MATH 777 MARKOV STOCHASTIC PROCESSES SPRING 2009

Lecture 4. Discrete Markov chains. Recurrent states and transient states.

Before I start this lecture note, let me give a problem for you to solve:

$\boxed{\mathbf{1}}$ Can you invent a discrete Markov chain so that for some state y the first return time to y has the following distribution?

(a) $P_y\{\tau_y = 1\} = f_{yy}^{(1)} = 1$; (b) $f_{yy}^{(2)} = 1$; (c) $f_{yy}^{(1)} = f_{yy}^{(2)} = 1/2$; (d) $f_{yy}^{(1)} = 1/3$, $f_{yy}^{(2)} = 2/3$; (e) $f_{yy}^{(1)} = 1/2$, $P_y\{\tau_y = \infty\} = 1/2$; (f) $f_{yy}^{(1)} = f_{yy}^{(2)} = f_{yy}^{(3)} = 1/3$; (g) $f_{yy}^{(k)} = \dfrac{1}{k(k+1)}$, $k = 1, 2, 3, \ldots$.

The easy way to handle the equation (3.31) is through generating functions.

If $a_0, a_1, a_2, \ldots, a_n, \ldots$ is a sequence of numbers, its generating function is defined as

$$A(s) = \sum_{n=1}^{\infty} a_n s^n. \tag{4.1}$$

For some s, the series (4.1) may diverge.

We will be denoting generating functions with capital letters corresponding to the lower-case letters denoting the sequences. So,

$$P_{xy}(s) = \sum_{n=0}^{\infty} p_{xy}^{(n)} s^n, \qquad F_{xx}(s) = \sum_{k=0}^{\infty} f_{xy}^{(k)} s^k. \tag{4.2}$$

The series defining $P_{xy}(s)$ converges at least for all s with $|s| < 1$, that for $F_{xy}(s)$ for $|s| \le 1$ (because $\sum_{k=0}^{\infty} f_{xy}^{(k)} = P_x\{\tau_y < \infty\} \le 1$).

If we multiply both sides of (3.31) by s^n and take the sum over n from 0 to ∞, we get, for $x \ne y$:

$$P_{xy}(s) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} f_{xy}^{(k)} p_{yy}^{(n-k)} \cdot s^n. \tag{4.3}$$

This double series converges absolutely, at least for $|s| < 1$, so we can change the order of summation:

$$\begin{aligned} P_{xy}(s) &= \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} f_{xy}^{(k)} p_{yy}^{n-k} \cdot s^n = \sum_{k=0}^{\infty} f_{xy}^{(k)} \cdot s^k \cdot \sum_{n=k}^{\infty} p_{yy}^{(n-k)} \cdot s^{n-k} \\ &= \sum_{k=0}^{\infty} f_{xy}^{(k)} \cdot s^k \cdot \sum_{m=0}^{\infty} p_{yy}^{(m)} \cdot s^m = F_{xy}(s) \cdot P_{yy}(s). \end{aligned} \tag{4.4}$$

For $x = y$ we have, instead of this, because $p_{yy}^{(0)} = 1$, while $\sum_{k=0}^{0} f_{yy}^{(k)} p_{yy}^{(0-k)} = 0$:

$$P_{yy}(s) = 1 + F_{yy}(s) \cdot P_{yy}(s). \tag{4.5}$$

While equality (4.4) involves three functions, $P_{xy}(s)$, $F_{xy}(s)$, and $P_{yy}(s)$, in (4.4) we have only two functions, and we can express one of them through the other:

$$F_{yy}(s) = 1 - \frac{1}{P_{yy}(s)}, \tag{4.6}$$

$$P_{yy}(s) = \frac{1}{1 - F_{yy}(s)}. \tag{4.7}$$

Of course, the question arises whether we can divide by $P_{yy}(s)$ or by $1 - F_{yy}(s)$. But $F_{yy}(s) < 1$ at least for all $s \in (-1, 1)$, and $P_{yy}(s) > 0$ at least for $s \in [0, 1)$; so equalities (4.6), (4.7) hold at least for all $s \in [0, 1)$.

