Stationary distributions and limiting probabilities

Dr. Guangliang Chen
This lecture is based on the following textbook sections:

- Section 4.4

and also the following lecture: https://www.stat.uchicago.edu/~yibi/teaching/stat317/2013/Lectures/Lecture5_4up.pdf

Outline of the presentation

- Stationary distributions
- Limiting probabilities
- Long-run proportions
Assume a Markov chain \( \{X_n : n = 0, 1, 2, \ldots \} \) with state space \( S \) and transition matrix \( P \).

Let \( \pi = (\pi_i)_{i \in S} \) be a row vector denoting a probability distribution on \( S \), i.e.,

\[
\pi_i \geq 0, \quad \sum_{i \in S} \pi_i = 1.
\]

**Def 0.1.** \( \pi \) is called a **stationary** (or equilibrium) distribution of the Markov chain if it satisfies

\[
\pi = \pi P, \quad (\pi \text{ is a left eigenvector corresponding to 1})
\]

or in entrywise form,

\[
\pi_j = \sum_{i \in S} \pi_i p_{ij}, \quad \text{for all } j \in S.
\]
Remark. $\pi^T$ is a (right) eigenvector of $P^T$ corresponding to the same eigenvalue 1:

$$P^T \pi^T = \pi^T.$$

Note that $1$ is a (right) eigenvector of $P$ corresponding to eigenvalue 1:

$$P1 = 1.$$

In general, a square matrix $A$ and its transpose have the same eigenvalues

$$\text{det}(\lambda I - A^T) = \text{det}(\lambda I - A)$$

but they do not have the same eigenvectors.
Theorem 0.1. Let \( \{X_n : n = 0, 1, 2, \ldots\} \) be a Markov chain with a stationary distribution \( \pi \). If \( X_n \sim \pi \) for some integer \( n \geq 0 \), then \( X_{n+1} \sim \pi \).

Remark. This implies that for the same \( n \), the future states \( X_{n+2}, X_{n+3}, \ldots \) all have the same distribution \( \pi \).

Proof. For any \( j \in S \),

\[
P(X_{n+1} = j) = \sum_{i \in S} P(X_{n+1} = j \mid X_n = i)P(X_n = i)
\]

\[= \sum_{i \in S} p_{ij} \pi_i = \pi_j. \quad \square\]
Example 0.1. Find the stationary distribution of the Markov chain below:

\[ P = \begin{pmatrix}
  0 & .9 & .1 & 0 \\
  .8 & .1 & 0 & .1 \\
  0 & .5 & .3 & .2 \\
  .1 & 0 & 0 & .9 \\
\end{pmatrix} \]

Answer: \( \pi = (.2788, .3009, .0398, .3805) \) by software. Alternatively, we can solve \( \pi P = \pi \) (along with the requirement \( \sum \pi_i = 1 \)) directly by hand:

\[
\begin{align*}
\pi_1 &= 0.8\pi_2 + 0.1\pi_4 \\
\pi_2 &= 0.9\pi_1 + 0.1\pi_2 + 0.5\pi_3 \\
\pi_3 &= 0.1\pi_1 + 0.3\pi_3 \\
\pi_4 &= 0.1\pi_2 + 0.2\pi_3 + 0.9\pi_4 \\
\end{align*}
\]

\[ \begin{align*}
\pi_1 &= 63/226 \\
\pi_2 &= 68/226 \\
\pi_3 &= 9/226 \\
\pi_4 &= 86/226 \\
\end{align*} \]
Existence (and uniqueness) of stationary distributions

**Theorem 0.2.** For any irreducible Markov chain with state space $S$ and transition matrix $P$, it has a stationary distribution $\pi = (\pi_j)$:

$$ \forall j \in S: \pi_j \geq 0, \quad \sum_{i \in S} \pi_i = 1, \quad \pi = \pi P. $$

if and only if the chain is positive recurrent.

Furthermore, if a solution exists, then it will be unique and for state $j$,

$$ \pi_j = \begin{cases} \lim_{n \to \infty} p_{ij}^{(n)}, & \text{if the chain is aperiodic} \\ \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)}, & \text{if the chain is periodic} \end{cases} $$
Remark. In the aperiodic case, $\pi_j$ is also the limiting probability that the chain is in state $j$, i.e.,

$$\pi_j = \lim_{n \to \infty} P(X_n = j).$$

To prove this, let $\alpha = (\alpha_i)_{i \in S}$ be the initial distribution of the chain. Then

$$P(X_n = j) = \sum_{i \in S} P(X_n = j \mid X_0 = i) P(X_0 = i)$$

$$= \sum_{i \in S} p_{ij}^{(n)} \alpha_i \quad \xrightarrow{n \to \infty} \quad \pi_j \sum_{i \in S} \alpha_i = \pi_j.$$
**Example 0.2** (Social mobility). Let $X_n$ be a family’s social class: 1 (lower), 2 (middle), 3 (upper) in the $n$th generation. This was modeled as a Markov chain with transition matrix

$$
\mathbf{P} = \begin{pmatrix}
.8 & .1 & .1 \\
.2 & .6 & .2 \\
.3 & .3 & .4
\end{pmatrix}
$$

