REPUBLIC OF CAMEROON
Peace - Work - Fatherland
THE UNIVERSITY OF BAMENDA
P.O. Box 39 Bambili

REPUBLIQUE DU CAMEROUN
Paix - Travail - Patrie
UNIVERSITE DE BAMENDA
B.P. : 39 Bambili

NATIONAL HIGHER POLYTECHNIC INSTITUTE

(NAHPI-SCHOOL OF ENGINEERING)

DEPARTMENT OF COMPUTER ENGINEERING

COME4106 Stochastic Processes and Simulation (3 Credits)

- Discrete and continuous stochastic processes.
- Markov chain models, martingale theory,
- Basic presentation of Brownian motion,
- Diffusion and jump processes.
- The convergence stability analysis of (discrete generation) Markov chains.
- Applications include statistical machine learning, operation research, mathematical biology, computational physics, as well as engineering sciences and financial mathematics.

Objective: By the end of this course students should be able to:

- State the defining properties of various stochastic process models.
- Sample on a computer any type of continuous or discrete time stochastic process.
- Identify appropriate stochastic process model(s) for a given research or applied problem.
- Provide logical and coherent proofs of important theoretic results.
- Apply the theory to model real phenomena and answer some questions in applied sciences.

Introduction

Example 1

Suppose we toss a six-sided die several times, and are interested in the number which appears at the n^{th}toss. Let x_1 denote the number which appears at the first toss, x_2, the number which appears at the second toss, and so on. It can be seen that, we can describe the outcome of this experiment by defining a family of random variables $\{X(n); n \in T\}$, where $T = \{1, 2, \cdots\}$ and $X(n)$ is the number which appears at the *nth* toss of the die. This family or collection of random variables, indexed by the parameter n, is an example of a ***stochastic process***. In this example, n is called the ***index parameter*** of the stochastic process, while T is called the ***index set*** of the stochastic process

Example 2

Consider the number of students admitted by The University of Bamenda each year, since 2010, the year the University was established. Let $X(2010)$ denote the number of students admitted in 1948, X(1949), the number of students admitted in 1949, and so on. It is clear that, we can represent the number of students admitted each year by the University, by a family of random variables $\{X(t); t \in T\}$, where $T = 2010, 2011, 2012, \cdots$. This family of random variables, indexed by the parameter t, is another example of a stochastic process. In this example, T is called the index set of the stochastic process.

Example 3

Consider the number of patients waiting in a hospital to see a doctor. Let $Q(t)$ denote the number of patients waiting at time t. $Q(t)$ can take the values $0, 1, 2, \cdots$. When a patient arrives to see the doctor, the value of $Q(t)$ increases by one, and when a patient departs, the value of $Q(t)$ decreases by one. Thus, as t varies, the value of $Q(t)$ changes. This family of random variables $\{Q(t); t \geq 0\}$, is another example of a stochastic process.

Definition: A **stochastic** or random **process** is a collection of random variables that take values in a set S, the state space. That is a collection of random variables that indexed by some mathematical set, **meaning** that each random variable of the **stochastic process** is uniquely associated with an element in the set.

A **stochastic simulation** is a **simulation** that traces the evolution of variables that can change stochastically (randomly) with certain probabilities. With a **stochastic** model we create a projection which is based on a set of random values.

A stochastic process is a collection of random variables that take values in a set S, the state space. The collection is indexed by another set T, the index set. The two most common index sets are the natural numbers = {0, 1, 2 …}, and the nonnegative real numbers $T = [0, \infty)$, which usually represent discrete time and continuous time, respectively.

A stochastic process means a function that develops itself over time in a partially random way, like, for example, the weather, the price of a share or the amount of waiting patients at a doctor's. We will study some important mathematical models for such functions, using both then probability theory, and computer simulations.

If a process comprise a sequence of experiments in which each experiment has a finite number of possible outcomes with given probabilities, then it is called stochastic or random process.

Suppose a certain business man has five photocopy machines and the number of machines in use is observed every 5 minutes for some period of time. Define x_1 to be the number of use during the first observation, x_2 to be the number of use during the second observation and $x_n, n \geq 1$ is to be the number of machines that are being used when the business man is observed for the n^{th} time. The sequence $\{X_1, X_2 \cdots\}$ is a discrete parameter state stochastic process and $x_n, n \geq 1$ is called the state of the process at n times.

The collection is indexed by another set T, the index set. The two most common index sets are the natural numbers = {0, 1, 2 …}, and the nonnegative real numbers $T = [0, \infty)$, which usually represent discrete time and continuous time, respectively.
The first index set gives a sequence of random variables (rvs) $\{X_0, X_1, X_2 \cdots\}$ and the second, a collection of random variables $\{X(t), t \geq 0\}$, one r.v. for each time t.

Stochastic (Random) processes are used to model random experiments that evolve in time:

(i) Received sequence/waveform at the output of a communication channel

(ii) Packet arrival times at a node in a communication network

(iii) Thermal noise in a resistor

(iv) Scores of an NBA team in consecutive games

(v) Daily price of a stock

(vi) Winnings or losses of a gambler

Earth movement around a fault line.

Stochastic (Random) processes are used to model random experiments that evolve in time:

◦ Received sequence/waveform at the output of a communication channel

◦ Packet arrival times at a node in a communication network

◦ Thermal noise in a resistor

◦ Scores of an NBA team in consecutive games

◦ Daily price of a stock

◦ Winnings or losses of a gambler

◦ Earth movement around a fault line

Discrete vs. Continuous-Time Processes

A) The random process {X(t) : t ∈ T } is said to be discrete-time if the index set T is countably infinite, e.g., {1,2, . . .} or {. . . , −2, −1,0,+1,+2, . . .}:

Example are;

i) The process is simply an infinite sequence of r.v.s $x_1, x_2, \cdots$

ii) An outcome of the process is simply a sequence of numbers

B) The random process {X(t) : t ∈ T} is said to be continuous-time if the index set T is a continuous set, e.g., $(0, \infty)$ or $(-\infty, \infty)$.

Example are;

i) The outcomes are random waveforms or random occurrences in continuous time

• A process $X_1, X_2, \cdots$ is said to be independent and identically distributed (i.i.d.) if it consists of an infinite sequence of independent and identically distributed random variables

Example are;:

A) Bernoulli process: $X_1, X_2, \cdots$ are i.i.d. Bern (p), 0 < p < 1, r.v.s. Model for random phenomena with binary outcomes, such as:

i) Sequence of coin flips

ii) Noise sequence in a binary symmetric channel

iii) The occurrence of random events such as packets (1 corresponding to an event and 0 to a non-event) in discrete-time

iv) Binary expansion of a random number between 0 and 1

B) Discrete-time white Gaussian noise (WGN) process: $X_1, X_2, \cdots$ are i.i.d. N (0, N) r.v.s. Model for:

i) Receiver noise in a communication system

ii) Fluctuations in a stock price

Useful properties of an iid process:

i) Independence: Since the r.v.s in an IID process are independent, any two events defined on sets of random variables with non-overlapping indices are independent

ii) Memorylessness: The independence property implies that the IID process is memoryless in the sense that for any time n, the future $X_{n+1}, X_{n+2}, \cdots$, is independent of the past $X_1, X_2, \cdots, X_n$

SOME EXAMPLES OF STOCHASTIC PROCESS

Example 4

Let *Y* (*t*) be the volume of water in a dam at time *t*. Then $\{Y(t), t \geq 0\}$,, is a stochastic process.

Example 5

Let *X*(*t*) be the average volume of water in a dam in year t. If the dam was constructed in say 1940, then$\{X(t), t \in T\}$, is a stochastic process, where $T = \{1940, 1941, \cdots\}$,.

Specification of stochastic processes

The main elements distinguishing between stochastic processes are the nature of its state space, its index set *T*, and the dependence relations among the random variables X (t).

(a) States

The values which *X*(*t*) can take, are called the states of *X* (*t*). In Example 1, the states of the stochastic process $X(n)$ are 1, 2, 3, 4, 5, and 6, and in Example 3, the states of the stochastic process $\{Q(t): \geq 0\}$ are $0, 1, 2, \cdots$. Changes in the values of the stochastic process $X(t)$ are called transitions between the states of $\{X(t)\}$. If $X(t) = 0$, then the process is said to be in state i at time t.

(b) The state space, S

Definition 2

The set of possible values of *X* (*t*) is called the state space of $\{X(t)\}$. It is denoted by *S*.

In Example 2.1,

$S = \{1, 2, \cdots, 6\}$ and in Example 3, $S = \{0, 1, 2, \cdots, 6\}$. If the state space of $\{X(t)\}$ is discrete, then $\{X(t)\}$ is called a discrete state stochastic process. If the state space of X(t) is continuous, then *X* (*t*) is called a continuous state stochastic

(c) The index set, *T*

If T is discrete, then $\{\mathbf{X(t)}; \mathbf{t} \in \mathbf{T}\}$ is called a discrete time stochastic process and in this case, it is customary to use n instead of t, and write X(n) or Xn instead of X(t). If T is continuous, then

$\{\mathbf{X(t)}\}$ is called a continuous time stochastic process. Examples .1 and 2 are discrete time stochastic processes while Examples 3 and 4 are continuous time stochastic processes. Example 5 is a continuous state and a discrete time stochastic process. Every stochastic process can be specified in terms of its state space S, and its index set T. The following four examples are given so that the reader gets a better understanding of the concept of a stochastic process, its state space and its index set.

Example 6

John and Ann play a coin tossing game. John agrees to pay Ann 100 Frs whenever the coin falls "heads" and Ann agrees to pay John 100 Frs whenever the coin falls "tails". Let S_n be the amount earned by Ann in n tosses of the coin.

(a) Determine the state space and the index set of the stochastic process $\{S_n\}$.

(b) If the coin is fair, show that $E\{S_n\} = 0$ and find $Var\{S_n\}$.

Solution

(a) Here, $\mathbf{n} = \{\mathbf{1}, \mathbf{2}, \mathbf{3}, \cdots\}$ and $\mathbf{S_n} = \{\mathbf{0}, \pm\mathbf{1}, \pm\mathbf{2}, \cdots \pm \mathbf{n}\}$. Thus, $\{\mathbf{S_n}\}$ is a discrete state stochastic process indexed by a discrete set.

$$X(t) = \begin{cases} 1 \; with\, probability \; \frac{1}{2} \\ -1 \; with\, probability \; \frac{1}{2} \end{cases}$$

Thus the $E\big(X(i)\big) = 1 \times \frac{1}{2} + -1 \times \frac{1}{2} = \frac{1}{2} - \frac{1}{2} = 0$

$E[\{X(i)\}^2] = (1)^2 \times \frac{1}{2} + (-1)^2 \times \frac{1}{2} = \frac{1}{2} + \frac{1}{2} = 1$

And so $Var\big(X(i)\big) = 1$. It follows that

$E\{S_n\} = \sum_{i=0}^{n} E\big(X(i)\big) = 0$ and $Var\big(X(i)\big) = \sum_{i=1}^{n} Var(X(i))$, by independence $= n$.

Example 7

Five green balls and 3 white balls are placed in two boxes A and B so that each box contains 4 balls. At each stage, a ball is drawn at random from each box and the two balls are interchanged. Let $X(n)$ denote the number of white balls in box A after the n^{th} draw. The state space and the index set of the stochastic process $\{X(n)\}$ are, respectively, $S = \{0, 1, 2, 3\}$ and

$T = \{1, 2, 3, \cdots,\}$

If Y_n is the total number of green balls in box A after the n^{th}draw, then the stochastic $\{Y_n\}$ process has the state space $S = \{0, 1, 2, 3\}$ and index set $T = \{1, 2, 3, \cdots,\}$. Both $\{X(n)\}$ and $\{Y_n\}$ are discrete state and discrete time stochastic processes

Example 8

Consider the number of telephone calls received at a telephone exchange at time This can be represented by the stochastic process $\{X(t): t \geq 0\}$, where $X(n)$ is the number of t. telephone calls received at time This is a discrete state and a continuous time stochastic process. The index set of each of the stochastic processes we have considered so far, is one dimensional. In the following example, we consider a stochastic process whose index set is two dimensional.

Example 9

Consider waves in oceans. Let $X(t)$ denote the height of the wave at the location t. We may regard the latitude and longitude coordinates of the wave as the value of t. We then have a stochastic process whose index set is not one dimensional.

Probability Vector and Stochastic Matrix

Probability Vector

A probability vector is a vector (i.e., a matrix with single column or row) where all the entries are non-negative and add up to exactly one. That is a row vector $P = (p_1, p_2, \cdots, p_n)$ is called a probability vector if $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$. It's sometime also called a stochastic vector.

Example The vector $\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$ represents the probability distribution of a random roll of a fait dice. Each entry represents your odds of rolling one face of the dice (i.e. you have a $\frac{1}{6}$ chance of rolling 1, 2, 3, 4, 5, 6).

Every n-dimensional probability vector has a mean $\frac{1}{n}$. The length of a probability vector is calculated with the formula:

$\sqrt{n\sigma^2 + \frac{1}{n}}$ where σ is the variance of all entries in the probability vector.

Example. Consider the following row vectors

$A = \left(\frac{2}{3},\ 0,\ -\frac{1}{3},\ \frac{1}{2}\right),\ B = \left(\frac{1}{2},\ \frac{1}{3},\ 0,\ \frac{1}{6}\right)$ and $C = \left(\frac{2}{3},\ \frac{1}{4},\ 0,\ \frac{1}{2}\right)$

Then

A is not a probability vector since its third component is negative

B is a probability vector

C is not a probability vector since the sum of its components is not 1

Stochastic Matrix

A square matrix $P = (p_{ij})$ is a stochastic matrix if each of its rows is a probability vector. If every row is a probability vector and every column is a probability vector, then the matrix is called a doubly stochastic matrix

Example.

Given the following matrices

a) $\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix}$ b) $\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & -\frac{3}{4} & \frac{1}{4} \end{bmatrix}$ c) $\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & 1 \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix}$ d) $\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}$

Identify each of the following matrices

Solution

$a)$ Is a doubly stochastic matrices

$b)$ is not a stochastic matrix since the entry on third row and second column is negative

$c)$ is not a stochastic matrix since the sum of the entries in the second row is not 1

$d)$ is a stochastic matrix

Show that if $A = (a_{ij}), i\,, j = 1, 2, \cdots, n$ is a stochastic matrix and $U = (u_1, u_2, \cdots, u_n)$ is a probability vector, then UA is a probability vector.

Proof

$$UA = (u_1, u_2, \cdots, u_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

$$= u_1a_{11} + \ u_2a_{21} + \cdots + u_na_{1n}, u_1a_{12} + \ u_2a_{22} + \cdots + u_na_{2n}, u_1a_{1n} + \ u_2a_{2n} + \cdots + u_na_{mn}$$

$$= u_1(a_{11} + a_{12} + \cdots + a_{1n}) + u_2(a_{21} + a_{22} + \cdots + a_{2n}) + u_n(a_{1n} + a_{n2} + \cdots + a_{mn})$$

$$= u_1.1 + u_2.1 + \cdots + u_1.1 = u_1 + u_2 + \cdots + u_1 = 1$$

Show that if A and B are stochastic matrices, then the product AB is a stochastic matrix and in general all powers $A^n, n > 2$ are stochastic matrices.

