A Note on the Brikhoff Ergodic Theorem

Nikola Sandrić

University of Zagreb

Probability and Analysis

Będlewo

May 15-19, 2017

Outline

Motivation and preliminaries

Let $\mathbf{M} = (\Omega, \mathcal{F}, \{\mathbb{P}^x\}_{x\in S}, \{\mathcal{F}_t\}_{t\in\mathbb{T}}, \{\theta_t\}_{t\in\mathbb{T}}, \{M_t\}_{t\in\mathbb{T}})$ be a Markov process with state space $(S, \mathcal{S})$. Here, $\mathbb{T}$ is the time set $\mathbb{Z}_+$ or $\mathbb{R}_+$.

Let $\mathbf{M} = (\Omega, \mathcal{F}, \{\mathbb{P}^x\}_{x\in S}, \{\mathcal{F}_t\}_{t\in\mathbb{T}}, \{\theta_t\}_{t\in\mathbb{T}}, \{M_t\}_{t\in\mathbb{T}})$ be a Markov process with state space $(S, \mathcal{S})$. Here, $\mathbb{T}$ is the time set $\mathbb{Z}_+$ or $\mathbb{R}_+$.

A measure $\pi(dy)$ on $\mathcal{S}$ is said to be invariant for $\mathbf{M}$ if

$$\int_S p^t(x, dy)\pi(dx) = \pi(dy), \qquad t \in \mathbb{T}.$$

Motivation and preliminaries

Let $\mathbf{M} = (\Omega, \mathcal{F}, \{\mathbb{P}^x\}_{x\in S}, \{\mathcal{F}_t\}_{t\in\mathbb{T}}, \{\theta_t\}_{t\in\mathbb{T}}, \{M_t\}_{t\in\mathbb{T}})$ be a Markov process with state space $(S, \mathcal{S})$. Here, $\mathbb{T}$ is the time set $\mathbb{Z}_+$ or $\mathbb{R}_+$.

A measure $\pi(dy)$ on $\mathcal{S}$ is said to be invariant for $\mathbf{M}$ if

$$\int_S p^t(x, dy)\pi(dx) = \pi(dy), \qquad t \in \mathbb{T}.$$

A set $B \in \mathcal{F}$ is said to be shift-invariant (for $\mathbf{M}$) if $\theta_t^{-1}B = B$ for all $t \in \mathbb{T}$. The shift-invariant σ-algebra $\mathcal{I}$ is a collection of all such shift-invariant sets.

Theorem (Birkhoff ergodic theorem)

Let $\mathbf{M}$ be a Markov process with invariant probability measure $\pi(dy)$. Then, for any $f \in L^p(S, \pi)$, $p \geq 1$, the following limit holds

$$\lim_{t\to\infty} \frac{1}{t} \int_{[0,t)} f(M_s)\tau(ds) = \mathbb{E}^{\pi}[f(M_0)|\mathcal{I}] \qquad \mathbb{P}^{\pi}\text{-a.s. and in } L^p(\Omega, \mathbb{P}^{\pi}),$$

where $\tau(dt)$ is the counting measure when $\mathbb{T} = \mathbb{Z}_+$ and Lebesgue measure when $\mathbb{T} = \mathbb{R}_+$.

A Markov process **M** is said to be ergodic if it possesses an invariant probability measure $\pi(dy)$ and if $\mathcal{I}$ is trivial with respect to $\mathbb{P}^{\pi}(d\omega)$, that is, $\mathbb{P}^{\pi}(B) = 0$ or 1 for every $B \in \mathcal{I}$.

Motivation and preliminaries

A Markov process **M** is said to be ergodic if it possesses an invariant probability measure $\pi(dy)$ and if $\mathcal{I}$ is trivial with respect to $\mathbb{P}^\pi(d\omega)$, that is, $\mathbb{P}^\pi(B) = 0$ or 1 for every $B \in \mathcal{I}$.

