Technological University Eindhoven

Department of Mathematics

Statistics and Operations Research Group

Memorandum COSOR 75-14

Stationary Markovian decision problems I

by

J. Wijngaard

Eindhoven, September 1975

The Netherlands
Stationary Markovian decision problems I

by

J. Wijngaard

1. Introduction

In the theory of discrete time Markovian decision processes there has been given a lot of attention to the existence of optimal policies. There are at least two ways to prove the existence of such an optimal policy. In the first place one can use the functional equation of dynamic programming. The existence of a solution of this equation guarantees, under some extra conditions, the existence of an optimal policy. Another way consists of the investigation of the compactness of the set of policies and the continuity of the costs (total, discounted or average) as a function of the policy. Here the fact is used that the minimum of a continuous function on a compact space exists. The first method is most frequently applied. For the total costs case the second method is used by Schäl [5,6]. In this paper the second method is used for the average costs case. We restrict ourselves to stationary policies.

This restriction seems to be not so severe. Most conditions which guarantee the existence of an optimal policy guarantee also the stationarity of this policy. The restriction to stationary policies implies that we may represent the problem by a set of pairs \((P_\alpha, r_\alpha)\), \(\alpha \in A\), where \(A\) is the set of all stationary policies, \(P_\alpha\) the Markov process under policy \(\alpha\) and \(r_\alpha(u)\) the one period costs starting in \(u\).

Let \(g_\alpha(u)\) be the average costs under policy \(\alpha\), starting in \(u\) (if existing). The problem now is to construct a topology on \(A\) such that \(A\) is compact and \(g_\alpha(u)\) is continuous on \(A\). It is not essential that \(A\) is the set of stationary policies. We may consider \(A\) to represent a set of indices only. In an example we shall show some difficulties arising in the proof of the continuity of \(g_\alpha\) in \(\alpha\).

Example. The state space consists of three elements, \(S = \{1, 2, 3\}\). Once in state 2 or 3 one must stay there, the costs being 0 and 10 each period. In state 1 one of the actions \(d \in [0, 1]\) can be chosen. The probability of a transition to the states 1, 2, 3 is \(1 - d - d^2, d, d^2\). The costs of each of these actions \(d\) in state 1 are 1. The average costs, starting in state 2 or 3, are
0 and 10, independently of the policy used in state 1. If one uses policy
d = 0, the average costs starting in state 1, \( g_0(1) \), are equal to 1 since
the system will never leave state 1. If one uses policy \( d > 0 \), the system
will certainly leave state 1 and will never return. In this case the average
costs starting in state 1, \( g_d(1) \), are equal to
\[
d \cdot 0 + d^2 \cdot 10 + (1-d-d^2)(d \cdot 0 + d^2 \cdot 10) + \ldots = 10 \frac{d}{1 + d}.
\]
Hence \( \inf \{ g_d(1) \} = 0 \) but this infimum is not attained since \( g_0(1) = 1 \).

There is no optimal policy. The average costs as function of \( d \) have a dis­
continuity in \( d = 0 \). This discontinuity corresponds to a discontinuity in
the number of ergodic sets. For \( d > 0 \) there are two ergodic sets, the sets
\{2\} and \{3\}, but for \( d = 0 \) the set \{1\} is also an ergodic set. The eigen­
values of the transition matrix corresponding to the policy \( d \) are 1 and
\( 1-d-d^2 \). For \( d = 0 \) the eigenvalues coincide.

These continuity problems can be investigated with the aid of the pertur­
bation theory of linear operators. Each Markov process in a finite state
space corresponds to a transition matrix. In a more general state space \( S \)
each Markov process corresponds to a linear operator in the space of all
complex valued bounded measurable functions on \( S \). As in the finite case the
point 1 is one of the eigenvalues of the operator. Now let \( \{(P_\alpha, r_\alpha)\} \), \( \alpha \in A \)
be a set of Markov processes with costs and assume that \( A \) is a metric space.
In the example we had \( A = [0, \frac{1}{2}] \). Although the transition matrix is continu­
ous in \( \alpha \) and the one-period costs even independent of \( \alpha \), the average costs
starting in state 1 have a discontinuity corresponding to a discontinuity in
the dimension of the eigenspace of eigenvalue 1. Apart from this discon­
tinuity the average costs are continuous. Using the perturbation theory of
linear operators it can be shown that this restricted continuity of \( g_\alpha(u) \)
holds if
- the cost functions \( r_\alpha \) are bounded and continuous in \( \alpha \),
- the Markov processes \( P_\alpha \) are quasi-compact and continuous in \( \alpha \).

Quasi-compactness of a Markov process is defined in terms of the correspond­
ing linear operator. Essential is that this operator has only a finite num­
ber of eigenvalues on the unit circle, each of these a root of unity, and
each with a finite dimensional eigenspace.
In section 3 we shall consider quasi-compact Markov processes and give some well known results about the spectrum of the corresponding linear operator. As a preliminary we give in section 2 some results from spectral and perturbation theory of linear operators in Banach spaces. In section 4 stationary Markovian decision problems are considered. Conditions for the existence of an optimal strategy are derived. The most important conditions are the boundedness of $r_\alpha$ on the state space and the quasi-compactness of $P_\alpha$. These conditions are weakened in [7].

2. Spectral and perturbation theory of quasi-compact operators

2.1. Spectral decomposition

Let $X$ be a complex Banach space and $T$ a bounded linear operator of $X$ in $X$. The resolvent set $\rho(T)$ of $T$ is the set of all complex numbers $\lambda$ such that $\lambda I - T$ is $1$-1 and onto ($I$ is the identity). If $\lambda \in \rho(T)$ then $R(\lambda; T) := (\lambda I - T)^{-1}$ exists and is bounded. The complement of $\rho(T)$ in $\mathbb{C}$ is called the spectrum of $T$ and will be denoted by $\sigma(T)$.

