Markov chains and (super)harmonic functions

Xinyu Mao

June 30, 2020

Abstract. In this note, I shall present two theorems related to (super)harmonic functions for Markov chains. The motivation comes from a brief introduction in our textbook(see 10.13 in [1]), in which the proof of the first theorem is given as an exercise. The second theorem, known as Choquet-Deny theorem, is about random walks on abelian groups. The elegant proof included here is an exercise in [2](Chapter 5, Exercise 1.7), which has a strong probabilistic flavor and involves serval results we have learned. During the process of writing this note, I also learned a lot about Markov chains from [3]. What's more, I found a very interesting survey about random walks on graphs: [4].

1 Time-homogeneous Markov chains

Let S be the set of *states* and $(S, \mathcal{G})$ be a measurable space.

If S is at most countable, a *time-homogeneous Markov chain* can be characterized by a probability measure on $(S, \mathcal{G})$, say μ, and a *stochastic* $S \times S$ matrix $(p_{ij} : i, j \in S)$ so that

$$p_{ij} \geq 0, \sum_{k \in S} p_{ik} = 1, \forall i, j \in S.$$

μ is said to be the *starting measure*.

Definition 1. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathsf{P}^\mu)$ be a filtered space. A adapted process $X = (X_n)_{n \in \mathbb{N}}$ is a *time-homogeneous Markov chain* with *transition probability* p if

$$\mathsf{P}^\mu(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) = \mu(i_0) p_{i_0 i_1} p_{i_1, i_2} \cdots p_{i_{n-1} i_n}. \tag{*}$$

We write P^μ instead of P to signify the dependence of P on μ.

The construction of $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathsf{P}^\mu)$**.** Take $(\Omega, \mathcal{F}) := \prod_{n\in\mathbb{N}}(S, \mathcal{G})$ and suppose that X_n is the coordinate map (i.e., $X_n(\omega) = \omega_n$). The existence of P^μ follows from Kolmogorov's Extension Theorem, and (*) guarantees the uniqueness of P^μ.

But what if we have uncountable number of states? We have to generalize the idea of stochastic matrix to *Markov kernel*.

Definition 2. Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{G})$ be measure spaces, A *Markov kernel* is a map $\kappa : \Omega \times \mathcal{G} \to [0, 1]$ satisfying

(i) $\forall G \in \mathcal{G}$, the map $\omega \mapsto \kappa(\omega, G)$ is $\mathcal{F}$-measurable;

(ii) $\forall \omega \in \Omega$, the map $G \mapsto \kappa(G, \omega)$ is a probability measure on $(S, \mathcal{G})$.

We call the space $(\Omega, \mathcal{F})$ *source* and $(S, \mathcal{G})$ is said to be *target*.

Let p be a Markov kernel with source $(\Omega, \mathcal{F})$ and target $(S, \mathcal{G})$. We now extend the definition to the case where S is uncountable.

Definition 3. A process $X = (X_n)_{n\in\mathbb{N}}$ on $(\Omega, \mathcal{F}, \mathsf{P}^\mu)$ is a *time-homogeneous Markov chain* with *transition probability* p and starting measure μ if for all $F_0, F_1, \ldots, F_n \in \mathcal{F}$,

$$\begin{aligned}&\mathsf{P}^\mu(X_0 \in F_0, X_1 \in F_1, \ldots, X_n \in F_n) = \\ &\int_{F_0} \cdots \int_{F_{n-1}} p(y_{n-1}, F_n)p(y_{n-2}, dy_{n-1}) \cdots p(y_0, dy_1)\mu(dy_0).\end{aligned}$$

2 Non-negative superharmonic functions

Suppose that S is **at most countable** and $\mathcal{G}$ is the power set of S for now.

For function $h : S \to \mathbb{R}$, define a new function Ph via

$$(Ph)(i) := \sum_{j\in S} h(j)p_{ij}.$$

Intuitively, the operator P averages h over all possible next states, with weight given by the Markov chain.

Definition 4. The function $h : S \to \mathbb{R}$ is P*-harmonic*(P*-superharmonic*) if $Ph = h$(if $Ph \le h$).

