Bernoulli **3**(4), 1997, 415–427

Stationary distribution of Markov chains in $\mathbb{R}^d$ with application to global random optimization

CHANG CHUNG YU DOREA

Departamento de Matemática, Universidade e Brasília, 70910-900 Brasília-DF, Brazil. e-mail: cdorea@mat.unb.br

Stationary and long-run distributions of a class of Markov chains with continuous state space $S \subset \mathbb{R}^d$ are studied. Applications to random search algorithms for global optimization are presented.

Keywords: continuous state space; global random optimization; stationary distribution

1. Introduction

We consider a Markov chain $\{X_n\}_{n\geqslant 0}$ that takes values on a bounded and measurable subset $S \subset \mathbb{R}^d$ with $m(S) > 0$ and whose transition probability function has the representation:

$$P(X_n \in B|X_0 = x) = P^{(n)}(B|x) = P_{\mathrm{c}}^{(n)}(B|x) + P_{\mathrm{d}}^{(n)}(B|x) \tag{1}$$

for $x \in S$ and $B \in \mathscr{B}^d$ (Borel subsets of $\mathbb{R}^d$ with Lebesgue measure). The absolutely continuous part,

$$P_{\mathrm{c}}^{(n)}(B|x) = \int_B p_{\mathrm{c}}^{(n)}(y|x)\,\mathrm{d}y \tag{2}$$

is such that $p_{\mathrm{c}}^{(n)}(y|x) = 0$ if $y \notin S$; similarly, the discrete part

$$P_{\mathrm{d}}^{(n)}(B|x) = \sum_{y\in B\cap S_x^{(n)}} p_{\mathrm{d}}^{(n)}(y|x) \tag{3}$$

is such that $p_{\mathrm{d}}^{(n)}(y|x) = 0$ for $y \notin S_x^{(n)}$ and $S_x^{(n)} \subset S$ for each $x \in S$ and $n \geqslant 1$.

Chains with continuous state space and the above representation arise naturally in random global optimization algorithms. A general scheme for random search algorithms can be described as follows: let $X_0 \in S$ be an initial random point and let $f\colon S \to \mathbb{R}$ be a function whose global optimum on S (maximum or minimum) is of interest; for each $x \in S$ let $g(y|x)$ denote a density function on $\mathbb{R}^d$; if X_n is the result of the algorithm at step n then at step $n+1$ one generates a random value Y_n according to the density g; next X_{n+1} is taken to be Y_n with an acceptance probability $a(Y_n|X_n)$ or $X_{n+1} = X_n$ with probability $1 - a(Y_n|X_n)$. It follows that the transition function of the algorithm can be written as:

1350–7265

$$P(B|x) = \int_B a(y|x)g(y|x)\,\mathrm{d}y + I_{(x\in B)}\left[1 - \int_S a(y|x)g(y|x)\,\mathrm{d}y\right] \tag{4}$$

where I denotes the indicator function. This type of Markovian algorithm will be detailed in Section 3.

Though there is an extensive literature on general state-space Markov processes, considerable mathematical background is required to understand them. In this paper we treat Markov chains in $\mathbb{R}^d$ by extending the notions of communicability and periodicity of states to subsets of $\mathbb{R}^d$. It will be shown that, as in the discrete state space, they play an important role in analysing the existence of stationary and long-run distributions,

$$\lim_{n\to\infty} P^{(n)}(B|x) = Q(B). \tag{5}$$

In Section 2 we derive conditions that will guarantee (5) and in Section 3 we apply these notions to random search algorithms studied by Dorea (1986; 1990) and to a continuous version of the 'simulated annealing' algorithm treated by Dekkers and Aarts (1991).

2. Continuous state space

Let $\{X_n\}_{n\geqslant 0}$ be a Markov process with values on S. Assuming that the nth-step transition function has the representation (1), we then have:

$$\begin{aligned} P^{(n+1)}(B|x) &= \int_S P^{(n)}(B|y)P(\mathrm{d}y|x) \\ &= \int_S P^{(n)}(B|y)p_\mathrm{c}(y|x)\,\mathrm{d}y + \sum_{y\in S_x} P^{(n)}(B|y)p_\mathrm{d}(y|x), \end{aligned}$$

where $S_x = \{y\colon y\in S,\ p_\mathrm{d}(y|x) > 0\}$. Using (2) and (3), we can write

$$\begin{aligned} P^{(n+1)}(B|x) = &\int_B \left[\int_S p_\mathrm{c}^{(n)}(z|y)p_\mathrm{c}(y|x)\,\mathrm{d}y + \sum_{y\in S_x} p_\mathrm{c}^{(n)}(z|y)p_\mathrm{d}(y|x)\right]\mathrm{d}z \\ &+ \int_S \left[\sum_{z\in B\cap S_y^{(n)}} p_\mathrm{d}^{(n)}(z|y)p_\mathrm{c}(y|x)\right]\mathrm{d}y + \sum_{y\in S_x}\sum_{z\in B\cap S_y^{(n)}} p_\mathrm{d}^{(n)}(z|y)p_\mathrm{d}(y|x). \end{aligned}$$

It follows that

$$p_\mathrm{c}^{(n+1)}(z|x) \geqslant \int_S p_\mathrm{c}^{(n)}(z|y)p_\mathrm{c}(y|x)\,\mathrm{d}y + \sum_{y\in S_x} p_\mathrm{c}^{(n)}(z|y)p_\mathrm{d}(y|x) \tag{6}$$

and

$$p_\mathrm{d}^{(n+1)}(z|x) \geqslant \sum_{y\in S_x} p_\mathrm{d}^{(n)}(z|y)p_\mathrm{d}(y|x). \tag{7}$$

