MARKOV CHAINS WITH APPLICATIONS

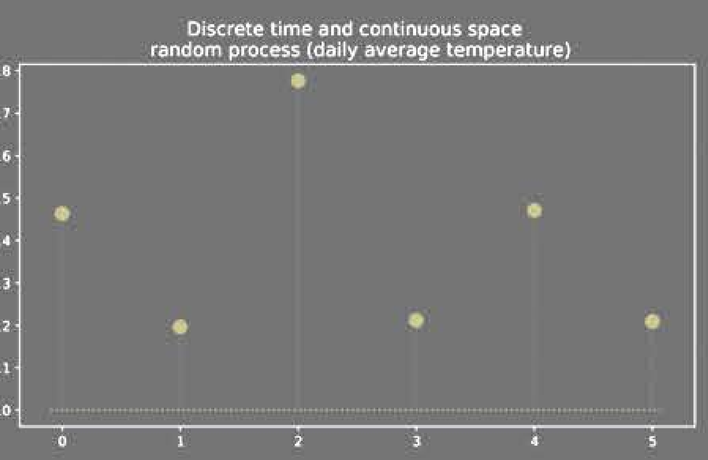
Işınsu ÇELEBİOĞLU

Supervisor: Inst. Dr. Barış ÇİÇEK

INTRODUCTION

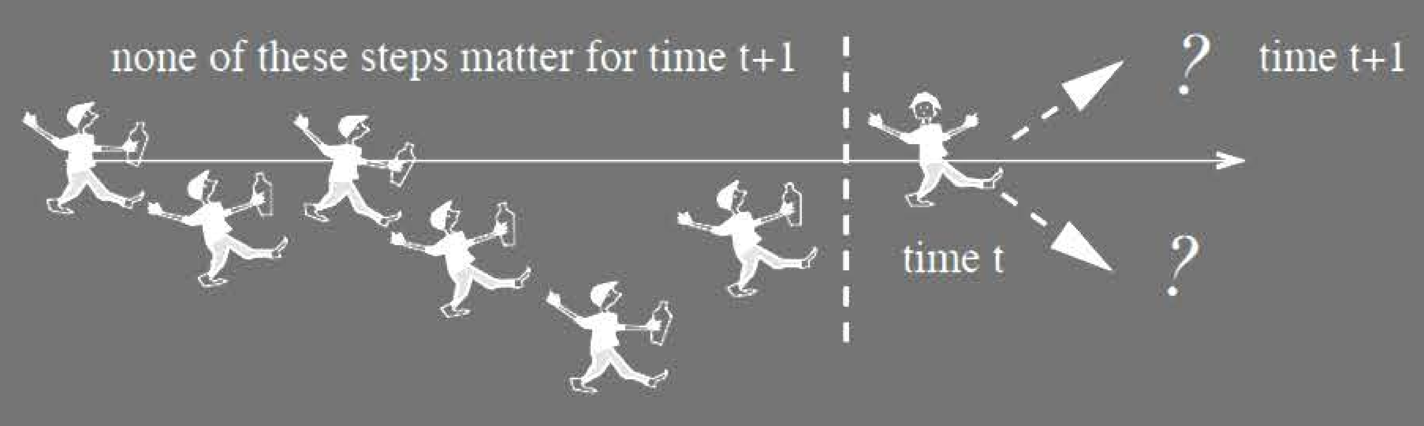
What is the Stochastic Processes?

Definition: A **stochastic process** is any process that describes the evolution over time of a random phenomenon. At each time $t \in [0,\infty)$ the system is in one state X_t, taken from a set S, the state space. One often writes such a process as $X = X_t : t \in [0,\infty)$



Random Walk

A **random walk** is a mathematical object, known as a stochastic or random processes, that describes a path that consists of a succession of random steps on some mathematical space. The repeating step is a move of one unit, left or right at random. The sum of the first t steps gives the position at time t.



What is the Markov Chain?

The Markov chain is the process $X_0, X_1, X_2, \ldots$.

Definition: The **state** of a Markov chain at time t is the **value of** X_t
For example, if $X_t = 6$, we say **the process is in state** 6 **at time** t

Definition: The **state space** of a Markov chain, S, is the set of values that each X_t can take.
For example, $S = \{1,2,3,4,5,6,7\}$.

Definition: A **trajectory** of a Markov chain is a particular set of values for $X_0, X_1, X_2, \ldots$.
For example, if $X_0 = 1$, $X_1 = 5$, and $X_2 = 6$, then the trajectory up to time $t = 2$ is $1, 5, 6$.

