Chapter 1

1.1 Basic definition

Definition 1.1

For a probability space $(\Omega, \mathcal{F}, P)$ *and a denumerable set* X. *A Markov chain is a sequence* $Z_n, n = 0, 1, 2\cdots$ *of random variables* $Z_n : \Omega \to X$ *with the following properties:*

a. For $x_i, 0 \leq i \leq n+1$ *if* $P(Z_n = x_N, \cdots, Z_0 = x_0)$*, we have*

$$P(Z_{n+1} = x_{n+1}|Z_n = x_n, \cdots, Z_0 = x_0) = P(Z_{n+1} = x_{n+1}|Z_n = x_n)$$

b. For all elements $x, y \in X$ *and* $m, n \in \mathbb{N}$ *such that* $P(Z_m = x) > 0, P(Z_n = x) > 0$*, we have*

$$P(Z_{m+1} = y|Z_m = x) = P(Z_{n+1} = y|Z_n = x)$$

If we write

$$p(x, y) \ = P_(Z_{n+1} = y|Z_n = x)$$

we obtain the transition matrix $P = (p(x, y))_{x,y \in X}$ *of* Z_n. *The initial distribution of* Z_n *is the probability measure* ν *on* X *defined by*

$$\nu(x) = P(Z_0 = x)$$

♣

Theorem 1.1

For any state space X *and transition matrix* P*, there is always a probability space, called the trajecetory space and rvs* Z_n *on it such that* Z_n *is a Markov chain with* (X, P). *We set*

$$\Omega = X^{\mathbb{N}}$$

and let $a_0, a_1, \cdots, a_k \in X$*, the cylinder with base* $a = (a_1, a_2, \cdots, a_k)$ *is the set*

$$C(a) = \{w = (x_1, \cdots) \text{ with } x_i = a_i, 0 \leq i \leq k\}$$

and define

$$P_\nu(C(a_0, a_1, \cdots, a_k) = \nu(a_0)p(a_0, a_1)\cdots p(a_{k-1}, a_k))$$

with $\mathcal{F}_n$ *the* σ*-algebra generated by* $C(a), a \in X^{k+1}, k \leq n$ *and* $Z_n(\omega) = x_n$ *Then*

a. P_ν *has a unique extension to a probability measure on* $\mathcal{F}$ *also denoted as* P_ν.

b. On the probability space $(\Omega, \mathcal{F}, P_\nu)$ *the projections* Z_n *define a Markov chain with state space* X*, initial distribution* ν *and transition matrix* P.

♡

Lemma 1.1

For a Markov chain (X, P)*, we know*

a. The number $p^n(x, y)$ *is the element at position* (x, y) *in the* n*-th power* P^n *of the transition matrix.*

b. $p^{(m+n)}(x, y) = \sum_{w \in X} p^{(m)}(x, w)p^{(n)}(w, y)$.

c. P^n *is a stochastic.*

♡

Proof a. We know for any initial ν and use the induction

$$\begin{aligned} P_\nu(Z_{k+n+1} = y|Z_k = x) &= \sum_{w \in X} P_\nu(Z_{k+n+1} = y, Z_{k+n} = w|Z_k = x) \\ &= \sum_{w \in X} P(Z_{k+n+1} = y|Z_{k+n} \\ &= w, Z_k = x)P(Z_{k+n} = w|Z_k = x) = \sum_{w \in X} p^{(n)}(x, w)p(w, y) \end{aligned}$$

and use the induction assumption that $p^{(n+1)(x,y)} = \sum_{w \in X} P^n(x, w)P(w, y) = P^{n+1}(x, y)$.

b. Change Z_{k+n} will be fine.
c. $\sum_{y\in X} P^{n+1}(x,y) = \sum_{y\in X, w\in X} P^n(x,w)P(w,y) = 1$ and we are done by induction.

Proposition 1.1

Let Z_n be a Markov chain on the state space X and let $0 \le n_1 < n_2 < \cdots < n_{k+1}$. Show that for $0 < m < n, x, y \in X$ and $A \in \mathcal{F}_m$ with $P_\nu(A) > 0$. Then if $Z_m(\omega) = x$ fir akk $\omega \in A$, then $P(Z_n = y|A) = p^{(n-m)}(x,y)$, deduce that if $x_1, \cdots, x_{k+1} \in X$ are such tthat $P_\nu(Z_{n_k} = x_k, \cdots, Z_{n_1} = x_1) > 0$, then

$$P_\nu(Z_{n_{k+1}=x_{k+1}}|Z_{n_k} = x_k, \cdots, Z_{n_1} = x_1) = P_\nu(Z_{n_{k+1}} = x_{k+1}|Z_{n_k} = x_k)$$

♠

Proof For any cylinder $C(a) \in \mathcal{F}_m$, we know

$$P_\nu(Z_n = y|C(a)) = p^{(n-m)}(x,y)$$

easily and we are done by $\pi - \lambda$ theorem.

The second conclusion is then trivial.

Definition 1.2

A real random variable is a measurable function $f : (\Omega, \mathcal{F}) \to (\mathbb{R}, \overline{\mathcal{B}})$. If the integral of f with respect to $P(\nu)$ exists, we denote it by

$$E_\nu(f) = \int_\Omega f dP_\nu$$

then we may define for a set $W \subset X$

$$v_n^W = 1_W(Z_n(\omega))$$

and $v^\Omega = \sum_{n\ge 0} v_n^W$, then we can compute

$$E_\nu(v_{[k,n]}^W) = \sum_{x\in X}\sum_{k}^{n}\sum_{y\in W} v(x)p^{(j)}(x,y)$$

♣

Definition 1.3

A stopping time is a random variable t taking its values in $\mathbb{N}_0 \cup \{\infty\}$ such that

$$(t \le n) \in \mathcal{F}_n$$

♣

Theorem 1.2

(Strong Markov property) Let Z_n be a Markov chain with initial distribtuion ν and transition matrix P on the state space X, and let t be a stopping time with $P_\nu(t < \infty) = 1$. Show that $(Z_{t+n})_{n\ge 0}$ defined by

$$Z_{t+n}(\omega) = Z_{t(\omega)+n}(\omega)$$

is again a Markov chain with transition matrix P and initial distribution

$$\nu'(x) = P_\nu(Z_t = x)$$

♡

Proof Notice

$$P_\nu(Z_{t+n+1} = x_{n+1}|Z_{t+n} = x_n, \cdots, Z_t = x_0) = \sum_{k\ge 0} P_\nu(t = k)p(x_n, x_{n+1}) = P_\nu(Z_{t+n+1} = x_{n+1}|Z_{t+n} = x_n)$$

and the time-homogenity is secured by the equality above easily. The initial distribution is also easy to be checked.

Definition 1.4

The hitting times $s^W = \inf\{n \ge 0, Z_n \in W\}$ and the first passage times $t^W = \inf\{n \ge 1, Z_n \in W\}$ and define $G(x,y) = E_x(v^y), F(x,y) = P_x(s^y < \infty)$ and $U(x,y) = P_x(t^y < \infty)$ and let $f^{(n)}(x,y) = P_x(s^y = n)$ and $u^{(n)}(x,y) = P_x(t^y = n)$.

We have

$$G(x,y)=\sum_{n\geq 0}p^{(n)}(x,y),\quad F(x,y)=\sum_{n\geq 0}f^{(n)}(x,y),\quad U(x,y)=\sum_{n\geq 0}u^{(n)}(x,y)$$

♣

Definition 1.5

(Factorization) Suppose we have a partition $\overline{X}$ of the state space X with the following property, for

$$x',y'\in\overline{X},p(x,y')=\sum_{y\in y'}p(x,y)$$

is constant for $x\in x'$. If this holds, we may consider $\overline{X}$ as a new satte space wtih transtion matrix $\overline{P}$ and then

$$\overline{p}(x',y')=p(x,y')$$

and the new Markov chain is the factor chain w.r.t. the given partition.

♣

Theorem 1.3

Let Z_n be a Markov chain on the state space X with transition matrix P, and let $\overline{X}$ be a partition of X with the natural projection $\pi : X\to\overline{X}$. Then $\pi(Z_n)$ is a Markov chain on $\overline{X}$ iff $p(x,y')=\sum_{y\in y'}p(x,y)$ is constant for $x\in x'$.