The spectral radius $r(T)$ of $T$ is defined by $r(T) := \sup_{\lambda \in \sigma(T)} |\lambda|$. It is well known that $r(T) = \lim_{n \to \infty} \|T^n\|^{1/n}$. A point $\lambda \in \sigma(T)$ such that $\lambda I - T$ is not $1$-1 is called an eigenvalue of $T$ and $N(\lambda I - T)$ is the corresponding eigenspace.

In the sequel the spectral decomposition theorem is important. As a preliminary to this decomposition theorem we recall the concept of a function of an operator as given in Dunford-Schwartz [2], VII.3.9.

Let $f$ be a complex valued function on $\mathbb{C}$ which is analytic on some neighbourhood of $\sigma(T)$. Let $U$ be an open set whose boundary $B$ consists of a finite number of rectifiable Jordan curves, oriented in the positive sense. Suppose that $U \subset \sigma(T)$ and that $U \cup B$ is contained in the domain of analyticity of $f$. The operator $f(T)$ is defined by

$$f(T) := \frac{1}{2\pi i} \int_0^B f(\lambda)R(\lambda; T)d\lambda.$$  

The operator $f(T)$ depends only on the values of $f$ on $\sigma(T)$. 

A spectral set is a subset of $\sigma(T)$ which is both open and closed in $\sigma(T)$. If $\alpha$ is a spectral set, then $\widetilde{\alpha} := \sigma(T) \setminus \alpha$ is also a spectral set. For each spectral set $\alpha$ it is possible to choose a function $f$ satisfying the conditions of the above definition with $f(\lambda) = 1$ on $\alpha$ and $f(\lambda) = 0$ on $\alpha$. For such a function $f$ the operator $f(T)$ is denoted by $E_{\alpha}(T)$, or shortly by $E_{\alpha}$. The range of $E_{\alpha}$ is denoted by $X_{\alpha}$.

The following properties are immediate consequences of [1], VII.3.10.

i) $E_{\alpha}^2 = E_{\alpha}$ ($E_{\alpha}$ is a projection).

ii) $E_{\alpha}T = TE_{\alpha}$, hence $Tx \in X_{\alpha}$ if $x \in X_{\alpha}$, $X_{\alpha}$ is invariant under $T$.

The restriction of $T$ to $X_{\alpha}$ is denoted by $T_{\alpha}$.

iii) $E_{\alpha} + E_{\widetilde{\alpha}} = I$ and $E_{\alpha}E_{\widetilde{\alpha}} = 0$. This implies $X = X_{\alpha} \oplus X_{\widetilde{\alpha}}$.

If $\lambda$ is an isolated point of $\sigma(T)$, then the set $\{\lambda\}$ is of course a spectral set. In this case we shall write $E_{\lambda}$ and $E_{\widetilde{\lambda}}$, instead of $E_{\{\lambda\}}$ and $E_{\{\widetilde{\lambda}\}}$.

A pole of $T$ of order $n$ is an isolated point of $\sigma(T)$ where the function $R(\cdot;T)$ has a pole of order $n$.

For the proof of the following two lemma's we refer to [2], VII.3.

**Lemma 2.1.** (Spectral decomposition theorem). Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be disjoint spectral sets such that $\sigma(T) = \bigcup_{i=1}^{n} \alpha_i$. Then the following properties hold:

i) $(X,T) = (X_{\alpha_1},T_{\alpha_1}) \oplus (X_{\alpha_2},T_{\alpha_2}) \oplus \cdots \oplus (X_{\alpha_n},T_{\alpha_n})$.

ii) $\sigma(T_{\alpha_i}) = \alpha_i$.

iii) $\lambda$ is a pole of $T_{\alpha_i}$ of order $n$ if and only if $\lambda \in \alpha_i$ and $\lambda$ is a pole of $T$ of order $n$.

**Lemma 2.2.** An isolated point $\lambda$ of $\sigma(T)$ is a pole of order $n$ if and only if $(\lambda I - T)^n E_{\lambda} = 0$ and $(\lambda I - T)^{n-1} E_{\lambda} \neq 0$. Furthermore, if $\lambda$ is a pole of $T$ of order $n$, then $\lambda$ is an eigenvalue of $T$ with index equal to $n$ (that means $n$ is the smallest integer such that $N((\lambda I - T)^n) = N((\lambda I - T)^{n+1})$) and

$$X_{\lambda} = N((\lambda I - T)^n), \quad X_{\widetilde{\lambda}} = R((\lambda I - T)^n).$$
In the next lemma we consider the existence of \( \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) \), where \( \lambda_i \) is a pole of \( T \) with \( |\lambda_i| = r(T) = \|T\| \).

**Lemma 2.3.** Let \( \|T\| = r(T) \). Assume that the spectrum of \( T \) consists of a finite number of poles \( \lambda_1, \ldots, \lambda_q \), on the circle with radius \( r(T) \) and of a set \( \alpha \) within this circle. Then

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) = E_i, \quad i = 1, \ldots, q.
\]

**Proof.** By lemma 2.1

\[
I = \sum_{j=1}^{q} E_j + E_{\alpha}.
\]

Hence

\[
\frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) = \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) \sum_{j=1}^{q} E_j + E_{\alpha} =
\]

\[
= \sum_{j=1}^{q} \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) E_j + \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) E_{\alpha}.
\]

By lemma 2.2, \( \lambda_j \) is an eigenvalue. Using \( |\lambda_j| = \|T\| \) it is easy to see that the index of \( \lambda_j \) is 1.

Therefore \( TE_j = \lambda_j E_j \), and

\[
\begin{align*}
(1) \quad \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) &= \sum_{j=1}^{q} \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) E_j + \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) E_{\alpha},
\end{align*}
\]

It is easy to see that

\[
(2) \quad \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) = \begin{cases} 0 & \text{if } j \neq i, \\ 1 & \text{if } j = i. \end{cases}
\]

By lemma 1.5, \( \sigma(T_{\alpha}) = \alpha \) and \( r(T_{\alpha}) < r(T) = |\lambda_i| \). Therefore

\[
R(\lambda_i; T_{\alpha}) = \sum_{n=0}^{\infty} \frac{T^n_{\alpha}}{\lambda_i^{n+1}} \quad \text{and} \quad \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left( \frac{T}{\lambda_i^k} \right) E_{\alpha} = 0.
\]

Together with (1) and (2) this implies
2.2. Quasi-compact operators

Quasi-compactness of an operator is closely related to compactness.

