⟨Research Paper⟩

Journal of the Korean
Statistical Society
Vol. 18, No.1, 1989

The Existence of a Unique Invariant Probability Measure for a Markov Process $X_{n+1}= f(X_n)+ \varepsilon_{n+1}$ +

Oesook Lee*

ABSTRACT

We consider a Markov process $\{X_n\}$ on $[0,\infty)^k$ which is generated by $X_{n+1}= f(X_n)+ \varepsilon_{n+1}$ where f is a continuous, nondecreasing concave function. Sufficient conditions for the existence of a unique invariant probability measure for $\{X_n\}$ are obtained.

1. Introduction

Dubins and Freedman(1966) have given necessary and sufficient conditions for the existence of a unique invariant probability measure for a Markov process generated by continuous nondecreasing functions on $[0,1]$.

Yahav (1976) has removed the restriction of compactness on the state space.

In this paper, we consider the Markov process which is generated by a stochastic difference equation.

$$X_{n+1}= f(X_n)+ \varepsilon_{n+1}, \quad n\geqq 0 \tag{1.1}$$

where $\varepsilon_n(n\geqq 1)$ is a sequence of independent, identically distributed(i.i.d.) random vectors on $S=[0,\infty)^k$, $K\geqq 1$ with common distribution P, $f= (f^{(1)}, f^{(2)},\cdots,f^{(k)})'$ is a continuous nondecreasing function on S into S such that each $f^{(i)}$ is concave and has first partial derivatives. X_0 can be taken arbitrary but independent of $\varepsilon_n(n\geqq 1)$. Also, assume S is equipped

+ This research was supported by KOSEF. Grant 883-0105-002-1.

* Department of Statistics, Ewha Womans University.

with the Euclidean norm $|\cdot|$, and $E|\varepsilon_1|^2<\infty$.

Sufficient conditions for the existence of a unique invariant probability are obtaind for such Markov process, extending earlier results of Yahav to multidimensional state space.

2. Notations and Lemmas

Let
$$f_{\varepsilon_n}(y)=f(y)+\varepsilon_n$$
$$g(y)=E[f_{\varepsilon_1}(y)],\ g^n(y)=g(g^{n-1}(y)),\ n\geqq 2.$$

Then we may express $X_n(y)$, which is the process generated by (1.1) whose initial distribution, distribution of X_0, is concentrated in y as

$$X_n(y)=f_{\varepsilon_n}(f_{\varepsilon_{n-1}}(\cdots(f_{\varepsilon_1}(y)))).$$

Clearly $g(y)$ is continuous, nondecreasing, and concave for each coordinate.

Here for $a,b\in S$, $a\leqq b$ means $a^{(i)}\leqq b^{(i)}$ for $1\leqq i\leqq k$, where $a^{(i)}$ is the i^{th} coordinate of a, and $a<b$ if $a\leqq b$ but $a\neq b$.

We make the following assumptions

(A-1) There exists $y_0>0$ such that

(i) $g(y_0)=y_0$, (ii) for $y<y_0$, $g(y)>y$, (iii) for $y>y_0$, $g(y)<y$.

(A-2) The eigenvalues of A are all less than one in magnitude, where

$$A=\left(\frac{\partial f^{(i)}}{\partial y^{(j)}}(y_0)\right),\qquad 1\leqq i,j\leqq k.$$

Lemma 1. $P(X_n(0)\leqq x)$ is nonincreasing in n and hence converges.

Proof. Let $\tilde{X}_n(0)=f_{\varepsilon_1}f_{\varepsilon_2}(\cdots(f_{\varepsilon_n}(0)))$. Then $\tilde{X}_n(0)$ is nondecreasing in n. Since $\tilde{X}_n(0)$ and $X_n(0)$ have the same distribution, the lemma follows. ∎

Lemma 2. Under the assumption (A–1), for every $y\geqq y_0$, $X_n(y)$ converges with probability 1 as $n\to\infty$.

