7. Markov Chains (Discrete-Time Markov Chains)

7.1. Introduction: Markov Property

7.2. Examples

- Two States
- Random Walk
- Random Walk (one step at a time)
- Gamblers' Ruin
- Urn Models
- Branching Process

7.3. Marginal Distribution of X_n

- Chapman-Kolmogorov Equations
- Urn Sampling
- Branching Processes
  - Nuclear Reactors
  - Family Names

7.4 Appendix: Notes on Matrices: I

7.1. Introduction: Markov Chains

Consider a system which can be in one of a countable number of states $1, 2, 3, \ldots$. The system is observed at the time points $n = 0, 1, 2, \ldots$.

Define X_n to be a random variable denoting the state of the system at "time" n. Suppose the history of the system up to time n is: $\{X_0, X_1, \ldots, X_n\}$. The probability distribution of X_{n+1} would ordinarily depend on the past history; i.e.

$$P\{X_{n+1}|X_0, X_1, \ldots, X_n\}.$$

The process is said to have the Markov property if

$$\boxed{P\{X_{n+1}|X_0, X_1, \ldots, X_n\} = P\{X_{n+1}|X_n\}}$$

$$\boxed{P\{X_{n+1}|X_0,\ldots,X_n\} = P\{X_{n+1}|X_n\}}$$

The stochastic process is called a Markov Chain. If the possible states are denoted by integers, then we have

$$P\{X_{n+1}=j|X_n=i, X_{n-1}=i_{n-1}, X_{n-2}=i_{n-2},\ldots,X_0=i_0\}$$

$$= P\{X_{n+1}=j|X_n=i\}$$

Define

$$p_{ij}(n) = P\{X_{n+1}=j|X_n=i\}$$

If S represents the state space and is countable, then the Markov Chain is called Time-Homogeneous if

$$p_{ij}(n) = p_{ij} \quad \text{for all } i,j \in S \text{ and } n \geq 0.$$

We will only be dealing with Time Homogeneous Markov Chains.

Note: Sometimes this process is referred to as a Discrete Time Markov Chain (DTMC).

Define $P = (p_{ij})$.
If S has m states, then $P = (p_{ij})$ $m \times m$ matrix.

P is often called the one-step transition probability matrix.

Definition: A matrix $P = (P_{ij})$ is called stochastic if

$$(i) \quad p_{ij} \geq 0 \qquad i, j \in S$$

$$(ii) \quad \sum_{j \in S} p_{ij} = \sum_{j=1}^{m} p_{ij} = 1 \text{ for all } i \in S.$$

$$
\begin{aligned}
X_0 &= \text{initial state} \\
a_i &= P\{X_0 = i\} = \text{Prob. of the initial state } X_0 = i.
\end{aligned}
$$

The probabilities a_i and $P = (p_{ij})$ completely determine the stochastic process.

Examples

$$
\begin{aligned}
P\{X_0 = i_0, X_1 = i_1\} &= P\{X_1 = i_1 | X_0 = i_0\} P\{X_0 = i_0\} \\
&= p_{i_0 i_1}\ a_{i_0} \\
P\{X_0 = i_0, X_1 = i_1, X_2 = i_2\} &= P\{X_0 = i_0\} \cdot P\{X_1 = i_1 | X_0 = i_0\} \\
&\cdot P\{X_2 = i_2 | X_1 = i_1\} \\
&= a_{i_0} p_{i_0 i_1}\ p_{i_1 i_2}
\end{aligned}
$$

7.2. Examples

Example: Two States

Suppose a person can be in one of two states — "healthy" or "sick". Let $\{X_n\}\ n = 0, 1, \ldots$ refer to the state at time n where

$$X_n = \begin{cases} 1 & \text{if healthy} \\ 0 & \text{if sick} \end{cases}$$

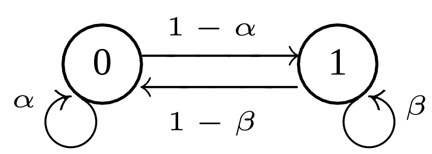
Define

$$P\{X_{n+1} = 0 | X_n = 0\} = \alpha$$
$$P\{X_{n+1} = 1 | X_n = 1\} = \beta$$

Transition Matrix

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

Transition Diagram



Ex. Independent Events

Let $\{X_n\}$ be iid with

$$P\{X_n = k\} = p_k \quad \text{for} \quad k = 0, 1, \ldots$$

and let the state space be $S = \{0, 1, 2, \ldots\}$

$$p_{jk} = P\{X_{n+1} = k | X_n = j\} = P\{X_{n+1} = k\} = p_k$$

$$P = \begin{bmatrix} p_0 & p_1 & p_2 & \ldots \\ p_0 & p_1 & p_2 & \ldots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