♡

Definition 1.6

For a sequence a_n, $\sum_{n=0}^{\infty}a_nz^n, sz\in\mathbb{C}$ is called the generating function of a_n. For a Markov chain (X,P), its Green function or Green Kernel is

$$G(x,y|z)=\sum_{n\geq 0}p^{(n)}(x,y)z^n$$

with the radius of convergence

$$r(x,y)=1/\limsup(p^{(n)}(x,y))^{1/n}\geq 1$$

Let $r=\inf\{r(x,y)|x,y\in X\}$ and $|z|<r$, we may form the matrix

$$G(z)=(G(x,y|z))_{x,y\in X}$$

then we know $G(z)=\sum_{n\geq 0}z^nP^n$ where the convergence is pointwise and we may know

$$(I-zP)G(z)=I$$

Similarly, we define

$$F(x,y|z)=\sum_{n\geq 0}f^{(n)}(x,y)z^n$$

$$U(x,y|z)=\sum_{n\geq 0}u^{(n)}(x,y)z^n$$

and denote $s(x,y)$ to be the radius of convergence of $U(x,y|z)$ and we have

$$s(x,y)\geq r(x,y)\geq 1$$

since $u^{(n)}(x,y)\leq p^n(x,y)$.

♣

Theorem 1.4

If X is finite then $G(x,y|Z)$ is a rational function in z.

♡

Proof We may consider

$$G(x,y|z)=\pm\frac{\det(I-zP|y,x)}{\det(I-zP)}$$

where $\det(I-zP|y,x)$ means the determinant of the maxtrix by deleting y row and x column informally.

Theorem 1.5

a. $G(x,x|z) = \dfrac{1}{1-U(x,x|z)}, |z| < r(x,x).$
b. $G(x,y|z) = F(x,y|z)G(y,y|z), |z| < r(x,y).$
c. $U(x,x|z) = \sum_y p(x,y)zF(y,x|z), |z| < s(x,x).$
d. If $y \neq x$ *then* $F(x,y|z) = \sum_\omega p(x,\omega)zF(\omega,y|z), |z| < s(x,y).$

Proof a. It is easy to check that

$$p^{(n)}(x,x) = \sum_{k=1}^{n} p^{(n-k)}(x,x)u^{(k)}(x,x)$$

then we have

$$G(x,x|z) = \sum_{n\geq 0} p^{(n)}(x,x)z^n = 1 + \sum_{n=1}^{\infty}\sum_{k=0}^{n} u^{(k)}(x,x)p^{(n-k)}z^n = 1 + U(x,x|z)G(x,x|z)$$

b. If $x = y$, then $F(x,y|z) = 1$ and we are done. So we assume $x \neq y$, then $F(x,y|z) = U(x,y|z)$ and we have

$$p^{(n)}(x,y) = \sum_{k=1}^{n} f^{(k)}(x,y)p^{(n-k)}(y,y) = \sum_{0\leq k\leq n} f^{(k)}(x,y)p^{(n-k)}(y,y)$$

for all $n \geq 0$ and then

$$G(x,y|z) = \sum_{n\geq 0} p^{(n)}(x,y)z^n = \sum_{n\geq 0}\sum_{0\leq k\leq n} f^{(k)}(x,y)z^k p^{(n-k)}(y,y)z^{n-k} = F(x,y|z)G(y,y|z)$$

which are all well-defined since $F = U$ now.

c. We know

$$u^{(n)}(x,x) = \sum_{y\in X} p(x,y)f^{(n-1)}(y,x)$$

so we know

$$\begin{aligned} U(x,x|z) &= \sum_{n\geq 1} u^{(n)}(x,x)z^n \\ &= \sum_{n\geq 1}\sum_{y\in X} p(x,y)f^{(n-1)}(y,x)z^n \\ &= \sum_{y\in X} p(x,y)zF(y,x|z) \end{aligned}$$

d. We know

$$f^{(n)}(x,y) = \sum_{w\omega\in X} p(x,\omega)f^{(n-1)}(y,x)$$

and hence

$$\begin{aligned} F(x,y|z) &= \sum_{n\geq 1} f^{(n)}(x,y)z^n \\ &= \sum_{n\geq 1}\sum_{\omega\in X} p(x,\omega)f^{(n-1)}(\omega,x)z^n \\ &= \sum_{\omega\in X} p(x,\omega)zF(\omega,x|z) \end{aligned}$$

Definition 1.7

Let Γ *be an oriented graph with vertex set* X*. For* $x, y \in X$*, a cut point between* x *and* y *is a vertex* $\omega \in X$ *such tha every path in* Γ *from* x *to* y *must pass through* ω*.*

Proposition 1.2

a. For all $x, \omega, y \in X$ and for real z with $0 \le z \le s(x,y)$ one has

$$F(x,y|z) \ge F(x,\omega|z)F(\omega,y|z)$$

b. Suppose that in the graph $\Gamma(P)$ of the Markov chain (X,P), the state ω is a cut point between x and $t \in X$. Then

$$F(x,y|z) = F(x,\omega|z)F(\omega,y|z)$$

for all $z \in \mathbb{C}$ with $|z| < s(x,y)$ and for $z = s(x,y)$.
c. x, y distinct and we will have

$$U(x,x|z) \ge F(x,y|z)F(y,x|z)$$

for $0 \le z \le s(x,x)$.

Proof a. We have

$$f^{(n)}(x,y) = P_x(s^y = n) \ge \sum_{k=0}^{n} P_x(s^y = n, s^\omega = k) = \sum_{k=0}^{n} f^{(k)(x,\omega)} f^{(n-k)}(\omega,y)$$

and let $x \to s(x,y)$ we are done.

b. The $\ge$ will be replaced with rquality and we are done.

Corollary 1.1

Suppose that ω is a cut point between x and y. Show that the expected time to reach y starting from x satisfies

$$E_x(s^y|s^y < \infty) = E_x(s^\omega|s^\omega < \infty) + E_\omega(s^y|s^y < \infty)$$

Proof Notice

$$E_x(s^y|s^y < \infty) = \sum_{n\ge 0} n \frac{P_\nu(s^y = n)}{\sum_{n\ge 0} P_\nu(s^y = n)} = F'(x,y|1-)/F(x,y|1)$$

and since

$$F'(x,y|z) = F'(x,\omega|z)F(\omega,y|z) + F(x,\omega|z)F'(\omega,y|z)$$

so we know that for any $|z| < 1$, we always have

$$F'(x,y|z)/F(x,y|z) = F'(\omega,y|z)/F(\omega,y|z) + F'(x,\omega|z)/F(x,\omega|z)$$

and then let $z \to 1-$ and we are done by observing that

$$F'(x,y|z-) = F'(x,y|z), F(x,y|1-) = F(x,y|1)$$

if the RHS' exist.

1.2 Irreducible classes

Definition 1.8

(X,P) will be a Markov chain. For $x, y \in X$ we write
a. $x \xrightarrow{n} y$ if $p^{(n)}(x,y) > 0$.
b. $x \to y$ if there is $n \ge 0$ such that $x \xrightarrow{n} y$.
c. $x \nrightarrow y$ if there is no $n \ge 0$ such that $x \xrightarrow{n} y$.
d. $x \leftrightarrow y$ if $x \to y$ and $y \to x$ and we call x and y communicate and easy to be checked an equivalence relation on X.
Then we call an equivalence class w.r.t. $\leftrightarrow$ as an irreducible class.

Lemma 1.2

We may define

$$C(x) \to C(y) \text{ iff } x \to y$$

which is well-defined and then we may know $\to$ is a partial order on the collection of all irreducible classes of (X, P).

♡

Proof Reflexivity: $x \xrightarrow{0} x$.

Transitivity and anti-symmetry are trivial.

Definition 1.9

The maximal elements (if exist) of the partial order $\to$ on the collection of the irreducible classes of (X, P) are called essential classes. A state x is essential iff $C(x)$ is essential.

♣

Proposition 1.3

Let $C \subset X$ be an irreducible class. Then the following statements are equivalent.

a. C is essential.

b. If $x \in C$ and $x \to y$ then $y \in C$.

c. If $x \in C$ and $x \to y$ then $y \in x$.

