Nonlinear Filtering of Diffusion Processes
A Guided Tour
by
Sanjoy K. Mitter
Department of Electrical Engineering and Computer Science and
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139

1. Introduction

In this paper we give a guided tour through the development of nonlinear filtering of diffusion processes. The important topic of filtering of point processes is not discussed in this paper.

There are two essentially different approaches to the nonlinear filtering problem. The first is based on the important idea of innovations processes, originally introduced by Rode and Shannon (and Kolmogoroff) in the context of Wiener Filtering problems and later developed by Kailath and his students in the late sixties for nonlinear filtering problems. This approach reaches its culmination in the seminal paper of Fujisaki-Kallianpur-Kunita (1972). A detailed account of this approach is now available in book form, e.g., Lipster-Shiryayev (1977) and Kallianpur (1980). The second approach can be traced back to the doctoral dissertation of Mortensen (1966), Duncan (1967) and the important paper of Sawai (1969). In this approach attention is focused on the unconditioned conditional density equation, which is a bilinear stochastic differential equation, and it derives its inspiration from measure space integration as originally introduced by Kac (1951) and Ray (1954). Mathematically, this view is closely connected to the path integral formulation of Quantum Physics due to Feynman (1965). For an exposition of this analogy see Mitter (1980, 1981). A detailed account of the second viewpoint can be found in the lectures given by Kunita, Pardoux and Mitter in the CIME Lecture Notes on Nonlinear Filtering and Stochastic Control (1982) and in Hazewinkel-Wiener (1981).

2. Basic Problem Formulation

To simplify the exposition we consider the situation where all processes are scalar-valued.

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a complete probability space and let \(\mathcal{F}_t, t \geq 0\) be an increasing family of sub \(\sigma\)-fields of \(\mathcal{F}\). Let \(\xi_t\) be an \(\mathcal{F}_t\)-adapted process, considered to be the signal process and consider the observation process \(y_t\) given by

\[
y_t = \int_0^t h_s \, ds + n_t, \tag{2.1}\n\]

where \(n_t\) is an \(\mathcal{F}_t\)-Weiner process and it is assumed that \(\sigma(n_t - n_s | t \geq s)\) is independent of the past of the joint signal-observation process \(\sigma(y_u, h_s | u \leq s)\). Information about the \(\xi\)-process is contained in \(h\) which satisfies

\[
\sum_{s \leq t} |h_s|^2 \, ds < \infty \quad \forall \ t \in [0, T]. \tag{2.2}\n\]

Let \(\mathcal{F}_t^y = \sigma(y_s | s \leq t)\). Then the filtering problem consists of computing

\[
\mathbb{E}(\xi_t | \mathcal{F}_t^y) = \xi_t, \tag{2.3}\n\]

where \(\xi\) is a bounded, continuous function (indeed any function such that the conditional expectation makes sense).

3. The Innovations Method

The fundamental paper of Fujisaki-Kallianpur-Kunita proceeds as follows:

Define the innovations process:

\[
y_t = y_t - \int_0^t h_s \, ds, \tag{3.1}\n\]

\[
h_s \triangleq \mathbb{E}(h_s | \mathcal{F}_s^y), \tag{3.2}\n\]

Then it can be shown that:

1. \(y_t\) is \(\mathcal{F}_t\)-Wiener
2. \ \[v_t - v_s \geq \xi_t \] is independent of \(\mathcal{F}_t^y\).
3. In general it is not true that:

   (Innovations Property) \(\mathcal{F}_t^y = \mathcal{F}_t\) (Cirelson counter-example).

   If it is assumed that \(h_t\) and \(n_t\) are independent then innovations property holds (cf. Allinger-Mitter [1981]).

   However even without the innovations property holding, it can be proved:

   Every square-integrable \(\mathcal{F}_t\)-martingale \(m_t\) can be represented as:

\[
m_t = m_0 + \int_0^t h_s \, ds, \tag{3.3}\n\]

where \(h_s\) is jointly measurable, adapted to
\[ d_\xi^Y_t = \int_0^t E[\xi_t^Y]ds + \xi_t^Y, \]

To proceed further, let us assume that \( \xi_t^Y \) is a continuous semimartingale.

