \Sigma$, $B := V \setminus A$. Define the function $Q$ on $V \times \Sigma$ by

$$Q(u, E) = \sum_{n=0}^{\infty} (P^n|_B P_A|_E)(u) \quad \text{for all } u \in V, \ E \in \Sigma.$$ 

Then $Q$ is a sub-transition probability on $V \times \Sigma$, the operator $Q$ on $B(V, \Sigma)$ is given by

$$\text{(1)} \quad (Qf)(u) = \sum_{n=0}^{\infty} (P^n|_B A)(u) \quad \text{for } u \in V, \ f \in B(V, \Sigma),$$

and the operator $Q$ on $M(V, \Sigma)$ by

$$\text{(2)} \quad (\mu Q)(E) = \sum_{n=0}^{\infty} (\mu P^n|_B A)(E) \quad \text{for } E \in \Sigma, \ \mu \in M(V, \Sigma).$$

Furthermore, $Q$ is a Markov process on $(V, \Sigma)$ if and only if

$$\lim_{n \to \infty} (P^n|_B l_V)(u) = 0 \quad \text{for all } u \in V.$$ 

**Proof.** We have $P_A = P - P_B$. Hence

$$\sum_{n=0}^{N} P^n|_B A|_V = \sum_{n=0}^{N} P^n|_B l_V - \sum_{n=0}^{N} P^n|_B A|_V = \sum_{n=0}^{N} P^n|_B l_V - \sum_{n=0}^{N} P^n|_B l_V = 1 - P^n|_B l_V.$$
which implies that \( Q(u,E) \leq Q(u,V) \leq 1 \) for \( u \in V, E \in \Sigma \). The measurability of \( Q \) as function of \( u \) and the \( \sigma \)-additivity as function of \( E \) are easy to verify. Hence \( Q \) is a sub-transition probability on \( V \times \Sigma \) and a transition probability if and only if \( \lim_{n \to \infty} (P^n_{B\mid V})(u) = 0 \) for all \( u \in V \). The equations (1) and (2) are direct consequences of the definition of \( Q \).

Let \( A \in \Sigma, B := V\setminus A \). The sub-Markov process \( Q \) on \((V, \Sigma)\) with sub-transition probability

\[
Q(u,E) := \sum_{n=0}^{\infty} (P^n_{B\mid A\mid E})(u), \quad u \in V, E \in \Sigma
\]

is called the sub-Markov process of \( P \) induced by \( A \).

It is clear that the restriction of \( Q \) to \( A \times \Sigma_A \) is a sub-transition probability on \( A \times \Sigma_A \). The sub-Markov process on \((A, \Sigma_A)\) corresponding to this sub-transition probability is called the embedded sub-Markov process of \( P \) on \( A \). By lemma 1 the sub-Markov process induced by \( A \) is a Markov process if and only if \( \lim_{n \to \infty} (P^n_{B\mid V})(u) = 0 \) for all \( u \in V \), that means, if the probability that the system will never reach the set \( A \) is zero.

The relationship between invariant sets, functions and measures of \( P \) and those of a process \( Q \) of \( P \) induced by a set \( A \in \Sigma \) is shown in the next lemma.

Lemma 2. Let \( A \in \Sigma, B := V\setminus A \). Assume that \( \lim_{n \to \infty} (P^n_{B\mid V})(u) = 0 \) for all \( u \in V \) and let \( Q \) be the embedded Markov process of \( P \) on \( A \). If \( \mu \in M(V, \Sigma) \) and \( f \in B(V, \Sigma) \) are invariant under \( P \), then \( \mu I_A Q = \mu I_A \) and \( Qf = f \). Conversely, if \( Qf = f \), then \( Pf = f \) and if \( E \) is an invariant set under \( Q \) then

\[
\mathbb{E} := \{u \mid Q(u,E) = 1\}
\]

is an invariant set under \( P \).

Proof. The proof of the invariance of \( \mu I_A \) and \( f \) under \( Q \) is straightforward using

\[
\mu I_A P^n_{B\mid A} = (\mu - \mu I_B) P^n_{B\mid A} = \mu P^n_{B\mid B\mid A} - \mu I_{B\mid B\mid A} = \mu I_{B\mid B\mid A} p^n_{B\mid B\mid A} - \mu I_{B\mid B\mid A} p^n_{B\mid B\mid A},
\]

and
Conversely, suppose \( Qf = f \). Then

\[
\begin{align*}
P^n_B f &= P^n_B f - P^{n+1}_B f = P^n_B f - P^n_B f. \\
\end{align*}
\]

Finally, let \( E \) be an invariant set under \( Q \) and let \( \overline{E} := \{ u \mid Q(u,E) = 1 \} \). From \( Q = P_A + P_B Q \) we conclude

\[
Q^n_A\setminus E = P^n_A\setminus E + P^n_B \overline{Q^n_A\setminus E} \geq P^n_A\setminus E + P^n_B \overline{Q^n_A\setminus E}.
\]

Since on \( \overline{E} \) we have \( Q^n_A\setminus E = 0 \), it follows that \( P^n_A\setminus E = 0 \) on \( \overline{E} \) and \( P^n_B \overline{Q^n_A\setminus E} = 0 \) on \( \overline{E} \). From \( Q^n_A\setminus E = 1 \) and the definition of \( \overline{E} \) we infer that on \( V\setminus E \) and in particular on \( B\setminus E \) we have \( Q^n_A\setminus E > 0 \). It follows that \( P^n_B\setminus E = 0 \) on \( \overline{E} \). Therefore \( P^n_{V\setminus E} = 0 \), \( P^n_{\overline{E}} = 1 \) on \( \overline{E} \).

3. Average costs

Let \( P \) be a Markov process on \((V,\mathcal{E})\) and \( r \) a nonnegative measurable function on \( V \). The average costs of \((P,r)\), starting in \( u \), \( g(u) \), are equal to

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} (P^k r)(u), \text{ if this limit exists. (The integral } \int_V P(u,ds)f(s), \text{ if existing, is denoted by } (Pf)(u).)
\]

In this section conditions sufficient for the existence of the average costs will be given and it will be shown that these costs can be written as the quotient of the recurrence costs and recurrence time to a set \( A \). This will be done by considering the equations

\[
x = Px \\
y = r - x + Py
\]

in the complex valued measurable functions \( x, y \) on \( V \). These equations are called the \((P,r)\)-equations. If \( P \) is quasi-compact and \( r \) is bounded, the existence of a solution of the \((P,r)\)-equations is a consequence of the strong ergodic theorem (see [4]). The solution is given by

\[
x := \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k r = g
\]
where $d$ is an integer such that $\lambda^d = 1$ for all eigenvalues $\lambda$ of $P$ on the unit circle. Now the existence of a solution of the $(P,r)$-equations will be proved under somewhat weaker conditions. The quasi-compactness of $P$ is replaced by the quasi-compactness of the embedded Markov process of $P$ on some set $A \in \mathcal{E}$. The boundedness of $r$ is replaced by the boundedness of the expected time and expected costs until the first recurrence to $A$. The function $x$ will again turn out to be equal to the average costs.

