AN $L_p$ THEORY FOR A MARKOV PROCESS WITH A SUB-INARIANT MEASURE

S. R. FOGUEL

Let a Markov process be given by its transition probability. If the process has a sub-invariant measure then it defines an operator (or a semigroup of operators) on $L_p$, $1 \leq p \leq \infty$. Using this representation we study:

1. The behaviour of $P^n(x, A)$ as $n \to \infty$.
2. Continuity and differentiability properties of $P_t(x, A)$.

(The notation is explained below.)

Kendall in [6] and [7] uses a similar method in the case where the state space is discrete and the process irreducible. Instead of irreducibility we assume a "Doeblin condition" and we do not restrict the state space. In [1], [2] and [3] the same problems were studied. In this paper we do not assume that the measure is invariant but only sub-invariant (this idea was employed in Kendall [6], [7]). Also we do not assume that the process is "honest"; we have $P(x, X) \leq 1$ instead of equality.

1. Definitions and notation. Let $(X, \Sigma, \mu)$ be a measure space (the State Space) where $\mu \geq 0$ but is not necessarily finite. Let $P(x, A)$ be a transition probability:

(a) $P(x, A)$ is defined for $x \in X$ and $A \in \Sigma$ and $0 \leq P(x, A) \leq 1$.
(b) For a fixed $x$ the function $P(x, \cdot)$ is a measure on $\Sigma$.
(c) For a fixed $A \in \Sigma$ the function $P(\cdot, A)$ is measurable.

The measure $\mu$ is assumed to satisfy

$$\mu(A) \geq \int P(x, A) \mu(dx).$$

(The measure $\mu$ is sub-invariant.)

Let

$$(Pf)(x) = \int f(y) P(x, dy).$$

Now if $f = \sum c_i I(A_i)$, where $I(A)$ denotes the characteristic function of $A$, and $\mu(A_i) < \infty$, then:

Received by the editors October 26, 1963.

1 This work was partially supported by NSF Grants GP87 and G14736 at Northwestern University.
The above argument shows that, on $L_1(X, \Sigma, \mu)$,

$$\|P\|_1 \leq 1.$$  

Since, clearly, $\|P\|_\infty \leq 1$ we can use the Riesz Convexity Theorem to conclude:

*For every $1 \leq p \leq \infty$ the operator $P$ on $L_p(X, \Sigma, \mu)$ has norm less than or equal to one.*

For our first problem the case $p = 2$ will be important. We shall use the following theorems on $L_2$.

Let

$$H_0 = \{f \mid f \in L_2 \text{ and weak lim } P^nf = 0\}.$$

$$H_1 = H_0^\perp.$$

$$K = \{f \mid f \in L_2 \text{ and } \|P^nf\| = \|P^*nf\| = \|f\| \text{ for all } n\}.$$

**Theorem 1.1.** The sets $H_0$, $H_1$ and $K$ are subspaces invariant under $P$ and $P^*$. The restriction of $P$ to $K$ is unitary and $H_1 \subset K$. The subspace $K$ is of the form $L_2(X, \Sigma_1, \mu)$ where $\Sigma_1$ is a $\sigma$-subfield of $\Sigma$. If $\sigma \in \Sigma_1$ then $P(\sigma) = I(\tau)$ where $\tau \in \Sigma_1$ and $P$ is an automorphism of $\Sigma_1$.

**Proof.** This theorem follows from Theorem 1.1 of [5] and Theorem 2 of [4]. We have only to show that $P^*$ satisfies Condition 3 of Definition 1 of [4]. Let $f \in L_2$ and $f \leq c$. Let $A$ be any set of finite measure, then

$$\int_A P^*f \, d\mu = \int f P(I(A)) \, d\mu \leq c \int P(x, A) \mu(dx) \leq c \mu(A).$$

Thus

$$P^*f \leq c \text{ a.e.}$$

**Lemma 1.2.** Let $H$ be a Hilbert space and $T$ an operator on $H$ with $\|T\| \leq 1$. For any $x \in H$:

$$\text{weak lim } T^nx = 0 \text{ if } (T^nx, x) \to 0 \text{ for some } k.$$  

((x, y) denotes the inner product in $H$.)