(3.1) \[ \xi_t^Y = \xi_0 + \int_0^t \xi_s^Y ds + v_t, \]

where \( v_t \) is a square integrable \( \mathcal{F}_t \)-martingale. Then

(3.2) \[ m_t = \xi_t^Y - \xi_0 - \int_0^t \xi_s^Y ds \]

is a \( \mathcal{F}_t \)-martingale and from the previous result

(3.3) \[ m_t = \int_0^t \phi_s dv_s, \]

where \( \phi_s \) can be identified as

(3.4) \[ \phi_s = \xi_s^H_s - \xi_s^N_s + D_s, \]

where \( D_s \) is a \( \mathcal{F}_s \)-predictable process and \( \int_0^t D_s ds = \langle v, n \rangle_t \), where \( \langle v, n \rangle_t \) denotes the quadratic variation.

In case \( h_\phi \) and \( n_\phi \) are independent there is an essential simplification using the innovations property.

For example, one would have the representation

(3.5) \[ \xi_t^Y = \xi_0 + \int_0^t \frac{1}{2} E(\xi_s^Y) dv_s + \int_0^t \int_0^s \frac{1}{2} E(\xi_s^Y) dv_s \]

To proceed further let us assume that \( \xi_t^Y \) is a Markov diffusion process satisfying the Ito equation

(3.6) \[ d\xi_t^Y = f(\xi_t^Y) dt + g(\xi_t^Y) d\xi_t^Y, \]

(3.7) \[ \phi_s = h(\xi_s^Y). \]

Let us also assume that \( h_\phi \) and \( n_\phi \) are independent (assumed throughout the rest of the paper). Then using the Ito differential rule for \( \phi \in C^2_t \), one gets

(3.8) \[ \phi_t^Y = \phi_0^Y + \int_0^t \phi_s^Y ds + \int_0^t \phi_s^Y dv_s, \]

where \( L \) is the generator of the diffusion process.


Let \( \Pi_t(\xi_t^Y) \) denote the conditional distribution of \( \xi_t^Y \) given \( \mathcal{F}_t \).

Let \( \phi \in C^2_t \) and denote by

(4.1) \[ \Pi_t(\phi) = \int \phi(y) \Pi_t(dy, \omega) \]

Then \( \Pi_t(\phi) \) satisfies the Nonlinear Stochastic Partial Differential Equation (Kushner-Stratonovich Equation)

(4.2) \[ \frac{d}{ds} \Pi_t(\phi) = \Pi_t(\phi) + \int_0^t \Pi_t(L\phi) ds \]

where \( L \) is the generator of the diffusion process.

5. Zakai Equation

Let \( \rho_t \) be a continuous stochastic process with values in the set of finite positive measures on \( \mathbb{R} \). Denote by

(5.1) \[ \rho_t(\phi) = \int \phi(y) \rho_t(dy). \]

Consider the equation

(5.2) \[ \rho_t(\phi) = \rho_0(\phi) + \int_0^t \rho_s \mathcal{L}_\phi ds + \int_0^t \rho_s \mathcal{N}_\phi d\xi_s \]

(weak form of Zakai equation)

Now it can be proved:

---

(1) The development here follows Kunita [1982].
If $\rho_t$ is a solution of (5.2) then $\rho_t(t) = \rho_{t}(t)$ is a solution of equation (4.2).

Moreover we have the Feynman-Kac formula:

$$\rho_t(\phi) = E_0 \left[ \phi(t) \rho_t(t) \exp \left( \int_0^t h_s \, dy_s - \frac{1}{2} \int_0^t h_s^2 \, ds \right) \right]$$

For later use (and throughout the rest of the paper) we shall consider the weak form of the Zakai equation in Stratonovich form:

$$\rho_t(\phi) = \rho_0(\phi) + \int_0^t \rho_s(\phi) \left( L - \frac{1}{2} \right) \lambda_s(\phi) \, ds + \int_0^t \rho_s(\phi) \, \lambda_s(\phi) \, dy_s$$

If the solution $\rho_t(dy)$ has a smooth density $\rho_t(y)$ it satisfies the Zakai equation:

$$\frac{d\rho_t}{dt} = \bigg( L - \frac{1}{2} \bigg) \rho_t(y) \, dt + h \rho_t(y) \, dy_t$$

where $* \text{ denotes formal adjoint.}$

$\rho_t(y)$ has the interpretation of an unnormalized density and is to be thought of as the "state" of the filter.