Proof

Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ then

$$AB = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

From the first row

$$a_{11}b_{11} + a_{12}b_{21} + a_{11}b_{12} + a_{12}b_{22} = a_{11}(b_{11} + b_{12}) + a_{12}(b_{21} + b_{22}) = a_{11} + a_{12} = 1$$

From the second row

$$a_{21}b_{11} + a_{22}b_{21} + a_{21}b_{12} + a_{22}b_{22} = a_{21}(b_{11} + b_{12}) + a_{22}(b_{21} + b_{22}) = a_{21} + a_{22} = 1$$

The result of the above product shows that for each row, a probability vector is multiplying a stochastic matrix and the result is a probability vector. Hence, for any power $A^n, n > 2$ are stochastic matrices.

Regular Stochastic Matrix

A stochastic matrix $P = (P_{ij})$ is said to be regular if they exit an n, such that all the entries of some powerP^n are positive.

Example

Determine whether any of this matrix is regular, $A = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

Solution

$\begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$ is a regular matrix

$\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}$ is not regular stochastic matric because of 1 and 0 in row one

Determine the unique probability

1) Recognize that for an $n \times n$ squares matrix, the space is $t = (t_1, t_2, \cdots, t_n)$

2) Solve simultaneously the vector $t = tP$

Example

Consider the regular matrix $T = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}$. Find the unique probability vector.

Solution

$t = tP$ We have

$$(t_1, t_2, t_3) = (t_1, t_2, 1 - t_1 - t_2) \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}$$

$$= \left[\frac{1}{2}t_1 + \frac{1}{2}t_2,\ \frac{1}{4}t_1 + 1 - \pi_1 - \pi_2, \frac{1}{4}t_1 + \frac{1}{2}t_2\right]$$

$$\begin{cases} t_1 = \frac{1}{2}t_1 + \frac{1}{2}t_2 \\ t_2 = \frac{1}{4}t_1 + 1 - t_1 - t_2 \\ 1 - t_1 - t_2 = \frac{1}{4}t_1 + \frac{1}{2}t_2 \end{cases} - \begin{cases} t_1 \frac{1}{2}t_1 + \frac{1}{2}t_2 = 0 \\ t_2 = \frac{3}{4}t_1 + 2t_2 = 1 \text{ or } t_1 = \frac{4}{11}, t_2 = \frac{4}{11} \\ \frac{5}{4}t_1 + \frac{3}{2}t_2 = 1 \end{cases}$$

Thus $t = \left(\frac{4}{11}, \frac{4}{11}, \frac{3}{11}\right)$ and hence $T = \lim_{n\to\infty} P^n = \begin{bmatrix} \frac{4}{11} & \frac{4}{11} & \frac{3}{11} \\ \frac{4}{11} & \frac{4}{11} & \frac{3}{11} \\ \frac{4}{11} & \frac{4}{11} & \frac{3}{11} \end{bmatrix}$

Markov Process

Markov process is one in which the future value is independent of the past values, given the present value. That is a class of stochastic processes in which the probability of the process being in a given state at a particular time is related only to the immediately preceding state of that process. Example, random walk process, Poisson process is a Markov process. Therefore, number of arrivals in $(0, t)$ is a Poisson process and hence a Markov process.

The Markov process may be defined as a stochastic process whose development may be treated as a series of transitions between certain values, or states, which have the property that the probability law of the future development of the process, once it is in a given state, depends only on the state and not on the past.

A real number x is said to be a state of a stochastic process $\{X(t), t \in T\}$ if there exists a time t in T such that the probability $P[x - h < X(t) < x + h]$ is positive for every $h > o$. The set of possible states of a stochastic process is called its state space.

Assume we are given an underlying probability space and a stochastic process $\{X(n), n \geq 0\}$ with each $X(n)$; $\Omega \rightarrow \{1, 2, \cdots, m\}$, where Ω is the sample space and the set $\{l, 2, \ldots, m\}$ is the state space. That is the process $\{X(n), n \geq 0\}$ is Markov independent if

$P[X_n/X_0, X_1, \cdots, X_n] = P[X_n/X_{n-1}]$

We may think of an organism moving in time (discrete interval observations) and at any time the organism is in one of states number1, number 2,... number m.

Thus $P(X(n)) = j)$ is the probability that at time observation n the organism is in state j.

A process that is Markov dependent is called a Markov chain if the state space is discrete and the parameter space is also discrete.

Example

Suppose there are 2 competing products, E and N, in the market, for example, Explorer and Netscape. After one year, some customers will keep using the same product but some will switch to using the other product. The proportion of the customers who keep using the same product or switch to use the other product after one year is summarized in the following table:

	E	N
Keep	$\frac{1}{4}$	$\frac{1}{3}$
Switch	$\frac{3}{4}$	$\frac{2}{3}$

Suppose $\frac{3}{5}$ of the customers use product E and the other $\frac{2}{5}$ of customers use product N in the beginning. Then,

a) What is the distribution after one year?

b) What is the distribution after two years?

c) What is the distribution as the market is said to be stable (i.e., the distribution of the market would be constant forever).

Solution**:**

Let the matrix E N

$$A = \begin{array}{c} \mathrm{E} \\ \mathrm{N} \end{array}\begin{bmatrix} 1/4 & 2/3 \\ 3/4 & 1/3 \end{bmatrix} = \begin{bmatrix} \mathrm{E} \to \mathrm{E} & \mathrm{N} \to \mathrm{E} \\ \mathrm{E} \to \mathrm{N} & \mathrm{N} \to \mathrm{N} \end{bmatrix}$$

The first row contains the proportions of which the customers keep using product E and the ones switch to using product E. The second row contains the proportions of which the customers switch to using product N and the ones keep using product N. Let

$x_0 = \begin{array}{c} \mathrm{E} \\ \mathrm{N} \end{array}\begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}$ be the vector of distribution in the beginning.

$$Ax_0 = \begin{bmatrix} 1/4 & 2/3 \\ 3/4 & 1/3 \end{bmatrix}\begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix} = \begin{array}{c} \mathrm{E} \\ \mathrm{N} \end{array}\begin{bmatrix} [(1/4).(3/5)] + [(2/3).(2/5)] \\ [(3/4).(3/5)] + [(1/3).(2/5)] \end{bmatrix}$$

$$= \begin{bmatrix} [\mathrm{E} \to \mathrm{E}] + [\mathrm{N} \to \mathrm{E}] \\ [\mathrm{E} \to \mathrm{N}] + [\mathrm{N} \to \mathrm{N}] \end{bmatrix} = \begin{array}{c} \mathrm{E} \\ \mathrm{N} \end{array}\begin{bmatrix} 25/60 \\ 35/60 \end{bmatrix}$$

After one year, $\frac{25}{50}$ of customers use product E while $\frac{35}{60}$ of customers use product N.

1 $\boldsymbol{A(Ax_0)} = \begin{bmatrix} \mathbf{1/4} & \mathbf{2/3} \\ \mathbf{3/4} & \mathbf{1/3} \end{bmatrix} \begin{bmatrix} \mathbf{25/60} \\ \mathbf{35/60} \end{bmatrix} = \begin{array}{c} \boldsymbol{E} \\ \boldsymbol{N} \end{array} \begin{bmatrix} [(\mathbf{1/4}).(\mathbf{25/60})] + [(\mathbf{2/3}).(\mathbf{35/60})] \\ [(\mathbf{3/4}).(\mathbf{25/60})] + [(\mathbf{1/3}).(\mathbf{35/60})] \end{bmatrix}$

$$= \begin{bmatrix} [\boldsymbol{E} \to \boldsymbol{E}] + [\boldsymbol{N} \to \boldsymbol{E}] \\ [\boldsymbol{E} \to \boldsymbol{N}] + [\boldsymbol{N} \to \boldsymbol{N}] \end{bmatrix} = \begin{array}{c} \boldsymbol{E} \\ \boldsymbol{N} \end{array} \begin{bmatrix} \mathbf{355/720} \\ \mathbf{365/720} \end{bmatrix}$$

After two year, $\frac{\mathbf{355}}{\mathbf{720}}$ of customers use product E while $\frac{\mathbf{355}}{\mathbf{720}}$ of customers use product N.

2. Suppose $\boldsymbol{s_x} = \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \end{bmatrix}$ is the distribution as the market is stable. Thus,

$$\boldsymbol{As_x} = \begin{bmatrix} \mathbf{1/4} & \mathbf{2/3} \\ \mathbf{3/4} & \mathbf{1/3} \end{bmatrix} \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \end{bmatrix} = \boldsymbol{x_s}$$

$$\Leftrightarrow \ \boldsymbol{Ax_s - x_s = Ax_s - I_2x_s = (A - I_2)x_s}$$

$$= \begin{bmatrix} \mathbf{-3/4} & \mathbf{2/3} \\ \mathbf{3/4} & \mathbf{-2/3} \end{bmatrix} \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \end{bmatrix} = \mathbf{0}$$

where I_2 is a 2×2 identity matrix. The solutions for the homogeneous system are

$$\boldsymbol{x_s} = \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \end{bmatrix} = \begin{bmatrix} \mathbf{8/9} \\ \mathbf{1} \end{bmatrix} \boldsymbol{t}, \boldsymbol{t} \in \boldsymbol{R}.$$

Since $\mathrm{x}_1 + \mathrm{x}_2 = 1 \Rightarrow \mathrm{x}_1 = \frac{8}{17}, \mathrm{x}_2 = \frac{9}{17}$.

Markov Chain

A Markov chain is a sequence of discrete states in time or space with fixed probabilities for the transition from one state to a given state in the next step in the chain. A Markov chain may be a series of transition from one state to another such that the probabilities associated with each transition depend only on the immediate past state of the process and not on how the process reached that state.

A stochastic process that is Markov dependent is said to be possess the Markovian property which is equivalent to the statement that the conditional probability of any future state $(X_n = j)$, is independent of the past states $(X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1})$, and the present state $(X_n = i)$, is independent of the past states. The Markov property asserts that the process is memory-less.

Definition;

If $\forall n, P\left[X_n = a_n / X_{n-1} = a_{n-1}, X_{n-2} = a_{n-2}, \ldots\ldots X_0 = a_0 \right] = P\left[X_n = a_n / X_{n-1} = a_{n-1} \right]$

then the process $\{X_n\}$, *n=0,1,2,….* is called a Markov chain.

The conditional probability $P\left[X_n = a_j / X_{n-1} = a_i \right]$ is called the one step transition probability from state a_i to state a_j at the n^{th} step.

General case:

Suppose there are n competing products, $\boldsymbol{C_1}, \boldsymbol{C_2}, \cdots \boldsymbol{C_n}$ in the market. A general Markov chain model can be employed. The proportion of the customers who keep using the same product or switch to use the other product after one year is summarized in the following table:

	C_1	C_2	$\cdots$	C_n
C_1(after one year)	P_{11}	P_{12}	$\cdots$	P_{1n}
C_2(after one year)	P_{21}	P_{22}	$\cdots$	P_{2n}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
C_n(after one year)	P_{n1}	P_{n2}	$\cdots$	P_{nm}

Also, $\boldsymbol{p_{1j} + p_{2j} + \cdots + p_{nj} = 1, j = 1, 2, \cdots, n}$

Let $\boldsymbol{C_1, C_2, \cdots, C_n}$

$$\boldsymbol{A} = \begin{array}{c} \boldsymbol{C_1} \\ \boldsymbol{C_2} \\ \vdots \\ \boldsymbol{C_n} \end{array} \begin{bmatrix} \boldsymbol{p_{11}} & \boldsymbol{p_{12}} & \cdots & \boldsymbol{p_{1n}} \\ \boldsymbol{p_{21}} & \boldsymbol{p_{22}} & \cdots & \boldsymbol{p_{2n}} \\ \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{p_{n1}} & \boldsymbol{p_{n2}} & \cdots & \boldsymbol{p_{nn}} \end{bmatrix},$$

be the matrix representing the proportions of which the customers keep using the original product and those switch to using the other products. Suppose, in the beginning, the proportions of the customers for using $\boldsymbol{C_1, C_2, \cdots, C_n}$ are $\boldsymbol{\pi_1, \pi_2, \cdots, \pi_n}$ and the vector of distribution is

$$\boldsymbol{x_0} = \begin{array}{c} \boldsymbol{C_1} \\ \boldsymbol{C_2} \\ \vdots \\ \boldsymbol{C_n} \end{array} \begin{bmatrix} \boldsymbol{\pi_1} \\ \boldsymbol{\pi_2} \\ \vdots \\ \boldsymbol{\pi_n} \end{bmatrix}, \boldsymbol{\pi_1 + \pi_2 + \cdots + \pi_n = 1}$$

Then, the vector of distribution after k years is

$A^k x_0 = A.A \cdots Ax_0, k = 1, 2, \cdots k - factors$

Suppose $\mathrm{x_s} = \begin{array}{c} \mathrm{C_1} \\ \mathrm{C_2} \\ \vdots \\ \mathrm{C_n} \end{array} \begin{bmatrix} \mathrm{x_1} \\ \mathrm{x_2} \\ \vdots \\ \mathrm{x_n} \end{bmatrix}$ is the vector of distribution as the market is stable. Then, $\mathrm{x_s}$ can be found by solving the following homogeneous linear system,

$$Ax_s = x_s \Leftrightarrow (A - I_2)x_s = \begin{array}{c} C_1 \\ C_2 \\ \vdots \\ C_n \end{array} \begin{bmatrix} p_{11} - 1 & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} - 1 & \cdots & p_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} - 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0$$

The Transition Matrix

A stochastic process is a mathematical model of a situation in the real world that evolves in time in a probabilistic fashion, i.e. we don't seem to be able to completely predict the future. The situation in the real world is often called a system. For the time being we model the system in discrete time as opposed to continuous time. For example, this might be because we observe the system at discrete time intervals. We let n = 0 be the initial time that we observe the system, n = 1 the next time and so on.

We are interested in calculating probabilities that the system is in various states at various times. Mathematically, we describe the system by a sequence of random variables

$X_0, X_1, ..., X_n, \ldots$ where

X_0 = the state of the system at time n = 0,

\X_1 = the state of the system at time n = 1,

. . .

X_n = the state of the system at time n

At any particular time the system is in one of s possible states that we number 1, 2, …, s. For the most part we shall be concerned with the case where there is a finite number of states, so s is finite. However, there are many important situations where the number of states is countably infinite, so the states are numbered by the positive integers 1, 2, …

Example 1

A tool and die shop has two milling machines. The system consisting of these two machines can be in any of the following states.

1. Working, i.e. both machines in working condition (W or 1)

2. Half-working, i.e. one machine is working and the other is broken (H or 2)

3. Broken, i.e. both machines are broken (B or 3)

A t the start of each day we check the two machines and classify the system state as W, H or B. If we do this for a sequence of four days, we might observe that the system state is W today, W tomorrow, H the next day and B the fourth day. In this case we would have

$X_0 = 1, X_1 = 1, X_2 = 2$ and $X_3 = 3$.

Unfortunately, we can't predict the state of the system in the future.

One of the most basic probabilities that one might want to know is the probability of a sequence of observations $x_0, x_1, \cdots, x_n$. We would denote this probability by

$P(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n)$

In Example 1 we might want to know the probability that the state is W today, W tomorrow, H the next day and B the fourth day.

This would be denoted by $P\{X_0 = 1, X_1 = 1, X_2 = 2,\ X_3 = 3\}$.

For a general stochastic process, one might provide the values of

$P(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n)$

for each sequence $x_0, x_1, \cdots, x_n$. This is a lot of data to provide. Things are a lot easier if we make the following two simplifying assumptions which we state as definitions.

Definition 1

A stochastic process $\{X_n: n = 1, 2, \ldots\}$ is called a Markov process if

(1) $\Pr\{X_{n+1} = x_{n+1} \mid X_0 = x_0,\ X_1 = x_1, ..., X_n = x_n\} = \Pr\{X_{n+1} = x_{n+1} \mid X_n = x_n\}$

for all n and $x_0, x_1, \ldots, x_{n+1}$.