**Proof.** If $((T^k)x, x) \to 0$, then by Theorem 3.1 of [5], weak lim$_{n \to \infty} (T^k)x = 0$. But then also
weak lim \( T^d T^{k_n} x = 0 \), \( d = 0, 1, \ldots, k - 1 \).

**Remark.** If \( \lambda \) is a finite positive measure on \([0, 2\pi]\) and \( c_n \) are its Fourier coefficients, then the lemma implies: \( c_{k_n} \to 0 \) iff \( c_n \to 0 \).

2. **Convergence of** \( P_n(x, A) \) as \( n \to \infty \). Let \( P_n(x, A) \) be defined as usual; then

\[
(P^nf)(x) = \int f(y) P^n(x, dy).
\]

**Theorem 2.1.** Let \( A \) be a set of finite measure. If, for some integer \( k \), the sequence \( P^{kn}(x, A) \) converges in measure, on \( A \), to zero, then \( P^n(x, A) \) converges in measure to zero, on every set of finite measure.

**Proof.** Our assumption implies

\[
\int_A P^{kn}(x, A) \mu(dx) \to 0.
\]

Thus, by Lemma 1.2, the sequences \( P^n(x, A) = P^n(I(A)) \) converge weakly, in the \( L_2 \) sense, to zero. Since \( P^n(x, A) \geq 0 \) this is the same as:

The sequence \( P^n(x, A) \) converges in measure, on every set of finite measure, to zero.

If one studies the transition probability on \((X, \Sigma, \mu)\), then, since \( P \) is an automorphism on \( \Sigma \), we have:

(d) For any \( A \in \Sigma \) there exist two sets \( B \) and \( C \) in \( \Sigma \) such that

\[
P(x, A) = 1 \text{ if } x \in B, \quad P(x, A) = 0 \text{ if } x \in B \text{ a.e.}
\]

\[
P(x, C) = 1 \text{ if } x \in A, \quad P(x, C) = 0 \text{ if } x \in A \text{ a.e.}
\]

The process on \((X, \Sigma, \mu)\) is a Deterministic Process: \( P^n(x, A) \) is either zero or one, a.e., for every \( x \in X \) and \( A \in \Sigma \).

Let

\[
X_0 = \{ x \mid P(x, X) < 1 \}.
\]

**Lemma 2.2.** If \( B \subset X_0 \) and \( A \in \Sigma \) then \( \mu(B \cap A) = 0 \).

**Proof.** By condition (d),

\[
I(A) = P(x, C).
\]

Thus, if \( x \in A \),

\[
P(x, X) \geq P(x, C) = 1 \text{ a.e. or a.e. } x \in B.
\]

**Theorem 2.3.** Let \( B \) be a set of finite measure such that if \( A \in \Sigma \) then \( \mu(B \cap A) = 0 \). Then
lim in measure of $P^n(x, B) = 0$ on every set of finite measure.

Proof. Since $I(B)$ is orthogonal to $K$ then, by Theorem 1.1, $I(B) \subset H_1 = H_0$.

We shall assume in the rest of this section that a "Doeblin Condition" holds, namely: there exists a positive measure $\lambda$, on $\Sigma$, and an $\epsilon > 0$ such that:

1. If $\mu(A) < \infty$ then $\lambda(A) < \infty$.
2. If $\lambda(A) < \epsilon$ then for some $n$, $P^n(x, A) < 1$ for all $x \in X$.

This kind of condition was used in [3]. If the state space $X$ is discrete, then by taking $\lambda$ to be the number of points in the set and $\epsilon = 1/2$ we see that the condition is trivially satisfied.

Lemma 2.4. The $\sigma$-field $\Sigma_1$ is generated by a collection of disjoint sets $\{\sigma_\alpha\}$.

Proof. It is enough to show that if $A \in \Sigma_1$ then it contains an atom of $\Sigma_1$. Now if $A$ contains $k$ disjoint sets of $\Sigma_1$ then $\epsilon \leq \lambda(A)/k$ since every set in $\Sigma_1$ must have a $\lambda$-measure greater than $\epsilon$. Thus the number of sets contained in $A$ is smaller than $\lambda(A)/\epsilon$.