**Definition 3.** Let $f$ be a nonnegative measurable real valued function on $\mathcal{V}$ and let $A$ be a measurable set. The Markov process $P$ is said to be $(A,f)$-recurrent if

i) $P^m_B f$ exists for all $m \in \mathbb{N}$, ($B := \mathcal{V}\setminus A$),

ii) the sum $\sum_{m=0}^{\infty} (P^m_B f)(u)$ exists for all $u \in \mathcal{V}$,

iii) the convergence of $\sum_{m=0}^{\infty} (P^m_B f)(u)$ is uniform on $A$ and $\sum_{m=0}^{\infty} (P^m_B f)(u)$ is bounded on $A$.

In the rest of this section we assume that $A$ is a fixed measurable set such that $P$ is $(A,\mathcal{V})$-recurrent and $(A,r)$-recurrent and further that the embedded Markov process $Q$ of $P$ on $A$ is quasi-compact, ($Q$ interpreted as a Markov process on $(A,\mathcal{E}_A)$). The $(A,\mathcal{V})$-recurrency implies that the embedded sub-Markov process of $P$ on $A$ is a Markov process.

Let $E_j, j = 1, \ldots, n$ be the maximal invariant sets of $Q$, $F := \bigcup_j E_j$, and $\Delta := A \setminus F$.

**Theorem 4.** The $(P,r)$-equations have a solution.

**Proof.** By the strong ergodic theorem the spectral radius of $Q_\Delta := Q_{|\Delta}$ is smaller than 1. Hence, each of the equations $x = Q^\infty_{E_j}1_{E_j} + Q_{\Delta}x$ in $B(A,\mathcal{E}_A)$ has a unique solution

$$
S_j := \sum_{k=0}^{\infty} Q_\Delta^k Q_{E_j} 1_{E_j}.
$$
Using \((Q_\Delta f)(u) = 0\) for all \(f \in \mathcal{B}(A, E_n)\), \(u \in E_j\), \(j = 1, 2, \ldots, n\), we get
\[g_j(u) = 1\] for \(u \in E_j\) and \(g_j(u) = 0\) for \(u \in E_i\) if \(i \neq j\). This means that
\[Q_{E_j} E_j = Q_{E_j} g_j = (Q - Q_\Delta) g_j\] and that \(g_j\) is a solution of the equation
\[x - Qx = 0\] in \(\mathcal{B}(A, E_n)\).

It is possible to extend \(g_j\) to a solution \(g_j^*\) of the equation \(x - Qx = 0\) in \(\mathcal{B}(V, E)\) by defining \(g_j^* := Q g_j\), where \(Q\) is used as an operator on \(\mathcal{B}(A, E_n)\) to \(\mathcal{B}(V, E)\).

Each function \(g_j^*, j = 1, \ldots, n\) is a solution of the equation \(x = P x\) since
\[P g_j^* = P B g_j^* + P A g_j^* = P B Q g_j^* + P A g_j^* = \sum_{\ell=1}^{\infty} P B^{\ell} P A g_j^* + P A g_j^* = Q g_j^* = g_j^* .\]

The problem is to choose a linear combination \(x\) of the \(g_j^*\) such that the equation \(y = r - x + P x\) has also a solution. Let \(Q_j\) be the restriction of \(Q\) to \(E_j \times E_j\). The \((A, r)\)-recurrency and \((A, I, V)\)-recurrency of \(P\) imply the boundedness on \(A\) of the functions \(\sum_{\ell=0}^{\infty} P B^\ell r\) and \(\sum_{\ell=0}^{\infty} P B^\ell g_j^*\), \(j = 1, \ldots, n\). For convenience we shall write \(T f\) instead of \(\sum_{\ell=0}^{\infty} P B^\ell f\) for \(f = r\) or \(f\) is bounded.

Notice that \(P B T f = \sum_{\ell=1}^{\infty} P B^\ell f\).

The restrictions of \(T r\) and \(T g_j^*\) to \(E_j\) are elements of \(\mathcal{B}(E_j, E_j)\). Therefore both \(\lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q_j^\ell T r\) and \(\lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q_j^\ell T g_j^*\) are elements of \(\text{N}(I - Q_j)\). Since \(\dim \text{N}(I - Q_j) = 1\) there is a constant \(c_j\) such that
\[\lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q_j^\ell (Tr - c_j T g_j^*) = 0 .\]

Using this it is straightforward to prove that
\[\lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q_j^\ell (Tr - T g_j^*) = 0\] on \(A\), where \(g^* := \sum_{j=1}^{n} c_j g_j^*\).

We shall show that the equation \(y = r - g^* + Py\) has a solution.

Let the integer \(d\) be such that \(\lambda^d = 1\) for all eigenvalues \(\lambda\) of \(Q\) on the unit circle. By the strong ergodic theorem the function \(f'\) on \(A\) defined by
\[f' := \frac{1}{d} \sum_{m=0}^{d-1} \lim_{k \to \infty} \sum_{\ell=0}^{kd-1+m} Q_j^\ell (Tr - T g_j^*)\]
is a solution of the equation \(y = Tr - T g_j^* + Qy\) in \(\mathcal{B}(A, E_n)\). The function \(f'\) can be extended to a function \(f\) on \(V\) by defining \(f := Tr - T g_j^* + Q f'\). The
function $f$ is a solution of the equation $y = r - g^* + Py$. This can be seen as follows:

$$Pf = P_B f + P_A f = P_B (Tr - Tg^* + Qf) + P_A f = P_B T(r - g^*) + P_B Qf + P_A T(r - g^*) + \sum_{k=1}^{\infty} P_B^k P_A f = P_B T(r - g^*) + Qf = P_B T(r - g^*) + f - T(r - g^*) = f + \sum_{k=0}^{\infty} P_B^k (r - g^*) - r + g^*.$$

Now some properties of the solution $g^*$, $f^*$ of the $(P,r)$-equations are given. The next lemma is preliminary.

