STOCHASTIC MODELLING AND APPLIED PROBABILITY

60

Alan Bain
Dan Crisan

Fundamentals of Stochastic Filtering

Springer

STOCHASTIC MODELLING AND APPLIED PROBABILITY

60

Alan Bain
Dan Crisan

Fundamentals of Stochastic Filtering

Springer

Stochastic Mechanics
Random Media
Signal Processing and Image Synthesis
Mathematical Economics and Finance
Stochastic Optimization
Stochastic Control
Stochastic Models in Life Sciences

Stochastic Modelling and Applied Probability

(Formerly:
Applications of Mathematics)

60

For other titles published in this series, go to
www.springer.com/series/602

Alan Bain · Dan Crisan

Fundamentals of Stochastic Filtering

Springer

Alan Bain
BNP Paribas
10 Harewood Av
London NW1 6AA
United Kingdom
alan.bain@bnpparibas.com

Dan Crisan
Department of Mathematics
Imperial College London
180 Queen's Gate
London SW7 2AZ
United Kingdom
d.crisan@imperial.ac.uk

ISSN: 0172-4568 Stochastic Modelling and Applied Probability
ISBN: 978-0-387-76895-3 e-ISBN: 978-0-387-76896-0
DOI 10.1007/978-0-387-76896-0

Library of Congress Control Number: 2008938477

Mathematics Subject Classification (2000): 93E10, 93E11, 60G35, 62M20, 60H15

Printed on acid-free paper

springer.com

Preface

Many aspects of phenomena critical to our lives can not be measured directly. Fortunately models of these phenomena, together with more limited observations frequently allow us to make reasonable inferences about the state of the systems that affect us. The process of using partial observations and a stochastic model to make inferences about an evolving system is known as *stochastic filtering*.

The objective of this text is to assist anyone who would like to become familiar with the theory of stochastic filtering, whether graduate student or more experienced scientist. The majority of the fundamental results of the subject are presented using modern methods making them readily available for reference. The book may also be of interest to practitioners of stochastic filtering, who wish to gain a better understanding of the underlying theory.

Stochastic filtering in continuous time relies heavily on measure theory, stochastic processes and stochastic calculus. While knowledge of basic measure theory and probability is assumed, the text is largely self-contained in that the majority of the results needed are stated in two appendices. This should make it easy for the book to be used as a graduate teaching text. With this in mind, each chapter contains a number of exercises, with solutions detailed at the end of the chapter.

The book is divided into two parts: The first covers four basic topics within the theory of filtering: the filtering equations (Chapters 3 and 4), Clark's representation formula (Chapter 5), finite-dimensional filters, in particular, the Beneš and the Kalman–Bucy filter (Chapter 6) and the smoothness of the solution of the filtering equations (Chapter 7). These chapters could be used as the basis of a one- or two-term graduate lecture course.

The second part of the book is dedicated to numerical schemes for the approximation of the solution of the filtering problem. After a short survey of the existing numerical schemes (Chapter 8), the bulk of the material is dedicated to particle approximations. Chapters 9 and 10 describe various particle filtering methods in continuous and discrete time and prove associated convergence results. The material in Chapter 10 does not require knowledge of stochastic integration and could form the basis of a short introductory course.

We should like to thank the publishers, in particular the senior editor, Achi Dosanjh, for her understanding and patience. Thanks are also due to various people who offered their support and advice during the project, in particular Martin Clark, Mark Davis and Boris Rozovsky. One of the authors (D.C.) would like to thank Robert Piché for the invitation to give a series of lectures on the subject in August 2006.

Part of the book grew out of notes on lectures given at Imperial College London, University of Cambridge and Tampere University of Technology. Special thanks are due to Kari Heine from Tampere University of Technology and Olasunkanmi Obanubi from Imperial College London who read large portions of the first draft and suggested many corrections and improvements.

Finally we would like to thank our families for their support, without which this project would have never happened.