It is irreducible, positive recurrent and aperiodic (i.e., ergodic). Thus, there is a unique stationary distribution:

$$
\pi = \left( \frac{6}{11}, \frac{3}{11}, \frac{2}{11} \right) = (0.5454, 0.2727, 0.1818),
$$

and the chain will converge to the stationary distribution.
Math 263, Stationary distributions and limiting probabilities

\[ P = \begin{pmatrix} .8 & .1 & .1 \\ .2 & .6 & .2 \\ .3 & .3 & .4 \end{pmatrix} \quad \rightarrow \quad P^{10} = \begin{pmatrix} 0.5471 & 0.2715 & 0.1814 \\ 0.5430 & 0.2745 & 0.1825 \\ 0.5441 & 0.2737 & 0.1822 \end{pmatrix} \]

\[ \rightarrow \quad P^{20} = \begin{pmatrix} 0.5455 & 0.2727 & 0.1818 \\ 0.5454 & 0.2727 & 0.1818 \\ 0.5454 & 0.2727 & 0.1818 \end{pmatrix} \]
Example 0.3. Consider the following Markov chain:

\[ P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \]

It is irreducible and positive recurrent, and thus has a unique stationary distribution:

\[ \pi = \left( \frac{1}{2}, \frac{1}{2} \right) \]

The chain does not converge to the stationary distribution because it is periodic with period 2: For any integer \( \ell \geq 0 \),

\[ P^{2\ell} = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P^{2\ell+1} = P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \]
However, the following identity is still true:

\[ \pi_j = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} p^{(k)}_{ij} \]
Example 0.4 (Gambler’s Ruin). The underlying Markov chain has three communicating classes \( \{0\}, \{1, \ldots, N-1\}, \{N\} \), and thus it is not irreducible.

However, the chain has two stationary distributions (corresponding to the two recurrent classes):

\[ \pi_1 = (1, 0, \ldots, 0, 0), \quad \pi_2 = (0, 0, \ldots, 0, 1) \]
When \( N = 4 \) and \( \omega = \frac{1}{2} \) (symmetric random walk),

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\
0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\quad \longrightarrow \quad
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
\frac{3}{4} & 0 & 0 & 0 & \frac{1}{4} \\
\frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\
\frac{1}{4} & 0 & 0 & 0 & \frac{3}{4} \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\]

What does this imply?
Example 0.5. The 1-dimensional symmetric random walk over $\mathbb{Z}$ must be null recurrent.

\[
\begin{array}{c}
\cdots \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad \cdots \\
\end{array}
\]

\[
\begin{array}{c}
\text{1} - p \quad p
\end{array}
\]

(This is a homework question, #39. Use proof by contradiction)
Consider the Markov chain defined on a finite, undirected, weighted graph \( G = \{V, E, W\} \), with state space \( S = V \) and transition matrix

\[
P = D^{-1}W, \quad D = \text{diag}(d), \quad d = W \cdot 1
\]

The chain is finite, and if the graph is connected, then the Markov chain must be irreducible and also positive recurrent. Accordingly, it possesses a unique stationary distribution.
Proposition 0.3. For any finite, connected graph, the induced Markov chain possesses the following unique stationary distribution

\[ \pi = \frac{1}{\text{Vol}(V)} \cdot d, \quad \text{where} \quad \text{Vol}(V) = \sum_{i \in V} d_i. \]

If the graph is also non-bipartite, then the chain always converges to the above stationary distribution.

Proof. First, we show that

\[ dP = dD^{-1}W = 1^T W = d \quad \rightarrow \quad \pi P = \pi. \]

Thus, \( \pi \) is a stationary distribution of the chain and it is also unique.
For the convergence part, we consider the following two cases:

(1) Bipartite graphs (no convergence, because $d = 2$)

(2) Non-bipartite graphs (convergence)
Long-run proportion of visits to a state

Theorem 0.4. For an irreducible, positive recurrent Markov chain with stationary distribution \( \pi = (\pi_j) \), \( \pi_j \) is also the long-run proportion of time that the chain is in state \( j \) (regardless of initial state \( i \)).
Proof. To see this, let

\[ I_n = 1_{X_n = j}, \quad \text{for all } n \geq 1 \]

and define

\[ T = \sum_{n=1}^{\ell} I_n \]

which represents the total number of visits to state \( j \) in \( \ell \) steps.