By expressing $P^{(n+1)}(B|x) = \int_S P(B|y)P^{(n)}(\mathrm{d}y|x)$ we also obtain the inequalities

$$p_{\mathrm{c}}^{(n+1)}(z|x) \geqslant \int_S p_{\mathrm{c}}(z|y)p_{\mathrm{c}}^{(n)}(y|x)\,\mathrm{d}y + \sum_{y \in S_x^{(n)}} p_{\mathrm{c}}(z|y)p_{\mathrm{d}}^{(n)}(y|x)$$

and

$$p_{\mathrm{d}}^{(n+1)}(z|x) \geqslant \sum_{y \in S_x^{(n)}} p_{\mathrm{d}}(z|y)p_{\mathrm{d}}^{(n)}(y|x).$$

Our main result (Theorem 1 below) shows that an irreducible and aperiodic chain possesses a long-run distribution if certain regularity conditions are met. And in this case it coincides with the unique stationary distribution. The following notation and notions will be needed:

$$p^{(n)}(y|x) = p_{\mathrm{c}}^{(n)}(y|x) + p_{\mathrm{d}}^{(n)}(y|x). \tag{8}$$

For $B \in \mathscr{B}^d$ let

$$B^+ = \{D\colon D \subset B,\ m(D) > 0\}, \qquad \text{if } m(B) > 0 \tag{9}$$

and

$$B^+ = \{D\colon D \neq \varnothing,\ D \subset B\} \qquad \text{if } B \text{ is countable.} \tag{10}$$

Let

$$S^1 = \{D\colon D \subset S,\ \text{either } m(D) > 0 \text{ or } D \text{ is countable}\}. \tag{11}$$

Definition 1. *Let* $\{A, B\} \subset S^1$. *We say that B is accessible from A if there exists* $n_B \geqslant 1$ *such that*

$$P^{(n_B)}(B'|x) > 0, \qquad \forall x \in A, \forall B' \in B^+$$

(which we write as $A \stackrel{n_B}{\rightarrow} B$*). We say that A and B are communicating subsets* ($A \leftrightarrow B$) *if* $A \stackrel{n_B}{\rightarrow} B$ *and* $B \stackrel{n_A}{\rightarrow} A$.

Definition 2. *We say that a chain is irreducible if there exists* $\{A_1, \ldots, A_k\} \subset S^1$ *such that* $\bigcup_{j=1}^k A_j = S$ *and* $A_i \leftrightarrow A_j$ *for all i and j.*

Definition 3. *Assume* $A \leftrightarrow A$. *We say that* d_A *is the period of A if* d_A *is the greatest common divisor of*

$$\mathscr{D}(A) = \{n\colon P^{(n)}(A'|x) > 0, \forall x \in A, \forall A' \in A^+\}.$$

If $d_A = 1$ *we say that the subset A is aperiodic.*

Some immediate properties can be derived:

Proposition 1. (a) *If* $A \stackrel{n_B}{\to} B$ *then, given* $\epsilon_0 > 0$, *there exists* $A_0 \in A^+$ *such that either*

$$P_{\mathrm{c}}^{(n_B)}(B|x) \geqslant \epsilon_0, \forall x \in A_0 \text{ or } P_{\mathrm{d}}^{(n_B)}(B|x) \geqslant \epsilon_0, \forall x \in A_0, \tag{12}$$

depending on whether $m(B) > 0$ *or* B *is countable.*

(b) *If* $A \stackrel{n_B}{\to} B$ *and* $B \stackrel{n_C}{\to} C$ *then* $A \stackrel{n_B+n_C}{\to} C$.

(c) *If* $A \leftrightarrow B$ *then they have the same period* ($d_A = d_B$).

(d) *If* d_A *is the period of* A *then there exists* $m_A \geqslant 1$ *such that for* $m \geqslant m_A$ *we have* $md_A \in \mathscr{D}(A)$.

Proof. (a) Assume $m(B) > 0$ and let $S_x^{(n_B)} = \{y\colon p_{\mathrm{d}}^{(n_B)}(y|x) > 0\}$. Since $B \backslash S_x^{(n_B)} \in B^+$ and $m(S_x^{(n_B)}) = 0$, $\forall x \in A$, we must have $P_{\mathrm{c}}^{(n_B)}(B'|x) > 0$, $\forall x \in A$ and $\forall B' \in B^+$. Given $\epsilon_0 > 0$, define for $k \geqslant 1$

$$D_k = \left\{x\colon x \in A,\ P_{\mathrm{c}}^{(n_B)}(B|x) < \frac{1}{k}\right\}.$$

Since $D_{k+1} \subset D_k$, $\bigcap_k D_k = \varnothing$ and $m(A) < \infty$ we have $\lim_{k\to\infty} m(D_k) = 0$. Let k_0 be large enough so that $\epsilon_0 \leqslant 1/k_0$ and $m(A \backslash D_{k_0}) > 0$ (if $m(A) > 0$) and $A \backslash D_{k_0} \neq \varnothing$ (if A is countable). Then $A_0 = A \backslash D_{k_0} \in A^+$ and for $x \in A_0$ we have $P_{\mathrm{c}}^{(n_B)}(B|x) \geqslant \epsilon_0$.

If B is countable clearly we have $P^{(n_B)}(B'|x) = P_{\mathrm{d}}^{(n_B)}(B'|x)$, $\forall B' \in B^+$. Now proceeding with arguments of the same type we obtain (12) by observing that $P_{\mathrm{d}}^{(n_B)}(B|x) > 0$, $\forall x \in A$.

