Review Handout 2: Markov chains

Math/Stat 491: Introduction to Stochastic Processes

Wellner; 11/18/2013

Part 1: Terminology and Definitions

1. Markov property: a discrete time stochastic process $X = \{X_n : n \geq 0\}$ is a *Markov process* if

$$P(X_{n+1} = y | X_n = x, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) = P(X_{n+1} = y | X_n = x).$$

2. probability transition matrix: $P(X_{n+1} = j | X_n = i) = P(i, j)$ for $i, j \in \mathbb{E}$.

3. $n-$step transition matrix: the $n-$step transition matrix is $P^m(i, j) = P(X_{m+n} = j | X_n = i)$, the $m-$th power of the 1-step transition matrix $\mathbf{P} = (P(i, j) : i, j \in \mathbb{E})$.

4. temporally homogeneous process: a Markov process is *temporally homogeneous* if the transition probability $P(X_{n+1} = j | X_n = i) = P(i, j)$ does not depend on n.

5. the *Chapman-Kolmogorov equation:* $P^{m+n}(i, j) = \sum_{k \in \mathbb{E}} P^m(i, k) P^n(k, j)$.

6. *first return time* T_j, $j \in \mathbb{E}$: $T_j = \min\{n \geq 1 : X_n = j\}$.

7. *strong Markov property*: If T is a stopping time and $\{X_n : n \geq 0\}$ is a Markov chain, then conditional on T and X_T the process $\{X_{T+k} : k \geq 0\}$ has the same distribution as the Markov chain itself started at the initial state X_T. (Informally: the process starts over at stopping times.)

8. *an absorbing state*: A state $j \in \mathbb{E}$ is an *absorbing state* if $P(j, j) = 1$.

9. *a state i communicates with a state j* and we write $i \to j$ if there is a positive probability of reaching j starting from i: $\rho_{i,j} = P_i(T_j < \infty) > 0$.

10. *transient state:* A state $j \in \mathbb{E}$ is *transient* if $\rho_{jj} = P_j(T_j < \infty) < 1$.

11. *recurrent state:* A state $j \in \mathbb{E}$ is *recurrent* if $\rho_{jj} = P_j(T_j < \infty) = 1$.

12. an *irreducible set of states:* A set $B \subset \mathbb{E}$ is *irreducible* if for any $i, j \in B$ we have $i \to j$ (and $j \to i$); that is, $\rho_{ij} \equiv P_i(T_j < \infty) > 0$; or, equivalently $P^n(i,j) > 0$ for some $n \geq 1$

13. a *closed set of states:* A set of states $A \subset \mathbb{E}$ is *closed* if for any $i \in A$ and $j \in A^c$, the transition probability $P(i,j) = 0$.

14. a *stationary distribution:* a probability distribution $\underline{\pi}$ on the state space $\mathbb{E}$ is a stationary distribution if $\underline{\pi}\mathbf{P} = \underline{\pi}$ (where $\underline{\pi}$ is a row vector on both sides of the equation).

15. a *stationary measure:* If $\underline{\pi}$ satisfies $\underline{\pi}\mathbf{P} = \underline{\pi}$ and $\pi_j \geq 0$, then $\underline{\pi}$ is a stationary measure. (It need not have total mass 1.)

16. the *period of a state* j is the greatest common divisor of the set $I_j = \{n \geq 1 : \; P^n(j,j) > 0\}$.

17. an *aperiodic Markov chain* is a Markov chain with all states having period 1.

18. the *detailed balance condition*: A probability distribution $\underline{\pi}$ on $\mathbb{E}$ satisfies the *detailed balance condition* if $\pi_i P(i,j) = \pi_j P(j,i)$ for all $i, j \in \mathbb{E}$.

19. a *doubly stochastic Markov chain*: both $\sum_{j \in \mathbb{E}} P(i,j) = 1$ for every $i \in \mathbb{E}$ *and* $\sum_{i \in \mathbb{E}} P(i,j) = 1$ for every $j \in \mathbb{E}$ hold.

20. the *Metroplis-Hastings* algorithm: given a transition probability matrix Q on $\mathbb{E}$ and a probability distribution $\underline{\pi}$ on $\mathbb{E}$, the new matrix $P(i,j) = Q(i,j)R(i,j)$ with

$$R(i,j) \equiv \min\left\{\frac{\pi(j)Q(j,i)}{\pi(i)Q(i,j)}, 1\right\}$$

    is the transition probability matrix of a Markov chain $\{X_n : \; n \geq 0\}$ on $\mathbb{E}$ which has $\underline{\pi}$ as its stationary distribution.