Proof. Let $\tilde{X}_n^{(i)}(y)=f_{\varepsilon_1}^{(i)}(f_{\varepsilon_2}(\cdots(f_{\varepsilon_n}(y)))$ and let $\mathcal{B}_n$ be the σ-field generated by $\varepsilon_1,\varepsilon_2,$

$\cdots\varepsilon_n$. Then Jensen's inequality shows that

$$\begin{aligned} E[\widetilde{X}_{n+1}^{(i)}(y) \mid \beta_n] &\leq f_{\varepsilon_1}^{(i)}(f_{\varepsilon_2}(\cdots(E(f_{\varepsilon_{n+1}}(y))) \\ &= f_{\varepsilon_1}^{(i)}(f_{\varepsilon_2}(\cdots(f_{\varepsilon_n}(g(y))) \quad \text{a.s.} \end{aligned}$$

Since $f_{\varepsilon_1}^{(i)}f_{\varepsilon_2}\cdots f_{\varepsilon_n}$ is nondecreasing, by (A-1)(iii) for any $y \leqq y_0$,

$$E[\widetilde{X}_{n+1}^{(i)}(y) \mid \beta_n] \leq \widehat{X}_n^{(i)}(y) \quad \text{a.s.}$$

Hence $\widetilde{X}_{n+1}^{(i)}(y)$ $(y \geqq y_0)$ is a nonnegative supermartingale with $E[\widetilde{X}_n^{(i)}(y)] \leq y^{(i)}$ for all n, which implies that for $y \geqq y_0$, $\widetilde{X}_n^{(i)}(y)$ converges a.s.. Hence the conclusion follows from the fact that $\widetilde{X}_n(y)$ and $X_n(y)$ have the same distribution. ■

Lemma 3. Let the assumption (A-1) hold and let $\varepsilon_1^{(i)}$ be not concentrated in one point for all $1 \leqq i \leqq k$. Then for $\delta > 0$ sufficiently small,

$$P[X_n(0) \geqq y_0 + \delta \text{ for some } n] = 1.$$

Proof. Note that

$$P[f_{\varepsilon_i}(y) \geqq g(y)] > 0 \text{ for all } i \geqq 1 \text{ and } y \in S.$$

Hence we have,

$$\begin{aligned} P[X_n(0) \geqq g^n(0)] &\geqq P[X_1(0) \geqq g(0), X_2(0) \geqq g^2(0), \cdots, X_n(0) \geqq g^n(0)] \\ &\geqq P[X_1(0) \geqq g(0)]\, p\,[X_2(0) \geqq g^2(0) \mid X_1(0) \geqq g(0)] \\ &\quad P[X_3(0) \geqq g^3(0) \mid X_1(0) \geqq g(0), X_2(0) \geqq g^2(0)] \\ &\quad \cdots\cdots\cdots\cdots\cdots\cdots \\ &\quad P[X_n(0) \geqq g^n(0) \mid X_1(0) \geqq g(0), \cdots, X_{n-1}(0) \geqq g^{n-1}(0)] \\ &\geqq P[f_{\varepsilon_1}(0) \geqq g(0)]\, P[f_{\varepsilon_2}(g(0) \geqq g^2(0)] \cdots\cdots \\ &\quad \cdots P[f_{\varepsilon_n}(g^{n-1}(0)) \geqq g^n(0)] > 0. \end{aligned}$$

Since $g^n(0) \uparrow y_0$, for any $\varepsilon > 0$, there exists $n_1(\varepsilon)$ such that $P[X_{n_1}(0) \geqq y_0 - \varepsilon] > 0$ and by lemma 1, we have for any $n \geqq n_1(\varepsilon)$,

$$P\,[X_n(0)\geqq y_0-\varepsilon]>0. \qquad (2.1)$$

Under the assumption $\varepsilon_1^{(i)}$ is not concentrated in one point, we have

$$P\,[f_{\varepsilon_1}^{(i)}(y_0)>y_0^{(i)}]>0 \qquad \text{for } 1\leqq i\leqq k. \qquad (2.2)$$