(b) Let $C' \in C^+$; then we have $B \stackrel{n_C}{\to} C'$. From (a), given $\epsilon_0 > 0$, there exists $B_0 \in B^+$ such that $P^{(n_C)}(C'|y) \geqslant \epsilon_0$ for $y \in B_0$. Since $B_0 \in B^+$ we also have $A \stackrel{n_B}{\to} B_0$. Now for $x \in A$,

$$P^{n_C+n_B}(C'|x) \geqslant \int_{B_0} P^{(n_C)}(C'|y) P^{(n_B)}(\mathrm{d}y|x) \geqslant \epsilon_0 P^{(n_B)}(B_0|x) > 0.$$

It follows that $A \stackrel{n_C+n_B}{\to} C$.

(c) Let $A \stackrel{n_B}{\to} B$ and $B \stackrel{n_A}{\to} A$. From (b) we have $A \stackrel{n_A+n_B}{\to} A$ so that $n_A + n_B \in \mathscr{D}(A)$. Now let $n_1 \in \mathscr{D}(B)$, we will show that $n_B + n_1 + n_A \in \mathscr{D}(A)$. Thus d_A divides n_1 since it divides $n_A + n_B$. It follows that $d_A \leqslant d_B$. Interchanging the roles of A and B, we conclude $d_A = d_B$.

Let $A' \in A^+$; since $B \stackrel{n_A}{\to} A'$, there exists $B_0 \in B^+$ such that $P^{(n_A)}(A'|z) \geqslant \epsilon_0 > 0$ on B_0. Since $n_1 \in \mathscr{D}(B)$ we have $B \stackrel{n_1}{\to} B_0$, and there exists $B_1 \in B^+$ such that $P^{(n_1)}(B_0|y) \geqslant \epsilon_1 > 0$ on B_1. Now for $x \in A$,

$$\begin{aligned} P^{(n_B+n_1+n_A)}(A'|x) &\geqslant \int_{(y\in B_1)} \int_{(z\in B_0)} P^{(n_A)}(A'|z) P^{(n_1)}(\mathrm{d}z|y) P^{(n_B)}(\mathrm{d}y|x) \\ &\geqslant \epsilon_0 \int_{(y\in B_1)} P^{(n_1)}(B_0|y) P^{(n_B)}(\mathrm{d}y|x) \\ &\geqslant \epsilon_0 \epsilon_1 P^{(n_B)}(B_1|x) > 0. \end{aligned}$$

Thus $n_B + n_1 + n_A \in \mathscr{D}(A)$.

(d) Note that if $r \in \mathscr{D}(A)$ and $s \in \mathscr{D}(A)$ then, using (a), we have $r + s \in \mathscr{D}(A)$. That is, $\mathscr{D}(A)$ is closed under addition. Then there exists m_A such that $md_A \in \mathscr{D}(A)$ for $m \geqslant m_A$; see Doob (1953, p. 176) or Parzen (1962, p. 262). □

Proposition 2. (a) *If* $A \overset{n_B}{\to} B$ *then there exist* $\delta_B > 0$, $A_0 \in A^+$ *and* $B_0 \in B^+$ *such that*

$$\inf\{p^{(n_B)}(y|x)\colon x \in A_0,\ y \in B_0\} \geqslant \delta_B. \tag{13}$$

(b) *If* $A \overset{n_A}{\leftrightarrow} A$ *then there exist* $\delta_A > 0$ *and* $A_0 \in A^+$ *such that*

$$\inf\{p^{(2n_A)}(y|x)\colon x \in A_0,\ y \in A_0\} \geqslant \delta_A. \tag{14}$$

(c) *If the chain is irreducible and aperiodic, then we can decompose* $S = S_c \cup S_d$ *where* S_d *is countable and* $S_c = S \backslash S_d$. *Moreover, there exists* $n_S \geqslant 1$ *such that* $S \overset{n_S}{\to} S_c$, *and if* $S_d \neq \varnothing$ *we also have* $S \overset{n_S}{\to} S_d$.

Proof. (a) First assume B is countable. Let $B_N = \{b_1, \ldots, b_N\} \subset B$. Since $A \overset{n_B}{\to} B_N$, from (12) we can write,

$$P_d^{(n_B)}(B_N|N) \geqslant \delta_0 > 0 \text{ on } A_0 \in A^+. \tag{15}$$

It follows that we must have $p_d^{(n_B)}(y_j|x) \geqslant \delta_0/N$ for some $y_j \in B_N$. Now (13) follows using (8) and taking $B_0 = \{y_j\} \in B^+$ and $\delta_B = \delta_0/N$.

Now assume $m(B) > 0$. If A is countable let $x_0 \in A$ and $A_0 = \{x_0\} \in A^+$. Since $P_c^{(n_B)}(B|x_0) = \int_B p_c^{(n_B)}(y|x_0)\,\mathrm{d}y > 0$, there exist $\delta_B > 0$ and $B_0 \in B^+$ such that on B_0 we have $p_c^{(n_B)}(y|x_0) \geqslant \delta_B$.

If $m(A) > 0$ then from (12),

$$P_c^{(n_B)}(B|x) \geqslant \delta_1 > 0 \text{ on } A_1 \in A^+. \tag{16}$$

Let $\delta_B = \delta_1/2m(B)$ and define

$$D = \{(x, y)\colon x \in A_1,\ y \in B,\ p_c^{(n_B)}(y|x) \geqslant \delta_B\}.$$