The Markov Property

A discrete time and discrete state space stochastic process is Markovian if and only if the conditional probabilities do not depend on $(X_0, \ldots, X_n)$ in full, but only on the most recent state X_n:

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

The likelihood of going to any next state at time $n+1$ depends only on the state we find ourselves in at time n. The system is said to have no memory.

Definition: Let $\{X_0, X_1, X_2, \ldots\}$ be a sequence of discrete random variables. Then $\{X_0, X_1, X_2, \ldots\}$ is a **Markov chain** if **it satisfies the Markov property:**

$$P(X_{t+1} = s|X_t = s_t, \ldots, X_0 = s_0) = P(X_{t+1} = s|X_t = s_t)$$

for all $t = 1,2,3,\ldots$ and **for all states** $s_0, s_1, \ldots, s_t, s$.

What Is The Transition Matrix and How We Can Get It?

The matrix describing the Markov chain is called the transition matrix. It is the most important tool for analyzing Markov chains.

MARKOV CHAINS

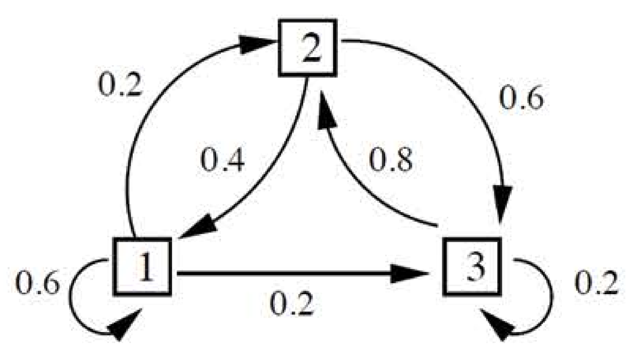
Transition Matrix

The transition matrix P must list all possible states in the state space S
P is a square matrix $(N \times N)$, because X_{t+1} and X_t both take values in the same state space S (of size N).
The rows of P should each sum to 1:

$$\sum_{j=1}^{N}(p_{ij}) = \sum_{j=1}^{N}(P(X_{t+1} = j|X_t = i)) = \sum_{j=1}^{N}(P_{X_t=i}(X_{t+1} = j))$$

This simply states that X_{t+1} must take one of the listed values.
The columns of P do not in general sum to 1.

Example: Let X_t be Purpose-flea's state at time t. $(t = 0, 1, ..)$



Find the transation matrix, P.

$$P = \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.4 & 0 & 0.6 \\ 0 & 0.8 & 0.2 \end{pmatrix}$$

Definition: Let $\{X_0, X_1, X_2, \ldots\}$ be a Markov chain with state space S where S has size N (possibly infinite). The **transition probabilities** of the Markov chains are

$$p_{ij} = P(X_{t+1} = j|X_t = i) \text{ for } i,j \in S,\ t = 0,1,2,\ldots$$

Distribution of X_t

Definition: Let $\{X_0, X_1, X_2, \ldots\}$ be a Markov chain with state space $S = \{1,2,\ldots,N\}$. Now each X_t is a random variable, so it has a probability distribution. We can write the probability distribution of X_t as an $N \times 1$ vector.

Notation: We will write $X_0 \sim \pi^T$ to denote that the row vector of probabilities is given by the row vector π^T.

Probability Distribution of X_1

Use the Partition Rule, conditioning on X_0:

$$\begin{aligned} P(X_1 = j) &= \sum_{i=1}^{N} P(X_1 = j|X_0 = i)P(X_0 = i) \\ &= \sum_{i=1}^{N} p_{ij}\pi_i \text{ by definitions} \\ &= \sum_{i=1}^{N} \pi_i p_{ij} \\ &= (\pi^T P)_j. \end{aligned}$$

Probability Distribution of X_2

Use the Partition Rule as before, conditioning again on X_0:

$$P(X_2 = j) = \sum_{i=1}^{N} P(X_2 = j|X_0 = i)P(X_0 = i) = \sum_{i=1}^{N}(P^2)_{ij}\pi_i = (\pi^T P^2)$$

The row vector $\pi^T P^2$ is therefore the probability distribution of X_2:

$$\begin{gathered} X_0 \sim \pi^T \\ X_1 \sim \pi^T P \\ X_2 \sim \pi^T P^2 \\ \vdots \\ X_t \sim \pi^T P^t \end{gathered}$$

