The minimal harmonic functions of sojourn processes of certain finite state Markov chains

Bolthausen, E
THE MINIMAL HARMONIC FUNCTIONS OF SOJOURN PROCESSES OF CERTAIN FINITE STATE MARKOV CHAINS

E. BOLTHAUSEN

ABSTRACT. To a finite state, discrete time Markov chain a new chain is constructed which counts the number of sojourns in each state. This sojourn process is a transient Markov chain. The minimal harmonic functions of this chain are identified.

1. Introduction. Let \( I \) be a finite set of \( s \) states and \( Z_n, n \in \mathbb{N} \) be a sequence of independent identically distributed \( I \)-valued random variables and let \( p_i = \Pr(Z_n = i) \). To \( Z_n \) we associate a sequence of random variables \( X_n = (x_1^{(n)}, \ldots, x_s^{(n)}) \) with values in \( \mathbb{N}_0^s (\mathbb{N}_0 \text{ the set of nonnegative integers}) \) by

\[
    x_i^{(n)} = \sum_{m=1}^{n} 1_i(Z_m). 
\]

Clearly, \( X_n \) is a transient Markov chain. D. Blackwell and D. Kendall calculated in [1] the Martin boundary of this chain. (See also [5, p. 394].) (Actually, Blackwell and Kendall calculated the Martin boundary of Pólya’s urn scheme, but it turns out that Pólya’s urn scheme is a \( h \)-path-transformation of \( X_n \).) The boundary can be identified with the simplex

\[
    \Delta_s = \left\{ t = (t_1, \ldots, t_s): 0 < t_i, \sum_{i=1}^{s} t_i = 1 \right\}. 
\]

The minimal nonnegative harmonic function corresponding to \( t \in \Delta_s \) is given by

\[
    \mathbb{N}_0^s \ni (x_1, \ldots, x_s) \rightarrow \prod_{i=1}^{s} \left( t_i/p_i \right)^{x_i}. 
\]

It is the purpose of this paper to generalize this result to the case where \( Z_n \) is a finite state Markov chain with all transition probabilities nonvanishing. To such a chain we associate \( X_n \) as above, but one has to take the joint variables \( Y_n = (X_n, Z_n) \) to get a Markov chain. We call \( Y_n \) the sojourn process of \( Z_n \). The minimal nonnegative harmonic functions again correspond bijective to points in \( \Delta_s \), but the connection is more complicated.

Received by the editors May 9, 1978 and, in revised form, November 27, 1978.

AMS (MOS) subject classifications (1970). Primary 60J05, 60J45; Secondary 60J50.

Key words and phrases. Markov chain, sojourn times, harmonic functions, Martin boundary.

© 1979 American Mathematical Society 0002-9939/79/0000-0479/$02.75
2. Notations and statements. Let $I = \{1, \ldots, s\}$ be as above and $P = (p_{ij})_{i,j \in I}$ be a matrix of transition probabilities on $I$ which all are nonvanishing. Let $\delta_j \in \mathbb{N}_0$ have components 0 except the $j$th which is 1. To $P$ we associate a transition probability matrix $Q$ on $\tilde{I} = \mathbb{N}_0^s \times I$: if $x, y \in \mathbb{N}_0^s$ then

$$q(x, i, (y, j)) = \begin{cases} p_{ij} & \text{if } y = x + \delta_j, \\ 0 & \text{else.} \end{cases}$$

We call a function $h: \tilde{I} \to \mathbb{R}^+$ harmonic if for all $(x, i) \in \tilde{I},$

$$h(x, i) = \sum_{j \in I} p_{ij} h(x + \delta_j, j) = \sum_{(y, i) \in \tilde{I}} q(x, i, (y, j)) h(y, j).$$

(We shall only consider nonnegative harmonic functions.) Let $\mathcal{H}$ denote the set of harmonic functions which are not identically 0. A harmonic function $h$ is called minimal if whenever $h' \in \mathcal{H}$ and $h' < h$ (i.e., $h'(x, i) < h(x, i)$ for all $(x, i) \in \tilde{I}$) then there is a constant $c > 0$ with $h' = ch$. Let $\mathcal{H}_e$ be the set of minimal harmonic functions.

We shall give a complete description of the elements in $\mathcal{H}_e$.