Note that for $x \in A_1$ and $D_x = \{y\colon (x, y) \in D\}$ we have $m(D_x) > 0$. If not, then for almost all y in D_x we have $p_c^{(n_B)}(y|x) < \delta_B$ and $P_c^{(n_B)}(B|x) \leqslant \delta_B m(B) \leqslant \delta_1/2$, which contradicts (16). Let m_2 denote the Lebesgue measure on $\mathbb{R}^{2d}$; then we have $m_2(D) = \int_{A_1} m(D_x)\,\mathrm{d}x > 0$. Thus there exists a rectangle $A_0 \times B_0$ with $m_2(A_0 \times B_0) > 0$, $A_0 \in A_1^+ \subset A^+$, $B_0 \in B^+$, and for $x \in A_0$ and $y \in B_0$ we have

$$p^{(n_B)}(y|x) \geqslant p_c^{(n_B)}(y|x) \geqslant \delta_B.$$

(b) From (a) there exist $A_1 \in A^+$ and $A_2 \in A^+$ such that

$$\inf\{p^{(n_A)}(y|x)\colon x \in A_1,\ y \in A_2\} \geqslant \delta_1 > 0.$$

By (12) there exists $A_0 \in A_2^+$ such that

$$P^{(n_A)}(A_1|y) \geqslant \delta_2 > 0 \text{ on } A_0.$$

Now for $x \in A_0$ and $y \in A_0$,

$$p^{(2n_A)}(y|x) \geqslant \int_{A_1} p^{(n_A)}(y|z)P^{(n_A)}(\mathrm{d}z|x) \geqslant \delta_1 P^{(n_A)}(A_1|x) \geqslant \delta_1\delta_2 > 0.$$

(c) Let $\{A_i, \ldots, A_k\}$ satisfy Definition 2. Since the chain is aperiodic, from Proposition 1 there exist $m_1, \ldots, m_k$ such that for $r \geqslant \max\{m_1, \ldots, m_k\}$ we have $r \in \mathscr{D}(A_j)$ for $j = 1, \ldots, k$.

Since $m(S) > 0$ not all A_j's can be countable. Without loss of generality, assume that $A_1, \ldots, A_\ell$ have positive measure ($\ell \leqslant k$). Let $S_\mathrm{c} = \bigcup_{r=1}^{\ell} A_r$ and $B \in S_\mathrm{c}^+$. Then for some j we have $B_j = B \cup A_j \in A_j^+$. For $i = 1, \ldots, k$ let n_i be such that $A_i \xrightarrow{n_i} B_j$. Since $m(B_j) > 0$, from the proof of (a) we have, for $i = 1, \ldots, k$,

$$\inf\{p_\mathrm{c}^{(n_i)}(y|x) \colon x \in A_i',\ y \in B_j'(i)\} \geqslant \delta_i > 0, \tag{17}$$

where $A_i' \in A_i^+$ and $B_j'(i) \in B_j^+$.

Now take n_S large enough so that $n_S - n_i \in \mathscr{D}(A_i)$ for $i = 1, \ldots, k$. From (2) and (17) we have, for $z \in A_i'$,

$$P_\mathrm{c}^{(n_i)}(B_j'(i)|z) \geqslant \delta_i m(B_j'(i)) > 0. \tag{18}$$

Let $\delta_B = \min_{1 \leqslant i \leqslant k}\{\delta_i m(B_j'(i)\}$ and $x \in S$. Then $x \in A_i$ for some i, and we have

$$P^{(n_S)}(B|x) \geqslant P^{(n_S)}(B_j'(i)|x) \geqslant \int_{A_i'} P_\mathrm{c}^{(n_i)}(B_j'(i)|z)P^{(n_S - n_i)}(\mathrm{d}z|x).$$

From (18) and the fact that $n_S - n_i \in \mathscr{D}(A_i)$, we have

$$P^{(n_S)}(B|x) \geqslant \delta_B P^{(n_S - n_i)}(A_i'|x) > 0. \tag{19}$$

Since (19) holds for all $B \in S_\mathrm{c}^+$, we have $S \xrightarrow{n_S} S_\mathrm{c}$.

If $\ell = k$ then the proof is completed by taking $S_\mathrm{d} = \varnothing$. If not, let $S_\mathrm{d} = \bigcup_{r=\ell+1}^{k} A_r$. Let $B \in S_\mathrm{d}^+$ and $\ell + 1 \leqslant j \leqslant k$ such that $B_j = B \cap A_j \in A_j^+$. The proof is exactly the same, except that p_c is replaced by p_d in (17), and P_c and $m(B_j'(i))$ by P_d and $\|B_j'(i)\|$ in (18) (where $\|\cdot\|$ denotes cardinality of the set). And we have $S \xrightarrow{n_S} S_\mathrm{d}$. □

Remark 1. Let $A \xrightarrow{n_B} B$. Then from the proofs of Propositions 1 and 2 we also have the following:

(a) If $m(B) > 0$ then

$$P_\mathrm{c}^{(n_B)}(B'|x) > 0, \forall x \in A, \forall B' \in B^+, \tag{20}$$

$$\inf\{p_\mathrm{c}^{(n_B)}(y|x) \colon x \in A_0,\ y \in B_0\} \geqslant \delta_B > 0, \tag{21}$$

with $A_0 \in A^+$ and $B_0 \in B^+$.

(b) If B is countable then (20) and (21) hold with P_d and p_d in place of P_c and p_c, respectively.

(c) If $A \overset{n_A}{\leftrightarrow} A$ and $m(A) > 0$, then

$$\inf\{p_\mathrm{c}^{(2n_A)}(y|x) \colon x \in A_0,\ y \in A_0\} \geqslant \delta_A > 0, \tag{22}$$

with $A_0 \in A^+$. If A is countable we have (22) with p_d in place of p_c.

Our next result requires the following condition:

Condition 1. *If* $m(E_k) \to 0$ *as* $k \to \infty$ *then:*

$$\lim_{k\to\infty} P_c(E_k|x) < 1, \qquad \textit{uniformly on } x. \tag{23}$$

Note that for each $x \in S$ *we always have* $P_c(E_k|x) \xrightarrow{k} 0$*. Condition 1 requires that the convergence be uniform on S. Also if* $p_c(y|x) \leqslant K < \infty$ *is bounded then* (23) *holds trivially, since* $P_c(E_k|x) \leqslant Km(E_k)$*.*

Theorem 1. *If a chain is irreducible and aperiodic, and if Condition 1 is satisfied, then it possesses a long-run distribution*

$$\lim_{n\to\infty} P^{(n)}(B|x) = Q(B), \qquad \forall B \in \mathscr{B}^d, \tag{24}$$

where Q is a probability on $(\mathbb{R}^d, \mathscr{B}^d)$*.*

Proof. The proof requires several steps and uses some of the techniques found in Doob (1953).