The place for this material should be in the previous lecture. Now let me start with what this was all for.

A point (state) $y \in X$ is called *recurrent* (with respect to our Markov chain) if $P_y\{\tau_y < \infty\} = 1$ (we return to the state y almost surely); and it is called *transient* if $P_y\{\tau_y < \infty\} < 1$ (we don't return with a positive probability).

$\boxed{\mathbf{2}}$ Let me introduce the random variables τ_y^k (the time of the k-th return to the state y) by $\tau_y^1 = \tau_y$, $\tau_y^k = \min\{n > \tau_y^{k-1} \colon \xi_n = y\}$ (and, of course, by definition $\tau_y^k = \infty$ if there are no such n).

Prove that for a recurrent state all τ_y^k are finite almost surely with respect to the probability P_y; and for a transient y, P_y-almost surely we return to the state y only finitely many times: almost surely all τ_y^k starting with some (random) k are equal to $+\infty$.

Theorem 4.1. *A state* $y \in X$ *is recurrent if and only if* $\sum_{n=0}^{\infty} p_{yy}^{(n)} = \infty$ (and it is transient if and only if the series $\sum_{n=0}^{\infty} p_{yy}^{(n)}$ converges).

Proof. Take $s \to 1^-$ in the equalities (4.6), (4.7). We have: $\lim_{s\to 1^-} P_{yy}(s) = \sum_{n=1}^{\infty} p_{yy}^{(n)}$, $\lim_{s\to 1^-} F_{yy}(s) = \sum_{n=1}^{\infty} f_{yy}^{(n)}$ (this follows from the monotone-convergence theorem for the Lebesgue integral, the sum being the integral with respect to the counting measure $\#$; but of course this can be easily proved in an elementary way). So $\sum_{n=1}^{\infty} f_{yy}^{(n)} = 1$ if and only if $\sum_{n=0}^{\infty} p_{yy}^{(n)} = \infty$.

Let us consider some examples.

For a finite ergodic Markov chain we know that $\lim_{n\to\infty} p_{yy}^{(n)} = p_y$. If the limiting probability p_y is positive for our state y, the series $\sum_{n=0}^{\infty} p_{yy}^{(n)}$ diverges, and the state y is recurrent.

On the other hand, we saw that the convergence $p_{xy}^{(n)} \to p_y$ is, for finite ergodic chains, exponentially fast; so if $p_y = 0$, the series $\sum_{n=0}^{\infty} p_{yy}^{(n)}$ converges, and the state y is transient.

A chain with an infinite state space (phase space) X that we considered as an example in Lecture Note 2008.23 – 24: the simplest random walk in $\mathbb{Z}^1$ (on the set of all integers). This chain can be constructed like this: we take independent integer-valued random variables ξ_0, η_1, η_2, ..., η_n, ..., with $P\{\eta_i = \pm 1\} = 1/2$; and define $\xi_1 = \xi_0 + \eta_1$, $\xi_2 = \xi_0 + \eta_1 + \eta_2$, ..., $\xi_n = \xi_0 + \sum_{i=1}^{n} \eta_i$. Is a state $y \in \mathbb{Z}^1$ recurrent or transient? (It is clear that if one y is recurrent, all other points in $\mathbb{Z}^1$ also are.)

We have:

$$p_{yy}^{(n)} = P_y\{y + \eta_1 + ... + \eta_n = y\} = P_y\{\eta_1 + ... + \eta_n = 0\} \tag{4.8}$$

(the probability of this event does not depend on y, so we could have dropped the subscript in it).