London
December 2007

Alan Bain
Dan Crisan

Contents

Notation

Spaces

- $\mathbb{R}^d$ – the d-dimensional Euclidean space.
- $\overline{\mathbb{R}^d}$ – the one-point compactification of $\mathbb{R}^d$ formed by adjoining a single point at infinity to $\mathbb{R}^d$.
- $\mathcal{B}(\mathbb{S})$ – the Borel σ-field on $\mathbb{S}$. That is the σ-field generated by the open sets in $\mathbb{S}$. If $\mathbb{S} = \mathbb{R}^d$ for some d, then this σ-field is countably generated.
- $(\mathbb{S}, \mathcal{S})$ – the state space for the signal. Unless otherwise stated, $\mathbb{S}$ is a complete separable metric space and $\mathcal{S}$ is the associated Borel σ-field $\mathcal{B}(\mathbb{S})$.
- $C(\mathbb{S})$ – the space of real-valued continuous functions defined on $\mathbb{S}$.
- $M(\mathbb{S})$ – the space of $\mathcal{B}(\mathbb{S})$-measurable functions $\mathbb{S} \to \mathbb{R}$.
- $B(\mathbb{S})$ – the space of bounded $\mathcal{B}(\mathbb{S})$-measurable functions $\mathbb{S} \to \mathbb{R}$.
- $C_b(\mathbb{S})$ – the space of bounded continuous functions $\mathbb{S} \to \mathbb{R}$.
- $C_k(\mathbb{S})$ – the space of compactly supported continuous functions $\mathbb{S} \to \mathbb{R}$.
- $C_k^m(\mathbb{S})$ – the space of compactly supported continuous functions $\mathbb{S} \to \mathbb{R}$ whose first m derivatives are continuous.
- $C_b^m(\mathbb{R}^d)$ – the space of all bounded, continuous functions with bounded partial derivatives up to order m. The norm $\|\cdot\|_{m,\infty}$ is frequently used with this space.
- $C_b^\infty(\mathbb{R}^d) = \bigcap_{m=0}^\infty C_b^m(\mathbb{R}^d)$.
- $D_{\mathbb{S}}[0,\infty)$ – the space of càdlàg functions from $[0,\infty) \to \mathbb{S}$.
- $C_b^{1,2}$ the space of bounded continuous real-valued funtions $u(t,x)$ with domain $[0,\infty) \times \mathbb{R}$, which are differentiable with respect to t and twice differentiable with respect to x. These derivatives are bounded and continuous with respect to (t,x).
- $C^l(\mathbb{R}^d)$ the subspace of $C(\mathbb{R}^d)$ containing functions φ such that $\varphi/\psi \in C_b(\mathbb{R}^d)$, where $\psi(x) = 1 + \|x\|$.
- $W_p^m(\mathbb{R}^d)$ – the Sobolev space of all functions with generalized partial derivatives up to order m with both the function and all its partial derivatives being L^p-integrable. This space is usually endowed with the norm $\|\cdot\|_{m,p}$.

- $\mathrm{SL}(\mathbb{R}^d) = \left\{\varphi \in C_b(\mathbb{R}^d) : \exists\, M \text{ such that } \varphi(x) \le M/(1+\|x\|),\ \forall x \in \mathbb{R}^d\right\}$
- $\mathcal{M}(\mathbb{S})$ – the space of finite measures over $(\mathbb{S}, \mathcal{S})$.
- $\mathcal{P}(\mathbb{S})$ – the space of probability measures over $(\mathbb{S}, \mathcal{S})$, i.e the subspace of $\mathcal{M}(\mathbb{S})$ such that $\mu \in \mathcal{P}(\mathbb{S})$ satisfies $\mu(\mathbb{S}) = 1$.
- $D_{M_F(\mathbb{R}^d)}[0,\infty)$ – the space of right continuous functions with left limits $a : [0,\infty) \to M_F(\mathbb{R}^d)$ endowed with the Skorohod topology.
- I – an arbitrary finite set $\{a_1, a_2, \ldots\}$.
- $P(I)$ – the power set of I, i.e. the set of all subsets of I.
- $\mathcal{M}(I)$ – the space of finite positive measures over $(I, P(I))$.
- $\mathcal{P}(I)$ – the space of probability measures over $(I, P(I))$, i.e. the subspace of $\mathcal{M}(I)$ such that $\mu \in P(I)$ satisfies $\mu(I) = 1$.

