LIDS-P-2223

SERIES EXPANSIONS FOR NONLINEAR FILTERS[1]

V. Borkar
Department of Electrical Engineering
Indian Institute of Science
Bangalore –560 012
INDIA

and

S.K. Mitter
Department of Electrical Engineering and Computer Science
and
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139
USA

December 1993

Abstract

We develop a series expansion for the unnormalized conditional estimate of nonlinear filtering. The expansion is in terms of a specific complete orthonormal system in the Hilbert space of square-integrable random variables on Wiener space. This expansion decouples the observation dependence from the dependence on the state dynamics and has an appealing recursive feature which suggests its use for developing approximate filters.

Key words: Nonlinear filtering, unnormalized estimate, series expansion, approximate filters

1 Introduction

Recently Mikulevicius and Rozovskii [4] proposed a spectral approach to nonlinear filtering. In this, they expand the unnormalized estimate of a function of the state into a Fourier series over a complete orthonormal system (CONS) of square integrable functionals of the observation process. A truncation of this series then serves as an approximate filter. The specific CONS used is the Wick polynomials of stochastic integrals of a CONS for $L_2[0, t]$ ($[0, t]$ being the observation interval) with respect to the observation process.

In this note, we follow exactly the same program, but with a different choice of CONS. The advantages are very significant. As in [4], the series expansion decouples the observation-dependent part of the filter (now captured in the aforementioned CONS) and the state dynamics-dependent part (which now becomes a deterministic problem). Our expressions for these are, however, much

[1]Research supported by the US Army Research Office under grant ARO DAAL03-92-G-0115 (Center for Intelligent Control Systems).

simpler than those of [4]. Furthermore, [4] requires a separate CONS of $L_2[0,t]$ for each t, whereas we fix a $T > 0$ and a CONS for $L_2[0,T]$ that works for all $t \in [0,T]$. There is also a simple scheme for repeating this procedure for successive time intervals of length T.

The paper is organized as follows: The next section introduces the estimation problem. Section III introduces our CONS for the Hilbert space of square integrable random variables on Wiener space. The series expansion in terms of this CONS is derived in Section IV. Section V discusses some computational issues.

2 Notation

Our signal process and observation process will be respectively, a d-dimensional diffusion process $X(\cdot) = [X_1(\cdot), \cdots, X_d(\cdot)]^T$ and an m-dimensional process $Y(\cdot) = [Y_1(\cdot), \cdots, Y_m(\cdot)]^T$, satisfying the stochastic differential equations

$$\begin{aligned} X(t) &= X_0 + \textstyle\int_0^t m(X(s))ds + \int_0^t \sigma(X(s))dW(s), \\ Y(t) &= \textstyle\int_0^t h(X(s))ds + W'(t), \end{aligned}$$

for $t \in [0,T]$. Here $m(\cdot) = [m_1(\cdot), \cdots, m_d(\cdot)]^T : \mathcal{R}^d \to \mathcal{R}^d, \sigma(\cdot) = [[\sigma_{ij}(\cdot)]]_{1 \leq i,j \leq d} : \mathcal{R}^d \to \mathcal{R}^{d \times d}$, $h(\cdot) : [h_1(\cdot), \cdots, h_d(\cdot)]^T : \mathcal{R}^d \to \mathcal{R}^m$ are bounded Lipschitz (for simplicity, this can be relaxed) and the least eigenvalue of $\sigma(\cdot)\sigma^T(\cdot)$ is uniformly bounded away from zero. $W(\cdot)$, $W'(\cdot)$ are respectively d and m-dimensional standard Wiener processes, with $(W(\cdot), W'(\cdot), X_0)$ independent. The law of X_0 will be denoted by π_0.

