MATH4240 Tutorial 7

2024-03-11

Example of Stationary Distributions

Recall the following theorems from lectures

Theorem 1. *An irreducible positive-recurrent Markov chain has a unique stationary distribution* $\pi(i) = 1/m(i)$ *for each state* i *where* $m(i) = E_i(T_i) < \infty$ *is the expected recurrent time of state* i.

In particular, an irreducible finite Markov chain has a unique stationary distribution[1].

Theorem 2. *If* P *is the transition matrix of a Markov chain with a finite state space, and*

- 1 *is a simple eigenvalue of* P
- *there exists a (left) eigenvector* π *corresponding to* 1 *with all entries nonnegative* $\pi_i \geq 0$
- *all other eigenvalues of* P *have moduli less than* 1

then π *is the unique stationary distribution of the chain, and* $P^n \xrightarrow{n\to\infty} (\pi, \pi, \dots, \pi)^\mathsf{T}$, *assuming* π *is normalized to have* $\sum \pi_i = 1$.

With these two theorems in mind, let us look at a few examples on stationary distributions.

Example 1. (HW4 Optional Q7) Recall that the Ehrenfest chain is a chain on state space $S = \{0, \dots, d\}$ with transition function $P(x, x+1) = 1 - x/d$, $P(x, x-1) = x/d$. What is its stationary distribution?

Note that Ehrenfest chain is an irreducible birth-and-death chain with forward probability $p_x = (d-x)/d$ and backward probability $q_x = x/d$.

Stationary Distribution for general Birth-and-Death Chain

It appears that formula for stationary distribution on a (general) birth-and-death chain is left as an exercise in the lecture notes. Here we will give a short proof on the result.

Recall that on a (general) birth-and-death chain, we have $r_x = P(x, x)$, $p_x = P(x, x+1)$, $q_x = P(x, x-1)$ with $q_0 = 0$. We will assume that the chain is irreducible by assuming that $p_x, q_x > 0$. Then on a stationary distribution π we have

$$\pi(0) = \pi(0)r_0 + \pi(1)q_1$$
$$\text{or} \quad q_1\pi(1) - p_0\pi(0) = 0$$
$$\pi(x) = \pi(x-1)p_{x-1} + \pi(x)r_x + \pi(x+1)q_{x+1} \quad \text{for } x \geq 1$$
$$\text{or} \quad q_{x+1}\pi(x+1) - p_x\pi(x) = q_x\pi(x) - p_{x-1}\pi(x-1)$$

Induction implies that $q_{x+1}\pi(x+1) - p_x\pi(x) = 0$ for all x, and so $\pi(x+1)/\pi(x) = p_x/q_{x+1}$. Another induction gives us

$$\pi(x) = \pi(0) \prod_{i=0}^{x-1} \frac{p_i}{q_{i+1}} = \pi(0)\pi_x$$

with $\pi_x = \prod_{i=0}^{x-1} \frac{p_i}{q_{i+1}} = (\prod_{i=0}^{x-1} p_i)/(\prod_{i=1}^{x} q_i)$ and $\pi_0 = 1$. Normalizing (if possible) gives us a stationary distribution $\pi(x) = \pi_x / \sum \pi_i$.

[1] You can actually show the existence and uniqueness of stationary distribution on a *finite* irreducible Markov chain with irreduciblity alone, without using the concept of expected recurrent time. (Try to show it!)

Using the formula we have

$$\pi_0 = 1 = \binom{d}{0}$$

$$\pi_x = (\prod_{i=0}^{x-1} p_i)/(\prod_{i=1}^{x} q_x) = (\prod_{i=0}^{x-1} d-i)/(\prod_{i=1}^{x} i) = (\frac{d!}{(d-x)!})/x! = \binom{d}{x} \quad \text{for } x \in \{1,\ldots,d\}$$

and $\sum_{i=0}^{d} \pi_i = \sum_{i=0}^{d} \binom{d}{i} = 2^d$, which implies that

$$\pi(x) = \pi_x / \sum_{i=0}^{d} \pi_i = 2^{-d}\binom{d}{x}$$

is of binomial distribution. With law of large number, this approaches to a normal distribution on $d \to \infty$, which matches what is known in statistical mechanics. (See also the *modified* Ehrenfest chain in textbook)