(a) Since the chain is irreducible and aperiodic, by Proposition 2 there exists $S_c \in S^+$ such that $S \xrightarrow{n_S} S_c$ (also $S_c \xrightarrow{n_S} S_c$). From (22) there exist $\delta_1 > 0$ and $S_c' \in S_c^+$ such that

$$\inf\{p_c^{(2n_s)}(y|x)\colon x \in S_c',\ y \in S_c'\} \geqslant \delta_1. \tag{25}$$

From (20) we have

$$P_c^{(n_S)}(S_c'|x) > 0, \qquad \forall x \in S. \tag{26}$$

Let $E_k = \{x\colon P_c^{(n_S)}(S_c'|x) < 1/k\}$; then, by (26) and the fact that $m(S) < \infty$, we have $m(E_k) \to 0$. From Condition 1, there exist $\epsilon_0 > 0$ and k_0 such that

$$P_c(E_{k_0}|x) \leqslant 1 - \epsilon_0, \qquad \forall x \in S. \tag{27}$$

Since $P_c^{(n_S)}(S_c'|z) \geqslant 1/k_0$ for $z \in S\backslash E_{k_0}$, using (6) and (27) we can write, for $x \in S$,

$$\begin{aligned} P_c^{(n_S+1)}(S_c'|x) &\geqslant \int_{S\backslash E_{k_0}} P_c^{(n_S)}(S_c'|z) P(\mathrm{d}z|x) \\ &\geqslant \frac{1}{k_0} P(S\backslash E_{k_0}|x) \geqslant \frac{\epsilon_0}{k_0}. \end{aligned} \tag{28}$$

Now take $D = S_c^+$ (thus $m(D) > 0$), $n_D = 3n_S + 1$ and $\delta_D = \delta_1\epsilon_0/k_0$. Then, using (6), (25) and (28), we have for $y \in D$ and $x \in S$,

$$\begin{aligned} p_c^{(3n_S+1)}(y|x) &\geqslant \int_D p_c^{(2n_S)}(y|z) p_c^{(n_S+1)}(z|x)\,\mathrm{d}z \\ &\geqslant \delta_1 P_c^{(n_S+1)}(S_c'|x) \geqslant \delta_D > 0. \end{aligned}$$

Thus there exist $\delta_D > 0$, $n_D \geqslant 1$ and $D \in S^+$ such that

$$\inf \{ p_c^{(n_D)}(y|x)\colon x \in S,\ y \in D\} \geqslant \delta_D. \tag{29}$$

(b) Let D and δ_D satisfy (29) and $\epsilon_D = \delta_D m(D)$, then

$$|P^{(kn_D)}(B|x) - P^{(kn_D)}(B|y)| \leqslant (1 - \epsilon_D)^k \tag{30}$$

$\forall B \in \mathscr{B}^d$, $\forall x \in S$, $\forall y \in S$ and $k \geqslant 1$.

From (1) and (29) we have

$$P^{(n_D)}(B|x) \geqslant \int_{B \cap D} p_c^{(n_D)}(y|x)\,\mathrm{d}y \geqslant \delta_D m(B \cap D)$$

and

$$P^{(n_D)}(B^c|x) \geqslant \delta_D m(B^c \cap D) = \epsilon_D - \delta_D m(B \cap D).$$

It follows that for $x \in S$,

$$\delta_D m(B \cap D) \leqslant P^{(n_D)}(B|x) \leqslant 1 - \epsilon_D + \delta_D m(B \cap D). \tag{31}$$

Using inequality (31) with y in place of x, we can write

$$P^{(n_D)}(B|x) - P^{(n_D)}(B|y) \leqslant 1 - \epsilon_D.$$

Interchanging the roles of x and y, we obtain

$$|P^{(n_D)}(B|x) - P^{(n_D)}(B|y)| \leqslant 1 - \epsilon_D. \tag{32}$$

For $k \geqslant 2$, let

$$L(\mathrm{d}z;\ x,\ y,\ k) = P^{((k-1)n_D)}(\mathrm{d}z|x) - P^{((k-1)n_D)}(\mathrm{d}z|y) \tag{33}$$

$$U = (L(\mathrm{d}z;\ x,\ y,\ k) \geqslant 0) \text{ and } V = (L(\mathrm{d}z;\ x,\ y,\ k) < 0).$$

And we can write

$$P^{(kn_D)}(B|x) - P^{(kn_D)}(B|y) = \int_U P^{(n_D)}(B|z) L(\mathrm{d}z;\ x,\ y,\ k) + \int_V P^{(n_D)}(B|z) L(\mathrm{d}z;\ x,\ y,\ k).$$

From (31) we have

$$\int_U P^{(n_D)}(B|z) L(\cdot) \leqslant (1 - \epsilon_D + \delta_D m(B \cap D)) \int_U L(\cdot)$$

and

$$\int_V P^{(n_D)}(B|z) L(\cdot) \leqslant \delta_D m(B \cap D) \int_V L(\cdot).$$

Since $\int_U L(\cdot) + \int_V L(\cdot) = 0$, we have

$$P^{(kn_D)}(B|x) - P^{(kn_D)}(B|y) \leqslant (1 - \epsilon_D) \int_U L(\cdot). \tag{34}$$

If $k = 2$ we have, from (32),

$$\int_U L(\cdot) = P^{(n_D)}(U|x) - P^{(n_D)}(U|y) \leqslant 1 - \epsilon_D.$$

Thus

$$P^{(2n_D)}(B|x) - P^{(2n_D)}(B|y) \leqslant (1 - \epsilon_D)^2.$$

Induction arguments and (34) give us (30).