In addition to the assumptions of the Birkhoff ergodic theorem, if **M** is ergodic then we conclude

$$\lim_{t\to\infty} \frac{1}{t}\int_{[0,t)} f(M_s)\tau(ds) = \int_S f(y)\pi(dy) \qquad \mathbb{P}^\pi\text{-a.s. and in } L^p(\Omega, \mathbb{P}^\pi).$$

Motivation and preliminaries

A Markov process $\mathbf{M}$ is said to be ergodic if it possesses an invariant probability measure $\pi(dy)$ and if $\mathcal{I}$ is trivial with respect to $\mathbb{P}^\pi(d\omega)$, that is, $\mathbb{P}^\pi(B) = 0$ or 1 for every $B \in \mathcal{I}$.

In addition to the assumptions of the Birkhoff ergodic theorem, if $\mathbf{M}$ is ergodic then we conclude

$$\lim_{t\to\infty} \frac{1}{t} \int_{[0,t)} f(M_s)\tau(ds) = \int_S f(y)\pi(dy) \qquad \mathbb{P}^\pi\text{-a.s. and in } L^p(\Omega, \mathbb{P}^\pi).$$

Question: Can we conclude the above relation for any initial distribution of $\mathbf{M}$?

Motivation and preliminaries

Meyn and Tweedie (2009) have shown that the following are equivalent:

(a) the above relation holds $\mathbb{P}^{\mu}$-a.s. for any $f \in L^p(S, \pi)$ and any $\mu \in \mathcal{P}(S)$

(b) $\mathbf{M}$ is a positive Harris recurrent Markov process.

Meyn and Tweedie (2009) have shown that the following are equivalent:

- **(a)** the above relation holds $\mathbb{P}^{\mu}$-a.s. for any $f \in L^p(S, \pi)$ and any $\mu \in \mathcal{P}(S)$
- **(b)** **M** is a positive Harris recurrent Markov process.

A Markov process **M** is called φ-irreducible if for the σ-finite measure $\varphi(dy)$ on $\mathcal{S}$, $\varphi(B) > 0$ implies

$$\int_{\mathbb{T}} p^t(x, B)\tau(dt) > 0, \qquad x \in S.$$

Motivation and preliminaries

The process **M** is called Harris recurrent if it is φ-irreducible, and $\varphi(B) > 0$ implies

$$\int_{\mathbb{T}} 1_{\{M_t \in B\}} \tau(dt) = \infty \qquad \mathbb{P}^x\text{-a.s. for all } x \in S.$$

The process **M** is called Harris recurrent if it is φ-irreducible, and $\varphi(B) > 0$ implies

$$\int_{\mathbb{T}} 1_{\{M_t \in B\}} \tau(dt) = \infty \qquad \mathbb{P}^x\text{-a.s. for all } x \in S.$$

It is well known that every Harris recurrent Markov process admits a unique (up to constant multiplies) invariant (not necessary probability) measure. If the invariant measure is finite, then the process is called positive Harris recurrent; otherwise it is called null Harris recurrent.

Motivation and preliminaries

If **M** is also aperiodic, Meyn and Tweedie (2009) have proved that (a), (b) and (c) are equivalent to

(d) **M** is strongly ergodic.

Motivation and preliminaries

If **M** is also aperiodic, Meyn and Tweedie (2009) have proved that (a), (b) and (c) are equivalent to

(d) **M** is strongly ergodic.

A Markov process **M** is said to be strongly ergodic if there exists $\pi \in \mathcal{P}(S)$ such that

$$\lim_{t\to\infty} d_{TV}(p^t(x,dy),\pi(dy)) = 0, \qquad x \in S,$$

where d_{TV} denotes the total variation metric on $\mathcal{P}(S)$ given by

$$d_{TV}(\mu(dy),\nu(dy)) := \frac{1}{2} \sup_{f\in B_b(S),\,|f|_\infty\le 1} \left| \int_S f(y)\mu(dy) - \int_S f(y)\nu(dy) \right|.$$

In the discrete-time case, Hernandez-Lerma and Lasserre (2000) have shown that if **M** has a unique invariant probability measure $\pi(dy)$, then either

- **(i)** $\lim_{t\to\infty} d_{TV}(p^t(x, dy), \pi(dy)) = 0$ π-a.e., or
- **(ii)** $\pi(dy) \perp \sum_{t=1}^{\infty} p^t(x, dy)$ π-a.e. and $p^t(x, dy)$ converges weakly to $\pi(dy)$.

