JB 79-10

Entitled:

"Nonlinear Filtering of Diffusion Processes:
A Generic Example"

Authors:

with G. L. Blankenship

Conference:

18th IEEE Conference on Decision and Control
Ft. Lauderdale, Florida
December 1979

NONLINEAR FILTERING OF DIFFUSION PROCESSES: A GENERIC EXAMPLE*

G. L. Blankenship and J. S. Baras
Electrical Engineering Department
University of Maryland
College Park, Maryland 20742

Abstract

An analytical procedure is described for treating a generic nonlinear filtering problem.

1. Problem Statement

In this paper we examine the "nonlinear filtering problem" for the apparently simple system

$$dx(t) = \frac{1}{2}\, x(t)dt + x(t)dw(t) \qquad (1.1a)$$

$$dy(t) = x(t)dt + dv(t) \qquad (1.1b)$$

$$x(0) = x_o,\ y(0) = 0,\ 0 \le t \le T$$

where x_o is a log normal random variable independent of the standard, real-valued Wiener processes w and v which are in turn mutually independent. The filtering problem is: Compute the probability density of x(t), $0 \le t \le T$, conditioned on $Y_t = \sigma\{y(s), s \le t\}$ the σ-algebra generated by y.

While this problem may appear overly specific, we feel that it exhibits most of the essential features (difficulties) of the corresponding filtering problem for the (vector) system

$$dx(t) = Ax(t)dt + \sum_{i=1}^{m} B_i x(t)dw_i(t)$$
$$dy(t) = Cx(t)dt + dv(t) \qquad (1.2)$$

involving state dependent noise and additive observation noise. We shall highlight those aspects of our analysis of (1.1) which are generic to the family (1.2).

Our analysis of (1.1) involves the following four steps:

(i) Introduce the new coordinates

$$z(t) = \ell n\, x(t) \qquad (1.3)$$

Then

$$dx(t) = dw(t)$$
$$dy(t) = e^{z(t)}dt + dv(t) \qquad (1.4)$$
$$z(0) = z_o = \ell n x_o,\ y(0) = 0,\ 0 \le t \le T.$$

We regard the estimation problem for (1.4) as equivalent to that for (1.1).

(ii) The conditional density p of z(t) given Y_t satisfies

$$p(t,z) = u(t,z) / [\int_{-\infty}^{\infty} u(t,\xi)d\xi] \qquad (1.5)$$

where the "unnormalized conditional density" u(t,z) satisfies (formally) the linear stochastic PDE

$$du(t,z) = \frac{1}{2}\frac{\partial^2}{\partial z^2}u(t,z)dt + e^z u(t,z)dy(t)$$

$$u(0,z) = p_o(z) = \text{density of } z_0, \text{(a normal density)}$$

$$0 \le t \le T.$$

The existence and uniqueness of solutions to (1.6) along with possible structural forms for the solutions are discussed in section 2.

(iii) Associated with (1.6) is the "backwards" stochastic PDE

$$dv(t,z) + \frac{1}{2}\frac{\partial^2}{\partial z^2} v(t,z) + e^z v(t,z)dy(t) = 0 \qquad (1.7)$$

$$v(T,z) = f,\ 0 \le t \le T.$$

which is adjoint to (1.6) in the sense that the trajectories of $\langle u(t), v(t)\rangle$ are constants. The solutions of (1.7) may be represented by a stochastic function space integral which generalizes the classical Feynman-Kac formula for deterministic, second order, parabolic PDE's.

(iv) Using a parabolic interpolation for Wiener paths and an appropriate quadrature formula, the stochastic Feynman-Kac formula is simply and accurately approximated by an n-fold integral. This provides a convenient basis for numerical solution of (1.6), and so, the filtering problem for (1.4) and (1.1). This approximation is described in section 3.

Remarks

(1) In the context of the filtering problem for

$$dx(t) = f(x(t))dt + g(x(t))dw(t) \qquad (1.8a)$$

$$dy(t) = h(x(t))dt + dv(t) \qquad (1.8b)$$

our step (i) corresponds to a transformation on (1.8a) which converts x(t) into a Wiener process (when this is possible). The transformation may be a simple coordinate change as above, or one implied by a Girsanov-Cameron-Martin transformation on the measure defined by the x process. The purpose of the transformation is to simplify the measure used in the path integral expression, or equivalently the infinitesimal generator.