The operator $P$ permutes the sets $\sigma_\alpha$. Let us write $P^n\sigma_\alpha$ for the set whose characteristic function is $P^n(I(\sigma_\alpha))$. For each $\sigma_\alpha$ there are two possibilities:

1. The set $\sigma_\alpha$ is cyclic: for some $k(\alpha)$, $P^{k(\alpha)}\sigma_\alpha = \sigma_\alpha$.
2. The sets $P^n\sigma_\alpha$ are disjoint.

(We used here the fact that

$$\mu(P^n\sigma_\alpha \cap P^k\sigma_\alpha) = \mu(P^{n-k}\sigma_\alpha \cap \sigma_\alpha)$$

which holds since $P$ is unitary on $K$.)

Theorem 2.5. Let $A \in \Sigma$.

(a) If $A \cap \sigma_\alpha = \emptyset$ for every $\alpha$ then the sequence $P^n(x, A)$ tends, in measure, to zero, on every set of finite measure.

(b) If $A \subset \sigma_\alpha$ where $\sigma_\alpha$ is not cyclic then the sequence $P^n(x, A)$ tends, in measure, to zero, on every set of finite measure.

(c) If $A \subset \sigma_\alpha$ and $P^n\sigma_\alpha = \sigma_\alpha$ and $B$ is a set of finite measure then:

$$\lim_{n \to \infty} \int_B P^{nk+d}(x, A) \mu(dx) = \mu(\sigma_\alpha)^{-1} \mu(B \cap P^d\sigma_\alpha),$$

$$d = 0, 1, \ldots, k - 1.$$

Proof. Case (a) was proved in Theorem 2.3. Case (b): Since the sets $P^n(\sigma_\alpha)$ are disjoint and

$$P^n(x, \sigma_\alpha) \leq P^n(x, \sigma_\alpha) = I(P^n(\sigma_\alpha)),$$
the sequence $P^n(x, A)$ tends weakly, in the sense of $L_2$, to zero, which implies (b). Case (c): Note that

$$I(A) = \frac{\mu(A)}{\mu(\sigma_u)} I(\sigma_u) + f,$$

where $f$ has its support in $\sigma_u$ and is orthogonal (in $L_2$) to $I(\sigma_u)$. Thus $f$ is orthogonal to $K$ and weak lim $P^n f = 0$ by Theorem 1.1. On the other hand,

$$P^n k + d I(\sigma_u) = I(P^n \sigma_u).$$

Thus

$$\int_B P^{n_k + d}(x, A) \mu(dx) = \frac{\mu(A)}{\mu(\sigma_u)} \mu(B \cap P^n \sigma_u) + \int_B P^{n_k + d} d \mu$$

and

$$\int_B P^{n_k + d} d \mu \rightarrow 0.$$

3. Markov process with continuous time parameter. Let the transition probability be $P_t(x, A)$, $0 \leq t$, where $P_t(x, A)$ satisfies conditions (a), (b) and (c) of the introduction and the Chapman-Kolmogoroff Equation:

$$P_{t+s}(x, A) = \int P_t(x, dy) P_s(y, A).$$

We shall replace equation (1.1) by

$$(3.1) \mu(A) \geq \int P_t(x, A) \mu(dx), \quad t \geq 0.$$  

It is easily seen that the transition probability will define a semigroup of contractions on $L_p$, $1 \leq p \leq \infty$, by

$$(P_t f)(x) = \int f(y) P_t(x, dy).$$

We shall study in this section continuity and differentiability properties of the semigroup $P_t$.

Let us assume that $P_0(x, A) = I(A)(x)$.

