2.2 Uniqueness

Remark 5. Consider a Markov chain $(X_0)_{t\geq 0}$ with finite state set S and transition probabilities K. Initializing the chain with a probability distribution μ on S, the consecutive states have distributions $\mu, \mu K, \mu K^2, \ldots$. Recall from the proof of the existence theorem that any convergent subsequence of the partial averages

$$\frac{1}{n}(\mu + \mu K + \cdots + \mu K^{n-1}) \tag{17}$$

converges to a stationary distribution. If the Markov chain happens to have only one stationary distribution, then it follows that the partial averages (17) converge to this unique stationary distribution, *independent of the initial distribution* μ.

In particular, if $f : S \to \mathbb{R}$ is a function of the state (an *observable*), it follows that

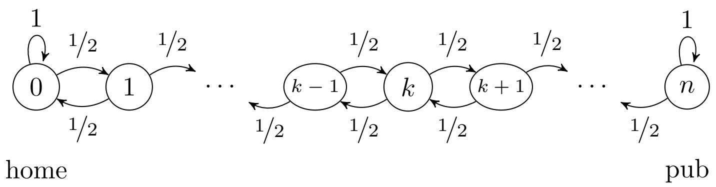
$$\mathbf{E}\left[\frac{1}{n}\left(f(X_0) + f(X_1) + \cdots + f(X_{n-1})\right)\right] \to \pi(f) , \tag{18}$$

independent of the distribution of X_0. (Recall the notation: $\pi(f) \triangleq \sum_{a\in S} \pi(a) f(a)$ is the expected value of f with respect to the distribution π.) Later we shall see that a much stronger statement holds: the averages

$$\frac{1}{n}\left(f(X_0) + f(X_1) + \cdots + f(X_{n-1})\right) \tag{19}$$

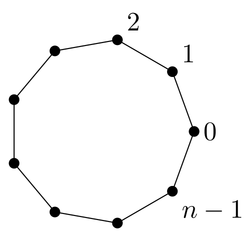
converge to $\pi(f)$ *almost surely* (i.e., with probability 1). In other words, the random variables $f(X_0), f(X_1), \ldots$ satisfy a version of the law of large numbers. This is the so-called ergodic theorem of Markov chains. ◇

Example 6 (Non-uniqueness)**.** The Markov chain of the drunkard's walk (a.k.a. the gambler's ruin)



has at least two stationary distributions: the distribution δ_{home} concentrated at "home", and the distribution δ_{pub} concentrated at the "pub". In fact, every convex combination $\lambda\delta_{\text{home}} + (1-\lambda)\delta_{\text{pub}}$ (for $0 \leq \lambda \leq 1$) is also a stationary distribution. ○

Example 7 (Uniqueness)**.** The simple random walk on a cycle $\mathbb{Z}_n$ of length n



has an obvious stationary distribution: the uniform distribution $\pi(x) \triangleq 1/n$. There is a simple argument showing that this is the only stationary distribution. For, let μ be a stationary distribution that is not uniform. Pick a vertex $a \in \mathbb{Z}_n$ such that

- $\mu(a)$ is maximum, and

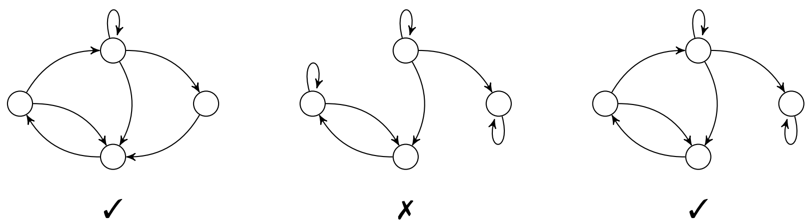
- $\mu(a-1) < \mu(a)$.

Then,

$$\mu K(a) = \underbrace{\mu(a-1)}_{<\mu(a)} \cdot \frac{1}{2} + \underbrace{\mu(a+1)}_{\leq \mu(a)} \cdot \frac{1}{2} < \mu(a) \ , \tag{20}$$

which means μ is not a stationary distribution. ○

The above examples (as well as others we have seen so far) might suggest that whether a (finite-state) Markov chain one or more stationary distribution depends not on the actual transition probabilities $K(a, b)$ but only on the combinatorics of the transitions between different states. Indeed, by looking at the underlying graph of probable transitions one could guess which Markov chain has a unique stationary distribution and which has more.



(By the underlying graph of a Markov chain we mean a directed graph with a vertex for each state and a directed edge $a \to b$ for each two states a and b with $K(a, b) > 0$.)

A Markov chain $(X_t)_{t\geq 0}$ is said to be *irreducible* if for every two states a, b there is an integer $n > 0$ such that

$$\mathbf{P}(X_{t+n} = b \,|\, X_t = a) = K^n(a, b) > 0 \ . \tag{21}$$

A Markov chain is irreducible if and only if its underlying graph is strongly connected.