(2) An evolution equation for the unnormalized conditional density of x(t) given Y_t in (1.8) was derived under strong conditions by Zakai [1] (and independently by Mortenson).

*This work was partially supported by O.N.R. under contract N00014-79-C-0808.

Pardoux [2][3] has proved existence and uniqueness of solutions of the Zakai equation under comparatively weak conditions which, however, require the function $h(\cdot)$ in (1.8b) to be bounded. Thus, Pardoux's results do not apply to (1.1) or (1.4). It is surprizingly difficult to prove existence and uniqueness in the latter case - see section 2.

(3) Since the infinitesimal generator of the z process in (1.4) is the Laplacean, the duality between (1.6) and (1.7) is very simple. The Zakai equation for the original system (1.1) is

$$d\phi(t,x)=\left(\frac{1}{2}\frac{\partial^2}{\partial x^2}[x^2\phi(t,x)]-\frac{1}{2}\frac{\partial}{\partial x}[x\phi(t,x)]\right)dt$$

$$+ x\phi(t,x)\, dy(t)$$

and it is evidently not self-dual. The coordinate change accomplishes this and transfers the complexity in the Zakai equation from the differential operator in the drift term into the coefficient in the stochastic term.

(4) Since the differential operator in (1.7) is the Laplacean, the measure in the function space integral which provides its solution (see equation (3.4)) is Wiener measure. This permits use of the simple interpolation and quadrature formulas refered to in step (iv). Again we feel that this property of the example is generic to the family of nonlinear filtering problems for which steps (i) - (iii) are possible.

(5) The stochastic Feynman-Kac formula for the solution of backwards stochastic PDE's was stated formally by Kushner in [4] and rigorously by Pardoux in [2]. In [4] Kushner gives a numerical procedure for evaluating the formula which is completely different from our approach.

2. Existence, Uniqueness, and Representations

Although (1.6) may appear innocent, a satisfactory existence and uniqueness result for its solutions is unavailable at present. The unbounded function e^z in the second term causes Pardoux's approach to fail. That is, if one introduces a weighted L_2 space, $L_2(d\mu)$, where

$$d\mu(z) = \exp(|z|)dz, \tag{2.1}$$

(the type cannot be higher than 2 since $p_o(z)$ in (1.6) is Gaussian), and defines

$$V = \{f : f\in L^2(d\mu),\ Df \in L^2\} \tag{2.2}$$

where D is the distributional derivative; then it is easy to show that the operator B = multiplication by e^z satisfies $B\in\mathcal{L}(V,H)$ with $H = L_2$. However, the coercivity property required by Pardoux does not hold. To date, other choices of V and H have also failed.

A different approach seems more promising. It is clear that if a solution exists, then it should be in L^1 with respect to z. It is also clear that we can assume y to be a Wiener process, since we can apply Girsanov's transformation and change measures to achieve this [2]. Let us denote by $\tilde{P}$ the new measure. This will not affect existence and uniqueness. Now if we let A be the operator $\partial^2/\partial z^2$, it is well known that it generates the semigroup

$$(e^{At}f)(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}}\, e^{-(z-a)^2/2t} f(a)\, da \tag{2.3}$$

which has a smoothing action. Since the initial condition is in $L_1(d\mu) \cap L_2(d\mu)$ we are looking for a weight different from (2.1), let us call it p(z), with the following propertities:

(a) If $V = \{f:\ f \in L_2(d\mu_p),\ Df \in L_2\}$ where $d\mu_p(z) = p(z)dz$, then $\exp(At)V \subset V$.

(b) Gaussian densities belong to V.

(c) There exists a Hilbert space H with $V \subset H \subseteq L_2$, with V dense in H.

(d) e^{At} is strongly continuous on H.

(e) The operator B = multiplication by e^z is closed on H and $\mathcal{D}(A) \subseteq \mathcal{D}(B) \subset V$.