This says the probability of the state being something at one observation depends only on the previous observation. The second assumption is that the quantity appearing on the right of (1) depends only on x_n and x_{n+1} and not on n.

Definition 2

A Markov process $\{X_n: n = 1, 2, \ldots\}$ is called stationary or time-homogeneous if

$\Pr\{X_{n+1} = j \mid X_n = i\} = p_{ij}$

for all n and all i and j where the numbers p_{ij} don't depend on n. A stationary Markov process is also called a Markov chain.

As an example of what (1) implies, suppose in Example 1 we wanted to know the probability that tomorrow the system state is B, i.e. both machines are broken. Consider the following two scenarios. In the first scenario, we are told that today the system state is H, i.e. one machine is working and one is broken. Then the probability the system state tomorrow is B would be $\Pr\{X_2 = 3 \mid X_1 = 2\}$ assuming today is time n = 1 and tomorrow is time n = 2. In the second scenario we are told that yesterday the state was W (both machines were working) and today the state is H (one working and one broken). Then the probability that the state tomorrow is B would be $\Pr\{X_2 = 3 \mid X_0 = 1\ X_1 = 2\}$ assuming yesterday is time n = 0. The Markov assumption would imply that knowing the system was in state W yesterday doesn't add any information to the fact that it is in H today, i.e.

$$\Pr\{X_2 = 3 \mid X_0 = 1, X_1 = 2\} \quad = \quad \Pr\{X_2 = 3 \mid X_1 = 2\}.$$

The probabilities p_{ij} appearing in (2) are called transition probabilities and the matrix

$$P \quad = \quad \begin{pmatrix} p_{11} \cdots p_{1s} \\ \\ \\ p_{s1} \cdots p_{ss} \end{pmatrix}$$

is called the transition matrix. If there is a countably infinite number of states then the transition matrix is countable infinite, i.e.

$$P \quad = \quad \begin{pmatrix} p_{11} \cdots p_{1s} \cdots \\ \\ \\ p_{s1} \cdots p_{ss} \cdots \\ \quad \cdots \qquad \cdots \end{pmatrix}$$

Corresponding to the transition matrix is a transition diagram where the states are represented by circles and the transitions are represented by arrows from one circle to another labeled by the transition probabilities.

Example 1 (continued). In the context of Example 1 suppose the transition matrix is

$$P \quad = \quad \begin{pmatrix} 0.81 & 0.18 & 0.01 \\ 0 & 0.9 & 0.1 \\ 0.6 & 0 & 0.4 \end{pmatrix}$$

This might correspond to the fact that each of the two machines might independently go from working to broken in the course of a day with probability 0.9 for each one separately. Furthermore, a serviceman isn't called until both machines are broken and the probability that he comes the next day is 0.6 and this probability remains constant even if it has been a number of days since you have called him for service.

The Powers of the Transition Matrix

The probabilities of future events in a Markov chain can be computed from the powers P^n of the transition matrix P. One way to compute P^n is to multiply P by itself n times. There is another way to compute P^n using the eigenvalues and eigenvectors of P that sheds light on the behavior of P^n for large n. In fact this method leads to a formula for P^n that shows that P^n approaches a limiting matrix, which we shall denote by P^∞, as n tends to ∞.

Example: The transition probability matrix of a Markov chain $\{X_n\}$, three states 1,2 and 3 is

$P = \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$ and the initial distribution is $p^{(0)} = (0.7, 0.2, 0.1)$ Find (i) $P[X_2 = 3]$

(ii) $P[X_3 = 2, X_2 = 3, X_1 = 3, X_0 = 2]$

$$\textbf{sol: } P^{(2)} = p^2 = \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}\begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$$

$$= \begin{bmatrix} 0.43 & 0.31 & 0.26 \\ 0.24 & 0.42 & 0.34 \\ 0.36 & 0.35 & 0.29 \end{bmatrix}$$

$$P[X_2 = 3] \qquad = \sum_{i=1}^{3} P[X_2 = 3/X_0 = i]P[X_0 = i]$$

$$= P[X_2 = 3/X_0 = 1]P[X_0 = 1] + P[X_2 = 3/X_0 = 2]P[X_0 = 1] + P[X_2 = 3/X_0 = 1]P[X_0 = 3]$$

$$= p_{13}^2 P[X_0 = 1] + p_{23}^2 P[X_0 = 2] + p_{33}^2 P[X_0 = 3]$$

$$= 0.26 \times 0.7 + 0.34 \times 0.2 + 0.29 \times 0.1$$

$$= 0.182 + 0.068 + 0.029$$

$$= 0.279$$

(ii) $P[X_3 = 2, X_2 = 3, X_1 = 3, X_0 = 2]$

$= P[X_3 = 2/X_2 = 3, X_1 = 3, X_0 = 2]P[X_2 = 3, X_1 = 3, X_0 = 2]$

Since $P(A \cap B) = P\left(\frac{A}{B}\right)P(B)$ by definition of conditional probability

$= P[X_3 = 2/X_2 = 3]P[X_2 = 3/X_1 = 3, X_0 = 2]P[X_1 = 3, X_0 = 2]$ (by Markov property)

$= P[X_3 = 2/X_2 = 3]P[X_2 = 3/X_1 = 3]P[X_1 = 3/X_0 = 2]P[X_0 = 2]$

$= p_{32}^{(1)}.p_{33}^{(1)}p_{23}^{(1)}.P[X_0 = 2]$

$= 0.4 \times 0.3 \times 0.2 \times 0.2$

$= 0.0048$

$\therefore P[X_3 = 2, X_2 = 3, X_1 = 3, X_0 = 2] = 0.0048$

Example 2.

In sunny Southern California, if it is sunny one day there is a 90% chance that it will be sunny the next day and a 10% chance that it will be cloudy. On the other hand if it cloudy on a particular day, there is a 60% chance that it will be sunny the next day and a 40% chance that it will be cloudy. We model this by a Markov chain with two states. State 1 corresponds to being sunny on a particular day and State 2 to being cloudy. We let n count the days with n = 0 being today. Also let X_n denote the state of day n, either sunny or cloudy. Assume that the assumptions for a Markov chain are satisfied.

The transition matrix is

$$\boldsymbol{P} = \begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \end{pmatrix} \qquad (1)$$

Then

$$\boldsymbol{P^n} = \begin{pmatrix} \Pr\{\text{sunny on day n} \mid \text{sunny today}\} & \Pr\{\text{cloudy on day n} \mid \text{sunny today}\} \\ \Pr\{\text{sunny on day n} \mid \text{cloudy today}\} & \Pr\{\text{cloudy on day n} \mid \text{cloudy today}\} \end{pmatrix}$$

i.e. P^n contains the probabilities of being sunny or cloudy *n* days from now given that it is sunny or cloudy today.

The formula for P^n that we are interested in involves the eigenvalues and eigenvectors of *P*. Recall that an *eigenvector v* of *P* has the following two properties

$$v \neq 0$$

$$Pv = \lambda v \quad \text{for some number } \lambda$$

λ is called an *eigenvalue* of P. It is not hard to show that an eigenvalue must satisfy the equation

$$\det(P - \lambda I) = 0$$

which is called the *characteristic equation*. For the matrix (1) in Example 1 one has

$$P - \lambda I = \begin{pmatrix} 0.9 - \lambda & 0.1 \\ 0.6 & 0.4 - \lambda \end{pmatrix}$$

In general to form $P - \lambda I$ from P one simply subtracts λ from the diagonal entries. The characteristic equation is

$$0 = \det(P - \lambda I) = \begin{vmatrix} 0.9 - \lambda & 0.1 \\ 0.6 & 0.4 - \lambda \end{vmatrix}$$

$$= (0.9 - \lambda)(0.4 - \lambda) - (0.1)(0.6)$$

$$= \lambda^2 - 1.3\lambda + 0.3 = (\lambda - 1)(\lambda - 0.3)$$

So the eigenvalues are

$\lambda_1 = 1$ and $\lambda_2 = 0.3$

In general, if P is an $n \times n$ matrix, then $\det(P - \lambda I)$ is an n^{th} degree polynomial, the characteristic equation is an n^{th} degree polynomial equation and there are n eigenvalues, although some of them may be repeated. It turns out that one is always an eigenvalue for the transition matrix of a Markov chain and the rest of the eigenvalues have absolute value less than or equal to one. In fact, usually the rest of the eigenvalues have absolute value strictly less than one.

One finds the eigenvectors v corresponding to the eigenvalue λ by solving the equation

$(P - \lambda I)v = 0.$

For the matrix (1) in Example 1 we find the eigenvectors as follows.

For $\lambda_1 = 1$ one has

$$P - \lambda I = \begin{pmatrix} -0.1 & 0.1 \\ 0.6 & -0.6 \end{pmatrix}$$

so an eigenvector $v = \begin{pmatrix} x \\ y \end{pmatrix}$ satisfies

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = (P - \lambda I)v = \begin{pmatrix} -0.1 & 0.1 \\ 0.6 & -0.6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -0.1x + 0.1y \\ 0.6x - 0.6y \end{pmatrix}$$

So $\quad 0 = -0.1x + 0.1y$

$$0 = 0.6x - 0.6y$$

These two equations are equivalent to $y = x$. So an eigenvector for $\lambda_1 = 1$ is any vector

$v = \begin{pmatrix} x \\ y \end{pmatrix}$ with $y = x$. So

$$v = \begin{pmatrix} x \\ x \end{pmatrix} = x\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

So any multiple of the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector for $\lambda_1 = 1$. In general any multiple of an eigenvector is also an eigenvector. For many applications, we only need one eigenvector of a given eigenvalue, so we pick one, say $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. In general, $v_1 = \begin{pmatrix} 1 \\ \dots \\ 1 \end{pmatrix}$ is always an eigenvector for the eigenvalue $\lambda_1 = 1$ for the transition matrix of a Markov chain. This is because $Pv_1 = v_1$ which is a consequence of the fact that the rows of P sum to 1.

For the second eigenvalue $\lambda_2 = 0.3$ one has

$$P - \lambda I = \begin{pmatrix} 0.6 & 0.1 \\ 0.6 & 0.1 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = (P - \lambda I)v = \begin{pmatrix} 0.6 & 0.1 \\ 0.6 & 0.1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0.6x + 0.1y \\ 0.6x + 0.1y \end{pmatrix}$$

$$0 = 0.6x + 0.1y$$

This equation is equivalent to $y = -6x$. So an eigenvector for $\lambda_2 = 0.3$ is any vector $v = \begin{pmatrix} x \\ y \end{pmatrix}$ with $y = -6x$. So

$$v = \begin{pmatrix} x \\ -6x \end{pmatrix} = x\begin{pmatrix} 1 \\ -6 \end{pmatrix}$$

So any multiple of the vector $\begin{pmatrix} 1 \\ -6 \end{pmatrix}$ is an eigenvector for $\lambda_2 = 0.3$.

So we can take $v_2 = \begin{pmatrix} 1 \\ -6 \end{pmatrix}$.

One result of the eigenvalues and eigenvectors of a matrix is a formula called the *diagonalization* of the matrix. Let

T = the matrix whose columns are the eigenvectors of P

$$= \begin{pmatrix} 1 & t_{12} & t_{1n} \\ 1 & t_{22} & t_{2n} \\ \dots & & \\ 1 & t_{n2} & t_{nn} \end{pmatrix} \quad \text{(in general)}$$

$$= \begin{pmatrix} 1 & 1 \\ 1 & -6 \end{pmatrix} \quad \text{(in the example)}$$

D = the matrix with the eigenvalues of P on the diagonal and zeros elsewhere

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ \dots & & \\ 0 & 0 & \lambda_n \end{pmatrix} \quad \text{(in general)}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 0.3 \end{pmatrix} \quad \text{(in the example)}$$

Then it is not hard to show that

$$P = TDT^{-1} \tag{2}$$

This formula is called the diagonalization of P. To see why it is true first consider PT. In general, the columns of PT are P times the columns of T. More precisely the j^{th} column of PT is P times the j^{th} column of T. In symbols we might write this as $(PT)_{\bullet,j} = P(T_{\bullet,j})$ where $T_{\bullet,j}$ denotes the j^{th} column of T and $(PT)_{\bullet,j}$ denotes the j^{th} column of T. However, $T_{\bullet,j}$ = the j^{th} eigenvector of P, so $P(T_{\bullet,j}) = \lambda_j T_{\bullet,j}$ where λ_j is the j^{th} eigenvalue of P. So, in the columns of the product PT consist of the eigenvectors of P multiplied by their eigenvalue. Now consider TD. We have $(TD)_{\bullet,j} = T(D_{\bullet,j})$. However, $D_{\bullet,j}$ is all zeros except for its j^{th} component. So $T(D_{\bullet,j}) = \lambda_j T_{\bullet,j}$. Therefore $PT = TD$. Multiplying on the left by T^{-1} gives the diagonalization formula (2).

The diagonalization formula (2) leads to a corresponding formula for P^n. One has

$$\begin{aligned}
P^n &= PP\cdots P && (P \text{ multiplied by itself } n \text{ times}) \\
&= (TDT^{-1})(TDT^{-1})\cdots(TDT^{-1}) && (TDT^{-1} \text{ multiplied by itself } n \text{ times}) \\
&= TD(T^{-1}T)D(T^{-1}T)\cdots(T^{-1}T)DT^{-1} && (\text{we group the } T^{-1}T \text{ together}) \\
&= TDIDI\cdots IDT^{-1} && (T^{-1}T = I) \\
&= TDD\cdots DT^{-1} && (D \text{ is multiplied by itself } n \text{ times})
\end{aligned}$$

So

$$P^n = TD^nT^{-1} \tag{3}$$

It is very easy to multiply diagonal matrices. One simply multiplies the diagonal elements and the off diagonal elements remain all zeros. In particular, D^n has the eigenvalues raised to the n^{th} power on the diagonal and zeros in the off diagonal, i.e.

D^n = the matrix with the $(\lambda_j)^n$ on the diagonal and zeros elsewhere

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & (\lambda_2)^n & 0 \\ \dots & & \\ 0 & 0 & (\lambda_n)^n \end{pmatrix} \quad \text{(in general)}$$

$$= \begin{pmatrix} (1)^n & 0 \\ 0 & (0.3)^n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & (0.3)^n \end{pmatrix} \quad \text{(in the example)}$$

So to get a formula for P^n we compute and multiply the matrices on the right side of (3). In the example we get

$$P^n = TD^nT^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -6 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (0.3)^n \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -6 \end{pmatrix}^{-1} \tag{4}$$

Recall the following useful formula to find the inverse of a 2×2 matrix.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

We swap the diagonal elements, negate the off diagonal elements and divide by the determinant. If we use this formula in (4) we get

$$P^n = \begin{pmatrix} 1 & 1 \\ 1 & -6 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (0.3)^n \end{pmatrix}\left(\frac{1}{-7}\right)\begin{pmatrix} -6 & -1 \\ -1 & 1 \end{pmatrix} = \left(\frac{1}{7}\right)\begin{pmatrix} 1 & (0.3)^n \\ 1 & (-6)(0.3)^n \end{pmatrix}\begin{pmatrix} 6 & 1 \\ 1 & -1 \end{pmatrix}$$

$$= \left(\frac{1}{7}\right)\begin{pmatrix} 6 + (0.3)^n & 1 - (0.3)^n \\ 6 - 6(0.3)^n & 1 + 6(0.3)^n \end{pmatrix} = \begin{pmatrix} \frac{6}{7} + \frac{1}{7}(0.3)^n & \frac{1}{7} - \frac{1}{7}(0.3)^n \\ \frac{6}{7} - \frac{6}{7}(0.3)^n & \frac{1}{7} + \frac{6}{7}(0.3)^n \end{pmatrix}$$

Therefore

$$Pr\{\text{sunny on day } n \mid \text{sunny today}\} = \frac{6}{7} + \frac{1}{7}(0.3)^n$$

$$Pr\{\text{cloudy on day } n \mid \text{sunny today}\} = \frac{1}{7} - \frac{1}{7}(0.3)^n$$

$$Pr\{\text{sunny on day } n \mid \text{cloudy today}\} = \frac{6}{7} - \frac{6}{7}(0.3)^n$$

$$Pr\{\text{cloudy on day } n \mid \text{cloudy today}\} = \frac{1}{7} + \frac{6}{7}(0.3)^n$$

Suppose that on a sunny day a restaurant makes a profit of \$1000 and on a cloudy day a restaurant makes a profit of \$200. Find the expected profit n days from now in each of the cases that it is sunny today and cloudy today.

