Markov Chains: lecture 2.

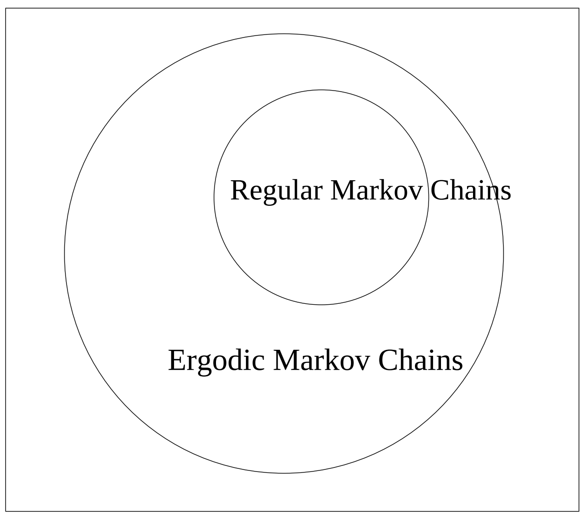
Ergodic Markov Chains

Defn: A Markov chain is called an ergodic or irreducible Markov chain if it is possible to eventually get from every state to every other state with positive probability.

Ex: The wandering mathematician in previous example is an ergodic Markov chain.

Ex: Consider 8 coffee shops divided into four groups. Suppose that each group of four has same geometry as wandering mathematician. Then this is not an ergodic Markov chain. Note, could have initial distribution with "weight" on all shops.

Defn: A Markov chain with finite state space is regular if some power of its transition matrix has only positive entries.



Remark: The above picture shows how the two classes of Markov chains are related. If P^n has all positive entries then

$$P(\text{going from } x \text{ to } y \text{ in } n \text{ steps}) > 0,$$

so a regular chain is ergodic.

To see that regular chains are a strict subclass of the ergodic chains, consider a walker going between two shops:

$$1 \leftrightarrows 2.$$

The transition matrix is given by

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then,

$$P^2 = P^4 = P^6 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad P = P^3 = P^5 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Hence P is ergodic, but not regular.

Fundamental Limit Theorem: Let P denote the transition matrix for a regular Markov chain with finite state space. To simplify notation assume $S = \{1, 2, ..., r\}$. Then

$$\lim_{n\to\infty} P^n = W,$$

where W is an r by r matrix, all rows of which are the same strictly positive probability vector

$$w = [p(1), \ldots, p(r)].$$

That is, $p(x) > 0$ for $x \in \{1, \ldots, r\}$ and $\sum_{x=1}^{r} p(x) = 1$. In particular, for all y we have

$$\lim_{n\to\infty} P^n(x, y) = p(y)$$

independent of x. We actually have the estimate

$$\max\{|P^n(x, y)) - p(y)| \ : \ x, y \in S\} \le Ce^{-Dn},$$

where C,D are positive finite constants independent of n.

Example: Consider the "coffee shop walker" from first lecture. The transition matrix was

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 0 \end{bmatrix}.$$

We have that

$$P^2 = \begin{bmatrix} 1/3 & 0 & 1/3 & 1/3 \\ 0 & 2/3 & 1/6 & 1/6 \\ 1/6 & 1/4 & \frac{5}{12} & 1/6 \\ 1/6 & 1/4 & 1/6 & \frac{5}{12} \end{bmatrix}, \quad P^3 = \begin{bmatrix} 0 & 2/3 & 1/6 & 1/6 \\ 2/9 & 1/6 & \frac{11}{36} & \frac{11}{36} \\ 1/12 & \frac{11}{24} & 1/6 & \frac{7}{24} \\ 1/12 & \frac{11}{24} & \frac{7}{24} & 1/6 \end{bmatrix},$$

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 1/8 & 3/8 & 1/4 & 1/4 \\ 1/8 & 3/8 & 1/4 & 1/4 \\ 1/8 & 3/8 & 1/4 & 1/4 \\ 1/8 & 3/8 & 1/4 & 1/4 \end{bmatrix}.$$

HOW DO WE FIND THESE LIMITS WHEN THEY EXIST?

Once we know $\lim_{n\to\infty} P^n = W$, where W has all rows the same, we can use this fact to actually compute W and provide an interpretation for the common row vector of W. This is taken up in the next theorem.

Theorem: Let P be a regular transition matrix for a finite state space Markov chain with state space $S = \{1, 2, 3, ..., r\}$ and assume that $\lim_{n\to\infty} P^n = W$, with common row w. Then the r-by-r system of linear equations given by $xP = x$ has a unique probability row vector solution, and this solution is the common row w. Furthermore, if v is an arbitrary probability row vector of length r, then

$$\lim_{n\to\infty} vP^n = w,$$

where w is the common row vector of W. Hence the long run probability of being in state y, namely

$$\sum_{x=1}^{r} v_x P^n(x,y),$$

is approximately w_y for all $y = 1, 2, ..., r$ no matter what initial probability distribution $v = [v_1, ..., v_r]$ we use. In addition, if w is the common row vector of W, then we also have

$$wP^n = w,$$

for any $n \geq 0$. Since the probability of being in x is w_x for all times $n = 0, 1, 2, ...$, the chain is in equilibrium if we start with initial distribution w.