♠

Proof (a) implies (b), we know $C(x)tC(y)$ and hence $C(x) = C(y)$. (b) implies (c) easily and we consider (c) implies (a), which can be shown by if $C(x) \to C(y)$, then we know $C(x) = C(y)$ and we are done.

Theorem 1.6

We call a set $B \subset X$ convex if $x, y \in B$ and $x \to w \to y$ implies $w \in B$. For $B \subset X$ finite,convex set containing no essential elements. Then there is $\epsilon > 0$ such that for each $x \in B$ and all but finitely many $n \in \mathbb{N}$

$$\sum_{y \in B} p^{(n)}(x, y) \le (1 - \epsilon)^n$$

♡

Proof B is a disjoint union of finite nonessential irreducible classes $C(x_1), \cdots, C(X_k)$ and assume $C(x_1), C(x_2), \cdots, C(x_j)$ are the maximal elements in the partial order $\to$ restricted on $C(x_i), 1 \le i \le k$. We know there is $v_i \in X$ such that $x_i \to v_i$ but $v_i \nrightarrow x_i$ for $1 \le i \le j$ with $v_i \in X - B$. For $x \in B$, $x \to x_i$ for some i and hence $x \to v_i$ while $v_i \nrightarrow x$ for some i. So we may find m_x such that

$$\sum_{y \in B} p^{(m_x)}(x, y) < 1$$

Let $m = \max\{m_x, x \in B\}$ and $x \in B$, we know

$$\sum_{y \in B} p^{(m)}(x, y) = \sum_{y \in B} \sum_{\omega \in X} p^{(m_x)}(m_x)(x, \omega) p^{(m - m_x)}(\omega, y) < 1$$

since B is finite, there is $\kappa > 0$ such that

$$\sum_{y \in B} p^{(m)}(x, y) \le 1 - \kappa$$

let $n \ge m$ and we assume $n = km + r$ and we know

$$\sum_{y \in B} p^{(n)}(x, y) = \sum_{w \in B} p^{(km)}(x, w) = \sum_{y \in B} p^{(k-1)m} \sum_{\omega \in B} p^{(m)}(y, \omega) \le \cdots \le (1 - \kappa)^k = (1 - \epsilon)^n$$

where $\epsilon = 1 - (1 - \kappa)^{1/2m}$.

Corollary 1.2

For C finite, non-essential irreducible class. The expected number of visits C starting from $x \in C$ is finite, i.e.

$$E_x(v^C) \le 1/\epsilon + M$$

Then we may know

$$P_x(\exists k, Z_n \in C \textit{ for all } n > k) = 1$$

since $P(v^C = \infty) = 0$.

♡

Corollary 1.3

If the set of all non-essential states in X *is finite, then the Markov chain reaches some essential class with probability one:*

$$P_x(s^{X_{ess}} < \infty) = 1$$

where X_{ess} *is the union of all essential classes.*

♡

Definition 1.10

For any subset A *of* X *we denote by* P_A *the restriction of* P *to* A

$$p_A(x,y) = p(x,y) \textit{ if } x,y \in A \textit{ and } p_A(x,y) = 0$$

which is substochastic, i.e. all row sums are less than 1. *We know*

$$p_A^{(n)}(x,y) = P_x(Z_n = y, Z_k \in A, 0 \leq k \leq n)$$

and define I_A *be* I *striction on* A

$$G_A(x,y|z) = \sum_{n\geq 0} p_A^{(n)}(x,y)z^n, \quad G_A(x,y) = G_A(x,y|1)$$

let $r_A(x,y)$ *be the radius of convergence of this power series, and* $r_A = \inf\{r_A(x,y), x,y \in A\}$. *If we write* $G_A(z) = (G_A(x,y|z))_{x,y\in A}$ *then*

$$(I_A - zP)G_A(z) = I_A$$

♣

Lemma 1.3

Suppose that $A \subset X$ *is finite and that for each* $x \in A$ *there is* $\omega \in X - A$ *such that* $x \to \omega$. *Then* $r_A > 1$. *In particular,* $G_A(x,y) < \infty$ *for all* $x,y \in A$.

♡

Proof Let out be a new state and equip the state space $A \cup \{out\}$ with the transition matrix Q given by

$$q(x,y) = p(x,y), \quad q(x,out) = 1 - p(x,A) \quad q(out,out) = 1, \quad q(out,x) = 0, x,y \in A$$

Then $Q_A = P_A$ the only essential state of the Markov chain $(A \cup \{out\}, Q)$ is $\{out\}$ and A is convex, so there is ϵ such that

$$\sum_{y\in A} Q^{(n)}(x,y) = \sum_{y\in A} p_A^{(n)}(x,y) \leq (1-\epsilon)^n$$

for all $x \in A$. And hence $r_A > 1/(1-\epsilon)$ and we are done.

Definition 1.11

The period of C *is the number*

$$d = d(C) = gcd(\{n > 0, p^{(n)}(x,x) > 0\})$$

where $x \in C$. *The number* $d(C)$ *does not depend on the specific choice of* $x \in C$.

♣

Proof Let $x,y \in C, x \neq y$. We write $d(x) = gcd(\mathbb{N}_x)$ where

$$\mathbb{N}_x = \{n > 0, p^{(n)}(x,x) > 0\}$$

we know there are $k,l > 0$ such that $p^{(k)}(x,y) > 0, p^{(l)}(y,x) > 0$. We have $p^{(k+l)}(x,x) > 0$ and hence $d(x)|k+l$.

Let $n \in \mathbb{N}_y$, then $p^{(k+n+l)}(x,x) \geq p^{(k)}(x,y)p^{(n)}(y,y)p^{(l)}(y,x) > 0$ and hence $d(x)|n$. The rest is easy to be checked.

Definition 1.12

If $d(C)=1$ then C is called an aperiodic class.

Lemma 1.4

Let C be an irreducible class and $d=d(C)$. For each $x\in C$ there is $m_x\in\mathbb{N}$ such that $p^{(md)}(x,x)>0$ for all $m\geq m_x$.

Proof It is easy to check $\mathbb{N}_x$ is closed under addition, and we may find $n_1,\cdots,n_l\in N_x$ and $a_1,\cdots,a_l\in\mathbb{Z}$ such that $d=\sum\limits_{i=1}^{l}a_in_i$.

Let $n^+=\sum\limits_{a_i>0}a_in_i$ and $n^-=\sum\limits_{a_i<0}(-a_i)n_i$ and hence $n^+,n^-\in\mathbb{N}_x$ and $d=n^+-n^-$. We set $k^+=n^+/d$ and $l^-=n^-/d$. Then $k^+-k^-=1$. We define

$$m_x=k^-(k^--1)$$

and it is easy to check for any $m\geq m_x$, $md\in N_x$.

Theorem 1.7

With respect to the matrix P_C^d, the irreducible class C decomposes into $d=d(C)$ irreducible, aperiodic classes $C_0,C_1,\cdots,C_{d-1}$ which are visited in cyclic order by the original Markov chain: if $u\in C_i,v\in C$ and $p(u,v)>0$ then $v\in C_{i+1}$ where $i+1$ is computed modulo d.
Schematically,

$$C_0\overset{1}{\to}C_1\overset{1}{\to}\cdots\overset{1}{\to}C_{d-1}\overset{1}{\to}C_0\overset{1}{\to}$$

x,y belong to the same $C_i\Leftrightarrow p^{(md)}(x,y)>0$ for some $m\geq 0$.

Proof Let $x_0\in C$ and since $p^{(m_0d)}(x_0,x_0)>0$ for some $m_0>0$, there are $x_1,\cdots,x_{d-1},x_d\in C$ such that

$$x_0\overset{1}{\to}x_1\overset{1}{\to}\cdots\overset{1}{\to}x_{d-1}\overset{1}{\to}x_d\overset{(m_0-1)d}{\to}x_0$$

and define

$$C_i=\{x\in C,x_i\overset{md}{\to}x\text{ for some }m\geq 0\}$$

a. C_i is the irreducible class of x_i with respect to P_C^d.

We know $x_i\in C_i$ and if $x\in C_i$ then $x\in C$ and $x\overset{n}{\to}x_i$ and hence $x\overset{P_C^d}{\leftrightarrow}$.

