0.1 Basic Limit Theorem of Markov Chains

(Lesson 8 of 3/4/20) Compiled: April 6, 2020

We are finally able to formalize and *prove* the intuitive fact that the *long-run* probability of returning to a state i is the *reciprocal* of the average return time of m_i: that is, if the system is at i every m_i steps, then it spends $1/m_i$ of the total time in i, and so if we inspect the state at a random step (of a *infinitely long* experiment) we will find the system at i with probability $1/m_i$. Moreover, we will also find the appropriate conditions that are necessary for this result to hold.

Consider a **recurrent** state i. As we have seen before, the first return time R_i can be defined as:

$$R_i = \min\{n \geq i; X_n = i\}$$

and is distributed according to:

$$f_{ii}^{(n)} = \mathbb{P}\{R_i = n|X_0 = i\} = \mathbb{P}\{X_n = i, X_\nu \neq i \,\forall \nu = 1, \dots, n|X_0 = i\}$$

Since the state i is recurrent, $f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)} = 1$, and so R_i is a finite-valued random variable. In other words, R_i can never be infinite, since the system will return to i *for sure*. More precisely, the probability of R_i being arbitrarily high *vanishes*.

The mean duration between visits to state i is the expectation of R_i:

$$m_i = \mathbb{E}[R_i|X_0 = i] = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

In other words, the system *on average* returns to i once every m_i units of time.

Note that the fact that R_i is a finite-valued random variable does **not** prevent m_i from being infinite. This in fact happens if $\mathbb{P}[R_i = n]$ decreases *sufficiently slowly* as $n \to \infty$.

Theorem 0.1.1. ***Basic limit theorem of Markov Chains***

(a) *Consider a **recurrent**, **irreducible**, **aperiodic** Markov chain. Let $P_{ii}^{(n)}$ be the probability of entering state i at the n-th transition, with $n \in \mathbb{N}$, given that the initial state is i ($X_0 = i$). Let $f_{ii}^{(n)}$ be the probability of first returning to state i at the n-th transition. Then:*

$$\lim_{n\to\infty} P_{ii}^{(n)} = \frac{1}{\sum_{n=1}^{\infty} n f_{ii}^{(n)}} = \frac{1}{m_i}$$

(b) Also:

$$\lim_{n\to\infty} P_{ji}^{(n)} = \pi_i = \lim_{n\to\infty} P_{ii}^{(n)} = \frac{1}{m_i}$$

for all states j.

The proof is referred to a later chapter.

Note that theorem **??** can be applied also to **aperiodic recurrent classes** in a Markov Chain. In fact, we noted that different classes can only be linked by *one-way* transitions - meaning that after leaving a class C, the system cannot return in it. So a *recurrent class* must necessarily be isolate, i.e. such that $P_{ij}^{(n)} = 0$ for all $i \in C$, $j \notin C$, and for all n. So we can consider the submatrix $\left\| \mathbf{P}_{ij} \right\|$, with $i, j \in C$, as the transition probability matrix of a separate **irreducible** Markov Chain, for which the basic limit theorem directly applies.

Depending on the finiteness of m_i, we distinguish two cases:

- If the average return time m_i is **finite**, then $\lim_{n\to\infty} P_{ii}^{(n)} > 0$, and the same applies to all states $j \leftrightarrow i$. This means that $\pi_j > 0$ for every j, and so all states in the *aperiodic recurrent class* continue to be visited in the long run. Classes with this property are said to be **positive recurrent**, or *strongly ergodic*.

- If $m_i = \infty$, then $\pi_j = 0$ for every j, then the class is said to be **null recurrent**, or *weakly ergodic*. In a sense, this is the *critical line* separating transient states from recurrent ones, where the system *will* return to i *surely*, but it needs *infinite* time to do so.

Since these are all class properties, there cannot be *positive recurrent* and *null recurrent* states in the same class. So, a state can be either **transient**, **positive recurrent** or **null recurrent**, with all the properties summarized in table **??**.

Type of State	$f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)}$	$\lim_{k\to\infty} \mathbb{P}[M > k \vert X_0 = i]$	$\mathbb{E}[M \vert X_0 = i]$	$m_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$	$\pi_i = \frac{1}{m_i}$
Transient	< 1	0	$\frac{f_{ii}}{1-f_{ii}} < \infty$	∞	0
Null Recurrent	1	1	∞	∞	0
Positive Recurrent	1	1	∞	$< \infty$	>0

Table (1) – Summary of the main properties for the different categories of states. f_{ii} is the probability of returning to i, M is the number of returns to i, m_i the average time between returns and π_i the probability of the system being in i in the long run.

