Math 20, Fall 2017

Edgar Costa

Week 9

Dartmouth College

Absorbing Markov Chains

- A state s_i of a Markov chain is called absorbing if it is impossible to leave it (i.e., $p_{ii} = 1$).

Absorbing Markov Chains

- A state s_i of a Markov chain is called absorbing if it is impossible to leave it (i.e., $p_{ii} = 1$).
- Markov chain is absorbing if it has at least one absorbing state, and if from every state it is possible to go to an absorbing state (not necessarily in one step).

Absorbing Markov Chains

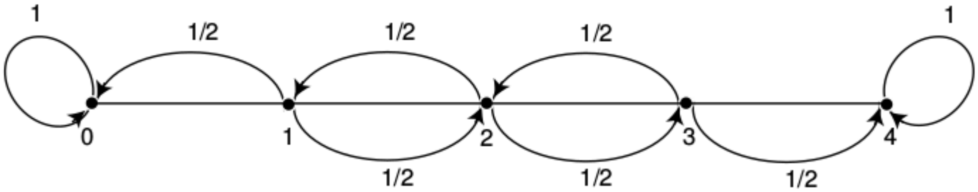
- A state s_i of a Markov chain is called absorbing if it is impossible to leave it (i.e., $p_{ii} = 1$).
- Markov chain is absorbing if it has at least one absorbing state, and if from every state it is possible to go to an absorbing state (not necessarily in one step).
- In an absorbing Markov chain, a state which is not absorbing is called transient.

Example - Drunkard's Walk

A man walks along a four-block stretch of Park Avenue. If he is at corner $1, 2$ or 3, he walks to the left or right with equal probability. He continues like this, until he reaches corner 4, which is a bar, or corner 0, which is his home. If he reaches either home or the bar, he stays there.

Example - Drunkard's Walk

A man walks along a four-block stretch of Park Avenue. If he is at corner $1, 2$ or 3, he walks to the left or right with equal probability. He continues like this, until he reaches corner 4, which is a bar, or corner 0, which is his home. If he reaches either home or the bar, he stays there.



The Transition Matrix - Drunkard's Walk

$$P = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \left(\begin{array}{ccccc} & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{array} \right) \end{array}$$

The Transition Matrix - Drunkard Walk

$$P = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Canonical Form

Canonical Form

- For an absorbing Markov chain, renumber the states so that the transient states come first.

Canonical Form

- For an absorbing Markov chain, renumber the states so that the transient states come first.
- If there are r absorbing states and t transient states, the transition matrix will have the following canonical form

Canonical Form

- For an absorbing Markov chain, renumber the states so that the transient states come first.
- If there are r absorbing states and t transient states, the transition matrix will have the following canonical form

$$P = \begin{array}{cc} & \begin{array}{cc} Trans & Absorb. \end{array} \\ \begin{array}{c} Trans \\ Absorb. \end{array} & \left(\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right) \end{array}$$

Canonical Form

- For an absorbing Markov chain, renumber the states so that the transient states come first.
- If there are r absorbing states and t transient states, the transition matrix will have the following canonical form

$$P = \begin{array}{c c} & \begin{array}{cc} \mathit{Trans} & \mathit{Absorb.} \end{array} \\ \begin{array}{c} \mathit{Trans} \\ \mathit{Absorb.} \end{array} & \left(\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right) \end{array}$$

- The first t states are transient and the last r states are absorbing.

Canonical Form

- For an absorbing Markov chain, renumber the states so that the transient states come first.
- If there are r absorbing states and t transient states, the transition matrix will have the following canonical form

$$P = \begin{array}{c} \\ Trans \\ Absorb. \end{array} \begin{array}{c} \begin{array}{cc} Trans & Absorb. \end{array} \\ \left(\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right) \end{array}$$

- The first t states are transient and the last r states are absorbing.
- I is an $r \times r$ identity matrix, 0 is an $r \times t$ zero matrix, R is a nonzero $t \times r$ matrix, and Q is an $t \times t$ matrix.

