STA447/STA2006 Stochastic Processes

Gun Ho Jang

Lecture on January 23, 2014

Note

This lecture note is prepared for the course STA447/STA2006 Stochastic Processes. This lecture note may contain flaws. Please consult text book or reference books for confidence.

- Gun Ho Jang

* indicates graduate level. So you may skip those parts.

2.3 Stationary Distribution

Definition 25. A stochastic process X_t is said to be *stationary* if $\{X_t\}$ and $\{X_{t+s}\}$ have the same distribution for any $s \geq 0$.

A (homogeneous) Markov chain X_t can be stationary if X_0 and X_1 have the save distribution. If X_0 and X_1 have the same distribution, then all X_t have the same distribution. For any fixed s. Let $T = s$ be a stopping time. X_0 and X_T have the same distribution and strong Markov property shows $\{X_t\}$ and $\{X_{T+t}\}$ have the same distribution.

Definition 26. A distribution π is called a *stationary distribution* if $\pi p = \pi$ so that $X_0 \equiv^d X_1$.

Example 28 (Two state Markov chain)**.**

$$\begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = \begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix}$$

Solves $\pi_1 = b/(a+b), \pi_2 = a/(a+b)$.

Example 29 (Weather chain)**.** Applying two state Markov chain for

$$\begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix} = \begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix}$$

we get $\pi_1 = 0.2/(0.4+0.2) = 1/3$ and $\pi_1 = 0.4/(0.4+0.2) = 2/3$.

Theorem 37. If a $k \times k$ transition matrix p is irreducible, then there exists a unique solution to $\pi p = \pi$ with $\sum_x \pi_x = 1$ and $\pi_x > 0$ for all $x \in S$.

Proof. Since the rank of $p - I$ is at most $k-1$, there exists a solution ν satisfying to $\nu p = \nu$. Let $r = [(I+p)/2]^{k-1}$. Then $\nu(I+p)/2 = \nu$ implies $\nu r = \nu$. For any x, y, there exists $p^{(l)}(x,y) > 0$ with $l \leq k-1$. Thus $r(x,y) > 0$.

Suppose there are two different signs among ν_x. Then $|\nu_y| = |\sum_x \nu_x r(x,y)| < \sum_x |\nu_x| r(x,y)|$ and $\sum_y |\nu_y| < \sum_y \sum_x |\nu_x| r(x,y) = \sum_x |\nu_x|$. It contradicts. Thus $\nu_x \geq 0$ for all x. The fact $\nu_y = \sum_x \nu_x r(x,y)$ implies $\nu_x > 0$. If there exists another solutions w, we can make a new solution $w' = aw + b\nu$ so that $\sum_x w'_x \nu_x = 0$. But both w' and ν are positive. Therefore the solution is unique. $\square$

Example 30 (Social mobility)**.** Let X_n be a familys social class in the n-th generation. States are 1: lower, 2: middle, or 3: upper. The transition probability is

	1	**2**	**3**
1	0.7	0.2	0.1
2	0.3	0.5	0.2
3	0.2	0.4	0.4

The stationary distribution $\pi = (\pi_1, \pi_2, \pi_3)$ satisfies $\pi\mathbf{1} = 1$ and $\pi p = \pi$, that is,

$$0.7\pi_1 + 0.3\pi_2 + 0.2\pi_3 = \pi_1, 0.2\pi_1 + 0.5\pi_2 + 0.4\pi_3 = \pi_2, 0.1\pi_1 + 0.2\pi_2 + 0.4\pi_3 = \pi_3, \pi_1 + \pi_2 + \pi_3 = 1.$$

The equations solve $3\pi_1 = 3\pi_2 + 2\pi_3$, $2\pi_1 = 5\pi_2 - 4\pi_3$and $\pi_1 = 1 - \pi_2 - \pi_3$. From the first two equations, $9\pi_2 = 16\pi_3$ and $\pi_1 = (22/9)\pi_3$. The last equation gives $(22/9 + 16/9 + 1)\pi_3 = 1$. Thus $\pi_3 = 9/47$ and $\pi = (22/47, 16/47, 9/47)$.

