MARCEL F. NEUTS

MATRIX-GEOMETRIC SOLUTIONS IN STOCHASTIC MODELS

AN ALGORITHMIC APPROACH

NEUTS

MATRIX-GEOMETRIC SOLUTIONS IN STOCHASTIC MODELS

JOHNS HOPKINS

Matrix-Geometric Solutions in Stochastic Models

An Algorithmic Approach

Marcel F. Neuts

The Johns Hopkins University Press
Baltimore and London

Printed in the United States of America

The Johns Hopkins University Press, Baltimore, Maryland 21218
The Johns Hopkins Press Ltd., London

Library of Congress Cataloging in Publication Data

NEUTS, MARCEL F
Matrix-geometric solutions in stochastic models.

(Johns Hopkins series in the mathematical sciences; 2)
Bibliography: pp. 310–28
Includes index.

1. Markov processes. 2. Queueing theory.
3. Matrices. I. Title. II. Series.
QA274.7.N48 519.2′33 80-8872

ISBN 0-8018-2560-1

Matrix-Geometric Solutions in Stochastic Models

Johns Hopkins Series in the Mathematical Sciences

1. *Nonparametric Probability Density Estimation,* Richard A. Tapia and James R. Thompson
2. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach,* Marcel F. Neuts

It is unworthy of excellent men to lose hours like slaves in the labor of calculation which could safely be relegated to anyone else if machines were used.

Gottfried Wilhelm Leibniz

Preface

This book is an outgrowth of a series of lectures I gave at The Johns Hopkins University, Baltimore, Maryland, from July 20 to July 24, 1979. It contains a variety of results on queues and other stochastic models with the unifying feature of ready algorithmic implementation. The material discussed in this book comes under the broader heading of *computational probability*, a subject area sufficiently young to require definition.

As I perceive it, computational probability is not primarily concerned with the algorithmic questions raised by the direct numerical computation of existing analytic solutions. Such questions are best considered within the framework of classical numerical analysis. It is the concern of the probabilist, however, to ensure that the solutions he obtains are in the best, most natural, form for numerical computation. The expression of this concern is of recent date. Before the era of the modern computer, much of the best effort in applied mathematics was aimed at obtaining insight into the behavior of formal models, while avoiding the drudgery of computation by primitive machinery. On the other hand, the early difficulty of computation has also allowed the development of a large number of formal solutions from which few, if any, qualitative conclusions may be drawn, and whose appropriateness for algorithmic implementation has not been seriously considered. There is, in fact, an attitude that still pervades most of the teaching and the research literature on applied probability today and that does not view algorithmic implementation as an integral, challenging part of the solution process. We view this attitude as a legacy of history, but not as a constructive one.

We therefore define computational probability as the study of stochastic models with a genuine added concern for algorithmic feasibility over a wide, realistic range of parameter values. We have imposed upon our own work and that of our students the requirement that careful and exhaustive computer studies be performed before research results are proposed as the solution to a given problem. The limitations on a specific algorithm are as real as are restrictive conditions on the validity of a theorem. They offer, as we shall see, the same challenge and stimulus to further research.

This self-imposed constraint has not been felt as a burden. It has, on the contrary, led us to seek alternatives to certain classical analytic methods that are often difficult and risky in their numerical implementation. Unifying structural properties of Markov chains have been identified, which, through purely probabilistic arguments, lead to highly stable numerical procedures. Rather than proceeding through purely formal manipulations, the analysis of many problems now runs parallel to the steps of the algorithms. The latter then acquire a significance, which is often of independent interest in the interpretation of numerical results. A number of specific algorithms and the interpretation of their intermediate steps are discussed in this book.

The reader who is acquainted with the literature on stochastic models may briefly miss in our treatment the prevalent transform solutions and the familiar arguments based on methods of complex analysis. The reason for this will soon be clear. Though useful tools in analytic investigations, these techniques may often be replaced by others that are better suited for numerical computation. For the models discussed in this book, these will be iterative matrix methods and classical systems of differential equations.

The oft-lamented *Laplacian curtain,* which covers the solution and hides the structural properties of many interesting stochastic models, is now effectively lifted, at least for broad classes of problems in the theory of queues. We feel that at this point the real difficulties of a detailed analysis of such models emerge. These are nearly all related to the vast increases in dimensionality which accompany even small increases in the complexity of the physical description of most models. The *curse of dimensionality,* so aptly named by R. Bellman in his book *Dynamic Programming* [13], also weighs heavily on applied probability.

About this matter, at least two things may be said. It has been stated that numerical procedures are always closely linked to the available computer technology and that matters of implementation will eventually be settled by those interested in specific answers. The implication that such issues need not preoccupy the theoretical investigator is usually unstated, yet very clear. The truism of the first statement hides the utter fallacy of the second. It is clear that with faster computers that are endowed with larger central memories we can handle problems of higher dimension and compute their solution in more detail and with greater ease. The dimensionality of the problems we would like to solve, however, far outruns even the most optimistic forecasts of the capabilities of future computers. There is also little evidence that persons working on specific applications can afford the time or the mathematical effort required to construct efficient and appropriate algorithms for their solutions. The algorithmic aspects of models usually lead to a complete rethinking of their mathematical structure and present challenges and technical difficulties on a par with those of the classical analytic procedures. They also lead to a more realistic view of what constitutes a meaningful and satisfactory solution. A case in point is the analysis of the $GI/PH/c$ queue, which is briefly discussed in Chapter 4. Our general theorems give a complete characterization of the stationary probability vector of this queue, but in terms of the solution of a nonlinear matrix equation involving matrices of order ν^c. This allows the computation of that vector for very small values of ν and c, but does not constitute in our view a meaningful and satisfactory solution to this problem. In this case, the limitations of the algorithm are brought about by the inherent complexity of the multi-server queue. The theoretical examination of these limitations will, we hope, shed light on qualitative results or approximate procedures that are capable of extending our study of systems with many servers and more versatile service time distributions.

