

*Stochastic Mechanics*  
*Random Media*  
*Signal Processing*  
*and Image Synthesis*  
*Mathematical Economics and Finance*  
*Stochastic Optimization*  
*Stochastic Control*

**Applications of  
Mathematics**

*Stochastic Modelling  
and Applied Probability*

**34**

*Edited by* I. Karatzas  
M. Yor

*Advisory Board* P. Brémaud  
E. Carlen  
W. Fleming  
D. Geman  
G. Grimmett  
G. Papanicolaou  
J. Scheinkman

# Applications of Mathematics

---

- 1 Fleming/Rishel, **Deterministic and Stochastic Optimal Control** (1975)
- 2 Marchuk, **Methods of Numerical Mathematics**, Second Edition (1982)
- 3 Balakrishnan, **Applied Functional Analysis**, Second Edition (1981)
- 4 Borovkov, **Stochastic Processes in Queuing Theory** (1976)
- 5 Liptser/Shiryayev, **Statistics of Random Processes I: General Theory** (1977)
- 6 Liptser/Shiryayev, **Statistics of Random Processes II: Applications** (1978)
- 7 Vorob'ev, **Game Theory: Lectures for Economists and Systems Scientists** (1977)
- 8 Shiryayev, **Optimal Stopping Rules** (1978)
- 9 Ibragimov/Rozanov, **Gaussian Random Processes** (1978)
- 10 Wonham, **Linear Multivariable Control: A Geometric Approach**, Third Edition (1985)
- 11 Hida, **Brownian Motion** (1980)
- 12 Hestenes, **Conjugate Direction Methods in Optimization** (1980)
- 13 Kallianpur, **Stochastic Filtering Theory** (1980)
- 14 Krylov, **Controlled Diffusion Processes** (1980)
- 15 Prabhu, **Stochastic Storage Processes: Queues, Insurance Risk, and Dams** (1980)
- 16 Ibragimov/Has'minskii, **Statistical Estimation: Asymptotic Theory** (1981)
- 17 Cesari, **Optimization: Theory and Applications** (1982)
- 18 Elliott, **Stochastic Calculus and Applications** (1982)
- 19 Marchuk/Shaidourov, **Difference Methods and Their Extrapolations** (1983)
- 20 Hijab, **Stabilization of Control Systems** (1986)
- 21 Protter, **Stochastic Integration and Differential Equations** (1990)
- 22 Benveniste/Métivier/Priouret, **Adaptive Algorithms and Stochastic Approximations** (1990)
- 23 Kloeden/Platen, **Numerical Solution of Stochastic Differential Equations** (1992)
- 24 Kushner/Dupuis, **Numerical Methods for Stochastic Control Problems in Continuous Time** (1992)
- 25 Fleming/Soner, **Controlled Markov Processes and Viscosity Solutions** (1993)
- 26 Baccelli/Brémaud, **Elements of Queuing Theory** (1994)
- 27 Winkler, **Image Analysis, Random Fields and Dynamic Monte Carlo Methods** (1995)
- 28 Kalpazidou, **Cycle Representations of Markov Processes** (1995)
- 29 Elliott/Aggoun/Moore, **Hidden Markov Models: Estimation and Control** (1995)
- 30 Hernández-Lerma/Lasserre, **Discrete-Time Markov Control Processes** (1995)
- 31 Devroye/Györfi/Lugosi, **A Probabilistic Theory of Pattern Recognition** (1996)
- 32 Maitra/Sudderth, **Discrete Gambling and Stochastic Games** (1996)
- 33 Embrechts/Klüppelberg/Mikosch, **Modelling Extremal Events** (1997)
- 34 Duflo, **Random Iterative Models** (1997)

Marie Duflo

# Random Iterative Models

Translated by Stephen S. Wilson



Springer

Marie Duflo

Université de Marne-la-Vallée

Equipe d'Analyse et de Mathématiques Appliquées

2, rue de la Butte Verte,

93166 Noisy-Le-Grand Cedex, France

*Managing Editors*

I. Karatzas

Departments of Mathematics and Statistics, Columbia University

New York, NY 10027, USA

M. Yor

Laboratoire de Probabilités, Université Pierre et Marie Curie

4 Place Jussieu, Tour 56, F-75230 Paris Cedex, France

Title of the French original edition: *Méthodes récursives aléatoires*

Published by Masson, Paris 1990

Cover picture: From a report on Prediction of Electricity Consumption drawn up in 1993 for E.D.F by Misiti M., Misiti Y., Oppenheim G. and Poggi J.M.