Example: Random Walk on Non-negative Real Line

Define $\{Z_n\}$ to be iid with $p_k = P\{Z_n = k\}$ for $k = 0, 1, 2, \ldots$

Define $X_0 = 0,\ X_n = \sum_{k=1}^{n} Z_k$

Then $\{X_n\}$ is a Markov Chain with state space $S = \{0, 1, 2, \ldots\}$;

$$P\{X_{n+1} = j | X_n = i\} = P\{Z_{n+1} = j - i\} = p_{j-i}$$

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccc} \underline{0} & \underline{1} & \underline{2} & \underline{3} & \cdots \end{array} \\ \begin{bmatrix} p_0 & p_1 & p_2 & p_3 & \cdots \\ 0 & p_0 & p_1 & p_2 & \cdots \\ 0 & 0 & p_0 & p_1 & \cdots \\ 0 & 0 & 0 & p_0 & \cdots \\ \vdots & & & & \end{bmatrix} \end{array}$$

Example: Random Walk (one step at a time)

$$P\{X_{n+1} = i+1 | X_n = i\} = p_i, P\{X_{n+1} = i+1 | X_n = j\} = 0 \text{ for } j \neq i$$

$$P\{X_{n+1} = i-1 | X_n = i\} = q_i, P\{X_{n+1} = i-1 | X_n = j\} = 0 \text{ for } j \neq i$$

$$P\{X_{n+1} = i \mid X_n = i\} = r_i = 1 - p_i - q_i$$

State Space: $S = \{0, 1, 2, \dots\}$

(i.) $q_0 = 0$ means that state 0 is reflecting barrier.

(ii.) If $r_0 = 1$, then once in state 0 it can never leave.

(iii.) If $p_N = 0 \Rightarrow S = \{0, 1, 2, \dots, N\}$

(iv.) If $p_N = 0$ and $r_N = 1 \Rightarrow N$ is absorbing ($r_N = 0$, N is reflecting barrier.)

Example: Gambler's Ruin

Gamblers: A, B have a total of N dollars

Game: Toss Coin

If $H \Rightarrow A$ receives \$1 from B

$T \Rightarrow B$ receives \$1 from A

$$P(H) = p, \quad P(T) = q = 1 - p$$

$X_n =$ Amount of money A has after n plays

$$P\{X_{n+1} = X_n + 1 | X_n\} = p$$

$$P\{X_{n+1} = X_n - 1 | X_n\} = q$$

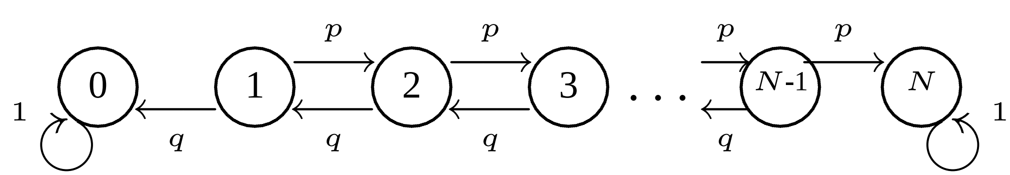
.....Game ends if $X_n = 0$ or $X_n = N$

State space= $\{0, 1, 2, \dots, N\}$

$$X_n \quad \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ N-1 \\ N \end{array} \overset{X_{n+1}}{\begin{array}{cccccccc} \underline{0} & \underline{1} & \underline{2} & \underline{3} & \cdots & \underline{N-2} & \underline{N-1} & \underline{N} \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 0 & q & 0 & p & \cdots & 0 & 0 & 0 \\ & & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{array}}$$

Transition Diagram for

Gambler's Ruin



Example: Urn Models (Ehrenfest Urn Model)

Two urns: A, B each containing N balls (Balls may be red or white).

Experiment consists of picking one ball at a time from each urn at random and placing them in the opposite urn.

$X_n =$ no. of white balls in urn A after n repetitions. Assume $X_0 = N$ (all white balls in A).