Definition 2.4. Let $T$ be a bounded linear operator of the complex Banach space $X$ in $X$. $T$ is said to be quasi-compact if there exists a compact operator $K$ of $X$ in $X$ and a positive integer $n$ such that $\|T^n - K\| < r(T)^n$.

There is an interesting relationship between quasi-compactness of $T$ and the structure of $\sigma(T)$. The proof of the next lemma can be found in [8].

Lemma 2.5. A bounded linear operator $T$ of $X$ in $X$ is quasi-compact if and only if $\sigma(T) \cap \{\lambda \mid |\lambda| = r(T)\}$ consists of a finite number of poles $\lambda_1, \ldots, \lambda_q$ such that the spaces $X_{\lambda_i}$ are finite dimensional.

2.3. Perturbation theory

In Markovian decision processes we are interested in sets of Markov processes, in the way the average costs changes as a function on the set of Markov processes. To investigate this problem we can use perturbation theory of linear operators.

Let $X$ be a complex Banach space, $A$ a metric space with metric $\rho$ and $T(a)$ a continuous operator valued function on $A$.

The set $S(A, \varepsilon)$ for $A \subset A$ or $A \subset C$ is defined as the set of all $e$ in $A$ or in $C$ such that the distance of $e$ to $A$ is less than $\varepsilon$. If $A$ consists of a single element $a$ we shall write $S(a, \varepsilon)$ instead of $S(\{a\}, \varepsilon)$.

The following two lemma's are consequences of [2], VII.6.3 and 6.7, and the fact that $T(.)$ is continuous on $A$.

Lemma 2.6. For each $\varepsilon > 0$ there is a $\delta > 0$ such that $a \in S(a_0, \delta)$ implies $\sigma(T(a)) \subset S(\sigma(T(a_0)), \varepsilon)$ and

$$\|R(\lambda; T(a)) - R(\lambda; T(a_0))\| < \varepsilon \quad \text{if} \quad \lambda \notin S(\sigma(T(a_0)), \varepsilon).$$
Lemma 2.7. Let $T(a)$ be a projection for all $a \in A$.
If $R(T(a_0))$ is $N$-dimensional, there is a $\delta > 0$ such that $R(T(a))$ is $N$-dimen-
sional for all $a \in S(a_0, \delta)$.

In the next lemma we give some results for the case where $T(a)$ is quasi-com-
 pact and has the point 1 as an eigenvalue.

Lemma 2.8. Let for all $a \in A$ the operator $T(a)$ be quasi-compact, $\|T(a)\| =
= r(T(a)) = 1$, and 1 is an eigenvalue of $T(a)$.

a) Let $a_0 \in A$. There is a $\delta > 0$ such that for all $a \in S(a_0, \delta)$

dimension $N(I - T(a)) \leq$ dimension $N(I - T(a_0))$.

b) Let $\{a_n\}_{n=1}^{\infty}$ be a sequence in $A$ converging to $a_0 \in A$, such that

$$\dim N(I - T(a_n)) = \dim N(I - T(a_0)) \text{ for all } n \in \mathbb{N}.$$ Let $S_\alpha := \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} T(a)^k$.
Then $\lim_{n \to \infty} S_{a_n} = S_{a_0}$.

c) Let $\beta$ be such that $0 < \beta < 1$ and for all $a \in A$ the spectrum of $T(a)$ does
not contain points of modulus between $\beta$ and 1. Then for all $a_0 \in A$ there
is a $\delta > 0$ such that for all $a \in S(a_0, \delta)$

$$\dim N(I - T(a)) = \dim N(I - T(a_0)).$$

Proof. Let $a_0 \in A$. The quasi-compactness of $T(a_0)$ implies the isolatedness
of the point 1 in $\sigma(T(a_0))$, there is an $\epsilon > 0$ such that $S(1, \epsilon) \cap \sigma(T(a_0)) = \{1\}.$
By lemma 2.6 there is a $\delta > 0$ such that for all $a \in S(a_0, \delta)$ the spectrum
$\sigma(T(a))$ contains no points $\lambda$ with $\frac{\epsilon}{3} < |1 - \lambda| < \frac{2\epsilon}{3}$. The quasi-compactness
of $T(a_0)$ implies the existence of a compact operator $K$ and an integer $n$ such
that $p := \|T(a_0)^n - K\| < 1$. Because of the continuity of $T(a)$ there is a
$\delta_1 > 0$ such that

$$\|T(a)^n - K\| < \frac{1 + p}{2} < 1 \text{ for all } a \in S(a_0, \delta_1).$$

Let $a \in S(a_0, \delta_1)$. By [2], VIII.8.2 each point $\lambda \in \sigma(T(a))$ with $|\lambda|^n > \frac{1 + p}{2}$
is an isolated point of $\sigma(T(a))$ and $X_\lambda(T(a))$ is finite dimensional. Hence,
$\lambda$ is a pole of $T_\lambda(a)$ and therefore, by lemma 2.1, a pole of $T(a)$.

Now we may assume without loss of generality that for $a \in S(a_0, \delta),$
$S(1, \frac{\epsilon}{3}) \cap \sigma(T(a))$ contains only poles of $T(a)$. 

Let $f$ be a function which is equal to 1 on $S(1, \frac{e}{3})$ and equal to 0 on $\mathbb{C}\setminus S(1, \frac{2e}{3})$.

Let $\sigma := S(1, \frac{e}{3}) \cap \sigma(T(a))$ and $\sigma_{\alpha} := \sigma \setminus \{1\}$.