If $u = (u_1, \ldots, u_s) \in (\mathbb{R}^+)^s - \{0\}$ let $v = (v_1, \ldots, v_s)$ be defined by $v_i = \sum_j p_{ij} u_j$. $u$ and $v$ will always be connected in this way. Clearly, from the assumption that $p_{ij} > 0$ for all $i, j \in I$, we have $v_j > 0$ for all $i$. To $u \in (\mathbb{R}^+)^s - \{0\}$ we associate a function $h_u: \tilde{I} \to \mathbb{R}^+$ by $h_u(x, i) = v_i \prod_{j=i}^{j=s} (u_j/v_j)^{\delta_j}$.

**Theorem 1.** $\mathcal{H}_e = \{h_u: u \in (\mathbb{R}^+)^s - \{0\}\}$.

The proof is deferred to §3.

We consider now the subsets $\mathcal{D} = \{h \in \mathcal{H}: h(0, 1) = 1\}$ and $\mathcal{D}_e = \mathcal{D} \cap \mathcal{H}_e$. It is easy to see that for $h \in \mathcal{H}$, $h(0, 1) > 0$. (In fact, if $h(0, 1) = 0$ then from $h(0, 1) = \sum p_{ij} h(\delta_j, j)$ one obtains $h(\delta_j, j) = 0$ for all $j$ and then $h(0, j) = 0$ for all $j$ and a straightforward conclusion gives $h = 0$.) So each $h \in \mathcal{H}$ has a unique representation $h = ch'$ where $c > 0$ and $h' \in \mathcal{D}$.

If $\mathcal{H}$ is equipped with the topology of pointwise convergence, $\mathcal{D}$ is a compact convex subset of $\mathcal{H}$ with $\mathcal{D}_e$ as the set of extremal points. Now $h_u \in \mathcal{D}$ if and only if $\sum_{j=1}^{j=s} p_{ij} u_j = 1$. Let $\tilde{\mathcal{H}}_e = \{u \in (\mathbb{R}^+)^s: \sum p_{ij} u_j = 1\}$, then the topology on $\tilde{\mathcal{H}}_e$ corresponds to the usual topology on $(\mathbb{R}^+)^s$ under the correspondence $\tilde{\mathcal{H}}_e \ni u \mapsto h_u \in \mathcal{D}_e$. From general Choquet representation theory one thus obtains (see e.g. [2]):

**Theorem 2.** To any $h \in \mathcal{D}$ there exists a unique probability measure $\mu$ on the Borel sets of $\tilde{\mathcal{H}}_e$ such that

$$h(x, i) = \int h_u(x, i) \mu(du)$$

for all $(x, i) \in I$.

One would like to have a probabilistic interpretation of the elements in $\tilde{\mathcal{H}}_e$. Let us first fix some notations: Let $\Omega = \mathcal{I}^{\mathbb{N}_0}$ equipped with the product $\sigma$-algebra which we denote by $\mathcal{A}$. For $i \in I$ let $P^i$ be the law on $(\Omega, \mathcal{A})$ which
governs the Markov chain with transition probabilities $p_{ij}$ and starting point $i \in I$. Let $Z_n, n \in \mathbb{N}_0,$ be the projections of $\Omega$ on its factors and let $X_n$ be defined as in the introduction ($X_0 = 0$). Then $Y_n = (X_n, Z_n)$ is a transient Markov chain with state space $\tilde{I}$ and transition probabilities $q_{ij}$. We call $Y_n, n \in \mathbb{N}_0,$ the sojourn process. Actually we can start $Y_n$ in any point $(x, i) \in \tilde{I}$ simply by replacing $X_n$ by $x + X_n$ and starting $Z_n$ in $i$. We denote by $Q(x,i)$ the thus induced law on $\tilde{I}^\mathbb{N}_0$. To $u \in \tilde{\Delta}$, we associate Markov chain transition probabilities $p_{ij}^u = p_{ij} u_i / v_i$. Let $\pi^u = (\pi_1^u, \ldots, \pi_n^u)$ be the stationary distribution of the $p_{ij}^u$, that is $\pi^u$ is the unique element in $\Delta_s$ (defined as in the introduction) with $p_{ij}^u = \sum_{i=1}^n \pi_i^u p_{ij}^u$. One thus obtains a mapping $\Psi: \tilde{\Delta} \rightarrow \Delta_s$ defined by $\Psi(u) = \pi^u$.