Note. $\pi_x \gamma_x p_x = p_0$

Example 2. Consider a chain on state space $S = \{1,\ldots,n\} = A \cup B$ with $A = \{1,\ldots,k\}$ and $B = \{k+1,\ldots,n\}$ with $1 \le k < n$ (so A, B are nonempty), with transition matrix being

$$P(x,y) = \begin{cases} 1/(n-k) & \text{if } x \in A,\ y \in B \\ 1/k & \text{if } x \in B,\ y \in A \\ 0 & \text{otherwise} \end{cases}$$

(so the transition graph is completely bipartite with uniform probability of transition). What is its stationary distribution?

Easy to see that the chain is irreducible, so there must be one (and only one) stationary distribution. Let it be π. We can note that

$$\pi(x) = \sum_{y\in B} \pi(y)P(y,x) = \frac{1}{k}\sum_{y\in B}\pi(y) \quad \text{for each } x \in A$$

$$\text{so} \quad \sum_{x\in A}\pi(x) = \frac{1}{k}\sum_{x\in A}\sum_{y\in B}\pi(y) = \sum_{y\in B}\pi(y)$$

Since $1 = \sum \pi(x) = \sum_{x\in A}\pi(x) + \sum_{y\in B}\pi(y)$, we must have $\sum_{x\in A}\pi(x) = \sum_{y\in B}\pi(y) = 1/2$. By symmetry, we can guess that

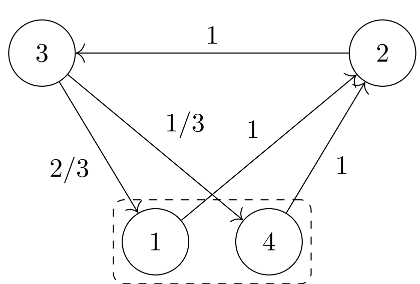
$$\pi(x) = \begin{cases} \frac{1}{2k} & \text{if } x \in A \\ \frac{1}{2(n-k)} & \text{if } x \in B \end{cases}$$

and it is easy to verify that π is indeed the stationary distribution of the chain.

Example 3. Consider a chain with transition matrix

$$P = \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

which transition diagram



Easy to see that the chain is irreducible, so there must be a (unique) stationary distribution. To find this distribution, we simply solve the equation $\pi = \pi P$, or

$$\begin{aligned} 1 &= \pi(1) + \pi(2) + \pi(3) + \pi(4) \\ \pi(1) &= \frac{2}{3}\pi(3) \\ \pi(2) &= \pi(1) + \pi(4) \\ \pi(3) &= \pi(2) \\ \pi(4) &= \frac{1}{3}\pi(3) \end{aligned}$$

which gives $\pi = (2/9, 1/3, 1/3, 1/9)$.

However, note that the eigenvalues of P are $0, 1, e^{i2\pi/3}, e^{-i2\pi/3}$, so there are eiganvalues other than 1 that has unit modulus, so the theorem about $\lim P^n$ does not apply. Indeed we can compute that

$$P^2 = \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}, \quad P^3 = \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{pmatrix} 2/3 & 0 & 0 & 1/3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2/3 & 0 & 0 & 1/3 \end{pmatrix} \end{array}, \quad P^4 = \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{array} = P$$

so $\lim_{n\to\infty} P^n$ does not exist[2].

As a side note, we can see that by permuting the indices and computing directly that

$$P^3 \sim \begin{array}{c} \begin{matrix} 1 & 4 & 2 & 3 \end{matrix} \\ \left(\begin{array}{cc|c|c} 2/3 & 1/3 & 0 & 0 \\ 2/3 & 1/3 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right) \end{array}, \quad \frac{1}{3}(P + P^2 + P^3) = \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{pmatrix} 2/9 & 1/3 & 1/3 & 1/9 \\ 2/9 & 1/3 & 1/3 & 1/9 \\ 2/9 & 1/3 & 1/3 & 1/9 \\ 2/9 & 1/3 & 1/3 & 1/9 \end{pmatrix} \end{array} = \begin{pmatrix} \pi \\ \pi \\ \pi \\ \pi \end{pmatrix}$$

and so with a simple argument $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P^k = (\pi, \pi, \pi, \pi)^{\mathsf{T}}$ as well[3]. This is a general phenomenon and will be covered (in detail) in later lectures.

