Markov Chains

Stochastic Processes - Lecture Notes

Fatih Cavdur

to accompany
Introduction to Probability Models
by Sheldon M. Ross

Fall 2015

Outline

Introduction

- Consider a stochastic process $\{X_n, n = 1, 2, \dots\}$ that takes on a finite number or countable number of possible values.
- Unless otherwise mentioned, the set of possible values are denoted by non-negative integers.
- If $X_n = i$, then, the process is said to be in state i at time n.
- Whenever the process is in state i, there is a fixed probability P_{ij} that it will next be in state j.

Introduction

We can then write P_{ij} as

$$P_{ij} = P\{X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0\}$$

for all states $i_0, i_1, \dots, i_{n-1}, i, j$ and $n \geq 0$. Such a stochastic process is known as a *Markov chain*.

In other words, for a Markov chain, the conditional distribution of any future state X_{n+1} given the present state X_n and the past states $X_{n-1}, X_{n-2}, \dots, X_0$, is independent of the past states and depends only on the present state.

Introduction

The value P_{ij} is the probability that, when in state i, the process will enter state j. We have that

$$P_{ij} \geq 0, \quad i, j \geq 0$$

and

$$\sum_{j=0}^{\infty} P_{ij} = 1, \quad i = 0, 1, \ldots$$

Introduction

We let $\mathbf{P}$ denote the matrix of one-step transition probabilities.

$$\mathbf{P} = \begin{vmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ \vdots & \vdots & \vdots & \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \vdots & \vdots & \vdots & \end{vmatrix}$$

Example 4.1

Suppose that the chance of rain tomorrow only depends on the weather conditions today, and also suppose that if it rains today, it will rain tomorrow with probability α and if it does not rain today, it will rain tomorrow with probability β. The TPM is then given by

$$\mathbf{P} = \begin{vmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{vmatrix}$$

Example 4.3

On any given day Gary is either cheerful (C), so-so (S), or glum (G). If he is cheerful today, then he will be C, S, or G tomorrow with respective probabilities 0.5, 0.4, 0.1. If he is feeling so-so today, then he will be C, S, or G tomorrow with probabilities 0.3, 0.4, 0.3. If he is glum today, then he will be C, S, or G tomorrow with probabilities 0.2, 0.3, 0.5.

Example 4.3

Letting X_n be the mood of Gary on day n, $\{X_n, n \geq 0\}$ is a 3-state MC with TPM

$$\mathbf{P} = \begin{vmatrix} .5 & .4 & .1 \\ .3 & .4 & .3 \\ .2 & .3 & .5 \end{vmatrix}$$

Example 4.4

Transforming a Process into an MC

Suppose that whether or not it rains today depends on previous weather conditions through the last two days. Specifically, suppose that if it has rained for the past two days, then it will rain tomorrow with probability 0.7; if it rained today but not yesterday, then it will rain tomorrow with probability 0.5; if it rained yesterday but not today, then it will rain tomorrow with probability 0.4; if it has not rained in the past two days, then it will rain tomorrow with probability 0.2.

Example 4.4

Transforming a Process into an MC

By defining our states as follows, we can express the problem as an MC:

- state 0: If it rained both today and yesterday
- state 1: If it rained today, but not yesterday
- state 2: If it rained yesterday, but not today
- state 3: If it did not rain today and yesterday

Example 4.4

Transforming a Process into an MC

We have an MC with the following TPM:

$$\mathbf{P} = \begin{vmatrix} .7 & .0 & .3 & .0 \\ .5 & .0 & .5 & .0 \\ .0 & .4 & .0 & .6 \\ .0 & .2 & .0 & .8 \end{vmatrix}$$

Chapman-Kolmogorov Equations

P_{ij} are the 1-step transition probabilities. We define the n-step transition probabilities,

$$P_{ij}^n = P\{X_{n+k} = j | X_k = i\}, \quad i, j, n \geq 0$$

The *Chapman-Kolmogorov equations* provide a method for computing n-step probabilities.