Lemma 1.5

If $x\to\omega\to y$ then $r(x,y)\leq\min\{r(x,\omega,r(\omega,y))\}$. In, particular, $x\to y$ implies $\min\{r(x,x),r(y,y)\}$.

Proof By assumption, there are a$k,l\geq 0$ such that $p^{(k)}(x,\omega)>0,p^{(l)}(\omega,y)>0$. Then

$$p^{(n+l)(x,y)}\geq p^{(n)}(x,\omega)p^{(l)}(\omega,y)$$

and hence

$$(p^{(n+l)}(x,y)^{1/(n+l)})^{(n+l)/n}\geq p^{(n)}(x,\omega)^{1/n}p^{(l)}(\omega,y)^{1./n}$$

then we know $r(x,y)\leq r(x,\omega)$ by $n\to\infty$ and the other one is similar.

1.3 Recurrence

Definition 1.13

Consider a Markov chain (X, P). A state $x \in X$ is called recurrent, if

$$U(x,x) = P_x(Z_n = x \text{ for some } n > 0) = 1$$

and transient otherwise.
Also define

$$H(x,y) = P_x(Z_n = y \text{ i.o.}), x, y \in X$$

Theorem 1.8

a. The state x is recurrent iff $H(x,x) = 1$.
b. The state x is transient iff $H(x,x) = 0$.
c. We have $H(x,y) = U(x,y)H(y,y)$.

Proof Define

$$H^{(m)}(x,y) = P_x(Z_n = y \text{ for at least } m \text{ times})$$

then

$$H^{(1)}(x,y) = U(x,y), \quad H(x,y) = \lim_{m\to\infty} H^{(m)}(x,y)$$

and

$$\begin{aligned} H^{(m+1)}(x,y) &= \sum_{k\geq 1} P_x(t^y = k, Z_n = y \text{ at least } m \text{ times after } n > k) \\ &= \sum f^{(k)}(x,y)H^{(m)}(y,y) \\ &= U(x,y)H^{(m)}(y,y) \end{aligned}$$

and hence $H^{(m)}(x,x) = U(x,x)^m$. And we are done.

Theorem 1.9

a. The state x is recurrent iff $G(x,x) = \infty$.
b. If x is recurrent and $x \to y$ then $U(y,x) = H(y,x) = 1$ and y is recurrent and hence x is essential.
c. If C is a finite essential class then all elements of C are recurrent.

Proof a. By the MCT,

$$U(x,x) = \lim_{z\to 1-} U(x,x|z) \text{ and } G(x,x) = \lim_{z\to 1-} G(x,x|z)$$

then we know

$$G(x,x) = \lim_{z\to 1-} \frac{1}{1-U(x,x|z)} = \begin{cases} \infty, & U(x,x) = 1 \\ \dfrac{1}{1-U(x,x)}, & U(x,x) < 1 \end{cases}$$

b. If $x \xrightarrow{n} y$ implies $U(y,x) = 1$, then if $x \xrightarrow{n} \omega \xrightarrow{1} y$ and since

$$1 = U(\omega,x) = p(\omega,x) + \sum_{v\neq x} p(\omega,v)U(v,x)$$

and hence

$$0 = \sum_{v\neq x} p(w,v)(1-U(v,x) \geq p(w,y))(1-U(y,x)) \geq 0$$

since $p(w,y) > 0$ and we have $U(y,x) = 1$ and hence $H(y,x) = U(y,x)H(x,x) = U(y,x) = 1$ for $y, x \to y$ by induction. And hence x is essential, for $k,l \geq 0$ we have $p^{(k)}(x,y) > 0$ and $p^{(l)}(y,x) > 0$, so we know

$$G(y,y) \geq \sum_n p^{(n)}(y,y) \geq p^{(l)}(y,x) \sum_{m=0}^{\infty} p^{(m)}(x,x)p^{(k)}(x,y) = \infty$$

c. Since C is essential and we know

$$P_x(Z_n \in C \text{ for all } n) = 1$$

and we know there is at least one $y \in C$ such that $Z_n(\omega) = y$ for infinitely many n and

$$1 = P_x(\exists y \in C, Z_n = y \text{ for infinitely many } n) \leq \sum_{y\in C} P_x(Z_n = y \text{ for infinitely many } n) = \sum_{y \in C} H(x,y)$$

so there must have $y \in C$ such that $0 < H(x,y) = U(x,y)H(y,y)$ and hence $H(y,y) = 1$, so there is an element in C is recurrent and hence all elements in C are recurrent.

Definition 1.14

Let x be a recurrent state of Markov chain (X,P), then t^x is P_x-a.s. finite and then the expected return time is

$$E_x(t^x) = \sum_{n\geq 1} nu^{(n)}(x,x) = U'(x,x|1-)$$

by MCT.

A recurrent state x is called positive recurrent if $E_x(t^x) < \infty$ and null recurrent if $E_x(t^x) = \infty$. ♣

Theorem 1.10

Suppose that x is positive recurrent and that $y \leftrightarrow x$, then also y is positive recurrent. Furthermore, $E_y(t^x) < \infty$. ♡

Proof We know y is recurrent and then we know the convergence radiuses of the Green function are $r(x,x) = r(y,y) = 1$ and

$$\frac{1-U(x,x|z)}{1-U(y,y|z)} = \frac{G(y,y|z)}{G(x,x|z)} \text{ for } 0 < z < 1$$

and by l'Hospital, we may know

$$\frac{E_x(t^x)}{E_y(t^y)} = \frac{U'(x,x|1-)}{U'(y,y|1-)} = \lim_{z\to 1-} \frac{G(y,y|z)}{G(x,x|z)}$$

there are $k,l > 0$ such that $p^{(k)}(x,y) > 0$ and $p^{(l)}(y,x) > 0$. Therefore, if $0 < z < 1$,

$$G(y,y|z) \geq \sum_{n=0}^{k+l-1} p^{(n)}(y,y)z^n + p^{(l)}(y,x)G(x,x|z)p^{(k)}(x,y)z^{k+l}$$

and hence $\lim_{z\to 1-} \dfrac{G(y,y|z)}{G(x,x|z)} \geq p^{(l)}(y,x)p^{(k)}(x,y) > 0$ and hence $E_y(t^y) < \infty$.

We still use the induction to show that $E_y(t^x) < \infty$. Let ω be $x \xrightarrow{n} \omega \xrightarrow{1} y$ and we know

$$U(\omega,x|z) = p(\omega,x)z + \sum_{v\neq x} p(\omega,v)z(U,x|z)$$

for $0 < z \leq 1$ and by MCT we may know that $U'(w,x|1-)$ finite implies that $U'(y,x|1-)$ is finite.

Theorem 1.11

Let C be a finite essential class of (X,P). Then C is positive recurrent. ♡

Proof Since the matrix P^n is stochastic, we have

$$\sum_{y\in X} G(x,y|z) = \sum_{n\geq 0}\sum_{y\in X} p^{(n)}(x,y)z^n = \frac{1}{1-z}$$

for $0 \leq z < 1$. By theorem 1.5.(b) we may know

$$\sum_{y\in X} F(x,y|z)\frac{1-z}{1-U(y,y|z)} = 1 \text{ for each } x \in X \text{ and } 0 \leq z < 1$$

Now consider $x \in C$ and we know C is recurrent. So

$$1 = \lim_{z \to 1-} \sum_{y \in C} F(x, y|z) \frac{1 - z}{1 - U(y, y|z)} = \sum \frac{1}{U'(y, y|1-)}$$

so there has to be $y \in C$ positive recurrent and we are done.

Definition 1.15

A measure ν on X is called invariant or stationary if $\nu P = \nu$. It is called excessive, if $\nu P \le \nu$ pointwise. And if $\nu(X) < \infty$ and we know an excessive measure is stationary.
We say that a set $A \subset X$ carries the measure ν on X, if the support of ν is contained in A. ♣

Theorem 1.12

Let C be an essential class of (X, P). Then C is positive recurrent iff it carries a stationary probability measure m_C. In this case, the latter is unique and given by

$$m_C(x) = \begin{cases} 1/E_x(t^x), & x \in C \\ 0, & otherwise \end{cases}$$

If (X, P) is an arbitrary Markov chain and ν is a stationary probability measure, then $\nu(Y) > 0$ implies that y is a positive recurrent state. ♡

Corollary 1.4

The Markov chain (X, P) admits stationary probability measures iff there are positive recurrent states. ♡

Proof Let ν be a stationary probability measure and we know $\nu(x) > 0$ implies x is positive recurrent. So there are positive recurrent essential classes. Conversely, it is clear that any convex combination of the m_i is a stationary probability measure.