Of course, this probability is equal to 0 if n is odd (the sum of an odd number of odd summands η_i is necessarily odd, and so not equal to 0: an impossible event); and for an even n it is a binomial probability: probability of $n/2$ successes (counting $\eta_i = 1$ as a success, and $\eta_i = -1$ as a failure) in n Bernoulli trials with probability of success 1/2:

$$p_{yy}^{(n)} = \frac{\binom{n}{n/2}}{2^n} = \frac{n!}{\big((n/2)!\big)^2 \cdot 2^n} \tag{4.9}$$

Using Stirling's formula for the asymptotics of the factorial $n! \sim \sqrt{2\pi n}\, n^n e^{-n}$ ($\sim$ meaning that the ratio of both quantities goes to 1 as $n \to \infty$), we obtain:

$$p_{yy}^{(n)} \sim \frac{\sqrt{2\pi n}\, n^n e^{-n}}{(\sqrt{\pi n}\,(n/2)^{n/2} e^{-n/2})^2 \cdot 2^n} = \frac{\sqrt{2/\pi}}{\sqrt{n}} \qquad (n \to \infty). \tag{4.10}$$

The series $\sum_{k=1}^{\infty} 1/\sqrt{k}$ diverges, so every state in this Markov chain is recurrent.

What about the two-dimensional simplest random walk (with transition probabilities given by formula (2008.23 – 24.20) or (2008.23 – 24.21)? As mentioned after these formulas, these two Markov chains are isomorphic to one another; but the chain with transition probabilities (2008.23 – 24.21) is easier to consider, because the horizontal component and the vertical component are independent; so we have:

$$p_{xy,\,xy}^{(n)} \begin{cases} = 0 & \text{if } n \text{ is odd,} \\ \sim \big(\frac{\sqrt{2/\pi}}{\sqrt{n}}\big)^2 = \frac{2/\pi}{n} & \text{if } n \text{ is even.} \end{cases} \tag{4.11}$$

The series of these quantities also diverges: again recurrence.

As for the three-dimensional simplest random walk, there are several non-isomorphic versions of it: one in which we go from a point $(x, y, z) \in \mathbb{Z}^3$ to the points $(x \pm 1, y, z)$, $(x, y \pm 1, z)$, $(x, y, z \pm 1)$ with probabilities 1/6; another where we go from (x, y, z) to $(x \pm 1, y \pm 1, z \pm 1)$ with probabilities 1/8. For this last, because of independence of the components, we have $p_{xyz,\,xyz}^{(n)} \sim \dfrac{(2/\pi)^{3/2}}{n^{3/2}}$ for even n, the series converges, and every point (x, y, z) is transient.

What can we say about the random walk (not "the simplest" one) with transition probabilities given by (2008.23 – 24.23), that is, about the Markov chain based on sums of independent identically distributed integer-valued random variables η_1, η_2, ..., η_n, ...? As we understand now, recurrence/transience of state 0 (and all other states) in this chain depends on whether the series $\sum_{n=1}^{\infty} P\{\eta_1 + ... + \eta_n = 0\}$ diverges or converges. Let us assume (as it was done in Lecture Note 2008.23 – 24) that $\xi_0 = 0$.

If $E\eta_k \neq 0$, we have by the strong law of large numbers, for every n_0,

$$\frac{\eta_{n_0+1} + \eta_{n_0+2} + ... + \eta_n}{n - n_0} \to^{\text{a. s.}} E\eta_k \quad (n \to \infty), \tag{4.12}$$

$$\eta_{n_0+1} + \eta_{n_0+2} + ... + \eta_n \to^{\text{a. s.}} \infty \tag{4.13}$$

or

$$\eta_{n_0+1} + \eta_{n_0+2} + ... + \eta_n \to^{\text{a. s.}} -\infty \tag{4.14}$$

(the first for $E\eta_k > 0$, the second for negative $E\eta_k$). Let us consider the case $E\eta_k > 0$.

Because of (4.13), there exists a negative integer $-C$ such that $P\{\eta_{n_0+1} + \eta_{n_0+2} + ... + \eta_n > -C \text{ for all } n > n_0\} > 1/2$.

Since $E\eta_k > 0$, there is some positive m such that $P\{\eta_k = m\} = p_m > 0$. Take n_0 so that $n_0 \cdot m > C$. Then we have:

$$\{\tau_0 = \infty\} \supseteq \{\eta_1 = \eta_2 = ... = \eta_{n_0} = m,\ \eta_{n_0+1} + \eta_{n_0+2} + ... + \eta_n > -C \text{ for all } n > n_0\}, \tag{4.15}$$

$$\begin{aligned} P_0\{\tau_0 = \infty\} \geq P_0\{&\eta_1 = \eta_2 = ... = \eta_{n_0} = m, \\ &\eta_{n_0+1} + \eta_{n_0+2} + ... + \eta_n > -C \text{ for all } n > n_0\} > p_m^{n_0} \cdot 1/2 > 0. \end{aligned} \tag{4.16}$$

So we see that in the case $E\,\eta_k \neq 0$ all states are transient.