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m, \quad i, j, n, m \geq 0$$

Chapman-Kolmogorov Equations

If we let $\mathbf{P}^{(n)}$ be the n-step TPM of P_{ij}^n, then, we can write

$$\mathbf{P}^{(n+m)} = \mathbf{P}^{(n)} \times \mathbf{P}^{(m)}$$

and thus, we have

$$\mathbf{P}^{(2)} = \mathbf{P}^{(1+1)} = \mathbf{P} \times \mathbf{P} = \mathbf{P}^2$$

and by induction

$$\mathbf{P}^{(n)} = \mathbf{P}^{(n-1+1)} = \mathbf{P}^{(n-1)} \times \mathbf{P} = \mathbf{P}^n$$

Example 4.8

Consider the TPM for Example 4.1, and by letting $\alpha = 0.7$ and $\beta = 0.4$, and compute the probability that it will rain 4 days from today given that it is raining today.

$$\mathbf{P} = \begin{vmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{vmatrix} \Rightarrow \mathbf{P} = \begin{vmatrix} .7 & .3 \\ .4 & .6 \end{vmatrix} \Rightarrow \mathbf{P}^2 = \begin{vmatrix} .61 & .39 \\ .52 & .48 \end{vmatrix}$$

We then have

$$\mathbf{P}^{(4)} = \mathbf{P}^{(2)} \times \mathbf{P}^{(2)} = \begin{vmatrix} .5749 & .4251 \\ .5668 & .4332 \end{vmatrix}$$

Example 4.9

Consider Example 4.4. Given that it rained on Monday and Tuesday, what is the probability that it will rain on Thursday?

$$\mathbf{P} = \begin{vmatrix} .7 & .0 & .3 & .0 \\ .5 & .0 & .5 & .0 \\ .0 & .4 & .0 & .6 \\ .0 & .2 & .0 & .8 \end{vmatrix} \Rightarrow \mathbf{P}^2 = \begin{vmatrix} .49 & .12 & .21 & .18 \\ .35 & .20 & .15 & .30 \\ .20 & .12 & .20 & .48 \\ .10 & .16 & .10 & .64 \end{vmatrix}$$

The desired probability is thus $P^2_{00} + P^2_{01} = 0.61$.

Example 4.10

A pensioner receives 2 (x 1,000 dollars) at the beginning of each month. The amount of money he needs to spend during a month is independent of the amount he has and is equal to i with probability P_i, $i = 1, 2, 3, 4$, and $\sum_i P_i = 1$. If he has more than 3 at the end of a month, he gives the amount greater than 3 to his son. If, after receiving his payment at the beginning of a month, he has a capital of 5, what is the probability that his capital is ever 1 or less at any time within the following 4 monhts?

Example 4.10

We let the states of the MC the amount he has at the end of each month. We then need to consider the states 1, 2 and 3, only. Why? The TPM, let it be $\mathbf{Q}$, is given by

$$\mathbf{Q} = \begin{vmatrix} 1 & 0 & 0 \\ P_3 + P_4 & P_2 & P_1 \\ P_4 & P_3 & P_1 + P_2 \end{vmatrix} \xrightarrow{P_i = .25} \mathbf{Q}^4 = \begin{vmatrix} 1 & 0 & 0 \\ \frac{222}{256} & \frac{13}{256} & \frac{21}{256} \\ \frac{201}{256} & \frac{21}{256} & \frac{34}{256} \end{vmatrix}$$

The desired probability is $Q^4_{3,1} = 201/256$.

Classification of States

- State j is *accessible* from state i if $P_{ij}^{n} > 0$ for some $n \geq 0$, i.e., $i \leftarrow j$.
- 2 states i and j that are accessible to each other are said to *communicate*, i.e., $i \leftrightarrow j$.
- State i communicates with state i, for all $i \geq 0$.
- If state i communicates with state j, then, state j communicates with state i.
- If state i communicates with state j, and state j communicates with state k, then, state i communicates with state k.

Classification of States

- 2 states that communicate with each other are said to be in the same *class*.
- As a consequence of the communication properties, any classes of states are either identical or disjoint.
- In other words, the concept of communication devides the state space up into a number of seperate classes.
- The MC is said to be *irreducible* if there is only 1 class, that is, if all states communicate with each other.