For $t \in [0,T]$, let $\mathcal{F}_t =$ the completion of $\bigcap_{s>t} \sigma(Y(y), y \leq s)$ with respect to the underlying probability measure P. The problem of nonlinear filtering is to compute $E[f(X(t))/\mathcal{F}_t]$ for $t \in [0,T]$, $f \in C_b(\mathcal{R}^d)$. An equivalent ('Zakai') formulation is as follows: For $t \geq 0$, define

$$\Gamma_t = \exp\left(\int_0^t < h(X(s)), dY(s) > -\frac{1}{2}\int_0^t ||h(X(s))||^2 ds\right).$$

Define a new probability measure P_0 on the underlying probability space by $P(A) = \int_A \Gamma_t dP_0$ for $A \in \sigma(X(y), Y(y), y \leq t)$. Under $P_0, Y(\cdot)$ is an m-dimensional standard Wiener process independent of $X(\cdot), W(\cdot)$. Letting $E_0[\cdot]$ denote the expectation/conditional expectation under P_0, one has

$$E[f(X(t))/\mathcal{F}_t] = E_0[f(X(t))\Gamma_t/\mathcal{F}_t]/E_0[\Gamma_t/\mathcal{F}_t], t \in [0,T].$$

Thus it suffices to compute the 'unnormalized' estimate $E_0[f(X(t))\Gamma_t/\mathcal{F}_t]$ for $t \in [0,T]$, $f \in C_b(\mathcal{R}^d)$. This can be viewed as $\int f dp(t)$ where $p(t)$ is a random finite nonnegative measure on $\mathcal{R}^d$, called the unnormalized conditional law of $X(t)$ given $\mathcal{F}_t$. Our aim is to develop a 'series expansion' for this quantity. For a more detailed account of nonlinear filtering, see [1].

We conclude the section with some notation for later use. For a finite signed measure μ on $\mathcal{R}^d$ and $f \in C_b(\mathcal{R}^d)$, let $\mu(f) = \int f d\mu$. For $x = [x_1, \cdots, x_d]$, let

$$A = \frac{1}{2}\sum_{ijk} \sigma_{ik}(x)\sigma_{jk}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_i m_i(x)\frac{\partial}{\partial x_i}.$$

Let $P' =$ the Wiener measure on $\Omega = C([0,T]; \mathcal{R}^m)$ endowned with its Borel σ-field $\mathcal{B}$ and $E'[\cdot]$ the expectation w.r.t. P'. Finally, let $\{e_i(\cdot)\}$ be a CONS for $L_2[0,T]$ and $\mathcal{L}_2 = L_2(\Omega, \mathcal{B}, P')$.

3 A CONS for $\mathcal{L}_2$.

In this section we view $Y(\cdot)$ as having been canonically realized on $(\Omega, \mathcal{B}, P')$ (i.e., $Y(\omega,t) = \omega(t)$ for $\omega \in \Omega$, $t \in [0,T]$). For $1 \le j \le m$, $i \ge 1$, $t \in [0,T]$, let

$$\Lambda_{ij}(t) = \exp(\int_0^t e_i(s)dY_j(s) - \frac{1}{2}\int_0^t e_i^2(s)ds).$$

Then it is easy to see that $(\Lambda_{ij}(t), \mathcal{F}_t)$, $t \in [0,T]$, are martingales. Let

$$\xi_0 = 1, \xi_{ij} = (\Lambda_{ij}(T) - 1)/\sqrt{e-1}, i \ge 1, 1 \le j \le m.$$

Theorem 3.1 *$\{\xi_0, \xi_{ij}, i \ge 1, 1 \le j \le m\}$ is a CONS for $\mathcal{L}_2$.*

Proof. The proof is in two parts.