Other notations

- $\|\cdot\|$ – the Euclidean norm, for $x = (x_i)_{i=1}^m \in \mathbb{R}^m$, $\|x\| = \sqrt{x_1^2 + \cdots + x_m^2}$. It is also applied to $d \times p$-matrices by considering them as $d \times p$ vectors, viz
$$\|a\| = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{p} a_{ij}^2}.$$
- $\|\cdot\|_\infty$ – the supremum norm; for $\varphi\colon \mathbb{R}^d \to \mathbb{R}$, $\|\varphi\|_\infty = \sup_{x\in\mathbb{R}^d} |\varphi(x)|$. In general if $\varphi : \mathbb{R}^d \to \mathbb{R}^m$ then
$$\|\varphi\|_\infty = \max_{i=1,\ldots m} \sup_{x\in\mathbb{R}^d} |\varphi^i(x)|.$$
The notation $\|\cdot\|_\infty$ is equivalent to $\|\cdot\|_{0,\infty}$. This norm is especially useful on spaces such as $C_b(\mathbb{R}^d)$, or $C_k(\mathbb{R}^d)$, which only contain functions of bounded supremum norm; in other words, $\|\varphi\|_\infty < \infty$.
- $\|\cdot\|_{m,p}$ – the norm used on the space W_p^m defined by
$$\|\varphi\|_{m,p} = \left(\sum_{|\alpha|\le m} \|D_\alpha \varphi(x)\|_p^p\right)^{1/p}$$
where $\alpha = (\alpha^1, \ldots, \alpha^d)$ is a multi-index and $D_\alpha\varphi = (\partial_1)^{\alpha^1} \ldots (\partial_d)^{\alpha^d}\varphi$.
- $\|\cdot\|_{m,\infty}$ is the special case of the above norm when $p = \infty$, defined by
$$\|\varphi\|_{m,\infty} = \sum_{|\alpha|\le m} \sup_{x\in\mathbb{R}^d} |D_\alpha\varphi(x)|\,.$$
- δ_a – the Dirac measure concentrated at $a \in \mathbb{S}$, $\delta_x(A) \equiv \mathbf{1}_A(x)$.
- $\mathbf{1}$ – the constant function 1.
- $\Rightarrow$ – used to denote weak convergence of probability measures in $\mathcal{P}(\mathbb{S})$; see Definition 2.14.

- μf, $\mu(f)$ – the integral of $f \in B(\mathbb{S})$ with respect to $\mu \in \mathcal{M}(\mathbb{S})$, i.e. $\mu f \triangleq \int_{\mathbb{S}} f(x)\mu(\mathrm{d}x)$.
- $a^\top$ is the transpose of the matrix a.
- $\mathbb{I}_d$ – the $d \times d$ identity matrix.
- $\mathbb{O}_{d,m}$ – the $d \times m$ zero matrix.
- $\mathrm{tr}(A)$ – the trace of the matrix A, i.e. if $A = (a_{ij})$, then $\mathrm{tr}(A) = \sum_i a_{ii}$.
- $[x]$ – the integer part of $x \in \mathbb{R}$.
- $\{x\}$ – the fractional part of $x \in \mathbb{R}$, i.e. $x - [x]$.
- $\langle M \rangle_t$ – the quadratic variation of the semi martingale M.
- $s \wedge t$ – for $s, t \in \mathbb{R}$, $s \wedge t = \min(s,t)$.
- $s \vee t$ – for $s, t \in \mathbb{R}$, $t \vee s = \max(s,t)$.
- $A \vee B$ – the σ-algebra generated by the union $A \cup B$.
- $A \triangle B$ – the symmetric difference of sets A and B, i.e. all elements that are in one of A or B but not both, formally $A \triangle B = (A \setminus B) \cup (B \setminus A)$.
- $\mathcal{N}$ – the collection of null sets in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

1

Introduction

1.1 Foreword

The development of mathematics since the 1950s has gone through many radical changes both in scope and in depth. Practical applications are being found for an increasing number of theoretical results and practical problems have also stimulated the development of theory. In the case of stochastic filtering, it is not clear whether this first arose as an application found for general theory, or as the solution of a practical problem.

Stochastic filtering now covers so many areas that it would be futile to attempt to write a comprehensive book on the subject. The purpose of this text is not to be exhaustive, but to provide a modern, solid and accessible starting point for studying the subject.

The aim of stochastic filtering is to estimate an evolving dynamical system, the signal, customarily modelled by a stochastic process. Throughout the book the signal process is denoted by $X = \{X_t,\ t \geq 0\}$, where t is the temporal parameter. Alternatively, one could choose a discrete time process, i.e. a process $X = \{X_t,\ t \in \mathbb{N}\}$ where t takes values in the (discrete) set $\{0, 1, 2, \ldots\}$. The former continuous time description of the process has the benefit that use can be made of the power of stochastic calculus. A discrete time process may be viewed as a continuous time process with jumps at fixed times. Thus a discrete time process can be viewed as a special case of a continuous time process. However, it is not necessarily effective to do so since it is much easier and more transparent to study the discrete case directly. Unless otherwise stated, the process X and all other processes are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The signal process X can not be measured directly. However, a partial measurement of the signal can be obtained. This measurement is modelled by another continuous time process $Y = \{Y_t,\ t \geq 0\}$ which is called the *observation* process. This observation process is a function of X and a measurement noise. The measurement noise is modelled by a stochastic process $W = \{W_t,\ t \geq 0\}$. Hence,

A. Bain, D. Crisan, *Fundamentals of Stochastic Filtering*,
DOI 10.1007/978-0-387-76896-0_1,