In the discrete-time case, Hernandez-Lerma and Lasserre (2000) have shown that if **M** has a unique invariant probability measure $\pi(dy)$, then either

(i) $\lim_{t\to\infty} d_{TV}(p^t(x,dy),\pi(dy)) = 0$ π-a.e., or

(ii) $\pi(dy) \perp \sum_{t=1}^{\infty} p^t(x,dy)$ π-a.e. and $p^t(x,dy)$ converges weakly to $\pi(dy)$.

Goal: To relax the notion of strong ergodicity and, under these new assumptions, conclude a version of the Birkhoff ergodic theorem which holds for any initial distribution of the process.

Main results

In the sequel we assume that $(S, \mathcal{S})$ is a Polish space with bounded (say by 1) metric d. In particular, $Lip(S) \subseteq C_b(S)$.

Main results

In the sequel we assume that $(S, \mathcal{S})$ is a Polish space with bounded (say by 1) metric d. In particular, $Lip(S) \subseteq C_b(S)$.

Wasserstein metric of order one on $\mathcal{P}(S)$ is defined by

$$d_W(\mu(dy), \nu(dy)) := \sup_{f \in Lip(S),\, |f|_{Lip} \le 1} \left| \int_S f(y)\mu(dy) - \int_S f(y)\nu(dy) \right|.$$

Main results

In the sequel we assume that $(S, \mathcal{S})$ is a Polish space with bounded (say by 1) metric d. In particular, $Lip(S) \subseteq C_b(S)$.

Wasserstein metric of order one on $\mathcal{P}(S)$ is defined by

$$d_W(\mu(dy), \nu(dy)) := \sup_{f \in Lip(S),\, |f|_{Lip} \le 1} \left| \int_S f(y)\mu(dy) - \int_S f(y)\nu(dy) \right|.$$

Recall, $d_W \le d_{TV}$, and d_W metrizes the weak convergence of probability measures. More precisely, $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(S)$ converges to $\mu \in \mathcal{P}(S)$ with respect to d_W if, and only if,

$$\lim_{n\to\infty} \int_S f(y)\mu_n(dy) = \int_S f(y)\mu(dy), \qquad f \in C_b(S).$$

Main results

Theorem

Assume that there is $\pi \in \mathcal{P}(S)$ satisfying

$$\lim_{t\to\infty} \sup_{s\in\mathbb{T}} \int_S d_W(p^t(y,dz),\pi(dz))p^s(x,dy) = 0, \qquad x \in S. \tag{1}$$

Then, for any $p \geq 1$, $f \in Lip(S)$ and $\mu \in \mathcal{P}(S)$,

$$\frac{1}{t}\int_{[0,t)} f(M_s)\tau(ds) \xrightarrow[t\nearrow\infty]{L^p(\Omega,\mathbb{P}^\mu)} \int_S f(y)\pi(dy), \tag{2}$$

where $\xrightarrow[t\nearrow\infty]{L^p(\Omega,\mathbb{P}^\mu)}$ denotes the convergence in $L^p(\Omega,\mathbb{P}^\mu)$.

Main results

Corollary

Assume the assumptions from the previous theorem. Then,

(i) (2) holds for all $f \in C_c(S)$ and $\mu \in \mathcal{P}(S)$. In particular, if S is compact, (2) holds for all $f \in C(S)$ and $\mu \in \mathcal{P}(S)$.

(ii) provided (S, d) is locally compact, (2) holds for all $f \in C_\infty(S)$ and $\mu \in \mathcal{P}(S)$.