Definition 1.16

For the following problems, assume X is finite and define

$$M(X) = \{\nu : X \to |R| \nu(x) \ge 0 \, for\ all\ x \in X\ and \sum_{x \in X} \nu(x) = 1\}$$

which is all probability distributions on X, which is a subset of $l^1(X)$, and $M(x)$ is closed in the metric $||\nu_1 - \nu_2||_1 = \sum_{x \in X} |\nu_1(x) - \nu_2(x)|$ and then P acts on $M(x)$ by $\nu \mapsto \nu P$. For $y \in X$ we define

$$a(y) = a(y, P) = \inf_{x \in X} p(x, y), \quad \tau = \tau(P) = 1 - \sum_{y \in X} a(y)$$

♣

Lemma 1.6

For all $\nu_1, \nu_2 \in M(X)$, we have

$$||\nu_1 P - \nu_2 P||_1 \le \tau(P) ||\nu_1 - \nu_2||_1$$

♡

Proof For each $y \in X$ we have

$$\begin{aligned} \nu_1 P(y) - \nu_2 P(y) &= \sum_{x \in X} (\nu_1(x) - \nu_2(x)) p(x, y) \\ &= \sum_{x \in X} |\nu_1(x) - \nu_2(x)| p(x, y) - \sum_{x \in X} \Big(|\nu_1(x) - \nu_2(x)| - (\nu_1(x) - \nu_2(x)) \Big) p(x, y) \\ &\le \sum_{x \in X} |\nu_1(x) - \nu_2(x)| p(x, y) - \sum_{x \in X} \Big(|\nu_1(x) - \nu_2(x)| - (\nu_1(x) - \nu_2(x)) \Big) a(y) \\ &= \sum_{x \in X} |\nu_1(x) - \nu_2(x)| \Big(p(x, y) - a(y) \Big) \end{aligned}$$

and hence

$$|\nu_1 P(y) - \nu_2 P(y)| \leq \sum_{x \in X} |\nu_1(x) - \nu_2(x)|(p(x,y) - a(y))$$

and we have

$$||\nu_1 P - \nu_2 P||_1 \leq \tau(P)||\nu_1 - \nu_2||_1$$

Theorem 1.13

For any Markov chain (X, P) is irreducible and such that $\tau(P^k) < 1$ for some $k \in \mathbb{N}$. Then P is aperiodic, recurrent and there is $\bar{\tau} < 1$ such that for each ν

♡

1.4 The ergodic theorem

Theorem 1.14

(Ergodic theorem) Let (X, P) be a positive recurent, irreducible Markov chain with stationary probability measure m. If $f : X \to \mathbb{R}$ is m-integrable that is $\int |f|dm = \sum_x |f(x)|m(x) < \infty$, then for any starting distribution,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(Z_n) = \int_X f dm, \quad a.s.$$

♡

Definition 1.17

Here we define

$$l^{(n)}(x,y) = P_X(Z_n = y, Z_k \neq x, 1 \leq k \leq n)$$

and the generating function

$$L(x,y|z) = \sum_{n \geq 0} l^{(n)}(x,y) z^n$$

♣

Lemma 1.7

If $y \to x$ or if y is a transient state, then $L(x,y) < \infty$.

♡

Proposition 1.4

For all $x, y \in X$, we have

$$G(x,y|z) = G(x,x|z)L(x,y|z)$$
$$U(x,x|z) = \sum_y L(x,y|z)p(y,x)z$$
$$L(x,y|z) = \sum_w L(x,w|z)p(w,y)z$$

♠

Chapter 2 Potential theory

2.1 Harmonic function

Definition 2.1

Let (X,P) be a finite, irreducible Markov chain. We choose and fix a subset $X^\circ \subset X$ called the interior, and $\partial X = X - X^\circ$, we suppose X° is connected i.e. P_{X° is irreducible.
We call a function $h : X \to \mathbb{R}$ harmonic on X° if $h(x) = Ph(x)$ for every $x \in X^\circ$, where $Ph(x) = \sum_{y\in X} p(x,y)h(y)$, which is also called mean value property. We denote by $\mathcal{H}(X^\circ) = \mathcal{H}(X^\circ, P)$ is the linear space of all functions on X and harmonic on X°. ♣

Lemma 2.1

(Maximum principle) Let $h \in \mathcal{H}(X^\circ)$ and $M = \max_X h(x)$, then there is $y \in \partial X$ such that $h(y) = M$.
If h is non-constant then $h(x) < M$ for every $x \in X^\circ$. ♡

Proof Here we may know if $x \in X^\circ$ and $h(x) = M$, then choose any $y \in X$ and we have

$$\begin{aligned} M = h(x) &= p^{(n)}(x,y)h(y) + \sum_{v\neq y} p^{(n)}(x,v)h(v) \\ &\le p^{(n)}(x,y)h(y) + (1-p^{(n)}(x,y))M \end{aligned}$$

where n such that $p^{(n)}(x,y) > 0$ and hence $h(y) = M$, which means h is then constant. And we are done.

Definition 2.2

Let $s = s^{\partial X}$, then $P_x(s^{\partial X} < \infty) = 1$ for any $x \in X$.
Then we may define

$$\nu_x(y) = P_x(s < \infty, Z_s = y), y \in \partial X$$

and then ν_x will become a probability distribution on ∂X, called the hitting distribution of ∂X. ♣

Proof Here we introduce $\tilde{P}$ which is defined by $\tilde{p}(x,y) = p(x,y)$ for $x \in X^\circ$ and $\tilde{p}(x,y) = \delta_x$ for $x \in \partial X$, then it is easy to check $h \in \mathcal{H}(X^\circ, P)$ iff $h \in \mathcal{H}(X^\circ, \tilde{P})$ and s is the same on (X,P) and $(X,\tilde{P})$. So consider s on $(X,\tilde{P})$, we know

$$P(s^{\partial X} < \infty) = 1$$

by corollary 1.3.

Theorem 2.1

(Solution of the Dirichlet problem) For every function $g : \partial X \to \mathbb{R}$ there is a unique function $h \in \mathcal{H}(X^\circ, P)$ such that $h(y) = g(y)$ for all $y \in \partial(X)$ which is given by

$$h(x) = \int_{\partial X} g d\nu_x$$

♡

Proof We firstly prove that the uniqueness of the solution, if $h, h' \in \mathcal{H}(X^\circ, P)$, then we know $h - h'$ should be the solution of the Dirichlet problem when $g = 0$ and by the maximum principle, we know $h - h' \le 0$ and $h' - h \le 0$ and we know $h = h'$.

Now we prove the existence of h, firstly we would like to show that $x \mapsto \nu_x(y)$ is harmonic, since

$$\begin{aligned}\sum_{v\in X} p(x,v)\nu_v(y) &= \sum_{v\in X} p(x,v)P_v(s<\infty, Z_s=y)\\ &= \sum_{v\in X} p(x,v)P_x(s<\infty, Z_s=y|Z_1=v)\\ &= \sum_{v\in X} P_x(s<\infty, Z_s=y, Z_1=v)\\ &= \nu_x(y)\end{aligned}$$

and hence $h=\int_{\partial x} g d\nu_x$ is actually a combination of harmonic functions with $h(y)=g(y)$ for $y\in\partial X$.

Definition 2.3

For a general finite Markov chain, we define the linear space of harmonic functions on X with

$$\mathcal{H}=\mathcal{H}(X,P)=\{h: X\to\mathbb{R}, h(x)=Ph(x), x\in X\}$$

♣

Theorem 2.2

Let (X,P) be a finite Markov chain, and denote its essential classes by $C_i, i\in I=\{1,\cdots,m\}$.
a. If h is harmonic on X, then h is constant on each C_i.
b. For each function $g: I\to\mathbb{R}$ there is a unique function $h\in\mathcal{H}(X,P)$ such that for all $i\in I$ and $x\in C_i$ one has $h(x)=g(i)$.