Proposition 5. *h is P-harmonic(P-superharmonic) if and only if $h(X_n)$ is a martingale(supermartingale).*

Proof. Just observe that $\mathsf{E}^\mu[h(X_{n+1})|\mathcal{F}_n] = (Ph)(X_n)$. $\square$

It is natural to consider the *hitting time* of state j:

$$T_j := \inf\{n \geq 1 : X_n = j\}.$$

We now focus on Markov chains with a very property that *it is possible to get to any state y from any state x*, including revisiting x(i.e., $y = x$). This property is called *reducible recurrent.*

Definition 6. A Markov chain if *reducible recurrent* if

$$f_{ij} := \mathsf{P}^i(T_j < \infty) = 1,$$

where P^i denotes P^μ when μ is the unit mass on i(i.e., $\mu = \delta_i$).

Now we can state our first theorem

Theorem 7. *A Markov chain X is recurrent reducible if and only if every non-negative P-superharmonic function is constant.*

Intuitive interpretation. We can think of a non-negative function $h : S \to \mathbb{R}^+$ as assigning 'energy' to each state, then the superharmonic property simply says "the energy is non-increasing in the walk of Markov chain X". Note that superharmonic property has nothing to do with the starting measure μ. Hence, one direction of Theorem 7 is easy to see: if X is recurrent reducible, that is, we can reach any state starting at any state with energy not increasing, then the energy of all states must be equal.

Proof of Theorem 7. Suppose that X is recurrent reducible and let $h : S \to \mathbb{R}^+$ be a P-superharmonic function. Since $h(X_n)$ is a supermartingale and each T_j is a stopping time, we have

$$h(j) = \mathsf{E}^i[f(X_{T_j})] \leq \mathsf{E}^i[f(X_0)] = h(i), \forall i, j \in S.$$

This establishes the 'only if' part.

Now assume that every non-negative P-superharmonic function is constant. By definition,

$$f_{ij} = \sum_{t=1}^{\infty} \mathsf{P}^i(T_j = t) = \sum_{t=1}^{\infty} \mathsf{P}^i(A_t),$$

where $A_t := \{T_j = t\} = \{\omega : X_t = j, X_s \neq j, \forall s \in [t]\}$. Note that $\mathsf{P}^i(A_1) = p_{ij}$ and for $t \geq 2$:

$$\begin{aligned}\mathsf{P}^i(A_t) &= \sum_{x_1 \in S: x_1 \neq j} \sum_{x_2 \in S: x_2 \neq j} \cdots \sum_{x_{t-1} \in S: x_{t-1} \neq j} p_{ix_1} p_{x_1 x_2} \cdots p_{x_{t-1} j} \\ &= \sum_{x_1 \in S: x_1 \neq j} p_{ix_1} \mathsf{P}^{x_1}(A_{t-1}).\end{aligned}$$

Summing over all t yields

$$f_{ij} = p_{ij} + \sum_{k \in S: k \neq j} p_{ik} \sum_{t=2}^{\infty} \mathsf{P}^k(A_{t-1}) = p_{ij} + \sum_{k \in S: k \neq j} p_{ik} f_{kj}. \tag{1}$$

Since $f_{jj} \leq 1$, we have

$$f_{ij} \geq \sum_{k \in S} p_{ik} f_{kj}. \tag{2}$$

Consider function $h^{(j)} : S \to \mathbb{R}^+, i \mapsto f_{ij}$. Eq. (2) says $h^{(j)}$ is superharmonic for all j, and thus f_{ij} is constant for fixed j. We know that $\sum_{j \in S} p_{ij} = 1$, and Eq. (1) forces $f_{ij} = 1$. This means X is reducible recurrent and finishes the proof. $\square$

3 Random walks on locally compact abelian group

Let $(\mathbb{G}, \mathcal{G}, \mu)$ be a probability space where $\mathbb{G}$ is a *locally compact abelian group*. This means we have a topology on G with good properties, and we will not look into this very deep. We attach the precise definition from Wikipedia[1] here for completeness:

Definition 8. A topological group is called *locally compact* if the underlying topological space is *locally compact* and *Hausdorff*.