Proof of orthonormality: Clearly $E'[\xi_0^2] = 1$ and

$$E'[\xi_{ij}\xi_0] = (E'[\Lambda_{ij}(T)] - 1)/\sqrt{e-1} = 0 \quad \forall\, i,j$$

by the martingale property of $\Lambda_{ij}(\cdot)$. For $i,k \ge 1, 1 \le j, \ell \le m$,

$$E'[\xi_{ij}\xi_{k\ell}] = (c_{ijk\ell} - 1)/(e-1)$$

where

$$C_{ijk\ell} = E'[\exp(\int_0^T e_i(s)dY_j(s) + \int_0^T e_k(s)dY_\ell(s) - \frac{1}{2}\int_0^T (e_i^2(s) + e_k^2(s))ds)].$$

For $j \ne \ell, Y_j(\cdot), Y_\ell(\cdot)$ are independent. Thus

$$C_{ijk\ell} = E'[\Lambda_{ij}(T)]E'[\Lambda_{k\ell}(T)] = 1.$$

Let $j = \ell$, $i \ne k$. Then, since $\int_0^T e_i(s)e_k(s)ds = 0$, we have

$$C_{ijkj} = E'[\exp(\int_0^T (e_i(s) + e_k(s))dY_j(s) - \frac{1}{2}\int_0^T (e_i(s) + e_k(s))^2 ds)] = 1.$$

In either case $E'[\xi_{ij}\xi_{k\ell}] = 0$. If $i = j, k = \ell$, then

$$\begin{aligned} C_{ijij} &= E'[\exp(2\int_0^T e_i(s)dY_j(s) - \int_0^T e_i^2(s)ds)] \\ &= eE'[\exp(\int_0^T (2e_i(s))dY_j(s) - \frac{1}{2}\int_0^T (2e_i(s))^2 ds)] \\ &= e. \end{aligned}$$

where we use the facts that $\int_0^T e_i^2(s)ds = 1$ and the expectation of the exponential martingale in question is 1. Thus

$$E'[\xi_{ij}^2] = (e-1)/(e-1) = 1.$$

Proof of completeness: Let $Z_{ij} = \int_0^T e_i(s)dY_j(s)$, $i \geq 1$, $1 \leq j \leq m$. Then

$$Y_j(t) = \sum_{i=1}^{\infty} z_{ij} \int_0^t e_i(s)ds, 1 \leq i \leq m,$$

where the right hand side converges in mean square (in fact, uniformly in t, a.s. [2]). Thus $\sigma(Y(t), t \in [0,T]) = \sigma(z_{ij}, 1 \leq j \leq m, i \geq 1)$. But

$$\xi_{ij} = (\exp(z_{ij} - \frac{1}{2}) - 1)/\sqrt{e-1},$$

implying

$$z_{ij} = \frac{1}{2} + \ell n(\sqrt{e-1}\xi_{ij} + 1)$$

$\forall\ i,j$. Hence $\sigma(\xi_0, \xi_{ij}, i \geq 1, 1 \leq j \leq m) = \sigma(Y(t), t \in [0,T])$. Let $Z \in \mathcal{L}_2$ be orthogonal to $\{\xi_0, \xi_{ij}, i \geq 1, 1 \leq j \leq m\}$. Then $E'[Z\xi_0] = E'[Z\xi_{ij}] = 0\ \forall\ i,j$, implying

$$E'[Z/\xi_0, \xi_{ij}, i \geq 1, 1 \leq j \leq m] = 0 = E'[Z/Y(t), t \in [0,T]] = Z.$$

Thus $\{\xi_0, \xi_{ij}, i \geq 1, 1 \leq j \leq m\}$ is complete. ∎

4 Series Expansion for the Unnormalized Estimate

By Theorem 3.1, we may write $p(t)(f), t \in [0,T], f \in C_b(\mathcal{R}^d)$ as

$$p(t)(f) = g_0(t)\xi_0 + \sum_{i=1}^{\infty} \sum_{j=1}^{m} g_{ij}(t)\xi_{ij} \tag{4.1}$$

where $g_0(t) = E_0[f(X(t))]$ and $g_{ij}(t) = E_0[\xi_{ij}p(t)(f)]\ \forall\ i,j$. Fix i,j and define a deterministic nonnegative measure-valued process $\eta_t^{ij}, t \in [0,T]$, by

$$\int q d\eta_t^{ij} = E_0[q(X(t))\Gamma_t\Lambda_{ij}(t)], t \in [0,T], q \in C_b(\mathcal{R}^d), 1 \leq j \leq m, i \geq 1.$$

Let $C_b^2(\mathcal{R}^d)$ = bounded twice continuously differentiable functions $\mathcal{R}^d \to \mathcal{R}$ with bounded first and second derivatives.