To compute conditional statistics, we need the state-output equation

$$\hat{\phi} = \frac{\int \phi(x) \rho_t(x,y) \, dx}{\int \rho_t(x,y) \, dx}$$

The fundamental problem of nonlinear filtering is the "invariant" study of equation (5.5). The analytic difficulty of this problem stems from the following:

(i) In most interesting situations the operator $x \rightarrow h(x)$ is unbounded.

(ii) The paths of the $y$-process are only Hölder continuous of exponent $<\frac{1}{2}$.

6. Pathwise Nonlinear Filtering

The ideas of this section are due to Clark [1978], Davis [1980], and Mitra [1980].

There is as yet no theory of nonlinear filtering where the observations are:

$$\gamma_t = h(y_t) + \eta_t$$

where $\eta_t$ is physical wide-band noise and hence smooth. Define $\gamma_t = \gamma_t$, and $\gamma_t = \eta_t$ where $\cdot$ denotes differentiation. Then (6.1) can be written as:

$$d\gamma_t = h(y_t) \, dt + d\eta_t,$$

or

$$\gamma_t = \int_0^t h(y_s) \, ds + \eta_t.$$
Equation (6.5), the pathwise filter equation, is now needed to be solved for each $y$, the observation (which can be taken to be a physical observation).

7. Existence and Uniqueness Results for the Zakai Equation.

In case the observation $h$ is bounded or linear, existence and uniqueness results for the Zakai equation has been given by PARDOUX [1982] by studying equation (5.5) directly. Existence, uniqueness and estimate of tail distributions for equation (5.5) including unbounded observations (in the scalar case all polynomial $h$) have been given by FLEMMING-MITTER [1982], using stochastic control arguments. This approach studies equation (6.5) by transforming it into a Bellman-Hamilton-Jacobi equation (cf. paper of FLEMMING this volume) using an exponential transformation and then showing that the measure constructed from $Q$ coincides with the measure given by formula (5.3).

For other literature on this problem see the bibliography in FLEMMING-MITTER [1982]. See also PARDOUX [1981] and MITTER [1982] for an interpretation of the exponential transformation in the context of nonlinear filtering. For related variational considerations see MITTER [1980], BISMUT [1981], HJIAB [1980].


How can one answer the question when two filtering problems have identical solutions? How can one decide whether the Zakai equation admits a finite-dimensional statistic?

The starting point of this analysis is the Zakai equation (5.5) in Stratonovich form. Consider the two operators:

$$
\mathscr{L}_0 = L^* - \frac{1}{2} h^2
$$

$$
\mathscr{L}_1 = h
$$

where the two operators are considered as formal differential operators on $C^0[\mathbb{R}]$. Denote by $\mathscr{F}$ the Lie algebra of operators generated by $\mathscr{L}_0$ and $\mathscr{L}_1$ under the standard bracket operation. This Lie algebra is invariant under (i) smooth change of coordinates $x \rightarrow \phi(x)$ and (ii) gauge transformations $\phi \rightarrow \phi \phi$ where $\phi$ is a $C^\infty$-function which is invertible. Indeed, if the "invariance group" of the Zakai equation is suitably defined then the above constitutes the largest invariance group of the Lie algebra $\mathscr{F}$. The insight here is that two filtering problems with isomorphic Lie algebras are likely to have the same filters. We say likely, since for a proof, analytic considerations such as the existence of a common dense set of analytic vectors must come into play. For a rigorous analysis in specific situations see OONS [1980].