Then for all $\alpha \in S(\sigma_0, \delta)$

$$f(T(a)) = E_{\sigma_{\alpha}}(T(a))$$

and

$$X_{\sigma_{\alpha}}(T(a)) = X_1(T(a)) \oplus X_{\sigma_{\alpha}}(T(a)).$$

By [2], VII.6.5 and lemma 2.7 there is a $\delta'$ with $0 < \delta' < \delta$ such that for all $\alpha \in S(\alpha_0, \delta')$

$$\dim X_{\sigma_{\alpha}}(T(a)) = \dim X_{\sigma_{\alpha}}(T(\alpha_0)) = \dim X_1(T(\alpha_0)).$$

Since $\|T(a)\| = 1$, the order of the pole 1 of $T(a)$ is 1. Hence, by lemma 2.2,

$$X_1(T(a)) = N(I - T(a)).$$

Using (1) and (2) we get for all $\alpha \in S(\alpha_0, \delta')$

$$\dim N(I - T(\alpha_0)) = \dim X_1(T(\alpha_0)) = \dim X_{\sigma_{\alpha}}(T(\alpha)) =$$

$$= \dim X_1(T(a)) + \dim X_{\sigma_{\alpha}}(T(\alpha)) =$$

$$= \dim N(I - T(\alpha)) + \dim X_{\sigma_{\alpha}}(T(\alpha)) \geq \dim N(I - T(a)).$$

This completes the proof of a).

If $\dim N(I - T(a)) = \dim N(I - T(\alpha_0))$ for some $\alpha \in S(\alpha_0, \delta')$, then $\sigma_{\alpha} = \emptyset$.

It follows that

$$\sigma_{\alpha} = \{1\} \quad \text{and} \quad f(T(a)) = E_{\sigma_{\alpha}}(T(a)) = E_1(T(a)) = S_{\alpha}.$$

The proof of b) is easily given by application of [2], VII.6.5.

The proof of c) is straightforward by choosing $\epsilon$ such that $1 - \epsilon > \beta$. □
3. Markov processes

3.1. Preliminaries

In the rest of this paper \((V, \Sigma)\) is assumed to be a measurable space. First the linear spaces \(B(V, \Sigma)\) and \(M(V, \Sigma)\) will be defined.

i) \(B(V, \Sigma)\), or shortly \(B\), is the space of all complex valued, bounded, measurable functions on \(V\). Let \(\|f\| := \sup_{u \in V} |f(u)|\) for all \(f \in B\). Then \(\|\cdot\|\) is a norm on \(B\) and with this norm \(B\) is a Banach space.

ii) \(M(V, \Sigma)\), or shortly \(M\), is the space of all complex valued measures on \(\Sigma\). Let, for all \(\mu \in M\) and \(E \in \Sigma\), \(v_\mu(E)\) be the total variation of \(\mu\) on \(E\) and define \(\|\mu\| := v_\mu(V)\) for all \(\mu \in M\). Then \(\|\cdot\|\) is a norm on \(M\) and with this norm \(M\) is a Banach space.

It is easy to see that \(M\) is isometrically isomorphic with a closed subspace of \(B^*\), the adjoint space of \(B\), and that \(B\) is isometrically isomorphic with a closed subspace of \(M^*\). The isomorphism \(\varphi\) from \(M\) in \(B^*\) is given by

\[
(\varphi \mu)(f) := \mu f := \int_V f(u)\mu(du).
\]

A sub-transition probability is a real valued function \(P\) on \(V \times \Sigma\) such that

i) for all \(u \in V\), \(P(u, \cdot)\) is a nonnegative measure on \(\Sigma\) with \(P(u, V) \leq 1\),

ii) for all \(E \in \Sigma\), \(P(\cdot, E) \in B(V, \Sigma)\).

A sub-transition probability \(P\) is called a transition probability if \(P(u, V) = 1\) for all \(u \in V\).

The (sub-)Markov process with (sub-)transition probability \(P\) is also denoted by \(P\).

A sub-Markov process \(P\) induces operators \(P_M\) and \(P_B\) in \(M\) and \(B\) by

\[
(\mu P_M)(\cdot) := \int_V P(u, \cdot)\mu(du) \quad \text{for all } \mu \in M
\]

\[
(P_B f)(\cdot) := \int_V f(v)P(\cdot, dv) \quad \text{for all } f \in B.
\]

The function \(\mu P_M\) on \(\Sigma\) is an element of \(M\) and the function \(P_B f\) on \(V\) is an element of \(B\). The operators \(P_B\) and \(P_M\) are linear.
Further it is easy to verify that
\[(\mu P^*_M)f = \mu(P_B f) \quad \text{for all } \mu \in M, \, f \in \mathcal{B} \, .\]

Let $P^*_B$ be the adjoint of $P_B$ and $P^*_M$ the adjoint of $P^*_M$. Relation (1) implies that the restriction of $P^*_B$ to the closed subspace of $B^*$, which is isometrically isomorphic with $M$, corresponds to the operator $P^*_M$ and it implies also that the restriction of $P^*_M$ to the closed subspace of $M^*$ which is isometrically isomorphic with $B$, corresponds to the operator $P_B$.

Therefore
\[
\begin{align*}
\sigma(P_B) &= \sigma(P^*_B) \supset \sigma(P^*_M) = \sigma(P^*_M) \supset \sigma(P_B), &\text{hence } &\sigma(P_B) = \sigma(P^*_M) \\
\Vert P_B \Vert &= \Vert P^*_B \Vert \supset \Vert P^*_M \Vert = \Vert P^*_M \Vert \supset \Vert P_B \Vert, &\text{hence } &\Vert P_B \Vert = \Vert P^*_M \Vert .
\end{align*}
\]

From $P(u, V) \leq 1$ for all $u \in V$ it follows that $\Vert P_B \Vert \leq 1$. If $P(u, V) = 1$ for all $u \in V$ then $P_B^*|_{V^*} = 1_{V^*}, \, \lambda \in \sigma(P_B)$ and $\Vert P_B \Vert = 1$. As the relations (2) and (3) one can prove that $\dim N(I - P_B) = \dim N(I - P^*_M)$.