If $X_n = i \Rightarrow$ i white and $N - i$ red in A

i red and $N - i$ white in B

$$P\{X_{n+1} = i + 1 | X_n = i\} = P\{\text{white ball from } B \text{ and red ball from } A\}$$

$$= \left(1 - \frac{i}{N}\right)^2 = p_{i,i+1} \quad i \neq 0, N$$

$$
\begin{aligned}
P\{X_{n+1} = i-1 | X_n = i\} &= P\{\text{white from } A \text{ and red from } B\} \\
&= \left(\frac{i}{N}\right)^2 = p_{i,i-1}
\end{aligned}
$$

$$
\begin{aligned}
P\{X_{n+1} = i | X_n = i\} &= P\{\text{white from } A \text{ and } B\} \\
&\quad + P\{\text{Red from } A \text{ and } B\} \\
&= 2\left(\frac{i}{N}\right)\left(1 - \frac{i}{N}\right) = p_{ii}
\end{aligned}
$$

Example: Branching Process

$X_n =$ no. of individuals in n^{th} generation beginning with $X_0 = 1$(1 individual)

$Y_{i,n} =$ no. of offspring of the i^{th} person in the n^{th} generation

$$X_{n+1} = Y_{1,n} + Y_{2,n} + \ldots + Y_{X_n,n} = \sum_{i=1}^{X_n} Y_{i,n}$$

Assume $\{Y_{i,n}\}$ are iid random variables.

$$p_{ij} = P\{X_{n+1} = j | X_n = i\} = P\{\textstyle\sum_{i=1}^{X_n} Y_{i,n} = j | X_n = i\}$$

$$= P\{\textstyle\sum_{r=1}^{i} Y_{r,n} = j\}$$

Process:$\{X_n\}$ is called a branching process

How long does it take for a family to become extinct?

What is distribution of size in the n^{th} generation?

7.3. Marginal Distribution of X_n

Define
$$a_j^{(n)} = P\{X_n = j\} = \sum_{i\in S} P\{X_n = j|X_0 = i\}P\{X_0 = i\}$$
$$= \sum_{i\in S} P\{X_n = j|X_0 = i\}a_i$$

$p_{ij}^{(n)}$ = Prob. of going from $i \to j$ in n steps

$p_{ij}^{(n)}$ = n-step transition probabilities

Th. Chapman-Kolmogorov Equations

$$p_{ij}^{(n)} = \sum_{r\in S} p_{ir}^{(k)}\, p_{rj}^{(n-k)}$$ Chapman-Kolmogorov Equations

where k is a fixed integer $0 \le k \le n$

Th. $P^{(n)} = (p_{ij}^{(n)}) = P^n$

Proof. $P\{X_0 = j | X_0 = i\} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$

$\Rightarrow$ $P^0 = I$. Also $P^1 = P$. Assume theorem is true for $n = k$. We will show it is true for $n = k + 1$.

$$P^{(k+1)} = P^{(k)} P = P^k P = P^{k+1}$$

Th. $a^{(n)} =$ row vector of $a_j^{(n)} = (a_1^{(n)}, a_2^{(n)}, \dots)$

$a^{(n)} = aP^n$

Proof. $a^{(n)} = a^{(0)} P^{(n)} = aP^n$

Urn Sampling (Continuation)

$$E(X_n|X_0) = \sum_{i=0}^{X_0} iP\{X_n = i|X_0\}$$

Expected number of white balls in urn A with n draws given X_0 = no. of white balls in A at start.

$$= (0, 0, \ldots, 1)P^n \begin{bmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ X_0 \end{bmatrix}$$

Suppose $X_0 = 10$

n	$E(X_n \mid X_0 = 10)$	n	$E(X_n \mid X_0 = 10)$
2	8.2	12	5.3
4	7.0	14	5.2
6	6.3	16	5.14
8	5.8	18	5.09
10	5.5	20	5.06

Ex. Branching Process (Continuation)

$$m_n = E(X_n), \;\; \sigma_n^2 = Var X_n, \;\; m = E(Y_{i,n}), \;\; \sigma^2 = V(Y_i, n)$$

$$m_n = E(X_n) = E\left(\sum_{i=1}^{X_{n-1}} Y_{i,n-1}\right) = mE(X_{n-1})$$

$$\Rightarrow \qquad m_n = m\, m_{n-1}$$

$$\boxed{m_n = m^n}, \quad m = E(Y_{i,n})$$

$$Var(X_n|X_{n-1}) = Var\left(\sum_{i=1}^{X_{n-1}} Y_{i,n}\right) = \sigma^2 X_{n-1}$$