(c) For $k \geqslant 1$, $m \geqslant 1$ and $x \in S$, we have

$$|P^{(kn_D+m)}(B|x) - P^{(kn_D)}(B|x)| \leqslant (1 - \epsilon_D)^k. \tag{35}$$

Since $\int_S P^{(m)}(\mathrm{d}y|x) = 1$ and $P^{(kn_D+m)}(B|x) = \int_S P^{(kn_D)}(B|y)P^{(m)}(\mathrm{d}y|x)$, we can write

$$P^{(kn_D+m)}(B|x) - P^{(kn_D)}(B|x) = \int_S [P^{(kn_D)}(B|y) - P^{(kn_D)}(B|x)]P^{(m)}(\mathrm{d}y|x).$$

and (35) follows from (30).

(d) $P^{(n)}(B|x)$ is a Cauchy sequence by (35). For $B \in B^d$ let $Q(B) = \lim_{n\to\infty} P^{(n)}(B|x)$, which is independent of x by (30). It is easy to verify that Q is σ-additive on $\mathscr{B}^d$ and since $Q(S) = 1$ it is a probability on $(\mathbb{R}^d, \mathscr{B}^d)$. □

Remark 2. (a) Under the hypothesis of Theorem 1 the long-run distribution Q necessarily has an absolutely continuous part. Note that from (29) we have $p_{\mathrm{c}}^{(n_D)}(y|x) \geqslant \delta_D > 0$, $\forall y \in D$ and $\forall x \in S$ with $m(D) > 0$. And from (6) for $y \in D$, $x \in S$,

$$p_{\mathrm{c}}^{(n_D+1)}(y|x) \geqslant \int_S p_{\mathrm{c}}^{(n_D)}(y|z)P(\mathrm{d}z|x) \geqslant \delta_D.$$

Thus for $D' \in D^+$ we have

$$\lim_{n\to\infty} P_{\mathrm{c}}^{(n)}(D'|x) \geqslant \delta_D m(D').$$

(b) Our next theorem shows that the results of Theorem 1 hold if we assume the following condition:

Condition 1′. *if* $m(E_k) \to 0$ *then* $\lim_{n\to\infty} P_{\mathrm{d}}(E_k|x) = 0$ *uniformly on* S.

Theorem 1′. *Assume that the chain is irreducible and aperiodic with* $S_{\mathrm{d}} \neq \varnothing$. *Then* (24) *holds if Condition 1′ is satisfied.*

Proof. From Proposition 2(c), if the chain is irreducible and aperiodic then $S = S_{\mathrm{c}} \cap S_{\mathrm{d}}$ with S_{d} countable and $S_{\mathrm{c}} = S \backslash S_{\mathrm{d}}$. Since $S_{\mathrm{d}} \neq \varnothing$ there exists $n_S \geqslant 1$ with $S \xrightarrow{n_S} S_{\mathrm{d}}$, and by (22) there exist $S'_{\mathrm{d}} \in S^+_{\mathrm{d}}$ and $\delta_1 > 0$ such that

$$\inf\{p_{\mathrm{d}}^{(2n_S)}(y|x)\colon x \in S'_{\mathrm{d}},\ y \in S'_{\mathrm{d}}\} \geqslant \delta_1.$$

And by (20) we have $P_{\mathrm{d}}^{(n_S)}(S'_{\mathrm{d}}|x) > 0$, $\forall x \in S$.

Let $E_k = \{x\colon P_{\rm d}^{(n_S)}(S'_{\rm d}|x) < 1/k\}$; then $m(E_k) \to 0$. From Condition 1′, given $\epsilon_0 > 0$, there exists k_0 such that $P_{\rm d}(E_{k_0}^{\rm c}|x) \geqslant \epsilon_0$ for $x \in S$. From (7) we have

$$P_{\rm d}^{(n_S+1)}(S'_{\rm d}|x) \geqslant \sum_{z\in(E_{k_0}^{\rm c}\cap S_x)} P_{\rm d}^{(n_S)}(S'_{\rm d}|z)p_{\rm d}(z|x)$$

$$\geqslant \frac{1}{k_0} P_{\rm d}(E_{k_0}^{\rm c}|x) \geqslant \frac{\epsilon_0}{k_0}.$$

Now let $D = S'_{\rm d}$, $n_D = 3n_S + 1$ and $\delta_D = \delta_1\epsilon_0/k_0$ and we have for, $y \in D$ and $x \in S$,

$$p_{\rm d}^{(3n_S+1)}(y|x) \geqslant \sum_{z\in D} p_{\rm d}^{(2n_S)}(y|z)p_{\rm d}^{(n_S+1)}(z|x)$$

$$\geqslant \delta_1 P_{\rm d}^{(n_S+1)}(S'_{\rm d}|x) \geqslant \frac{\delta_1\epsilon_0}{k_0} = \delta_D > 0.$$

Thus there exist $\delta_D > 0$, $n_D \geqslant 1$ and $D \in S_{\rm d}^{+}$ such that

$$\inf\{p_{\rm d}^{(n_D)}(y|x)\colon x \in S,\ y \in D\} \geqslant \delta_D. \tag{36}$$

It follows that for $B \in B^d$

$$P^{(n_D)}(B|x) \geqslant \sum_{(y\in B\cap D\cap S_x^{(n_D)})} p_{\rm d}^{(n_D)}(y|x) \geqslant \delta_D\|D\cap B\|$$

and

$$\delta_D\|D\cap B\| \leqslant P^{(n_D)}(B|x) \leqslant 1 - \delta_D\|D\| + \delta_D\|D\cap B\|.$$

Since $0 < \delta_D\|D\| < 1$, using the same arguments as in Theorem 1 we obtain (24). □

Theorem 2. *Let $\{E_1, \ldots, E_k\} \subset S^1$ be mutually communicating and aperiodic subsets of S. For $E = \bigcup_{i=1}^{k} E_i$, let $F = S\backslash E$. Assume that $F \neq \varnothing$, $m(E) > 0$, Condition 1 holds and that for some r and n_F we have $F \overset{n_F}{\rightarrow} E_r$. Then the chain has a long-run distribution.*

Proof. Since $m(E) > 0$ we may assume $m(E_i) > 0$ for $i = 1, \ldots, \ell$ and E_i countable for $i = \ell + 1, \ldots, k$. Let $E_{\rm c} = \bigcup_{i=1}^{\ell} E_i$ and $E_{\rm d} = \bigcup_{i=\ell+1}^{k} E_i$.