Example 31. The following function computes the stationary distribution of closed and irreducible Markov chain having finite state space.

```
solve_stationary <- function(p) {
k <- nrow(p);
pi <- - solve(t(p)[1:(k-1),1:(k-1)]-diag(k-1), t(p)[1:(k-1),k]);
pi <- c(pi,1); pi <- pi/sum(pi);
return(pi);
}
## Example
p <- matrix(c(0.7,0.3,0.2,0.2,0.5,0.4,0.1,0.2,0.4),3,3);
solve_stationary(p);
# [1] 0.4680851 0.3404255 0.1914894
```

2.4 Periodicity

Definition 27. The *period* of a state x is the greatest common divisor (g.c.d.) of n's with $p^{(n)}(x, x) > 0$. A Markov chain X_t is said to be *aperiodic* if all states have period 1.

Note. Notation: Let $I_x = \{n \geq 1 : p^{(n)}(x, x) > 0\}$.

Example 32 (Ehrenfest chain)**.** If $N = 2$, then the transition matrix is

	0	**1**	**2**	**3**
0	0	1	0	0
1	1/3	0	2/3	0
2	0	2/3	0	1/3
3	0	0	1	0

It is easy to see that $I_x = \{2, 4, 6, \dots\}$ for all $x = 0, 1, 2, 3$. Hence all states have period 2.

Proposition 38. *I_x is closed under addition, that is, $i, j \in I_x$ implies $i + j \in I_x$.*

Proof. By the definition, $i, j \in I_x$ implies $p^{(i)}(x, x), p^{(j)}(x, x) > 0$. Hence $p^{(i+j)}(x, x) \geq p^{(i)}(x, x)p^{(j)}(x, x) > 0$ and $i + j \in I_x$. $\square$

Example 33. If $p(x, x) > 0$, then $1 \in I_x$, $2 = 1 + 1 \in I_x$, $3 = 2 + 1 \in I_x$, $\dots$. Hence $I_x = \{1, 2, \dots\}$ and the period of x is 1. If $p(x, x) > 0$ for all x, then X_t is aperiodic.

Proposition 39. *If state x has period $d > 0$, then there exists $n_0 \geq 1$ such that $p^{(nd)}(x, x) > 0$ for all $n \geq n_0$.*

Proof. Note that if g.c.d. of I_x is d, then there exist a positive integer l, $i_1, \dots, i_l \in I_x$ and l integers $\alpha_1, \dots, \alpha_l$ such that $\alpha_1 i_1 + \cdots + \alpha_l i_l = d$. Let $j_m = i_m/d$, $p_m = \max(0, \alpha_m)$, $m_m = \max(0, -\alpha_m)$ where p_m and m_m are positive and negative part of integer α_m. It is easy to see that $p_1 j_1 + \cdots + p_l j_l = m_1 j_1 + \cdots + m_l j_l + 1$. Let $k = m_1 j_1 + \cdots + m_l j_l$. Then $kd, (k+1)d \in I_x$.
Claim: Let $J_x = \{i/d : i \in I_x\}$. If $k, k+1 \in J_x$, then $n \in J_x$ for all $n \geq k^2 - k$.
Fix $n \geq k^2 - k$. Let b be the remainder of n divided by k and a be the largest integer not bigger than n/k, that is, $n = ak + b$ where $a \geq k-1, 0 \leq b < k$. Then we can write $n = ak + b = (a-b)k + b(k+1)$. Since J_x is closed under addition, $n \in J_x$. □

Proposition 40. If x and y are mutually communicate, that is, $x \to y$ and $y \to x$, then x and y have the same period.

Proof. Let c, d be the periods of x and y. From $\rho_{xy}, \rho_{yx} > 0$, there exists $k, l > 0$ such that $p^{(k)}(x,y), p^{(l)}(y,x) > 0$. Then, $p^{(k+m)}(x,x) = p^{(k)}(x,y)p^{(l)}(y,x) > 0$ $p^{(k+m)}(y,y) = p^{(l)}(y,x)p^{(k)}(x,y) > 0$ imply $k+l$ is multiple of both c and d. For any $m > 0$ with $p^{(m)}(y,y) > 0$, we get $p^{(k+l+m)}(x,x) \geq p^{(k)}(x,y)p^{(m)}(y,y)p^{(l)}(y,x) > 0$. Hence $k+l+m \in I_x$ and should be multiple of c, that means, l is a multiple of c. Hence d is a multiple of c. By changing x and y, c is multiple of d. Hence c and d are the same. □

2.5 Limit Behaviour

Proposition 41. (a) If a state x is transient, then $p^{(n)}(y,x) \to 0$ for all y.
(b) If π is a stationary distribution, then $\pi(x) = 0$ for any transient state x.