Recall $VarZ = E_Y Var(Z|Y) + Var_Y E(Z|Y)$

In our example $Z = X_n, \;\; Y = X_{n-1}$

$$Var(X_n|X_{n-1}) = Var(\sum_1^{X_{n-1}} Y_{i,n-1}|X_{n-1}) = X_{n-1}\sigma^2, \text{ if } X_{n-1} \text{ fixed}$$

$$
\begin{aligned}
E(X_n|X_{n-1}) &= E(\textstyle\sum_1^{X_{n-1}} Y_{i,n-1}|X_{n-1}) = X_{n-1}m \\
\therefore\ VarX_n &= \sigma^2 E(X_{n-1}) + Var(X_{n-1}m) \\
\sigma_n^2 &= \sigma^2 m_{n-1} + m^2 VarX_{n-1}
\end{aligned}
$$

$$\boxed{\sigma_n^2 = \sigma^2 m_{n-1} + m^2 \sigma_{n-1}^2}$$

$$\sigma_n^2 = \sigma^2 m_{n-1} + m^2 \sigma_{n-1}^2$$

$$m_n = m^n$$

<u>Case 1:</u> $m = 1\ (\sigma_0^2 = 0)$

$$\sigma_n^2 = \sigma^2 + \sigma_{n-1}^2$$

$$\Rightarrow \quad \sigma_1^2 = \sigma^2,\ \ \sigma_2^2 = 2\sigma^2,\ \ \sigma_3^2 = 3\sigma^2$$

$$\boxed{\sigma_n^2 = n\sigma^2} \qquad \text{if } m = 1$$

Case 2: $m \neq 1$

$$\sigma_n^2 = \sigma^2 m^{n-1} + m^2 \sigma_{n-1}^2$$

$$\sigma_1^2 = \sigma^2 \quad (\sigma_0^2 = 0)$$

$$\sigma_2^2 = \sigma^2 m + m^2 \sigma_1^2 = \sigma^2 m \left[\frac{m^2 - 1}{m - 1}\right]$$

$$\sigma_3^2 = \sigma^2 m^2 + m^2 \sigma_2^2 = \sigma^2 m^2 + m^2 \left[\sigma^2 m \left(\frac{m^2 - 1}{m - 1}\right)\right]$$

$$= \sigma^2 m^2 \left[\frac{m^3 - 1}{m - 1}\right]$$

$$\vdots \qquad \vdots$$

$$\boxed{\sigma_n^2 = \sigma^2 m^{n-1} \left[\frac{m^n - 1}{m - 1}\right] \quad m \neq 1}$$

Use of Generating Functions

$$
\begin{aligned}
G(z) &= \sum_{n=1}^{\infty} \sigma_n^2 z^n \quad (\sigma_0^2 = 0) \\
\sigma_n^2 &= \sigma^2 m^{n-1} + m^2 \sigma_{n-1}^2 \\
\sum_{1}^{\infty} \sigma_n^2 z^n &= \sigma^2 \sum_{1}^{\infty} m^{n-1} z^n + m^2 \sum_{n=1}^{\infty} \sigma_{n-1}^2 z^n \\
G(z) &= \sigma^2 z \sum_{n=1}^{\infty} (mz)^{n-1} + m^2 z \sum_{n=1}^{\infty} \sigma_{n-1}^2 z^{n-1} \\
G(z) &= \sigma^2 z \frac{1}{1 - mz} + m^2 z G(z) \\
G(z)[1 - m^2 z] &= \sigma^2 z / (1 - mz) \\
G(z) &= \sigma^2 z / (1 - m^2 z)(1 - mz)
\end{aligned}
$$

$$
\begin{aligned}
G(z) &= \sigma^2 z/(1-m^2 z)(1-mz) \\
&= \sigma^2 z \left\{ \sum_{r=0}^{\infty} (m^2 z)^r \sum_{s=0}^{\infty} (mz)^s \right\} \\
&= \sigma^2 z \left\{ \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} m^{2r+s} z^{r+s} \right\}, \quad n = r+s \quad 0 \le r \le n \\
&= \sigma^2 z \sum_{n=0}^{\infty} z^n m^n \sum_{r=0}^{n} m^r = \sigma^2 z \sum_{n=0}^{\infty} z^n m^n \left(\frac{1-m^{n+1}}{1-m} \right) \\
&= \sigma^2 \sum_{0}^{\infty} z^{n+1} m^n \left(\frac{1-m^{n+1}}{1-m} \right) \\
\Rightarrow \quad \sigma^2_{n+1} &= \sigma^2 m^n \left(\tfrac{m^{n+1}-1}{m-1} \right) \quad \text{or} \quad \sigma^2_n = \sigma^2 m^{n-1} \left(\tfrac{m^n - 1}{m-1} \right) \\
m_n &= m^n
\end{aligned}
$$