A positive recurrent aperiodic class behaves, when taken by itself, the same as a *regular* Markov chain, and so the same result about the limiting distribution still holds:

Theorem 0.1.2. ***Limit distribution of a positive recurrent aperiodic class***.
In a positive recurrent aperiodic class with states $j \in \mathbb{N}$, we have:

$$\lim_{n\to\infty} P_{jj}^{(n)} = \pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij} \qquad \sum_{i=0}^{\infty} \pi_i = 1$$

and $\boldsymbol{\pi}$ is uniquely determined by the set of equations:

$$\pi_i \geq 0, \sum_{i=0}^{\infty} \pi_i = 1 \qquad \pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij} \quad j \in \mathbb{N} \tag{1}$$

In general, any set $(\pi_j)_{i=0}^{\infty}$ satisfying (**??**) is called a **stationary probability distribution**.

Note, however, that theorem **??** is actually more general than the result we got for regular Markov chain, as here we are not assuming a *finite* number of states. In fact, we cannot employ the same proof - because there we needed to exchange the order of two sums, which needs justification in the infinite case. Of course, we could use Fubini's theorem to generalize the previous arguments, but in this case is actually more instructive to restart from first principles.

Proof. As before, we need to prove two things: that the limiting probabilities π_j are indeed the solution of (**??**), and that this solution is unique.

Existence:

1. We start from the normalization condition for rows in the n-step transition matrix:

$$1 = \sum_{j=0}^{+\infty} P_{ij}^{(n)} \qquad \forall n$$

   Since all the addends are non-negative, the total sum cannot be lower than any *truncated* sum:

$$1 = \sum_{j=0}^{+\infty} P_{ij}^{(n)} \geq \sum_{j=0}^{M} P_{ij}^{(n)} \qquad \forall n, M$$

   We then take the limit $n \to \infty$, and bring it inside the sum, since it is over a *finite* number of elements:

$$1 \geq \lim_{n\to\infty} \sum_{j=0}^{M} P_{ij}^{(n)} = \sum_{j=0}^{M} \lim_{n\to\infty} P_{ij}^{(n)} = \sum_{j=0}^{M} \pi_j \qquad \forall M$$

Finally we take also the $M \to \infty$ limit, leading to:

$$\lim_{M\to\infty} \sum_{j=0}^{M} \pi_j = \sum_{j=0}^{+\infty} \pi_j \leq 1 \tag{2}$$

So we have found that the sum of π_j converges. Clearly, we would like to prove that it is *exactly* 1.

2. Again we start from a known relationship:

$$P_{ij}^{(m+n)} = \sum_{k=0}^{+\infty} P_{ik}^{(m)} P_{kj}^{(n)} \qquad \forall m, n$$

Again we can truncate the sum to write an inequality:

$$P_{ij}^{(m+n)} \geq \sum_{k=0}^{M} P_{ik}^{(m)} P_{kj}^{(n)} \qquad \forall m, n, M$$

Taking the limit $m \to \infty$ and taking it into the finite sum:

$$\begin{aligned}\pi_j = \lim_{m\to\infty} P_{ij}^{(m+n)} \geq \lim_{m\to\infty} \sum_{k=0}^{M} P_{ik}^{(m)} P_{kj}^{(n)} &= \sum_{k=0}^{M} \lim_{m\to\infty} P_{ik}^{(m)} P_{kj}^{(n)} = \\ &= \sum_{k=0}^{M} \pi_k P_{kj}^{(n)} \quad \forall M, n\end{aligned}$$

Finally, we take also $M \to \infty$, leading to:

$$\pi_j \geq \sum_{k=0}^{+\infty} \pi_k P_{kj}^{(n)} \qquad \forall n \tag{3}$$

3. We want to show that (**??**) hold as an equality, and we do this by contradiction. Suppose that there exist an index j for which (**??**) holds *strictly*:

$$\exists j \colon \pi_j > \sum_{k=0}^{+\infty} \pi_k P_{kj}^{(n)}$$

Summing over j, the inequality remains strict:

$$\sum_{j=0}^{+\infty} \pi_j > \sum_{j=0}^{+\infty} \sum_{k=0}^{+\infty} \pi_k P_{kj}^{(n)} \tag{4}$$

Let's evaluate this last sum. Again, we first *truncate* the inner sum to the first M elements, so that we can exchange the two sums and obtain an inequality that remains valid also in the limit $M \to \infty$:

$$\sum_{j=0}^{+\infty}\sum_{k=0}^{+\infty}\pi_k P_{kj}^{(n)} \geq \sum_{j=0}^{+\infty}\sum_{k=0}^{M}\pi_k P_{kj}^{(n)} = \sum_{k=0}^{M}\pi_k \sum_{j=0}^{+\infty} P_{kj}^{(n)} = \sum_{k=0}^{M}\pi_k \quad \forall M$$