First, we will show that there exists $n'_F \geqslant 1$ such that

$$F \overset{n'_F}{\rightarrow} E_{\rm c} \text{ and } F \overset{n'_F}{\rightarrow} E_{\rm d} \qquad (\text{if } E_{\rm d} \neq \varnothing). \tag{37}$$

Since the E_i are communicating and aperiodic subsets we can take m large enough so that $E_r \overset{m}{\rightarrow} E_i$ for $i = 1, \ldots, k$. Since $F \overset{n_F}{\rightarrow} E_r$ we have (37) by setting $n'_F = n_F + m$ and using Proposition 1.

Using aperiodicity again, there exists $n_E \geqslant 1$ such that

$$S \overset{n_E}{\rightarrow} E_{\rm c} \text{ and } S \overset{n_E}{\rightarrow} E_{\rm d} \qquad (\text{if } E_{\rm d} \neq \varnothing). \tag{38}$$

Now $m(E_c) > 0$ and $E_c \overset{n_E}{\to} E_c$. Using exactly the same type of argument as in the proof of Theorem 1, we show that there exist $D \in E^+$, $\delta_D > 0$ and $n_D \geqslant 1$ such that

$$\inf\{p_c^{(n_D)}(y|x)\colon x \in S,\ y \in D\} \geqslant \delta_D.$$

Following the proof of Theorem 1, we have (24). □

3. Applications

Consider the problem of estimating the global minimum of $f\colon S \to \mathbb{R}$, that is,

$$y_{\min} = \min_{x \in S}\{f(x)\} \text{ or } S_{\min} = \{x\colon x \in S,\ f(x) = y_{\min}\}. \tag{39}$$

Assume that S is bounded with $m(S) > 0$, the global minimum $y_{\min}$ is finite, f is continuous in a neighbourhood of each minimum point $x_{\min} \in S_{\min}$ and the minimum points are interior points of S.

The following random search algorithm will be used: let $X_0 \in S$ be an initial random point; for each $x \in S$ let $g(\cdot|x)$ be a density function on $\mathbb{R}^d$; for $k \geqslant 0$ let X_k denote the value of the algorithm at step k; at step $k+1$ a random value Y_k is generated according to the density $g(\cdot|X_k)$ and we define

$$X_{k+1} = \begin{cases} Y_k & \text{with probability } a(Y_k|X_k) \\ X_k & \text{with probability } 1 - a(Y_k|X_k). \end{cases}$$

It follows that the Markov chain $\{X_n\}_{n \geqslant 0}$ has the transition probability function given by $P(B|x) = P_c(B|x) + P_d(B|x)$, with

$$P_c(B|x) = \int_B p_c(y|x)\,dy, \qquad p_c(y|x) = a(y|x)g(y|x) \tag{40}$$

and

$$P_d(B|x) = \sum_{y \in B \cap \{x\}} p_d(y|x), \qquad p_d(x|x) = 1 - \int_S p_c(y|x)\,dy, \tag{41}$$

and $p_d(y|x) = 0$ if $y \neq x$.

Note that the second step transition is given by

$$P^{(2)}(B|x) = \int_S P(B|y)p_c(y|x)\,dy + P(B|x)p_d(x|x),$$

and writing $P(B|x) = \int_B p_c(y|x)\,dy + I_{(x\in B)}p_d(x|x)$ we have

$$P_c^{(2)}(B|x) = \int_B \left[\int_S p_c(z|y)p_c(y|x)\,dy + p_d(z|z)p_c(z|x) + p_c(z|x)p_d(x|x)\right] dz$$

and

$$P_d^{(2)}(B|x) = I_{(x\in B)}p_d^2(x|x).$$

In general we have

$$p_{\mathrm{d}}^{(n)}(y|x) = p_{\mathrm{d}}^{n}(y|x)$$

and

$$p_{\mathrm{c}}^{(n)}(y|x) = \int_S p_{\mathrm{c}}^{(n-1)}(y|z)p_{\mathrm{c}}(z|x)\,\mathrm{d}z + p_{\mathrm{d}}^{n-1}(y|y)p_{\mathrm{c}}(y|x) + p_{\mathrm{d}}(y|y)p_{\mathrm{c}}^{(n-1)}(y|x).$$

Note that, in this case, inequality (6) is strict and we have equality in (7). Two types of algorithms will be analysed.

Algorithm 1. Take $g(y) = g(y|x)$ independent of x and the acceptance probability to be $a(y|x) = I_{(f(y)\leqslant f(x))}I_{(y\in S)}$.

Algorithm 2. Take $a(y|x) = \min\{1, \exp\{-c(f(y) - f(x))\}\}$, where $c > 0$ is a constant.