Proof. (a) If x is transient, then $\infty > \mathbb{E}_y N_x = \sum_{n=1}^{\infty} p^{(n)}(y,x)$. Hence $p^{(n)}(y,x) \to 0$.
(b) From $\pi = \pi p^n$, we get $\pi(x) = \sum_{y \in \mathcal{S}} \pi(y) p^{(n)}(y,x)) = \sum_{y \in \mathcal{S}} \pi(y)(1/n) \sum_{k=1}^{n} p^{(k)}(y,x) \to \sum_{y \in \mathcal{S}} \pi(y) \times 0 = 0$. Hence $\pi(x) = 0$ for all transient state x. □

Note (Cesàro's Sum)**.** Let x_n be a sequence of real numbers that converges to x and v_n be a nondecreasing sequence diverging to infinity with $v_0 = 0$, that is, $x_n \to x$ and $v_n \nearrow \infty$. Then, $v_n^{-1} \sum_{k=1}^{n} (v_k - v_{k-1}) x_k \to x$. [A proof can be found in lecture note 1.]

Applying Cesàro's sum with $v_n = n$ and $x_n = p^{(n)}(x,y)$, we get $(1/n) \sum_{k=1}^{n} p^{(k)}(x,y) \to 0$.

Definition 28. A nonnegative function μ is called an *invariant measure* if $\mu p = \mu$ and $\mu \neq 0$, that is, $\sum_x \mu(x) p(x,y) = \mu(y)$ for all y. Let $N_t(y)$ be the number of visits to y up to time t, that is, $N_t(y) = \sum_{n \leq t} 1(X_n = y)$.

Theorem 42. Let X_n be an irreducible and recurrent Markov chain having p as its transition matrix. Define $\mu_x(y) = \sum_{n=0}^{\infty} P_x(X_n = y, T_x > n)$ for any x, y. Then,
(a) μ_x is an invariant measure satisfying $0 < \mu_x(y) < \infty$.
(b) $\mu_z(y) = \mu_x(y)/\mu_x(z)$ for any x, y, z.
(c) $\sum_y \mu_x(y) = \mathbb{E}_x T_x$.

Proof. (a) By the definition, $\mu_x(y) = \sum_{n=0}^{\infty} P_x(X_n = y, T_x > n)$ is the expected number of visits to y before returning to x. Note that $\mu_x(x) = 1$ and

$$\begin{aligned}
\sum_y \mu_x(y)p(y,x) &= \mu_x(x)p(x,x) + \sum_{y \neq x}\sum_{n=0}^{\infty} P_x(T_n = y, T_x > n)p(y,x) \\
&= p(x,x) + \sum_{n=1}^{\infty}\sum_{y\neq x} P_x(X_n = y, T_x > n)p(y,x) \\
&= p(x,x) + \sum_{n=1}^{\infty}\sum_{y\neq x} P_x(X_n = y, X_{n+1} = x, T_x > n) \\
&= P_x(X_1 = x) + \sum_{n=1}^{\infty} P_x(X_1 \neq x, \ldots, X_n \neq x, X_{n+1} = x) \\
&= P_x(T_x = 1) + \sum_{n=1}^{\infty} P_x(T_x = n+1) = P_x(T_x < \infty) = 1 = \mu_x(x).
\end{aligned}$$