In the rest of this paper both the operator $P_B$ in $B$ and the operator $P^*_M$ in $M$ are denoted by $P$; $P$ to the left of a function is the operator in $B$, $P$ to the right of a measure is the operator in $M$.

A set $E \in \Sigma$ is called invariant under $P$ if $P(u, E) = 1$ for all $u \in E$. If $\mu P = \mu$, $\mu$ is called an invariant measure, if $Pf = f$, $f$ is an invariant function. We saw already that the function $1_V$ is invariant for all Markov processes.

A special case of a sub-Markov process is the process $I_A$ for $A \in \Sigma$, which is determined by the sub-transition probability

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

Application of the corresponding operator in $B$ is multiplying by the characteristic function of $A$, $I_A f = 1_A f$, $f \in \mathcal{B}$, and the corresponding operator in $M$ is given by $(\mu I_A)(\cdot) = (A \cap \cdot)$, $\mu \in M$.

Let $P$ and $Q$ be sub-Markov processes on $(V, \Sigma)$. The sub-Markov process $PQ$ is defined by $(PQ)(u, E) := (P(Q|_{E^*}))(u), \, u \in V, \, E \in \Sigma$. For the process $PQ$ the operator in $B$ is given by $(PQ)(f) = P(Qf)$ and the operator in $M$ by $\mu(PQ) = (\mu P)Q$. If $R$ is another sub-Markov process then $(PQ)R = PQR$. 
3.2. Quasi-compact Markov processes

Using the relationship between the spaces $\mathcal{B}$ and $\mathcal{M}$ one can prove that the operator $P$ in $\mathcal{B}$ is quasi-compact if and only if the operator $P$ in $\mathcal{M}$ is quasi-compact (see [8]).

Now a Markov process $P$ on $(\mathcal{V}, \Sigma)$ is called quasi-compact if the operator $P$ (in $\mathcal{B}$ or in $\mathcal{M}$) is quasi-compact.

In Neveu [4] it is shown that quasi-compactness of a Markov process is equivalent with the Doeblin condition.

The following lemma is known as the strong ergodic theorem. A proof based on the spectral decomposition theorem can be found in Yosida and Kakutani [9], see also [8].

Lemma 3.1. Let $P$ be a quasi-compact Markov process and suppose $\dim N(I - P) = n$. Then there exists a unique set of probabilities $\{\pi_1, \ldots, \pi_n\}$ such that

i) $\pi_1, \ldots, \pi_n$ are invariant under $P$

ii) there are pairwise disjoint invariant sets $E_1, \ldots, E_n$ such that $\pi_i(E_i) = 1$ for $i = 1, \ldots, n$.

If $\mu$ is a probability with $\mu(E_i) = 1$ then

$$\pi_i = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu P^k.$$ 

There is a real number $\beta$, $0 < \beta < 1$, and an integer $N$ such that $\|P^k f\| \leq \beta^k \|f\|$ for all $k > N$ and all functions $f \in \mathcal{B}$ which are $\pi_1$-almost everywhere equal to zero, $i = 1, \ldots, n$.

For each $i = 1, \ldots, n$ there is an integer $d_i$ and pairwise disjoint sets $E_{i1}, \ldots, E_{id_i}$ such that

$$P(u, E_{ik}) = 1 \quad \text{for } u \in E_{ik-1}, \quad k = 2, \ldots, d_i$$

$$P(u, E_{i1}) = 1 \quad \text{for } u \in E_{id_i}.$$ 

If $d$ is a common multiple of all $d_i$, then $\lambda^d = 1$ for all eigenvalues of $P$ on the unit circle.
Now let the conditions of lemma 3.1 be satisfied and let $\pi_1, \ldots, \pi_n$ and $E_1, \ldots, E_n$ be as in this lemma. By lemma 2.3, $S := \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k$ exists. It is easy to see that $S$ is a Markov process satisfying $PS = SP = S$. Hence $S(u, \cdot)$ is an invariant probability of $P$ for all $u \in V$. Define the sets $F_1, \ldots, F_n$ by $F_i := \{u \in V \mid S(u, \cdot) = \pi_i(\cdot)\}$. The sets $F_i$ are pairwise disjoint and since $F_i \ni E_i$ we have $\pi_i(F_i) = 1$. For $u \in F_i$ we have

$$1 = S(u, E_i) = (PS)(u, E_i) = \int_V P(u, ds) S(s, E_i).$$

It follows that $S(\cdot, E_i) = 1$, $P(u, \cdot)$-almost everywhere, hence $P(u, F_i) = 1$.

The sets $F_1, \ldots, F_n$ are called the maximal invariant sets of $P$.

For all functions $f \in B$ the function $Sf = \int_V S(u, dv)f(v)$ is constant on $F_i$, $i = 1, \ldots, n$.

4. Stationary Markovian decision problems

4.1. Preliminaries

A Markov process with costs on $(V, \Sigma)$ is a pair $(P, r)$ where $P$ is a Markov process on $(V, \Sigma)$ and $r$ is a nonnegative measurable function on $V$, the cost-function. The

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} (P^k r)(u),$$

if existing, is called the average costs in $u$.

A stationary Markovian decision problem (SMD) on $(V, \Sigma)$ is a set of Markov processes with costs, $\{(P_\alpha, r_\alpha)\}$, $\alpha \in A$. The elements of $A$ are called strategies. The average costs of $(P_\alpha, r_\alpha)$ in $u$ are denoted by $g_\alpha(u)$.

The strategy $\alpha_0 \in A$ is called optimal if $g_{\alpha_0}(u) \leq g_\alpha(u)$ for all $\alpha \in A$, $u \in V$.