Lemma 8. *Any stationary distribution for an irreducible Markov chain is (strictly) positive.*

Proof. Let $(X_t)_{t\geq 0}$ be an irreducible Markov chain initialized according to a stationary distribution π. Let b be an arbitrary state. Choose a state a such that $\mathbf{P}(X_0 = a) = \pi(a) > 0$. By irreducibility, there is a time $n \geq 0$ such that $\mathbf{P}(X_n = b \,|\, X_0 = a) > 0$. It follows that

$$\begin{aligned} \pi(b) = \mathbf{P}(X_n = b) &\geq \mathbf{P}(X_0 = a, X_n = b) && (22) \\ &= \mathbf{P}(X_0 = a)\,\mathbf{P}(X_n = b \,|\, X_0 = a) > 0 \ . && (23) \end{aligned}$$

□

Theorem 9. *An irreducible Markov chain has at most one stationary distribution.*

Remark 10. The argument used in Example 7 can be adapted to prove the above theorem in case of finite-state irreducible Markov chains (see [LPW08], Exercise 1.13). Various other proofs are possible, each with its own insight. Below, we shall take an alternative approach that has the advantage of providing an explicit probabilistic expression for the unique stationary distribution. Yet another proof is by defining a notion of "entropy" associated to each probability distribution and showing that the entropy increases unless the chain is in equilibrium. ◇

For each state x, consider the average time m_x it takes for the chain to return to x if started from x. If the chain is irreducible, every state x is visited over and over again, and the gap between every two consecutive visits is on average m_x. The law of large numbers suggests that on the long run, the chain is spending roughly $1/m_x$ of its time in state x.

To be specific, let us denote by T_x the first time $t > 0$ that the chain visits a state x, that is

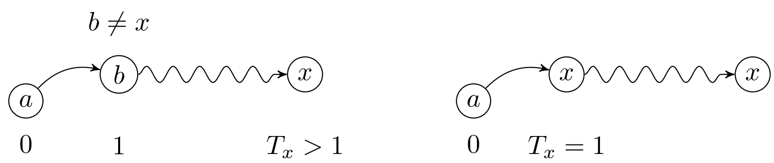
$$T_x \triangleq \inf\{t > 0 : X_t = x\} \ . \tag{24}$$

Theorem 11. *Let $(X_t)_{t \geq 0}$ be an irreducible Markov chain and π a stationary distribution. Then,*

$$\pi(x) = \frac{1}{\mathbf{E}_x\, T_x} \tag{25}$$

for every $x \in S$.

Proof. Let us initialize the chain with the stationary distribution π and write the expected hitting time of x by conditioning on the initial state and the state after on step. There are two cases, based on whether the chain hits x in its first step or not:



$$\mathbf{E}[T_x \mid X_0 = a, X_1 = b] = \begin{cases} 1 + \mathbf{E}_b\, T_x & \text{if } b \neq x, \\ 1 & \text{if } b = x. \end{cases} \tag{26}$$

Therefore,

$$\mathbf{E}\, T_x = \mathbf{E}[\mathbf{E}[T_x \mid X_0, X_1]] = 1 + \sum_a \pi(a) \underbrace{\sum_{b \neq x} K(a,b)\, \mathbf{E}_b\, T_x}_{\sum_b K(a,b)\, \mathbf{E}_b\, T_x - K(a,x)\, \mathbf{E}_x\, T_x} \ , \tag{27}$$

which leads to

$$\mathbf{E}\, T_x = 1 + \underbrace{\sum_a \sum_b \pi(a) K(a,b)\, \mathbf{E}_b\, T_x}_{\sum_b \pi(b)\, \mathbf{E}_b\, T_x} - \underbrace{\sum_a \pi(a) K(a,x)}_{\pi(x)} \mathbf{E}_x\, T_x \ . \tag{28}$$

It follows that

$$\cancel{\mathbf{E}\, T_x} = 1 + \cancel{\sum_b \pi(b)\, \mathbf{E}_b\, T_x} - \pi(x)\, \mathbf{E}_x\, T_x \ , \tag{29}$$

hence proving $\pi(x)\, \mathbf{E}_x\, T_x = 1$. ∼∼∼∼∼∼∼∼ □

Remark 12. If an irreducible Markov chain (with finite or countable state set) has a stationary distribution, Theorem 11 along with Lemma 8 imply that the expected return time $\mathbf{E}_x\, T_x$ of each state x is finite.

An irreducible Markov chain for which the expected return times are not finite does not need to have a stationary distribution. (An example is the simple random walk on $\mathbb{Z}$.) However, if every state does have a finite expected return time (such a chain is said to be *positively recurrent*), then an adaptation of the proof of Theorem 11 can be used to show that

$$\pi(x) \triangleq \frac{1}{\mathbf{E}_x\, T_x} \tag{30}$$

indeed defines a stationary distribution. ∼∼∼∼∼∼∼ ◇