If $y \neq x$, $\mu_x(y) = \sum_{n=1}^{\infty} P_x(X_n = y, T_x > n)$. Note that $P_x(X_1 = y, T_x > 1) = p(x,y)1(y \neq x)$ and for $n \geq 2$, $P_x(X_n = y, T_x > n) = \sum_{z\neq x} P_x(X_n = y, X_{n-1} = z, T_x > n) = \sum_{z\neq x} P_x(X_{n-1} = z, T_x > n-1)P_x(X_n = y \mid X_{n-1} = z) = \sum_{z\neq x} P_x(X_{n-1} = z, T_x > n-1)p(z,y)$. Thus

$$\begin{aligned}
\mu_x(y) &= \sum_{n=0}^{\infty} P_x(X_n = y, T_x > n) = p(x,y) + \sum_{n=2}^{\infty}\sum_{z\neq x} P_x(X_{n-1} = z, T_x > n-1)p(z,y) \\
&= p(x,y) + \sum_{z\neq x}\sum_{n'=1}^{\infty} P_x(X_{n'} = z, T_x > n')p(z,y) = p(x,y) + \sum_{z\neq x}\mu_x(z)p(z,y) = \sum_z \mu_x(z)p(z,y).
\end{aligned}$$

Hence μ_x is a nontrivial invariant measure.

The irreducibility implies $\rho_{xy} > 0$ and hence there exists $k > 0$ such that $p^{(k)}(x,y) > 0$. The relation $\mu = \mu p = \mu p^k$ implies $\mu_x(y) = \sum_z \mu_x(z)p^{(k)}(z,y) \geq \mu_x(x)p^{(k)}(x,y) = p^{(k)}(x,y) > 0$. Thus $\mu(y) > 0$ for all y. Similarly, there exists l such that $p^{(l)}(y,x) > 0$. Then, $1 = \mu_x(x) = \sum_z \mu_x(z)p^{(l)}(z,x) \geq \mu_x(y)p^{(l)}(y,x)$ implies $\mu_x(y) \leq 1/p^{(l)}(y,x) < \infty$.

(b) Exercise or see Bremaud.

(c) $\sum_y \mu_x(y) = \mu_x(x) + \sum_{y\neq x}\mu_x(y) = 1 + \sum_{y\neq x}\sum_{n=0}^{\infty} P_x(X_n = y, T_x > n) = 1 + \sum_{n=1}^{\infty} P_x(X_n \neq x, T_x > n) = \sum_{n=1}^{\infty} P_x(T_x = n) + \sum_{n=1}^{\infty} P_x(T_x > n) = \sum_{n=1}^{\infty} P_x(T_x \geq n) = \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} P(T_x = k) = \sum_{k=1}^{\infty}\sum_{n=1}^{k} P_x(T_x = k) = \sum_{k=1}^{\infty} kP_x(T_x = k) = \mathbb{E}_x T_x$. □

Note. If $\sum_y \mu_x(y) < \infty$, then $\pi(y) = \mu_x(y)/\sum_z \mu_x(z)$ becomes a stationary distribution. By part (b), these stationary distributions are the same. The term $\mathbb{E}_x T_x$ plays very important role. If $\mathbb{E}_x T_x < \infty$ for a x, then $\mathbb{E}_y T_y < \infty$ for all y. Besides there exists the unique stationary distribution. Particularly $\pi(y) = \mu_y(y)/\sum_z \mu_y(z) = 1/\mathbb{E}_y T_y$ for y.

Corollary 43. Let X_t be an irreducible recurrent homogeneous Markov chain. If $\mathbb{E}_x T_x < \infty$, then $\mathbb{E}_y T_y < \infty$ and there exists the unique stationary distribution $\pi(x) = 1/\mathbb{E}_x T_x > 0$.

Definition 29. A recurrent state x is *positive recurrent* if $\mathbb{E}_x T_x < \infty$ and is *null recurrent* otherwise.

The above definition comes from $1/\mathbb{E}_x T_x$ which is positive if x is positive recurrent and which is zero if x is null recurrent.

Theorem 44. An irreducible homogeneous Markov chain X_n is positive recurrent if and only if it has a stationary distribution π. Moreover, if there exists a stationary distribution, then it is unique and positive ($\pi > 0$).