**PROPOSITION 3.** $\Psi$ is a homeomorphism.

The proof is deferred to §3.

It is now easy to see that the $h$-path transformation (see, e.g., [5]) of the $Y_n$ process by $h_\mu$ is exactly the sojourn process associated with the transition probabilities $p_{ij}^u$. So to any probability measure $\mu$ on $\Delta_s$ there corresponds exactly one element in $\tilde{\Delta}$, namely $h^{-1}(\mu)$, such that the $h$-path transformation of the $Y_n$ process with this element is the sojourn process of a chain with stationary distribution $\mu$.

The present work has relations to the investigations of harmonic functions and Martin boundaries of the so called space-time process $(Z_n, n)$, where $Z_n, n \in \mathbb{N}_0$, are recurrent Markov chains (much more general than ours, see [6]). Indeed, $n$ is simply $\sum_{i=1}^n \chi_i^{(n)}$ in our notation. So our sojourn processes obviously carry more information than the space-time processes.

There is also a connection with certain conditionings investigated by Darroch and Seneta [3]. Let $I_0$ be a proper nonvoid subset of $I$ and let $T = \inf\{n \in \mathbb{N}: Z_n \notin I_0\}$. Then it is easy to see that $Z_0, \ldots, Z_n$ conditioned on $\{T > n\}$ is a nonhomogeneous Markov chain. Darroch and Seneta have shown that for $n \rightarrow \infty$ the transition probabilities of this chain converge to stationary transition probabilities $q_{ii}$ on $I_0$, that is, for $l \in I_0$ and each $k \in \mathbb{N}_0$,

$$\lim_{n \rightarrow \infty} P^l(X_{n+1} = j|X_n = i, T > n) = q_{ij}.$$ 

One can show that $q_{ij} = p_{ij}^u$ where $u = \Psi^{-1}(\mu)$ where $\mu$ is a probability on $I_0$ such that the Donsker-Varadhan information $I(\mu)$ (see [4]) is minimal among probability measures which are carried by $I_0$.

Now $T > n$ is simply a condition on the sojourn times of $Z_n$: $\sum_{i \notin I_0} \chi_i^{(n)} = 0$. One may consider other conditionings: Let $\Delta_0$ be a closed subset of $\Delta_s$ and $l \in I$ such that $Q^{(0,l)}(X_n / n \in \Delta_0) > 0$ for sufficiently large $n$. Then one can take elementary conditionings on $\{X_n / n \in \Delta_0\}$. In general, $Z_0, \ldots, Z_n$ is not a Markov chain, but $Y_0, \ldots, Y_n$ is. One may conjecture that, as in the Darroch-Seneta case, if $I(\mu)$ attains a minimum on $\Delta_0$ which is uniquely
attained in μ₀, then
\[ \lim_{n \to \infty} \mathcal{Q}^{(0,i)}(Y_{k+1} = (y,j)|Y_k = (x,i), X_n/n \in \Delta_0) = q^{u,(x,i),(y,j)}(\mu_0) \]
where \( q^{u,(x,i),(y,j)} \) are constructed as above from the \( p^{u,y}_{y} \) and where \( u = \Psi^{-1}(\mu_0) \).
As far as I know, such conditionings have not been considered in the literature as yet.

3. Proofs. For \( x \in \mathbb{N}_0 \) we associate a map \( \phi_x : \mathcal{Y} \to \mathcal{Y} \) by \( (\phi_x h)(y,i) = h(x+y,i) \). It is obvious that \( \phi_x h \in \mathcal{Y} \) if \( h \in \mathcal{Y} \). We simply write \( 0_1 \) for \( 4_{*} \). For \( i \in I \) let \( T_i : (\hat{I})^\mathbb{N}_0 \to \mathbb{N}_0 \) be defined by
\[
\tau_i((x_0, i_0), (x_1, i_1), \ldots) = \inf \{ n \in \mathbb{N}_0 : x_{n,i} > 1 \}
\]
where \( x_{n,i} \) is the \( i \)th component of \( x_n \). Clearly \( \tau_i \) is a stopping time for the \( Y_n \) process and \( Q^{(x,i)}(\tau_i < \infty) = 1 \) for all \( (x,j) \in \hat{I} \) and \( Q^{(x,i)}(\tau_i = 0) = 1 \) if \( x_i > 1 \). For \( h \in \mathcal{Y} \) and \( i \in I \) let \( \Lambda_i h \) be the function \( \hat{I} \to \mathbb{R}^+ \) defined by
\[
(\Lambda_i h)(x,j) = E^{(x,i)}(h(Y_{\tau_i} - \delta_i))
\]
where \( (x, l) - \delta_i = (x - \delta_i, l) \).