Since $f_{\varepsilon_1}(\cdot)$ is continuous at y_0,(2.1) with sufficiently small $\varepsilon>0$ and (2.2) imply that, for sufficiently small $\delta>0$, there exists n_2 such that

$$P\,[X_{n_2}(0)\geqq y_0+\delta]>0. \qquad (2.3)$$

Now,

$$\begin{aligned}
&P\,[X_n(0)<y_0+\delta \text{ for all } n]\\
\leqq\ &P\,[X_{mn_2}(0)<y_0+\delta \text{ for every } m=1,2,3,\cdots,k]\\
\leqq\ &P\,[X_{n_2}(0)<y_0+\delta]\cdot P\,[X_{2n_2}(0)<y_0+\delta\mid X_{n_2}(0)<y_0+\delta]\\
&\cdots\cdots P\,[X_{kn_2}(0)<y_0+\delta\mid X_{jn_2}(0)<y_0+\delta,\ \text{for } j=1,2,\cdots,k-1]\\
\leqq\ &\{P\,[X_{n_2}(0)<y_0+\delta]\}^k.
\end{aligned}$$

The above inequalities hold for all $k\geqq 1$ and therefore the required result $P\,[X_n(0)<y_0+\delta \text{ for all } n]=0$ follows from (2.3). ■

Let $\rho(S)$ be the set of all probability measures on S and let $p^{(n)}(y,B)=P\,[X_n(y)\in B]$, $B\in\beta(S)$, $n=1,2,3,\cdots$ where $\beta(S)$ denotes the Borel σ-field of S. On $\rho(S)$, define the bounded Lipschitzian distance of Dudley (1966):

$$\|\mu-\nu\|_{BL}=\sup\{\,|\textstyle\int fd\mu-\int fd\nu|\ :f\in BL\}\quad(\mu,\nu\in\rho(S))$$

where $BL=\{f:f:S\to R \text{ such that } |f(x)-f(y)|\leqq 1\ \forall x,y \text{ and } |f(x)-f(y)|\leqq|x-y|\ \forall x,y\}$. It is known that $\|\cdot\|_{BL}$ metrizes the weak* topology on $\rho(S)$.

Lemma 4. Let $X_{n+1}=AX_n+\varepsilon_{n+1}$ where ε_n is i.i.d. with taking values in S and $E|\varepsilon_1|^2<\infty$, and A $k\times k$ matrix whose eigenvalue λ all satisfy $|\lambda|<1$. Then there exists a unique invariant distribution for $\{X_n\}$ and for any y in S

$$\lim_{n\to\infty} E\,[X_n(y)]=\Big(\sum_{n=0}^{\infty}A^n\Big)E(\varepsilon_1)=(I-A)^{-1}E(\varepsilon_1).$$

Proof. Let C be any bounded set in S. Then one has for all $y_1, y_2 \in C$,

$$\| p^{(n)}(y_1,dz) - p^{(n)}(y_2,dz) \|_{BL} = \sup\{ | Ef(X_n(y_1)) - Ef(X_n(y_2)) | : f \in BL\}$$
$$\leq E[| X_n(y_1) - X_n(y_2) | \wedge 1]$$
$$\leq \| A^n \| \cdot \text{diam } C \to 0 \text{ as } n \to \infty. \tag{2.4}$$

Similarly for all n,m,

$$\| p^{(n+m)}(0,dz) - p^{(n)}(0,dz) \|_{BL} \leq E[| \widetilde{X}_{n+m}(0) - \widetilde{X}_n(0) | \wedge 1]$$
$$= E[| A^n(A^{m-1}\varepsilon_{n+m} + \cdots + \varepsilon_{n+1}) | \wedge 1]$$
$$\leq P[| A^{m-1}\varepsilon_{n+m} + \cdots + \varepsilon_{n+1} | > M]$$
$$+ P[\text{diam } A^n(\overline{B(0,M)}) > \delta] + \delta$$

where $\widetilde{X}_n(0) = f_{\varepsilon_1}(f_{\varepsilon_2}(\cdots(f_{\varepsilon_n}(0))))$ and $\overline{B(0,M)} = \{X \in S : |X| \leq M\}$.