For Algorithm 2 we assume the same type of hypothesis as in Dekkers and Aarts (1991) (but weaker relative to the objective function f and the set of minimum points $S_{\min}$): (i) if $m(A) > 0$ then $\int_A g(y|x)\,\mathrm{d}y > 0$, $\forall x \in S$; (ii) if $m(E_k) \xrightarrow{k} 0$ then $\int_{E_k} g(y|x)\,\mathrm{d}y \xrightarrow{k} 0$ uniformly on x; and (iii) $\int_S g(y|x)\,\mathrm{d}y = 1$ for all $x \in S$ and $g(y|x) = g(x|y)$.

We will show that the hypothesis of Theorem 2 is satisfied and the long-run distribution is given by

$$Q(B) = \int_B \alpha \mathrm{e}^{-c(f(y)-y_{\min})}\,\mathrm{d}y \qquad \text{with } \alpha^{-1} = \int_S \mathrm{e}^{-c(f(y)-y_{\min})}\,\mathrm{d}y. \tag{42}$$

For $\epsilon > 0$ define

$$\eta(\epsilon) = \{x\colon x \in S, |x - x_0| \leqslant \epsilon \text{ for some } x_0 \in S_{\min}\}. \tag{43}$$

Let $y_{\min}(\epsilon) = \inf\{f(x)\colon x \in S\backslash\eta(\epsilon)\}$ and

$$B(\epsilon) = \eta(\epsilon) \cap \{x\colon x \in S, f(x) \leqslant y_{\min}(\epsilon)\}. \tag{44}$$

Since f is continuous in a neighbourhood of each minimum point we have $m(B(\epsilon)) > 0$. We will show that $B(\epsilon) \overset{1}{\leftrightarrow} B(\epsilon)$ and $S\backslash B(\epsilon) \xrightarrow{1} B(\epsilon)$. This, together with (ii), verifies the conditions of Theorem 2. Thus the long-run distribution exists and coincides with the unique stationary distribution. To prove (42) it is enough to show that the stationary density is given by $q(y) = \alpha\exp\{-c(f(y) - y_{\min})\}$. And this can be done by verifying that q satisfies

$$q(y) = \int_S p_{\mathrm{c}}(y|x)q(x)\,\mathrm{d}x + q(y)p_{\mathrm{d}}(y|y).$$

To prove $B(\epsilon) \overset{1}{\leftrightarrow} B(\epsilon)$, first note that $f(y) - f(x) \leqslant y_{\min}(\epsilon) - y_{\min}$ for $x \in B(\epsilon)$ and $y \in B(\epsilon)$. It follows that $a(y|x) \geqslant \delta_\epsilon = \exp\{-c(y_{\min}(\epsilon) - y_{\min})\}$. Now let $B' \in B^+(\epsilon)$ and $x \in B(\epsilon)$; then by (40) and (i) we have

$$P_{\mathrm{c}}(B'|x) = \int_{B'} a(y|x)g(y|x)\,\mathrm{d}y \geqslant \delta_\epsilon \int_{B'} g(y|x)\,\mathrm{d}y > 0.$$

To prove that $S\backslash B(\epsilon)\xrightarrow{1} B(\epsilon)$, note that for $z\in S\backslash B(\epsilon)$ and $y\in B(\epsilon)$ we have $f(y)\leqslant f(z)$ so that $a(y|z)=1$. And by (i),

$$P(B'|z)\leqslant\int_{B'}g(y|z)\,\mathrm{d}y>0,\qquad \forall B'\in B^{+}(\epsilon).$$

As for Algorithm 1, we assume that x_0 is the unique minimum point and that $g(y)>0$ in a neighbourhood of x_0. An atypical situation arises: $S_{\min}\overset{1}{\leftrightarrow}S_{\min}$ but $S_{\min}$ is not accessible from any other subset of S (for all $n\geqslant 1$ we have $P^{(n)}(\{x_0\}|x)$ equal to 0 if $x\neq x_0$ and equal to 1 if $x=x_0$). Now let $B(\epsilon)$ be defined by (44) and $\epsilon>0$ small enough so that $g(y)>0$ on $B(\epsilon)$. Then we can show that $S\backslash B(\epsilon)\xrightarrow{1} B(\epsilon)$. In this case one can prove directly that the long-run distribution Q is the probability mass at x_0. Note that for all $n\geqslant 1$ and $\epsilon>0$ we have $P^{(n)}(B(\epsilon)|x_0)=1$. And for $x\neq x_0$ and $q_\epsilon=\int_{B(\epsilon)}g(y)\,\mathrm{d}y$ we have $P(B^{\mathrm{c}}(\epsilon)|x)=1-q_\epsilon$. Using induction arguments it is easy to show that, for $x\neq x_0$,

$$P^{(n)}(B^{\mathrm{c}}(\epsilon)|x)=\int_{B^{\mathrm{c}}(\epsilon)}P^{(n-1)}(B^{\mathrm{c}}(\epsilon)|y)P(\mathrm{d}y|x)=(1-q_\epsilon)^n. \tag{45}$$

From (45), if $\eta(\epsilon)$ is an ϵ-neighbourhood of x_0, we have

$$\lim_{n\to\infty}P^{(n)}(\eta(\epsilon)|x)=1,\qquad \forall x\in S.$$

It follows that $X_n\to x_0$ in probability and $Q(\{x_0\})=1$.

Acknowledgement

The research for this paper was partially supported by CNPq-Brasil.

References

Dekkers, A. and Aarts, E. (1991). Global optimization and simulated annealing. *Math. Programming*, **50**, 367–393.

Doob, J.L. (1953). *Stochastic Processes*. New York: Wiley.

Dorea, C.C.Y. (1986). Limiting distribution for random optimization methods. *SIAM J. Control Optim.*, **24**, 76–82.

Dorea, C.C.Y. (1990). Stopping rules for a random optimization method. *SIAM J. Control Optim.*, **28**, 841–850.

Parzen, E. (1962). *Stochastic Processes*. San Francisco: Holden-Day.

Received October 1995 and revised July 1996.