Introducing Markov Chains Models to Undergraduates

Ann-Lee Wang
University of Malaya, Institute of Mathematical Sciences
50603 Kuala Lumpur, Malaysia
j2wang@umcsd.um.edu.my

1. Introduction

Learning Markov chains requires a variety of skills that are taught in probability theory and mathematics classes. A second year undergraduate possesses these skills to a greater or lesser degree. He/she finds that learning Markov chains involves the understanding of quite a number of new concepts and the applications of skills that he/she may or may not have being taught previously.

In general, most students do not find Markov chains models an easy topic. There are many obstacles and difficulties in the learning of Markov chains (Wang 2001).

In this paper, the introduction of discrete time and discrete state space Markov chains models with stationary transition probabilities to undergraduates who are in their second year in the university is discussed. In Sections 2, 3 and 4, attempts are made to give some general and specific strategies for the teaching of Markov chains models. A classroom strategy for working out tutorial questions is described in Section 5. In Section 6, some views of students who have taken the course are given. Section 7 lists some concluding remarks.

2. General strategies

The models used in this course are mainly finite state space Markov chains. As a general rule, the subject matters taught need of necessity to be rather simple and brief.

In learning Markov chains models, students need to be able to recall some of their knowledge of probability and mathematics and to use their knowledge to work out problems in new and unfamiliar situations. Students usually need time to understand and assimilate new concepts that are inherent in these new situations. One way to ensure the understanding of these concepts is to teach each concept repeatedly but with some variation and in a slightly different context each time the concept is repeated. Tutorial questions are then used as a form of reinforcement of the concept learnt. In this way, when teaching a new concept or topic, one would: **R**ecall, **R**epeat and **R**einforce.

In addition, new concepts have to be taught in small doses. These small building blocks will make it easier for the students to grasp and retain new concepts.

Most students have difficulties in translating a word problem into probability statements that may lead them to consider Markov chains models. The number of word problems used in a course may be kept to a minimum by asking different questions of the same Markov chain model as the course proceeds. In so doing, one will also be teaching the various topics in small building blocks, always adding on to the existing building blocks.

When the needs arise, students are reminded of the expressions for the sum of geometric series and its variations. Techniques for dealing with recursive equations and difference equation are demonstrated as the need arise. Other mathematical techniques such as generating function methods are also taught in context. All topics are taught with as little use of nomenclature as possible. The idea is to explain the concepts without burdening the students with too many technical terms.

3. Visual displays

A Markov chain is a random process evolving in time in accordance with the transition probabilities of the Markov chain. Students have to be made aware of the time element in a Markov chain. Some pictorial representations or diagrams may be helpful to students. Only two visual displays will be discussed in this paper. These visual displays are sample path diagram and transition graph.

A sample path diagram is similar to a tree diagram that is usually taught in an introductory probability course. In this diagram, starting from an initial state all the possible sample paths of the Markov chain are drawn for a small value of the time parameter, n. Usually n is taken to be four or five. This sample path diagram displays the possible progression of the Markov chain for n steps starting from an initial state. This display may help to clarify to the students the dependent nature of the Markov chain. Students may use this sample path diagram to evaluate the probabilities of occurrence of a particular sample path. Some state and class properties may be apparent from the sample path diagrams. Single and longer string of sample path may be simulated and shown to the students.

Another useful visual display is transition graph. This graph is based on the one-step transition probabilities of the Markov chain that are usually displayed as a transition probability matrix. In this type of graph, states are numbered and each state is written, with a small circle around it, on a piece of paper. The states are spread out on the paper so that they are distinguishable and lines may be drawn to link them. Positive transition probability between two states is indicated by a line joining the two states. An arrow is used to indicate the direction of positive transition probability. Once all the lines and arrows are drawn, students may use the graph to help them to partition the state space into equivalence classes, to identify absorbing states, transient states and the periodic behaviour of certain state or equivalence class.