Example 4.11

Find the classes of the following MC.

$$\mathbf{P} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{vmatrix}$$

We see that the MC is irreducible.

Example 4.12

Find the classes of the following MC.

$$\mathbf{P} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

The classes of the MC are $\{0, 1\}$, $\{2\}$ and $\{3\}$.

Classification of States

For any state i, we let f_i be the probability that, starting in state i, the process will ever reenter state i. We say,

- state i is *recurrent* if $f_i = 1$,
- state i is *transient* if $f_i < 1$.

Classification of States

We also say that

- If state i is recurrent, then, starting in state i, the MC will reenter state i infinitely often.
- If state i is transient, then, starting in state i, the number of time periods that the MC will be in state i is has a geometric distribution with mean $1/(1 - f_i)$.
- Hence, if state i is recurrent iif, starting in state i, the expected number of time periods that the MC is in state i is infinite.

Classification of States

For the number of periods that the MC is in state i, we can write

$$I_n = \begin{cases} 1, & X_n = i \\ 0, & X_n \neq i \end{cases} \Rightarrow E\left(\sum_{n=0}^{\infty} I_n | X_0 = i\right) = \sum_{n=0}^{\infty} E(I_n | X_0 = i)$$

$$= \sum_{n=0}^{\infty} P\{X_n = i | X_0 = i\}$$

$$= \sum_{n=0}^{\infty} P_{ii}^n$$

Classification of States

We have thus proven the following:

Proposition

State i is recurrent if

$$\sum_{n=1}^{\infty} P_{ii}^n = \infty$$

State i is trainsient if

$$\sum_{n=1}^{\infty} P_{ii}^n < \infty$$

Classification of States

Corollary

If state i is recurrent, and if state i communicates with state j, then, state j is recurrent.

Example

Find the recurrent and transient states for the following MC:

$$\mathbf{P} = \begin{vmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix}$$

Since all states communicate, and since this is a finite MC, all states must be recurrent.

Example

Find the recurrent and transient states for the following MC:

$$\mathbf{P} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{2} \end{vmatrix}$$

Limiting Probabilities

In Example 4.8, we have

$$\mathbf{P} = \begin{vmatrix} .7 & .3 \\ .4 & .6 \end{vmatrix} \Rightarrow \mathbf{P}^2 = \begin{vmatrix} .61 & .39 \\ .52 & .48 \end{vmatrix} \Rightarrow \mathbf{P}^{(4)} = \begin{vmatrix} .5749 & .4251 \\ .5668 & .4332 \end{vmatrix}$$

If we continue, we obtain the following matrix. What do you note about it?

$$\mathbf{P}^{(8)} = \mathbf{P}^{(2)} \times \mathbf{P}^{(2)} = \begin{vmatrix} .572 & .428 \\ .570 & .430 \end{vmatrix}$$

Limiting Probabilities

- State i is said to have a period of d if $P_{ii}^n = 0$ whenever n is not divisible by d, and d is the largest integer with this property.
  - A state with period 1 is said to be aperiodic.
  - It can be shown that periodicity is a class property.
- If state i is recurrent, then, it is said to be positive recurrent if, starting in i, the expected time until the MC returns to state i is finite.
  - It can be shown that positive recurrence is a class property.
  - While there exist recurrent states that are not positive recurrent, it can be shown that in a finite-state MC all recurrent states are positive recurrent.
- Positive recurrent, aperiodic states are called ergodic.