The proportion of visits to state \( j \) in \( \ell \) steps is

\[ \frac{T}{\ell} = \frac{1}{\ell} \sum_{n=1}^{\ell} I_n, \]
and we would like to show that it converges to $\pi_j$ on average:

$$E \left[ \frac{T}{\ell} \bigg| X_0 = i \right] = \frac{1}{\ell} \sum_{n=1}^{\ell} E[I_n \mid X_0 = i]$$

$$= \frac{1}{\ell} \sum_{n=1}^{\ell} 1 \cdot P(I_n = 1 \mid X_0 = j) + 0 \cdot P(I_n = 0 \mid X_0 = j)$$

$$= \frac{1}{\ell} \sum_{n=1}^{\ell} P(X_n = j \mid X_0 = j)$$

$$= \frac{1}{\ell} \sum_{n=1}^{\ell} p_{ij}^{(n)} \quad \ell \to \infty \quad \pi_j.$$
Example 0.6. Three out of every four trucks on the road are followed by a car, while only one out of every five cars is followed by a truck. What fraction of vehicles on the road are trucks?

\[
C \quad C \quad T \quad C \quad C \quad C \quad T \quad T \quad C \quad C \quad C \quad C \quad T \quad C \quad C \quad C
\]
Solution. Let $X_n$ be the type of the $n$th vehicle, $T$ (for truck) or $C$ (for car), when counting from one end of the road to the other end. Then \( \{X_n, n \geq 1\} \) is a Markov chain with state space \( S = \{T, C\} \) and corresponding transition matrix

\[
P = \begin{bmatrix}
\frac{1}{4} & \frac{3}{4} \\
\frac{1}{5} & \frac{4}{5}
\end{bmatrix}
\]

Since the chain is irreducible and positive recurrent, it has a unique stationary distribution \( \pi = (\pi_T, \pi_C) \) given by

\[
\pi_T = \pi_T \cdot \frac{1}{4} + \pi_C \cdot \frac{1}{5}, \quad \pi_T + \pi_C = 1 \quad \rightarrow \quad \pi_T = \frac{4}{19}, \quad \pi_C = \frac{15}{19}
\]

The fraction of trucks on the road is the long-run proportion \( \pi_T = \frac{4}{19} \).
**Theorem 0.5.** For any irreducible, positive recurrent Markov chain, with stationary distribution \( \pi = (\pi_j) \), we must have

\[
\pi_j = \frac{1}{m_{jj}} \quad \text{for all } j \in S,
\]

where \( m_{jj} \) represents the mean recurrence time of state \( j \):

\[
m_{jj} = E(N_j \mid X_0 = j).
\]

**Remark.** This theorem implies that \( \pi_j > 0 \) for all positive recurrent states \( j \) in an irreducible chain (as \( m_{jj} < \infty \) for all \( j \)). Note that \( \pi_j \) can also be interpreted as the long-run proportion of the chain being in state \( j \) here.
Proof. To see this, consider

\[ T = \sum_{n=1}^{\ell} I_n, \quad I_n = 1_{X_n=j} \]

which represents the total number of visits to state \( j \) in \( \ell \) time steps.

Denote by \( N_j^1, \ldots, N_j^T \) the individual recurrence times in the \( \ell \) time steps:
Then
\[ N_j^1 + \cdots + N_j^T \leq \ell < N_j^1 + \cdots + N_j^T + N_j^{T+1}, \]

where \( N_j^{T+1} \) represents the additional number of time steps that will be needed by the chain to enter state \( j \) again (after the first \( T \) visits).

Taking conditional expectation \( \mathbb{E} [ \cdot \mid X_0 = j] \) of left-hand side gives that

\[
\mathbb{E} \left( N_j^1 + \cdots + N_j^T \middle| X_0 = j \right) = \mathbb{E} \left[ \mathbb{E} \left( N_j^1 + \cdots + N_j^T \middle| X_0 = j, T \right) \middle| X_0 = j \right] \\
= \mathbb{E} \left[ T \cdot \mathbb{E} \left( N_j^1 \middle| X_0 = j \right) \middle| X_0 = j \right] \\
= \mathbb{E} \left[ T \cdot m_{jj} \middle| X_0 = j \right] \\
= m_{jj} \cdot \mathbb{E} \left[ T \middle| X_0 = j \right]
\]
Similarly,

\[
E\left(N_j^1 + \cdots + N_j^T + N_j^{T+1} \mid X_0 = j \right) = m_{jj} \cdot E\left[T + 1 \mid X_0 = j \right]
\]

\[
= m_{jj} + m_{jj} \cdot E\left[T \mid X_0 = j \right]
\]