Consider the random walk on $\mathbb{G}$: the process $Y_n := \sum_{i=1}^n X_i$, where $X_1, X_2, \ldots$ are I.I.D random variables with law μ(i.e., $\mu(X_i \in A) = \mu(A), \forall A \in \mathcal{G}$). The process Y certainly fits our model of Markov chain: the transition probability is

$$p : \Omega \times \mathcal{G} \to [0, 1], (\omega, G) \mapsto \mu(G). \tag{3}$$

Since the number of states might be uncountable, we need to generalize the definition of harmonic function. In fact, we now define the operator P as

$$(Ph)(x) := \int_{\mathbb{G}} h(x + y) p(x, dy).$$

[1] `https://en.wikipedia.org/wiki/Locally_compact_abelian_group`

In the case of Eq. (3), $(Ph)(x) = \int_{\mathbb{G}} h(x+y)\mu(dy)$. Now we can say a function h is P-harmonic if $Ph = h$ as we did in the previous section.

Definition 9. The *support* of μ is the complement of the union of all open sets which are μ-null sets, i.e., the smallest closed set C such that $\mu(\mathbb{G} \setminus C)$=0.

Theorem 10 (Choqeut-Deny theorem)**.** *All bounded P-harmonic functions are constant a.e. on each coset of $\mathbb{G}_\mu$, where $\mathbb{G}_\mu$ is the subgroup generated by the support of μ.*

Proof. Let h be a P-harmonic function and fix some $x \in \mathbb{G}$. Let Y be the Markov chain above with starting measure δ_x. Note that $M_n := h(Y_n) = h(x + X_1 + X_2 + \cdots X_n)$ is a martingale. Since h is bounded, M_n is $\mathcal{L}^p$ bounded, and thus M is an UI martingale. According to Doob's Forward Convergence Theorem, $M_\infty := \lim M_n$ exists and $M_n \to M_\infty$ in $\mathcal{L}^1$ by the UI property.

Observe that M_∞ is permutation invariant(this requires $\mathbb{G}$ is abelian!), and thus M_∞ is a constant a.e. by Hewitt-Savage 0-1 Law. Then $M_n = \mathsf{E}^x(M_\infty|\mathcal{F}_n)$ implies M_n is constant a.e. as well. In particular, $M_1 = h(x + X_1)$ is constant a.e., and we know that $\mathsf{E}^x[h(x + X_1)] = h(x)$ since h is harmonic. Hence, $h(x + X_1) = h(x)$ a.e., which means $h(x + x') = h(x)$ for almost all $x' \in \mathbb{G}_\mu$. Since x is chosen arbitrarily, h is constant a.e. on each coset of $\mathbb{G}_\mu$. $\square$

Remark. Why we can only assert h is constant on cosets of $\mathbb{G}_\mu$ instead of the whole state space $\mathbb{G}$? This is because our random walk relies on μ, and some states cannot be visited starting from a given state x.

The following theorem from analysis is a direct consequence of Theorem 10.

Theorem 11 (Liouville's theorem for harmonic functions)**.** *A bounded harmonic function h on $\mathbb{R}^n$ is constant.*

In the context of analysis, a function $h : \mathbb{R}^n \to \mathbb{R}$ is *harmonic* if it satisfies Laplace's equation

$$\nabla^2 h = \frac{\partial^2 h}{\partial^2 x_1} + \frac{\partial^2 h}{\partial^2 x_2} + \cdots + \frac{\partial^2 h}{\partial^2 x_n} = 0.$$

Proof of Theorem 11. Let h be a bounded harmonic function on $\mathbb{R}^n$. We uses the *mean value property* of harmonic function without presenting a proof: If $B(x, r)$ is a ball with center x and radius r, then

$$h(x) = \frac{1}{w_n r^n} \int_{B(x,r)} h \; dV, \tag{4}$$

where w_n is the area of the unit sphere in $\mathbb{R}^n$. In other words, the value at the center of the ball is given by the average value in the ball.

Let μ be Lebesgue measure on the ball $B(0, w_n^{-n})$. Then Eq. (4) implies that h is P-harmonic with respect to the random walk with Markov kernel in Eq. (3). Note that the support of μ is the entire ball $B(0, w_n^{-n})$, which generates the whole group $\mathbb{R}^n$. By Theorem 10, h is constant on $\mathbb{R}^n$. $\square$

Remark. I like this kind of proof, for it seems that we learned some deterministic truth by looking into randomness. Another proof that impresses me with the similar spirit is the proof of Doob's Forward Convergence Theorem using Doob's Upcrossing Lemma.

References

[1] David Williams. *Probability with martingales.* Cambridge university press, 1991.

[2] Daniel Revuz. *Markov chains.* Elsevier, 2008.

[3] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[4] László Lovász et al. Random walks on graphs: A survey.