If $\mu(A) < \infty$ and $P_t I(A) = P_t(x, A)$ is continuous in the $L_p$ topology for $1 \leq p < \infty$, then

$$(3.2) \int_A P_t(x, A) \mu(dx) \rightarrow \mu(A),$$
since $I(A) \in L_q$ where $1/p + 1/q = 1$. Conversely, if (3.2) holds then

$$\|P_t(x, A) - I(A)\|^p = \int |P_t(x, A) - I(A)|^p d\mu \leq \int |P_t(x, A) - I(A)| d\mu,$$

since $0 \leq P_t(x, A) \leq 1$. Thus

$$\|P_t(x, A) - I(A)\|^p \leq \int_A (1 - P_t(x, A)) d\mu + \int_{X-A} P_t(x, A) d\mu = \mu(A) - \int_A P_t(x, A) d\mu + \int_{X-A} P_t(x, A) d\mu.$$

Now

$$\int_A P_t(x, A) d\mu + \int_{X-A} P_t(x, A) d\mu = \int_P P_t(x, A) d\mu \leq \mu(A).$$

Thus

$$\|P_t(x, A) - I(A)\|^p \leq 2(\mu(A) - \int_A P_t(x, A) \mu(dx)).$$

We shall assume that (3.2) holds for every set $A \in \Sigma$ with $\mu(A) < \infty$, and thus the semigroup $P_t$ is strongly continuous for $1 \leq p < \infty$ and $t \geq 0$.

Note that if the process was given by the functions $x_1$ then

$$\int_A P_t(x, A) \mu(dx) = \text{Probability}(x_1 \in A \cap x_0 \in A).$$

If $P_t(x, A)$ is strongly differentiable for $t=0$ in the $L_p$ sense $1 \leq p < \infty$, then

$$\frac{1}{t} \left( \mu(A) - \int_A P_t(x, A) \mu(dx) \right) = \frac{1}{t} \int_A (I(A) - P_t(x, A)) \mu(dx)$$

would be bounded since $I(A) \in L_q$ $(1/p + 1/q = 1)$.

If one assumes that

$$t^{-p} \left[ \mu(A) - \int_A P_t(x, A) \mu(dx) \right], \quad p > 1,$$

is bounded, then, by (3.3), the quotient

$$t^{-1}(P_t(x, A) - I(A))$$
is bounded in the $L_p$ norm. But $L_p$ is a reflexive space and boundedness implies differentiability. In this case it is easy to see that
\[
\frac{d}{dt} \int_A P_t(x, A) \mu(dx) \bigg|_{t=0} = 0,
\]
\[
\frac{d}{dt} \int_B P_t(x, A) \mu(dx) \bigg|_{t=0} = 0 \quad \text{if } B \subset X - A.
\]

We do not get any information on $(d/dt)\int_B P_t(x, A) \mu(dx)$, $B \subset A$.

**Definition.** Let $D^* \subset L_\infty(X, \Sigma, \mu) = L^*_\infty$ be the set
\[
\{ f \ | \ f \in L_\infty \text{ and } \lim\| P_t^* f - f \|_\infty = 0 \}.
\]

**Theorem 3.1.** The set $D^*$ is a closed subspace of $L_\infty$, invariant under $P_t^*$. For every $g \in L_1$,
\[
\| g \|_1 = \sup \left\{ \left| \int g f \mu \right| \ | f \in D^* \text{ and } \| f \|_\infty = 1 \right\}.
\]

This theorem is a consequence of Theorems 2.1, 2.2 and 3.1 of [8].

Throughout the rest of this paper $A$ will denote a fixed set of finite measure such that
\[
i(A) - \int P_t(x, A) \mu(dx)\]

is bounded.

**Theorem 3.2.** If $f \in D^*$, then the function $\int f(x) P_t(x, A) \mu(dx)$ is differentiable for every $t \geq 0$.

**Proof.** Since $P_t^* f \in D^*$ it is enough to prove differentiability at $t=0$. Now, by semigroup theory, the function $P_t^* g$ is strongly differentiable on a dense subset of $D^*$. On the other hand, for any $h$ with $\| h \|_\infty \leq 1$,
\[
t^{-1} \left| \int h(x) P_t(x, A) \mu(dx) - \int_A h(x) \mu(dx) \right|
\leq t^{-1} \left( \mu(A) - \int A P_t(x, A) \mu(dx) \right).
\]

Thus the quotient
\[
t^{-1} \left( \int h(x) P_t(x, A) \mu(dx) - \int_A h(x) \mu(dx) \right)
\]
is uniformly bounded on $\|k\|_{\infty} \leq 1$ and tends to a limit on a dense subset on $D^*$. Thus the limit exists for every $f \in D^*$.