Library of Congress Cataloging-in-Publication Data

Duflo, Marie.

[Méthodes récursives aléatoires. English]

Random iterative models / Marie Duflo ; translated by Stephen S. Wilson.

p. cm. -- (Applications of mathematics ; 34)

Includes bibliographical references (p. - ) and index.

ISBN 978-3-642-08175-0 ISBN 978-3-662-12880-0 (eBook)

DOI 10.1007/978-3-662-12880-0

1. Iterative methods (Mathematics) 2. Stochastic processes.

3. Adaptive control systems--Mathematical models. I. Title.

II. Series.

QA297.8.D8413 1997

003'.76'015114--dc21

96-45470

CIP

---

Mathematics Subject Classification (1991): 60F05/15, 60G42, 60J10/20, 62J02/05, 62L20, 62M05/20, 93E12/15/20/25

---

ISSN 0172-4568

ISBN 978-3-642-08175-0

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag Berlin Heidelberg GmbH.

Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1997

Originally published by Springer-Verlag Berlin Heidelberg New York in 1997

Softcover reprint of the hardcover 1st edition 1997

Typeset from the translator's L<sup>A</sup>T<sub>E</sub>X files with Springer-T<sub>E</sub>X style files

SPIN: 10070196

41/3143 - 5 4 3 2 1 0 - Printed on acid-free paper

## **Preface**

Be they random or non-random, iterative methods have progressively gained sway with the development of computer science and automatic control theory.

Thus, being easy to conceive and simulate, stochastic processes defined by an iterative formula (linear or functional) have been the subject of many studies. The iterative structure often leads to simpler and more explicit arguments than certain classical theories of processes.

On the other hand, when it comes to choosing step-by-step decision algorithms (sequential statistics, control, learning, ...) recursive decision methods are indispensable. They lend themselves naturally to problems of the identification and control of iterative stochastic processes. In recent years, know-how in this area has advanced greatly; this is reflected in the corresponding theoretical problems, many of which remain open.

### **At Whom Is This Book Aimed?**

I thought it useful to present the basic ideas and tools relating to random iterative models in a form accessible to a mathematician familiar with the classical concepts of probability and statistics but lacking experience in automatic control theory. Thus, the first aim of this book is to show young research workers that work in this area is varied and interesting and to facilitate their initiation period. The second aim is to present more seasoned probabilists with a number of recent original techniques and arguments relating to iterative methods in a fairly classical environment.

Very diverse problems (prediction of electricity consumption, production control, satellite communication networks, industrial chemistry, neurons, ...) lead engineers to become interested in stochastic algorithms which can be used to stabilize, identify or control increasingly complex models. Their experience and the diversity of their techniques go far beyond our aims here. But the third aim of the book is to provide them with a toolbox containing a quite varied range of basic tools.

Lastly, it seems to me that many lectures on stochastic processes could be centred around a particular chapter. The division into self-contained parts described below is intended to make it easy for undergraduate or postgraduate students and their teachers to access selected and relevant material.

## Contents

The overall general foundations are laid in Part I. The other three parts can be read independently of each other (apart from a number of easily locatable references and optional examples). This facilitates partial use of this text as research material or as teaching material on stochastic models or the statistics of processes.

### Part I. Sources of Recursive Methods

Chapter 1 presents the first mathematical ideas about sequential statistics and about stochastic algorithms (Robbins–Monro). An outline sketch of the theory of martingales is given together with certain complementary information about recursive methods.

Chapter 2 summarizes the theory of convergence in distribution and that of the central limit theorem for martingales, which is then applied to the Robbins–Monro algorithm. The AR(1) autoregressive vectorial model of order 1 is studied in detail; this model will provide the essential link between the following three parts.

Despite its abstract style, the development of this book has been heavily influenced by dialogues with other research workers interested in highly specific industrial problems. Chapter 3 gives an all-too-brief glimpse of such examples.

### Part II. Linear Models

The mathematical foundations of automatic control theory, which were primed in Chapter 2 based on the AR(1) model, are developed here.

Chapter 4 discusses the concepts of causality and excitation for ARMAX models. The importance of transferring the excitation of the noise onto that of the system is emphasized and algebraic criteria guaranteeing such a transfer are established.

Identification and tracking problems are considered in Chapter 5, using classical (gradient and least squares) or more recent (weighted least squares) estimators.

