

Lecture 37: Properties of Markov Chains

Charles A. Bouman

Purdue University

Friday, November 20, 2026

Outline

- 1 Existence of a stationary distribution versus computing it
- 2 Conditioning on the past of a Markov chain
- 3 Subsampling a Markov chain
- 4 The time-reverse of a Markov chain
- 5 From the steady-state joint distribution to detailed balance
- 6 Reversible Markov chains
- 7 Detailed balance implies full balance
- 8 Example 13.3: the birth-death process

Core idea

Detailed balance is the practical route to a stationary distribution.

Not every Markov chain is reversible, but when a distribution π solves the detailed-balance equations, it also solves the full-balance equations, and so it is the stationary distribution.

Today's Big Ideas

- A Markov chain conditioned on a set of past samples depends on those samples only through the most recent one.
- Subsampling a Markov chain at a fixed spacing gives another Markov chain, and a homogeneous chain stays homogeneous.
- Running a Markov chain backward in time also gives a Markov chain, because Bayes' rule reverses the direction of a conditional dependence.
- A chain is reversible when its backward transition probabilities equal its forward ones, which is the same as saying that any two states exchange probability at equal rates in both directions.
- A distribution that satisfies the detailed-balance equations also satisfies the full-balance equations, so solving detailed balance is a practical way to find the stationary distribution of an ergodic chain.

The birth-death process is today's worked example. Chapter 15 builds samplers the same way, by choosing transition probabilities that satisfy detailed balance for a desired posterior distribution.

Existence is not the same as computation

Last lecture's limit theorem gives relatively simple conditions to establish that a Markov chain has a stationary distribution.

- However, while it may be known that a stationary distribution exists, it may be **very difficult to compute the solution of the full-balance equations** in order to determine the precise form of that distribution.
- Recall the full-balance equations from Lecture 36:

$$1 = \sum_{i \in \Omega} \pi_i, \quad \pi_j = \sum_{i \in \Omega} \pi_i P_{i,j}.$$

- Fortunately, Markov chains have a variety of properties that can be extremely useful in both quantifying and understanding their convergence.

This section presents a sampling of some of the most useful properties.

Conditioning on the past

Our first property is that the conditional distribution of X_n given some past set of samples, X_r for $r < k$, depends only on the most recent sample, X_k .

Property: Conditioning on the past of a Markov chain

Let $\{X_n\}_{n=0}^{\infty}$ be a Markov chain taking values on Ω , and let $k < n$ be two integers. Then

$$P\{X_n = x_n | X_r \text{ } r \leq k\} = P\{X_n = x_n | X_k\} .$$

- Intuitively, the present sample depends only on the most recent sample that is available in the Markov chain.

Subsampling: the derivation

This property can be used recursively to show that any subsampling of a Markov chain must also be a Markov chain. So, for example, if $Y_n = X_{dn}$, where d is a positive integer, then we have that

$$\begin{aligned} P\{Y_n = y_n | Y_r, r < n\} &= P\{X_{dn} = x_{dn} | X_{dr}, r < n\} \\ &= \mathbb{E}[P\{X_{dn} = x_{dn} | X_r, r \leq d(n-1)\} | X_{dr}, r < n] \\ &= \mathbb{E}[P\{X_{dn} = x_{dn} | X_{d(n-1)}\} | X_{dr}, r < n] \\ &= P\{X_{dn} = x_{dn} | X_{d(n-1)}\} \\ &= P\{Y_n = y_n | Y_{n-1}\} . \end{aligned}$$

The next slide lists the property used at each equality.

Subsampling: which properties were used

- Equality 1: the definition $Y_n = X_{dn}$.
- Equality 2: the *filtration property of conditional expectation* (Chapter 2).
- Equality 3: the *conditioning on the past* property of the previous slide.
- Equality 4: the corollary of the *known quantities* property (Chapter 2),

$$\mathbb{E}[\mathbb{E}[X|Y, Z] | Y] = \mathbb{E}[X|Y] = \mathbb{E}[\mathbb{E}[X|Y] | Y, Z] .$$

- Equality 5: the definition $Y_n = X_{dn}$ again.

