

Lecture 36: Stationary Distributions of Markov Chains

Charles A. Bouman

Purdue University

Wednesday, November 18, 2026

Outline

- 1 Markov chains as a computational tool
- 2 The evolution equation $\pi^{(n+1)} = \pi^{(n)}P$
- 3 Steady state: $\pi^\infty = \pi^\infty P$, and the eigenvalue of one
- 4 Three things that can go wrong: trapped, periodic, dispersed
- 5 Example 13.2: three counterexamples
- 6 Communicating states and irreducible Markov chains
- 7 Periodic states and aperiodic chains
- 8 The Limit Theorem for Markov Chains
- 9 The full-balance equations; ergodic Markov chains

Core idea

A well-behaved Markov chain reaches a stable stationary distribution after the transient behavior dies away. We need tools both for guaranteeing that a steady-state distribution is reached and for determining precisely what that distribution is.

Today's Big Ideas

- The distribution of the state at the next time is the current distribution multiplied by the transition matrix, so the distribution at any future time is found by raising that matrix to a power.
- If the chain settles into a distribution that does not depend on where it started, then that distribution must be left unchanged by one more step of the chain.
- Not every Markov chain settles down. The distribution can stay trapped in part of the states, it can cycle forever, or it can spread out until every state has probability zero.
- Three conditions rule out those three failures: finitely many states, the ability to get from any state to any other, and no fixed cycle length.
- A chain that meets those conditions is called ergodic. It has one stationary distribution, it converges to that distribution from any starting point, and expectations can be computed as time averages along a single run.

Markov chains are also a computational tool

In addition to using Markov chains as models of data, they can also serve as a very powerful computational tool.

- In Chapter 15 we will use **Monte Carlo Markov chain** techniques to generate samples from very general posterior distributions.
- Similar methods can also be used to numerically compute expectations/integrals, or to solve optimization problems.

So this section is not a detour. It is the theory that makes the stochastic simulation methods of Chapter 15 work.

The framework: ergodic Markov chains

In order to provide a framework for these methods, we will first need to introduce the theory of **ergodic Markov chains**.

The key concept

A well-behaved Markov chain should eventually reach a stable **stationary distribution** after the transient behavior dies away.

So we will need tools for two separate jobs:

- **guaranteeing** that such a steady-state distribution is reached, and
- **determining precisely** what that steady-state distribution is.

How the marginal density evolves

Again, let $\{X_n\}_{n=0}^{\infty}$ be a homogeneous Markov chain with marginal density given by

$$\pi_j^{(n)} = P\{X_n = j\} ,$$

and transition probabilities given by

$$P_{i,j} = P\{X_n = j | X_{n-1} = i\} .$$

Then a fundamental property of Markov chains is that the marginal probability density for time $n + 1$ can be expressed as

$$\pi_j^{(n+1)} = \sum_{i \in \Omega} \pi_i^{(n)} P_{i,j} .$$

Vector-matrix form

When the number of states is finite, then we can more compactly express these ideas using vector-matrix notation.

- Let $\pi^{(n)}$ be a $1 \times M$ **row vector** containing the marginal densities.
- Let P be an $M \times M$ matrix containing the transition probabilities.

Then in matrix notation, we have that

$$\pi^{(n+1)} = \pi^{(n)} P .$$

Repeated application of this equation results in

$$\pi^{(n+m)} = \pi^{(n)} P^m ,$$

where P^m denotes the matrix raised to the m^{th} power. So we can compute the marginal distribution at any time in the future by repeatedly multiplying by the transition matrix P . In particular,

$$\pi^{(n)} = \pi^{(0)} P^n .$$

What if the chain reaches a steady state?

Intuitively, we might expect that as n grows large, the Markov chain should reach some type of steady state.

If that were to hold, then we must have that $\lim_{n \rightarrow \infty} P^n$ exists and that

$$\pi^\infty = \lim_{n \rightarrow \infty} \pi^{(0)} P^n = \pi^{(0)} P^\infty .$$

Notice that so far this says nothing about whether the answer depends on where the chain started.

Forgetting the initial distribution

Moreover, we might expect that the steady-state behavior should not depend on the initial distribution, $\pi^{(0)}$. So if this holds, then we know two important things.

First, each row of P^∞ must equal π^∞ , so we have that

$$P^\infty = \mathbf{1}\pi^\infty ,$$

where $\mathbf{1}$ is an $M \times 1$ vector of ones.

Second, π^∞ must solve the equation

$$\pi^\infty = \pi^\infty P .$$

An eigenvalue observation

This last equation also indicates that the matrix P has an eigenvalue of one, with an associated eigenvector of π^∞ .

Three things that can go wrong

Unfortunately, it is **not guaranteed** that a homogeneous Markov chain will reach a steady-state distribution that is invariant to the initial distribution. As a matter of fact, it is relatively easy to produce counterexamples.

Fortunately, the things that can go wrong can be grouped into three basic categories, and we can provide technical conditions to guard against these cases:

- The density can become **trapped in disconnected states**.
- The density can become **periodic**.
- The density can **disperse**.

Next we present three corresponding counterexamples illustrating these three behaviors.

Example 13.2: trapped in disconnected states

Let the Markov chain take values in the set $\{-1, 1\}$. Let

$$X_{n+1} = X_n ,$$

and let the initial distribution be given by $\pi^{(0)}$.

In this case, it is obvious that the steady-state distribution will always be given by

$$\pi_i^\infty = \lim_{n \rightarrow \infty} P\{X_n = i\} = \pi_i^{(0)} .$$

Therefore, the steady-state distribution is **not independent of the initial distribution**.

