The IMA Volumes
in Mathematics
and Its Applications

Volume 10

Wendell Fleming
Pierre-Louis Lions
Editors

Stochastic Differential Systems, Stochastic Control Theory and Applications

Springer-Verlag
World Publishing Corporation

Wendell Fleming Pierre-Louis Lions
Editors

Stochastic Differential Systems, Stochastic Control Theory and Applications

With 10 Illustrations

Springer-Verlag
World Publishing Corporation

Wendell Fleming
Division of Applied Mathematics
Brown University
Providence, Rhode Island 02912, USA

Pierre-Louis Lions
Ceremade
Universite Paris-Dauphine
Place de Lattre de Tassigny
75775 Paris Cedex 16
France

AMS Classifications: 60HXX, 93EXX

Library of Congress Cataloging in Publication Data
Stochastic differential systems, stochastic control theory, and applications.
(The IMA volumes in mathematics and its applications ; v. 10)
"Proceedings of a Workshop on Stochastic Differential Systems, Stochastic Control Theory, and Applications, held at IMA June 9–19, 1986"—Pref.
Bibliography: p.
1. Stochastic systems–Congresses. 2. Differentiable dynamical systems–Congresses. 3. Control theory–Congresses. I. Fleming, Wendell Helms, 1928– II. Lions, P. L. (Pierre-Louis) III. Workshop on Stochastic Differential Systems, Stochastic Control. Theory, and Applications (1986 : University of Minnesota, Institute for Mathematics and its Applications) IV. Series.
QA402.S8463 1988 003 87–28380

Reprinted by World Publishing Corporation, Beijing, 1990
for distribution and sale in The People's Republic of China only

ISBN 0-387-96641-2 Springer-Verlag New York Berlin Heidelberg
ISBN 3-540-96641-2 Springer-Verlag Berlin Heidelberg New York
ISBN 7-5062-0542-4

The IMA Volumes in Mathematics and Its Applications

Current Volumes:

Volume 1: Homogenization and Effective Moduli of Materials and Media
Editors: Jerry Ericksen, David Kinderlehrer, Robert Kohn, and J.-L. Lions

Volume 2: Oscillation Theory, Computation, and Methods of Compensated Compactness
Editors: Constantine Dafermos, Jerry Ericksen, David Kinderlehrer, and Marshall Slemrod

Volume 3: Metastability and Incompletely Posed Problems
Editors: Stuart Antman, Jerry Ericksen, David Kinderlehrer, and Ingo Müller

Volume 4: Dynamical Problems in Continuum Physics
Editors: Jerry Bona, Constantine Dafermos, Jerry Ericksen, and David Kinderlehrer

Volume 5: Theory and Applications of Liquid Crystals
Editors: Jerry Ericksen and David Kinderlehrer

Volume 6: Amorphous Polymers and Non-Newtonian Fluids
Editors: Constantine Dafermos, Jerry Ericksen, and David Kinderlehrer

Volume 7: Random Media
Editor: George Papanicolaou

Volume 8: Percolation Theory and Ergodic Theory of Infinite Particle Systems
Editor: Harry Kesten

Volume 9: Hydrodynamic Behavior and Interacting Particle Systems
Editor: George Papanicolaou

Volume 10: Stochastic Differential Systems, Stochastic Control Theory and Applications
Editors: Wendell Fleming and Pierre-Louis Lions

Forthcoming Volumes:

1986–1987: Scientific Computation
Numerical Simulation in Oil Recovery
Computational Fluid Dynamics and Reacting Gas Flows
Numerical Algorithms for Modern Parallel Computer Architectures
Mathematical Aspects of Scientific Software
Atomic and Molecular Structure and Dynamics

FOREWORD

This IMA Volume in Mathematics and its Applications

STOCHASTIC DIFFERENTIAL SYSTEMS, STOCHASTIC CONTROL THEORY AND APPLICATIONS

is the proceedings of a workshop which was an integral part of the 1986-87 IMA program on STOCHASTIC DIFFERENTIAL EQUATIONS AND THEIR APPLICATIONS. We are grateful to the Scientific Committee:

Daniel Stroock (Chairman)
Wendell Fleming
Theodore Harris
Pierre-Louis Lions
Steven Orey
George Papanicolaou

for planning and implementing an exciting and stimulating year-long program. We especially thank Wendell Fleming and Pierre-Louis Lions for organizing an interesting and productive workshop in an area in which mathematics is beginning to make significant contributions to real-world problems.