If $m > 1, \;\; m_n \to \infty$ as $n \to \infty$

If $m < 1, \;\; m_n \to 0$ as $n \to \infty$

If $m = 1, \;\; m_n = m$ always

Application: Nuclear Reactors

A neutron (0^{th} generation) is introduced into a fissionable material. If it hits a nucleus it will produce a random number of new neutrons (1^{st} generation). This process continues as each new neutron behaves like the original neutron.

$$X_n = \text{ No. of neutrons after } n \text{ collisions}$$

$$\boxed{m_n = m^n}$$

If $m > 1$, each neutron produces on average more than one neutron and reaction is explosive—(nuclear explosion or meltdown).

If $m < 1$, reaction eventually dies out.

In nuclear power station, $m > 1$ to reach “hot stage”. Once hot, moderator rods are inserted to remove neutrons and reduce m. Hence reactor is controlled. The moderator rods are continually removed and inserted to keep temperature in a given range. (Heat is converted to electricity).

Application: Family Names

Consider only male offspring who will carry family name. If $m < 1$, family name will eventually die out as $m^n \to 0$. Males in historical times would keep marrying until a wife could produce a male heir.
i.e. $P\{X_n \geq 1\} = 1 \Rightarrow m \geq 1$.

7.4 Appendix: Notes on Matrices: I

Let A : $n \times n$ matrix

x_i : $n \times 1$ vector

Eigenvalues:

$|A - \lambda I| = 0$ Polynomial in λ of degree n. The eigenvalues $\lambda_1, \ldots, \lambda_n$ are the zeros of the polynomial.

Eigenvectors

If $Ax_i = \lambda_i x_i \quad i = 1, \ldots, n$ then $x_i(n \times 1)$ are the right eigenvectors associated with λ_i.

If $y_i' A = \lambda_i y_i' \quad i = 1, \ldots, n$ then $y_i(n \times 1)$ are the left eigenvectors associated with λ_i.

$$\Rightarrow x_i' y_j = 0, \quad i \neq j$$

Proof: $Ax_i = \lambda_i x_i, \;\; y_j' A x_i = \lambda_j y_j' x_i = \lambda_i y_j' x_i$

If $y_j' x_i \neq 0$, then $\lambda_i = \lambda_j$ which is false. Hence $y_j' x_i = 0$.

Scale x_i, and y_i so that $x_i' y_i = 1$

Define

$$
\begin{aligned}
X^{n \times n} &= [x_1, x_2, \ldots, x_n] \\
Y^{n \times n} &= \begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{bmatrix}
\end{aligned}
$$

Therefore $AX = XD$ and $YA = DY$ where $D =$diag$(\lambda_1, \ldots, \lambda_n)$. We can write $A = XDX^{-1} = Y^{-1}DY$. Hence $X = Y^{-1}$.

Since

$$
\begin{aligned}
A &= XDX^{-1} \\
A^2 &= XDX^{-1}XDX^{-1} = XD^2X^{-1} \\
A^m &= XD^mX^{-1}, \;\; D^m = \text{diag}\,(\lambda_1^m, \ldots, \lambda_n^m)
\end{aligned}
$$

Idempotent Decomposition $\boxed{A^m = \sum_{i=1}^{n} \lambda_i^m x_i y_i' = \sum_{i=1}^{n} \lambda_i^m E_i}$

$$E_i = x_i y_i' \text{ and } E_i^2 = E_i, \;\; E_i E_j = 0 \;\; i \neq j$$

If A is stochastic $\underline{1}'A = \underline{1}'$ (columns add to unity), then $\lambda = 1$ is the largest eigenvalue.

$$P = \sum_1^n \lambda_i E_i, \qquad P^m = \sum_1^n \lambda_i^m E_i$$

as $m \to \infty$, $\lim_{m \to \infty} P^m = E_1 = y_1 x_1'$