(f) for each $t > 0$ there exists $K(t) < \infty$ such that

$$||Be^{At}f||_H \leq K(t)||f||_H \quad , \quad f \in \mathcal{D}(A)$$

with $\int_0^1 K(t)dt$ finite.

If such a weight exists, we can show that:

(i) (1.6) has a unique solution in $L_2(\Omega;C(0,T;H))$ for every finite T.

(ii) The solution of (1.6) is adapted to $Y_t=\sigma\{y(s), s\leq t\}$.

Furthermore, we can show, via a generalized Ito rule for (1.6) that this solution can be represented as the "stochastic" Feynman-Kac path integral

$$u(t,z) = E_z\{\exp\left[\int_0^t e^{x(s)}dy(s) - \frac{1}{2}\int_0^t e^{2x(s)}ds\right] \cdot p_o(x(t))\} \tag{2.4}$$

where x(s) is a standard Wiener process and E_z is the function space Wiener integral over all paths starting from z.

3. Numerical Evaluation of the Conditional Density.

From [2][3] the Zakai equation has a dual backward stochastic PDE. In our case the duality between (1.7) and (1.6) is particularly simple; it implies that $\langle u(t),v(t)\rangle$ is constant where $\langle\cdot,\cdot\rangle$ is the $L_2(R)$ inner product. The solution of the backward equation has the form

$$v(t,z)=\tilde{E}[f(z(T))Y_T^t \,|\, \mathcal{F}_T^t,\ z(t) = z] \tag{3.1}$$

Here expectation is with respect to Wiener measure $\tilde{P}$, $\mathcal{F}_T^t = \sigma\{y(s) - y(t),\ t \leq s \leq T\}$, and

$$Y_T^t = \exp\left[\int_t^T e^{z(s)}dy(s) - \frac{1}{2}\int_t^T e^{2z(s)}ds\right] \tag{3.2}$$

If $u(t,z)$ is the solution to the forward equation (1.6) then the duality implies that

$$\langle u(T),v(T)\rangle = \langle u(0),v(0)\rangle = \int_{R^n} p_o(z)\tilde{E}\{f(z(T))\gamma_T^0 \,|\mathcal{F}_T^0,\ z(0)=z\} \qquad (3.3)$$

and from this that for any $f \in L_1$

$$\langle u(T),f\rangle = \tilde{E}\{f(z(T))\ \gamma_T^0\ |\mathcal{F}_T^0\} \qquad (3.4)$$

Since T is arbitrary, this provides a solution to the filtering problem. Of course, one is left with the problem of computing the function space integral in (3.4). Note that (3.4) is the integrated version (with respect to z) of (2.4), and that any computational scheme applied to (3.4) can be equally well applied to (2.4).

To make contact with the literature in mathematical physics, we rewrite the integral (3.4) in the form

$$I = \int_C F(z,y)\ dW(z) \qquad (3.5)$$

where $C = C([0,T])$, dW is Wiener measure on C, and

$$F(z,y) = f(z(t))\exp[\int_0^t e^{z(s)}dy(s) - \frac{1}{2}\int_0^t e^{2z(s)}ds\ ] \qquad (3.6)$$

The Wiener integral I in (3.5) is defined as the limit

$$I = \lim_{\substack{\max|t_j-t_{j-1}|\to 0\\ 0\le j\le n}} \int_R \overset{(n)}{\cdots}\int_R dx_1\cdots dx_n F(z_{t,x},y) \prod_{j=1}^n \frac{\exp[-(x_j-x_{j-1})^2/2(t_j-t_{j-1})]}{[2\pi(t_j-t_{j-1})]^{n/2}} \qquad (3.7)$$

where $0 < t_1 < t_2 < \cdots < t_n = T$ and $z_{t,x}$ is a polygonal function of t on [0,T] that passes through x_j at t_j, $j = 1,2,\ldots,n$.

