Math 5652: Introduction to Stochastic Processes: Fall 2014

Appendix B. Markov Chains.

B.1. General properties.

Let $X_0, X_1, X_2, \dots$ be a Markov chain with values in S (the state space) and transition probabilities $p(x,y)$. For an arbitrary function h on S, define

$$Lh(x) := \sum_y p(x,y)h(y) - h(x) \quad \text{for} \quad x \in S. \tag{B.1}$$

This is a linear operator, i.e

$$L(h_1 + h_2) = Lh_1 + Lh_2, \quad \text{and} \quad L(ch) = c \cdot Lh \quad \text{for} \quad c = \text{const}. \tag{B.2}$$

In addition, we always have

$$L(1) = \sum_y p(x,y) - 1 = 0. \tag{B.3}$$

We will use the following facts which are true for general Markov chains.

Theorem 1. *Let Γ be a subset of the state space S, such that its complement $S \setminus \Gamma$ is finite. Denote $T = \inf\{n \geq 0 : X_n \in \Gamma\}$, and assume $P_x(T < \infty) = 1$ for all $x \in S \setminus \Gamma$. Obviously $T = 0$ on Γ. Let f and g be given functions defined correspondingly on $S \setminus \Gamma$ and Γ. Then the function*

$$h(x) := E_x(Z) \quad \text{on} \quad S, \quad \text{where} \quad Z := \sum_{k=0}^{T-1} f(X_k) + g(X_T), \tag{B.4}$$

and $E_x(Z) := E(Z \mid X_0 = x)$ is a unique solution of the system

$$Lh = -f \quad \text{on} \quad S \setminus \Gamma, \qquad h = g \quad \text{on} \quad \Gamma. \tag{B.5}$$

Proof. If $X_0 = x \in \Gamma$, then $T = 0$ and $h(x) = E_x g(X_0) = g(x)$. If $X_0 = x \in S \setminus \Gamma$, then

$$h(x) = E_x(Z) = \sum_y P_x(X_1 = y) \cdot E_x(Z \mid X_1 = y) = \sum_y p(x,y) \cdot \big[f(x) + h(y)\big] = f(x) + \sum_y p(x,y)h(y),$$

and the equality $Lh(x) = -f(x)$ follows.

In order to show that there exists a unique solution of (B.5), denote $M := \#(S \setminus \Gamma)$ – the number of points in $S \setminus \Gamma$. Since the values $h = g$ on Γ are known, the system (B.5) consists of M linear equations with M unknowns $h(x),\ x \in S \setminus \Gamma$. From Linear Algebra it is known that the existence and uniqueness for this system is equivalent to the uniqueness of trivial solution $h \equiv 0$ for the corresponding homogeneous system

$$Lh = 0 \quad \text{on} \quad S \setminus \Gamma, \qquad h = 0 \quad \text{on} \quad \Gamma. \tag{B.6}$$

Suppose this is not the case. Then

$$0 < A := \max_{S \setminus \Gamma} |h| = |h(x_0)| \quad \text{for some} \quad x_0 \in S \setminus \Gamma. \tag{B.7}$$

The equality $Lh = 0$ can be rewritten as $h = ph$ with $M \times M$ matrix $p = [p(x,y)], x, y \in S \setminus \Gamma$ and the column vector h of length M with components $h(y), y \in S \setminus \Gamma$. By iteration, we get

$$h = ph = p^2 h = \cdots = p^n h = \cdots . \tag{B.8}$$

By our assumptions $P_{x_0}(T < \infty) > 0$, which implies that

$$p^n(x_0, y_0) > 0 \quad \text{for some natural} \quad n \quad \text{and} \quad y_0 \in \Gamma. \tag{B.9}$$

From $h(y_0) = 0$ it follows

$$h(x_0) = p^n h(x_0) = \sum_y p^n(x_0, y)\, h(y) = \sum_{y \neq y_0} p^n(x_0, y)\, h(y).$$

By (B.7) and (B.9),

$$0 < A = |h(x_0)| \le \sum_{y \neq y_0} p^n(x_0, y) \cdot |h(y)| \le A \cdot \sum_{y \neq y_0} p^n(x_0, y) = A \cdot \big(1 - p^n(x_0, y_0)\big) < A.$$

This contradiction shows that the system (B.6) has a unique solution $h \equiv 0$. Theorem is proved. □

The following two corollaries can be considered as generalizations of Theorems 1.27 and 1.28 in [D].