George R. Sell

Hans Weinberger

PREFACE

This volume is the Proceedings of a Workshop on Stochastic Differential Systems, Stochastic Control Theory, and Applications held at IMA June 9-19, 1986. The Workshop Program Committee consisted of W.H. Fleming and P.-L. Lions (co-chairmen), J. Baras, B. Hajek, J.M. Harrison, and H. Sussmann.

The Workshop emphasized topics in the following four areas.

(1) Mathematical theory of stochastic differential systems, stochastic control and nonlinear filtering for Markov diffusion processes. Connections with partial differential equations.

(2) Applications of stochastic differential system theory, in engineering and management science. Adaptive control of Markov processes. Advanced computational methods in stochastic control and nonlinear filtering.

(3) Stochastic scheduling, queueing networks, and related topics. Flow control, multiarm bandit problems, applications to problems of computer networks and scheduling of complex manufacturing operations.

(4) Simulated annealing and related stochastic gradient algorithms. Connections with statistical mechanics and combinational optimization.

This choice of topics was deliberately made to obtain a mix of traditional areas of stochastic control theory, and topics arising in newer areas of application. The papers included in this volume represent a diversity of approaches and points of view. They emphasize variously underlying mathematical theory, modelling issues and questions of computational implementation.

We would like to take this opportunity to extend our gratitude to the staff of the IMA, Professors H. Weinberger and G.R. Sell, Mrs. Pat Kurth, and Mr. Robert Copeland for their assistance in arranging the workshop. Special thanks are due to Mrs. Patricia Brick, and Mrs. Kaye Smith for their preparation of the manuscripts. We gratefully acknowledge the support of the Division of Mathematical Sciences and the Engineering and Computer Science Division of the National Science Foundation and the U.S. Army Research Office.

W.H. Fleming
P.-L. Lions

CONTENTS

Foreword x

Preface xiii

Optimality of "full bang to reduce predicted miss" for some partially observed stochastic control problems

V.E. Beneš, R.W. Rishel 1

On some approximation techniques in non-linear filtering

A. Bensoussan 17

Applications of Homogenization Theory to the control of flexible structures

Gilmer L. Blankenship 33

Control of Markov chains with long-run average cost criterion

Vivek S. Borkar 57

Automatic study in stochastic control

J.P. Chancelier, C. Gomez, J.P. Quadrat, A. Sulem 79

Some results on Kolmogoroff equations for infinite dimensional stochastic systems

G. Da Prato 87

Hamilton-Jacobi equations with constraints

I. Capuzzo-Dolcetta 99

An approximate minimum principle for a partially observed Markov chain

Robert J. Elliott 107

Generalized solutions in the optimal control of diffusions

Wendell H. Fleming and Domokos Vermes 119

Consistency of maximum likelihood and pseudo-likelihood estimators for Gibbs Distributions

B. Gidas 129

Brownian models of queueing networks with heterogeneous customer populations
J. Michael Harrison 147

Non-linear filtering - the degenerate case
U.G. Haussmann 187

The asymptotic behaviour of the maximum likelihood estimates for a class of diffusion processes.
Kurt Helmes 205

The filtering problem for infinite dimensional stochastic processes
G. Kallianpur and R.L. Karandikar 215

Stochastic control under finite-fuel constraints
Ioannis Karatzas 225

Recent advances in the theory of stochastic adaptive control
P.R. Kumar 241

Almost optimal controls for wideband noise driven systems
Harold J. Kushner 255

Asymptotic solutions of bandit problems
T.L. Lai 275

Viscosity solutions of second-order equations, stochastic control and stochastic differential games
P.-L. Lions and P.E. Souganidis 293