(iii) provided $\pi(dy)$ is an invariant measure for $\mathbf{M}$ and (S, d) is locally compact,

$$\frac{1}{t}\int_{[0,t)} f(M_s)\tau(ds) \xrightarrow[t\nearrow\infty]{L^p(\Omega,\mathbb{P}^\pi)} \int_S f(y)\pi(dy), \quad p \geq 1,\ f \in L^p(S,\pi).$$

Main results

As in the strong ergodicity case, the relation in (1) implies that

(i) $\pi(dy)$ is the only measure satisfying (1);

(ii) if $\pi(dy)$ is invariant for **M**, then it is its unique invariant measure.

Main results

As in the strong ergodicity case, the relation in (1) implies that

(i) $\pi(dy)$ is the only measure satisfying (1);

(ii) if $\pi(dy)$ is invariant for $\mathbf{M}$, then it is its unique invariant measure.

Question: Does (1) imply invariance of $\pi(dy)$?

Main results

As in the strong ergodicity case, the relation in (1) implies that

(i) $\pi(dy)$ is the only measure satisfying (1);

(ii) if $\pi(dy)$ is invariant for **M**, then it is its unique invariant measure.

Question: Does (1) imply invariance of $\pi(dy)$?

If **M** is a Feller process the answer is yes. Another condition ensuring that (1) implies invariance of $\pi(dy)$ is contractivity of **M** with respect to d_W: for all $t \in \mathbb{T}$ and $\mu, \nu \in \mathcal{P}(S)$,

$$d_W\left(\int_S p^t(x,dy)\mu(dx), \int_S p^t(x,dy)\nu(dx)\right) \le d_W(\mu(dy),\nu(dy)).$$

Main results

Observe that for any $t \in \mathbb{T}$ and $x \in S$,

$$d_W(p^t(x,dy),\pi(dy)) \le \sup_{s\in\mathbb{T}} \int_S d_W(p^t(y,dz),\pi(dz))p^s(x,dy).$$

Main results

Observe that for any $t \in \mathbb{T}$ and $x \in S$,

$$d_W(p^t(x,dy),\pi(dy)) \le \sup_{s\in\mathbb{T}} \int_S d_W(p^t(y,dz),\pi(dz))p^s(x,dy).$$

Thus, it is tempting to conclude that (1) might be replaced by

$$\lim_{t\to\infty} d_W(p^t(x,dy),\pi(dy)) = 0, \qquad x \in S. \tag{3}$$

Main results

Observe that for any $t \in \mathbb{T}$ and $x \in S$,

$$d_W(p^t(x,dy),\pi(dy)) \le \sup_{s\in\mathbb{T}} \int_S d_W(p^t(y,dz),\pi(dz))p^s(x,dy).$$

Thus, it is tempting to conclude that (1) might be replaced by

$$\lim_{t\to\infty} d_W(p^t(x,dy),\pi(dy)) = 0, \qquad x \in S. \tag{3}$$

Theorem

Assume that there is $\pi \in \mathcal{P}(S)$ satisfying (3), and for every $f \in Lip(S)$ and $t \in \mathbb{T}$ the function

$$F_{f,t}(x) := \int_S f(y)p^t(x,dy), \qquad x \in S,$$

is also in $Lip(S)$ with $|F_{f,t}|_{Lip} \le C_f$, where the constant C_f depends only on $f(x)$. Then, for any $f \in Lip(S)$ and $\mu \in \mathcal{P}(S)$, **M** satisfies the relation in (2).

Main result

As a direct consequence of the contraction property we conclude that for every $f \in Lip(S)$ and $t \in \mathbb{T}$,

$$F_{f,t} \in Lip(S) \qquad \text{and} \qquad |F_{f,t}|_{Lip} \le |f|_{Lip}.$$

Main result

As a direct consequence of the contraction property we conclude that for every $f \in Lip(S)$ and $t \in \mathbb{T}$,

$$F_{f,t} \in Lip(S) \qquad \text{and} \qquad |F_{f,t}|_{Lip} \leq |f|_{Lip}.$$

Corollary

Assume that **M** is contractive with respect ot d_W, and there is $\pi \in \mathcal{P}(S)$ satisfying (3). Then, for any $f \in Lip(S)$ and $\mu \in \mathcal{P}(S)$, **M** satisfies the relation in (2).