Limiting Probabilities

Theorem

For an irreducible ergodic MC,

$$\lim_{n\to\infty} P_{ij}^n$$

exists and is independent of i. Furthermore, if we let

$$\pi_j = \lim_{n\to\infty} P_{ij}^n, j \geq n$$

then, π_j is the unique solution of

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, j \geq 0 \text{ and } \sum_{j=0}^{\infty} \pi_j = 1$$

Example

The limiting probabilities of the following MC:

$$\mathbf{P} = \begin{vmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{vmatrix}$$

$$\pi_0 = \alpha\pi_0 + \beta\pi_1$$
$$\pi_1 = (1-\alpha)\pi_0 + (1-\beta)\pi_1$$
$$\pi_0 + \pi_1 = 1$$

$$\pi_0 = \frac{\beta}{1+\beta-\alpha} \text{ and } \pi_1 = \frac{1-\alpha}{1+\beta-\alpha}$$

Example

The limiting probabilities of the following MC:

$$\mathbf{P} = \begin{vmatrix} .5 & .4 & .1 \\ .3 & .4 & .3 \\ .2 & .3 & .5 \end{vmatrix}$$

$$\pi_0 = .5\pi_0 + .3\pi_1 + .2\pi_2$$
$$\pi_1 = .4\pi_0 + .4\pi_1 + .3\pi_2$$
$$\pi_2 = .1\pi_0 + .3\pi_1 + .5\pi_2$$
$$\pi_0 + \pi_1 + \pi_2 = 1$$

$$\pi_0 = \frac{21}{62}, \pi_1 = \frac{23}{62} \text{ and } \pi_2 = \frac{18}{62}$$

Example

Suppose that a production process changes states in accordance with an irreducible, positive recurrent MC with TPs P_{ij}, $i, j = 1, ..., n$, and suppose that certain of the states are considered acceptable and the remaining unacceptable. Let A denote the acceptable states and A^c the unacceptable ones. If the production process is said to be "up" when in an acceptable state and "down" when in an unacceptable state, determine

- the rate at which the production process goes from up to down (that is, the rate of breakdowns);
- the average length of time the process remains down when it goes down;
- the average length of time the process remains up when it goes up.

Example

Let π_k, $k = 1, \ldots, n$ be the long-run proportions. For $i \in A$ and $j \in A^c$, the rate at which the MC enters j from i is

$$\pi_i P_{ij}$$

and so the rate at which the production process j from an unacceptable state is

$$\sum_{i \in A} \pi_i P_{ij}$$

and hence, the rate at which breakdown occurs is

$$\sum_{j \in A^c} \sum_{i \in A} \pi_i P_{ij}$$

Example

If we let $\bar{U}$ and $\bar{D}$ be the time that the MC remains up and down when it goes up and down, respectively. Since there singe breakdown every $\bar{U} + \bar{D}$ time units on the average, it follows that the rate at which breakdown occurs is

$$\frac{1}{\bar{U} + \bar{D}}$$

and then,

$$\frac{1}{\bar{U} + \bar{D}} = \sum_{j \in A^c} \sum_{i \in A} \pi_i P_{ij}$$

Example

Now, consider the proportion of time the MC is up which is

$$\frac{\bar{U}}{\bar{U} + \bar{D}} = \sum_{i \ inA} \pi_i$$

and hence,

$$\bar{U} = \frac{\sum_{i \in A} \pi_i}{\sum_{j \in A^c} \sum_{i \in A} \pi_i P_{ij}}$$

and

$$\bar{D} = \frac{1 - \sum_{i \in A} \pi_i}{\sum_{j \in A^c} \sum_{i \in A} \pi_i P_{ij}} = \frac{\sum_{i \in A^c} \pi_i}{\sum_{j \in A^c} \sum_{i \in A} \pi_i P_{ij}}$$

Example

For example, if we have 1 and 2 are up and 3 and 4 are down states

$$\mathbf{P} = \begin{vmatrix} .25 & .25 & .50 & .00 \\ .00 & .25 & .50 & .25 \\ .25 & .25 & .25 & .25 \\ .25 & .25 & .00 & .50 \end{vmatrix}$$

$$\begin{aligned} \pi_1 &= .25\pi_1 + .25\pi_2 + .50\pi_3 \\ \pi_2 &= .25\pi_1 + .25\pi_2 + .25\pi_3 + .25\pi_4 \\ \pi_3 &= .50\pi_1 + .50\pi_2 + .25\pi_3 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1 \end{aligned}$$