♡

Proof a. We know for any $x\in C_i, x\to y$ iff $y\in C_i$ and then if $M_i=\max_{C_i} h=h(x), x\in C_i$, then for any $y\in C_i$, we know

$$h(x)=\sum_{y\in X} p^{(n)}(x,y)h(y)\le \sum_{v\in C_i, v\neq y} p^{(n)}(x,y)M+p^{(n)}(x,y)h(y)$$

for any $n, y\in C_i$ and we are done.

b. Let prove the uniqueness at first, if h, h' are harmonic functions on X, then assume $M=\max_X h$ and be obtained at $x\in X-X_{ess}$, then we know since $P_x(s<\infty)$ by corollary 1.3. where $s=s^{X_{ess}}$, then there will be an $y\in X_{ess}$ such that

$$M=h(x)\le p^{(n)}(x,y)h(y)+(1-p^{(n)}(x,y))M$$

and hence the maximum has to be obtained at X^{ess} and the rest is easy to be checked.

Now we define $\nu_x(i)=P_x(s<\infty, Z_s\in C_i)$ which will be an harmonic function since

$$\sum_{y\in X} p(x,y)P_y(s<\infty, Z_s\in C_i)=\nu_x(i)$$

and it is easy to check that

$$h(x)=\sum_{i\in I} g(i)\nu_x(i)$$

will be a solution.

2.2 Infinite cases

In the section we assume P is irreducible on X.

Definition 2.4

All functions $f: X\to\mathbb{R}$ are assumed to be P-integrable (which is a subspace) i.e.

$$\sum_{y\in X} p(x,y)|f(y)|<\infty$$

for all $x\in X$.

A real function h on X is called harmonic if $h(x) = Ph(x)$ and superharmonic if $h(x) \geq Ph(x)$ for every $x \in X$. Addition to $\mathcal{H}$, we define

$$\mathcal{H}^+ = \{h \in \mathcal{H}, h(x) \geq 0\} \quad \mathcal{H}^\infty = \{h \in \mathcal{H}, h \text{ is bounded on } X\}$$

and lett $\mathcal{S} = \mathcal{S}(X, P)$ the space of all superharmonic functions and similarly $\mathcal{S}^+, \mathcal{S}^\infty$ ♣

Lemma 2.2

(Maximum principle) (Assume $|X| > 1$) If $h \in \mathcal{H}(X, P)$ and there is $x \in X$ such that $h(x) = M = \max_X h$, then h is constant, where P is substochatic. Furthermore, if $M \neq 0$ then P is stochastic. ♡

Proof We still have

$$M \leq \sum_{y \neq x'} p^{(n)}(x, y)M + p^{(n)}(x, x')h(x') \leq (1 - p^{(n)}(x, x'))M + p^{(n)}(x, x')h(x')$$

and hence $h = M$, if $M \neq 0$. we know the equality has to be reached by P is stochastic.

Lemma 2.3

a. If $h \in \mathcal{S}^+$ then $P^n h \in S^+$ for each n, and either $h = 0$ for $h > 0$.
b. If $h_i, i \in I$ is a family of superharmonic functions and $h(x) = \inf_I h_i(x)$ defines a P-integrable function if I is finite or h_i is bounded below, then also h is superharmonic. ♡

Proof a. Firstly, the P-integrability of h implies that of Ph since

$$\sum_{y \in X} p(x, y)|Ph(y)| \leq \sum_{y \in X, w \in X} p(x, y)|h(y)| < \infty$$

and by induction $P^n h \in \mathcal{S}^+$, and it is easy to check that $P^n h \leq h$ by $f \geq g$ imples $Pf \geq Pg$, for each 0 and so if $h(x) = 0$ for some x, then h will be 0.

b. We know $Ph \leq Ph_i \leq h_i$ implies $Ph \leq h$.

For the P-integrability, we may use the MCT for the first cases for h^- and Fatou for h^+. On the other case h^- is easier.

Lemma 2.4

If (X, P) is transient, then for each $y \in X$, the function $G(\cdot, y)$ is superharmonic and positive. There is at most one $y \in X$ for which $G(\cdot, y)$ is a constant function. If P is stochastic, then $G(\cdot, y)$ is non-constant for every y. ♡

Proof We know

$$PG(x, y) = \sum_{w \in X} p(x, w)G(w, y) = G(x, y)$$

and

$$PG(y, y) = \sum_{w \in X} p(y, w)G(w, y) = G(y, y) - 1$$

and hence $G(\cdot, y) \in \mathcal{S}^+$. Suppose $y_1, y_2 \in X$ and $y_1 \neq y_2$ such that $G(\cdot, y_i)$ are constant, then

$$F(y_1, y_2) = G(y_1, y_2)/G(y_2, y_2) = 1, F(y_2, y_1) = 1$$

and then $F(y_1, y_1) \geq F(y_1, y_2)F(y_2, y_1) \geq 1 = 1$ and y_1 is recurrent, which is a contradiction.

If P is stochastic, since $G(\cdot, y)$ is strictly superharmonic and there will be a contradiction since constant function is harmonic.

Theorem 2.3

(X, P) is recurrent iff every nonnegative superharmonic function is constant. ♡

Proof (Here notice (X, P) is either transient or recurrent since it is irreducible).

a. Suppose that (X, P) is recurrent, we show that $\mathcal{S}^+ = \mathcal{H}^+$, let $h \in \mathcal{S}^+$, we have

$$g = h - Ph$$

is non-negative and P-integrable. We have

$$\sum_{k=0}^{n} P^k g = h - P^{n+1}(x)$$

If $g(y) > 0$ for some y, then

$$\sum_{k=0}^{n} p^{(k)}(x, y)g(y) \leq \sum_{k=0}^{h} P^k g(x) \leq h(x)$$

and then we have

$$G(y, y) \leq h(y)/g(y) < \infty$$

which is a contradiction since y is recurrent. So $g = 0$ and hence h is harmonic.

Then consider for any $h \in \mathcal{S}^+ = \mathcal{H}^+$, let $x, y \in X$ and define $g(v) = \min_{h(v),h(x)}$, then we know

$$Pg(y) = \sum_{x \in X} p(y, x)g(x) \leq Ph(y)$$

if $h(y) \leq h(x)$ and the RHS is less than $h(x)$ since P is substochastic, so g is subharmonic and hence harmonic, then g should be constant and hence for any $y \neq x$ $h(y) \geq h(x)$ and then we know h is constant.

b. If (X, P) is transient, then since all the superharmonic functions are constant, then it has to be $|X| = 1$ which is a contradiction.

Definition 2.5

Here we assume the invariant measure must satisfy nonnegative and

$$vP(y) = \sum_{x \in X} v(x)p(x, y) < \infty$$

Recall we call a measure on X is invariant or stationary if $v = vP$ and excessive or superinvariant $v = vP$. We denote $I^+ = I^+(X, P)$ and $E^+ = E^+(X, P)$ the cones of all invariant and excessive measures. ♣

Proposition 2.1

a. If $v \in E^+$ then $vP^n \in E^+$ for each n and either $v = 0$ or $v(x) > 0$ for every x.

b. If $v_i, i \in I$ is a family of excessive measures, then also $v(x) = \inf_I v_i(x)$ is excessive.

c. If (X, P) is transient, then for each $x \in X$, the measure $G(x, \cdot)$ defined by $y \mapsto G(x, y)$ is excessive. ♠

Proof a. Here we know

$$vP^{(n)}(x) = \sum_{y \in X} p^{(n)}(y, x)v(y) \leq v(x)$$

and hence if $v(x) = 0$, then $v(y) = 0$ since (X, P) irreducible.

b. $vP \leq v_i P \leq v_i$.

c. A ‘’

2.3 Induced Markov chains

Definition 2.6

Suppose (X, P) is irreducible and substochastic. Let $A \subset X$ and we may define

$$p^A(x, y) = P_x(t^A < \infty, Z_{t^A} = y)$$

where $p^A(x, y) = 0$ if $y \notin A$. Then we may know $P^A = (p^A(x, y))$ is substochastic and (A, P^A) is called the Markov chain induced by (X, P) on A.