Let us consider the case of $E\eta_k = 0$. Here we'll use characteristic functions (see Lectures 2008.14, 2008.15).

For an integer-valued random variable ζ its characteristic function $f_\zeta(t)$ is given by

$$f_\zeta(t) = \sum_{j=-\infty}^{\infty} P\{\zeta = j\} \cdot e^{ijt}. \tag{4.17}$$

This is a (2π)-periodic function given by its Fourier series (in the complex form). The probabilities $P\{\zeta = j\}$ are the Fourier coefficients, and they are found as

$$P\{\zeta = j\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ijt} f_\zeta(t)\, dt \tag{4.18}$$

(see (2008.15.22)).

If the characteristic function of one summand η_k is $f(t)$ (a (2π)-periodic function), the characteristic function of the random variable $\eta_1 + ... + \eta_n$ is $f(t)^n$; and by formula (4.18) with $j = 0$ we have:

$$P\{\eta_1 + ... + \eta_n = 0\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)^n\, dt \qquad \Big[= \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re} f(t)^n\, dt\Big]. \tag{4.19}$$

We need to find out whether the series $\sum_{n=0}^{\infty} p_{yy}^{(n)} = \sum_{n=0}^{\infty} \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} f(t)^n \, dt$ converges or diverges.In the lecture I said that "under some conditions", this sum is equal to $\dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} \sum_{n=0}^{\infty} f(t)^n dt$, and then the problem was easily solved; but under *what* conditions?

In this lecture note I'll go another way. We can find $P_{yy}(s) = \sum_{n=0}^{\infty} \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} f(t)^n \cdot s^n \, dt$ for $s \in [0,\, 1)$, and then take $s \to 1^-$. We have: $\sum_{n=0}^{\infty} \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} |f(t)|^n \cdot s^n \, dt \le \sum_{n=0}^{\infty} s^n < \infty$, and by Fubini's Theorem we can change the order of summation and integration:

$$\begin{aligned} P_{yy}(s) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n=0}^{\infty} [s \cdot f(t)]^n \, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{1 - s f(t)} \, dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \operatorname{Re} \frac{1}{1 - s f(t)} \, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 - s \cdot \operatorname{Re} f(t)}{\big(1 - s \cdot \operatorname{Re} f(t)\big)^2 + s^2 \big(\operatorname{Im} f(t)\big)^2} . \end{aligned} \tag{4.20}$$

As $s \to 1^-$, the integrand converges to $\operatorname{Re} \dfrac{1}{1 - f(t)}$; but does it follow that the integral converges to $\dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} \operatorname{Re} \dfrac{1}{1 - f(t)} \, dt$? If we can find an integrable function $g(t)$ dominating all integrands, yes; but we'll see that in the case that we are interested in there is no such dominating integrable function.

Let us restrict ourselves to the case of $E\eta_k = 0$, $0 < E\eta_k^2 = \operatorname{Var}(\eta_k) < \infty$, and the characteristic function $f(t) \ne 1$ for $t \in [-\,\pi,\, \pi]$.

Of course, the characteristic function $f(t)$ is equal to 1 at $t = 0$; it may be equal to 1 for finitely many other t's in the interval $(-\,\pi,\, \pi]$.

$\boxed{\mathbf{3}}$ Give an example of an integer-valued random variable η not equal to a constant almost surely, for which the characteristic function is equal to 1 in absolute value at three points in the interval $(-\,\pi,\, \pi]$.

Solving Problem $\boxed{\mathbf{3}}$, you understand just how restrictive our requirement of $f(t) \ne 1$ for $t \in [-\,\pi,\, \pi] \setminus \{0\}$ is and how to handle the case when this requirements is not satisfied.