Combining them together, we have

\[
\begin{align*}
m_{jj} \cdot E[T \mid X_0 = j] & \leq \ell < m_{jj} + m_{jj} \cdot E[T \mid X_0 = j] \\
m_{jj} \cdot \frac{1}{\ell} E[T \mid X_0 = j] & \leq 1 < m_{jj} \left( \frac{1}{\ell} + \frac{1}{\ell} E[T \mid X_0 = j] \right)
\end{align*}
\]
We next derive an expression for $E[T \mid X_0 = j]$:

$$E[T \mid X_0 = j] = \sum_{n=1}^{\ell} E(I_n \mid X_0 = j)$$

$$= \sum_{n=1}^{\ell} 1 \cdot P(I_n = 1 \mid X_0 = j) + 0 \cdot P(I_n = 0 \mid X_0 = j)$$

$$= \sum_{n=1}^{\ell} P(X_n = j \mid X_0 = j)$$

$$= \sum_{n=1}^{\ell} p^{(n)}_{jj}$$
It follows that

\[ m_{jj} \cdot \frac{1}{\ell} \sum_{n=1}^{\ell} p_{jj}^{(n)} \leq 1 < m_{jj} \cdot \left( \frac{1}{\ell} + \frac{1}{\ell} \sum_{n=1}^{\ell} p_{jj}^{(n)} \right) \]

Letting \( \ell \to \infty \) yields that

\[ m_{jj} \cdot \pi_j \leq 1 \leq m_{jj} \cdot (0 + \pi_j) \]

So we must have

\[ m_{jj} \cdot \pi_j = 1, \quad \text{and thus} \quad \pi_j = \frac{1}{m_{jj}}. \]
Remark. If the chain is irreducible but null recurrent, then $m_{jj} = \infty$ for all states $j$. Such a Markov chain may have no stationary distribution $\pi$ (e.g., the 1D symmetric random walk over $\mathbb{Z}$).

However, we can still talk about the long-run proportion of the chain being in state $j$:

$$\frac{T}{\ell} = \frac{1}{\ell} \sum_{n=1}^{\ell} I_n, \quad \text{as } \ell \to \infty.$$ 

Starting with the inequality

$$N_j^1 + \cdots + N_j^T \leq \ell \quad \text{for all } \ell$$
we take conditional expectation \( E[\cdot \mid X_0 = j] \) and repeat the same steps to obtain that

\[
m_{jj} \cdot E[T \mid X_0 = j] \leq \ell \quad \text{for all } \ell
\]

or equivalently,

\[
m_{jj} \cdot E[T/\ell \mid X_0 = j] \leq 1 \quad \text{for all } \ell
\]

Because state \( j \) is null recurrent (\( m_{jj} = \infty \)), we must have

\[
E[T/\ell \mid X_0 = j] = 0 \quad \text{for all } \ell
\]

This shows that the long run proportion of visits to state \( j \) is zero. Thus, if \( \pi_j \) represents the long-run proportion of state \( j \) (instead of a stationary probability), then the formula \( \pi_j = \frac{1}{m_{jj}} \) is still valid.
**Theorem 0.6.** Positive recurrence is a class property. That is, if state \( j \) is positive recurrent, and state \( j \) communicates with state \( k \), then state \( k \) is also positive recurrent.

**Proof.** (We cannot use the stationary distribution as we do not know whether it exists; we’ll consider long-run proportions instead)

First, there exists a positive integer \( n \) such that

\[
p_{jk}^{(n)} > 0
\]

Since state \( j \) is positive recurrent, the long-run proportion is

\[
\pi_j = 1/m_{jj} > 0
\]
For any positive integer $t$ and state $i$, we have

$$p_{ik}^{(t+n)} \geq p_{ij}^{(t)} \cdot p_{jk}^{(n)}$$

and also

$$\frac{1}{\ell} \sum_{t=1}^{\ell} p_{ik}^{(t+n)} \geq \left( \frac{1}{\ell} \sum_{t=1}^{\ell} p_{ij}^{(t)} \right) \cdot p_{jk}^{(n)}$$

Letting $\ell \to \infty$, we obtain that

$$\pi_k \geq \pi_j \cdot p_{jk}^{(n)} > 0$$

where $\pi_k$ represents the long-run proportion of visits to state $k$. It follows that

$$m_{kk} = \frac{1}{\pi_k} < \infty$$

and thus state $k$ is also positive recurrent.