Lemmma 1. If \( h \in \mathcal{Y} \) then \( \Lambda_i h \in \mathcal{Y} \).

Proof. Let \( (x, l) \in \hat{I} \),
\[
\sum_{(y,j) \in I} q^{(x,i),(y,j)}(\Lambda_i h)(y,j)
\]
\[ = \sum_{j \in I} p^{y}_{y} E^{(x+\delta_j,i)} h(Y_{\tau_i} - \delta_i) = E^{(x,i)} h(Y_{\tau_i} - \delta_i) \]
where \( \tau_i = \max(\tau_i, 1) \). Now \( x_i = 0 \) if and only if \( \tau_i > 1 Q^{(x,i)} \) a.s. So if \( x_i = 0 \) the above expression becomes
\[
E^{(x,i)} h(Y_{\tau_i} - \delta_i) = (\Lambda_i h)(x, l)
\]
and if \( x_i > 1 \) it equals \( E^{(x,i)} h(Y_{\tau_i} - \delta_i) \) which by harmonicity is
\[
h((x - \delta_i, l)) = E^{(x,i)} h(Y_{0} - \delta_i) = E^{(x,i)}(Y_{\tau_i} - \delta_i) = (\Lambda_i h)(x, l). \]
So the lemma is proved.
Clearly, one has for \( h \in \mathcal{Y} \),
\[
(\Lambda_i \phi_i h)(x,j) = E^{(x,j)} h(Y_{\tau_i}) < h(x,j)
\]
the last inequality because \( h(Y_n), n \in \mathbb{N}_0 \) is a nonnegative martingale and \( \tau_i \) is a stopping time.

\[
(\Lambda_i \phi_i h)(x,j) = h.
\]

Lemmma 2. If \( h \in \mathcal{Y}_e \) then \( \phi_i h = 0 \) or \( \phi_i h \in \mathcal{Y}_e \).

Proof. From (3.1), \( \Lambda_i \phi_i h < h \) so \( \Lambda_i \phi_i h = c h \) for some \( c \in [0, 1] \). If \( c < 1 \) then it follows from the fact that \( (\Lambda_i \phi_i h)(x,j) = h(x,j) \) for all \( (x,j) \) with \( x_i > 1 \) that \( h(x,j) = 0 \) for \( x_i > 1 \), so \( \Lambda_i \phi_i h = 0 \) and therefore from (3.2), \( \phi_i h = 0 \). If \( c = 1 \), let \( h' \in \mathcal{Y}, h' < \phi_i h \). It follows that \( h = \Lambda_i \phi_i h > \Lambda_i h' \) so
there is a \( c' \in [0, 1] \) with \( \Lambda_i h' = ch \) and so \( h' = \phi_i \Lambda_i h' = c' \phi_i h \), and it is proved that \( \phi_i h \) is extremal.

**Lemma 3.** Let \( h \in \mathcal{K}_e \) and \( x, y, z \in N_0 \) with \( \sum_{i \in I} x_i y_i = 0 \) and \( \phi_h \neq 0 \) and \( \phi_j h \neq 0 \). Then \( \phi_{x+y} h \neq 0 \).

**Proof.** We proceed by induction on \( |x| = \sum_{i \in I} x_i \). If \( |x| = 0 \) is clear. If \( i \) is such that \( y_i = 0 \) and \( \sum x_i y_i = 0 \) and \( \phi_{x+y} h \neq 0 \) then
\[
\Lambda_i \phi_{x+y} h = \phi_y \phi_x \Lambda_i \phi_i h = \phi_{y+x} h
\]
by the proof of Lemma 2. So \( \phi_{x+y} h \neq 0 \).