### Part III. Nonlinear Models

The first part of Chapter 6 describes the concept of ‘stability’ of an iterative Markov Fellerian model. Simple criteria ensuring the almost sure weak convergence of empirical distributions to a unique stationary distribution are obtained. This concept of stability seems to me, pedagogically and practically, much more manageable than the classical notion of recurrence; moreover, many models (fractals, automatic control theory) can be stable without being recurrent. A number of properties of rates of convergence in distribution and almost sure convergence complete this chapter.

The identification and tracking problems resolved in Chapter 5 for the linear case are much more difficult for functional regression models. Some partial solutions are given in Chapter 7, largely using the recursive kernel estimator.

#### **Part IV. Markov Models**

Paradoxically, Part IV of this book is the most classical. It involves a brief presentation of probabilistic topics described in greater detail elsewhere, placing them in the context of the preceding chapters.

The general theory of the recurrence of Markov chains is finally given in Chapter 8. Readers will note that, in many cases, it provides a useful complement to the stability theory of Chapter 7, but at the cost of much heavier techniques (and stronger assumptions about the noise).

On the subject of learning, Chapter 9 outlines the theory of controlled Markov chains and on-average optimal controls. The chapter ends with a number of results from the theory of stochastic approximation introduced in Chapter 1: the ordinary differential equation method, Markovian perturbation, traps, applications to visual neurons and principal components analysis.

#### **What You Will Not Find**

Since the main aim was to present recursive methods which are useful in adaptive control theory, it was natural to emphasize the almost sure properties (laws of large numbers, laws of the iterated logarithm, optimality of a strategy for the average asymptotic cost, . . .). Convergence in distribution is thus only discussed in outline and the principles of large deviations are not touched upon.

Iterative Markov models on finite spaces, the simulation of a particular model with a given stationary distribution and simulated annealing are currently in vogue, particularly in image processing circles. Although they come under the umbrella of ‘random iterative models’, they are not dealt with here.

These gaps have been partially filled in my recent book ‘Algorithmes Stochastiques’, 1966, Springer-Verlag.

#### **History**

The history of this book dates back to the end of the 1980s. It was developed at that time within the statistical research team of the Université Paris-Sud, in particular, by the automatic control team. Its contents have been enriched by numerous exchanges with the research workers of this team and its composition has been smoothed by several years of post-graduate courses. The first French edition of this book was published by Masson in 1990.

When, Springer-Verlag decided to commission an English translation in 1992, I felt it was appropriate to present a reworked text, taking into account the rapid evolution of some of the subjects treated. This book is a translation of that adaptation, which was carried out at the beginning of 1993 (with a number of additions and alterations to the Bibliography).

## **Acknowledgments**

It is impossible to thank all those research workers and students at the Université Paris-Sud and at the Université de Marne-la-Vallée where I have worked since 1993, who have contributed to this book through their dialogue. Their contributions will be acknowledged in the Bibliography.

Three research workers who have read and critically reviewed previous drafts deserve special mention: Bernard Bercu, Rachid Senoussi and Abderhamen Touati.

Lastly, Dr Stephen Wilson has been responsible for the English translation. He deserves hearty thanks for the intelligent and most useful critical way in which he achieved it.



# Notation

## Numbering System

- Within a chapter, a continuous numbering system is used for the Exercises on the one hand and for the Theorems, Propositions, Corollaries, Lemmas and Definitions on the other hand. The references indicate the chapter, section and number: Theorem 1.3.10 (or Exercise 1.3.10) occurs in Section 3 of Chapter 1 and is the tenth of that chapter.
- $\square$  marks the end of a Proof;  $\diamond$  marks the end of the statement of an Exercise or a Remark.

## Standard Mathematical Symbols

- *Abbreviations.* Constant is abbreviated to const. and  $\ln(\ln x)$  to LL.
- *Sets.*  $\mathbb{N}$  = integers  $\geq 0$ ;  $\mathbb{Z}$  = relative integers;  $\mathbb{Q}$  = rational numbers;  $\mathbb{R}$  = real numbers;  $\mathbb{C}$  = complex numbers.  
 $\mathbf{1}_A$  is the characteristic function for  $A$ ,