If we measure profit in \$100's then $r(1) = 10$ and $r(2) = 6$ and the profit vector is

$$r = \begin{pmatrix} 10 \\ 2 \end{pmatrix}.$$

The vector of expected profits on day n in each of the cases that it is sunny today and cloudy today is

$$P^n r \;=\; \left(\frac{1}{7}\right)\begin{pmatrix} 6 + (0.3)^n & 1 - (0.3)^n \\ 6 - 6(0.3)^n & 1 + 6(0.3)^n \end{pmatrix}\begin{pmatrix} 10 \\ 2 \end{pmatrix} \;=\; \left(\frac{1}{7}\right)\begin{pmatrix} 62 + 8\,(0.3)^n \\ 62 - 48\,(0.3)^n \end{pmatrix}$$

So

$$\text{Expected profit on day } n \text{ if it is sunny today} \;=\; \frac{62 + 8\,(0.3)^n}{7}$$

$$\text{Expected profit on day } n \text{ if it is cloudy today} \;=\; \frac{62 - 48\,(0.3)^n}{7}$$

Problem 1. A store sells 50" television sets. Let D be the demand for this item on any given day. Suppose $Pr\{D = 0\} = 0.3$, $Pr\{D = 1\} = 0.4$, $Pr\{D = 2\} = 0.2$, $Pr\{D = 4\} = 0.1$ and there is no chance they will sell more than four sets in any day. Suppose the demand for sets on any given day is independent of the demand on all the other days. If the store runs out of sets on any day then they order two more from the wholesaler and these are delivered overnight so that the inventory at the start of the next day is two. Let X_n be the inventory at the start of day n. We let n count the days with $n = 0$ being today. So X_n is either 1 or 2. This is a Markov chain with two states. State 1 corresponds to having one set at the start of a day and State 2 to having two sets. The transition matrix is

$$P \;=\; \begin{pmatrix} Pr\{1 \text{ set tomorrow} \mid 1 \text{ set today}\} & Pr\{2 \text{ sets tomorrow} \mid 1 \text{ set today}\} \\ Pr\{1 \text{ set tomorrow} \mid 2 \text{ sets today}\} & Pr\{2 \text{ sets tomorrow} \mid 2 \text{ sets today}\} \end{pmatrix}$$

$$= \begin{pmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{pmatrix}$$

(a) Find the characteristic equation for P.

Answer. $0 = \begin{vmatrix} 0.3 - \lambda & 0.7 \\ 0.4 & 0.6 - \lambda \end{vmatrix} = \lambda^2 - 0.9\lambda - 0.1.$

(b) Find the eigenvalues of P.

Answer. $\lambda^2 - 0.9\lambda - 0.1 = (\lambda - 1)(\lambda + 0.1)$, so the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 0.1$.

(c) For each eigenvalue of P, find an eigenvector.

Answer. For $\lambda_1 = 1$ one has $P - \lambda I = \begin{pmatrix} -0.7 & 0.7 \\ 0.4 & -0.4 \end{pmatrix}$ so an eigenvector $v = \begin{pmatrix} x \\ y \end{pmatrix}$ satisfies $\begin{pmatrix} 0 \\ 0 \end{pmatrix} =$

$(P - \lambda I)v = \begin{pmatrix} -0.7x + 0.7y \\ 0.4x - 0.4y \end{pmatrix}$.

So $0 = -0.7x + 0.7y$ and $0 = 0.4x - 0.4y$.

These imply $y = x$. So $v = \begin{pmatrix} x \\ x \end{pmatrix} = x\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Let's take $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. For $\lambda_2 = -0.1$ one has $P - \lambda I = \begin{pmatrix} 0.4 & 0.7 \\ 0.4 & 0.7 \end{pmatrix}$,

so $\begin{pmatrix} 0 \\ 0 \end{pmatrix} = (P - \lambda I)v = \begin{pmatrix} 0.4x + 0.7y \\ 0.4x + 0.7y \end{pmatrix}$. So $0 = 0.4x + 0.7y$ or $y = -\frac{4}{7}x$.

So $v = \begin{pmatrix} x \\ -4y/7 \end{pmatrix} = \frac{x}{7}\begin{pmatrix} 7 \\ -4 \end{pmatrix}$. Let's take $v_2 = \begin{pmatrix} 7 \\ -4 \end{pmatrix}$.

(d) Find the entries of P^n.

Answer. $P^n = TD^nT^{-1} = \begin{pmatrix} 1 & 7 \\ 1 & -4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (-0.1)^n \end{pmatrix}\begin{pmatrix} 1 & 7 \\ 1 & -4 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 7 \\ 1 & -4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (-0.1)^n \end{pmatrix}$

$\left(\frac{1}{-7}\right)\begin{pmatrix} -4 & -7 \\ -1 & 1 \end{pmatrix} = \left(\frac{1}{11}\right)\begin{pmatrix} 1 & (7)(-0.1)^n \\ 1 & (-4)(-0.1)^n \end{pmatrix}\begin{pmatrix} 4 & 7 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} \frac{4}{11} + \frac{7}{11}(-0.1)^n & \frac{7}{11} - \frac{7}{11}(-0.1)^n \\ \frac{4}{11} - \frac{4}{11}(-0.1)^n & \frac{7}{11} + \frac{4}{11}(-0.1)^n \end{pmatrix}$

(e) Find the probabilities of 1 or 2 sets on day n given there is 1 or 2 sets today.

Answer. P (1 set on day n | 1 set today) $= \frac{4}{11} + \frac{7}{11}(-0.1)^n$

P (2 sets on day n | 1 set today) $= \frac{7}{11} - \frac{7}{11}(-0.1)^n$

P (1 set on day n | 2 sets today) $= \frac{4}{11} - \frac{4}{11}(-0.1)^n$

P (2 sets on day n | 2 sets today) $= \frac{7}{11} + \frac{4}{11}(-0.1)^n$

Example 2. (see Problem 1 of section 2.1.1) At the start of each day an office copier is checked and its condition is classified as either good (1), poor (2) or broken (3). Suppose the transition matrix from one state to another in the course of day is

$$P = \begin{pmatrix} 0.8 & 0.06 & 0.14 \\ 0 & 0.4 & 0.6 \\ 0.32 & 0.08 & 0.6 \end{pmatrix}$$

To get the eigenvalues we must solve the equation

$$0 = \det(P - \lambda I) = \begin{vmatrix} 0.8 - \lambda & 0.06 & 0.14 \\ 0 & 0.4 - \lambda & 0.6 \\ 0.32 & 0.08 & 0.6 - \lambda \end{vmatrix}$$

Using mathematical software (see section 2.2.2) one obtains

$$\det(P - \lambda I) = -\lambda^3 + 1.8\lambda^2 - 0.9472\lambda + 0.1472 = -(\lambda - 1)(\lambda - 0.51)(\lambda - 0.29)$$

where we have rounded the numbers in the last formula to two decimal places. So the eigenvalues are

$\lambda_1 = 1 \quad \lambda_2 = 0.51 \quad \lambda_3 = 0.28$

Again using software, one finds the eigenvectors

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad v_2 = \begin{pmatrix} -0.284 \\ 0.942 \\ 0.178 \end{pmatrix} \quad v_3 = \begin{pmatrix} 0.064 \\ -0.942 \\ 0.185 \end{pmatrix}$$

So

$$P^n = TD^nT^{-1} = \begin{pmatrix} 1 & -0.284 & 0.064 \\ 1 & 0.942 & -0.942 \\ 1 & 0.178 & 0.185 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & (0.51)^n & 0 \\ 0 & 0 & (0.28)^n \end{pmatrix} \begin{pmatrix} 1 & -0.284 & 0.064 \\ 1 & 0.942 & -0.942 \\ 1 & 0.178 & 0.185 \end{pmatrix}^{-1}$$

Again using software, one obtains

$$P^n = \begin{pmatrix} 0.55 + (0.52)(0.51)^n - (0.078)(0.28)^n & 0.1 - (0.05)(0.51)^n - (0.05)(0.28)^n & 0.35 - (0.47)(0.51)^n + (0.12)(0.28)^n \\ 0.55 - (1.74)(0.51)^n + (1.19)(0.28)^n & 0.1 + (0.18)(0.51)^n + (0.72)(0.28)^n & 0.35 + (1.56)(0.51)^n - (1.91)(0.28)^n \\ 0.55 - (0.33)(0.51)^n - (0.22)(0.28)^n & 0.1 + (0.03)(0.51)^n - (0.14)(0.28)^n & 0.35 + (0.29)(0.51)^n + (0.36)(0.28)^n \end{pmatrix}$$

For example,

Pr { the copier is in good condition n days from now | it is broken today }$= (P^n)_{31}$

$$= 0.55 - (0.33)(0.51)^n - (0.22)(0.28)^n$$

Problem 2. Let $P = \begin{pmatrix} 1 - p & p \\ q & 1 - q \end{pmatrix}$

(a) Show that $\det(P - \lambda I) = \lambda^2 - (2 - p - q)\lambda + (1 - p - q)$

(b) Show that the eigenvalues of P are $\lambda_1 = 1$ and $\lambda_2 = 1 - p - q$

(c) Show that the eigenvectors of P are $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} p \\ -q \end{pmatrix}$

(d) Show that $P^n = \frac{1}{p+q}\begin{pmatrix} q + p(1-p-q)^n & p - p(1-p-q)^n \\ q - q(1-p-q)^n & p + q(1-p-q)^n \end{pmatrix}$

In order to complete the probabilistic description of the system we need to be given the initial probabilities that the system is in the various states at time 0, i.e.

$$p_j^{(0)} = Pr\{X_0 = j\}$$

The vector $p^{(0)} = (p_1^{(0)}, ..., p_s^{(0)})$ is called the *initial probability vector*. We think of this as a row vector as opposed to a column vector.

Example 1 (continued). Consider Example 1. If time $n = 0$ corresponds to today, then the probabilities of the system being in each state would depend on whether we had already checked the condition of the milling machines already today. Suppose we had and had observed that one machine was working and one was broken so the state is H today. Then the initial probability vector would be

$$p^{(0)} = (0, 1, 0)$$

corresponding to the fact that we know with perfect certainty that at the start of today the state was 2. On the other hand, if we had not already checked the condition of the two machines today, then the initial probability vector might depend on your experience with the two machines. Maybe, based on this experience, you might feel that the initial probability vector is the following.

$$p^{(0)} = (0.7, \quad 0.25, \quad 0.05)$$

Let's use this initial probability vector in the computations below.

In order to get the probability of two consecutive observations i and then j we multiply $p_i^{(0)}$ by p_{ij}, i.e. (3)

$$\Pr\{ X_0 = i,\ X_1 = j \} = \Pr\{ X_0 = i \} \frac{\Pr\{ X_0 = i,\ X_1 = j \}}{\Pr\{ X_0 = i \}}$$

$$= p_i^{(0)} \Pr\{ X_1 = j \mid X_0 = i \} = p_i^{(0)} p_{ij} \qquad (3)$$

Example 1 (continued). In the context of Example 1 using the above values, the probability that the system is in the broken state today and working state tomorrow is

$$Pr\{ X_0 = 3,\ X_1 = 1 \} = p_3^{(0)} p_{31} = (0.05)(0.6) = 0.03$$

Let us denote the probability that the state is j at time $n = 1$ by $p_j^{(1)}$, i.e.

$$p_j^{(1)} = Pr\{X_1 = j\}.$$

We can calculate this by summing (3) over the initial states, i.e.

$$p_j^{(1)} = Pr\{ X_1 = j \} = \sum_{i=1}^{s} Pr\{ X_0 = i, X_1 = j \} = \sum_{i=1}^{s} p_i^{(0)} p_{ij} = p^{(0)} \text{ (}j^{\text{th}}\text{ column of } P\text{)}$$

The vector $p^{(1)} = (p_1^{(1)}, \ldots, p_s^{(1)})$ of probabilities at time $n = 1$ can be found by matrix multiplication.

$$p^{(1)} = p^{(0)} P$$

In the context of 1 using the above values one has

$$p^{(1)} = p^{(0)} P = (0.7, \ 0.25, \ 0.05) \begin{pmatrix} 0.81 & 0.18 & 0.01 \\ 0 & 0.9 & 0.1 \\ 0.6 & 0 & 0.4 \end{pmatrix}$$

$= ((0.7)(0.81) + (0.25)(0) + (0.05)(0.6), \ (0.7)(0.18) + (0.25)(0.9) + (0.05)(0), \ (0.7)(0.01) + (0.25)(0.1) + (0.05)(0.4))$

$= (0.567 + 0 + 0.03, \ 0.126 + 0.225 + 0, \ 0.007 + 0.025 + 0.02)$

$= (0.597, \ 0.351, \ 0.052)$

Consider the probability of getting three consecutive observations,

i.e. $X_0 = i, X_1 = j, X_2 = k$. One has

$$Pr\{ X_0 = i, X_1 = j, X_2 = k\} = Pr\{ X_0 = i, X_1 = j \} Pr\{X_2 = k \mid X_0 = i, X_1 = j\}$$

Using the Markov assumption (1) this can be written as

$$Pr\{ X_0 = i, X_1 = j, X_2 = k\} = Pr\{ X_0 = i, X_1 = j \} Pr\{X_2 = k \mid X_1 = j\}$$

Using the time invariance assumption (2) this becomes

$$Pr\{ X_0 = i, X_1 = j, X_2 = k\} = Pr\{ X_0 = i, X_1 = j \} p_{jk}$$

Finally, using (3) we get

$$\Pr\{ X_0 = i, X_1 = j, X_2 = k\} = p_i^{(0)} p_{ij} p_{jk} \qquad (4)$$

In the context of Example 1 using the above values, the probability that the system is in the *Working* state today, state H tomorrow and state B the day after tomorrow would be $Pr\{ X_0 = 1, X_2 = 2, X_3 = 3 \} = p_1^{(0)} p_{12} p_{23} = (0.7)(0.18)(0.1) = 0.0126$

In a similar fashion one can show

$$\Pr\{ X_0 = i_0,\ X_1 = i_1, \ldots,\ X_n = i_n\} = p_{i_0}^{(0)} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} \qquad (5)$$

If we sum (4) over j we get

$$\Pr\{ X_0 = i, X_2 = k\} = p_i^{(0)} (P^2)_{ik} \qquad (6)$$

Dividing by $Pr\{ X_0 = i\}$ gives

$$\Pr\{ X_2 = k \mid X_0 = i\} = (P^2)_{ik} \qquad (7)$$

i.e. the entries of P^2 are the probabilities of being in various states at time $n = 2$ given one is in various states at time $n = 0$.