**Lemma 5.** Let

$$Tr := \sum_{k=0}^{\infty} P_B^k r \quad \text{and} \quad Tl_{1V} := \sum_{k=0}^{\infty} P_B^k l_{1V}.$$

Then $P^m Tr$ and $P^m Tl_{1V}$ exist for all $m \in \mathbb{N}$ and

$$\lim_{m \to \infty} \frac{1}{m} (P^m Tr)(u) = \lim_{m \to \infty} \frac{1}{m} (P^m Tl_{1V})(u) = 0 \quad \text{for all} \quad u \in V.$$

**Proof.** Substitution of $P = P_A + P_B$ in $P^{m+1}$ yields

$$P^{m+1} = P_{m+1}^m P_A + P_{m+1}^m P_B = P_{m+1}^m P_A + P_{m+1}^m P_B + \cdots + \sum_{k=0}^{m-1} P_B^k P_A + P_B^k = \sum_{k=0}^{m} P_B^k P_A + P_B^k.$$

Hence

$$P^{m+1} Tr = \sum_{k=0}^{m} P_B^k P_A Tr + \sum_{k=0}^{m} P_B^k Tr + \sum_{k=m+1}^{\infty} P_B^k Tr \leq \sum_{k=0}^{m} P_B^k P_A Tr + Tr.$$

The existence of $P^{m+1} Tr$ is implied by the existence of $Tr$ and the boundedness of $Tr$ on $A$. The existence of $P^{m+1} Tl_{1V}$ is proved similarly. For each $\epsilon > 0$ there is an integer $N_\epsilon$ such that
For \( m > N \) we have
\[
F^{m+1} = \sum_{k=0}^{N} F^{m-k} P_{A} \sum_{k=N+1}^{m} P^{k} P_{B} + \sum_{k=0}^{m} P^{k} P_{A} \sum_{k=0}^{m} P^{k} P_{B} + \sum_{m=+1}^{\infty} P^{k} P_{B}.
\]

Let \( \| \text{Tr} \|_{A} := \sup_{u \in A} (\text{Tr}(u)) \). Then
\[
(P^{m+1} \text{Tr})(u) \leq (N + 1) \| \text{Tr} \|_{A} + (m - N) \epsilon + \sum_{j=m+1}^{\infty} (P^{j} \text{Tr})(u).
\]

Using standard arguments we can show that \( \lim_{m \to \infty} \frac{1}{m} (P^{m} \text{Tr})(u) = 0 \) for all \( u \in V \).

That \( \lim_{m \to \infty} \frac{1}{m} (P^{m} \text{Tr})(u) = 0 \) can be proved similarly.

**Theorem 6.** Let the functions \( g^{*} \) and \( f \) be as constructed in the proof of theorem 4. Then \( P^{m} f \) exists for all \( m \in N \), \( \lim_{m \to \infty} \frac{1}{m} P^{m} f = 0 \), and \( g^{*} = g \) (the average 

**Proof.** The functions \( g^{*} \) and \( f \) on \( V \) were defined in the following way:
\[
g^{*} := \sum_{j=1}^{n} c_{j} g_{j}^{*}, \quad \text{where } g_{j}^{*} := Qg_{j} \text{ and } g_{j} \in \mathcal{B}(A, \Sigma_{A});
\]
\[
f := \text{Tr} - Tg^{*} + Qf', \quad \text{where } f' \in \mathcal{B}(A, \Sigma_{A}).
\]

Hence \( g^{*} \) and \( Qf' \) are bounded. By lemma 5 \( P^{m} f \) exists for all \( m \in N \) and
\[
(1) \quad \lim_{m \to \infty} \frac{1}{m} (P^{m} f) = 0.
\]

Repeated substitution of \( f(u) = r(u) - g^{*}(u) + (Pf)(u) \) in its right-hand side yields
\[
f(u) = \sum_{l=0}^{m-1} (P^{l} r)(u) - \sum_{l=0}^{m-1} (P^{l} g^{*})(u) + (P^{m} f)(u).
\]

Hence
\[
\frac{f(u) - (P_m f)(u)}{m^2} = \frac{1}{m} \sum_{\lambda=0}^{m-1} (P^\lambda r)(u) - g^*(u)
\]
and by (1)
\[
g^*(u) = \lim_{m \to \infty} \frac{1}{m} \sum_{\lambda=0}^{m-1} (P^\lambda r)(u) = g(u) \quad \text{for all } u \in V.
\]

Now we consider the solution \((g_1, f_1)\) of the \((P, r)\)-equations. As for the solution \((g^*, f)\) we can prove
\[
g_1(u) = \lim_{m \to \infty} \frac{1}{m} \sum_{\lambda=0}^{m-1} (P^\lambda r)(u).
\]
Further the function \(f - f_1\) satisfies \(f - f_1 = P(f - f_1)\), hence by lemma 2, \(f - f_1 = Q(f - f_1)\).

For \(u \in E_j\) it is possible to write the average costs in a somewhat different way. To show this we need the following lemma.

**Lemma 7.** For all \(f \in B(V, \Sigma)\) and for all \(m \in \mathbb{N}\) the following relation holds,
\[(B := V \setminus A)
\[
(P_B^m Qf)(u) = (P_B^m Qf_{E_j})(u), \quad u \in E_j, \ j = 1, 2, \ldots, n.
\]

**Proof.** It is sufficient to prove the assertion for nonnegative functions, namely, each \(f \in B(V, \Sigma)\) can be written as \(f = f_1 - f_2 + i(f_3 - f_4)\), where the functions \(f_1, f_2, f_3, f_4\) are nonnegative elements of \(B\). Now assume that \(f\) is a nonnegative function in \(B\). Substitution of \(Qf_E = \sum_{\lambda=0}^{\infty} P_B^\lambda f_E\) in \(P_B^m Qf_E\) yields
\[
(1) \quad \sum_{\lambda=0}^{\infty} P_B^\lambda f_E \leq \sum_{\lambda=0}^{\infty} P_B^\lambda f_E = Qf_E \quad \text{for all } E \in \Sigma.
\]
Furthermore
\[
(2) \quad (Qf)(u) = (Qf_{E_j})(u) \quad \text{for } u \in E_j, \ j = 1, \ldots, n.
\]
Let \(E_j := V \setminus E_j\), then \(f = f_{E_j} + f_{E_j}^\perp\). By (1) and (2)
\[ \left( P^{n} B_{E}^{Q} f_{E}^{j} \right)(u) \leq \left( Q f_{E}^{j} \right)(u) = 0 \quad \text{for} \ u \in E. \]