**Lemma 4.** \( i \in \mathcal{K}_e \).

**Proof.** Let \( h \in \mathcal{K}_e \), \( h < 1 \). Fix a state \( i \in I \) and let \( \hat{h}_i : N_0^i \to R^+ \) be defined by \( \hat{h}_i(x) = h(x, i) \). Let \( \tau_0 = 0 \) and \( \tau_k \) be the time of the \( k \)th hitting of \( i \) by the chain \( Z_n \), \( n \in N \). It is well known that \( X_{\tau_k} - X_{\tau_{k-1}} \), \( k \in N \) are independent identically distributed \( N_0^i \) valued random variables under the law \( Q^{(y, i)} \) for every \( y \in N_0^i \). As \( \{h(Y_n), n \in N_0\} \) is a bounded martingale it follows that \( \{\hat{h}_i(X_n), k \in N_0\} \) is a bounded martingale, so \( \hat{h}_i \) is bounded harmonic for the random walk \( X_n \), \( k \in N_0 \). Our assumptions on the \( p_{ij} \) clearly imply that the support of the law of \( X_{\tau_k} - X_0 \) generates \( Z^+ \) as an Abelian group. So it follows from the Choquet-Deny theorem (see [7, p. 135]) that \( \hat{h}_i \) is a constant, say \( c_i \). From harmonicity of \( h \) it follows that \( c_i = \sum_{j \in I} p_{ij} c_j \) for all \( i \in I \), so all \( c_i \) are the same. So \( h \) is constant and it follows that \( 1 \in \mathcal{K}_e \).

**Proof of Theorem 1.** The proof splits in two parts:

(I) For each \( u \in (R^+)^I - \{0\} \), \( h_u \in \mathcal{K}_e \).

(II) If \( h \in \mathcal{K}_e \) there is an \( u \in (R^+)^I - \{0\} \) with \( h = h_u \).

**Proof of (I).** Let \( u \in (R^+)^I - \{0\} \) and \( N_u = \{i \in I: u_i = 0\} \). For \( i, j \in N_u \) let \( \tilde{p}_{ij} = p_{ij} / v_i \). The \( \tilde{p}_{ij} \), \( i, j \in N_u \) form a Markov transition probability matrix on \( N_u \).

Let \( h \in \mathcal{K}_e \), \( h < h_u \). Clearly \( h(x, i) = 0 \) if \( x_j > 1 \) for some \( j \in N_u \). It is then easy to see that \( h / h_u \) restricted to elements \( (x, i) \), \( i \notin N_u \) and \( x_j = 0 \) for \( j \in N_u \) is bounded harmonic for the space time process associated to the \( \tilde{p}_{ij} \). From Lemma 4 (applied to \( \tilde{p} \)) there is a \( c \in (0, 1] \) with \( h(x, i) = ch_u(x, i) \) for all \( (x, i) \) satisfying the above restriction but clearly then for all \( (x, i) \in \bar{I} \) as for the other elements both sides are 0. So \( h_u \) is minimal.

**Proof of (II).** Let \( h \in \mathcal{K}_e \). It is easy to see that \( h(0, l) \neq 0 \) for each \( l \in I \). Indeed if \( h(0, l) = 0 \) for some \( l \) \( h \) would be identically 0.

Let \( N_h = \{i \in I: \phi_h = 0\} \) and \( d = \sum_{i \in N_h} \delta_i \). From Lemma 3, \( \phi_d h \neq 0 \). Let \( i \notin N_k \) and \( i_1, \ldots, i_k \) be an enumeration of the elements of \( N_h^* \) with \( i_k = i \). Then for \( x \in N_0^l \),
\[
h(x, i) > p_{i_1} p_{i_2} \cdots p_{i_{k-1}} h(x + d, i).
\]
It follows that there is a \( c > 0 \) such that \( h(x, i) > ch(x + d, i) \) for all \( x \) and \( i \notin N_k \). But then for all \( (x, j) \in \bar{I} \),
\[ h(x,j) = \sum_{i \in N_h^c} p_{ij} h(x + \delta_i, i) \geq c \sum_{i \in N_h^c} p_{ij} h(x + d + \delta_i, i) = ch(x + d, i) \]
that is \( h > c\phi_d h \). As \( \phi_d h \neq 0 \) there is a \( c_1 > 0 \) such that \( h = c_1 \phi_d h \) and then \( h = c_1^2 \phi_d h \).