In the context of Example 1 using the above values one has

$$P^2 = PP = \begin{pmatrix} 0.81 & 0.18 & 0.01 \\ 0 & 0.9 & 0.1 \\ 0.6 & 0 & 0.4 \end{pmatrix}\begin{pmatrix} 0.81 & 0.18 & 0.01 \\ 0 & 0.9 & 0.1 \\ 0.6 & 0 & 0.4 \end{pmatrix} = \begin{pmatrix} 0.6621 & 0.3078 & 0.0301 \\ 0.06 & 0.81 & 0.13 \\ 0.726 & 0.108 & 0.166 \end{pmatrix}$$

For example, the probability the system is in state H the day after tomorrow given it is state W today is $(P^2)_{12} = 0.3078$.

Let us denote the probability that the state is k at time $n = 2$ by $p_k^{(2)}$, i.e.

$$p_k^{(2)} = Pr\{X_2 = k\}$$

We can calculate this by summing (6) over the initial states, i.e.

$$p_k^{(2)} = Pr\{ X_2 = k \} = \sum_{i=1}^{s} Pr\{ X_0 = i,\ X_2 = k \} = \sum_{i=1}^{s} p_i^{(0)} (P^2)_{ik} = p^{(0)} (k^{\text{th}} \text{ column of } P^2)$$

The vector $p^{(2)} = (p_1^{(2)}, \ldots, p_s^{(2)})$ of probabilities at time $n = 2$ can be found by matrix multiplication.

$$p^{(2)} = p^{(0)} P^2$$

In the context of Example 1 using the above values one has

$$p^{(2)} = p^{(0)} P^2 = (0.7, 0.25, 0.05) \begin{pmatrix} 0.6621 & 0.3078 & 0.0301 \\ 0.06 & 0.81 & 0.13 \\ 0.726 & 0.108 & 0.166 \end{pmatrix} = (0.51477, 0.42336, 0.06187)$$

We can extend this argument to any number of days in the future. One has

$$\Pr\{ X_n = k \mid X_0 = i\} = (P^n)_{ik} \qquad (8)$$

i.e. the entries of P^n are the probabilities of being in various states at time n given one is in various states at time $n = 0$.

Let us denote the the probability that the state is k at time n by $p_k^{(n)}$, i.e.

$p_k^{(n)} \;=\; \Pr\{X_n = k\}$

If we multiply (8) by $p_i^{(0)}$ we get

$$p_k^{(n)} \;=\; \Pr\{X_n = k\} \;=\; p_i^{(0)}(P^n)_{ik} \qquad (9)$$

If we sum this over the initial states we get

$$p^{(n)} \;=\; p^{(0)}\,P^n$$

For example P^5 has the transition probabilities in going from one day to five days later.

$$P^5 \;=\; \begin{pmatrix} 0.81 & 0.18 & 0.01 \\ 0 & 0.9 & 0.1 \\ 0.6 & 0 & 0.4 \end{pmatrix}^5 \;=\; \begin{pmatrix} 0.426708 & 0.496892 & 0.0763992 \\ 0.225203 & 0.652634 & 0.122162 \\ 0.53029 & 0.405366 & 0.0643433 \end{pmatrix}$$

Rewards, profits and costs.

Often one can assign a profit or cost (or some other type of reward) to being in each state at a particular time. Let $r_i = r(i)$ be the profit or cost for being in state i at a particular time. We think of this as a column vector instead of a row vector. Then the vector

$$r \;=\; \begin{pmatrix} r_1 \\ r_2 \\ \cdots \\ r_s \end{pmatrix}$$

is the "reward" or "cost" vector of profits or costs at a particular time for all the states. Now suppose

$$p \;=\; (Pr\{X_n = 1\}, Pr\{X_n = 2\}, \ldots, Pr\{X_n = s\}) \;=\; (p_1, p_2, \ldots, p_s)$$

are the probabilities of being in each of the states at time n. Then $r(X_n)$ is a random variable and its expected value is

$$E(\, r(X_n)\,) = Pr\{X_n = 1\}\; r(1) + Pr\{X_n = 2\}\; r(2) + \cdots + Pr\{X_n = s\}\; r(s)$$

$$= \;p_1\, r_1 + p_2\, r_2 + \cdots + p_s\, r_s \;=\; (p_1, p_2, \ldots, p_s)\begin{pmatrix} r_1 \\ r_2 \\ \cdots \\ r_s \end{pmatrix} \;=\; pr$$

= product of the probability vector p and reward vector r

Example 1 (continued). In the context of Example 1, suppose when the system is state H on a certain day it costs the office an extra 2000 Frs in lost productivity and when the system is in state B on a certain day it costs the office an extra 10000 Frs in lost productivity. Then if measure lost productivity in 1000's Frs one has $r(1) = 0$, $r(2) = 2$ and $r(3) = 10$ and the cost vector is

$$r = \begin{pmatrix} 0 \\ 2 \\ 10 \end{pmatrix}$$

Suppose the probabilities that the system is in each of the three states today is

$p^{(0)} = (0.7,\ \ 0.25,\ \ 0.05)$

Then the expected cost of lost productivity today is

$$E(r(X_0)) = p^{(0)}r = (0.7,\ \ 0.25,\ \ 0.05)\begin{pmatrix} 0 \\ 2 \\ 10 \end{pmatrix}$$

$$= (0.7)(0) + (0.25)(2) + (0.05)(10) = 0 + 0.5 + 0.5 = 1$$

i.e. \$1000. The expected cost of lost productivity tomorrow is

$$E(r(X_1)) = p^{(1)}r = p^{(0)}Pr = (0.7,\ \ 0.25,\ \ 0.05)\begin{pmatrix} 0.81 & 0.18 & 0.01 \\ 0 & 0.9 & 0.1 \\ 0.6 & 0 & 0.4 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \\ 10 \end{pmatrix}$$

$$= (0.7,\ \ 0.25,\ \ 0.05)\begin{pmatrix} 0.46 \\ 2.8 \\ 4 \end{pmatrix} = 0.322 + 0.7 + 0.2 = 1.222$$

i.e. 1222 Frs. The vector

$$Pr = \begin{pmatrix} 0.81 & 0.18 & 0.01 \\ 0 & 0.9 & 0.1 \\ 0.6 & 0 & 0.4 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \\ 10 \end{pmatrix} = \begin{pmatrix} 0.46 \\ 2.8 \\ 4 \end{pmatrix}$$

is of interest. It is the vector of expected costs lost productivity tomorrow given the system is in each of the three states today, i.e. $(Pr)_i = E(r(X_1) \mid X_0 = i)$.

More generally, the expected reward or cost for day n is $p^{(0)}P^n r$ and $P^n r$ is the vector of expected rewards or costs for day n given that the system is in each of the states today, , i.e. $(P^n r)_i = E(r(X_n) \mid X_0 = i)$

In Example 1, $P^2 r$ has the costs of lost productivity for the day after tomorrow for each of the possible system conditions today.

$$P^2 r = \begin{pmatrix} 0.6621 & 0.3078 & 0.0301 \\ 0.06 & 0.81 & 0.13 \\ 0.726 & 0.108 & 0.166 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \\ 10 \end{pmatrix} = \begin{pmatrix} 0.9166 \\ 2.92 \\ 1.876 \end{pmatrix}$$

And $p^{(0)}P^2 r$ is the expected cost of lost productivity for the day after tomorrow.

$$p^{(0)}P^2 r = (0.7,\ 0.25,\ 0.05)\begin{pmatrix} 0.9166 \\ 2.92 \\ 1.876 \end{pmatrix} = 1.46542$$

i.e. 1465.42 Frs.

Problem 1. An office copier is either in

1. good condition (G or 1),
2. poor condition (P or 2), or
3. broken (B or 3).

At the start of each day we check the condition of the copier. Suppose the condition of the copier on the start day n is X_n with $n = 0$ corresponding to today. Suppose $X_0, X_1, \ldots, X_n$ is a time homogeneous Markov chain. with transition matrix $P = \begin{pmatrix} 0.8 & 0.06 & 0.14 \\ 0 & 0.4 & 0.6 \\ 0.32 & 0.08 & 0.6 \end{pmatrix}$ and initial probability vector $p^{(0)} = (0.7, 0.25, 0.05)$. Suppose when the copier is in poor condition on a certain day it costs the office an extra \$50 in lost productivity and when the copier broken on a certain day it costs the office an extra 200 Frs in lost productivity so the cost vector is $r = \begin{pmatrix} 0 \\ 50 \\ 200 \end{pmatrix}$.

(a) What is the probability that the copier is in bad condition today and good condition tomorrow?

Answer. $Pr\{ X_0 = 3,\ X_1 = 1 \} = p_3^{(0)} p_{31} = (0.05)(0.32) = 0.016$.

(b) Find $p^{(1)}$.

Answer.

$$p^{(1)} = p^{(0)} P = (0.7,\ \ 0.25,\ \ 0.05)\begin{pmatrix} 0.8 & 0.06 & 0.14 \\ 0 & 0.4 & 0.6 \\ 0.32 & 0.08 & 0.6 \end{pmatrix} = (0.576,\ \ 0.146,\ \ 0.278).$$

(c) What is the probability that the copier is in good condition today, poor condition tomorrow and broken the day after tomorrow?

Answer. $Pr\{ X_0 = 1,\ X_2 = 2,\ X_3 = 3 \} = p_1^{(0)} p_{12} p_{23} = (0.7)(0.6)(0.6) = 0.0252$

(d) Find P^2 and the probability the copier is in poor condition the day after tomorrow given it is in good condition today.

Answer.

$$P^2 = \begin{pmatrix} 0.8 & 0.06 & 0.14 \\ 0 & 0.4 & 0.6 \\ 0.32 & 0.08 & 0.6 \end{pmatrix} \begin{pmatrix} 0.8 & 0.06 & 0.14 \\ 0 & 0.4 & 0.6 \\ 0.32 & 0.08 & 0.6 \end{pmatrix} = \begin{pmatrix} 0.6848 & 0.0832 & 0.232 \\ 0.192 & 0.208 & 0.6 \\ 0.448 & 0.0992 & 0.4528 \end{pmatrix}.$$

The probability the copier is in poor condition the day after tomorrow given it is in good condition today is $(P^2)_{12} = 0.0832$.

(e) Find $p^{(2)}$.

Answer. $$p^{(2)} = p^{(0)} P^2 = (0.7, 0.25, 0.05) \begin{pmatrix} 0.6848 & 0.0832 & 0.232 \\ 0.192 & 0.208 & 0.6 \\ 0.448 & 0.0992 & 0.4528 \end{pmatrix}$$

$$= (0.54976, 0.1152, 0.33504).$$

(f) Find P^5.

Answer. $$P^5 = \begin{pmatrix} 0.6848 & 0.0832 & 0.232 \\ 0.192 & 0.208 & 0.6 \\ 0.448 & 0.0992 & 0.4528 \end{pmatrix}^5 = \begin{pmatrix} 0.5715 & 0.0994 & 0.3291 \\ 0.4933 & 0.1092 & 0.3975 \\ 0.5409 & 0.1023 & 0.3568 \end{pmatrix}$$

(g) Find the expected cost for today.

Answer. $$E(r(X_0)) = p^{(0)}r = (0.7, 0.25, 0.05) \begin{pmatrix} 0 \\ 50 \\ 200 \end{pmatrix} = 22.50 \text{ Frs}$$

(h) Find the expected cost for tomorrow.

Answer. $$E(r(X_1)) = p^{(1)}r = p^{(0)}Pr = (0.7, 0.25, 0.05) \begin{pmatrix} 0.8 & 0.06 & 0.14 \\ 0 & 0.4 & 0.6 \\ 0.32 & 0.08 & 0.6 \end{pmatrix} \begin{pmatrix} 0 \\ 50 \\ 200 \end{pmatrix}$$

$$= (0.7, 0.25, 0.05) \begin{pmatrix} 31 \\ 140 \\ 124 \end{pmatrix} = \$62.90$$

(i) Find the vector of expected costs for tomorrow given the copier is in each of the three states today.

Answer. $$Pr = \begin{pmatrix} 0.8 & 0.06 & 0.14 \\ 0 & 0.4 & 0.6 \\ 0.32 & 0.08 & 0.6 \end{pmatrix} \begin{pmatrix} 0 \\ 50 \\ 200 \end{pmatrix} = \begin{pmatrix} 31 \\ 140 \\ 124 \end{pmatrix}$$

(j) Find the vector of expected costs for the day after tomorrow given the copier is in each of the three states today and the expected cost for the day after tomorrow.

Answer. $$P^2r = \begin{pmatrix} 0.6848 & 0.0832 & 0.232 \\ 0.192 & 0.208 & 0.6 \\ 0.448 & 0.0992 & 0.4528 \end{pmatrix} \begin{pmatrix} 0 \\ 50 \\ 200 \end{pmatrix} = \begin{pmatrix} 50.56 \\ 130.40 \\ 95.52 \end{pmatrix}.$$

Expected cost for the day after tomorrow = $p^{(0)}P^2 r$ = (0.7, 0.25,0.05) $\begin{pmatrix} 0 \\ 50 \\ 200 \end{pmatrix}$ = 72.77 Frs.

If the transition probability matrix of a given Markov chain is given by $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$. Suppose the initial probability distribution is given by $(0, 0,\ 1)$, find the probability distribution after three steps

Solution

(a) $p^{(1)} = p^{(0)}p = (0, 0, 1) \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$

$p^{(2)} = p^{(1)}p = \left(\frac{1}{2}, \frac{1}{2}, 0\right) \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \left(0, \frac{1}{2}, \frac{1}{2}\right)$

$p^{(3)} = p^{(2)}p = \left(0, \frac{1}{2}, \frac{1}{2}\right) \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right)$

$p_{S_1}^{(3)} = \frac{1}{4}, p_{S_2}^{(3)} = \frac{1}{4}, p_{S_3}^{(3)} = \frac{1}{2}$

Absorbing Markov Chain

This is a type of Markov chain in which when a certain state is reached, it is impossible to leave that state. Such states are called **absorbing states**, and a Markov Chain that has at least one such state is called an **Absorbing Markov chain.**

Definition: An **absorbing Markov chain** is Markov chains in which it is impossible to leave some states, and any state could (after some number of steps, with positive probability) reach such a state. It follows that all non-absorbing states in an absorbing Markov chain are transient.

An **absorbing state** is a state I in a Markov chain such that $P(X_{t+1} = i \mid X_t = i) = 1$. Note that it is not sufficient for a Markov chain to contain an absorbing state (or even several!) in order for it to be an absorbing Markov chain. It must also have all other states eventually reach an absorbing state with probability 1.

If the chain has t transient states and s absorbing states, the transition matrix P for a time-homogeneous absorbing Markov chain may then be written as

Suppose you have the following transition matrix.

$$\begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} .1 & .3 & .6 \\ 0 & 1 & 0 \\ .3 & .2 & .5 \end{bmatrix} \end{array}$$

The state S_2 is an absorbing state, because the probability of moving from state S_2 to state S_2 is 1. Which is another way of saying that if you are in state S_2, you will remain in state S_2.

In fact, this is the way to identify an absorbing state. If the probability in row i and column i,P_{ii}, is 1, then state S_i is an absorbing state.

Identify an absorbing state

A state S is an absorbing state in a Markov chain in the transition matrix if

- The row for state S has one 1 and all other entries are
- The entry that is 1 is on the main diagonal (row = column for that entry), indicating that we can never leave that state once it is entered.

Example

Consider transition matrices A, B, C for Markov chains shown below. Which of the following Markov chains have an absorbing state?

$$A = \begin{bmatrix} 0.3 & 0.7 & 0 \\ 0 & 1 & 0 \\ 0.2 & 0.3 & 0.5 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 0.1 & 0.3 & 0.4 & 0.2 \\ 0 & 0,2 & 0.1 & 0.7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Solution

$$A = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} .3 & .7 & 0 \\ 0 & 1 & 0 \\ .2 & .3 & .5 \end{bmatrix} \end{array}$$ has S_2 as an absorbing state.