For $\mu$ a nonnegative measure on $\Sigma$, the strategy $\alpha_0$ is called $\mu$-optimal if

$$\mu g_{\alpha_0} := \int_V g_{\alpha_0}(u) \mu(du) \leq \mu g_\alpha \quad \text{for all } \alpha \in A.$$ 

An important property of an SMD is its completeness.
Definition 4.1. Let \((P_1, r_1)\) and \((P_2, r_2)\) be Markov processes with costs on \((V, \Sigma)\). For each \(F \in \Sigma\) the Markov process with costs \((P_1 P_2, r_2 r_2)\) is defined by:
\[
(P_1 P_2)(u, E) = P_1(u, E), \quad (r_2 r_2)(u) = r_1(u) \quad \text{for } u \in F, \quad E \in \Sigma,
\]
\[
(P_1 P_2)(u, E) = P_2(u, E), \quad (r_2 r_2)(u) = r_2(u) \quad \text{for } u \in V \setminus F, \quad E \in \Sigma.
\]

Definition 4.2. An SMD \(\{(P_\alpha, r_\alpha)\}, \alpha \in A\) is complete if for all \(\alpha_1, \alpha_2 \in A\) and for all \(F \in \Sigma\) there is an \(\alpha \in A\) such that
\[
(P_\alpha, r_\alpha) = (P_{\alpha_1} P_{\alpha_2}, r_{\alpha_1} r_{\alpha_2})
\]

The strategy \(\alpha\) is denoted by \(\alpha_1 \alpha_2\).

Intuitively speaking an SMD is complete if for all \(F \in \Sigma\) and for all \(\alpha_1, \alpha_2 \in A\) it is allowed to apply strategy \(\alpha_1\) on \(F\) and strategy \(\alpha_2\) on \(V \setminus F\).

We will derive conditions for the existence of an optimal strategy, using the fact that the minimum of a continuous function on a compact space exists. This is worked out in the next subsections.

4.2. Continuity

An SMD \(\{(P_\alpha, r_\alpha)\}, \alpha \in A\) on \((V, \Sigma)\) is considered such that \(P_\alpha\) is quasi-compact for all \(\alpha \in A\) and \(r_\alpha\) is bounded on \(V\), uniform in \(\alpha\). Let for all \(\alpha \in A\), \(n_\alpha\) be the dimension of \(N(I - P_\alpha)\), \(E_{\alpha j}\) for \(j = 1, \ldots, n_\alpha\) the maximal invariant sets of \(P_\alpha\) and \(\pi_{\alpha j}\), \(j = 1, \ldots, n_\alpha\), the corresponding invariant probabilities.

Let \(S_\alpha := \lim \frac{1}{m} \sum_{m=0}^{m} P_\alpha^m\). The average costs \(g_\alpha\) are equal to \(S_\alpha r_\alpha\) and are constant on \(E_{\alpha j}\), \(j = 1, \ldots, n_\alpha\), these constants are denoted by \(g_{\alpha j}\). The existence of a metric \(\rho\) on \(A\) is assumed such that
\[
\lim_{\rho(\alpha, \alpha_0) \to 0} \|P_\alpha - P_{\alpha_0}\| = 0 \quad \text{for all } \alpha_0 \in A
\]
\[
\lim_{\rho(\alpha, \alpha_0) \to 0} |\pi_{\alpha j} r_\alpha - \pi_{\alpha_0 j} r_{\alpha_0}| = 0 \quad \text{for all } \alpha_0 \in A, \ j = 1, \ldots, n_\alpha.
\]

We are interested in the question if these conditions imply the continuity of \(g_{\alpha j}\) and \(\pi_{\alpha j}\) in \(\alpha\).
The subsets $A_n$ of $A$ are defined by $A_n := \{ \alpha \in A \mid \dim N(I - P_\alpha) = n \}$.

**Lemma 4.3.** Let $\mu$ be an arbitrary nonnegative measure on $\Sigma$. The function $\mu g_\alpha$ is continuous in $\alpha$ on the subsets $A_n$ of $A$.

**Proof.** Let $\alpha_0 \in A_n$ and $\mu$ is a positive measure on $\Sigma$. We have

$$\mu g_\alpha = \mu(S_\alpha r_\alpha) = (\mu S_\alpha) r_\alpha.$$ 

Hence

$$|\mu g_\alpha - \mu g_{\alpha_0}| = |(\mu S_\alpha) r_\alpha - (\mu S_{\alpha_0}) r_{\alpha_0}| \leq |(\mu S_\alpha - \mu S_{\alpha_0}) r_\alpha| +$$

$$+ |\mu S_{\alpha_0}(r_\alpha - r_{\alpha_0})|.$$ 

By lemma 2.8b, $S_\alpha$ is continuous in $\alpha$ on $A_n$. Together with the boundedness of $r_\alpha$, uniform on $A$, this implies

$$\lim_{\rho(\alpha, \alpha_0) \to 0} |(\mu S_\alpha - \mu S_{\alpha_0}) r_\alpha| = 0 \quad (\alpha \in A_n).$$

Since $\mu S_{\alpha_0}$ is a linear combination of $\pi_{\alpha_0}, \ldots, \pi_{\alpha_0}$, the assumption

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

implies

$$\lim_{\rho(\alpha, \alpha_0) \to 0} |\mu S_{\alpha_0}(r_\alpha - r_{\alpha_0})| = 0.$$ 

This completes the proof. \qed

**Remark.** This result implies the continuity of $g_\alpha$ as function on $A_1$. However, the condition $\lim_{\rho(\alpha, \alpha_0) \to 0} \|P_\alpha - P_{\alpha_0}\| = 0$ is unnecessarily strong. This condition can be replaced by $\lim_{\rho(\alpha, \alpha_0) \to 0} \|P^k(\alpha - P_{\alpha_0})\| = 0$ for some $k \geq 1$, (see [8]).
Since $S\alpha$ is continuous in $\alpha$ on $\mathcal{A}_1$ the function $\pi_{\alpha_1}$ is also continuous in $\alpha$ on $\mathcal{A}_1$. If $\{\alpha_k\}_{k=1}^\infty$ is a sequence in $\mathcal{A}_1$ with limit $\alpha_0 \in A_1$ then

$$\lim_{k \to \infty} \|\pi_{\alpha_k} - \pi_{\alpha_0}\| = 0.$$ 

The case where $\alpha_0$ is not necessarily an element of $A_1$ is considered in the next lemma, the proof of which can be found in [8], section 4.1.