Corollary 2. *For an arbitrary state $x_0 \in \Gamma$, the function $h(x) = P_x(X_T = x_0)$ is a unique solution of the system*

$$Lh = 0 \quad \text{on} \quad S \setminus \Gamma, \qquad h(x_0) = 1, \quad h = 0 \quad \text{on} \quad \Gamma \setminus \{x_0\}.$$

Corollary 3. *The function $h(x) = E_x(T)$ is a unique solution of the system*

$$Lh = -1 \quad \text{on} \quad S \setminus \Gamma, \qquad h = 0 \quad \text{on} \quad \Gamma.$$

B.2. Random walk $X_{n+1} = X_n + Y_{n+1}$ with bounded i.i.d. Y.

Let $Y_1, Y_2, \dots, Y_n, \dots$ be independent identically distributed (i.i.d.) random variables with distribution

$$a_k = P(Y = k), \quad -k_1 \le k \le k_2, \quad \sum_{k=-k_1}^{k_2} a_k = 1, \tag{B.10}$$

where k_1 and k_2 are non-negative integers. We exclude the trivial case $a_0 = P(Y = 0) = 1$, and assume $a_{-k_1} > 0,\ a_{k_2} > 0$. Then automatically $K = k_1 + k_2 \ge 1$. Consider the *random walk*

$$X_0 = x, \quad X_{n+1} = X_n + Y_{n+1}, \qquad a < x < b, \tag{B.11}$$

with integers $a, b, x,\ b - a \ge 2$. The first exit time of X_n out of the interval (a, b) is a stopping time

$$T = \inf\{n \ge 0 : X_n \le a \quad \text{or} \quad X_n \ge b\}.$$

We can have $X_T \le a$ (with positive probability) only if $k_1 \ge 1$, and $X_T \in \Gamma_1 = \{a, a-1, \dots, a - k_1 + 1\}$. Similarly, we can have $X_T \ge b$ only if $k_2 \ge 1$, and $X_T \in \Gamma_2 = \{b, b+1, \dots, b + k_2 - 1\}$. In any case $X_T \in \Gamma = \Gamma_1 \cup \Gamma_2$, and the set Γ consists of $K = k_1 + k_2 \ge 1$ points (states). We may have $\Gamma_1 = \emptyset$ (if $k_1 = 0$) or $\Gamma_2 = \emptyset$ (if $k_2 = 0$), but at least one of these two sets is not empty.

Under these assumptions, Theorem 1 can be reformulated as follows.

Theorem 4. *Denote $S = \{\text{integers } x :\ a - k_1 < x < b + k_2\}$, $\Gamma := \{x \in S :\ x \le a \text{ or } x \ge b\}$. Then for arbitrary functions f on $S \setminus \Gamma = \{\text{integers } x :\ a < x < b\}$ and g on Γ, the function*

$$h(x) = E_x\Big[\sum_{k=0}^{T-1} f(X_k) + g(X_T)\Big] \tag{B.12}$$

is a unique solution of the system (B.5).

From (B.10) and (B.11) it follows

$$p(x,y) = P(X_{n+1} = X_n + Y_{n+1} = y \mid X_n = x) = P(Y_{n+1} = y - x) = a_{y-x}. \tag{B.13}$$

Then for $h(x) := r^x$ with $r = \text{const} \neq 0$ in (B.1), we have

$$L(r^x) = \sum_y p(x,y)\, r^y - r^x = \sum_y a_{y-x} r^y - r^x = \sum_k a_k r^{x+k} - r^x = r^{x-k_1} q(r), \tag{B.14}$$

where

$$q(r) = a_{-k_1} + a_{1-k_1} r + \cdots + (a_0 - 1) r^{k_1} + \cdots + a_{k_2} r^K \tag{B.15}$$

is a polynomial of degree $K = k_1 + k_2 \geq 1$. The properties of this polynomial are quite similar to those of the characteristic polynomial in the theory of linear differential equations with constant coefficients. Here we list (without proof) some of them.

(I) Write $q(r)$ as a product of linear factors:

$$q(r) = c \cdot \prod_j (r - r_j)^{m_j} \quad \text{with} \quad \sum_j m_j = K,$$

where r_j are distinct roots (real or complex) of $q(r)$. Note that by (B.3), one of these roots $r_1 = 1$. Moreover, since $q(0) = a_{-k_1} \neq 0$, we have $r_j \neq 0$ for all j, so that r_j^x is defined in the usual algebraic sense for any integer x. We claim that each of functions

$$\{h_1(x), \ldots, h_K(x)\} = \{x^\mu r_j^x : \ 1 \leq j \leq l,\ 0 \leq \mu \leq m_j - 1\} \tag{B.16}$$

satisfies $Lh(x) = 0$ for all integer x. The total number of these functions is $m_1 + \cdots + m_l = K$. For any pair of mutually conjugate roots r_0 and $\overline{r_0}$ (they must have same multiplicity m_0), one can replace complex functions r_0^x and $\overline{r_0}^x$ by the real and imaginary parts $\mathrm{Re}(r_0^x)$ and $\mathrm{Im}(r_0^x)$.