If we are in state S_2, we cannot leave it. From state S_2, we can not transition to state S_1 or S_3; the probabilities are 0.

The probability of transition from state S_2 to state S_2 is 1.

$$B = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$ does not have any absorbing states.

From state S_1, we always transition to state S_2. From state S_2 we always transition to state S_3. From state S_3 we always transition to state S_1. In this matrix, it is never possible to stay in the same state during a transition.

$$C = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} .1 & .3 & .4 & .2 \\ 0 & .2 & .1 & .7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$ has two absorbing states, S_3 and S_4.

From state S_3, you can only remain in state S_3, and never transition to any other states. Similarly from state S_4, you can only remain in state S_4, and never transition to any other states.

Next we define an absorbing Markov Chain

Absorbing Markov Chain

A Markov chain is an absorbing Markov Chain if

- It has at least one absorbing state

- From any non-absorbing state in the Markov chain, it is possible to eventually move to some absorbing state (in one or more transitions).

Example

Consider transition matrices C and D for Markov chains shown below. Which of the following Markov chains is an absorbing Markov Chain?

$$C = \begin{bmatrix} 0.1 & 0.3 & 0.4 & 0.2 \\ 0 & 0,2 & 0.1 & 0.7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0 & 0 & 0 & 0.3 & 0.7 \\ 0 & 0 & 0 & 0.6 & 0.4 \end{bmatrix}$$

Solution

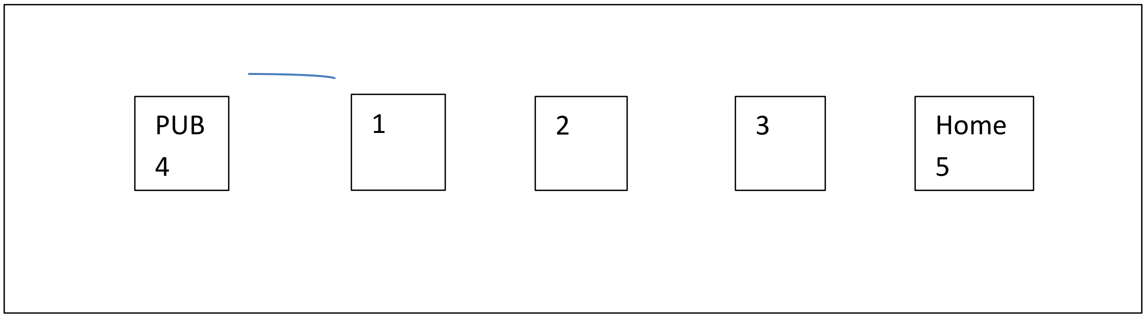
C is an absorbing Markov Chain but D is not an absorbing Markov chain.

Matrix C has two absorbing states, S_3 and S_4, and it is possible to get to state S_3 and S_4 from S_1 and S_2.

Matrix D is not an absorbing Markov chain. It has two absorbing states, S_1 and S_2, but it is never possible to get to either of those absorbing states from either S_4 or S_5. If you are in state S_4 or $S_{5,}$ you always remain transitioning between states S_4 or S_5 and can never get absorbed into either state S_1 or S_2

Example

A simple example of an absorbing Markov chain is the drunkard's walk of length $n + 2$. In the drunkard's walk, the drunkard is at one of n intersections between their house and the pub. The drunkard wants to go home, but if they ever reach the pub (or the house), they will stay there forever. However, at each intersection along the way, there is a probability p, typically $\frac{1}{2}$, that the drunkard will become confused and, losing their sense of direction, return to the previous intersection.



Example

A man walks along a three block portion of Main St. His house is at one end of the three block section. A bar is at the other end of the three block section. Each time he reaches a corner he randomly either goes forward one block or turns around and goes back one block. If he reaches home or the bar, he stays there. The four states are Home (H), Corner 1(C_1), Corner 2 (C_2) and Bar (B). Write the transition matrix and identify the absorbing states. Find the probabilities of ending up in each absorbing state depending on the initial state.

Solution

The transition matrix is written below.

$$T = \begin{array}{c} \\ H \\ C_1 \\ C_2 \\ B \end{array} \begin{array}{c} \begin{array}{cccc} H & C_1 & C_2 & B \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ .5 & 0 & .5 & 0 \\ 0 & .5 & 0 & .5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Home and the Bar are absorbing states. If the man arrives home, he does not leave. If the man arrives at the bar, he does not leave. Since it is possible to reach home or the bar from each of the other two corners on his walk, this is an absorbing Markov chain.

We can raise the transition matrix T to a high power, *n*

. One we find a power T^n that remains stable, it will tell us the probability of ending up in each absorbing state depending on the initial state.

$$T^{90} = \begin{array}{c c} & \begin{array}{cccc} H & C_1 & C_2 & B \end{array} \\ \begin{array}{c} H \\ C_1 \\ C_2 \\ B \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 1/3 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \quad \text{and} \quad T^{91} = \begin{array}{c c} & \begin{array}{cccc} H & C_1 & C_2 & B \end{array} \\ \begin{array}{c} H \\ C_1 \\ C_2 \\ B \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 1/3 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

$T^{91} = T^{90}$; the matrix does not change as we continue to examine higher powers. We see that in the long-run, the Markov chain must end up in an absorbing state. In the long run, the man must eventually end up at either his home or the bar.

The second row tells us that if the man is at corner C_1, then there is a 2/3 chance he will end up at home and a 1/3 chance he will end up at the bar.

The third row tells us that if the man is at corner C_2, then there is a 1/3 chance he will end up at home and a 2/3 chance he will end up at the bar.

Once he reaches home or the bar, he never leaves that absorbing state.

Note that while the matrix T^n for sufficiently large *n* has become stable and is not changing, it does not represent an equilibrium matrix. The rows are not all identical, as we found in the regular Markov chains that reached an equilibrium.

We can write a smaller "solution matrix" by retaining only rows that relate to the non-absorbing states and retaining only the columns that relate to the absorbing states.

Then the solution matrix will have rows C_1 and C_2, and columns H and B. The solution matrix is

Solution Matrix: $\begin{array}{c} \\ C_1 \\ C_2 \end{array}\begin{array}{c} \begin{array}{cc} H & B \end{array} \\ \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix} \end{array}$

The first row of the solution matrix shows that if the man is at corner C_1, then there is a 2/3 chance he will end up at home and a 1/3 chance he will end up at the bar.

The second row of the solution matrix shows that if the man is at corner C_2, then there is a 1/3 chance he will end up at home and a 2/3 chance he will end up at the bar.

The solution matrix does not show that eventually there is 0 probability of ending up in C1 or C2, or that if you start in an absorbing state H or B, you stay there. The smaller solution matrix assumes that we understand these outcomes and does not include that information.

Example 10

A gambler has 3,000 Frs, and she decides to gamble 1,000Frs at a time at a Black Jack table in a casino in Olorunti. She has told herself that she will continue playing until she goes broke or has 5,000 Frs. Her probability of winning at Black Jack is .40. Write the transition matrix, identify the absorbing states, find the solution matrix, and determine the probability that the gambler will be financially ruined at a stage when she has 2,000 Frs.

Solution

The transition matrix is written below. Clearly the state 0 and state 5K are the absorbing states. This makes sense because as soon as the gambler reaches 0, she is financially ruined and the game is over. Similarly, if the gambler reaches 5,000 Frs, she has promised herself to quit and, again, the game is over. The reader should note that $p_{00} = 1$, and $p_{55} = 1$.

Further observe that since the gambler bets only 1,000 Frs at a time, she can raise or lower her money only by 1,000 Frs at a time. In other words, if she has 2,000Frs now, after the next bet she can have 3,000 Frs with a probability of .40 and 1,000 Frs with a probability of .60.

$$
\begin{array}{c|cccccc}
 & 0 & 1K & 2K & 3K & 4K & 5K \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1K & .60 & 0 & .40 & 0 & 0 & 0 \\
2K & 0 & .60 & 0 & .40 & 0 & 0 \\
3K & 0 & 0 & .60 & 0 & .40 & 0 \\
4K & 0 & 0 & 0 & .60 & 0 & .40 \\
5K & 0 & 0 & 0 & 0 & 0 & 1
\end{array}
$$

To determine the long term trend, we raise the matrix to higher powers until all the non-absorbing states are absorbed. This is the called the **solution matrix.**

$$
\begin{array}{c|cccccc}
 & 0 & 1K & 2K & 3K & 4K & 5K \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1K & 195/211 & 0 & 0 & 0 & 0 & 16/211 \\
2K & 171/211 & 0 & 0 & 0 & 0 & 40/211 \\
3K & 135/211 & 0 & 0 & 0 & 0 & 76/211 \\
4K & 81/211 & 0 & 0 & 0 & 0 & 130/211 \\
5K & 0 & 0 & 0 & 0 & 0 & 1
\end{array}
$$

The solution matrix is often written in the following form, where the non-absorbing states are written as rows on the side, and the absorbing states as columns on the top.

$$
\begin{array}{c|cc}
 & 0 & 5K \\
\hline
1K & 195/211 & 16/211 \\
2K & 171/211 & 40/211 \\
3K & 135/211 & 76/211 \\
4K & 81/211 & 130/211
\end{array}
$$

The table lists the probabilities of getting absorbed in state 0 or state 5K starting from any of the four non-absorbing states. For example, if at any instance the gambler has 3,000 Frs, then her probability of financial ruin is 135/211 and her probability reaching 5K is 76/211.

Example

Solve the Gambler's Ruin Problem of the example given above without raising the matrix to higher powers, and determine the number of bets the gambler makes before the game is over.

Solution

In solving absorbing states, it is often convenient to rearrange the matrix so that the rows and columns corresponding to the absorbing states are listed first. This is called the **Canonical form.** The transition matrix of Example 1 in the canonical form is listed below.

$$
\begin{array}{c}
 \\ 0 \\ 5K \\ 1K \\ 2K \\ 3K \\ 4K
\end{array}
\begin{array}{c}
\begin{array}{cccccc} 0 & 5K & 1K & 2K & 3K & 4K \end{array} \\
\left[\begin{array}{cc|cccc}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
\hline
.60 & 0 & 0 & .40 & 0 & 0 \\
0 & 0 & .60 & 0 & .40 & 0 \\
0 & 0 & 0 & .60 & 0 & .40 \\
0 & .40 & 0 & 0 & .60 & 0
\end{array}\right]
\end{array}
$$

The canonical form divides the transition matrix into four sub-matrices as listed below.

$$
\begin{array}{c}
 \\ \text{Absorbing states} \\ \text{Non-absorbing states}
\end{array}
\begin{array}{c}
\begin{array}{cc} \text{Absorbing} & \text{Non-absorbing} \end{array} \\
\left[\begin{array}{c|c}
I_n & O \\
\hline
A & B
\end{array}\right]
\end{array}
$$

The matrix $F = (I - B)^{-1}$ is called the fundamental matrix for the absorbing Markov chain, where I_n is an identity matrix of the same size as B. The i, j-th entry of this matrix tells us the average number of times the process is in the non-absorbing state *j* before absorption if it started in the non-absorbing state *i*

The matrix $F = (I - B)^{-1}$for our problem is listed below.

$$F = \begin{array}{c} \\ 1K \\ 2K \\ 3K \\ 4K \end{array} \begin{array}{c} \begin{array}{cccc} 1K & 2K & 3K & 4K \end{array} \\ \begin{bmatrix} 1.54 & .90 & .47 & .19 \\ 1.35 & 2.25 & 1.18 & .47 \\ 1.07 & 1.78 & 2.25 & .90 \\ .64 & 1.07 & 1.35 & 1.54 \end{bmatrix} \end{array}$$

You can use your calculator, or a computer, to calculate matrix F.

The Fundamental matrix F helps us determine the average number of games played before absorption.

According to the matrix, the entry 1.78 in the row 3, column 2 position says that the gambler will play the game 1.78 times before she goes from 3KFrs to 2K Frs. The entry 2.25 in row 3, column 3 says that if the gambler now has 3K Frs, she will have 3K Frs on the average 2.25 times before the game is over.

We now address the question of how many bets will she have to make before she is absorbed, if the gambler begins with 3KFrs?

If we add the number of games the gambler plays in each non-absorbing state, we get the average number of games before absorption from that state. Therefore, if the gambler starts with $3K, the average number of Black Jack games she will play before absorption is

1.07+1.78+2.25+0.90=6.0

.That is, we expect the gambler will either have 5,000 Frs or nothing on the 7th bet.

Lastly, we find the solution matrix without raising the transition matrix to higher powers. The matrix FA gives us the solution matrix.

$$FA = \begin{bmatrix} 1.54 & 0.90 & 0.47 & 0.19 \\ 1.35 & 2,25 & 1.18 & 0.47 \\ 1.07 & 1.78 & 2.25 & 090 \\ 0.64 & 1.07 & 1.35 & 1.54 \end{bmatrix} \begin{bmatrix} 0.6 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0.4 \end{bmatrix} = \begin{bmatrix} 0.92 & 0.08 \\ 0.81 & 0.19 \\ 0.64 & 0.36 \\ 0.38 & 0.62 \end{bmatrix}$$

which is the same as the following matrix we obtained by raising the transition matrix to higher powers.

$$\begin{array}{c c} & \begin{array}{cc} 0 & 5K \end{array} \\ \begin{array}{c} 1K \\ 2K \\ 3K \\ 4K \end{array} & \begin{bmatrix} 195/211 & 16/211 \\ 171/211 & 40/211 \\ 135/211 & 76/211 \\ 81/211 & 130/211 \end{bmatrix} \end{array}$$

Example

At a professional school, students need to take and pass an English writing/speech class in order to get their professional degree. Students must take the class during the first quarter that they enroll. If they do not pass the class they take it again in the second semester. If they fail twice, they are not permitted to retake it again, and so would be unable to earn their degree.

Students can be in one of 4 states: passed the class (P), enrolled in the class for the first time (C), retaking the class (R) or failed twice and cannot retake (F). Experience shows 70% of students taking the class for the first time pass and 80% of students taking the class for the second time pass. Write the transition matrix and identify the absorbing states. Find the probability of being absorbed eventually in each of the absorbing states.

Solution

The absorbing states are P (pass) and F (fail repeatedly and cannot retake). The transition matrix T is shown below.

$$T=\begin{array}{c} \\ P \\ C \\ R \\ F \end{array}\begin{array}{c} \begin{array}{cccc} P & C & R & F \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ .7 & 0 & .3 & 0 \\ .8 & 0 & 0 & .2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

If we raise the transition matrix T to a high power, we find that it remains stable and gives us the long-term probabilities of ending up in each of the absorbing states.

$$T^{30}=\begin{array}{c} \\ P \\ C \\ R \\ F \end{array}\begin{array}{c} \begin{array}{cccc} P & C & R & F \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ .94 & 0 & 0 & .06 \\ .8 & 0 & 0 & .2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Of students currently taking the class for the first time, 94% will eventually pass. 6% will eventually fail twice and be unable to earn their degree. Of students currently taking the class for the second time time, 80% will eventually pass. 20% will eventually fail twice and be unable to earn their degree.