**Lemma 4.4.** Let $\{\alpha_k\}_{k=1}^\infty$ be a sequence in $A_1$ converging to $\alpha_0 \in A$, such that $\lim_{k \to \infty} \pi_{\alpha_k} (E_{\alpha_0})$ exists for all $j = 1, 2, \ldots, n_{\alpha_0}$. Let for all $k$ and $j$ with $\pi_{\alpha_k} (E_{\alpha_0}) > 0$ the measures $\pi_{\alpha_k}^j$ be defined by

$$\pi_{\alpha_k}^j (E) := \frac{\pi_{\alpha_k} (E)}{\pi_{\alpha_k} (E_{\alpha_0})}, \quad E \in \Sigma.$$ 

Then for all $j$ with $\lim_{k \to \infty} \pi_{\alpha_k} (E_{\alpha_0}) > 0$

$$\lim_{k \to \infty} \|\pi_{\alpha_k}^j - \pi_{\alpha_0}^j\|_{E_{\alpha_0}} = 0,$$

(where $\|\cdot\|_{E_{\alpha_0}}$ is the norm of the measure restricted to $E_{\alpha_0}$).

Notice that $\pi_{\alpha_k}^j$ is the probability $\pi_{\alpha_k}^j$ conditioned to being in $E_{\alpha_0}$.

### 4.3. Existence of optimal strategies

We consider the same SMD as in the preceding subsection. In lemma 4.3 the continuity of $g_{\alpha_1}$ on $A_1$ was proved, hence if $A = A_1$ and $A$ is compact an optimal strategy exists. In the next lemma a straightforward generalisation of this result is formulated.

**Lemma 4.5.** Let $A$ be compact and $A_n$ closed in $A$ for all $n \in \mathbb{N}$. If $n_{\alpha}$ is bounded on $A$, a $\mu$-optimal strategy exists for each nonnegative measure $\mu$. 

One can prove that in case the SMD is complete, the \( \mu \)-optimality of \( \alpha_0 \) implies that \( g_{\alpha_0}(u) \leq g_{\alpha}(u) \), \( \mu \)-almost everywhere, for all \( \alpha \in A \) (see [8]).

The following condition for closedness of \( A_n \) in \( A \) is a direct consequence of lemma 2.8c.

**Lemma 4.6.** If there is a \( \beta, 0 < \beta < 1 \), such that for all \( \alpha \in A \) the spectrum of \( P_\alpha \) has no points of absolute value between \( \beta \) and 1, then the set \( A_n \) is closed in \( A \) for all \( n \in \mathbb{N} \).

If for all \( \alpha \in A \) there is an \( \alpha_1 \in A_1 \) such that \( g_{\alpha_1}(u) \leq g_{\alpha}(u) \), \( u \in V \), then \( A_1 \) is said to *dominate* \( A \). This case is easier to analyse than the most general one. We shall formulate conditions for \( A_1 \) to dominate \( A \).

**Definition 4.7.** The SMD is called *communicative* if for all \( \alpha \in A \) and \( j = 1, \ldots, n_\alpha \) there is an \( \alpha_1 \in A_1 \) such that \( \pi_{\alpha_1}(E_{\alpha_1}) > 0 \).

This concept is introduced by Bather [1] for a finite state space and used by Hordijk [3] for a countable state space.

**Lemma 4.8.** If the SMD is complete and communicative, \( A_1 \) dominates \( A \).

**Proof.** Let \( \alpha \in A \). Choose \( j_0 \) such that \( g_{\alpha j_0} = \min_{j=1, \ldots, n_\alpha} g_{\alpha j} \). The communicativity implies the existence of an \( \alpha_1 \in A_1 \) such that \( \pi_{\alpha_1}(E_{\alpha j_0}) > 0 \). Let \( \alpha_2 := \alpha E_{\alpha j_0} \alpha_1 \) (possible by the completeness). Then \( \alpha_2 \in A_1 \), \( E_{\alpha j_0} \) is the only maximal invariant set, and \( g_{\alpha_2}(u) = g_{\alpha j_0} \) for \( u \in E_{\alpha j_0} \).

Let \( A := E_{\alpha j_0} \) and \( B := V \setminus A \). Instead of \( P_{B} \) and \( P_{A} \) we write \( P_{B} \) and \( P_{A} \). Using \( g_{\alpha_2} = P_{A} g_{\alpha_2} \), we get

\[
P_n P_{\alpha_2}^n = P_{\alpha_2}^n (P_{\alpha_2} - P_{A} B) g_{\alpha_2} = P_{\alpha_2}^n B g_{\alpha_2} - P_{\alpha_2}^{n+1} B g_{\alpha_2}.
\]

Hence

\[
\sum_{n=0}^{\infty} P_{\alpha_2}^n P_{\alpha_2} B g_{\alpha_2} = g_{\alpha_2}.
\]
Since \( g_{a_2} = g_{a_j_0} \) on \( A \), this implies

\[
g_{a_2} = g_{a_j_0} \cdot \sum_{n=0}^{\infty} p^n_{a_2B} p_{a_2A}^n v = g_{a_j_0} \{ 1_v - \lim_{n \to \infty} p^n_{a_2B} 1_v \}.
\]

But \( \lim_{n \to \infty} (p^n_{a_2B} 1_v)(u) \) is the probability that starting in \( u \) the process \( p_{a_2B} \) will never reach the set \( A \) and because of \( \pi_{a_2}^1(\{E_{a_j_0}\}) > 0 \) this probability is equal to 0. Therefore \( g_{a_2}(u) = g_{a_j_0} \) for all \( u \in V \), which completes the proof.