Let now \( l, i \notin N_h^c \) and let \( i_1, \ldots, i_{2k-1} \) be a sequence of elements in \( N_h^c \) where each element of \( N_h^c \) appears exactly twice except \( i \) which appears once and further \( i_{2k-1} = l \). Then

\[
(\phi_i h)(x, l) = h(x + \delta_i, l) > p_{li_1} p_{i_1i_2} \cdots p_{i_{2k-1}i_{2k-1}} h(x + 2d, l)
\]
for each \( x \in N_0^c \). So it follows as above that there is a \( c_2 > 0 \) with \( \phi_i h \geq c_2 \phi_d h \). From \( h = c_1^2 \phi_d h \) it follows that \( \phi_d h \neq 0 \). So from Lemma 2 it follows that there is a constant \( \lambda_i > 0 \) with \( \phi_i h = \lambda_i h \). This obviously remains true for \( i \in N_h \) and \( \lambda_i = 0 \).

It follows that for \( (x, l) \in \tilde{I}, h(x, l) = \lambda_1^3 \lambda_2^2 \cdots \lambda_5^5 h(0, l) \). We further have the relation

\[ h(0, i) = \sum_{j \in I} p_{ij} h(\delta_j, j) = \sum_{j \in I} p_{ij} \lambda_j h(0, j) \]
so putting \( u_i = \lambda_i h(0, i), u = (u_1, \ldots, u_s) \) one easily obtains \( h = h_u \).

**Proof of Proposition 3.** As \( \Delta_x \) and \( \Delta_x \) both are compact, it suffices to show that \( \Psi \) is continuous and bijective. Clearly the \( \bar{p}_y \) depend continuously on \( u \in \Delta_x \) and the expectation of the recurrence times depend continuously on the transition probabilities. So the coefficients \( \pi_u \) of the stationary distribution being the inverses of the expected recurrence times depend continuously on \( u \). We now prove that \( \Psi \) is onto. First, if \( \pi \in \Delta_x \) has all components different from 0 it follows from Lemma 2.4 of [4] that there is a \( u \in \Delta_x \) with \( \pi = \pi_u \). If now \( \pi \) has some components 0, say \( \pi_1 = \cdots = \pi_r = 0 \) \( (1 \leq r < s) \). Then let \((\bar{p}_{y})_{i,j=r+1,\ldots,s}\) be defined by

\[
\bar{p}_{ij} = \frac{p_{ij}}{\sum_{k=r+1}^s p_{ik}}
\]
and let \( u = (u_{r+1}, \ldots, u_s) \) with \( \sum_{j=r+1}^s p_{ij} u_j = 1 \) be such that \( p_{ij}^u \) has stationary distribution \( (\pi_{r+1}, \ldots, \pi_s) \) on \( \{r+1, \ldots, s\} \). Then \( u \) \( (0, \ldots, 0, u_{r+1}, \ldots, u_s) \) does it.

It remains to show that \( \Psi \) is 1-1. Let \( u, u' \in \Delta_x \) with \( \pi_u = \pi_{u'} \). We may assume that \( u_i \neq 0 \) for all \( i \in I \). Now \( p_{ij}^u = (p_{ij}^u)^{u'/u} \) where \( (u'/u)_i = u'_i / u_i \). So it follows from Lemmas 2.5 and 2.7 of [4] that

\[
\sum_j \log\left((u_j'/u_j)\left(\sum_k p_{jk}^u (u_k'/u_k)\right)^{-1}\right)\pi_j^u = 0
\]
or

\[
\sum_j \pi_j^u \log(u_j'/u_j) = \sum_j \pi_j^u \log\left(\sum_k p_{jk}^u (u_k'/u_k)\right).
\]
Now if \( u \neq u' \) then \( \log(\Sigma_k p_{jk}(u'_k / u_k)) > \Sigma_k p_{jk} \log(u'_k / u_k) \) so

\[
\sum_j \pi_j^u \log(u'_j / u_j) > \sum_j \pi_j^u \sum_k p_{jk} \log(u'_k / u_k) = \sum_k \pi_k^u \log(u'_k / u_k)
\]

which is a contradiction.

REFERENCES


Fachgruppe Statistik, Universität Konstanz, Postfach 7733 D-775 Konstanz, Germany