The solution matrix contains the same information in abbreviated form

$$\text{Solution Matrix}=\begin{array}{c} \\ C \\ R \end{array}\begin{array}{c} \begin{array}{cc} P & F \end{array} \\ \begin{bmatrix} .94 & .06 \\ .80 & .20 \end{bmatrix} \end{array}$$

Note that in this particular problem, we don’t need to raise T to a “very high” power. If we find T^2, we see that it is actually equal to T^n for higher powers *n*. T^n becomes stable after two transitions; this makes sense in this problem because after taking the class twice, the student must have passed or is not permitted to retake it any longer. Therefore the probabilities should not change any more after two transitions; by the end of two transitions, every student has reached an absorbing state.

Absorbing Markov Chains

1. A Markov chain is an absorbing Markov chain if it has at least one absorbing state. A state i is an absorbing state if once the system reaches state i, it stays in that state; that is, $P_{ii} = 1$
2. If a transition matrix T for an absorbing Markov chain is raised to higher powers, it reaches an absorbing state called the solution matrix and stays there. The *i, j*-th entry of this matrix gives the probability of absorption in state *j* while starting in state
3. Alternately, the solution matrix can be found in the following manner.

a. Express the transition matrix in the canonical form as below.

$T = \begin{bmatrix} I_n & 0 \\ A & B \end{bmatrix}$ Where I_n is an identity matrix, and **0** is a matrix of all zeros.

b) The fundamental matrix $F = (I - B)^{-1}$. The fundamental matrix helps us find the number of games played before absorption.

c) FA is the solution matrix, whose *i, j*-th entry gives the probability of absorption in state *j* while starting in state *i*.

4) The sum of the entries of a row of the fundamental matrix gives us the expected number of steps before absorption for the non-absorbing state associated with that row.

Advantages of Markov models

a) Markov models are relatively easy to derive from successional data
b) The computational requirements are modest and the analysis is mathematically tractable
c) The basic transition matrix summarizes the fundamental parameter of dynamical systems
d) The Markov model does not require deep insight into the mechanism of dynamic change, but it can help indicate areas where such insight would be useful and as such, act as a guide as well as stimulant for further research.

Disadvantages of Markov models

a) The validation of Markov chain models depends on observation and protection of system behavior over time, and is therefore frequently difficult, and sometimes may even be impossible to have this for a long period of time.

b) The Markov type model assumes that the conditional probability of transition from one state to another does not depend on n i.e.$P[X_n = j / X_{n-1} = i] = P_{ij}$ is independent on n. This does not apply in all situations.

c) C It does not take into consideration the duration of stay in a state before transition takes place. This has necessitated the development of semi-Markov models which is a generalization of the simple Markov model.

Classification of states of a Markov Chain:

If $P_{ij}^{(n)} > 0$ for some *n* and for all *i* and *j*, then every state can be reached from every other state. When this condition is satisfied, the Markov chain is said to be irreducible. The transition probability matrix of an irreducible chain is an irreducible matrix. Otherwise, the chain is said to be non-irreducible or reducible.

State *i* of a Markov chain is called a return state if $P_{ij}^{(n)} > 0$ for some *n*>1.

The period d_i of a return state *i* is defined as the greatest common divisor of all *m* such that $P_{ii}^{(m)} > 0 i.e., d_i = GCD\{m : P_{ii}(m) > 0\}$ State *i* is said to be periodic with period d_i if $d_i > 1$ and aperiodic if $d_i = 1$.

Obviously, state *i* is aperiodic if $p_{ii} \neq 0$. The probability that the chain returns to state *i* , having started from state *i* ,for the first time at the n^{th} step is denoted by $\boldsymbol{f_{ii}^{(n)}}$ and called the first return time probability or the recurrence time probability. $\{n, f_{ii}^{(n)}\} n = 1,2,3,\ldots$ is the distribution of recurrence times of the state *i*.

If $F_{ii} = \sum_{n=1}^{\infty} f_{ii} = 1$, the return to state *i* is certain.

$\mu_{ii} = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$ is called the mean recurrence time of the state *i*

A state *i* is said to be persistent or recurrent if the return to state *i* is certain i.e., if $F_{ii} = 1$. The state *i* is said to be transient if the return to the state *i* is uncertain, i.e., if $F_{ii} < 1$.The state *i* is said to be non-null persistent if it's mean recurrence time μ_{ii} is finite and null persistent if $\mu_{ii} = \infty$

A non-null persistent and aperiodic state is called ergodic.

Chapman-Kolmogorov Equation

If P is transition probability matrix of a homogeneous Markov Chain, then the n-step transition probability matrix

$$P^{(n)} = P^n \, (i.e) \left\{ P_{ij}{}^{(n)} \right\} = \left\{ P_{ij} \right\}^n$$

Example

A man either drives a car or catches a train to got to office each day. He never goes 2 days in a row by train, but if he drives one day, then the next day he is just as likely to drive again as he is to travel by train. Now suppose that on the first day of the week, the man tossed a fair die and drove to work iff a 6 appeared. Find

(i) the probability that he takes a train on the third day and

(ii) the probability that he drives to work in the long run

solution

: The travel pattern is a Markov chain, with state space = (train,car)

The TPM of the chain is $P = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

The initial state probability distribution is $p^{(1)} = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} \end{pmatrix}$ since

P(traveling by car)=P(getting 6 in the toss of the die)= $\frac{1}{6}$

P(traveling by train)= $\frac{5}{6}$

$$p^{(2)} = p^{(1)} P = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{12} & \frac{11}{12} \end{pmatrix}$$

$$p^{(3)} = p^{(2)}P = \begin{pmatrix} \frac{1}{12} & \frac{11}{12} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{11}{24} & \frac{13}{24} \end{pmatrix}$$

Therefore, P(the man travels by train on the third day)= $\frac{11}{24}$ Let $\pi = (\pi_1, \pi_2)$ be the limiting form of the state probability distribution or stationary state distribution of the Markov chain.

By the property of π, $\pi P = \pi$

$$(\pi_1, \pi_2) \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = (\pi_1, \pi_2)$$

$$\frac{1}{2}\pi_2 = \pi_1 \text{-------------------(1)}$$

$$\pi_1 + \frac{1}{2}\pi_2 = \pi_2 \text{-------------(2)}$$

Equations (1) and (2) are one and the same.

Therefore, consider (1) or (2) with $\pi_1 + \pi_2 = 1$ since π is a probability distribution.

$$\frac{1}{2}\pi_2 = 1 - \pi_2$$

$$\frac{1}{2}\pi_2 + \pi_2 = 1$$

$$\frac{3}{2}\pi_2 = 1$$

$$\pi_2 = \frac{2}{3}$$

$$\pi_1 + \frac{2}{3} = 1$$

$$\pi_1 = 1 - \frac{2}{3} = \frac{1}{3}$$

Therefore, P[the man travels by car in the long run]= $\frac{2}{3}$

2Three boys A, B and C are throwing a ball to each other. A always throws the ball to B and B always throws the ball to C, but C is just as likely to throw the ball to B as to A. Show that the process is Markovian. Find the transition matrix and classify the states.
solution:

The transition probability matrix of the process $\{X_n\}$is given below.

$$P = states\, of\ X_{n-1}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

States of X_n depends only on states of X_{n-1} but not on states of $X_{n-2}, X_{n-3}, \ldots$ or earlier states. Therefore, $\{X_n\}$is a Markov chain.

Now, $P^2 = P.P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

$$P^3 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}; P^4 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}; P^5 = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \end{bmatrix}; P^6 = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{8} & \frac{3}{8} \end{bmatrix} \text{ and so on}$$

$p_{11}^{(3)} > 0, p_{13}^{(2)} > 0, p_{21}^{(2)} > 0, p_{22}^{(2)} > 0, p_{33}^{(2)} > 0$ and all other $p_{ij}^{(1)} > 0$. Therefore, the chain is irreducible.

We note that $p_{ii}^{(2)}, p_{ii}^{(3)}, p_{ii}^{(5)}, p_{ii}^{(6)}$ etc are >0 for *i=2,3*and GCD of 2,3,5,6,….=1

Therefore, the states 2 and 3 are periodic with period 1 i.e., aperiodic

We note that $p_{11}^{(2)}, p_{11}^{(3)}, p_{11}^{(5)}, p_{11}^{(6)}$ etc., aare>0 and GCD of 3,5,6,…=1.Therefore, the state 1 is periodic with period 1, i.e, aperiodic.

Since the chain is finite and irreducible all its states are nonnull persistent.

Moreover, all the states are ergodic.

Semi-Markov Process

A semi-Markov process is equivalent to a Markov renewal process in many aspects, except that a state is defined for every given time in the semi-Markov process, not just at the jump times.
The semi-Markov process is an actual stochastic process that evolves over time.

The Bernoulli Process

• The Bernoulli process is an infinite sequence $X_1, X_2, \cdots$ of i.i.d. Bern(p) r.v.s

• The outcome from a Bernoulli process is an infinite sequence of 0s and 1s

• A Bernoulli process is often used to model occurrences of random events; $X_n = 1$ if an event occurs at time n, and 0, otherwise

• Three associated random processes of interest:

◦ Binomial counting process: The number of events in the interval [1, n]

◦ Arrival time process: The time of event arrivals

◦ Interarrival time process: The time between consecutive event arrivals

• We discuss these processes and their relationships

Binomial Counting Process

• Consider a Bernoulli process $X_1, X_2, \cdots$ with parameter p

• We are often interested in the number of events occurring in some time interval

• For the time interval [1, n], i.e., i = 1, 2, , . . . , n, we know that the number of occurrences

$w_n = n\ X_i = 1, X_n \sim B(n, p)$

• The sequence of r.v.s $w_1, w_2, \cdots$ is referred to as a Binomial counting process

• The Bernoulli process can be obtained from the Binomial counting process as:

$X_n = w_{n-1}, for\ n = 1, 2, \cdots$

where $w_0 = 0$

Markov chain

Example: Exponential R.V.

Exponential r.v. X with parameter has PDF $p(X < h) = 1 - e^{-h}, h > 0$

$$p[X \leq t + h/X > t] = p[X \leq h]$$

$$= 1 - e^{-\lambda h}$$

$$= 1 - \left[1 - \lambda h + \sum_{n=2}^{\infty} \frac{(\lambda h)^n}{n!}\right]$$

$$= \lambda h + 0(h)$$

Poisson Process

1. The process $\{X(t), t > 0\}$ has independent increment $t_0 < t_1 < t_2 < t_3$ ….

Increments $X(t_1) - X(t_0), X(t_2) - X(t_1)$ ...

These increment are independent i.e. the distribution of $X(t - h) - X(t) = R(h)$ depends only on h.

2. Consider a very small interval of time $(t, t + h)$

$P\left[X(t, t + h) - X(t) = {}^{1}/_{X(t)}\right] = \lambda h + 0(h)\ Where \lim_{h-0} \frac{0(h)}{h} = 0$

3. $P\left[X(t + h) = {}^{0}/_{X(t)} = 0\right] = 1 - \ \lambda h + 0(h)$

Note: The value $0(h)$is so small that could be taken equal to zero

To obtain the distribution of the process

Define $P_k(t) = P[X(t) = k]$and consider the event that $X(t + h) = k$

1. $X(t) = k\ and\ X(t + h) - X(t) = 0$

2. $X(t) = k - 1 \; and \; X(t+h) - X(t) = 1$
3. $X(t) = k - i \; and \; X(t+h) - X(t) = i, i \geq 2$

$\Rightarrow P_k(t+h) = P_k(t)P_0(h) + P_{k+1}(t)P_0(h) + \sum_{i=2}^{k} P_{k-i}(t)P_i(h)$

$P_k(t)[1 - \lambda h + 0(h)] + P_{k+1}(t)[\lambda h + 0(h)] + \sum_{i=2}^{k} P_{k-i}(t)0(h)$

Note: $0(h)$ multiply by any quantity remains an $0(h)$

$P_k(t+h) = P_k(t) - \lambda h P_k(t) + 0(h) + \lambda h P_{k-1}(t) + 0(h)$

Rearrange, divide by h and take limit as h tends to zero, we obtain

$P_k^1(t) = -\lambda P_k(t) + \lambda P_{k-1}(t), K = 1,2, \ldots$

Now consider the Solution of the equation 1.0 when $k = 0$

$P_0^1(t) = -\lambda P_0(t) \; -\lambda = \frac{P_0^1(t)}{P_0(t)} \Rightarrow \log P_0(t) = -\lambda t + c \; Since \int \frac{dt}{P_0(t)} = -\int \lambda dt$

$P_0(t) = e^{-\lambda t + c}$

Recall $P_0(0) = P[X(0) = 0] = 1, 1 = e^c \; that \; is \; c = 0$

$P_0(t) = e^{-\lambda t}$

$when \; k = 1, we \; obtain$

$P_1^1(t) = -\lambda P_1(t) + \lambda P_0(t) \; -\lambda P_1(t) + \lambda e^{-\lambda t}$

$P_1(t) = \lambda t e^{-\lambda t}$

In general,

$P_k^1(t) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}, k = 0,1,2,3 \ldots$

$which \; shows \; that \; the \; arrival \; process \; has \; poisson \; distribution$

With E(t) $= \lambda t \; and \; t) \; V(t) = \lambda t$

SERVICE PROCESS

There are no branches in the service process.

Suppose μ is the service rate

$Pr(T > t) = e^{-\mu t}$

$Pr(T < t) = 1 - e^{-\mu t}$

To obtain the pdf,

$f(t) = \frac{d}{dt}[F(t)] = \mu e^{-\mu t}, t > 0$

The service process has exponential distribution with parameter μ

$\mathrm{E(t)} = \frac{1}{\mu}$, and $\mathrm{V(t)} = \frac{1}{\mu^2}$

Arrival process

Let X is a patient's arrival process then, X has a Poisson distribution with parameter (λ).