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

- *Sequences.* If  $(u_n)$  is a real monotonic sequence,  $u_\infty$  is its limit, either finite or infinite.  
If  $(u_n)$  and  $(v_n)$  are two positive sequences,  $u_n = O(v_n)$  (resp.  $o(v_n)$ ) means that  $(u_n/v_n)$  is bounded (resp. tends to 0).
- *Vectors.*  $u$ ,  ${}^t u$ ,  ${}^* u$ ,  $\langle u, v \rangle$ ,  $\|u\|$  – see Section 4.2.1.
- *Matrices*  $d \times d$ .  $A = (A^{ij})$ ,  $I$  or  $I_d$  identity,  ${}^t A$ ,  ${}^* A$ ,  $\text{Tr } A$ ,  $\|A\|$ ,  $\det A$  – see Section 4.2.1;  $\rho(A)$  – see Section 2.3.1.
- *Positive Hermitian Matrices.*  $\lambda_{\min} C$ ,  $\lambda_{\max} C$ ,  $\sqrt{C}$ ,  $C^{-1}$ ,  $C^1 \leq C^2$  – see Section 4.2.1;  $C^1 \otimes C^2$  – see Section 6.3.2.  
Norm of a rectangular matrix  $B$ ,  $\|B\|$  – see Section 4.2.1.
- *Excitation of a Sequence of Vectors*  $y = (y_n)$ .  $c_n(y) = \sum_{k=0}^n y_k {}^* y_k$ . We also set (see Section 4.2)  $s_n(y) = \sum_{k=0}^n \|y_k\|^2$ ,

$$f_n(y) = {}^* y_n (c_n(y))^{-1} y_n \quad \text{and} \quad g_n(y) = {}^* (y_n (c_{n-1}(y))^{-1} y_n.$$

- *Functions.* If  $\phi$  is differentiable from  $\mathbb{R}^p$  to  $\mathbb{R}^q$ , we denote its Jacobian matrix by  $D\phi$ . When  $q = 1$ ,  $\nabla\phi = {}^tD\phi$  is its gradient.
- *Lipschitz function.*  $\text{Li}(r, s)$  – Section 6.3.2.
- *ODE* – Section 9.2.

## Standard Probabilistic Symbols

- *Measure.*  
 $(\Omega, \mathcal{A}, P)$  probability space;  $\mathbb{F} = (\mathcal{F}_n)$  filtration – see Section 1.1.5;  
 $(E^n, \mathcal{E}^{\otimes n}) = (E, \mathcal{E})^n$ ;  $\mathcal{B}_E$  Borel  $\sigma$ -field for  $E$ .  
 For  $f$  measurable from  $(\Omega, \mathcal{A})$  to  $(E, \mathcal{E})$  and  $\Gamma \in \mathcal{E}$ , we denote  $\{f \in \Gamma\} = \{\omega; f(\omega) \in \Gamma\}$ .  
 For two sequences of positive random variables  $(\alpha_n)$  and  $(\beta_n)$ , we denote

$$\begin{aligned} \{\alpha_n = O(\beta_n)\} &= \{\omega; \alpha_n(\omega) = O(\beta_n(\omega))\} \\ \{\alpha_n = o(\beta_n)\} &= \{\omega; \alpha_n(\omega) = o(\beta_n(\omega))\} \end{aligned}$$

a.s. = almost surely

$\langle M \rangle$  = increasing process, hook of a martingale – see Sections 1.3.1, 2.1.3 and 4.3.2.

- *Convergence.*  
 $\xrightarrow{\text{a.s.}}$  = converges almost surely  
 $\xrightarrow{P}$  = converges in probability  
 $\xrightarrow{\mathcal{L}}$  = converges in distribution.

## Symbols for Linear Models

- *Models.* ARMAX, ARMA, ARX, AR, MA – see Section 4.1.1; RMA – see Section 5.4.1
- *Estimators.* LS, RLS – Section 5.2.1; SG – Section 5.3.1; WLS – Section 5.3.2; ELS, AML – Section 5.4.1
- $R$  for the delay operator – Section 4.1.1