Example 4. Consider the symmetric 1D random walk on $S = \mathbb{Z}$ with transition function $P(x, x+1) = P(x, x-1) = 1/2$, which we know from tutorial 4 is an irreducible recurrent chain. What is the stationary distribution?

Suppose $\pi : \mathbb{Z} \to [0, 1]$ is a stationary distribution. Then π satisfies $\sum \pi(i) = 1$ and for all i,

$$\begin{aligned} \pi(i) &= \pi(i+1)P(i+1, i) + \pi(i+1)P(i+1, i) \\ &= \frac{1}{2}(\pi(i-1) + \pi(i+1)) \\ \text{and so} \quad \pi(i+1) - \pi(i) &= \pi(i) - \pi(i-1) \end{aligned}$$

This gives $\pi(i) = \pi(0) + i(\pi(1) - \pi(0))$ for all $i \in \mathbb{Z}$. Since $0 \le \pi(i) \le 1$ for all i, we must have $\pi(1) - \pi(0) = 0$ and so $\pi(i) = \pi(0)$ for all i (which is expected from translation symmetry). This contradicts the condition that $0 < \sum \pi(i) = 1 < \infty$ no matter what $\pi(0) \in [0, 1]$ is, so no such π exists.

The same argument also tells us that asymmetric 1D random walk with transition function $P(x, x+1) = p$, $P(x, x-1) = 1 - p$, $p \in [0, 1]$ has no stationary distribution.

---

[2] While the number 3 can be found by computing all P^n (or computing all eigenvalues), it can also be found efficiently with a graph traversal, just like irreducibility with Tarjan's algorithm. For more detail, see e.g. the article by Jarvis and Shier and the NetworkX Python package for an implementation. Unfortunately these algorithms don't seem to produce a certificate (e.g. a circuit for irreducibility) on the properties.

[3] For an irreducible recurrent chain, $P - \lim \frac{1}{n}\sum_k P^k$ plays the role of Q in absorbing chain.

Although all constant functions $f : \mathbb{Z} \to \mathbb{R}$ satisfy $f = fP$ (as row vector)[4], none of them corresponds to a *probability* distribution (as pmf). Whether we can normalize these functions (to have a stationary *distribution*) requires additional condition on the chain (i.e. positive recurrent).

Example 5. Consider the chain on $S = \mathbb{N}$ with transition function $P(x, x+1) = p$, $P(x, 0) = 1 - p$ for all $x \in \mathbb{N}$ on some $p \in (0, 1)$. The chain is trivially irreducible and recurrent. What is the stationary distribution?

A stationary distribution must satisfy $\sum_{n=0}^{\infty} \pi(n) = 1$ and

$$\pi(0) = \sum_{n=0}^{\infty} \pi(n) P(n, 0) = 1 - p$$
$$\pi(i) = \pi(i-1) P(i-1, i) = p\pi(i-1) \quad \text{for } n \geq 1$$

Solving this gives $\pi(n) = (1-p)p^n$ for $n \geq 0$ when $p \in (0, 1)$ (which we can verify that this is indeed a stationary distribution).

When $p = 0$ the only recurrent state is 0, so easy to see that the only stationary distribution is $\pi(n) = \delta_{n0}$ (which matches the result above); when $p = 1$ the chain has no recurrent state, so no stationary distribution exists (neither exists a nontrivial function f with $f = fP$[5]).

Bounded Convergence Theorem

When dealing with a chain with (countably) infinite states, we often need to interchange the order of sum and limit. The following theorem provides justification on when we can do so.