Here the irreducibility of (X, P) implies irreducibility of the induced chain. ♣

Proof For $x, y \in A$ there are $n > 0$ and $x_1, \cdots, x_{n-1} \in X$ such that $p(x, x_1)p(x_1, x_2)\cdots p(x_{n-1}, y) > 0$ and let $x_{i_k} \in A$ and we know $p^A(x_{i_k}, x_{i_{k+1}}) \leq p^A(x_{i_k}, x_{i_{k+1}})$.

Definition 2.7

If P^A is tochastic, then we call A is recurrent for (X, P) ♣

Lemma 2.5

If A is recurrent for (X, P) then

$$P_x(t^A < \infty) = 1, \textit{ for all } x \in X$$

♡

Proof We know

$$P_x(t^A < \infty) = \sum_{y \in A} p(x, y) + \sum_{y \in X-A} p(x, y)P_y(t^A < \infty)$$

If we have $P_y(t^A < \infty) = 1$, then we know $h(x) = P_x(t^A < \infty)$ and hence to be a constant on (X, P).

Theorem 2.4

If $A \subset B \subset X$, then $(P^B)^A = P^A$. ♡

Proof We should give an interpretation of Z_n^B and define $w_N^B(\omega) = k$ if $n \leq v^B(\omega)$ and k is the instant of the n-th return visit to B, then $Z_n^B = Z_{w_n^B}$ if $n \leq v^B$.

Let t_B^A be the stopping time of the first visit of (Z_n^B) in A. Since $A \subset B$, we have for any $\omega \in \Omega$, $t^A(\omega) = \infty$ iff $t_B^A(\omega) = \infty$ and $t^A(\omega) \geq t^B(\omega)$. Hence, if $t^A(\omega) < \infty$, we know

$$Z_{t_B^A(\omega)^B}(\omega) = Z_{t^A(\omega)}(\omega)$$

so for $x, y \in A$, we have

$$(p^B)^A(x, y) = P_x(t_B^A < \infty, Z_{t_B^A}^B = y) = P_x(t^A < \infty, Z_{t^A} = y) = p^A(x, y).$$

by consider ω.

Definition 2.8

For $A, B \subset X$, define the restriction of P to $A \times B$ by $P_{A,B} = (p(x, y))_{x \in A, y \in B}$. ♣

Lemma 2.6

$P^A = P_A + P_{A,X-A}G_{X-A}P_{X-A,A}$ ♡

Proof Notice for $x, y \in A$, we have

$$p^A(x, y) = p(x, y) + \sum_{v \in X-A} p(x, v)P_v(t^A < \infty, Z_{t^A} = y)$$

and then

$$P_v(t^A < \infty, Z_{t^A} = y) = \sum_{w \in X-A} P_v(t^A < \infty, Z_{t^A - 1} = \omega, Z_{t^A} = y)$$

$$\sum_{w \in X-A} \sum_{n \geq 1} P_v(t^A = n, Z_{n-1} = w, Z_n = y)$$

$$= \sum_{w \in X-A} G_{X-A}(v, w)p(w, y)$$

and we have

$$p^A(x, y) = p(x, y) + \sum_{v \in X-A} \sum_{w \in X-A} p(x, v)G_{X-A}(v, w)p(w, y)$$

Theorem 2.5

Let $v \in E^+(X,P)$, $A \subset X$ and v_A the restriction of v to A. Then $v_A \in E^+(A, P^A)$. ♡

Proof For $x \in A$, then

$$v_A(x) = v(x) \geq vP(x) = v_A P_A(x) + v_{X-A} P_{X-A,A}(x)$$

and hence

$$v_A \geq v_A P_A + v_{X-A} P_{X_A,A}$$

and similarly

$$v_{X-A} \geq v_{X-A} P_{X-A} + v_A P_{A,X-A}$$

and multiply $\sum\limits_{k=0}^{n-1} P^k_{X-A}$ to RHS and we obtain

$$v_{X-A} \sum_{k=0}^{n-1} P^k_{X-A} \geq v_{X-A} P^n_{X-A} + v_A P_{A,X-A} (\sum_{k=0}^{n-1} P^k_{X-A})$$

and hence

$$v_{X-A} \geq v_A P_{A,X-A} (\sum_{k=0}^{n-1} P^k_{X-A})$$

for every $n \geq 1$. And we know

$$v_A P_{A,X-A} (\sum_{k=0}^{n-1} P^k_{X-A}) \to v_A P_{A,X-A} G(X-A)$$

since $I/(I - P_{X-A}) = G(X-A)$ and then

$$v_A \geq v_A P_A + v_A P_{A,X-A} G(X-A) P_{X-A,A} = v_A P^A$$

2.4 Potentials, Riesz decomposition

Definition 2.9

For this section, we assume (X,P) is irreducible and transient, which means

$$0 < G(x,y) < \infty$$

for all $x, y \in X$.

A G-integrable function $f : X \to \mathbb{R}$ is one that satisfies $\sum\limits_y G(x,y)|f(y)| < \infty$ for each $x \in X$. In this case, $g(x) = Gf(x) = \sum\limits_{y \in X} G(x,y) f(y)$ is called the potential of f, while f is called the charge of g. The support of f is $\{x \in X, f(x) \neq 0\}$.

We may know $(I - G)^{-1}$ convergent. ♣

Lemma 2.7

a. If g is the potential of f, then $f = (I-P)g$. Furthermore, $P^n g \to 0$ pointwise.

b. If f is non-negative, then $g = Gf \in \mathcal{S}^+$ and g is harmonic on $X - supp(f)$ that is $Pg(x) = g(x)$ for every $x \in X - supp(f)$. ♡

Proof a. Suppose that $f \geq 0$ firstly, then we know

$$PGf(x) = \sum_{y \in X} p(x,y) \sum_{w \in X} G(w,y) f(y) = GPf = \sum_{n \geq 1} P^n f = Gf - f$$

since

$$Gf = \sum_{y \in X} \sum_{n \geq 0} P^{(n)}(x,y) f(y) = \sum_{n \geq 0} P^n f$$

by MCT. And hence Gf is superharmonic and harmonic on $X - supp(f)$. Then notice

$$P^n g(x) = GP^n f(x) = \sum_{k=n}^{\infty} f(x)$$

has to be convergent to 0. For general f, decompose it as f^+ and f^- will be fine.

Theorem 2.6

(Riesz decomposition theorem) If $u \in \mathcal{S}^+$ then there are a potential $g \in Gf$ and a function $h \in \mathcal{H}^+$ such that

$$u = Gf + h$$

The decomposition is unique.

Proof Since $u \geq 0$ and $u \geq u$, for every $x \in X$ and every $n \geq 0$, we know

$$P^n u(x) \geq P^{n+1} u(x) \geq 0$$

Therefore, there is the limit function

$$h(x) = \lim_{n\to\infty} P^n u(x)$$

where

$$Ph(x) = P(\lim_{n\to\infty} P^n u)(x) = \lim_{n\to\infty} P^{n+1} u(x) = h(x)$$

by DCT since u is P-integrable. Then let $f = u - Pu$ and then we know

$$u - h = Gf$$

Then let us prove the uniqueness, we consider $u = g_1 + h_1$ another decomposition, then $P^n = P^n g_1 + h_1$ and then we know $P^n u \to h_1$ since $P^n g_1 \to 0$ and we are done.

Corollary 2.1

a. If g is a non-negative potential then the only funciton $h \in \mathcal{H}^+$ with $g \geq h$ is $h = 0$.
b. If $u \in \mathcal{S}^+$ and there is a potential $g = Gf$ with $g \leq u$, then u is the potential of a non-negative function.

Proof a. $h = P^n h \leq P^n g \to 0$ pointwise.
b. Trivial.

Theorem 2.7

(Approximation theorem) If $h \in \mathcal{S}^+(X, P)$ then there is a sequence of potentials $g_n = Gf_n, f_n \geq 0$ such that $g_n(x) \leq g_{n+1}(x)$ for x and n, and

$$\lim_{n\to\infty} g_n(x) = h(x)$$

Notice here we do not use that h is G-integrable.

