8.1 Review

1. Stationary distributions

2. The basic limit theorem for Markov chains

8.2 Outline

1. Finite ergodic Markov chains
2. Some real-world examples
3. Finite absorbing chains
8.3 Finite ergodic Markov chains

We saw in the last lecture that Markov chains which are first-order, time homogenous, irreducible and aperiodic often have stationary distributions which are unique and equal to their limiting distributions. Such Markov chains are called **Ergodic** Markov chains.

When they further have a finite number of states we call them finite Ergodic Markov chains. These Markov chains always have unique stationary distributions, and will be the focus of the rest of the lecture.

**Example 1:** For a finite ergodic Markov chain the stationary distribution is given by the expression

\[ \pi = e(I + E - P)^{-1} \]

where \( e \) is a row vector of ones, \( E \) is a matrix of all ones, \( I \) is the identity matrix.
We know that the stationary distribution satisfies $\pi P = \pi$. As a result it is often referred to as an \textbf{equilibrium} distribution. If the distribution of $X_n$ is $\pi$, then for any $m \geq n$, $X_m$ will also have the distribution $\pi$.

The multi-chain interpretation: For a finite ergodic chain, one can imagine running many chains in parallel from different initial distributions over $X_0$. In this case we will have that:

$$\lim_{n \to \infty} \mathbb{E}[\text{proportion of chains in state } i \text{ at time } n] = \pi_i$$

\textbf{Mean first-passage times:} Let $n_{ij}$ be the expected number of steps to reach state $j$ (for the first time), given that it started in state $i$. In the same vein, $n_{ii}$ is the expected number of steps to return to state $i$ given that you start in state $i$. These are sometimes referred to as mean first-passage and mean first-return times.

\textbf{Example 2:} Show that

$$n_{ij} = 1 + \sum_{k \neq j} P_{ik} n_{kj}.$$
This result can be written in matrix form:

\[ \mathbf{N} = \]

**Example 3:** Show that \( \pi_i = \frac{1}{\hat{n}_{ii}} \). Discuss the practical implications of this result. Is the result intuitive? Why or why not?
8.4 Some real-world examples

The point of this portion of the lecture is just to see some Markov chains in action. It is more to build some intuition about how to set things up as a Markov chain and think about problems in that way but not necessarily to see all the math in action.

Example 4: This is an example from Minh’s book on “Applied probability models” which I recommend, especially for examples of this kind. Land in an urban area can be used for a variety of purposes. The land uses can be divided into 10 categories. (1) low-density residential, (2) high-density residential, (3) office, (4) general commercial, (5) automobile commercial, (6) parking, (7) warehousing, (8) industry, (9) transportation, (10) vacant land.

Using historical data, one can estimate the conversion probabilities. Particularly, using data on the conversions in Toronto between 1952 and 1962, we can arrive at the following transition matrix.

\[
P = \begin{bmatrix}
0.13 & 0.34 & 0.10 & 0.04 & 0.04 & 0.22 & 0.03 & 0.02 & 0 & 0.08 \\
0.02 & 0.41 & 0.05 & 0.04 & 0 & 0.04 & 0 & 0 & 0 & 0.44 \\
0 & 0.07 & 0.43 & 0.05 & 0.01 & 0.28 & 0.14 & 0 & 0 & 0.02 \\
0.02 & 0.01 & 0.09 & 0.30 & 0.09 & 0.27 & 0.05 & 0.08 & 0.01 & 0.08 \\
0 & 0 & 0.11 & 0.07 & 0.70 & 0.06 & 0 & 0.01 & 0 & 0.05 \\
0.08 & 0.05 & 0.14 & 0.08 & 0.12 & 0.39 & 0.04 & 0 & 0.01 & 0.09 \\
0.01 & 0.03 & 0.02 & 0.12 & 0.03 & 0.11 & 0.38 & 0.21 & 0.01 & 0.08 \\
0.01 & 0.02 & 0.02 & 0.03 & 0.03 & 0.08 & 0.18 & 0.61 & 0 & 0.02 \\
0.01 & 0.18 & 0.14 & 0.04 & 0.10 & 0.39 & 0.03 & 0.03 & 0.08 & 0 \\
0.25 & 0.08 & 0.03 & 0.03 & 0.05 & 0.15 & 0.22 & 0.13 & 0 & 0.06 \\
\end{bmatrix}
\]

Find the limiting distribution and discuss briefly policy implications.