On the memory length of the optimal nonlinear filter
E. Mayer-Wolf, M. Zakai and O. Zeitouni 311

Implementation issues for Markov decision processes
Armand M. Makowski and Adam Shwartz 323

Navigating and stopping multi-parameter bandit processes
Avi Mandelbaum 339

Bounded variation control of a damped linear oscillator under random disturbances
Jose-Luis Menaldi 373

The support of the law of a filter in C^∞ topology
M. Chaleyat-Maurel and D. Michel 395

Existence of densities for statistics in the cubic sensor problem
Daniel Ocone 409

Piecewise linear filtering
E. Pardoux and C. Savona 429

Quick simulation of excessive backlogs in networks of queues
Shyam Parekh and Jean Walrand 439

On some perturbation problems in optimal stopping and impulse control
Maurice Robin .... 47[illegible]

Optimal control of jump-markov processes and viscosity solutions
Halil Mete Soner .... 50[illegible]

An introduction to singular stochastic control
Steven E. Shreve 51

Scheduling, routing, and flow control in stochastic networks
Shaler Stidham, Jr. .... 52[illegible]

Product expansions of exponential Lie Series and the discretization of stochastic differential equations
H.J. Sussmann 563

A survey of large time asymptotics of simulated annealing algorithms
John N. Tsitsiklis 583

Stochastic scheduling on parallel processors and minimization of concave functions of completion times
Richard R. Weber 601

Optimality of "full bang to reduce predicted miss" for some partially observed stochastic control problems

V. E. Beneš

AT&T Bell Laboratories
Murray Hill, New Jersey 07974
and
Columbia University
New York, New York 10027

R. W. Rishel

University of Kentucky
Lexington, KY 40506

ABSTRACT

For final value stochastic control problems, the "predicted miss" of the title is the expected final position, conditional on the cumulative information to date, if no further control is exerted. For partially observed problems with bounded control, similar to some proposed by Åström, we use PDE methods to show that predicted miss defines the switching surfaces for an optimal "bang-bang" control law whose gist, simply stated, is to reduce predicted miss. The surfaces in question are explicitly calculated.

1. Introduction

We report on a simple example of an approach which we believe is effective for a large class of problems. Consider the "partially observed" control task: For ϕ convex, even, of polynomial growth, and maximal curvature at 0, minimize $E\phi(y_T)$ subject to dynamics $dy_t = (z+u_t)dt + dw_t$, with T a final time, z an unobserved symmetric r.v. of known distribution, w_t a Wiener process (noise) independent of z, and u_t causal in the past of the observed y-process, and bounded by 1 in magnitude. Similar control problems under various constraints are to be found in [1],[8]. The main results are as follows.

An extension of the "predicted miss" idea used in [2] also yields an optimal law in the present, *partially observed* case. The predicted miss is the conditional expected final value of y_T, given information $Y_t = \sigma\{y_s,\ 0 \le s \le t\}$ up to t, if no control were exerted from t on. Here the predicted miss is calculated as

$$y_t + (T-t)\hat{z}_t$$

where $\hat{z}_t = E\{z\,|Y_t\}$. This conditional expectation is readily calculated by nonlinear filtering from the statistics of z. The optimal law takes the form

$$\begin{aligned} u_t^{\text{opt}} &= -\text{sgn}(y_t + (T-t)\hat{z}_t) \\ &= -\text{sgn}(\text{predicted miss})_{\text{at } t}, \end{aligned}$$

and so represents an extension of the old advice to "bang so as to reduce predicted miss." The switching curves are given by the "rotating lines" $y+(T-t)\hat{z} = 0$.

2. Innovations formulation

For simplicity, we carry through the development for a r.v. z that is Bernoulli: $z = \pm 1$ with probability $\frac{1}{2}$ each. The general case has exactly the same structure, modulo finiteness of $E \exp\{zx - \frac{1}{2} z^2 t\}$, and minor technical conditions.