Example

By solving the equations,

$$\pi_1 = \frac{3}{16}, \pi_2 = \frac{1}{4}, \pi_3 = \frac{14}{48} \text{ and } \pi_4 = \frac{13}{48}$$

we have the rate of breakdowns as

$$\pi_1(P_{13} + P_{14}) + \pi_2(P_{23} + P_{24}) = \frac{9}{32}$$

Hence,

$$\bar{U} = \frac{14}{9} \quad \text{and} \quad \bar{D} = 2$$

Some Applications

- The Gambler's Ruin Problem
- A Model for Algorithmic Efficiency
- Using a Random Walk to Analyze a Probabilistic Algorithm

The Gambler's Ruin Problem

Consider a gambler who at each play of the game has probability p of winning one unit and probability $q = 1 - p$ of losing one unit. Assume that each play is independent. What is the probability that, starting with i units, the gambler will reach N units before reaching zero units?

The Gambler's Ruin Problem

Let X_n denote the player's fortune at time n, then, the process $\{X_n, n = 0, 1, \dots\}$ is an MC with transition probabilities

$$P_{00} = P_{NN} = 1$$

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \ \ i = 1, \dots, N-1$$

This MC has 3 classes as $\{0\}$, $\{1, 2, \dots, N-1\}$, and $\{N\}$, the second one is transient whereas the others are recurrent.

The Gambler's Ruin Problem

Let P_i, $i = 0, \dots, N$, denote the probability that, starting with i, the gambler's fortune will eventually reach N. By conditioning on the outcome of the first play, we have

$$\begin{aligned} P_i &= pP_{i+1} + qP_{i-1} \\ pP_i + qP_i &= pP_{i+1} + qP_{i-1} \end{aligned}$$

and thus,

$$P_{i+1} - P_i = \frac{q}{p}(P_i - P_{i-1})$$

for $i = 1, \dots, N-1$.

The Gambler's Ruin Problem

Since $P_0 = 0$, we have that

$$P_{i+1} - P_i = \frac{q}{p}(P_i - P_{i-1}) \Rightarrow P_2 - P_1 = \frac{q}{p}(P_1 - P_0) = \left(\frac{q}{p}\right) P_1$$
$$P_3 - P_2 = \frac{q}{p}(P_2 - P_1) = \left(\frac{q}{p}\right)^2 P_1$$

and thus,

$$P_i - P_{i-1} = \frac{q}{p}(P_{i-1} - P_{i-2}) = \left(\frac{q}{p}\right)^{i-1} P_1$$

and

$$P_N - P_{N-1} = \frac{q}{p}(P_{N-1} - P_{N-2}) = \left(\frac{q}{p}\right)^{N-1} P_1$$

The Gambler's Ruin Problem

Adding the first $i-1$ equations, we get

$$P_i - P_1 = P_1 \left[\left(\frac{q}{p}\right) + \left(\frac{q}{p}\right)^2 + \ldots + \left(\frac{q}{p}\right)^{i-1}\right]$$

or

$$P_i = \begin{cases} \frac{1-(q/p)^i}{1-(q/p)}, & \text{if } \frac{q}{p} \neq 1 \\ iP_i, & \text{if } \frac{q}{p} = 1 \end{cases}$$

Since $P_N = 1$, we get

$$P_1 = \begin{cases} \frac{1-(q/p)}{1-(q/p)^N}, & \text{if } p \neq \frac{1}{2} \\ \frac{1}{N}, & \text{if } p = \frac{1}{2} \end{cases} \Rightarrow P_i = \begin{cases} \frac{1-(q/p)^i}{1-(q/p)^N}, & \text{if } p \neq \frac{1}{2} \\ \frac{i}{N}, & \text{if } p = \frac{1}{2} \end{cases}$$

The Gambler's Ruin Problem

Note that, as $N \to \infty$, we have

$$P_i = \begin{cases} 1 - \left(\frac{q}{p}\right)^i, & \text{if } p > \frac{1}{2} \\ 0, & \text{if } p \leq \frac{1}{2} \end{cases}$$

As a result, we conclude that if $p > .5$, we have a positive probability that the gambler's fortune will increase indefinitely whereas otherwise he will lose all of his money.