21. a *reflecting random walk*: a Markov chain with state space $\mathbb{E} = \{0, 1, 2, \ldots\}$ and transition probability matrix given by $P(i, i+1) = p$ for $i \geq 0$, $P(i, i-1) = 1 - p$ for $i \geq 1$, and $P(0,0) = 1 - p$ for some $p \in (0,1)$.

22. a *Galton-Watson branching process:* A Markov chain with state space $\mathbb{E} = \{0, 1, 2, \ldots\}$ and with transition probability matrix given by $P(i, j) = P(\sum_{k=1}^{i} Y_k = j)$ where $Y_1, Y_2, \ldots$ are independent and identically distributed random variables with offspring distribution $\{p_k\}$ on $\{0, 1, 2, \ldots\}$.

23. a *renewal chain:* a Markov chain with state space $\mathbb{E} = \{0, 1, 2, \ldots\}$ and transition probability matrix given by $P(0, i) = p_i$ for $i \geq 0$ with $\sum_{i=0}^{\infty} p_i = 1$ and $P(i, i-1) = 1$ for $i > 0$.

24. an *aging chain:* a Markov chain with state space $\mathbb{E} = \{0, 1, 2, \ldots\}$ with transition matrix given by $P(i, i+1) = p_i$ for each i, $P(i, 0) = 1 - p_i$, and $P(i, j) = 0$ for $j \notin \{i+1, 0\}$.

Part 2: Results and theorems

1. Theorem: If $C \subset \mathbb{E}$ is a finite closed and irreducible set, then all states in C are recurrent.

2. Theorem: If the state space $\mathbb{E}$ of a Markov chain is finite, then $\mathbb{E}$ can be decomposed as a disjoint union of the transient states T and a finite number of closed, irreducible, and recurrent sets of states.

3. Theorem: A state $j \in \mathbb{E}$ is recurrent if and only if $\sum_{n=1}^{\infty} P^n(j, j) = E_j[N(j)] = \infty$.

4. martingales connected with Markov chains: If $\mathbf{P}f = \lambda f$ then $M_n = \lambda^{-n} f(X_n)$ is a martingale.

5. *Theorem: (existence of a stationary distribution, finite state space):* If the state space $\mathbb{E}$ is finite and $\mathbf{P}$ is irreducible, then there is a unique solution to $\underline{\pi}\mathbf{P} = \underline{\pi}$ with $\sum_j \pi_j = 1$ and we have $\pi_j > 0$ for all j.

6. *Proposition:* If $\{X_n : n \geq 0\}$ is a Markov chain with finite state space, $N = \#$ of states in $\mathbb{E}$, and $\mathbf{P}$ is doubly stochastic, then $\underline{\pi} = (1/N)\underline{1} = (1/N, \ldots, 1/N)$ is a stationary distribution.

7. *Proposition:* A Markov chain with transition matrix given by the Metropolis-Hastings algorithm satisfies the detailed balance condition.

8. *Convergence Theorem 1:* If I, A, and S hold, then $P^n(i, j) \to \pi_j$ for all $i, j \in \mathbb{E}$.

9. *Theorem:* If I and R hold, then there is a stationary measure with $\mu(j) > 0$ for all j.

10. *Convergence Theorem 2:* Suppose I and R hold. If

$$N_n(y) = \sum_{k=1}^{n} 1\{X_k = y\},$$

then

$$n^{-1} N_n(y) \to_p \frac{1}{E_y T_y}$$

and this also holds with probability 1.

11. *Convergence theorem corollary:* If I and S (and hence also R) holds, then

$$n^{-1} N_n(y) \to_p \frac{1}{E_y T_y} = \pi_y,$$

and this also holds with probability 1.

12. *Convergence Theorem 3:* Suppose that I and S hold and that $\sum_j |f(j)|\pi(j) = E_\pi |f(X)| < \infty$. Then

$$n^{-1} \sum_{k=1}^{n} f(X_k) \to_p E_\pi f(X) = \sum_j f(j)\pi(j).$$

and this also holds with probability 1.