By a finite-dimensional filter for the conditional statistic $\bar{\psi}(\nu_t)$ we mean a stochastic dynamical system,

$$
\dot{\psi}_t = \alpha(\psi_t) dt + \beta(\psi_t) d\nu_t
$$

where $\alpha$ and $\beta$ are smooth vector fields on some finite-dimensional smooth manifold, and a state-output equation

$$
\psi_t = \gamma(\psi_t), \quad \text{with $\gamma$ a smooth real-valued function.}
$$

The idea of studying the Lie algebra is independently due to BROCKETT [1981] (cf. the bibliography of earlier Brockett papers cited there) and MITTER [1980] (cf. the bibliography of earlier paper of Mitter cited there) and the idea of implementing the Lie algebra $\mathscr{F}$ as a Lie Algebra of vector fields is due to Brockett. The first examples of finite-dimensional filters for nonlinear filtering problems were constructed by BENES [1981] using functional integral methods. For a generalisation of Benes results see OONS-BARAS-MARCUS [1982].

In most situations the Lie algebra $\mathscr{F}$ is infinite dimensional (cf. IGUSA [1981]) and in many situations simple. If the Lie algebra $\mathscr{F}$ is infinite dimensional it does not necessarily mean that a finite-dimensional filter does not exist. For a precise result in this direction see HAZEWINKEL-MARCUS [1982], SUSSMANN [1981].


There are other topics of interest and the theory represented by the second point of view is far from complete.

(i) Asymptotic Expansions: see BLANKENSHIP-LIU-MARCUS [1982]. Much work remains to be done here.

(ii) Lower Bounds on Nonlinear Filtering: see article of BOBROWSKY-ZAKAI in HAZEWINKEL-MARCUS [1981], (and the bibliography cited there). The stochastic control interpre-
(iii) Filtering on Manifolds: see DUNCAN [1977] for a beginning.
(iv) Smoothness of densities: see KUNITA [1982].
(v) For a partial solution to the vector "cubic-sensor" problem, see DELFOUR-MITTER [1982].

REFERENCES

7. CIME Lectures on Nonlinear Filtering and Stochastic Control (1982), to be published as Springer Lecture Notes, eds. S.K. Mitter & A. Moro.
8. CIME Lectures on Nonlinear Filtering and Stochastic Control (1982), to be published as Springer Lecture Notes, ed. A. Moro.
Note on uniqueness of semigroup associated with Bellman operator

Makiko Kisho

Department of Mathematics, Kobe University
Kobe, 657, Japan

1. Introduction. As a mathematical formulation of Bellman principle, we define a nonlinear semigroup, using the value function of stochastic optimal control. [2, 4, 7] The generator of this semigroup is the Bellman operator. The purpose of this note is to consider the uniqueness of semigroup with generator of Bellman operator, appealing to results of integral solution (Benilan solution) [1] of Cauchy problem.

First we recall our nonlinear semigroup. Let Ω, be a compact convex subset of $\mathbb{R}^n$, called a control region. $B(t), t \geq 0$, denotes an n-dimensional Brownian motion on a probability space $(\Omega, F, P)$. Any $\Gamma$-valued $\sigma_t(\Omega)$-progressive measurable process is called an admissible control. $\mathcal{A}$ denotes the totality of admissible controls. For $V \in \mathcal{A}$, we consider the following controlled stochastic differential equation

\begin{equation}
X(t) = X(t), U(t) \, dB(t) + \gamma(X(t), U(t)) \, dt
\end{equation}

\begin{align}
X(0) = \eta \in \mathbb{R}^n
\end{align}

we assume that $n \times n$ symmetric non-negative definite matrix $a(x, u)$ and n-vector $\gamma(x, u)$ satisfy the following conditions (A1) and (A2).

\begin{align}
&\text{(A1)} \quad |a(x, u)| \leq b \quad \text{for any } x, u \\
&\text{(A2)} \quad |a(x, u) - a(y, u)| \leq K|x - y| + p(|u|)
\end{align}

where $b$ and $K$ are constants and $p$ continuous on $[0, +\infty)$ with $p(0) = 0$. Then there exists a unique solution $X(t) = X(t; x, U)$, called the response for $U$. The value function is defined as

\begin{equation}
\gamma(t, x, \varphi) = \sup_{U \in \mathcal{A}} \mathbb{E}_x \left[ e^{-\frac{t}{\tau}} \int_0^\tau a(X(s), U(s)) \, ds + f(X(s), U(s)) \, ds \right]
\end{equation}