The Gambler's Ruin Problem

Example

SupposeMax and Patty decide to flip pennies; the one coming closest to the wall wins. Patty, being the better player, has a probability 0.6 of winning on each flip. (a) If Patty starts with 5 pennies and Max with 10, what is the probability that Patty will wipe Max out? (b) What if Patty starts with 10 and Max with 20?

We have, $i = 5$, $N = 15$ and $p = .6$, the desired probability for part (a) is then given by

$$P_i = \frac{1 - (q/p)^i}{1 - (q/p)^N} = \frac{1 - (2/3)^5}{1 - (2/3)^{15}} \approx 0.87$$

and for part (b),

$$P_i = \frac{1 - (q/p)^i}{1 - (q/p)^N} = \frac{1 - (2/3)^{10}}{1 - (2/3)^{30}} \approx 0.98$$

A Model for Algorithmic Efficiency

Consider the following LP:

$$\min \mathbf{cx} : \mathbf{Ax} \leq \mathbf{b}, \mathbf{x} \geq 0$$

where $\mathbf{c}$ is a $1 \times n$ row vector, $\mathbf{x}$ is a $n \times 1$ column vector, $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{b}$ is a is a $n \times 1$ column vector.
We thus have an LP with n variables and m constraints. For such an LP, assuming that $n > m$, we can find a BFS and the optimal solution $\mathbf{x}$ by setting $n - m$ terms equal to 0 corresponding to the extreme points of the feasible space.

A Model for Algorithmic Efficiency

The famous simplex algorithm solves the LP by moving along the extreme points of the feasible space, and thus, we can have N such extreme points for an LP with n variables and m constraints as

$$N = \binom{n}{m}$$

What do you think about the performance of the simplex algorithm?

A Model for Algorithmic Efficiency

To analyze the performance of the simplex algorithm, consider an MC model to show how the algorithm moves along the extreme points.

- Assume that, if at any time, the algorithm is at the jth best extreme point, then, after the next pivot, the next extreme point will be equally likely to be any of the $j-1$ best points.
- We can then show that the time from the Nth best to the best extreme point has approximately, for large N, a normal distribution with mean and variance equal to the natural logarithm of N.

A Model for Algorithmic Efficiency

Consider an MC for which $P_{11} = 1$ and

$$P_{ij} = \frac{1}{i-1}, \; j = 1, \ldots, i-1, \; i > 1$$

and if we let T_i be the time (number of iterations) to go from state i to state 1, we can write

$$E(T_i) = 1 + \frac{1}{i-1}\left(\sum_{j=1}^{i-1} E(T_j)\right)$$

Starting with $E(T_1) = 0$, we can write by induction that

$$E(T_i) = \sum_{j=1}^{i-1} \frac{1}{j}$$

A Model for Algorithmic Efficiency

To obtain a more complete description, we can write

$$T_N = \sum_{j=1}^{N-1} I_j$$

where

$$I_j = \begin{cases} 1, & \text{if the process ever enter } j \\ 0, & \text{otherwise} \end{cases}$$

A Model for Algorithmic Efficiency

We can then write the followings.

Proposition

$I_1, \ldots, I_{N-1}$ are independent and

$$P_{I_j=1} = \frac{1}{j},\ 1 \leq j \leq N-1$$

Corollary

$$E(T_N) = \sum_{j=1}^{N-1} \frac{1}{j} \quad \text{and} \quad \text{Var}(T_N) = \sum_{j=1}^{N-1} \frac{1}{j}\left(1 - \frac{1}{j}\right)$$

and for N large, T_N has approximately normal with mean and variance $\log N$.