The innovations formulation serves to change the original partially observed problem into a completely observed one by introducing various corditional moments and their dynamics. In this case the observation equation takes the form

$$dy_t = (\hat{z}_t + u_t)dt + dv_t$$

where $d\nu$ = innovations (a Brownian motion on $Y.$), and

$$\hat{z}_t = E\{z \,|Y_t\} = E\{z \,|y_s,\ 0 \le s \le t\}$$

is the current least squares estimate of z. The class A of admissible controls is that of all $Y.$-nonanticipating controls [3]. We calculate, using $\hat{z}_0 = 0$, that by the Kallianpur-Striebel formula

$$\hat{z}_t = \frac{E^z z \exp\left\{\int_0^t (z+u_s)dy_s - \frac{1}{2}\int_0^t (z+u_s)^2 ds\right\}}{E^z \exp\left\{\text{ditto, as in numerator}\right\}}$$

$$= \frac{E^z z \exp\left\{z(y_t-y_0) - z\int_0^t u_s ds\right\}}{E^z \exp\left\{\text{ditto, as in numerator}\right\}}$$

$$= \tanh\left(y_t - y_0 - \int_0^t u_s ds\right).$$

Since we must solve the problem for starts other than $\hat{z}_0 = 0$, it is convenient to set $\hat{z}_t = \tanh \xi_t$, with $\xi_0 = \tanh^{-1}\hat{z}_0$, and dynamics

$$d\xi_t = \tanh \xi_t \, dt + d\nu_t$$

The associated $\hat{z}_t$ dynamics are $d\hat{z}_t = (1-\hat{z}_t^2)d\nu_t$.

3. Intuitive discussion and guess

The term $\hat{z}_t dt$ driving dy_t is a drift. Its overall contribution to y_T is just

$$\int_t^T \hat{z}_s ds$$

Thus the expected final position, $E\{y_T|Y_t\}$, given information up to t, if no further control is exercised ($u_s \equiv 0$ in (t,T)), is just

$$
\begin{aligned}
r_t &= E\{y_T \mid Y_t\} \\
&= y_t + E\left\{\int_t^T \hat{z}_s ds \mid Y_t\right\} + E\left\{\nu_T - \nu_t \mid Y_t\right\} \\
&= y_t + \int_t^T E\left\{\hat{z}_s - \hat{z}_t \mid Y_t\right\} ds + \hat{z}_t(T-t) \\
&= y_t + \hat{z}_t(T-t)
\end{aligned}
$$

because $\hat{z}_t$ and ν_t are both martingales on Y_t. The process r_t is the predicted position (or "miss," since we are aiming at 0). See [2].

We claim that the right thing to do is "full bang to reduce the predicted miss r_t." This guess is in analogy to a completely observed case described in [2], but it can be given its own cogency in the present context, as follows: r_t satisfies the stochastic DE

$$
\begin{aligned}
dr_t &= dy_t + (T-t)d\hat{z}_t - \hat{z}_t dt \\
&= \left[1+(T-t)(1-\tanh^2\xi_t)\right] d\nu_t + u_t dt
\end{aligned}
$$

with ξ developing independently of u (but not of ν!) according to $d\xi_t = \tanh \xi_t + d\nu$. The final values r_T and y_T are the same, so how could we do better than to keep r_t near zero? Thus we are led to guess an optimal law in feedback form:

$$u_t^{\text{opt}} = -\text{sgn}\, r_t$$

This law gives a linear switching curve $y+(T-t)\hat{z} = 0$ in $(y, \hat{z})$-space, with slope changing linearly in time.

Remark: When we first started on this problem, our initial guess, presented at the 1986 IMA Workshop in Minneapolis, was that $u = -\text{sgn}\, y_t$ was optimal. It was primarily the strenuous objections of H. J. Sussman to this incorrect conjecture that guided us to find the *right* answer. He supplied an involved but convincing counter-example, which suggested the crystalline doable example $z = 1$ a.s.; the last led to the predicted miss. Sussman's role as critic and midwife is deeply appreciated.

In the light of the remark just made we see that $(T-t)\hat{z}$ is the "offset" the bang-bang law $-\text{sgn}\, y_t$ needs to be optimal and use the available information. A similar offset has appeared in Davis and Clark's [5] analysis of asymmetric predicted miss problems. We sketch a proof of the optimality of this offset law by PDEs, then given the details.