## Symbols for Nonlinear Models

ARF – Section 6.2.3; ARXF – Section 6.2.4; ARCH – Section 6.3.3

# Table of Contents

<b>Preface</b> .....	V
<b>Notation</b> .....	IX

---

## Part I. Sources of Recursive Methods

---

<b>1. Traditional Problems</b> .....	3
1.1 Themes .....	3
1.1.1 Dosage: Robbins–Monro Procedure .....	3
1.1.2 Search for a Maximum: Kiefer–Wolfowitz Procedure ....	4
1.1.3 The Two-armed Bandit .....	5
1.1.4 Tracking .....	5
1.1.5 Decisions Adapted to a Sequence of Observations .....	6
1.1.6 Recursive Estimation .....	7
1.2 Deterministic Recursive Approximation .....	11
1.2.1 Search for a Zero of a Continuous Function .....	11
1.2.2 Search for Extrema .....	13
1.3 Random Sequences and Series .....	14
1.3.1 Martingales .....	14
1.3.2 Stopping and Inequalities .....	16
1.3.3 Robbins–Siegmund Theorem .....	18
1.3.4 Laws of Large Numbers .....	19
1.3.5 Laws of Large Numbers and Limits of Maxima .....	23
1.3.6 Noise and Regressive Series .....	24
1.4 Stochastic Recursive Approximation .....	29
1.4.1 Robbins–Monro Algorithm .....	29
1.4.2 Control .....	30
1.4.3 The Two-armed Bandit .....	32
1.4.4 Tracking .....	33
1.4.5 Recursive Estimation .....	35

<b>2. Rate of Convergence</b> .....	39
2.1 Convergence in Distribution .....	39
2.1.1 Weak Convergence on a Metric Space .....	39
2.1.2 Convergence in distribution of random vectors .....	40
2.1.3 Central Limit Theorem for Martingales .....	45
2.1.4 Lindeberg's Condition .....	47
2.1.5 Applications .....	48
2.2 Rate of Convergence of the Robbins–Monro Algorithm .....	52
2.2.1 Convergence in distribution of the Robbins–Monro Algorithm .....	52
2.2.2 Rate of Convergence of Newton's Estimator .....	57
2.3 Autoregressive Models .....	59
2.3.1 Spectral Radius .....	59
2.3.2 Stability .....	61
2.3.3 Random Geometric Series .....	64
2.3.4 Explosive Autoregressive Model .....	68
2.3.5 Jordan Decomposition .....	70
<b>3. Current Problems</b> .....	75
3.1 Linear Regression .....	75
3.1.1 Multiple Regression .....	75
3.1.2 Time Series .....	76
3.1.3 Tuning .....	77
3.2 Nonlinear Regression .....	79
3.2.1 A Tank .....	79
3.2.2 Prediction of Electricity Consumption .....	80
3.3 Satellite Communication: Markov Models .....	82
3.4 Neurons: Learning .....	85
3.4.1 From Neurobiology to Learning Algorithms .....	85
3.4.2 Artificial Neural Networks .....	86

---

**Part II. Linear Models**

---

<b>4. Causality and Excitation</b> .....	89
4.1 ARMAX Models .....	89
4.1.1 Definitions and Notation .....	89
4.1.2 Companion Matrix .....	91
4.1.3 Causality and Stability .....	92
4.2 Excitation .....	95
4.2.1 Positive Hermitian Matrices .....	95
4.2.2 Sequences and Series of Positive Hermitian Matrices ....	97
4.2.3 Excitation of a Sequence of Vectors .....	100
4.2.4 Inversion of the Excitation .....	101
4.3 Laws of Large Numbers .....	104

4.3.1	Preliminary Calculations . . . . .	104
4.3.2	Vector Martingales . . . . .	106
4.3.3	Multidimensional Regressive Series . . . . .	108
4.3.4	Counterexamples . . . . .	111
4.4	Transfers of Excitation . . . . .	116
4.4.1	Transfers of the Noise Excitation . . . . .	116
4.4.2	Irreducibility . . . . .	119
4.4.3	Transfers of Excitation to an ARMAX Model . . . . .	120
4.4.4	Excitation of an AR(1) Model . . . . .	124
4.4.5	Excitation of an ARMA Model . . . . .	129
<b>5.</b>	<b>Linear Identification and Tracking . . . . .</b>	<b>133</b>
5.1	Predict, Estimate, Track . . . . .	133
5.1.1	Estimators and Predictors . . . . .	133
5.1.2	Tracking . . . . .	135
5.2	Identification of Regression Models . . . . .	136
5.2.1	Least-squares Estimator . . . . .	136
5.2.2	Identification of an AR( $p$ ) Model . . . . .	140
5.3	Tracking by Regression Models . . . . .	144
5.3.1	Gradient Estimator . . . . .	144
5.3.2	Weighted Least-squares Estimator . . . . .	145
5.3.3	Tracking . . . . .	147
5.4	ARMAX Model: Estimators and Predictors . . . . .	150
5.4.1	Description of Estimators and Predictors . . . . .	150
5.4.2	Fourier Series: Passivity . . . . .	152
5.4.3	Consistency and Prediction Errors . . . . .	158
5.5	ARMAX Model: Identification and Tracking . . . . .	166
5.5.1	Identification of an ARMA Model . . . . .	166
5.5.2	Identification of an ARMAX Model . . . . .	168
5.5.3	Tracking: Optimality . . . . .	169
5.5.4	Tracking: Optimality and Identification . . . . .	173