Proof. Sufficiency ($\Longrightarrow$). Obvious from Theorem 6 and the note following.
Necessity ($\Longleftarrow$). If there exists a transient state x then, $\pi(x) = 0$ by Proposition 5, which violates the assumption $\pi > 0$. Hence all states are recurrent. If there exists a null recurrent state x, then all states are null recurrent and there is no stationary distribution. It also violates the existence assumption of a stationary distribution. Thus all states should be positive recurrent. $\square$

Theorem 45. Let X_n be an irreducible homogeneous Markov chain. Then,
(a) $\lim_{n\to\infty} N_n(t)/n = 1/\mathbb{E}_y T_y$ almost surely.
(b) If there exists a stationary distribution π and further X_n is aperiodic, then $\lim_{n\to\infty} p^{(n)}(x,y) \to \pi(y)$.

Proof. (a) Let $T_y^0 = 0$ for convenience. Note that $N_n(y) = \sum_{i=1}^n 1(X_i = y) = \sum_{k=1}^\infty k1(T_y^k \le n < T_y^{k+1})$. Let k_n be the k such that $T_y^k \le n < T_y^{k+1}$. Then $k_n/T_y^{k_n+1} < N_n(y)/n \le k_n/T_y^{k_n}$. Since y is recurrent, that is, $P_y(T_y < \infty) = 1$, we get $k_n \to \infty$ as $n \to \infty$. By the strong Markov property, the distributions of $T_y^2 - T_y^1, T_y^3 - T_y^2, \ldots$ are independent and identically distributed. Hence $T_y^k/k \to \mathbb{E}_y T_y$ almost surely by the strong law of large numbers. Then $k_n/T_y^{k_n+1}, k_n/T_y^{k_n} \to 1/\mathbb{E}_y T_y$ almost surely. Therefore $N_n(y)/n \to 1/\mathbb{E}_y T_y$ almost surely.
(b) If $p^{(n)}(x,y)$ converges, then it converges to $\lim_{n\to\infty}(1/n)\sum_{k=1}^n p^{(k)}(x,y) = \lim_{n\to\infty} \mathbb{E}_x N_n(y)/n = \pi(y)$. If there exists a stationary distribution, then $\pi = \pi p$ and π is an eigen vector having eigen value 1. If X_t is aperiodic, then the second largest absolute eigen value is less than 1. That is for any distribution μ, $|(\mu-\pi)p| \le c|\mu - \pi|$ for some $0 \le c < 1$. Hence $|(\mu-\pi)p^n| \le c|(\mu-\pi)p^{n-1}| \le c^n|\mu-\pi| \le 2c^n \to 0$. By taking $\mu(x) = 1$ and $\mu(z) = 0$ for all $z \neq x$, we get $p^{(n)}(x,y) - \pi(y) \to 0$ as $n \to \infty$. Equivalently, $p^{(n)}(x,y) \to \pi(y)$ as $n \to \infty$. $\square$

Theorem 46 (Ergoric theorem)**.** Suppose X_n is an irreducible positive recurrent homogeneous Markov chain. If $\sum_x |f(x)|\pi(x) < \infty$, then $(1/n)\sum_{k=1}^n f(X_k) \to \sum_x f(x)\pi(x) = \mathbb{E}_\pi f(X_0)$ as $n \to \infty$.

Proof. Let $k_n = N_n(x)$ be the k satisfying $T_x^k \le n < T_x^{k+1}$. Then $Y_1 = \sum_{j=T_x^1+1}^{T_x^2} f(X_j), \ldots, Y_k = \sum_{j=T_x^k+1}^{T_x^{k+1}} f(X_j)$ are i.i.d. with mean $\mathbb{E}_x \sum_{j=1}^{T_x} f(X_j) = \sum_y \mu_x(y) f(y)$. Then

$$\frac{1}{n}\sum_{j=1}^n f(X_j) = \frac{N_n(x)}{n}\frac{1}{N_n}\Big[\sum_{j=1}^{N_n-1} Y_j + \sum_{j=1}^{T_x} f(X_j) + \sum_{j=T_x^k+1}^{n} f(X_j)\Big] \to (\mathbb{E}_x T_x)^{-1}\sum_y \mu_x(y)f(y) = \sum_y \pi(y)f(y) = \mathbb{E}_\pi f(X_0).$$

$\square$