Let $\Gamma \subseteq \Sigma$ be such that every set in $\Gamma$ has a finite measure, and every set of finite measure, in $\Sigma$, is a countable union of sets in $\Gamma$.

Let $f \in L_\infty$:

$$\|P_t f - f\| = \sup \left\{ \mu(B)^{-1} \left| \int_B (P_t f - f) \, d\mu \right| \mid B \subseteq \Gamma \right\}$$

$$= \sup \left\{ \mu(B)^{-1} \left| \int f(x) (P_t(x, B) - I(B)) \mu(dx) \right| \mid B \subseteq \Gamma \right\}.$$

In particular, $I(C) \in D^*$ if

$$\sup \left\{ \mu(B)^{-1} \int_C P_t(x, B) \mu(dx) \mid B \subseteq \Gamma \text{ and } B \cap C = \emptyset \right\} \to 0, \quad t \to 0,$$

and

$$\sup \left\{ (1 - \mu(B)^{-1}) \int_C P_t(x, B) \mu(dx) \mid B \subseteq \Gamma \text{ and } B \subset C \right\} \to 0, \quad t \to 0.$$

For a discrete process this implies $P_t(i, j)$ is differentiable for $t \geq 0$ if

$$t \to (1 - P_t(j, j))$$

is bounded and

$$\sup \left\{ \mu_k^{-1} \mu_t(i, k) \mid k, k \neq i \right\} \to 0, \quad t \to 0,$$

where $\mu_k$ is the sub-invariant measure.

For our last result we shall need the notation of [3]. Let us add the assumptions:

$$P(x, X) = 1 \quad \text{for all } x \in X,$$

$$\mu(A) = \int P_t(x, A) \mu(dx).$$

One can define a measure space $(\Omega, T, \nu)$ and a class of measurable functions $x_t$ on $\Omega$ to $X$, where $t \geq 0$, such that: If $T(t)$ is given by

$$T(t)f(x_0(\omega)) = f(x_t(\omega)),$$

$E_0$ is the self-adjoint projection, in $L_2(\Omega, T, \nu)$, on $L_2(\Omega, \nu_0^{-1}(\Sigma), \nu)$, then:
\[ \mu(A) = \nu(x_0^{-1}(A)), \]
\[ E_0 T(t) I(x_0 \in A)(\omega) = P_t(x_0(\omega), A), \]
\[ E_0 T(t)f(x_0(\cdot))(\omega) = \int f(y) P_t(x_0(\omega), dy). \]

**Theorem 3.3.** Let \( f \in L_1(\mu) \cap L_\infty(\mu) \) be such that
\[ \|f(x(\omega)) - f(x_0(\omega))\|_\infty \to 0, \quad t \to 0; \]
then
\[ \int f(y) P_t(y, A) \mu(dy) \]
is differentiable for \( t \geq 0 \).

**Proof.** It is enough to show that \( f \in D^* \). Now for every \( g \in L_1(\nu) \cap L_\infty(\nu) \),
\[ (E_0 T(t))^* g = T(t)^{-1} E(t) g, \]
where \( E(t) \) is the self-adjoint projection on \( L_2(\Omega, x_t^{-1}(\Sigma), \nu) \) (see Theorem 1.1 of \[2\]). Thus
\[ \| (E_0 T(t))^* f(x_0(\omega)) - f(x_0(\omega)) \|_\infty = \| T(t)^{-1} (E(t)f(x_0(\omega)) - T(t)f(x_0(\omega))) \|_\infty. \]
But \( T(t)^{-1} \) is an isometry on \( L_\infty \); thus
\[ \| (E_0 T(t))^* f(x_0(\omega)) - f(x_0(\omega)) \|_\infty = \| E(t)(f(x_0(\omega)) - T(t)f(x_0(\omega))) \|_\infty \leq \| f(x_0(\omega)) - f(x_t(\omega)) \|_\infty. \]

**Bibliography**


Northwestern University and Hebrew University, Jerusalem, Israel