Example 13.2: a periodic chain

Let the Markov chain take values in the set $\{-1, 1\}$. Let $X_0 = 1$, and let

$$X_{n+1} = -X_n .$$

In this case, it can be easily shown that

$$P\{X_n = 1\} = \begin{cases} 1 & \text{if } n \text{ is even} , \\ 0 & \text{if } n \text{ is odd} \end{cases} .$$

Therefore, $\lim_{n \rightarrow \infty} P\{X_n = 1\}$ **does not exist**.

Example 13.2: the density disperses

Let the Markov chain take on integer values. Let $X_0 = 0$, and let

$$X_{n+1} = X_n + B_n ,$$

where B_n is an i.i.d. sequence of random variables with

$$P\{B_n = 1\} = P\{B_n = -1\} = 1/2 .$$

In this case, it can be shown that for all integers i ,

$$\pi_i^\infty = \lim_{n \rightarrow \infty} P\{X_n = i\} = 0 .$$

Therefore, the density function **disperses to zero**.

Definition: Communicating States

So if such simple examples of homogeneous Markov chains can result in counterexamples, how can we ensure convergence to a steady state? Of course, this is a huge field of research and many results exist, but we will provide some simple conditions that when satisfied ensure ergodic behavior. We start with two definitions used to guard against disconnected states.

Definition: Communicating States

The states $i, j \in \Omega$ of a discrete-time discrete-state homogeneous Markov chain are said to **communicate** if there exist integers $n > 0$ and $m > 0$ such that $P_{i,j}^n > 0$ and $P_{j,i}^m > 0$.

Intuitively, two states communicate if it is possible to transition between the two states.

Definition: Irreducible Markov Chain

It is easily shown that communication is an **equivalence property**, so it partitions the set of states into disjoint subsets, and only states within a subset communicate with each other. This leads to a natural definition for irreducible Markov chains.

Definition: Irreducible Markov Chain

A discrete-time discrete-state homogeneous Markov chain is said to be **irreducible** if for all $i, j \in \Omega$, i and j communicate.

So a Markov chain is irreducible if it is possible to change from any initial state to any other state in finite time.

Definition: Periodic State

The next condition is used to guard against periodic repetition of states that can last indefinitely and will not allow convergence to steady state.

Definition: Periodic State

The **period** of a state $i \in \Omega$ is the value $d(i)$, where $d(i)$ is the largest integer so that $[P^n]_{i,i} = 0$ whenever n is not divisible by $d(i)$. If $d(i) > 1$, then we say that the state i is **periodic**.

- It can be shown that states of a Markov chain that communicate must have the **same period**.
- Therefore, all the states of an irreducible Markov chain must have the same period.
- We say that an irreducible Markov chain is **aperiodic** if all the states have period one.

Theorem: Limit Theorem for Markov Chains

Let $X_n \in \Omega$ be a discrete-state discrete-time homogeneous Markov chain with transition probabilities $P_{i,j}$ and the following additional properties:

- Ω is a finite set.
- The Markov chain is irreducible.
- The Markov chain is aperiodic.

Then there exists a unique stationary distribution π , which for all states i is given by

$$\pi_j = \lim_{n \rightarrow \infty} [P^n]_{i,j} > 0 ,$$

and which is the unique solution to the following set of equations:

$$\begin{aligned} 1 &= \sum_{i \in \Omega} \pi_i , \\ \pi_j &= \sum_{i \in \Omega} \pi_i P_{i,j} . \end{aligned}$$

The full-balance equations, and ergodic chains

The two relations of the theorem are called the **full-balance equations**.

- Any probability density which solves the full-balance equations is guaranteed to be the stationary distribution of the Markov chain.
- Furthermore, in the limit as $n \rightarrow \infty$, the Markov chain is guaranteed to converge to this stationary distribution **independent of the initial state**.
- Markov chains that have the property $\pi_j = \lim_{n \rightarrow \infty} [P^n]_{i,j} > 0$ are said to be **ergodic**.

Why we will care

It can be shown that for ergodic Markov chains, expectations of state variables can be replaced by **time averages**, which will be very useful in later sections.

Summary

- The marginal density evolves as $\pi_j^{(n+1)} = \sum_{i \in \Omega} \pi_i^{(n)} P_{i,j}$, or in matrix form $\pi^{(n+1)} = \pi^{(n)} P$, so $\pi^{(n)} = \pi^{(0)} P^n$.
- If a steady state exists and does not depend on $\pi^{(0)}$, then $P^\infty = \mathbf{1}\pi^\infty$ and $\pi^\infty = \pi^\infty P$; so P has an eigenvalue of one with eigenvector π^∞ .
- Three counterexamples show that a chain need not settle down: $X_{n+1} = X_n$ keeps the initial distribution, $X_{n+1} = -X_n$ has no limit, and the ± 1 random walk gives $\pi_i^\infty = 0$ for every state i .
- Two states communicate if $P_{i,j}^n > 0$ and $P_{j,i}^m > 0$ for some $n, m > 0$, a chain is irreducible if every pair of states communicates, and an irreducible chain is aperiodic if every state has period $d(i) = 1$.
- Limit Theorem: Ω finite, irreducible, and aperiodic imply a unique stationary distribution $\pi_j = \lim_{n \rightarrow \infty} [P^n]_{i,j} > 0$, the unique solution of the full-balance equations.
- Such chains are called **ergodic**, and for them expectations can be replaced by time averages.

Reading: Section 13.4. Next lecture: Section 13.5, Properties of Markov Chains.