Examples

Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. random variables satisfying $\mathbb{P}(X_n = 0) = \mathbb{P}(X_n = 1/2) = 1/2$. Define

$$M_{n+1} := \frac{1}{2}M_n + X_{n+1}, \qquad n \geq 0,\ M_0 \in [0,1].$$

Clearly, $\mathbf{M}$ is a Markov process with state space $([0,1], \mathcal{B}([0,1]))$ and transition function $p(x, dy) := \mathbb{P}(X_1 + x/2 \in dy),\ x \in [0,1]$. Also, it is easy to see that $\mathrm{Leb}(dy)$ is invariant for $\mathbf{M}$ and $\mathbf{M}$ is ergodic with respect to $\mathrm{Leb}(dy)$. However, observe that, since $\mathrm{Leb}(dy)$ is singular with respect to $p(x, dy)$, $\mathbf{M}$ is not strongly ergodic.

Examples

Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. random variables satisfying $\mathbb{P}(X_n = 0) = \mathbb{P}(X_n = 1/2) = 1/2$. Define

$$M_{n+1} := \frac{1}{2}M_n + X_{n+1}, \qquad n \geq 0,\ M_0 \in [0,1].$$

Clearly, $\mathbf{M}$ is a Markov process with state space $([0,1], \mathcal{B}([0,1]))$ and transition function $p(x,dy) := \mathbb{P}(X_1 + x/2 \in dy)$, $x \in [0,1]$. Also, it is easy to see that $\mathrm{Leb}(dy)$ is invariant for $\mathbf{M}$ and $\mathbf{M}$ is ergodic with respect to $\mathrm{Leb}(dy)$. However, observe that, since $\mathrm{Leb}(dy)$ is singular with respect to $p(x,dy)$, $\mathbf{M}$ is not strongly ergodic. Let $d(x,y) := |x-y|$, $x, y \in [0,1]$. Now, by a straightforward computation we get that for all $n \geq 1$ and $x_1, x_2 \in [0,1]$,

$$d_W(p^n(x_1,dy), p^n(x_2,dy)) \leq \frac{d(x_1,x_2)}{2^n}.$$

Examples

Let $\mathcal{C} := C([-1,0],\mathbb{R})$. For $t \geq 0$ and a function $f(s)$ defined on $[t-1,t]$, we write $f^t(s) := f(s+t)$, $s \in [-1,0]$. Consider the following stochastic functional differential equation (the so-called stochastic delay equation)

$$dM_t = b(M^t)dt + \sigma(M^t)dB_t, \qquad t \geq 0,\ M^0 \in \mathcal{C}, \tag{4}$$

where $b : \mathcal{C} \to \mathbb{R}$ and $\sigma : \mathcal{C} \to \mathbb{R}$. Now, by taking $b(f^t) = -f^t(0) = -f(t)$ and $\sigma(f^t) = g(f^t(-1)) = g(f(t-1))$, where $g(u)$ is bounded, Lipschitz continuous, strictly positive and strictly increasing, Hairer, Mattingly and Scheutzow 2011 have shown that

(i) the equation in (4) admits a unique strong solution $\mathbf{M} := \{M^t\}_{t\geq 0}$ which is a strong Markov and Feller process with state space $(\mathcal{C},\mathcal{B}(\mathcal{C}))$;

Examples

(ii) **M** possesses a unique invariant probability measure $\pi(dy)$ which is singular with respect to the transition function $p^t(x, dy)$ of **M** (hence, **M** is not strongly ergodic);

(iii) there are $\delta > 0$, $c > 0$ and Borel function $C : \mathcal{C} \longrightarrow [0, \infty)$, such that **M** is contractive with respect to d_W,

$$d_W(p^t(x, dy), \pi(dy)) \leq C(x)e^{-ct}$$

and

$$\sup_{s \geq 0} P_s C(x) < \infty, \quad t \geq 0,\ x \in \mathcal{C},$$

where $d(x, y) := 1 \wedge |x - y|_\infty/\delta$, $x, y \in \mathcal{C}$.

Thank you for your attention!