A Model for Algorithmic Efficiency

In a simplex implementation, for n and m large, we have, by Stirling's approximation, that

$$N = \binom{n}{m} \approx \frac{n^{n+1/2}}{(n-m)^{n-m+1/2} m^{m+1/2} \sqrt{2\pi}}$$

and letting $c = n/m$,

$$\log N \approx \left(mc + \frac{1}{2}\right) \log (mc) + \left[m(c-1) + \frac{1}{2} \log [m(c-1)]\right]$$
$$- \left(m + \frac{1}{2}\right) \log m - \frac{1}{2} \log (2\pi)$$

A Model for Algorithmic Efficiency

We then write

$$\log N \approx \left(mc + \frac{1}{2}\right) \log (mc) + \left[m(c-1) + \frac{1}{2} \log [m(c-1)]\right]$$
$$- \left(m + \frac{1}{2}\right) \log m - \frac{1}{2} \log (2\pi)$$
$$\approx m \left[c \log \frac{c}{c-1} + \log (c-1)\right]$$

and when c is large

$$\lim_{x \to \infty} \log N \approx m[1 + \log (c-1)]$$

For instance, for $n = 8,000$ and $m = 1,000$, the number of transitions would be $3,000 \pm 2\sqrt{3000} = 3,000 \pm 110$ approximately, 95% of the time.

Mean Time in Transient States

For a finite-state MC, let $T = \{1, 2, \ldots, t\}$ be the set of transient states, and

$$\mathbf{P}_T = \begin{bmatrix} P_{11} & P_{12} & \ldots & P_{1t} \\ \vdots & \vdots & \ldots & \vdots \\ P_{t1} & P_{t2} & \ldots & P_{tt} \end{bmatrix}$$

For transient states i and j, let s_{ij} denote the expected number of time periods that the MC is in j given that it started in i, and let $\delta_{ij} = 1$ when $i = j$ and let it be 0 otherwise. We then write

$$s_{ij} = \delta_{ij} + \sum_k P_{ik} s_{kj} = \delta_{ij} + \sum_k^t P_{ik} s_{kj}$$

Mean Time in Transient States

If we let

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} & \ldots & s_{1t} \\ \vdots & \vdots & \ldots & \vdots \\ s_{t1} & s_{t2} & \ldots & s_{tt} \end{bmatrix}$$

we can then write

$$(\mathbf{I} - \mathbf{P}_T)\mathbf{S} = \mathbf{I} \Rightarrow \mathbf{S} = (\mathbf{I} - \mathbf{P}_T)^{-1}$$

Example

In the gambler's ruin problem with $p = .4$ and $N = 7$, find the expected amount of time that the gambler has 5 units given that he started with 3 units.

$$\mathbf{P}_T = \begin{bmatrix} .0 & .4 & .0 & .0 & .0 & .0 \\ .6 & .0 & .4 & .0 & .0 & .0 \\ .0 & .6 & .0 & .4 & .0 & .0 \\ .0 & .4 & .6 & .0 & .4 & .0 \\ .0 & .4 & .0 & .6 & .0 & .4 \\ .0 & .4 & .0 & .0 & .6 & .0 \end{bmatrix}$$

Example

In the gambler's ruin problem with $p = .4$ and $N = 7$, find the expected amount of time that the gambler has 5 units given that he started with 3 units.

$$\mathbf{S} = (\mathbf{I} - \mathbf{P}_T)^{-1} = \begin{bmatrix} 1.61 & 1.02 & 0.63 & 0.37 & 0.19 & 0.08 \\ 1.54 & 2.56 & 1.58 & 0.92 & 0.49 & 0.19 \\ 1.42 & 2.37 & 3.00 & 1.75 & 0.92 & 0.37 \\ 1.25 & 2.08 & 2.63 & 3.00 & 1.58 & 0.63 \\ 0.98 & 1.64 & 2.08 & 2.37 & 2.56 & 1.02 \\ 0.59 & 0.98 & 1.25 & 1.42 & 1.54 & 1.61 \end{bmatrix}$$

We then have $s_{35} = 0.92$.

Branching Process

- Consider a population consisting of individuals able to produce offspring of the same kind.
- Assume that each individual will have produced j new offspring with probability P_j, $j \geq 0$, independently of the numbers produced by the others.
- X_n is the size of the nth generation.
- It follows that $\{X_n, n \geq 0\}$ is an MC with its state space as the set of non-negative integers.