Canonical Form

- Recall that the entry $p_{ij}^{(n)}$ of the matrix P^n is the probability of being in the state s_j after n steps, when the chain is started in state s_i.
- where

$$P^n = \begin{array}{cc} & \begin{array}{cc} Trans & Absorb. \end{array} \\ \begin{array}{c} Trans \\ Absorb. \end{array} & \left(\begin{array}{c|c} Q^n & ? \\ \hline 0 & I \end{array}\right) \end{array}$$

Canonical Form

- Recall that the entry $p_{ij}^{(n)}$ of the matrix P^n is the probability of being in the state s_j after n steps, when the chain is started in state s_i.
- where

$$
P^n = \begin{array}{c c} & \begin{array}{cc} \mathit{Trans} & \mathit{Absorb.} \end{array} \\ \begin{array}{c} \mathit{Trans} \\ \mathit{Absorb.} \end{array} & \left(\begin{array}{c|c} Q^n & ? \\ \hline 0 & I \end{array}\right) \end{array}
$$

- What is the probability that the process will be absorbed?

Probability of Absorption

Theorem

In an absorbing Markov chain, the probability that the process will be absorbed is 1.

Probability of Absorption

Theorem

In an absorbing Markov chain, the probability that the process will be absorbed is 1. In other words,

$$\lim_{n \to +\infty} Q^n = 0.$$

Probability of Absorption

Theorem

In an absorbing Markov chain, the probability that the process will be absorbed is 1. In other words,

$$\lim_{n \to +\infty} Q^n = 0.$$

$$P^n = \begin{array}{c c} & \begin{array}{cc} \mathit{Trans} & \mathit{Absorb.} \end{array} \\ \begin{array}{c} \mathit{Trans} \\ \mathit{Absorb.} \end{array} & \left(\begin{array}{c|c} Q^n & ? \\ \hline 0 & I \end{array}\right) \end{array}$$

How many steps until the process gets absorbed?

Write

$$P^n = \begin{array}{c} \\ \mathit{Trans} \\ \mathit{Absorb.} \end{array} \begin{array}{c} \begin{array}{cc} \mathit{Trans} & \mathit{Absorb.} \end{array} \\ \left(\begin{array}{c|c} Q^n & B_n \\ \hline 0 & I \end{array}\right) \end{array},$$

then

$$(I + Q + Q^2 + \cdots + Q^{n-1})R$$

How many steps until the process gets absorbed?

Write

$$P^n = \begin{array}{c c} & \begin{array}{cc} \text{Trans} & \text{Absorb.} \end{array} \\ \begin{array}{c} \text{Trans} \\ \text{Absorb.} \end{array} & \left(\begin{array}{c|c} Q^n & B_n \\ \hline 0 & I \end{array}\right) \end{array},$$

then

$$(I + Q + Q^2 + \cdots + Q^{n-1})R \longrightarrow B =?$$

Equivalently,

$$I + Q + Q^2 + \cdots + Q^{n-1} \longrightarrow ?$$

The Fundamental Matrix

Theorem

For an absorbing Markov chain the matrix $I - Q$ is invertible and

$$(I - Q)^{-1} = N = I + Q + Q^2 + \cdots$$

For an absorbing Markov chain the matrix $N = (I - Q)^{-1}$ is called the **fundamental matrix** for P.

$$P^n = \begin{pmatrix} Q^n & (I + Q + Q^2 + \cdots + Q^{n-1})R \\ 0 & I \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & N \cdot R \\ 0 & I \end{pmatrix} \quad (\text{as } n \to +\infty)$$

Can you interpret the entries of $B = N \cdot R$ and N?

$B_{i,j}$ = probability of being absorbed by the state s_j, given that it started in s_i.

$N_{i,j}$ = number of expected times that the process is in the transient state s_j, given that it started in s_i.

Theorem

Write

$$N = (I - Q)^{-1} = I + Q + Q^2 + \cdots + Q^n + \cdots$$

and

$$B = N \cdot R.$$

- $N_{i,j}$ = number of expected times that the process is in the transient state s_j, given that it started in s_i.
- $B_{i,j}$ = probability of being absorbed by the state s_j, given that it started in s_i.

Theorem

Write

$$N = (I - Q)^{-1} = I + Q + Q^2 + \cdots + Q^n + \cdots$$

and

$$B = N \cdot R.$$

- $N_{i,j}$ = number of expected times that the process is in the transient state s_j, given that it started in s_i.
- $B_{i,j}$ = probability of being absorbed by the state s_j, given that it started in s_i.

Proof idea: How did you compute the expected value of the geometric distribution?