Hence

\[ \left( P^{n} B_{E}^{Q} f \right)(u) = \left( F P^{n} Q f_{E}^{j} \right)(u) \quad \text{for} \ u \in E, \ j = 1, \ldots, n. \]

---

**Theorem 8.** For \( u \in E, \ j = 1, \ldots, n \)

\[
\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (P r_{E}^{j})(u) = \frac{\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (Q r_{E}^{j})(u)}{\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (Q T l_{E}^{j})(u)}.\]

---

**Proof.** Let \( g^{*} \) be as constructed in the proof of theorem 4. Then \( g^{*}(u) = c_{j} \) for \( u \in E, \) where

\[
c_{j} = \frac{\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (Q r_{E}^{j})(u)}{\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (Q T g_{E}^{j})(u)}.\]

But for \( u \in E, \)

\[
(Q r_{E}^{j})(u) = (Q r_{E}^{j})(u)
\]

and

\[
(Q T g_{E}^{j})(u) = (Q T g_{E}^{j})(u).
\]

Further

\[
T g_{E}^{j} = T Q g_{E}^{j} \quad \text{and} \quad T l_{E}^{j} = T Q l_{E}^{j}.
\]

By lemma 7

\[
(T Q g_{E}^{j})(u) = (T Q l_{E}^{j})(u) = (T Q l_{E}^{j})(u) \quad \text{for} \ u \in E. \]

Hence \( (T g_{E}^{j})(u) = (T l_{E}^{j})(u) \) for \( u \in E. \) This completes the proof.

A more general result of this type is given by de Leve [2], part II, lemma 1.57.
4. Stationary Markovian decision problems

In this section we consider a stationary Markovian decision problem \( \{(P_\alpha, r_\alpha)\} \), \( \alpha \in A \) on \((V, E)\) (for a definition see [5]). In [5] it was assumed that \( P_\alpha \) is quasi-compact and \( r_\alpha \) bounded. Now we assume the existence of a measurable set \( A \) such that

i) for all \( \alpha \in A \) the Markov process \( P_\alpha \) is \((A, 1_V)\)-recurrent and \((A, r_\alpha)\)-recurrent,

ii) the embedded Markov process \( Q_\alpha \) of \( P_\alpha \) on \( A \) is quasi-compact for all \( \alpha \in A \), \((Q_\alpha \) is interpreted as a Markov process on \((A, E_A))\),

iii) the functions \( \sum_{n=0}^{\infty} p^n_{\alpha B} 1_V \) and \( \sum_{n=0}^{\infty} p^n_{\alpha B} r_\alpha \) on \( A \), with \( B = V \setminus A \), are uniformly bounded on \( A \).

Let for all \( \alpha \in A \), \( n_\alpha \) be the dimension of \( N(I - Q_\alpha) \), \( E_{\alpha j} \) for \( j = 1, \ldots, n_\alpha \) the maximal invariant sets of \( Q_\alpha \), and \( \pi_{\alpha j} \) the invariant probabilities of \( Q_\alpha \) with support \( E_{\alpha j} \). Let

\[
E_\alpha := \sum_{j=1}^{n_\alpha} E_{\alpha j} \quad \text{and} \quad S_\alpha := \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} Q_\alpha^{j}.
\]

By theorem 6 the average costs of \( (P_\alpha, r_\alpha) \) starting in \( u \),

\[
g_\alpha(u) := \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \langle P_\alpha^j r_\alpha \rangle(u)
\]

exist for all \( \alpha \in A \) and \( u \in V \). In theorem 8 we proved

\[
g_\alpha(u) = \frac{(S_\alpha T_\alpha r_\alpha)(u)}{(S_\alpha T_\alpha 1_V)(u)} \quad \text{for} \ u \in E_\alpha,
\]

where

\[
T_\alpha f = \sum_{n=0}^{\infty} p^n_{\alpha B} f.
\]

Hence \( g_\alpha \) is constant on \( E_{\alpha j} \) for \( j = 1, \ldots, n_\alpha \). These constants are denoted by \( g_{\alpha j} \). Let \( \rho \) be a metric on \( A \) such that

iv) \( \lim_{\rho(\alpha, \alpha_0)\to0} \|Q_\alpha - Q_{\alpha_0}\| = 0 \) for all \( \alpha_0 \in A \),

v) \( \lim_{\rho(\alpha, \alpha_0)\to0} |\pi_{\alpha j}(T_\alpha r_\alpha) - \pi_{\alpha_0 j}(T_{\alpha_0} r_{\alpha_0})| = 0 \)
for all \( \alpha_0 \in A \) and \( j = 1, \ldots, n_{\alpha_0} \).

\[
\lim_{\rho(\alpha, \alpha_0) \to 0} |\pi_{\alpha_0}^j(T_{\alpha} \mathbf{1}_V) - \pi_{\alpha_0}^j(T_{\alpha_0} \mathbf{1}_V)| = 0
\]

for all \( \alpha_0 \in A \) and \( j = 1, \ldots, n_{\alpha_0} \).

For \( A = V \) these assumptions are identical to the assumptions made in [5], section 4.

In the next two subsections we consider the continuity of \( g_\alpha \) and the existence of an optimal strategy.

4.1. Continuity of \( g_\alpha \)

Let \( A_n \) be the subset of \( A \) with all \( \alpha \) such that \( n_\alpha = n \). In [5] lemma 2.9, it was shown that the assumption iv) implies the continuity of \( \mathbf{S}_\alpha \) as operator valued function on \( A_n \). This is used in the next lemma.

**Lemma 9.** Let \( n \in \mathbb{N} \) and \( \alpha_0 \in A_n \). Then there is a \( \delta > 0 \) such that for all \( \alpha \in A_n \) with \( \rho(\alpha, \alpha_0) < \delta \) and for all \( i = 1, 2, \ldots, n \), \( \pi_{\alpha_0}^i(E_{\alpha j}) > 0 \) for precisely one \( j \in \{1, 2, \ldots, n\} \).