$f(x) = \frac{e^{-\lambda}}{x!}\lambda^{x}, \;\; x = o, 1, 2 \ldots$

$\mathrm{E(x)} = \lambda, \; \mathrm{v(x)} = \lambda$

Proof: using MGF approach

$\mathrm{M_X(t)} = \mathrm{E(e^{tx})} = \sum_{Rx} f(x).\mathrm{e^{tx}} = \sum_{Rx} \frac{e^{-\lambda}}{x!}\lambda^{x}.\mathrm{e^{tx}} = \sum_{Rx} \frac{e^{-\lambda}}{x!}.(\lambda \mathrm{e^t})^{x}$

$= \frac{e^{-\lambda}}{e^{-\lambda \mathrm{e^t}}} \sum_{Rx} \frac{e^{-\lambda}}{x!}.(\lambda \mathrm{e^t})^{x}.e^{-\lambda \mathrm{e^t}} = e^{-\lambda}.e^{\lambda \mathrm{e^t}}$

$\{\mathrm{M_X(t)}\} = e^{\lambda(\mathrm{e^t}-1)}$

$\mathrm{E(x)} = \frac{\boldsymbol{d}}{\boldsymbol{dt}}\{\mathrm{M_X(t)}\}/_{\boldsymbol{t=0}}$

$\lambda \mathrm{e^t}.e^{\lambda(\mathrm{e^t}-1)}/_{t=0} = \lambda$

$v(x) = M_X^{11}(t) - \{\mathrm{M^1{}_X(t)}\}^2/_{\mathrm{t=0}}$

$(\lambda \mathrm{e^t})^2 + \lambda \mathrm{e^t}.e^{\lambda(\mathrm{e^t}-1)} - \lambda^2$

$= \lambda^2 + \lambda - \lambda^2 = \lambda$

3.7.2 Service process

Let y be the service process, then y follows exponential distribution with paraeter μ,

$f(y) = \mu e^{-\mu y}, 0 < y < \infty$

$E(x) = \frac{1}{\mu}, v(y) = \frac{1}{\mu^2}$

Proof: using MGF since y is continuous,

$M_y(t) = E(e^{ty}) = \int_0^\infty e^{ty}.\, f(y)dy = \int_0^\infty e^{ty}.\mu e^{-\mu y}dy = \int_0^\infty \mu e^{ty-\mu y}. = \;\; \int_0^\infty \mu \mathrm{e}^{-y(\mu-t)}dy$

Using integration by part,

$M_y(t) = \; \mu[\mu - \mathrm{t}]^{-1}$

$now\ to\ find\ the\ mean\ , E(y) = M_y^1(0) = -\mu[\mu - \mathrm{t}]^{-2}.-1 = \frac{1}{\mu}$

$v(y) = M_y^{11}(0) - [M_y^1(0)]^2$

$= -2\mu[\mu - \mathrm{t}]^{-3}.-1 = 2\mu[\mu - \mathrm{t}]^{-3} = \frac{2\mu}{\mu^3} - [M_y^1(0)]^2 = \frac{2\mu}{\mu^3} - \frac{1}{\mu^2} = \frac{2\mu-\mu}{\mu^3} = \frac{1}{\mu^2}$

This is the variance of the distribution

Birth-Death Process

Continuous-time, discrete-space stochastic process $\{N(t), t > 0\}$, $N(t) \in \{0, 1, ...\}$

N(t): population at time t

- $P(N(t+h) = n+1 \mid N(t) = n) = \lambda_n h + o(h)$
- $P(N(t+h) = n-1 \mid N(t) = n) = \mu_n h + o(h)$
- $P(N(t+h) = n \mid N(t) = n) = 1-(\lambda_n + \mu_n) h + o(h)$
- λ_n - birth rates
- μ_n - death rates, $\mu_0 = 0$

Q: what is $P_n(t) = P(N(t) = n)$? $n = 0,1,.$

- Similar to Poisson process drift equation

$$\frac{dP_n(t)}{dt} = P_{n-1}(t)\lambda_{n-1} + P_{n+1}(t)\mu_{n+1} - (\lambda_n + \mu_n)P_n(t)' \qquad n = 1, 2, \cdots$$

$$\frac{dP_0(t)}{dt} = P_1(t)\mu_1 - (\lambda_0 + \mu_0)P_0(t)'$$

Initial Condition: $\boldsymbol{P_n(0)}$

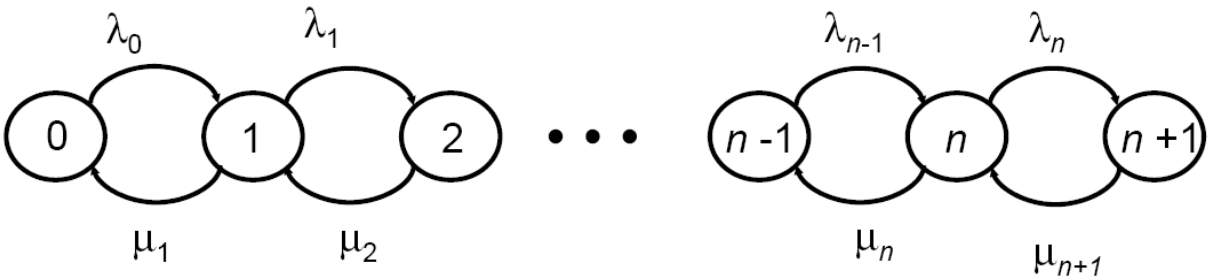
- If $\mu_i = 0$, $\lambda_i = \lambda$, then B-D process is a Poisson process

Stationary Behavior of B-D Process

- Most real systems reach equilibrium as $t \rightarrow \infty$
  - No change in $P_n(t)$ as t changes
  - No dependence on initial condition
- $P_n = \lim_{t \rightarrow \infty} P_n(t)$
- Drift equation becomes:

$$(\lambda_n + \mu_n)\, P_n = \lambda_{n-1}\, P_{n-1} + \mu_{n+1}\, P_{n+1}$$

Transition State Diagram



- Balance Equations:
  - Rate of trans. into n = rate of trans. out of n

$$\lambda_{n-1} P_{n-1} + \mu_{n+1} P_{n+1} = (\lambda_n + \mu_n) P_n, \qquad n \geq 1$$

$$\mu_1 P_1 = \lambda_0 P_0,$$

  - Rate of trans. to left = rate of trans. to right

$$\lambda_{n-1} P_{n-1} = \mu_n P_n$$

- Probability requirement:

$$\sum_{n=0}^{\infty} P_n = 1$$

Markov Process

- Prob. of future state depends only on present state
- $\{X(t), t>0\}$ is a MP if for any set of time $t_1<\cdots<t_{n+1}$ and any set of states $x_1<\cdots<x_{n+1}$
  - $P(X(t_{n+1})=x_{n+1}|X(t_1)=x_1, \cdots X(t_n)=x_n\}$

$= P(X(t_{n+1})=x_{n+1}| X(t_n)=x_n\}$

B-D process, Poisson process are MP

Markov Chain

- Discrete-state MP is called Markov Chain (MC)
  - Discrete-time MC
  - Continuous-time MC
- First, consider discrete-time MC

$$P_{ij} = P(X_{n+1} = j \mid X_n = i), i, j = 0, 1, \ldots; n \geq 0$$

Define transition prob. matrix:

$$P = [P_{ij}]$$

Chapman-Kolmogorov Equation

- What is the state after n transitions?

A: define

$$P_{ij}^n = P(X_{n+m} = j \mid X_m = i),\ n \geq 0,\ i, j \geq 0$$

$$\begin{aligned} P_{ij}^{n+m} &= P(X_{n+m} = j \mid X_0 = i), \\ &= \sum_{k=0}^{\infty} P(X_{n+m} = j, X_n = k \mid X_0 = i), \end{aligned}$$

$$= \sum_{k=0}^{\infty} P(X_{n+m} = j \mid X_n = k, X_0 = i) P(X_n = k \mid X_0 = i),$$

$$= \sum_{k=0}^{\infty} P(X_{n+m} = j \mid X_n = k) P(X_n = k \mid X_0 = i),$$

$$= \sum_{k=0}^{\infty} P_{kj}^{m} P_{ik}^{n}$$

- If MC has n state

$$P_{ij}^{2} = \sum_{k=1}^{n} P_{ik} P_{kj} \Rightarrow [P_{ij}^{2}] = P \cdot P$$

- Define n-step transition prob. matrix:

$$P^{(n)} = [P_{ij}^{n}]$$

C-K equation means:

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

$$P^{(n)} = P \cdot P^{(n-1)} = P^{n}$$

Markov Chain

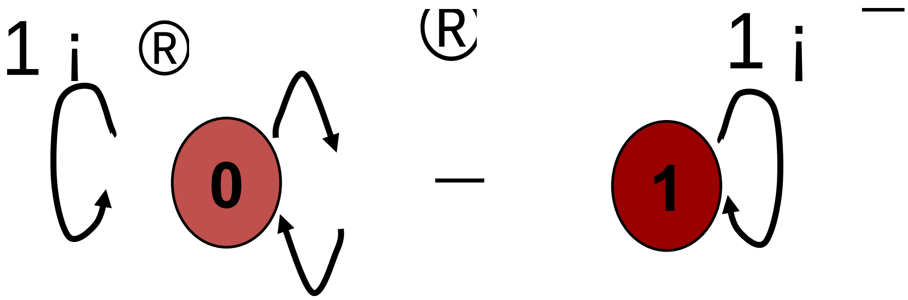
- Irreducible MC:
    - If every state can be reached from any other states
- Periodic MC:
    - A state i has period k if any returns to state i occurs in multiple of k steps
    - k=1, then the state is called aperiodic
    - MC is aperiodic if all states are aperiodic
- An irreducible, aperiodic finite-state MC is ergodic, which has a stationary (steady-state) prob. distr.

¼ = (¼0 ; ¼1 ; ¢¢¢ ; ¼n)

$$\pi = \pi P,$$

$$\pi\mathbf{1} = 1$$

where $\mathbf{1} = (1 \ \ldots)^T$

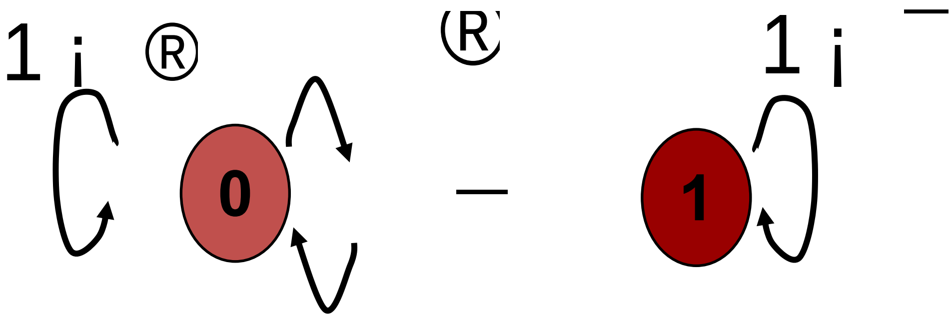


- Markov on-off model (or 0-1 model)
- Q: the steady-state prob.?

P = 1 ¡ ® ® ; ¯ 1 ¡ ¯

¼0 = (1 ¡ ®)¼0 + ¯¼1 ; ¼1 = ®¼0 + (1 ¡ ¯)¼1) ¼0 = ¯/(®+¯) ; ¼1 = ®/(®+¯)

- Use balance equation:



- Rate of trans. to left = rate of trans. to right

$$\alpha \pi_0 = \beta \pi_1$$

$$\pi_0 + \pi_1 = 1$$

$$\Rightarrow \begin{cases} \pi_0 = \dfrac{\beta}{\alpha+\beta} \\ \pi_1 = \dfrac{\alpha}{\alpha+\beta} \end{cases}$$

Discrete-Time MC State Staying Time

- X_i: the number of time steps a MC stays in the same state i
- $P(X_i = k) = P_{ii}^{k-1}(1-P_{ii})$
    - X_i follows geometric distribution
    - Average time: $1/(1-P_{ii})$
- In continuous-time MC, the staying time is?
    - Exponential distribution time

Homogeneous Continuous-Time Markov Chain

- P(X(t+h)=j|X(t)=i) = ¸$_{ij}$h +o(h)

- We have the properties:

$$P(X(t+h)=i \mid X(t)=i) = 1 - \sum_{j \neq i} \lambda_{ij} h + o(h)$$

$$P(X(t+h) \neq i \mid X(t)=i) = \sum_{j \neq i} \lambda_{ij} h + o(h)$$

  - The state holding time is exponential distr. with rate

$$\lambda = \sum_{j \neq i} \lambda_{ij}$$

    - Why?
    - Due to the summation of independent exponential distr. is still exponential distr.

Steady-State

- Ergodic continuous-time MC
- Define $\pi_i = P(X=i)$
- Consider the state transition diagram
  - Transit out of state i = transit into state i

$$\pi_i \sum_{j \neq i} \lambda_{ij} = \sum_{j \neq i} \pi_j \lambda_{ji}$$

$$\sum_i \pi_i = 1$$

Infinitesimal Generator

- Define Q = [q_{ij}] where

$$q_{ij} = \begin{cases} -\nu_i = -\sum_{k \neq i} \lambda_{ik} & \text{when } i = j \\ \lambda_{ij} & i \neq j \end{cases}$$

- Q is called infinitesimal generator

$$\pi = [\pi_1 \ \pi_2 \ \cdots]$$

$$\pi Q = 0$$

$$\pi \mathbf{1} = 1$$

Discrete vs. Continues MC

- Discrete
  - Jump at time tick
  - Staying time: geometric distr.
  - Transition matrix P
  - Steady state:

$$\pi = \pi P,$$
$$\pi \mathbf{1} = 1$$

  - State transition diagram:
    - Has self-jump loop
    - Probability on arc
- Continuous

- Jump at continuous t
- Staying time: exponential distr.
- Infinitesimal generator Q
- Steady state:
- State transition diagram:
  - No self-jump loop
  - Transition rate on arc

Semi-Markov Process

- X(t): discrete state, continuous time
  - State jump: follow Markov Chain Z_n
  - State i holding time: follow a distr. $Y^{(i)}$
- If $Y^{(i)}$ follows exponential distr. ¸
  - X(t) is a continuous-time Markov Chain

Steady State

- Let ¼'$_j$=lim$_{n\to 1}$ P(Z_n=j)
- Let ¼$_j$ = lim$_{t\to 1}$ P(X(t)=j)

$$\pi_j = \frac{\pi'_j E[Y^{(j)}]}{\sum_{i\in S} \pi'_i E[Y^{(i)}]}, \quad j \in S$$

CLASSIFICATION OF GENERAL STOCHASTIC PROCESSES

INTRODUCTION

The main elements of distinguishing stochastic processes are in the nature of the statespace, the index parameter *T*, and the dependence relations among the random variables X_t.

State Space S

This is the space in which the possible values of each X_t lie. In the case that S =(0,1,2,…), we refer to the process as integer valued, or alternatively as a discrete state process. If S = the real line (−∞, ∞), then we call X_t a real-valued stochastic process. If S is the Euclidean *k* spaced then X_t is said to be a *k*-vector process.

The choice of state space is not uniquely specified by the physical situation being described, although usually one particular choice may stand out as most appropriate.

Index Parameter T

If $T = (0,1,2,\cdots)$, then we state that X_t is a discrete time stochastic process. Often when T is discrete we shall write X_n instead of X_t. If $T = [0,\infty)$, then X_t is called a continuous time process.

CLASSICAL TYPE OF STOCHASTIC PROCESSES

We now describe (first briefly) then in detail, some of the classical types of stochastic processes characterized by different dependence relationships among X_t. Unless otherwise stated, we take $T = [0,\infty)$ and assume the random variables X_t are real valued.

Process with Stationary Independent Increments

If the random variables $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, X_{t_4} - X_{t_3} \cdots, X_{t_n} - X_{t_{n}-1}$ are independent for all choices of $t_1, t_2, t_3, \cdots, t_n$ satisfying $t_1 < t_2 < t_3 < \cdots < t_n$ then we say that X_t is a process with independent increments.

If the index set contains a smallest index X_0, it is also assumed X_{t_1}, $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \cdots, X_{t_n} - X_{t_{n}-1}$ are independent. If the index set is discrete that is T = (0,1,…), then a process with independent increments reduces to a sequence of independent random variables $Z_0 = X_0, Z_i = X_i - X_{i-1}, i = 1,2,3,\cdots$,

in the sense that knowing the individual probabilities/distributions of $Z_0, Z_1, \cdots$ enables us to determine the joint distributions of any finite set ofX_i, in fact that of $X_i = Z_1 + Z_2 + \cdots + Z_i, i = 1,2,3,\cdots$

REMARKS/DEFINITIONS

1. If the distribution of the increments $X(t_1 + h) - X(t_1)$ depends only on the length h of the interval and not on the time t_1, the process is said to have stationary increments.

2. For a process with stationary increments, the distribution of $X(t_1 + h) - X(t_1)$ is the same as the distribution of $X(t_2 + h) - X(t_2)$, no matter the values oft_1, t_2 and h.

3. We now state a theorem. If a process $\{X(t), t \in T\}$, where $T = [0, \infty)$ or $T = (1, 2, \cdots)$ has stationary independent increments and has a finite mean, then it is true that:

(a) $E(X_t) = M_0 + M_1 t$ where $M_0 = E(X_0)$ and $M_1 = E(X_1) - M_0$

(b) $\sigma_t^2 = \sigma_0^2 + \sigma_1^2 t$ where $\sigma_0^2 = E[(X_0 - M_0)^2]$ and $\sigma_1^2 = E[(X_1 - M_1)^2]$

4. Both the Brownian motion process and the Poisson process have stationary independent increments.