Drunkard's Walk example

$$P = \begin{array}{c c} & \begin{array}{ccccc} 1 & 2 & 3 & 0 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 0 \\ 4 \end{array} & \left(\begin{array}{ccc|cc} 0 & 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \end{array}$$

Identify Q and $I - Q$ and $N = (I - Q)^{-1}$

Drunkard's Walk example

$$N = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{pmatrix} \end{array}$$

and

$$B = NR =$$

Drunkard's Walk example

$$N = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{pmatrix} \end{array}$$

and

$$B = NR = \begin{array}{c c} & \begin{array}{cc} 0 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix} \end{array}$$

Theorem

Let t_i be the expected number of steps before the chain is absorbed, given that the chain starts in state s_i, and let t be the column vector whose ith entry is t_i. Then

$$t = Nc,$$

where c is a column vector all of whose entries are 1.

Summary: Absorbing Markov Chains

- $P = \begin{array}{c c} & \begin{array}{cc} \textit{Trans} & \textit{Absorb.} \end{array} \\ \begin{array}{c} \textit{Trans} \\ \textit{Absorb.} \end{array} & \left(\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right) \end{array}$

- The probability of the process being absorbed is 1.
- $\lim_{n \to +\infty} P^n = \lim_{n \to +\infty} \begin{pmatrix} Q^n & (I + Q + Q^2 + \cdots + Q^{n-1})R \\ 0 & I \end{pmatrix} = \begin{pmatrix} 0 & N \cdot R \\ 0 & I \end{pmatrix}$
- N tells us about the expected number steps in a certain state until absorption or the total time to absorption
- $N \cdot R$ tells us the probability by which state the process will be absorbed

Ergodic Markov Chains

- A Markov chain is called an **ergodic** chain if it is possible to go from every state to every state (not necessarily in one move).

Ergodic Markov Chains

- A Markov chain is called an **ergodic** chain if it is possible to go from every state to every state (not necessarily in one move).
- A Markov chain is called a **regular** chain if some power of the transition matrix has only positive elements.
  In other words, for some n, it is possible to go from any state to any state in exactly n steps.

Ergodic Markov Chains

- A Markov chain is called an **ergodic** chain if it is possible to go from every state to every state (not necessarily in one move).
- A Markov chain is called a **regular** chain if some power of the transition matrix has only positive elements.
  In other words, for some n, it is possible to go from any state to any state in exactly n steps.
- Every regular chain is ergodic. On the other hand, an ergodic chain is not necessarily regular. Example?

Examples

- $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is ergodic but not regular!
- an absorbing chain cannot be regular!
- $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$ is regular

Examples

- $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is ergodic but not regular!
- an absorbing chain cannot be regular!
- $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$ is regular, as $P^2 = \begin{pmatrix} \frac{7}{16} & \frac{3}{16} & \frac{3}{8} \\ \frac{3}{8} & \frac{1}{4} & \frac{3}{8} \\ \frac{3}{8} & \frac{3}{16} & \frac{7}{16} \end{pmatrix}$

Fundamental Limit Theorem for Regular Chains

Theorem

If P is the transition matrix for a regular Markov chain, then $\lim_{n\to+\infty} P^n$ exists. Let

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

then W is a matrix where all rows are the same vector w. The vector w is a strictly positive probability vector (i.e., the components are all positive and they sum to one).

Fundamental Limit Theorem for Regular Chains

Theorem (hard)

If P is the transition matrix for a regular Markov chain, then $\lim_{n \to +\infty} P^n$ **exists**. Let

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

then W is a matrix where all rows are the same vector w. The vector w is a strictly positive probability vector (i.e., the components are all positive and they sum to one).

Theorem (easy)

- $wP = w$;
- any row vector v such that $vP = v$ is a constant multiple of w;
- w is the unique probability vector such that $wP = w$.

Land of Oz

Back to the Land of Oz. Recall, $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$.

Find $\lim_{n\to+\infty} P^n$.

Land of Oz

Back to the Land of Oz. Recall, $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$.

Find $\lim_{n \to +\infty} P^n$.

- Find a vector v such that $v = vP$.

Land of Oz

Back to the Land of Oz. Recall, $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$.

Find $\lim_{n \to +\infty} P^n$.

- Find a vector v such that $v = vP$.
- To make your life easier, assume $v_1 = 1$.

Land of Oz

Back to the Land of Oz. Recall, $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$.

Find $\lim_{n \to +\infty} P^n$.