Consider the set \( A_{n_\delta} := \{ \alpha \in A_n \mid \rho(\alpha, \alpha_0) < \delta \} \). Let for all \( \alpha \in A_{n_\delta} \) and \( i \in \{1, 2, \ldots, n\} \) the integer \( i_\alpha \) be defined by \( \pi_{\alpha_0}^i(E_{\alpha j}) > 0 \). Then for all \( i = 1, 2, \ldots, n \) the functions \( \pi_{\alpha_0}^i(E_{\alpha j} \setminus E_{\alpha i}) \), \( \| \pi_{\alpha_0}^i - \pi_{\alpha i} \| \), and \( |g_{\alpha_0}^i - g_{\alpha i}^j| \) on \( A_n \) converge to 0 if \( \rho(\alpha, \alpha_0) \) converges to 0.

**Proof.** The continuity of \( \mathbf{S}_\alpha \) on \( A_n \) implies the existence of a \( \delta > 0 \) such that \( \| \mathbf{S}_\alpha - \mathbf{S}_{\alpha_0} \| < \frac{1}{4} \) for all \( \alpha \in A_n \) with \( \rho(\alpha, \alpha_0) < \delta \).

Let \( i \in \{1, 2, \ldots, n\} \) and let \( \nu_i \) be some probability on \( E_A \) with \( \nu_i(E_{\alpha 0 i}) = 1 \), then \( \pi_{\alpha_0}^i(E_{\alpha 0}) = (\nu_i \mathbf{S}_\alpha)(E_{\alpha 0}) = 1 \). Hence \( (\nu_i \mathbf{S}_\alpha)(E_{\alpha 0}) > \frac{3}{4} \) for all \( \alpha \in A_n \) with \( \rho(\alpha, \alpha_0) < \delta \). But

\[
(\nu_i \mathbf{S}_\alpha)(E_{\alpha 0}) = (\nu_i \mathbf{S}_\alpha)(E_{\alpha 0}) \cap \bigcup_{j=1}^n (E_{\alpha j})
\]

and therefore
for all $\alpha \in A_n$ with $\rho(\alpha, \alpha_0) < \delta$. This implies the existence of at least one $j \in \{1, 2, \ldots, n\}$ such that $(\vee_{i \in A_0} (E_{aj}^j) = \pi_{\alpha_0}^j (E_{aj}^j) > 0$ for $\alpha \in A_n$ with $\rho(\alpha, \alpha_0) < \delta$.

Suppose that for some $\alpha \in A_n$ with $\rho(\alpha, \alpha_0) < \delta$ there are two $j$'s, $j_1$ and $j_2$ such that $\pi_{\alpha_0}^j (E_{aj}^j) > 0$. Let the probabilities $\nu_{ij_1}$ and $\nu_{ij_2}$ on $\Sigma_A$ be given by

$$
\nu_{ij_1}(E) = \frac{\pi_{\alpha_0}^j (E \cap E_{aj_1}^j)}{\pi_{\alpha_0}^j (E_{aj_1}^j)} \quad \text{and} \quad \nu_{ij_2}(E) = \frac{\pi_{\alpha_0}^j (E \cap E_{aj_2}^j)}{\pi_{\alpha_0}^j (E_{aj_2}^j)}.
$$

Then $\nu_{ij_1} \alpha_0^j = \nu_{ij_2} \alpha_0^j = \pi_{\alpha_0}^j$. Using $\nu_{ij_1} \alpha_0^j (E_{aj_1}^j) = 1$ and $\nu_{ij_2} \alpha_0^j (E_{aj_2}^j) = 1$ it is easy to see that

$$
\pi_{\alpha_0}^j (E_{aj_1}^j) = (\nu_{ij_1} \alpha_0^j) (E_{aj_1}^j) > 1
$$

and

$$
\pi_{\alpha_0}^j (E_{aj_2}^j) = (\nu_{ij_2} \alpha_0^j) (E_{aj_2}^j) > 1.
$$

The disjunctness of $E_{aj_1}^j$ and $E_{aj_2}^j$ implies $\pi_{\alpha_0}^j (E_{aj_1}^j \cup E_{aj_2}^j) > 1$ which contradicts the fact that $\pi_{\alpha_0}^j$ is a probability. This completes the proof of the first part of the lemma.

Now let for all $\alpha \in A_n$ and $i \in \{1, 2, \ldots, n\}$ the integers $i_{\alpha}^j$ be such that $\pi_{\alpha_0}^j (E_{ai_{\alpha}^j}) > 0$. The probability $\nu_{ii_{\alpha}^j}$ on $\Sigma$ is given by

$$
\nu_{ii_{\alpha}^j}(E) := \frac{\pi_{\alpha_0}^j (E \cap E_{ai_{\alpha}^j})}{\pi_{\alpha_0}^j (E_{ai_{\alpha}^j})}.
$$

Then $\pi_{ai_{\alpha}^j} = \nu_{ii_{\alpha}^j} \alpha_0^j$ and $\pi_{\alpha_0}^j = \nu_{ii_{\alpha}^j} \alpha_0^j$. Hence $\|\pi_{ai_{\alpha}^j} - \pi_{\alpha_0}^j\| \leq \|S_\alpha - S_\alpha_0\|$ and therefore

$$
\lim_{\rho(\alpha, \alpha_0) \to 0} \|\pi_{ai_{\alpha}^j} - \pi_{\alpha_0}^j\| = 0.
$$

Furthermore $\pi_{ai_{\alpha}^j} (E_{ai_{\alpha}^j} \setminus E_{ai_{\alpha}^j}) = 0$ and hence
For \( j = 1, \ldots, n \) we have
\[
\lim_{\rho(\alpha, \alpha_0) \to 0} \pi_0^\alpha (E_\alpha \setminus E_\alpha) = 0.
\]

For \( j = 1, \ldots, n \) we have
\[
\mathcal{S}_{\alpha} \mathcal{T}_{\alpha} (u) = \pi_{\alpha j} \mathcal{T}_{\alpha} (u) \text{ for } u \in \mathcal{E}_{\alpha j}.
\]

Using (1), the uniform boundedness of \( \mathcal{T}_{\alpha} \) and \( \mathcal{T}_{\alpha}^* \) and the continuity assumptions made at the beginning of this section, we get
\[
\lim_{\rho(\alpha, \alpha_0) \to 0} \|Q_{\alpha} - Q_{\alpha_0}\| = 0.
\]

This result implies the continuity of \( g_{\alpha j} \) on \( A_1 \). However, the condition
\[
\lim_{\rho(\alpha, \alpha_0) \to 0} \|Q_{\alpha} - Q_{\alpha_0}\| = 0
\]
is unnecessarily strong. It can be replaced by (see [4])
\[
\lim_{\rho(\alpha, \alpha_0) \to 0} \|Q_{\alpha}^k (Q_{\alpha} - Q_{\alpha_0})\| = 0 \text{ for some } k \geq 1.
\]

Now we shall prove that \( \mu g_{\alpha} \) is continuous on \( A_\alpha \) for each nonnegative measure \( \mu \).