This result guarantees the existence of an optimal strategy if \( A_1 \) is compact and if the SMD is complete and communicative.

In the next theorem it will be shown that the compactness of \( A \) is sufficient even if \( A_1 \) is not closed in \( A \).

**Theorem 4.9.** Let the SMD be complete and communicative. If \( A \) is compact then an optimal strategy exists.

**Proof.** Let \( g := \inf_{a \in A_1} g_{a_1} \). The compactness of \( A \) implies the existence of a sequence \( \{a_k\} \) in \( A_1 \) such that \( \lim g_{a_k} = g \) and \( \lim a_k = a_0 \) for some \( a_0 \in A \).

Instead of \( E_{a_0} \) we shall write \( E_j \). Let \( \Delta := V \setminus \cup_{j=1}^{n_0} E_j \). Without loss of generality we may assume that \( \lim_{k \to \infty} \pi_{a_k}^1(\{E_j\}) \) exists for all \( j = 1, 2, \ldots, n_0 \). These limits are denoted by \( \pi_j \). We have

\[
g_{a_k} = \pi_{a_k} S_{a_k} \pi_{a_k} = \pi_{a_k} S_{a_k} \pi_{a_k} = \int_{\Delta} r_{a_k}(u) \pi_{a_k}^1(du) + \sum_{j=1}^{n_0} \int_{E_j} r_{a_k}(u) \pi_{a_k}^1(du).
\]

Further

\[
\pi_{a_k} = \pi_{a_k} P_{a_k} = \pi_{a_k} P_{a_0} + \pi_{a_k} (P_{a_k} - P_{a_0})
\]
The restriction of the measure $\pi_{\alpha_k}^{1\Delta \alpha_0}$ to $\Sigma_{\Delta}$ is identical to the restriction of the measure $\pi_{\alpha_k}^{1\Delta \alpha_0}$ to $\Sigma_{\Delta}$. Hence $\pi_{\alpha_k}^{1\Delta \alpha_0} = \pi_{\alpha_k}^{1\Delta \alpha_0} + \pi_{\alpha_k}^{(P_{\alpha_k} - P_{\alpha_0})}$ (as equation in $(\Delta_{\alpha}, \Sigma_{\alpha})$). The convergence of $\alpha_k$ to $\alpha_0$ implies

$$\lim_{k \to \infty} \| \pi_{\alpha_k}^{1\Delta \alpha_0} - P_{\alpha_0} \|_{\Delta} = 0.$$ 

By the strong ergodic theorem the spectral radius of $I_{\Delta} P_{\alpha_0}$ is smaller than 1. Therefore $\lim_{k \to \infty} \| \alpha_k \|_{\Delta} = 0$ and $\lim_{k \to \infty} \int_{\Delta} r_{\alpha_k}^{1\Delta \alpha_0} (du) = 0$. Now we have to consider

$$\int_{\Delta} r_{\alpha_k}^{1\Delta \alpha_0} (du).$$ 

If $\pi_j = 0$,

$$\lim_{k \to \infty} \int_{\Delta} r_{\alpha_k}^{1\Delta \alpha_0} (du) = 0.$$ 

Let $\pi_j > 0$. By lemma 4.4

$$\lim_{k \to \infty} \{ \int_{\Delta} r_{\alpha_k}^{1\Delta \alpha_0} (du) - \pi_j \int_{\Delta} r_{\alpha_k}^{1\Delta \alpha_0} (du) \} = 0.$$ 

Using the assumption

$$\lim_{\rho(a_j a_0) \to 0} | \pi_{\alpha_0}^{1\Delta \alpha_0} - \pi_{\alpha_0}^{1\Delta \alpha_0} | = 0$$

we get

$$\lim_{k \to \infty} \pi_j \{ \int_{\Delta} r_{\alpha_k}^{1\Delta \alpha_0} (du) - \pi_j \int_{\Delta} r_{\alpha_k}^{1\Delta \alpha_0} (du) \} = 0$$

and hence

$$\lim_{k \to \infty} \int_{\Delta} r_{\alpha_k}^{1\Delta \alpha_0} (du) = \pi_j \int_{\Delta} r_{\alpha_0}^{1\Delta \alpha_0} (du).$$ 

But

$$g_{\alpha_0}^{\pi_j} = \pi_{\alpha_0}^{1\Delta \alpha_0} = \pi_{\alpha_0}^{1\Delta \alpha_0} r_{\alpha_0}^{1\Delta \alpha_0} = \pi_{\alpha_0}^{1\Delta \alpha_0} r_{\alpha_0}^{1\Delta \alpha_0} = \int_{\Delta} r_{\alpha_0}^{1\Delta \alpha_0} (du).$$
Hence

\[ g = \lim_{k \to \infty} g_{\alpha_k} = \sum_{j=1}^{n_{\alpha_0}} \pi_j \cdot g_{\alpha_0}^j. \]

We had \( \lim_{k \to \infty} \| \pi_{\alpha_k} \|_\Delta = 0 \), which implies that \( \lim_{k \to \infty} \pi_{\alpha_k}(\Delta) = 0 \) and

\[ \sum_{j=1}^{n_{\alpha_0}} \pi_j = \sum_{j=1}^{n_{\alpha_0}} \lim_{k \to \infty} \pi_{\alpha_k}(E_j) = 1. \]

Now, by (1),

\[ \min_{j=1,2,\ldots,n_{\alpha_0}} \{ g_{\alpha_0}^j \} \leq g. \]

This implies, by lemma 4.8, the existence of a strategy \( \alpha_1 \in A_1 \) such that

\[ g_{\alpha_1} \leq \lim_{j=1,\ldots,n_{\alpha_0}} \{ g_{\alpha_0}^j \} \leq g. \]

This strategy \( \alpha_1 \) is optimal.

\[ \square \]

References