Given $\varepsilon > 0$, let $\delta = \varepsilon / 3$. Then by Chebyshev's inequality, we may choose $M = M_\varepsilon$ such that

$$P[| A^{m-1}\varepsilon_{n+m} + A^{m-2}\varepsilon_{n+m-1} + \cdots + \varepsilon_{n+1} | > M_\varepsilon] < \varepsilon / 3, \; m = 1,2,3,\cdots$$

Since diam $A^n(\overline{B(0,M_\varepsilon)}) \leq \| A^n \| \cdot M_\varepsilon$, $P[\text{diam } A^n(\overline{B(0,M_\varepsilon)}) > \varepsilon / 3] \to 0$ as $n \to \infty$.

Hence for all sufficiently large n,

$$\| p^{(n+m)}(0,dz) - p^{(n)}(0,dz) \|_{BL} < \varepsilon, \; m = 1,2,3,\cdots$$

Since $(\rho(s), \| \cdot \|_{BL})$ is a complete metric space, it follows that there exists a probability measure, say π such that

$$\| p^{(n)}(0,dz) - \pi(dz) \|_{BL} \to 0 \quad \text{as } n \to \infty. \tag{2.5}$$

(2.4) and (2.5) imply the uniform convergence on C to π. Since, in this case, $\int f(z) p^{(1)}(y,dz)$ is a bounded continuous function whenever f is bounded continuous, the weak convergence of $p^{(n)}(y,dz)$ to $\pi(dz)$ for all y implies that π is the unique invariant probability. Hence the proof for the first part of lemma 4 is completed.

Now consider

$$E[X_n(y)] = A^n y + E[\sum_{j=1}^{n} A^{n-j}\varepsilon_j] = A^n y + (\sum_{j=0}^{n-1} A^j) E(\varepsilon_1).$$

But $\lim_{n\to\infty} E[X_n(y)]=(\sum_{j=0}^{\infty} A^j) E(\varepsilon_1)$, since $A^n \to 0$ as $n\to\infty$.

The proof follows from the fact that $\sum_{j=0}^{\infty} A^j = (I-A)^{-1}$ if eigenvalue λ of A all satisfy $|\lambda|<1$. ■

From the concavity of $f^{(i)}$, we have

$$f^{(i)}(y) \leqq \nabla f^{(i)}(y_0)(y-y_0) + f^{(i)}(y_0)$$

where $\nabla f^{(i)}(y_0)$ is the gradient of f (Roberts and Vererg(1973)), and hence we can write

$$f(y) \leqq A(y-y_0) + f(y_0)$$

where A is a $k\times k$ matrix given in (A-2).

Define $\hat{f}(y)= A(y-y_0) + f(y_0)$. Then $f(y)\leqq\hat{f}(y)$ and $\hat{f}(y_0)= f(y_0)$.

Now let $\{X_n^*(y) : n\geqq 1\}$ be the Markov process generated by

$$X_{n+1}^* = \hat{f}(X_n^*) + \varepsilon_{n+1} = A(X_n^* - y_0) + f(y_0) + \varepsilon_{n+1}$$

with $X_0^* = y$.

Lemma 5. Under the assumptions (A-2) and $g(y_0) = y_0$, $\{X_n^*(y): n\geqq 1\}$ is a Markov process with a unique limiting stationary distribution and $\lim_{n\to\infty} E[X_n^*(y)] = y_0$ for all $y\in S$.