**Lemma 10.** Let \( \mu \) be a nonnegative measure on \( \Sigma_A \). Then the function \( \mu g_{\alpha} \) is continuous on \( A_\alpha \) for all \( n \in \mathbb{N} \).
Proof. We have $g_\alpha = P_\alpha g_\alpha$, hence by lemma 2, $g_\alpha = Q_\alpha g_\alpha$. Therefore

$$\int_{A} g_\alpha(u)\mu(du) = \int_{A} (Q_\alpha g_\alpha)(u)\mu(du) = \int_{A} (S_\alpha g_\alpha)(u)\mu(du) = \int_{A} g_\alpha(u)(\mu S_\alpha)(du).$$

The measure $\mu S_\alpha$ is a linear combination of the $\pi_{\alpha j}$, $j = 1, \ldots, n_\alpha$. So $(\mu S_\alpha)(A\setminus E_\alpha) = 0$ and

$$\int_{A} g_\alpha(u)(\mu S_\alpha)(du) = \int_{E_\alpha} g_\alpha(u)(\mu S_\alpha)(du).$$

Let $n \in \mathbb{N}$ and $\alpha_0, \alpha \in A_n$. Then

$$\int_{A} g_\alpha(u)\mu(du) - \int_{A} g_{\alpha_0}(u)\mu(du) = \int_{E_\alpha} g_\alpha(u)\mu(S_\alpha - S_{\alpha_0})(du) +$$

$$+ \int_{E_\alpha} (g_\alpha(u) - g_{\alpha_0}(u))(\mu S_{\alpha_0})(du) + \int_{E_\alpha} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du) -$$

$$- \int_{E_{\alpha_0}} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du).$$

The continuity of $S_\alpha$ as an operator valued function on $A_n$ and the uniform boundedness of $g_\alpha$ imply

$$\lim_{\rho(\alpha, \alpha_0) \to 0} \int_{E_\alpha} g_\alpha(u)(S_\alpha - S_{\alpha_0})(du) = 0.$$ 

Using $(\mu S_{\alpha_0})(V\setminus E_{\alpha_0}) = 0$ we get

$$\int_{E_\alpha} (g_\alpha(u) - g_{\alpha_0}(u))(\mu S_{\alpha_0})(du) = \int_{E_\alpha \cap E_{\alpha_0}} (g_\alpha(u) - g_{\alpha_0}(u))(\mu S_{\alpha_0})(du)$$

and

$$\int_{E_{\alpha_0}} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du) = \int_{E_{\alpha_0} \cap E_{\alpha_0}} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du).$$
Hence

\[ \int_{E_\alpha} (g_\alpha(u) - g_\alpha(u))(uS_{a_0})(du) = \sum_{j=1}^{n} \int_{E_{a_0} \cap E_\alpha} (g_\alpha(u) - g_\alpha(u))(uS_{a_0})(du) \]

and

\[ \int_{E_\alpha} g_\alpha(u)(uS_{a_0})(du) - \int_{E_{a_0}} g_\alpha(u)(uS_{a_0})(du) = - \int_{E_{a_0} \setminus E_\alpha} g_\alpha(u)(uS_{a_0})(du) . \]

We complete the proof by application of lemma 9 on (1) and (2), using that \( uS_{a_0} \) is a linear combination of the \( \pi_{a_0} \).

4.2. Existence of optimal strategies

In lemma 9 we proved the continuity of \( g_{a_1} \) on \( A_1 \). So, if \( A = A_1 \) and \( A \) is compact then an optimal strategy exists. If \( A_1 \neq A \) but if \( A_1 \) is dominating \( A \) (see [5]) we may restrict our attention to the set \( A_1 \), which is easier to analyse since \( g_\alpha \) is constant on \( V \) for \( \alpha \in A_1 \). To formulate conditions sufficient for \( A_1 \) to dominate \( A \), some new concepts are needed.

**Definition 11.** The SMD is called \( A \)-communicative if for all \( \alpha \in A \) and \( j = 1, \ldots, n_\alpha \) there is an \( \alpha_1 \in A_1 \) such that \( \pi_{\alpha_1}(E_{\alpha_j}) > 0 \).

Notice that \( A \)-communicativeness is equivalent to communicativeness as defined in [5] if \( A = V \).

**Definition 12.** Let \( \alpha \in A \) and \( i \in \{1, 2, \ldots, n_\alpha\} \). The set

\[ \overline{E}_{ai} := \{ u \in V \mid Q_{\alpha}(u, E_{ai}) = 1 \} \]

is called the extension of \( E_{ai} \).

Notice that \( E_{ai} \subseteq \overline{E}_{ai} \) and that by lemma 2, \( \overline{E}_{ai} \) is an invariant set of \( P_\alpha \).

**Lemma 13.** Let the SMD be complete (see [5]) and \( A \)-communicative. Then \( A_1 \) dominates \( A \).
Proof. Let \( a \in A \). Choose \( j_0 \) such that \( g_{a_{j_0}} = \min_{j = 1, 2, \ldots, n_A} g_{a_j} \). The A-communicativeness implies the existence of a strategy \( a_1 \in A_1 \) such that \( \pi_{a_1}^I(E_{a_{j_0}}) > 0 \).

Let \( C := E_{a_{j_0}} \) and \( a_2 := \alpha a_1 \) (apply strategy on \( C \) and strategy \( a_1 \) outside of \( C \)). Since \( \pi_{a_1}^I(C) > 0 \), under strategy \( a_2 = \alpha a_1 \) the system will never reach set \( C \) and the induced sub-Markov process \( Q_{a_2}^I \) of \( P_{a_2} \) on \( C \) is a Markov process. Further \( a_2 \in A_1 \) and by lemma 2, \( g_{a_2} := Q_{a_2}^I g_{a_2} \). The invariance of \( C = E_{a_{j_0}} \) under \( P_a \) implies \( g_{a_2} (u) = g_{a_{j_0}} \) for \( u \in C \). Hence

\[
g_{a_2} (u) = (Q_{a_2}^I g_{a_2}) (u) = \sum_{n=0}^{\infty} (P_{a_2}^n | C_n \cap C | C_n \cap C) g_{a_2} (u) = g_{a_{j_0}} \pi (u) \text{ for all } u \in V,
\]

which completes the proof. \( \square \)

The following theorem is an extension of theorem 4.9 of [5].