---

### Part III. Nonlinear Models

---

<b>6.</b>	<b>Stability . . . . .</b>	<b>181</b>
6.1	Stability and Recurrence . . . . .	181
6.1.1	Stability and Recurrence of a System . . . . .	181
6.1.2	Markov Chain . . . . .	183
6.1.3	Stationary Distribution . . . . .	186
6.2	Lyapunov's Method . . . . .	189
6.2.1	Stabilization Criteria . . . . .	189
6.2.2	Stability Criteria . . . . .	191
6.2.3	Stable ARF( $p$ ) Models . . . . .	193
6.2.4	Markov Representations . . . . .	196

6.3	Lipschitz Mixing .....	199
6.3.1	Stability of Iterative Lipschitz Models .....	199
6.3.2	Lipschitz Mixing .....	200
6.3.3	Stable Iterative Model of Order $p$ .....	205
6.4	Rates .....	208
6.4.1	Law of the Iterated Logarithm .....	209
6.4.2	Rates for a Stable AR( $p$ ) Model .....	213
6.4.3	Convergence in Distribution of Continuous Processes .....	216
6.4.4	Functional Central Limit Theorem .....	218
6.4.5	Uniform Laws of Large Numbers .....	218
<b>7.</b>	<b>Nonlinear Identification and Control</b> .....	<b>227</b>
7.1	Estimation of the Stationary Distribution of a Stable Model .....	227
7.1.1	Empirical Estimators .....	227
7.1.2	Regularized Empirical Estimators .....	228
7.1.3	Estimation of the Density of the Stationary Distribution ..	230
7.2	Estimation of a Regression Function .....	238
7.2.1	Empirical Estimators of a Regression Function .....	238
7.2.2	Regression with a Stable Explicative Variable .....	241
7.2.3	Identification of a Stable ARF( $p$ ) Model .....	244
7.2.4	Regression with a Stabilized Explicative Variable .....	245
7.2.5	Prediction Errors .....	247
7.3	Controlled Markov Chain .....	251
7.3.1	Modelling and Examples .....	251
7.3.2	Likelihood of a Controlled Markov Chain .....	254
7.3.3	Stabilization .....	256
7.3.4	Optimal control .....	260
7.3.5	Optimal Quadratic Cost of an ARX(1,1) Model .....	263

---

**Part IV. Markov Models**

---

<b>8.</b>	<b>Recurrence</b> .....	<b>269</b>
8.1	Markov Chain .....	269
8.1.1	Data and Definitions .....	269
8.1.2	Markov Properties .....	271
8.1.3	Return Times .....	272
8.1.4	Coupling .....	274
8.2	Recurrence and Transience .....	276
8.2.1	Concept of Recurrence .....	276
8.2.2	Atomic Markov Chains .....	278
8.2.3	Random Walks on $\mathbb{Z}$ .....	281
8.2.4	From Atoms to Small Sets .....	286
8.3	Rate of Convergence to the Stationary Distribution .....	291

8.3.1	Convergence of Transition Probabilities to the Stationary Distribution .....	292
8.3.2	Central Limit Theorem .....	295
8.3.3	Orey's Theorem .....	298
8.3.4	Riemannian and Geometric Recurrence .....	299
<b>9.</b>	<b>Learning .....</b>	<b>305</b>
9.1	Controlled Markov Chains .....	305
9.1.1	Markov Property .....	305
9.1.2	Return Times and Recurrence .....	306
9.1.3	Average Costs .....	310
9.2	Stochastic Algorithms .....	315
9.2.1	Regions of Attraction of a Differential Equation .....	316
9.2.2	The Ordinary Differential Equation Method .....	317
9.2.3	Markovian Perturbations .....	321
9.3	Search for a Strongly Attractive Point .....	327
9.3.1	Gradient Estimator for a Linear Model .....	327
9.3.2	Strongly Attractive Point .....	331
9.3.3	Visual Neurons .....	334
9.3.4	Search for the Minimum of a Potential .....	337
9.4	How Algorithms Avoid Traps .....	340
9.4.1	Negligible Sets for Regressive Series .....	340
9.4.2	Recursive Principal Components Analysis .....	343
<b>Bibliography</b>	.....	<b>349</b>
<b>Index</b>	.....	<b>381</b>