Proof. Let $\varepsilon_n^* = f(y_0) - Ay_0 + \varepsilon_n$. Then

$$X_n^*(y) = A^n y + \sum_{j=1}^{n} A^{n-j}\varepsilon_j^*.$$

By assumption (A−2) and lemma 4, we have

$$\lim_{n\to\infty} E[X_n^*(y)] = (I-A)^{-1}E(\varepsilon_1^*).$$

But $E(\varepsilon_1^*) = f(y_0) + E(\varepsilon_1) - Ay_0 = y_0 - Ay_0 = (I-A)y_0$. Hence $\lim_{n\to\infty} E[X_n^*(y)] = y_0$. ■

Lemma 6. Under the assumptions (A−1) and (A−2), we have for every $\delta>0$

$$P[X_n(y) \leqq y_0 + \delta \text{ i.o.}] = 1.$$

Proof. First show that $P[X_n^*(y) \leqq y_0 + \delta \text{ i.o.}] = 1$.
Recall that for any y in S, $\lim_{n\to\infty} E[X_n^*(y)] = y_0$.
Suppose $P[X_n^*(y) \leqq y_0+\delta] = 0$. Then

$$E(X_n^*(y)) = \int_{(X_n^*(y)>y_0+\delta)} X_n^*(y)\,dP + \int_{(X_n^*(y)\leq y_0+\delta)} X_n^*(y)\,dP$$
$$> y_0+\delta$$

which is not true for sufficiently large n. Hence

$$P[X_n^*(y) \leqq y_0+\delta] > 0 \tag{2.6}$$

for sufficiently large n. Since $\{X_n^*(y) \leqq y_0+\delta \text{ i.o.}\}$ is an invariant tail event and $X_n^*(y)$ is a Markov process with a unique limiting stationary distribution, (2.6) implies $P[X_n^*(y) \leqq y_0+\delta \text{ i.o.}] = 1$. The proof follows from the relation $X_n(y) \leqq X_n^*(y)$ a.s. ■

3. Main Theorem

Theorem. Let the assumptions (A−1) and (A-2) hold. If $\varepsilon_1^{(i)}$ is not concentrated in one point for all $1 \leqq i \leqq k$, then the process $\{X_n(y): n \geqq 1\}$ has a unique invariant probability measure.

Proof. Let $H_y^n(x) = P[X_n(y) \leqq x]$

$$F_y(x) = \lim_{n\to\infty} H_y^n(x).$$

Let $\delta>0$ be fixed.

Define $\tau_1=\inf\{n : X_n(0)\geq y_0+\delta\}$. Then by lemma 3, $P(\tau_1<\infty) = 1$.

Now we can write

$$\begin{aligned} P[X_n(0)\leq x,\ \tau_1\leq n] &= \sum_{i=1}^{n} P[X_n(0)\leq x,\ \tau_1=i] \\ &= \sum_{i=1}^{n} P[X_n(0)\leq x \quad \tau_1=i]\cdot P(\tau_1=i) \\ &= \sum_{i=1}^{n} \left\{ \int_{y\geq y_0+\delta} H^{n-i}_y(x)\, dQ_1(y \mid \tau_1=i)\right\}\cdot P(\tau_1=i) \\ &\leq \sum_{i=1}^{n} H^{n-i}_{y_0+\delta}(x)\cdot P(\tau_1=i) \end{aligned} \tag{3.1}$$

where $Q_1(y \mid \tau_1=i)=P[X_{\tau_1}(0)\leq y \mid \tau_1=i]$ and

$$Q_1(y \mid \tau_1=i)=0 \text{ whenever } P(\tau_1=i)=0.$$

The last inequality follows from $H^{n-i}_y(x)\leq H^{n-i}_{y_0+\delta}(x)$ for all $y\geq y_0+\delta$. By lemma 1 and lemma 2, the limits of the first and last term of (3.1) exist, and hence we have