**Theorem 14.** Let the SMD be complete and A-communicative. If \( A \) is compact then an optimal strategy exists.

**Proof.** Let \( g := \inf_{a \in A} g_a \). The compactness of \( A \) implies the existence of a sequence \( \{a_k\} \) in \( A_1 \) converging to \( a_0 \in A \) such that \( \lim_{k \to \infty} g_{a_k} = g \). Without loss of generality we may assume that \( \pi_j := \lim_{k \to \infty} \pi_{a_k} (E_{a_{0,j}}) \) exists for all \( j = 1, \ldots, n_{a_0} \). We have

\[
g_{a_k} = \frac{\pi_{a_k} (T_{a_k} \cap C)}{\pi_k (T_{a_k} \cap C)} \text{ for all } k = 1, 2, 3, \ldots.
\]

As in the proof of theorem 4.9 of [5] we can show that
\[
\lim_{k \to \infty} \pi_{\alpha_k} (T_{\alpha_k} r_{\alpha_k}) = \sum_{j=1}^{n_{\alpha_k}} \pi_j r_j,
\]

where \( r_j := \pi_{\alpha_0 j} (T_{\alpha_0} r_{\alpha_0}) \) and

\[
\lim_{k \to \infty} \pi_{\alpha_k} (T_{\alpha_k} v) = \sum_{j=1}^{n_{\alpha_k}} \pi_j t_j,
\]

where \( t_j := \pi_{\alpha_0 j} (T_{\alpha_0} v) \). Hence

\[
g = \frac{\sum_{j=1}^{n_{\alpha_0}} \pi_j r_j}{\sum_{j=1}^{n_{\alpha_0}} \pi_j t_j}
\]

and therefore

\[
\min_{j=1, \ldots, n_{\alpha_0}} \left\{ \frac{r_j}{t_j} \right\} \leq g.
\]

But \( g_{\alpha_0 j} = \frac{r_j}{t_j} \) for \( j = 1, 2, \ldots, n_{\alpha_0} \), which implies that

\[
\min_{j=1, \ldots, n_{\alpha_0}} \{ g_{\alpha_0 j} \} \leq g.
\]

By lemma 13 there is an \( \alpha \in A_1 \) such that

\[
g_{\alpha} (u) \leq \min_{j=1, \ldots, n_{\alpha_0}} \{ g_{\alpha_0 j} \} \text{ for all } u \in V.
\]

The strategy \( \alpha \) is optimal.

4.3. Extensions and remarks

In subsection 4.2 we derived conditions for the existence of an optimal strategy. Optimality of some strategy \( \alpha_0 \) implies of course the \( \mu \)-optimality of this strategy for each nonnegative measure \( \mu \) on \( E \). Conversely, if \( \alpha_0 \) is \( \mu \)-optimal and the SMD is complete, then \( \alpha_0 \) is optimal \( \mu \)-almost everywhere, that
means \( g_{\alpha 0} (u) \leq g_{\alpha} (u) \), \( \mu\)-almost everywhere on \( V \) for all \( \alpha \in \Lambda \). See [4] for a proof.

To verify the conditions i - v of this section (section 4), it can be usefull to introduce the spaces \( B_\omega \) and \( M_\omega \).

Let \( \omega \) be a positive measurable function on \( V \) with \( \inf_{u \in V} \omega(u) > 0 \). The space \( B_\omega \) is the space of all complex valued measurable functions \( f \) on \( V \) such that \( \frac{f}{\omega} \in B \). With the norm \( \| f \| := \frac{f}{\omega} \) this space is a Banach space.

The space \( M_\omega \) is the space of all measures \( \mu \) such that the measure \( \mu_\omega \) defined by

\[
\mu_\omega (E) = \int_E \omega(u) \mu(du), \quad E \in \Sigma
\]

is an element of \( M \). With the norm \( \| \mu \| := \| \mu_\omega \| \) this space is a Banach space.

For an application of this idea to inventory problems, see [4].

The methods described in this section can also be applied to semi-Markovian decision problems. It is sufficient that the average costs can be written as the quotient of the expected recurrence costs and the expected recurrence time to a fixed set \( A \).

5. **Countable state space**

In this section some results of the preceding one are applied to the case where \( V \) is countable and \( \Sigma \) is the \( \sigma \)-field of all subsets of \( V \).

In the next lemma it will be shown that the conditions i), ii), iii), iv), and v), stated in section 4, are implied by some simpler ones.

**Lemma 15.** Let the following conditions be satisfied.

a) The functions \( r_\alpha \) are bounded on \( V \) for all \( \alpha \in \Lambda \) and the boundedness is uniform on \( \Lambda \).

b) There is a metric \( \rho \) on \( \Lambda \) such that \( P_\alpha (u, v) \) and \( r_\alpha (u) \) are continuous in \( \alpha \) for all \( u, v \in V \). (Instead of \( P_\alpha (u, \{ v \}) \) we write \( P_\alpha (u, v) \).

c) There is a finite subset \( A \) of \( V \) such that the sum

\[
\sum_{n=0}^{\infty} \left( P^n_{\alpha B, V} \right)(u)
\]
with $B := V \setminus A$, exists for all $u \in V$, $\alpha \in A$, and the convergence is uniform on $A$ for all $u \in A$.

Then the conditions i), ii), iii), iv), and v) are satisfied.

For the proof of this lemma we need the following result.

**Lemma 16.** Let $\rho$ be a metric on $A$ such that $P_\alpha(u,v)$ is continuous as function on $A$, for all $u,v \in V$. Let $\{f_\alpha\}, \alpha \in A$ be a set of complex valued functions, bounded on $V$ uniform on $A$ and let $f_\alpha(u)$ be continuous in $\alpha$ for all $u \in V$.

Then $(P_{Gf_\alpha})(u)$ is continuous in $\alpha$ for all $u \in V$, $G \in \Sigma$.