Theorem 3 (Bounded Convergence Theorem / Tannery's theorem)**.** *Let $\{f_n : \mathbb{Z}^+ \to \mathbb{R}\}$ be a sequence of functions on positive integers that satisfies*

- *(uniform bound) there exists $M > 0$ such that $|\, f_n(x) \,| \leq M$ for all n, x*
- *(pointwise convergence) on each x, $\lim_{n\to\infty} f_n(x)$ exists (and is finite)*

Let $p(x) \geq 0$ on $x \in \mathbb{Z}^+$ be such that $\sum_x p(x) < \infty$. Then we can interchange sum and limit:

$$\lim_{n\to\infty} \sum_x p(x) f_n(x) = \sum_x p(x) \lim_{n\to\infty} f_n(x)$$

Proof. Let $f(x) = \lim_{n\to\infty} f_n(x)$ be the pointwise limit of f_n. On each x we have $|\, f_n(x) \,| \leq M$ and $f_n(x) \to f(x)$, so $|\, f(x) \,| \leq M$ and thus $|\, p(x)f(x) \,| \leq Mp(x)$, which implies that $\sum |\, p(x)f(x) \,| \leq M \sum p(x) = M < \infty$ and so $\sum p(x)f(x) \in \mathbb{R}$ is well-defined.

Let $\epsilon > 0$. Since $\sum p(x) < \infty$, there exists X such that $\sum_{x=X+1}^{\infty} p(x) < \frac{\epsilon}{4M}$. This implies that

$$\left| \sum_{x=X+1}^{\infty} p(x) f_n(x) \right| \leq \sum_{x=X+1}^{\infty} p(x) \,|\, f_n(x) \,| \leq M \sum_{x=X+1}^{\infty} p(x) < \epsilon/4$$

for all n and similarly $\left| \sum_{x=X+1}^{\infty} p(x) f(x) \right| < \epsilon/4$.

Furthermore, as $f_n(x) \to f(x)$ for all $x \in \{1, \ldots, X\}$, there exists N such that $|\, f_n(x) - f(x) \,| < \epsilon/2$ for all $n \geq N$ and for all $x \in \{1, \ldots, X\}$.

[4]In fact, such function exists in general: on every *recurrent* state x of a (possibly infinite) chain you can always construct a function $f_x : S \to [0, \infty)$ with $f_x(x) = 1$ and $f = fP$: (Durrett, *Probability Theory and Examples*, Edition 3, Ch. 5 Thm. 4.3) $f_x(y)$ is the expected number of visit of y from x before returning to x. See also the example on Ehrenfest chain (and general birth-and-death chain).

[5]However, there may exist such nontrivial function on *certain* transient chains. I do not know any exact classification on when this happens.

So for each $n \geq N$,

$$
\begin{aligned}
& \left| \sum_x p(x) f_n(x) - \sum_x p(x) f(x) \right| \\
&= \left| \sum_{x=1}^{X} p(x) f_n(x) + \sum_{x=X+1}^{\infty} p(x) f_n(x) - \sum_{x=1}^{X} p(x) f(x) - \sum_{x=X+1}^{\infty} p(x) f(x) \right| \\
&\leq \left| \sum_{x=1}^{X} p(x) f_n(x) - \sum_{x=1}^{X} p(x) f(x) \right| + \left| \sum_{x=X+1}^{\infty} p(x) f_n(x) \right| + \left| \sum_{x=X+1}^{\infty} p(x) f(x) \right| \\
&\leq \sum_{x=1}^{X} p(x) \mid f_n(x) - f(x) \mid + \frac{\epsilon}{4} + \frac{\epsilon}{4} \\
&\leq \frac{\epsilon}{2} \sum_{x=1}^{\infty} p(x) + \frac{\epsilon}{2} = \epsilon
\end{aligned}
$$

which implies that $\lim_{n \to \infty} \sum_x p(x) f_n(x) = \sum_x p(x) f(x)$. □

With some modification we can actually show that the same result holds if $\mid f_n(x) \mid \leq M(x)$ and $\sum_x p(x) M(x) < \infty$. These are all special cases of *Lebesgue's dominated convergence theorem*, for which the statement and the proof can be found in most standard textbooks on measure theory (and MATH4050).