4. Types of examples and problems

The examples and problems used in this course have to be simple and as unambiguous as possible. Simple coin tossing examples and simple coloured balls in boxes examples (i.e. examples involving the withdrawing, replacing and inter-changing of the balls in various manners) are usually devoid of ambiguity and are on the whole non-emotional, non-political and non-cultural. These examples will help students understand new concepts without the extra difficulties of translating word problems into Markov chain models. Unfortunately, students find this type of examples boring. They consider simplified versions of real-world random phenomena interesting. Simple examples, boring or otherwise, may be taken from Taylor and Karlin (1994), Bhat (1984), Chung (1978) and Karlin and Taylor (1975).

5. Tutorial classroom strategy

At the beginning of the course, students are asked to form themselves into groups consisting of three to five students. Those students who cannot form a group are arbitrarily grouped. Each group is responsible for writing out a brief solution to a tutorial question on the blackboard. Students from other groups will then be questioned as to whether the solution is correct, or whether there are other ways of solving the question. Suggestions and further questions from the students are encouraged. When no students can provide the correct solution, hints are given to the students. The group responsible for the question will then have to work on the question again and to write out the solution on the blackboard the following week.

This strategy is used to promote active participation by the students in the course, to create motivation and interest in the subject and to ensure the students actually read and study lecture notes.

Most of the tutorial questions are written in such a way that the method of solution is the same or similar to examples given in lectures. Usually some slight variations are written into the tutorial questions to make them less of a straightforward application of methods or materials taught in the lectures. Hints are sometimes stated at the end of a question. However, my impression is that students on the whole ignore the hints provided. So instead of writing hints at the end of a question, hints are incorporated into the questions. This usually means breaking down each question into a few small parts whose solutions lead to the required answer. Not many students appreciate this method either.

The fundamental matrix of a Markov chain is not taught in this course. Sets of simultaneous equations that have to be solved in tutorial questions are such that they can be solved easily by hand. Students are, however, expected to be able to find eigenvalues and eigenvectors. Matrices with easily evaluated eigenvalues and eigenvectors are used in the course.

Problems that require computer packages and softwares for the enumeration of certain pertinent values are not be used in this course.

6. **Students' perspectives**

The principal concern of my students is to sign onto an easy course so that they will be able to obtain very good grades. Most of them were very disappointed that they have to work so hard for this course.

Good students find the parts dealing with conditional probabilities and evaluation of transition probabilities easy to follow. As a student put it, she is still "on familiar ground". Once the lectures turn to topics like recurrence, first passage probability, absorption probabilities and the use of generating function methods, these students find that they have to study each and every single line of the lectures carefully before they can grasp the material taught.

7. **Concluding remarks**

Learning Markov chains models for the first time tax students' abilities to grasp new concepts in probabilities and to apply techniques that they have learned (or not yet learned) in other probability and mathematics courses. There is a need to make an introductory course in Markov chains as simple as possible. Time has to be allocated to reviews the necessary probability theory and mathematical techniques used in the course. Exercises that are numerically tractable without the use of computer software should be use.

Last but not least, one should remind students constantly to think and also to practise, practise and do more practice in solving problems on their own.

REFERENCES

Bhat, U. N. (1984) *Elements of applied stochastic processes, Second edition*. New York: Wiley.

Chung, K. L. (1975) *Elementary probability theory with stochastic processes*, New York: Springer-Verlag.

Karlin, S. and Taylor, H. M. (1975) *A first course in stochastic processes*, Second Edition, San Diego: Academic Press.

Taylor, H. M. and Karlin, S. (1994) *An introduction to stochastic modeling*, Boston: Academic Press.

Wang, A.-L. (2001) How much can be taught about stochastic processes and to whom? In C. Batanero (Ed.), *Training researchers in the use of statistics*, 73-85, Granada: International Association for Statistical Education and International Statistical Institute.

SUMMARY

Learning Markov chains models requires some knowledge of probability theory, calculus, matrix algebra and a general level of mathematical maturity. While trying to understand Markov chains models, students usually encounter many obstacles and difficulties. In this paper, some suggestions regarding the teaching of introductory Markov chains models to undergraduates are discussed.