Week 1

02407 Stochastic Processes
2021-8-30
BFN/bfn

Probability theory underlies modern statistics including machine learning. Besides this, probability theory has many applications in its own right, in areas as diverse as biology, health care, and operations management. A comprehensive knowledge of probability theory is necessary to properly understand statistical and machine learning models. For instance, reinforcement learning is based on the theory of Markov Decision processes.

We start with a short recap of elements from probability theory. Then we introduce the concept of a stochastic process and the first specific example, the discrete Markov chain. The process is evaluated on a discrete set, typically thought of as time epochs, and takes values in a discrete set, typically the integers. The range of the random variables is usually referred to as the state space. A Markov chain could be used as a model for the number of cars a car dealer has in stock at the beginning of each day.

Must read and nice to read

The definition and basic properties of Markov chains are introduced in Section 3.1. Higher order transition probabilities are defined in Section 3.2, while the important concept of first step analysis is dealt with in Section 3.4. A number of classical and generic Markov chain models are presented in sections 3 and 5.

Some important specific definitions and results

Markov property (3.1) p.79

$\Pr\{X_{n+1}=j|X_0=i_0,X_1=i_1,\ldots,X_n=i_n\}=\Pr\{X_{n+1}=j|X_n=i_n\}$

One step transition probabilities (3.2) p.79

$P_{ij}^{n,n+1}=\Pr\{X_{n+1}=j|X_n=i_n\} \qquad P_{ij}=\Pr\{X_{n+1}=j|X_n=i_n\} \qquad \boldsymbol{P}=\|P_{ij}\|$

joint distribution (3.8) p.80

$\Pr\{X_0=i_0,X_1=i_1,\ldots,X_n=i_n\}=p_{i_0}P_{i_0i_1}P_{i_1i_2}\cdots P_{i_{n-1}i_n}$

n-step transition probabilities (3.10) p.83

$P_{ij}^{(n)}=\Pr\{X_{n+m}=j|X_m=i\}$

Chapmann-Kolmogorov (3.11) p.83

$$P_{ij}^{(n)}=\sum_{k=0}^{\infty}P_{ik}P_{kj}^{(n-1)}$$

n step transition probability matrix (3.12) p.84

$\boldsymbol{P}^{(n)}=\boldsymbol{P}^n$

First step analysis, T time to absorption p.100

$$u_i=\Pr\{X_T=0|X_0=i\}=\sum_j \Pr\{X_T=0|X_0=i,X_1=j\}\cdot P_{ij}=\sum_j u_j\cdot P_{ij}$$

The Markov property is essential. Most of the theory of stochastic processes is constructed using this assumption. The transition probabilities are functions of time in the general version of (3.2). However, one usually assumes the second version with time-invariant transition probability. This is due to the very limited closed form results which can be obtained in the general case. The Chapmann-Kolmogorov equations are due to the Markov property and reappears in various ways in more general settings. Finally, probably the most advanced topic of the first week, first step analysis is a powerful technique which we will see reappear later in other forms.

Some typos

Page 89, first row of the transition matrix $\mathbf{P}$ in line 12 should have been -1 instead of -0.