(II) Let r_0 be a root of $q(r)$ of multiplicity $m_0 \geq 0$ ($m_0 = 0$ if $q(r_0) \neq 0$). Then the equation

$$Lh_0(x) = r_0^x P(x), \quad \text{where} \quad P(x) \quad \text{is a polynomial,} \tag{B.17}$$

has a solution

$$h_0(x) = r_0^x x^{m_0} Q_0(x), \quad \text{where} \quad Q_0(x) \quad \text{is a polynomial of degree} \quad \deg Q_0 = \deg P. \tag{B.18}$$

(III) The set of functions in (B.16) is independent on Γ, i.e. the equality

$$c_1 h_1 + \cdots + c_K h_K = 0 \quad \text{on} \quad \Gamma \tag{B.19}$$

is only possible if $c_1 = \cdots = c_K = 0$.

Given these properties, we get the following algorithm of solving the system (B.5) for Markov chains in (B.10)–(B.11) in the case $f(x) = r_0^x P(x)$, where $P(x)$ is a polynomial.

Step 1. Using (I), find linearly independent solutions $h_1, \ldots, h_k$ of $Lh = 0$ in the form (B.16).

Step 2. Using (II), find a particular solution h_0 of $Lh_0 = -f$ in the form (B.18).

Step 3. Write the desired solution in the form

$$h = h_0 + c_1 h_1 + \cdots + c_K h_K \tag{B.20}$$

and find the unknown constants $c_1, \ldots, c_K$ from K conditions $h(x) = g(x),\ x \in \Gamma$.

Problem 1. Consider a symmetric random walk

$$X_0 = 0, \quad X_{n+1} = X_n + Y_{n+1} \quad \text{for} \quad n \ge 0,$$

where $Y_0, Y_1, Y_2, \ldots$ are independent with distribution $P(Y = 1) = P(Y = -1) = \frac{1}{2}$.
For fixed $N \ge 1$, find the expectation

$$E(X_1^2 + X_2^2 + \cdots + X_T^2), \quad \text{where} \quad T = \inf\{n \ge 0 : |X_n| = N\}.$$

Solution. By Theorem 4, the function

$$h(x) = E_x(X_0^2 + X_1^2 + \cdots + X_{T-1}^2)$$

is a unique solution of the system

$$Lh(x) = \frac{1}{2}\,[h(x+1) + h(x-1)] - h(x) = -x^2 \quad \text{for} \quad x = 0, \pm 1, \ldots, \pm(N-1); \qquad h(\pm N) = 0.$$

We have

$$L(r^x) = \frac{1}{2}\,[r^{x+1} + r^{x-1})] - r^x = r^{x-1}q(r), \quad \text{where} \quad q(r) = \frac{1}{2}(r-1)^2.$$

The polynomial $q(r)$ has the only root $r_1 = 1$ of multiplicity $m_1 = 2$. Following the above procedure, we get two independent solutions $h_1(x) \equiv 1$ and $h_2(x) = x$ of $Lh = 0$. Next, one can find a particular solution of $Lh_0 = -x^2$ in the form $h_0 = x^2 Q_0(x)$, where $Q_0(x) = a_0 + a_1 x + a_2 x^2$. Since

$$\begin{aligned}
L(x^2) &= \frac{1}{2}\,[(x+1)^2 + (x-1)^2] - x^2 = 1, \\
L(x^3) &= \frac{1}{2}\,[(x+1)^3 + (x-1)^3] - x^3 = 3x, \\
L(x^4) &= \frac{1}{2}\,[(x+1)^4 + (x-1)^4] - x^4 = 6x^2 + 1,
\end{aligned}$$

we obtain $h_0(x) = \frac{1}{6}(x^2 - x^4)$. Now we can write

$$h = h_0 + c_1 h_1 + c_2 h_2 = \frac{1}{6}\,(x^2 - x^4) + c_1 + c_2 x,$$

and find c_1 and c_2 from the equalities $h(\pm N) = 0$, i.e. $c_1 = \frac{1}{6}\,(N^4 - N^2)$, $c_2 = 0$. Finally,

$$\begin{aligned}
E_0(X_1^2 + X_2^2 + \cdots + X_T^2) &= h(0) + E_0(X_T^2) = h(0) + N^2 \\
&= \frac{1}{6}\,(N^4 - N^2) + N^2 = \frac{1}{6}\,(N^4 + 5N^2).
\end{aligned}$$

References

[D] Richard Durrett, *Essentials of Stochastic Processes*, 2nd Edition, Springer, 2012.