**Proof.** Choose $u \in V$, $G \in \Sigma$, $\alpha_0 \in A$. Let $\epsilon > 0$. There is a finite set $F_\epsilon$ such that $P_\alpha(u,F_\epsilon) > 1 - \epsilon$. The continuity of $P_\alpha(u,F_\epsilon)$ implies the existence of a $\delta > 0$ such that $P_\alpha(u,V \setminus F_\epsilon) < 2\epsilon$ for all $\alpha \in A$ with $\rho(\alpha,\alpha_0) < \delta$. We have

$$
(P_{Gf_\alpha})(u) = \int_{G \setminus F_\epsilon} P_\alpha(u,ds)f_\alpha(s) + \int_{F_\epsilon \cap G} P_\alpha(u,ds)f_\alpha(s)
$$

and

$$
(P_{Gf_\alpha})(u) - (P_{Gf_\alpha_0})(u) = \int_{G \setminus F_\epsilon} P_\alpha(u,ds)f_\alpha(s) - \int_{G \setminus F_\epsilon} P_{\alpha_0}(u,ds)f_{\alpha_0}(s) + \int_{F_\epsilon \cap G} (P_\alpha(u,ds) - P_{\alpha_0}(u,ds))f_\alpha(s) + \int_{F_\epsilon \cap G} P_{\alpha_0}(u,ds)(f_\alpha(s) - f_{\alpha_0}(s)).
$$

The rest of the proof is obvious.

Now we can give the proof of lemma 15.

**Proof of lemma 15.** The conditions i) and iii) are direct consequences of the conditions a) and c), condition ii) is implied by the finiteness of the set $A$ ($Q_\alpha$ is even compact). To prove iv) it is sufficient to prove the continuity of $Q_\alpha(u,E)$ in $\alpha$ for all $u \in A$, $E \in \Sigma_A$. This is easily done by using the expression

$$
Q_\alpha(u,E) = \sum_{n=0}^{\infty} (P_{nB}^\alpha P_{\alpha A} 1_E)(u).
$$
Namely, condition c) implies that for all $\varepsilon > 0$ there is an integer $N$ such that

$$\sum_{n=N}^{\infty} (P^n_{\alpha B_{\alpha A} E})(u) < \varepsilon$$

for all $u \in A$, $\alpha \in \mathcal{A}$, $E \in \Sigma$. The continuity of

$$\sum_{n=0}^{N-1} (P^n_{\alpha B_{\alpha A} E})(u)$$

in $\alpha$ follows from Lemma 16. The rest of the proof of iv) is straightforward. That condition v) is also satisfied can be shown similarly, using the continuity of $r_\alpha(u)$ in $\alpha$.

The following theorem is a direct consequence of Lemma 15 and Theorem 14.

Theorem 17. Let the conditions a), b), and c) of Lemma 15 be satisfied and let the SMD be complete and $A$-communicative. If $A$ is compact then an optimal strategy exists.

For the case of a countable $V$ we shall relate our results to those of some others. Ross [3] and Hordijk [1] investigate the existence of a stationary policy which is average optimal in the class of all policies (Ross) or in the class of all Markov policies (Hordijk) of a Markov decision process. If only the stationary policies are allowed the Markov decision process corresponds to a complete SMD $\{(P_\alpha, r_\alpha)\}$, $\alpha \in A$, where $A$ is the set of all stationary policies. The existence of an optimal strategy of this SMD implies the existence of a stationary policy which is optimal only in the class of all stationary policies. It is important to be conscious of this fact in relating our results to those of Hordijk and Ross.

The conditions of Hordijk, in our terminology, are as follows:

1) the functions $r_\alpha(\cdot)$ are bounded on $V$, uniform on $A$;
2) the simultaneous Doeblin condition is satisfied: there is a finite set $A$, a pos. number $c$, and an integer $n$ such that $P^n_\alpha(u, A) \geq c$ for all $u \in V$, $\alpha \in A$;
3) there is a metric $\rho$ on $A$ such that $A$ is compact and
4) for all \( u, v \in V \) the functions \( r_\alpha(u) \) and \( P_\alpha(u, v) \) are continuous in \( \alpha \);
5) the SMD is comminicate. (The simultaneous Doeblin condition implies the quasi-compactness of \( P_\alpha \) for all \( \alpha \in A \), so we may speak indeed about communicativeness.)

The most striking difference with the conditions of theorem 17 is the simultaneous Doeblin condition. Instead of this condition we require condition c) of lemma 15: there is a finite set \( A \subseteq V \) such that the sum

\[
\sum_{n=0}^{\infty} (P^n_{\alpha B \mid V})(u),
\]

where \( B := V \setminus A \), exists for all \( u \in V \) and \( \alpha \in A \), and the convergence is uniform on \( A \) for all \( u \in A \).

The simultaneous Doeblin condition implies the convergence of

\[
\sum_{n=0}^{\infty} (P^n_{\alpha B \mid V})(u), \text{ uniform on } V \times A.
\]

Ross [3] gives the following conditions:
- for all \( u \in V \) the set \( A(u) \) of all possible actions in \( u \) is finite;
- the functions \( r_\alpha(\cdot) \) are bounded on \( V \), uniform on \( A \);
- there exists a state \( v \in V \), an integer \( N > 0 \), and a sequence of discount factors \( \{ \beta_n \} \), \( 0 < \beta_n < 1 \), such that \( \lim_{n \to \infty} \beta_n = 1 \) and \( M_{uv}(R_{\beta_n}) < N \) for all \( u \in V \), \( n \in N \), where \( M_{uv}(R_{\beta_n}) \) is the mean time to go from state \( u \) to state \( v \) when using the \( \beta_n \)-discounted optimal policy \( R_{\beta_n} \).

The finiteness of \( A(u) \) makes the compactness and continuity conditions superfluous.

The last condition of Ross states a very strong recurrency, (recurrency to a point \( v \in V \)) for a subset \( \{ R_n \} \), \( n = 1, 2, \ldots \) of the set of all stationary deterministic policies. This condition guarantees the quasi-compactness of the Markov process under policy \( R_n \) and also \( R_n \in A \) (only one invariant probability). In condition c) of lemma 15 a weaker recurrency is stated (recurrency to a set \( A \)), but for all strategies \( \alpha \in A \). The A-communicativeness assumed in theorem 17 implies that \( A \) dominates \( A \).
In a set of conditions different from the just mentioned one Hordijk [1] section 5, also requires recurrency to a point. This set of conditions is more directly related to the conditions i - v of section 4 of this paper, with $V$ countable and $A$ consisting of one point. The conditions guarantee the continuity of the recurrence costs to $A$ and the recurrence time to $A$ as function of $\alpha$. The boundedness of $r_\alpha$ and the quasi-compactness of $P_\alpha$ is not required.

References