Theorem: Let $\{X_0, X_1, X_2, \ldots\}$ be a Markov chain with $N \times N$ transition matrix P. If the probability distribution of X_0 is given by the $1 \times N$ row vector π^T, then the probability distribution of X_t, is given by the $1 \times N$ row vector $\pi^T P^t$. That is,

$$X_0 \sim \pi^T \Rightarrow X_t \sim \pi^T P^t$$

Equilibrium

If $\{X_0, X_1, X_2, \ldots\}$ is a Markov chain transition matrix P, then

$$X_0 \sim \pi^T \Rightarrow X_{t+1} \sim \pi^T P$$

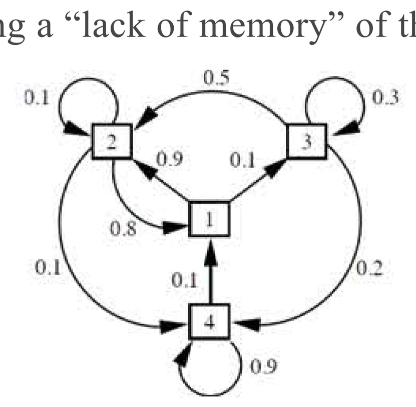
If $\pi^T P = \pi^T$, we say that the distribution π^T is an equilibrium distribution.

Theorem: Let $\{X_0, X_1, \ldots\}$ be a Markov chain with transition matrix P. Suppose that π^T is an equilibrium distribution for the chain. If $X_t \sim \pi^T$ for any t, then $X_{t+r} \sim \pi^T$ for all $r \geq 0$.
Once a chain has hit an equilibrium distribution**, it stays there for ever.**

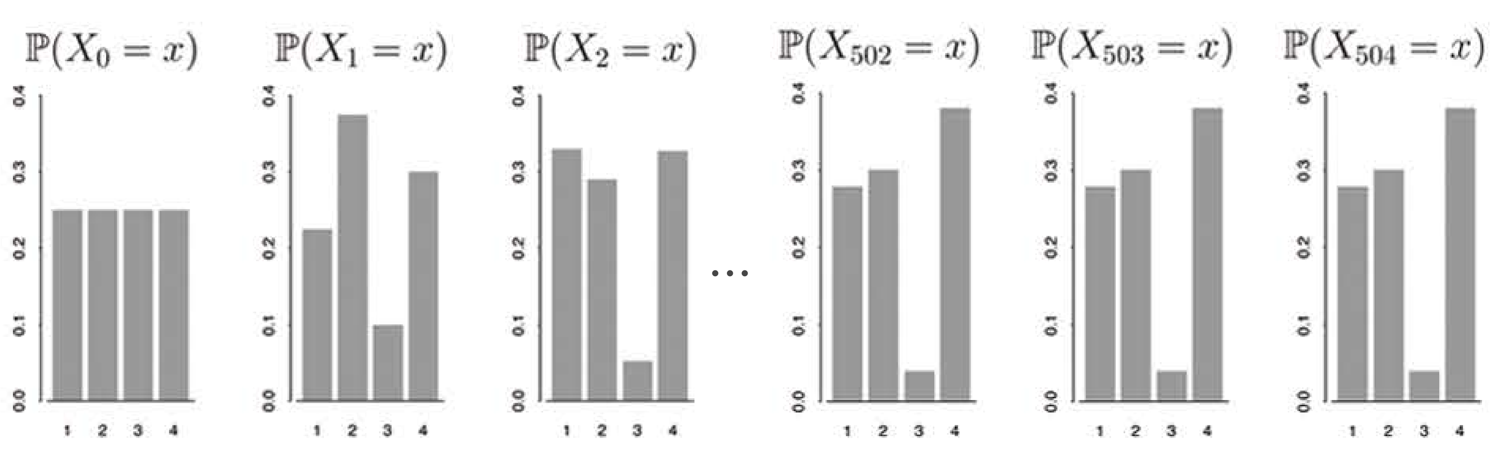
Long-term Behaviour

One thing that *could* happen over time is that the distribution $P(X_n = i)$ of the Markov chain could gradually settle down towards some "equilibrium" distribution. Further, perhaps that long-term equilibrium might not depend on the initial distribution, but the effects of the initial distribution might eventually almost disappear, exhibiting a "lack of memory" of the start of the process.

$$P = \begin{pmatrix} 0.0 & 0.9 & 0.1 & 0.0 \\ 0.8 & 0.1 & 0.0 & 0.1 \\ 0.0 & 0.5 & 0.3 & 0.2 \\ 0.1 & 0.0 & 0.0 & 0.9 \end{pmatrix}$$



Suppose we start at time 0 with $X_0 \sim (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$



$X_t \sim (0.28,\ 0.30,\ 0.04,\ 0.38)$ as $t \to \infty$.

References

[1] Rachel Fewster, "COURSE NOTES STATS 325 Stochastic Process", Department of Statistics University of Auckland, 2015.

[2] J. R. Norris, "Cambridge Series in Statistical and Probabilistic Mathematics",Cambridge University Press, 1997.