The high dimensionality of most interesting problems offers the opportunity for investigations of a different nature. As in mathematical programming, many problems of practical significance have special structural properties that may be used to construct workable algorithms, even when the problem is inherently of high dimension. Where these models have been treated earlier, such special, simplifying structures have usually been obscured by the formalism of transform methods. In this book, there are many instances where I discuss how particular structural features can be utilized to simplify and expedite numerical computation. These examples are far from the final word on this subject. They should serve to illustrate the type of simplifying structure to look for in specific problems. Such examples further illustrate the need for the probabilist to be very closely involved with the algorithmic analysis of a problem. One may hardly expect simplifications, which require a thorough understanding of the problem, to be discovered by a programmer. Such an involvement is also necessary for another reason. Statements regarding the feasibility or limitations of algorithms have the same standing as theorems and should be subject to the same scrutiny and scientific critique. The profusion of trivial numerical examples and of unwarranted algorithmic claims in the "applied" literature on stochastic models suggests, however, that such standards are not yet widely held.

This book must not be regarded as a systematic treatment of all of computational probability, or even of its applications in the theory of queues. Insofar as it represents my work and that of my associates, it deals only with results obtained since 1975. A significant portion of its contents are of very recent date, and some useful algorithms are presented for the first time here. In order to keep its length within reasonable bounds, we have limited the book to the discussion of one class of structured Markov chains only, and to a broad selection of its applications. We have given detailed references to procedures for other types of Markov chains, whose structure we have found to be very useful for algorithmic purposes. The presentation of the book is also intentionally open-ended. Such important approaches as, for example, the use of interactive computation in the design of stochastic systems, are represented only by simple, but illustrative, examples. We are confident, however, that as the algorithmic approach gains wider currency, much more will be heard about this methodology.

The discussion in this book is also intentionally limited to the algorithmic aspects of the *stable versions in steady-state* of the stochastic models under consideration. This was done for several reasons. The mathematical structure of the models is particularly useful in the discussion of steady-state features. The computation of transient solutions is, even for simple models, a belabored task that may only rarely be simplified. It requires in the best instances the solution of very large systems of differential equations. Both the computational effort and the interpretation of numerical results of transient solutions depend crucially on the choice of the initial conditions. It is next to impossible to give a unified discussion of all types of transient behavior which may arise. In steady-state analysis, the consideration of initial conditions may, of course, be avoided. Provided it is carried out in sufficient detail, the steady-state solution sheds ample light on many behavioral features of the model. In such cases where transient solutions to, for example, a stable queue are desirable for a variety of initial conditions, it is still advisable to obtain detailed information on the steady-state distributions. In performing the belabored computations for the transient phase, the latter will then serve to indicate when the effect of the initial conditions has been attenuated to the point that further numerical solution of the time-dependent equations becomes uninformative.

The organization of this book is as follows. Chapter 1 contains a systematic discussion of the general properties of a class of Markov chains and processes that in the positive recurrent case have a matrix-geometric invariant probability vector. Chapter 2 is a self-contained treatment of the properties of phase type distributions. These distributions arise from a generalization of Erlang's method of stages in a form that is particularly well-suited for numerical computation. Chapter 3 deals with the case of block-tridiagonal transition probability matrices. For such matrices, which frequently arise in applications, much more detailed results may be obtained. Chapter 4 treats the *GI/PH/*1 queue and several of its variants and generalizations. A lengthy discussion of semi-Markovian arrival processes and their applications is given in section 4.2. Chapters 5 and 6 deal with an eclectic variety of models suggested by diverse applications. In these chapters, we concentrate primarily on the formulation and formalization of models and on the interpretation of numerical results.

In concluding this preface, I fulfill the pleasant task of acknowledging the support of several institutions and the cooperation of many persons, both in the research effort and in the writing effort that have led to this book.

Purdue University and, since 1976, the University of Delaware have provided me with an environment and an atmosphere conducive to learning and scholarly pursuit. The Air Force Office of Scientific Research, through grant numbers AFOSR-72-2350 and AFOSR-77-3236, and the National Science Foundation, through grant numbers APR 74-15256 and ENG-7908351, have generously supported my research and that of research associates and graduate students.

Lecture series presented during short-term visits to the Indian Institute of Management, Calcutta, and to Clemson University, Clemson, South Carolina, prior to the delivery of the Mathematical Sciences Lectures at The Johns Hopkins University in 1979 have helped me in finding greater unity and conciseness in much diverse material.

To Professor Moshe Barash of Purdue University go special thanks for broadening my views of applications in technology, notably in manufacturing processes. Dr. Guy Latouche of the Free University of Brussels has generously shared his insights into the problems of computer modeling. To Dr. V. Ramaswami of Drexel University, and to Professor Ohoe Kim of Towson State University, I am grateful for many constructive suggestions and comments on the mathematical results in this book. Dr. Alan F. Karr of The Johns Hopkins University has ably organized the lecture series given there in July, 1979.

D. M. Lucantoni, S. Chakravarthy, and S. Kumar, currently graduate students at the University of Delaware, have lightened our task through their enthusiasm for the subject and through a careful reading of the manuscript. The typescript was prepared by Karen L. Tanner, whose conscientious work is deeply appreciated.

Matrix-Geometric Solutions in Stochastic Models