Property: Subsampling of a Markov chain

Let $\{X_n\}_{n=0}^{\infty}$ be a Markov chain taking values on Ω , and let $Y_n = X_{dn}$, where d is a positive integer. Then Y_n is a Markov chain. Furthermore, if X_n is a homogeneous Markov chain, then Y_n is also a homogeneous Markov chain.

Reversing time

Perhaps surprisingly, it turns out that the concepts of reversibility are often extremely useful in determining the stationary distributions of ergodic Markov chains. So we will need to derive some important conditions related to reversible Markov chains.

Our first observation is that the time-reverse of any Markov chain results in another process with the Markov property.

Property: Time-reverse of a Markov chain

Let $\{X_n\}_{n=0}^N$ be a Markov chain taking values on Ω , and let $Y_n = X_{N-n}$ for $n = 0, \dots, N$. Then Y_n is a Markov chain.

- Perhaps this result is surprising given the causal definition of a Markov chain. But, in fact, there is hidden symmetry in the definition that arises through the ability to reverse conditional dependence using Bayes' rule.

The joint distribution in steady state

The time-reverse property provides a powerful tool for us to analyze the convergence of a class of Markov chains.

Consider the case when X_n is a homogeneous Markov chain which is in steady state with the stationary distribution π . In general, we can write the joint distribution of (X_n, X_{n-1}) as

$$\begin{aligned} P\{X_n = j, X_{n-1} = i\} &= P\{X_n = j | X_{n-1} = i\} P\{X_{n-1} = i\} \\ &= P\{X_{n-1} = i | X_n = j\} P\{X_n = j\} . \end{aligned}$$

So using this equality and the definitions of π_i and $P_{i,j}$, we have that

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

The transition probabilities of the reversed chain

Now we know that the time-reversed process is also a Markov chain, so let its transition probabilities be denoted by

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

- Notice that from the form of the previous equation, $Q_{j,i}$ is not a function of n , so the time-reversed Markov chain must also be **homogeneous**.

Plugging back in, we get an expression which must hold for any homogeneous Markov chain in steady state:

$$\pi_i P_{i,j} = \pi_j Q_{j,i} .$$

The detailed-balance equations

Now if in addition the Markov chain is reversible, then we know that the transition probabilities of the process running forward in time must be the same as the transition probabilities running backward in time. Formally, this means that $P_{i,j} = Q_{i,j}$.

Substituting this equality into $\pi_i P_{i,j} = \pi_j Q_{j,i}$ yields the **detailed-balance equations**:

$$\pi_i P_{i,j} = \pi_j P_{j,i} .$$

What the equations say

The detailed-balance equations specify that the rate of transitions from state i to state j equals the rate of transitions from state j to i . This is always the case when a Markov chain is reversible and in steady state.

Definition: reversible Markov chain

This result leads to the following very important definition that extends to both homogeneous and nonhomogeneous Markov chains.

Definition: Reversible Markov Chain

A Markov chain with transition probabilities $P\{X_n = j | X_{n-1} = i\}$ is said to be reversible under the stationary distribution π_i if the following detailed-balance equations hold:

$$\sum_{i \in \Omega} \pi_i = 1$$

$$\forall i, j \in \Omega, \quad P\{X_n = j | X_{n-1} = i\} \pi_i = P\{X_n = i | X_{n-1} = j\} \pi_j$$

Why detailed balance is worth having

It is not always possible to find a probability density π_i which is a solution to the detailed-balance equations because not all Markov chains are reversible.

The key fact

But when you can find a solution, then that solution must also be a solution to the full-balance equations, which means that this is the stationary distribution of an ergodic Markov chain.

To see why this is true, we simply sum the detailed-balance equations over the index i :

$$\begin{aligned}\sum_{i \in \Omega} P\{X_n = j | X_{n-1} = i\} \pi_i &= \sum_{i \in \Omega} P\{X_n = i | X_{n-1} = j\} \pi_j \\ &= \pi_j \sum_{i \in \Omega} P\{X_n = i | X_{n-1} = j\} \\ &= \pi_j .\end{aligned}$$

Example 13.3: the birth-death process

A **birth-death process** is a classic example of a reversible Markov chain. In a birth-death process the state of the Markov chain, X_n , takes on values in the set $\{0, 1, 2, \dots\}$.