Important Fact: Since w is the unique probability vector satisfying $x = xP$ where P is regular and finite, we can use this to solve for w, and hence compute

$$\lim_{n\to\infty} P^n(u,v)$$

for all $u, v = 1, 2, ..., r$, since this limit equals $w_v = v^{th}$ term in w. That is, $x = xP$ is an r by r system of linear equations in the variables $x_1, ..., x_r$ and the solution is the probability vector w.

Example: Consider the Markov chain with transition matrix

$$P = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$$

and state space $S = \{1, 2, 3\}$. Then

$$P^2 = \begin{bmatrix} 7/16 & 3/16 & 3/8 \\ 3/8 & 1/4 & 3/8 \\ 3/8 & 3/16 & 7/16 \end{bmatrix}$$

so P is regular and $\lim_{n\to\infty} P^n = W$ exists. To find w we solve $xP = x$ subject to x being a probability vector, i.e. $x_1 + x_2 + x_3 = 1$ and all $x_i \geq 0$. Then

$$\begin{aligned} x_1/2 + x_2/2 + x_3/4 &= x_1 \\ x_1/4 + 0 \times x_2 + x_3/4 &= x_2 \\ x_1/4 + x_2/2 + x_3/2 &= x_3. \end{aligned}$$

Solving this 3 by 3 linear system we get $w = x = [2/5, 1/5, 2/5]$ and hence $\lim_{n\to\infty} P^n(x,y) = 2/5$ for $y = 1$ and 3, and $\lim_{n\to\infty} P^n(x,y) = 1/5$ for $y = 2$.

Remark: For nonnegative ergodic chains, the fundamental limit theorem may fail, as can be seen when

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then the chain is ergodic, but

$$P^n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ for } n \text{ even and } P^n = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ for } n \text{ odd.}$$

However, for ergodic Markov chains with finite state space there is a unique stationary probability vector w such that $wP = w$. That is, we have the following theorem.

Theorem 1. *Let P be the transition matrix for an ergodic Markov chain. Then there is a unique probability vector w such that $w = wP$.*

Hence, using w as the initial distribution of the chain, the chain has the same distribution for all times since $w = wP^n$ for any $n \geq 1$. For the example we've been using of a chain that is ergodic but not regular, $w = [1/2, 1/2]$.

For a regular Markov chain, the initial distribution w which satisfies

$$wP^n = w$$

can be interpreted as the long run probability vector for being in the various states, i.e.

$$\lim_{n\to\infty} p^n(i,j) = w_j \text{ for } j = 1, 2, ..., r$$

when $w = [w_1, \ldots, w_r]$. However, the limits of the individual n step probabilities do not necessarily exist for ergodic chains. However, the following averages do hold:

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{p^k(i,j)}{n+1} = w_j \qquad \text{for } i, j = 1, 2, ..., r,$$

where $w = [w_1, \ldots, w_r]$ is the stationary probability vector for P. This follows by using the following result, which is a weak law of large numbers for Markov chains.

Theorem 2. *Let w be the stationary initial distribution of an ergodic Markov chain. For $m = 0, 1, 2, \ldots$, let $Y_m = 1$ if the mth step is in state j and zero otherwise. Let*

$H_j^n = (Y_0 + \cdots + Y_n)/(n+1) =$ *average # of times in state j in the first $n+1$ steps* .

Then, for every $\epsilon > 0$

$$\lim_{n\to\infty} P(|H_j^n - w_j| > \epsilon) = 0,$$

independent of the starting distribution.

Exercises:

1. Consider a Markov chain transition matrix

$$P = \begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 3/4 & 0 & 1/4 \\ 0 & 1 & 0 \end{bmatrix}.$$

   (a) Show the this is a regular Markov chain.
   (b) If the process is started in state 1, find the probability that it is in state 3 after two steps.
   (c) Find the limiting probability vector w.
   (d) Find $\lim_{n\to\infty} p^n(x,y)$ for all $x, y \in S$. Why do you know these limits exist?

2. Find the fixed stationary probability vector for

$$P = \begin{bmatrix} 3/4 & 1/4 & 0 \\ 0 & 2/3 & 1/3 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}.$$

3. Consider the Markov chain on $S = \{0, 1, 2, 3, 4\}$ which moves a step to the right with probability 1/2 and to the left with probability 1/2 when it starts at $1, 2, 3$. If it is at 0, then assume it moves to 1 with probability 1 and if it as at 4 it moves to 3 with probability 1. Is this chain ergodic? Is it regular?