Proof Define

$$R^A[h](x) = \inf\{u(x), u \in \mathcal{S}^+, u(y) \geq h(y) \text{ for all } y \in A\}$$

and $R^A[h] \leq h$. In particular, we have

$$R^A[h](x) = h(x)$$

for $x \in A$. And by lemma 2.3. we know $R^A[h](x) \in \mathcal{S}^+$. Let A be a finite subset X. Let $f_0(x) = h(x)$ if $x \in A$ and $f_0(x) = 0$. f_0 is non-negative and finitely supported. Then Gf_0 exists and finite on X, with $Gf_0 \geq f_0$. So Gf_0 is a superharmonic function since $PGf_0 = GPf_0 \leq Gf_0$ and with $Gf_0 \geq h$ on A. So we know $R^A[h] \leq Gf_0$.

So we know $R^A[h]$ has to be a potential and then let B be another finite subset of X containing A. Then $R^B[h] \geq R^A[h]$. Let A_n be an increasing sequence of finite subsets of X such that $X = \bigcup_n A_n$ and let $g_n = R^{A_n}[h]$ then we know $g_n \leq h$ but $g_n = h$ on A_n.

Definition 2.10

For $A \subset X, x, y \in X$, we define

$$F^A(x,y) = \sum_{n=0}^{\infty} P_x(Z_n = y, Z_j \notin A \, for \, 0 \le j < n)\chi_A(y)$$

and

$$L^A(x,y) = \sum_{n=0}^{\infty} P_x(Z_n = y, Z_j \notin A \, for \, 0 < j \le n)\chi_A(x)$$

And for P and an excessive measure v, define the v-reversal $\hat{P}$ of P as (to secure $\hat{p}$ is substochastic)

$$\hat{p}(x,y) = v(y)p(y,x)/v(x)$$

♣

Proposition 2.2

a. We have

$$\hat{L}^A(x,y) = \frac{v(y)F^A(y,x)}{v(x)}, \quad \hat{F}^A(x,y) = \frac{v(y)L^A(y,x)}{v(x)}$$

b. $x \in A \implies F^A(x,\cdot) = \delta_x, y \in A \implies L^A(\cdot, y) = 1_y$.

♠

Proof a. We have

$$\begin{aligned}
\hat{L}^A(x,y) &= \sum_{n\ge 0}\sum \hat{P}_x(Z_n = y, Z_j = x_j, 0 \le j < n)\chi_A(x) \\
&= \sum_{n\ge 0}\sum v(y)p(y,\cdot)\cdots p(\cdot,x)/v(x) \\
&= v(y)\sum_{n\ge 0} P_y(Z_n = x, Z_j \notin A)\chi_A(x)/v(x) \\
&= v(y)F^A(y,x)/v(x)
\end{aligned}$$

and the rest is similar.

b. $x \in A$, then $F^A(x,y) = P_x(Z_0 = y)$. And the other one is similar.

Lemma 2.8

a. $G = G_{X-A} + F^A G$.
b. $G = G_{X-A} + GL^A$.
c. $F^A G = GL^A = G - G_{X-A}$.

♡

Proof We know

$$\begin{aligned}
p^{(n)}(x,y) &= P_x(Z_n = y, s^A > n) + P_x(Z_n = y, s^A \le n) \\
&= p^{(n)}_{X-A}(x,y) + \sum_{v\in A}\sum_{k=0}^{n} P_x(Z_n = y, s^A = k, Z_k = v) \\
&= p^{(n)}_{X-A}(x,y) + \sum_{v\in A}\sum_{k=0} P_x(s^A = k, Z_k = v)p^{(n-k)}(v,y)
\end{aligned}$$

then we have

$$G(x,y) = G_{X-A}(x,y) = \sum_{v\in A}(\sum_{k=0}^{\infty} P_x(s^A = k, Z_k = v))(\sum_{n=0}^{\infty} p^{(n)}(v,y))$$

and hence

$$G(x,y) = G_{X-A}(x,y) + \sum_{v\in X} F^A(x,v)G(v,y)$$

The rest is to enumerate the last time of visiting A.

Lemma 2.9

$P^A = P_{A,X}F^A = L^A P_{X,A}$.

Proof We know

$$\begin{aligned} p^A(x,y) &= p(x,y) + \sum_{v\in X-A} p(x,v)P_v(s^A < |infty, Z_{s^A} = y) \\ &= \sum_{v\in A} p(x,v)\delta_v(y) + \sum_{v\in X-A} p(x,v)F^A(v,y) \\ &= \sum_{v\in X} p(x,v)F^A(v,y) \end{aligned}$$

Then let $v = 1$ and we have

$$p^A(x,y) = \hat{p}(y,x) = \sum_{v\in X} \hat{p}(y,v)\hat{F}(v,x) = \sum_{v\in X} L(x,v)p(v,y)$$

and we are done.(Ensured by proposition 2.1. c)

Lemma 2.10

a. If $h \in \mathcal{S}^+(X,P)$, then $F^Ah(x) = \sum_{y\in A} F^A(x,y)h(y)$ if finite and

$$F^Ah(x) \leq h(x)$$

b. If $v \in E^+(X,P)$, then $vL^A(y) = \sum_{x\in A} v(x)L^A(x,y)$ is finite and

$$vL^A(y) \leq v(y)$$

Proof By approximation theorem, we may find $g_n = Gf_n$ such that $g_n \uparrow h$ on X. The f_n can be chosen to have finite support. So

$$F^A g_n = F^A G f_n = G f_n - G_{X-A} f_n \leq g_n \leq h$$

and hence $F^Ah \leq h$ by MCT.

For the other conclusion, we know

$$vL^A(y) = \sum_{x\in A} v(x)L^A(x,y) = \sum_{x\in A} \hat{F}^A(y,x)v(y) \leq v(y)$$

Definition 2.11

Reduced measure on A

$$R^A[v](x) = \inf\{\mu \in E^+, \mu(y) \geq v(y), y \in A\}$$

Theorem 2.8

a. If $h \in \mathcal{S}^+$ then $R^A[h] = F^Ah$. In particular, $R^A[h]$ is harmonic in every point of $X - A$ while $R^A[h] = h$ on A.

b. If $v \in E^+$ then $R^A[v] = vL^A$. In particular, $R^A[v]$ is invariant in every point of $X - A$ while $R^A[v] = v$ on A.

Proof a. For $x \in X - A$ and $y \in A$, we factorize and then

$$F^A(x,y) = p(x,y) + \sum_{v\in X-A} p(x,v)F^A(v,y) = \sum_{v\in X} p(x,v)F^A(v,y)$$

then

$$F^Ah(x) = \sum_{y\in A} F^A(x,y)h(y) = \sum_{v\in X, y\in X} p(x,v)F^A(v,y)h(y) = P(F^Ah)(x)$$

then for $x \in A$

$$P(F^Ah)(x) = \sum PF^A(x,y)h(y) = P^Ah(x) \leq h(x)$$

and it is easy to check $F^Ah = h$ on A. sp we know $F^A \in \{u \in \mathcal{S}^+, u \geq h, y \in A\}$ then $R^A[h] \leq F^Ah$. Then for

$u \in Sar^+$ and $u \geq h$ on A, we know

$$u(x) \geq \sum_{y \in A} F^A(x,y)u(y) \geq F^A h(x)$$

and we are done.

b. For $x \in X$ we have $L^A(x,y) = 0$ and then

$$vL^A P(y) = \sum_{x \in A, w \in A} v(x)L^A(x,w)P(w,y) = \sum_{x \in A} v(x)L^A P(x,y) = vP^A \leq v(y)$$

for $y \in A$ and for $x \in X - A$, we have

$$vL^A P(x) = \sum_{y \in A, w \in A} v(y)L^A(y,w)P(w,x) = 0 = vL^A(x)$$

and then since $vL^A(y) = v(y)$ for all $y \in A$, so we are done.

Definition 2.12

Define the potential of an excessive measure v by vG.
If f is a non-negative G-integrable function on X, then the balayee of f is the function $f^A = L^A f$.
If μ is a non-negative, G-integrable measure on X, then the balayee of μ is the measure $\mu^A = \mu F^A$. ♣

Theorem 2.9

Let f be a non-negative, G-integrable function on X with support A. If $h \in \mathcal{S}^+$ is such that $h(x) \geq Gf(x)$ for every $x \in A$, then $h \geq Gf$ on the whole of X. ♡

Proof We know

$$h(x) \geq F^A h(x) \geq \sum_{y \in A} F^A(x,y)Gf(y) = F^A Gf(x) = Gf^A(x) = Gf(x)$$

for every x since $f^A = f$.