4. A PDE argument: sketch

$r_T = y_T$ a.s., so we use the variables r_t and ξ_t moving according to

$$d\xi_t = \tanh \xi_t dt + dv_t, \qquad \xi_0 = \tanh^{-1}\hat{z}_0$$

$$dr_t = u_t dt + \left[1+(T-t)(1-\tanh^2\xi_t)\right] dv_t,$$
$$r_0 = y_0 + T \tanh^{-1}\hat{z}_0,$$

and look at the Hamilton-Jacobi-Bellman equation

$$v_t + \frac{1}{2} v_{\xi\xi} + [1+(T-t)(1-\tanh^2\xi)]v_{\xi r}$$

$$+ \frac{1}{2} [1+(T-t)(1-\tanh^2\xi)]^2 v_{rr} + v_\xi \tanh \xi - |v_r| = 0$$

$$v|_T = \phi(r)$$

The characteristic form is only nonnegative; but by regularization and approximation to $|\cdot|$ by smooth even functions, we can use $\operatorname{sgn} \phi'(r) = \operatorname{sgn} r$ and the maximum principle to show that $\operatorname{sgn} v_r = \operatorname{sgn} r$. Cf. [6], [7]. Thus $r = 0$ gives the switching curve, i.e.

$$u_t^{\text{opt}} = -\operatorname{sgn}(y_t + (T-t)\hat{z}_t)$$

describes the optimal law: control so as to reduce the predicted final position if no further control were used. Optimality follows by a classical verification arguments [3].

5. Details of PDE argument

To make use of parabolic regularization we adjoin to the probability space an independent Brownian motion (b_t, B_t) and allow information about b_t to be used for control. That is, the class A of admissible controls now consists of those which are nonanticipating with respect to the filtration $Y_t \vee B_t$.

For $\epsilon \geq 0$ consider the ϵ-processes defined uniquely in law by the equations

$$
\begin{aligned}
(1)\qquad dr_t^\epsilon &= u_t\,dt + [1+(T-t)(1-\tanh^2 \xi_t^\epsilon)]\,dv_t \\
d\xi_t^\epsilon &= \tanh \xi_t^\epsilon\,dt + dv_t + \epsilon^{1/2}\,db_t \\
r_0^\epsilon &= y_0 + T\tanh^{-1}\hat{z}_0 \\
\xi_0^\epsilon &= \tanh^{-1}\hat{z}_0 .
\end{aligned}
$$

The requisite law can be obtained for any admissible u by Girsanov's theorem [3], giving rise to an expectation E^u. Set

$$
\begin{aligned}
J^\epsilon(u) &= J^\epsilon(u, r, \xi) = E^u_{r\xi}\,\phi(r_T^\epsilon) \\
V^\epsilon &= \inf_{u \in A} J^\epsilon(u).
\end{aligned}
$$

Writing, for brevity,

$$\sigma(t, x) = 1 + (T-t)[1-\tanh^2 x]$$

the dynamic programming equation for minimizing J^ϵ is just

$$(2)\qquad v_t + \frac{1}{2}\sigma^2(t,\xi)v_{rr} + \sigma(t,\xi)v_{\xi r} + \frac{1}{2}(1+\epsilon)v_{\xi\xi} + (\tanh\xi)v_\xi - |v_r| = 0$$

with final time condition $v|_T = \phi$; in operator form

$$(3)\qquad \left(\frac{\partial}{\partial t} + A^\epsilon\right)v = |v_r|.$$

We recall [3] the distinction between nonanticipating controls and feedback controls. The formula $u_t = -\operatorname{sgn} r^\epsilon$ ostensibly defines a feedback law for equation (1). More precisely, let $(r^{*\epsilon}, \xi^{*\epsilon})$ be the vector strong solution (process) of the stochastic DEs above for the explicit feedback law $u_t = -\operatorname{sgn} r_t^{*\epsilon}$; there is then a causal "solution map" S^ϵ such that