Branching Process

- Note that state 0 is a recurrent state.
- Also note that if $P_0 > 0$, all other states are transient.
- We can hence conclude that, if $P_0 > 0$, then, the population will either die out or its size will converge to infinity.

Branching Process

We define the mean and variance of the number of offspring of a single individual as follows:

$$\mu = \sum_{j}^{\infty} jP_j \quad \text{and} \quad \sigma^2 = \sum_{j=0}^{\infty} (j-\mu)^2 P_j$$

We define X_n as above where Z_i represents the number of offspring of the ith individual of the $(n-1)$st generation.

$$X_n = \sum_{i=1}^{X_{n-1}} Z_i$$

Branching Process

By conditioning on the number of individuals on the previous generation, we finally conclude that

$$\pi_0 = \sum_{j=0}^{\infty} \pi_0^j P_j$$

where we can show that when $\mu > 1$, π_0 satisfies the above expression, and when $\mu \geq 1$, $\pi_0 = 1$.

Example

If $P_0 = 1/2$, $P_1 = 1/4$ and $P_2 = 1/4$, what is π_0?
Since $\mu = 3/4$, we have that $\pi_0 = 1$.

If $P_0 = 1/4$, $P_1 = 1/4$ and $P_2 = 1/2$, what is π_0?
We have

$$\pi_0 = \frac{1}{4} + \frac{1}{4}\pi_0 + \frac{1}{2}\pi_0^2 \Rightarrow \pi_0 = \frac{1}{2}$$

Time Reversible MCs

Consider a stationary ergodic MC with TPs P_{ij} and stationary probabilities π_i. Suppose that starting at some time we trace the sequence of states going backward in time. It turns out that this sequence is iteself an MC with TPs

$$\begin{aligned} Q_{ij} &= P\{X_m = j | X_{m+1} = i\} \\ &= \frac{P\{X_m = j, X_{m+1} = i\}}{P\{X_{m+1} = i\}} \\ &= \frac{P\{X_m = j\} P\{X_{m+1} = i | X_m = j\}}{P\{X_{m+1} = i\}} \\ &= \frac{\pi_j P_{ji}}{\pi_i} \end{aligned}$$

Time Reversible MCs

If, for the reversed process which is also an MC with TPS,

$$Q_{ij} = \frac{\pi_j P_{ji}}{\pi_i}$$

we have,

$$Q_{ij} = P_{ij} \Rightarrow \pi_i P_{ij} = \pi_j P_{ji}$$

then, the MC is said to be time reversible.

MC Monte Carlo Methods

Let $\mathbf{X}$ be discrete random vector, and let the PMF of it be given by $P\{\mathbf{X} = x_j\}$, $j \geq 1$, and assume that we want to find, for some function h,

$$\theta = E[h(\mathbf{X})] = \sum_{j=1}^{\infty} h(x_j) P\{\mathbf{X} = x_j\}$$

If it is computationally difficult to evaluate the function, we often use simulation to approximate θ, which is mostly the Monte Carlo simulation.

Markov Decision Process

- Consider a process that is observed at discrete time points to be in any of the M possible states. After observing the state of the process, an action must be chosen where we let A be the set of all possible actions.
- If the process is in state i at time n and action a is chosen, then, the next state of the system is determined according to the transition probabilities $P_{ij}(a)$.
- The TPs are functions of only of the present states and the subsequent actions.

$$P\{X_{n+1} = j | X_0, a_0, X_1, a_1, \ldots, X_n = i, a_n = a\} = P_{ij}(a)$$

Hidden MCs

For an MC with TPs, P_{ij} and initial state probabilities $p_i = P\{X_1 = i\}$, $i \geq 0$, suppose that there is a finite set of signals, and that a signal from the set is emitted each time the MC enters a state. If S_n is the nth signal emitted, then,

$$\begin{aligned} P\{S_1 = s | X_1 = j\} &= p(s|j) \\ P\{S_n = s | X_1, S_1, \ldots, X_n = j\} &= p(s|j) \end{aligned}$$

Such a model is called a hidden MC model.

Thanks. Questions?