With each new time increment, the value of X_n is either incremented (i.e., birth occurs) or decremented (i.e., death occurs), or the state remains unchanged. Mathematically, this can be expressed with the following transition probabilities:

$$P_{i,j} = \begin{cases} \lambda & \text{if } j = i + 1, \\ \mu & \text{if } j = i - 1 \text{ and } i > 0, \\ 1 - \lambda - \mu & \text{if } j = i \text{ and } i > 0, \\ 1 - \lambda & \text{if } j = i \text{ and } i = 0, \\ 0 & \text{otherwise} \end{cases}.$$

Here $0 < \lambda < 1$ is the probability of a birth and $0 < \mu < 1$ is the probability of a death, where we also assume that $\lambda + \mu < 1$.

Example 13.3: solving detailed balance

In this case, the detailed-balance equations have a simple solution. Let π_i be the steady-state probability that the Markov chain is in state i . Then the detailed-balance equations require that

$$\pi_i \lambda = \pi_{i+1} \mu$$

for all $i \geq 0$.

This relation implies that the recursion $\pi_{i+1} = \frac{\lambda}{\mu} \pi_i$ must hold, so we know that the solution must have the form

$$\pi_i = \frac{1}{z} \left(\frac{\lambda}{\mu} \right)^i,$$

where z is a normalizing constant.

Example 13.3: the stationary distribution

When $\frac{\lambda}{\mu} < 1$, then the value of z may be calculated as

$$z = \sum_{i=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^i = \frac{1}{1 - \frac{\lambda}{\mu}} .$$

So the stationary distribution is given by

$$\pi_i = \left(\frac{\lambda}{\mu}\right)^i \left(1 - \frac{\lambda}{\mu}\right) .$$

- Since this distribution solves the detailed-balance equations, it must also solve the full-balance equations.
- This implies that it is a stationary distribution of the Markov chain and that the Markov chain is ergodic and reversible.

Example 13.3: when there is no steady state

However, when $\frac{\lambda}{\mu} \geq 1$, then the only solution to the detailed-balance equations has the form $\pi_i = 0$.

- In this case, the Markov chain is not ergodic, which is to say that it does not reach a steady state.

Intuition

When $\lambda \geq \mu$, the rate of births is greater than the rate of deaths, and the state grows toward infinity, so it cannot reach a steady state.

Where this is used

Detailed balance is the tool we will use to build stochastic simulation algorithms.

- In Chapter 15, samplers are designed so that their stationary distribution is a desired posterior distribution. The design works by choosing transition probabilities that satisfy the detailed-balance equations for that posterior.

Summary

- Conditioning on the past: $P\{X_n = x_n | X_r, r \leq k\} = P\{X_n = x_n | X_k\}$ for $k < n$.
- Subsampling: if $Y_n = X_{dn}$ then Y_n is a Markov chain, and it is homogeneous when X_n is homogeneous.
- Time reversal: if $Y_n = X_{N-n}$ then Y_n is a Markov chain. In steady state its transition probabilities are $Q_{j,i} = P\{X_{n-1} = i | X_n = j\}$, and $\pi_i P_{i,j} = \pi_j Q_{j,i}$.
- Reversible means $P_{i,j} = Q_{j,i}$, which gives the detailed-balance equations $\pi_i P_{i,j} = \pi_j P_{j,i}$: the rate of transitions from i to j equals the rate from j to i .
- Summing detailed balance over i gives full balance, so a detailed-balance solution is the stationary distribution of an ergodic Markov chain.
- Birth-death process: $\pi_i \lambda = \pi_{i+1} \mu$ gives $\pi_i = (\lambda/\mu)^i (1 - \lambda/\mu)$ when $\lambda/\mu < 1$, and no steady state when $\lambda/\mu \geq 1$.

Reading: Chapter 13.5. **Next lecture:** Chapter 14 highlights, Sections 14.1–14.4.