Using MATLAB, we can find:

\[
\pi = \begin{bmatrix}
0.0502 & 0.0839 & 0.1186 & 0.0816 & 0.1507 & 0.1899 & 0.1142 & 0.1154 & 0.0042 & 0.0914 \\
\end{bmatrix}
\]
Example 5: This is another example from Minh’s book. The expected first reaching times of Markov chains were used to study social mobility, especially as reflected in father-child occupation transitions. Beshers and Laumann (1967) (the paper is in your folder) defined the occupational categories as: (1) top professional, top business; (2) semiprofessional, middle business; (3) clerical, sales; (4) skilled; and (5) semiskilled and unskilled. Analyzing some classical social mobility data for the British, they obtained the following expected first reaching time matrix:

\[
N = \begin{bmatrix}
44.00 & 26.39 & 4.57 & 3.89 & 6.01 \\
63.26 & 24.18 & 4.24 & 3.47 & 5.68 \\
71.34 & 31.61 & 4.66 & 2.69 & 4.84 \\
73.57 & 33.45 & 5.65 & 2.44 & 4.08 \\
74.76 & 34.24 & 6.20 & 2.67 & 3.21
\end{bmatrix}
\]

Show that the transition matrix \( P \) for the above example is approximately

\[
P = \begin{bmatrix}
0.3880 & 0.1460 & 0.2635 & 0.1406 & 0.0619 \\
0.1070 & 0.2671 & 0.3466 & 0.2065 & 0.0729 \\
0.0270 & 0.0639 & 0.3468 & 0.4011 & 0.1612 \\
0.0089 & 0.0240 & 0.1980 & 0.4734 & 0.2956 \\
0.0001 & 0.0110 & 0.1243 & 0.3789 & 0.4859
\end{bmatrix}
\]

As a side note: The authors use many small chains (i.e. father-son chains) to calculate \( N \). This may or may not be reasonable in this context. Possibly in future lectures we will try to reason more critically about the design of studies of this form. For this lecture: how did I calculate \( P \)?
8.5 Finite absorbing chains

A (first-order, time homogenous) Markov chain with a finite number of states such that each state is either absorbing or transient, is called a finite absorbing chain.

Often, absorbing states are end points in the process being modeled, but it is convenient to think of the chain being forever “stuck” in these states.

Suppose we permute the states so that the absorbing states come first. Then the transition matrix for a finite absorbing chain can be written as

\[
P = \begin{bmatrix} I & 0 \\ S & T \end{bmatrix}.
\]

Why is this the case?

For a Markov chain we can define the expected number of visits to state \( j \) given the chain starts in state \( i \). We have seen this quantity before and showed previously that

\[
E[\text{number of visits to state } j | X_0 = i] = \sum_{k=0}^{\infty} P_{ij}^k.
\]

Let us denote this quantity \( v_{ij} \), and put the \( v_{ij} \)s into a matrix \( V \).
For a finite absorbing chain what can we say about the structure of $V$? Particularly, notice that we can write

$$V = \begin{bmatrix} A & 0 \\ B & U \end{bmatrix}.$$ 

What can one say about the form of $A$, $B$ and $U$?

The matrix $U$ is known as the **fundamental matrix**, of the finite absorbing chain.

**Example 6:** Let $\mathcal{T}$ denote the collection of all the transient states of a finite absorbing chain. Show that for any pair $i, j \in \mathcal{T}$, we have that

$$v_{ij} = \mathbb{1}(i = j) + \sum_{k \in \mathcal{T}} v_{kj} P_{ik}.$$
Example 7: Use the result of the previous example, to express the fundamental matrix via a matrix equation.

Example 8: Recall the rat maze example.

Notice that this is a finite absorbing chain. Calculate its fundamental matrix.
Example 9: How can we calculate the expected number of steps until absorption when working with a finite absorbing chain?