An Investigation into Stochastic Processes for Modelling Human-Generated Data

A 20-CP 3RD YEAR PROJECT

Author:
Tim Jones

Supervisor:
Dr. Gordon Ross

Rendered on Sunday 31$^{\text{st}}$ March, 2013

Contents

Chapter 1

Some Basic Stochastic Processes

1.1 Markov Chains

1.1.1 A Formal Definition

A stochastic process [1, p590] is a collection $\{X_t | t \in T\}$ (sometimes $X(t)$ for continuous T) of random variables. These collections may be indexed arbitrarily, but tend to be used to describe the evolution of some random series of events, using T to be some representation of either discrete or continuous time, for instance X_t may be the number of observed emissions from a radioactive source after t minutes or the lisence plate of the tth car to go past a speed camera.

For Markov Chains, we usually set $T = \{0, 1, 2, ...\} = \mathbb{N}$ or $T = \{x | x \geqslant 0\} = \mathbb{R}^+$. If $T = \mathbb{N}$ and each X_t is a discrete random variable over the set S, then we say that X obeys the Markov Property iff

$$\forall t \in T, s \in S \quad \mathbb{P}(X_{t+1} = s | X_0, X_1, ... X_t) = \mathbb{P}(X_{t+1} = s | X_t)[2]$$

Similarly, if $T = \mathbb{R}^+$, we have a continuous analogue to the Markov Property;

$$\forall \delta > 0, s \in S \quad \mathbb{P}(X(t + \delta) = s | \{X(\tau), \tau < t\}) = \mathbb{P}(X(t + \delta) = s | X_t)$$

If a stochastic process obeys the Markov Property, then we say that it is a Markov Chain. If a stochastic process obeys the continuous analogue to the Markov Property, then it is referred to as a Continuous Time Markov Process (CTMP). We refer to S as the state-space, each $s \in S$ is a state, and X_t represents the state at time t. We usually write the Markov Property as "given the present, the future is conditionally independent of the past" [2].

For CTMPs, we usually define the transition rate matrix Q such that

$$\forall i, j \in S, \delta > 0 \quad \mathbb{P}(X(t + \delta) = j | X(t) = i) = q_{ij} \delta + o(\delta)$$

Where $o(\delta)$ is some function such that $o(\delta) \to 0$ as $\delta \to 0$. This gives us that

$$\forall i, j \in S, \delta > 0 \quad \mathbb{P}(X(t + \delta) = j | X(t) = i) = (e^{\delta Q})_{ij}$$

That is, the i, j^{th} element of the matrix exponential of δQ.

1.1.2 An Intuitive Interpretation

Whilst the above defines Markovian Processes, it fails to describe them in any intuitive way. Before attempting to use them, it is important to be able to deal with both the continuous and discrete-time Markov Chains in an intuitive way. My personal preferred method is to use edge-weighted directed graphs[3]. Each state in S is given a node in the graph, and for each $i, j \in S$ the edge (i, j) is given weight p_{ij} in the discrete case and q_{ij} in the continuous case.

We can then say that a discrete Markov Chain will hop from node to node at each time step with probabilities defined by the weights of the edges between the current node and its neighbors. A CTMP is

slightly more complex; on arriving in a state i, all neighboring states j generate a random time $T_j \; Exp(q_{ij})$. We let $k = \arg\min_{j \in S} T_j$, and say that the CTMP will remain in state i for T_k units of time, then jump to state k.

1.1.3 The Poisson Process

The simplest form of CTMP is the homogeneous Poisson Process, where $S = \mathbb{N}$ and $\forall i \in S, q_{ii} = -\lambda, q_{i,i+1} = \lambda$. We call λ the rate of this process. If N(t) is a Poisson process, we then have that

$$\forall i \in \mathbb{N}, t, \delta \in \mathbb{R}^+, \mathbb{P}(N(t+\delta) = i+1 | N(t) = i) = \lambda\delta + o(\delta)$$

The graph for this process is similarly simple

We can also sacrifice the Markov Property to define an inhomogeneous poisson process. Everything remains from before, except rather than having a rate parameter $\lambda \in \mathbb{R}^+$, we have a rate function $\lambda : \mathbb{R}^+ \to \mathbb{R}^+$, where

$$\forall i \in \mathbb{N}, t, \delta \in \mathbb{R}^+, \mathbb{P}(N(t+\delta) = i+1 | N(t) = i) = \lambda(t)\delta + o(\delta)$$

Since this process is no longer Markovian, it makes little sense to attempt to graph it, however it remains a powerful tool.

1.2 Hidden Markov Models

The Hidden Markov Model is an extention of the Markov Chain used for fitting various stochastic models.

With discrete time and state space, rather than having a directly observable Markov Chain, we assume that the process we observe has an underlying unobservable Markov Chain. We then define an observation space O, and a probability distrubtion, p, on $S \times O$. Before jumping out of state i, the Markov Chain emits observation o with probability $p(i, o)$.

The traditional example is that of the "unfair casino". Imagine the dealer at the casino has two six-sided dice. Die one is fair, but die two is weighted such that it never shows a 6, and shows the numbers 1 to 5 with equal probabilities. The dealer will clandestinely swap the dice with probability 0.1 before rolling. Here, our state space is $\{1, 2\}$ and our observation space is $\{1, 2, 3, 4, 5, 6\}$. We can represent this as follows;

We can only observe the results of the dice rolls, but what we're interested in is the underlying model. Firstly, we want to know what die the dealer is likely to be using at any given time - if I've seen the dealer roll a 6, I know he was using the fair die then, so he will use the fair die again with probability 0.9, for example. Further to this, suppose we didn't know the above parameters. We also want to estimate them - how many dice does the casino have, how does the dealer swap them, how do these dice roll?

These questions wil be discussed later on.

Chapter 2

Potential Modelling Techniques

Bibliography

[1] The Development of Rigor in Mathematical Probability (1900-1950)
Joseph L. Doob
The American Mathematical Monthly, Vol. 103, No. 7 (Aug. - Sep., 1996), pp. 586-595
`http://www.jstor.org/stable/2974673`

[2] Weisstein, Eric W. "Markov Chain." From MathWorld–A Wolfram Web Resource.
`http://mathworld.wolfram.com/MarkovChain.html`

[3] Weisstein, Eric W. "Graph." From MathWorld–A Wolfram Web Resource.
`http://mathworld.wolfram.com/Graph.html`