$$F_0(x)\leq F_{y_0+\delta}(x).$$

But since $H_0^n(x)\geq H^n_{y_0+\delta}(x)$ for all n, we get

$$F_0(x)=F_{y_0+\delta}(x). \tag{3.2}$$

Now for any y, $0\leq y\leq y_0+\delta$,

$$H_0^n(x)\geq H_y^n(x)\geq H^n_{y_0+\delta}(x).$$

Taking limits on $H_0^n(x)$, $H^n_{y_0+\delta}(x)$, and using the equation (3.2) we have for all y, $0\leq y\leq y_0+\delta$,

$$F_0(x)=F_y(x)$$

Let $y_1>y_0+\delta$ and define $\tau_2=\inf\{n : X_n(y_1)\leq y_0+\delta\}$. Then by lemma 6, $P(\tau_2<\infty)=1$.

Now, consider

$$P[X_n(y_1) \le x,\ \tau_2 \le n] = \sum_{j=1}^{n} P[X_n(y_1) \le x,\ \tau_2 = j]$$
$$= \sum_{j=1}^{n} \{ \int_0^{y_1 < y_0+\delta} H_y^{n-1}(x) dQ_2(y_1 \mid \tau_2 = j) \} \cdot P(\tau_2 = j)$$
$$\ge \sum_{j=1}^{n} H_{y_0+\delta}^{n-1}(x) \cdot P(\tau_2 = j)$$

where $Q_2(y_1 \mid \tau_2 = j) = P[X_{\tau_2}(y) \le x \mid \tau_2 = j)$

By lemma 2, we have $F_{y_1}(x) \geqq F_{y_0+\delta}(x)$. But for all n, $H_{y_1}^n(x) \leqq H_{y_0+\delta}^n(x)$,

and hence for any y, $y \geqq y_0 + \delta$,

$$F_y(x) = F_{y_0+\delta}(x) =$$

Hence for any $y \in S$,

$$F_y(x) = F_0(x). \tag{3.3}$$

To prove the stationarity of $F_0(x)$, consider

$$H_0^{n+1}(x) = \int_S H_y^1(x)\, dH_0^n(y).$$

Since $f(\cdot)$ is continuous, $H_y^1(x)$ is continuous and uniformly bounded, and hence, by the Helly−Bray lemma, we get

$$F_0(x) = \int_S H_y^1(x)\, dF_0(y). \tag{3.4}$$

Now let π be the probability measure on S corresponding to $F_0(x)$, that is

$$\pi([0,x]) = F_0(x).$$

Then (3.3) says for any x in S, the n−step transition probability function $p^{(n)}(y,[0,x]) = H_y^n(x)$ converges to $\pi([0,x])$ for all y.

We may rewrite (3.4) as

$$\pi([0,x]) = \int_S p(y, [0,x])\, \pi(dy), \quad x \in S$$

which implies that π is an invariant probability for $\{X_n\}$.

To prove the uniqueness of π, let ν be another invariant probability measure for $\{X_n\}$. Then

$$\nu([0,x])= \int_S p^{(n)} (y, [0,x])\, \nu(dy), \text{ for all n.}$$

Taking limits in above equation, we have

$$\nu([0,x])= \int_S \pi([0,x])\, \nu(dy) =\pi([o,x]), \quad x \in s$$

which implies $\nu=\pi$.

Hence we complete the proof. ■

Acknowledgement : Author is greatful to the refree for a careful reading of manuscript and for numerous helpful comments and suggestions.

References

(1) P. Billingsley (1979). *Probability and measure* JohnWiley, New York.

(2) L.E. Dubins and D.A. Freedman(1966). Invariant probalilities for certain Markov processes. *Annals of Mathematical Statistics,* Vol. 37, 837−847.

(3) R.M. Dudley (1966). Convergence of Baire measures. *Studia Mathematica.* 27, 251−268.

(4) A.W. Roberts and D.E. Varberg (1973). *Convex functions.* Academic press, New York.

(5) J.A. Yahav (1976). On a Markov process generated by nondecreasing concave functions. *Stochastic processes and their applications* 4, 41−54.