Another chance to justify the limit passage in the integral is using the monotone-convergence theorem. Is the function $\dfrac{1 - s \cdot \operatorname{Re} f(t)}{\big(1 - s \cdot \operatorname{Re} f(t)\big)^2 + s^2 \big(\operatorname{Im} f(t)\big)^2}$ non-decreasing in s? Let us denote, for shortness, $\operatorname{Re} f(t) = a$, $\operatorname{Im} f(t) = b$. Is the function $\dfrac{1 - as}{(1 - as)^2 + b^2 s^2}$ non-decreasing in s for $s \in [0,\, 1)$? Let us differentiate:

$$\frac{d}{ds} \frac{1 - as}{(1 - as)^2 + b^2 s^2} = \frac{(a - 2b^2 s)(1 - as)^2 - ab^2 s^2}{\big((1 - as)^2 + b^2 s^2\big)^2} . \tag{4.21}$$

If $1/2 \le a \le 1$, and $|b| \le (1 - a)/2$, the numerator in (4.21) is greater or equal

$$(1 - a)^2 (a - 2b^2) - ab^2 \ge (1 - a^2)^2 \cdot 3/8 - b^2 > 0. \tag{4.22}$$

We don't know whether the conditions $1/2 \le a \le 1$ and $|b| \le (1-a)/2$ are satisfied for $a = \mathrm{Re} f(t)$ and $b = \mathrm{Im} f(t)$ for all $t \in [-2\pi, 2\pi]$, but as $t \to 0$ we have: $a = \mathrm{Re} f(t) \to 1$, $1 - \mathrm{Re} f(t) = 1 - \dfrac{\mathrm{Var}(\eta_k)}{2} \cdot t^2 + o(t^2) > C \cdot t^2/2$ for sufficiently small t with some $C > 0$, $b = \mathrm{Im} f(t) = o(t^2)$ as $t \to 0$. So the inequalities $1/2 \le a \le 1$, and $|b| \le (1-a)/2$ are satisfied for t in some interval $(-\delta, \delta)$, $\delta > 0$.

On the rest of the interval $[-\pi, \pi]$, that is, on $[-\pi, \pi] \setminus (-\delta, \delta)$, the integrands are dominated by some constant; so we have:

$$\lim_{s \to 1^-} \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re} \frac{1}{1 - s\, f(t)}\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re} \frac{1}{1 - f(t)}\, dt. \tag{4.23}$$

Since $f(t) = 1 - \dfrac{\mathrm{Var}(\eta_k)}{2}\, t^2 + o(t^2)$ as $t \to 0$, the integral (4.23) diverges, and we have recurrence.

In the r-dimensional case, if $\boldsymbol{\eta}_k$ are independent identically distributed random vectors with the expectation $E\boldsymbol{\eta}_k = 0$ and a non-degenerate covariance matrix $E(\boldsymbol{\eta}_k \cdot \boldsymbol{\eta}_k^T)$, we have the same; the integral $\dfrac{1}{(2\pi)^r} \displaystyle\int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} \mathrm{Re} \frac{1}{1 - f(\boldsymbol{t})}\, dt_1 \dots dt_r$ diverges in the two-dimensional case, from which recurrence follows, and converges in the case of three dimensions and more: transience.

$\boxed{\mathbf{4}}$ Check that the (2π)-periodic function $f(t) = 1 - |\sin(t/2)|$ is a characteristic function of a discrete distribution. For this, find its Fourier coefficients p_j and check that they are nonnegative, $\sum_{j=-\infty}^{\infty} p_j = 1$.

$\boxed{\mathbf{5}}$ Does the series $\sum_{j=-\infty}^{\infty} |j| \cdot p_j$ converge or diverge?

$\boxed{\mathbf{6}}$ Is $\sum_{n=0}^{\infty} \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} f(t)^n\, dt = \infty$ or $< \infty$? Are the states in the Markov chain with transition matrix (2008.23 – 24.23) with p_j of Problem $\boxed{\mathbf{4}}$ recurrent or transient?