- Find a vector v such that $v = vP$.
- To make your life easier, assume $v_1 = 1$.
- Rescale, to get the probability vector.

$$w = (2/5, 1/5, 2/5)$$

Equilibrium starting state

We might also reinterpret w as the **equilibrium state** as for all n we have

$$wP^n = w.$$

If we start with a probability distributio given by w, then the probability of being in the various states after n steps is still given by w.

Equilibrium starting state

We might also reinterpret w as the **equilibrium state** as for all n we have

$$wP^n = w.$$

If we start with a probability distributio given by w, then the probability of being in the various states after n steps is still given by w.

Theorem

For a regular Markov chain,

- there is a unique probability vector w such that $wP = w$ and w is strictly positive.
- Any row vector such that $vP = v$ is a multiple of w.

Equilibrium starting state

We might also reinterpret w as the **equilibrium state** as for all n we have

$$wP^n = w.$$

If we start with a probability distributio given by w, then the probability of being in the various states after n steps is still given by w.

Theorem

For ~~a regular~~ an ergodic Markov chain,

- there is a unique probability vector w such that $wP = w$ and w is strictly positive.
- Any row vector such that $vP = v$ is a multiple of w.

Equilibrium starting state

We might also reinterpret w as the **equilibrium state** as for all n we have

$$wP^n = w.$$

If we start with a probability distributio given by w, then the probability of being in the various states after n steps is still given by w.

Theorem

For ~~a regular~~ an ergodic Markov chain,

- there is a unique probability vector w such that $wP = w$ and w is strictly positive.
- Any row vector such that $vP = v$ is a multiple of w.

Proof?

Equilibrium starting state

We might also reinterpret w as the **equilibrium state** as for all n we have

$$wP^n = w.$$

If we start with a probability distributio given by w, then the probability of being in the various states after n steps is still given by w.

Theorem

For ~~a regular~~ an ergodic Markov chain,

- there is a unique probability vector w such that $wP = w$ and w is strictly positive.
- Any row vector such that $vP = v$ is a multiple of w.

Proof? Examples?

Law of Large Numbers for Ergodic Markov Chains

Theorem

Let P be the transition matrix for an ergodic chain. Let

$$A_n = \frac{I + P + P^2 + \cdots + P^{n-1}}{n}.$$

Then

$$\lim_{n \to +\infty} A_n = W,$$

where W is a matrix all of whose rows are equal to the unique fixed probability vector w for P.

Theorem

Let P be the transition matrix for an ergodic chain. Let

$$A_n = \frac{I + P + P^2 + \cdots + P^{n-1}}{n}.$$

Then

$$\lim_{n \to +\infty} A_n = W,$$

where W is a matrix all of whose rows are equal to the unique fixed probability vector w for P.

How to interpret A_n?

Law of Large Numbers for Ergodic Markov Chains

Theorem

Let $H_j^{(n)}$ be the proportion of times in n steps that an ergodic chain is in state s_j. Then for any $\epsilon > 0$,

$$P\left(|H_j^{(n)} - w_j| > \epsilon\right) \to 0 \ ,$$

independent of the starting state s_i.

Law of Large Numbers for Ergodic Markov Chains

Theorem

Let $H_j^{(n)}$ be the proportion of times in n steps that an ergodic chain is in state s_j. Then for any $\epsilon > 0$,

$$P\left(|H_j^{(n)} - w_j| > \epsilon\right) \to 0 \,,$$

independent of the starting state s_i.

"idea": Let $X^{(m)}$ be the random variable that is 1 if the mth step is to state s_j and 0 otherwise, given that we started in state s_i. $E[X^{(m)}] = 1 \cdot p_{ij}^{(m)} + 0 \cdot (1 - p_{ij}^{(m)})$

$$H^{(n)} = \frac{X^{(0)} + X^{(1)} + X^{(2)} + \cdots + X^{(n)}}{n+1}$$

$$E[H^{(n)}] = \frac{1 + p_{ij} + \cdots p_{ij}^{(n)}}{n+1} \longrightarrow w_j$$

Exercise

Consider the Markov chain with general 2×2 transition matrix

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}.$$

1. Under what conditions is P absorbing?
2. Under what conditions is P ergodic but not regular?
3. Under what conditions is P regular?
4. Find the fixed probability vector w for the cases that this makes sense.
5. With $a = b = 1$, show that P^n does not converge to W, but $A_n = \frac{I+P+P^2+\cdots+P^{n-1}}{n}$ does.