And in particular:

$$\sum_{j=0}^{+\infty}\sum_{k=0}^{+\infty}\pi_k P_{kj}^{(n)} \geq \lim_{M\to\infty}\sum_{k=0}^{M}\pi_k = \sum_{k=0}^{+\infty}\pi_k$$

Substituting in (**??**) we get:

$$\sum_{j=0}^{+\infty}\pi_j > \sum_{j=0}^{+\infty}\sum_{k=0}^{+\infty}\pi_k P_{kj}^{(n)} \geq \sum_{k=0}^{+\infty}\pi_k$$

which is absurd, as no quantity can be strictly greater than itself. So, by contradiction it must be:

$$\pi_j = \sum_{k=0}^{+\infty}\pi_k P_{kj}^{(n)} \quad \forall n$$

Setting $n = 1$ we obtain part of the thesis we wish to prove.

4. All that's left is to deal with the normalization property, i.e. show that (**??**) holds as an equality.

   First, note that $|P_{kj}^{(n)}| \leq 1 \; \forall n$ (they are **uniformly bounded**). This, along with the convergence of $\sum_{k=0}^{+\infty}\pi_k \leq 1$ (**??**) allows to bring the limit inside the sum in the following:

$$\pi_j = \lim_{n\to\infty}\sum_{k=0}^{+\infty}\pi_k P_{kj}^{(n)} = \sum_{k=0}^{+\infty}\pi_k \lim_{n\to\infty} P_{kj}^{(n)} = \left(\sum_{k=0}^{+\infty}\pi_k\right)\pi_j$$

   Since $\pi_j > 0$ (because the chain is positive recurrent by hypothesis), we can divide both sides by π_j, leading to:

$$\sum_{k=0}^{+\infty}\pi_k = 1$$

   This finally proves the existence of the solution of (**??**).

Uniqueness.
Let $\boldsymbol{x}$ be a solution, i.e. such that:

$$x_j = \sum_{i=0}^{+\infty} x_i P_{ij}; \qquad \sum_{i=0}^{+\infty} x_i = 1 \tag{5}$$

We then proceed as we did for regular MCs, rewriting the x_i in the rhs of (**??**) by using (**??**) itself. Then we apply again the trick of *truncating* the inner sum to exchange the sums:

$$\begin{aligned} x_j &= \sum_{i=0}^{+\infty} \left(\sum_{k=0}^{+\infty} \pi_k P_{ki} \right) P_{ij} \geq \sum_{i=0}^{+\infty} \left(\sum_{k=0}^{M} x_k P_{ki} \right) P_{ij} = \\ &= \sum_{k=0}^{M} x_k \sum_{i=0}^{+\infty} P_{ki} P_{ij} = \sum_{k=0}^{M} x_k P_{kj}^{(2)} \quad \forall M \end{aligned}$$

And this holds also in the limit $M \to \infty$. Therefore:

$$x_j \geq \sum_{k=0}^{+\infty} x_k P_{kj}^{(2)}$$

And by iterating this argument:

$$x_j \leq \sum_{k=0}^{+\infty} x_k P_{kj}^{(n)} \qquad \forall n \tag{6}$$

To prove that this is indeed an equality, we proceed as in point 3 of the previous proof, and assume that there is an index j for which (**??**) is strict. By a similar reasoning, this leads to a contradiction:

$$\sum_{j=0}^{+\infty} x_j > \sum_{k=0}^{+\infty} x_k$$

Therefore (**??**) must hold as an equality:

$$x_j = \sum_{k=0}^{+\infty} x_k P_{kj}^{(n)} \qquad \forall n$$

Finally, letting $n \to \infty$ and using the same argument as in point 4 to bring the limit inside the sum:

$$x_k = \lim_{n\to\infty} \sum_{k=0}^{+\infty} x_k P_{kj}^{(n)} = \sum_{k=0}^{+\infty} x_k \lim_{n\to\infty} P_{kj}^{(n)} = \left(\sum_{k=0}^{+\infty} x_k \right) \pi_j$$

and since $\sum_{k=0}^{+\infty} x_k = 1$ we have $x_j = \pi_j$, thus concluding the proof. □

Solving (**??**) suffices to say that an **aperiodic** Markov chain is *positive recurrent*. Conversely, proving that (**??**) does not admit solutions, means that the *aperiodic* Markov chain *is not* positive recurrent.
For a general Markov chain, however, a *stationary distribution* is not necessarily the same as the *limiting distribution*. In fact, if a limiting distribution exists, then it is stationary (i.e. it is the solution of (**??**)), but the converse is not true:

sometimes (**??**) can be solved, but the Markov chain is periodic and so that solution is clearly not the limiting distribution (that does not exist).
The simplest example of this kind of behaviour is given by:

$$\mathbf{P} = \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$$

This is a non-regular Markov chain that always cycles between states 0 and 1, thus presenting no *limiting distribution*. However it admits a *stationary distribution*, which is $\boldsymbol{\pi} = (1/2, 1/2)^T$. In fact:

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \times \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Example 1 (Stationary distribution for a random walk)**:**

Consider the following random walk:

$$\mathbf{P} = \begin{Vmatrix} 0 & 1 & 0 & \cdots & \cdots \\ q_1 & 0 & p_1 & \cdots & \cdots \\ 0 & q_2 & 0 & p_2 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{Vmatrix}$$

If $q_i, p_i > 0$, then the chain is **irreducible**, with $d(i) = 2$ (to return to the same state i, we need an *equal number* of steps in one direction, and in the opposite one, thus all $P_{ii}^{(n)}$ with n odd are 0). So it is **not** *aperiodic*, meaning that theorem **??** cannot be applied. However, there is a *stationary distribution*, i.e. we can solve:

$$\boldsymbol{x} = \boldsymbol{x}\mathbf{P} \Leftrightarrow x_i = \sum_{j=0}^{+\infty} x_j P_{ji} = p_{i-1}x_{i-1} + q_{i+1}x_{i+1} \qquad i > 0 \tag{7}$$

under the normalization:

$$\sum_{i=0}^{+\infty} x_i = 1$$

and with $p_0 = 1$, and $x_0 = q_1 x_1$.

Recall that we previously solved a similar equation while doing first-step analysis:

$$\boldsymbol{u} = \mathbf{P}\boldsymbol{u} \tag{8}$$

However we cannot directly apply the same method, as in (**??**) $\mathbf{P}$ multiplies the vector *from the left* and not from the right as in (**??**).

However, the idea is similar. We start by solving the first equation:

$$x_0 = q_1 x_1 \Rightarrow x_1 = \frac{x_0}{q_1}$$

And substitute in the second one:

$$x_1 = x_0 + q_2 x_2 \Rightarrow x_2 = \frac{x_1 - x_0}{q_2} = \frac{(1 - q_1)\, x_1}{q_2} = \frac{p_1\, x_0}{q_1 q_2} \tag{9}$$

where we used the *row-normalization* of $\mathbf{P}$, for which $p_i + q_i = 1$.
Repeating one more time:

$$x_2 = p_1 x_1 + q_3 x_3 \Rightarrow x_3 = \frac{x_2 - p_1 x_1}{q_3} \underset{(\mathbf{??})}{=} \frac{p_1 x_1 (1 - q_2)}{q_2 q_3} = \frac{p_1 p_2 x_0}{q_1 q_2 q_3}$$

From that we can *guess* the form of the general solution:

$$x_i = x_0 \frac{p_{i-1} p_{i-2} \cdots p_1}{q_i q_{i-1} \cdots q_i} \underset{p_0 = 1}{=} x_0 \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}} \qquad i > 0 \tag{10}$$

and substitute it back in (**??**) to check if it is right:

$$\begin{aligned} p_{i-1} x_{i-1} + q_{i+1} x_{i+1} &= p_{i-1} \frac{p_{i-2} \cdots p_1}{q_{i-1} \cdots q_1} + \cancel{q_{i+1}} \frac{p_i \cdots p_1}{\cancel{q_{i+1}} \cdots q_1} = \\ &= \frac{p_{i-1} \cdots p_1}{q_i \cdots q_1} \underbrace{(q_i + p_i)}_{1} = x_i \end{aligned}$$

All that's left is to *fix* the value of x_0 by imposing the normalization:

$$\sum_{i=0}^{+\infty} x_i = x_0 \sum_{i=0}^{+\infty} \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}} = 1 \Rightarrow x_0 = \left(\sum_{i=0}^{+\infty} \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}} \right)^{-1} \tag{11}$$

(with the convention that a product with *no elements*is equal to 1, the *neutral element* of the product: $\prod_{k=0}^{0} (\cdots) = 1$).

The *stationary* solution (**??**) exists only if the infinite sum in (**??**) converges to a non-zero finite value. If it were diverging, then $x_0 = 0$, and so all $x_i = 0\ \forall i$, meaning that the normalization constraint is not respecting.

Suppose that $p_k \equiv p$ and $q_k \equiv q$, i.e. the probabilities of moving in one or the other direction are *independent* of the system's state. In this case we can directly inspect the convergence of the sum in (**??**):

- If $p < q$, the sum converges, and the chain is *positive recurrent*. Intuitively, in this case the system "tends to return over its steps", thus visiting the same states over and over.

- If $p \geq q$, the sum diverges and no solution exists, meaning that the chain is not positive recurrent. Intuitively, in this case the system tends to "escape" towards infinity, always visiting new *transient* states.