Integrals of the form (3.5) with no y dependence arise in physics as representations of the solutions of the deterministic Schrodinger equation

$$\frac{\partial}{\partial t}u(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}u(t,x) + V(x)u(t,x) \qquad (3.8)$$

$$u(0,x) = f(x)$$

That is,

$$u(t,x) = E_x[\exp(\int_0^t V(x(s))ds)\ f(x(t))] \qquad (3.9)$$

where x is a path of Brownian motion and E_x is the Wiener (function space) integral over all paths starting at x in R. There has been a substantial effort devoted to computing Wiener integrals, both analytically and numerically [6]-[8]. For example, Chorin [8] has shown that if $G:R\to R$ and $V:R\to R$ are smooth, then

$$\int_C G[\int_0^T V(x(s))ds]\ dW = \pi^{-n/2}\int_{R^n} [G(\sum_{i=1}^n \frac{1}{n}V(x_{i-1} + vT/(2n)^{1/2}))]\ \cdot$$

$$\cdot\exp(-u_1^2 - u_2^2 - \cdots -u_{n-1}^2 - v^2)\ du_1\cdots du_{n-1}dv + O(n^{-2})$$

where $x_{i-1} = (u_1 + \cdots + u_{i-1})T/\sqrt{n}$. The $O(n^{-2})$ error makes the approximation very accurate.

In the case when $dy(t) = m(t)dt + r(t)dw(t)$ is an Ito process with m,r square integrable, we have been able to show that

$$\int_C \exp[\int_0^T V(x(s))dy(s)\ ]\ dW = \pi^{-n/2}\int_{R^n} [\exp(\sum_{i=1}^n V(x_{i-1} + vT/\sqrt{2n})\ \Delta y_{i-1})] \cdot\exp(u_1^2 - u_2^2 - \cdots - u_{n-1}^2 - v^2)du_1\cdots du_{n-1}dv + e_n \qquad (3.11)$$

where x_{i-1} is as above and $\Delta y_{i-1} = y(iT/n) - y((i-1)T/n)$. (Accurate formulas for computing Δy_i are available in [9], among other sources.) The error e_n satisfies

$$[E_y(e_n^2)]^{1/2} = O(n^{-3/2}),$$

where E_y is expectation over the distribution of y. It is larger than that in (3.10) since $(E\Delta y^2)^{1/2} \approx n^{-1/2}$ rather than n^{-1}.

Using the approximation (3.11) on (3.5) gives the following result:

$$I = \tilde{E}\,[f(z(t))\ \gamma_t^0 |\ \mathcal{F}_t^0\ ]$$

$$= \int_C [\exp(\int_0^t e^{x(s)}dy(s) - \frac{1}{2}\int_0^t e^{2x(s)})f(x(t))]dW$$

$$= \pi^{-n/2}\int_{R^n} [\exp(\sum_{i=1}^n \exp(x_{i-1} + vt/\sqrt{2n})\ \Delta y_{i-1} - \frac{1}{2}\sum_{i=1}^n \exp(2(x_{i-1} + vt/\sqrt{2n}))(t/n))$$

$$\cdot f(x_{n-1} + vt/\sqrt{2n})]\exp(-u_1^2 - \cdots - u_{n-1}^2 - v^2) \cdot du_1\cdots du_{n-1}dv + O(n^{-3/2})$$

Efforts to make this formula recursive in $y(it/n)$ have not yet been successful. The simplicity and accuracy of the approximation are its chief advantages. An expansion for the conditional density $u(t,z)$ in (2.4) similar to (3.12) may be obtained by substititing p_o for f and interpolating over paths x starting at $x(0)=z$.

References.

1. M.Zakai, Z.Wahr. Verw. Geb.,LL(1969), p.230-243.
2. E.Pardoux, to appear in Stochastics.
3. _____, this conference.
4. H.Kushner, Prob. Methods Approx. Stoc. Con., 1977.
5. M.Kac, Trans. Amer.Math.Soc., 65(1949), p. 1-13.
6. R.Cameron, Duke Math. J., 18(1951), p. 111-130.
7. C.Morette, Phys. Rev., 81(1951), p. 848-852.
8. A.Chorin, Math. Comp., 27(1973), p. 1-15.
9. N.Rao,et.al., SIAM J. Control, 12(1974), p. 124-139.