Lemma 4.1 *For $q \in C_b^2(\mathcal{R}^d)$, $\eta_t^{ij}(q)$, $t \in [0,T]$, satisfies*

$$\eta_t^{ij}(q) = \pi_0(q) + \int_0^t \eta_s^{ij}(Aq)ds + \int_0^t \eta_s^{ij}(h_j q)e_i(s)ds.$$

Proof. This follows easily on applying the Ito differentiation rule to $q(X(t))\Gamma_t\Lambda_{ij}(t)$ and taking expectations. ∎

Corollary 4.1 *For $t > 0, \eta_t^{ij}$ has a density $\varphi^{ij}(t,\cdot)$ with respect to the Lebesgue measure on $\mathcal{R}^d$. Furthermore, the latter satisfies the parabolic p.d.e.*

$$\frac{\partial \varphi^{ij}}{\partial t}(t,x) = A^* \varphi^{ij}(t,x) + \varphi^{ij}(t,x) h_j(x) e_i(t), t \in [0,T], \tag{4.2}$$

$$\lim_{t \downarrow 0} \int q(x)\varphi(t,x)dx = \pi_0(q) \quad \forall\, q \in C_b(\mathcal{R}^d), \tag{4.2'}$$

where A^* is the formal adjoint of A.

Proof. Under our hypotheses on m, σ, this follows from standard p.d.e. theory [3] and Lemma 4.1. ∎

Remark 4.1 *The above solution of (4.2), (4.2') is unique in a class of functions satisfying suitable regularity and growth conditions [3]. We omit the details.*

Let $\tilde{\varphi}^{ij}(t,y,x), t \in [0,T], x, y \in \mathcal{R}^d$, denote the solution of (4.2) with initial conditions

$$\lim_{t \downarrow 0} \int q(x)\tilde{\varphi}^{ij}(t,y,x)dx = q(y) \quad \forall\, q \in C_b(\mathcal{R}^d)$$

(i.e., its 'fundamental solution' [3]). Then

$$\varphi^{ij}(t,x) = \pi_0(\tilde{\varphi}^{ij}(t,\cdot,x)).$$

Standard p.d.e theory [3] also ensures a transition density $\tilde{p}(t,y,x), t > 0, x, y \in \mathcal{R}^d$, for the homogeneous Markov process $X(\cdot)$. Let $p(t,x) = \pi_0(\tilde{p}(t,\cdot,x))$ denote the probability density of $X(t)$ for $t > 0$.

Theorem 4.1 $p(t)(f) = \int f(x)\psi(t,x)dx$, *where*

$$\psi(t,x) = p(t,x) + \sum_{i=1}^{\infty} \sum_{j=1}^{m} \frac{\xi_{ij}}{\sqrt{e-1}} (\varphi^{ij}(t,x) - p(t,x)), t > 0, x \in \mathcal{R}^d.$$

Proof. For $i \geq 1, 1 \leq j \leq m$,

$$g_{ij}(t) = E_0[\xi_{ij} p(t)(f)] = (\int f d\eta_t^{ij} - E[f(X(t))])/\sqrt{e-1}.$$

The claim follows from (4.1) and the preceding corollary. ∎

5 Computational Aspects

The p.d.e.s for $\varphi^{ij}, \tilde{\varphi}^{ij}$ are deterministic. In particular, $\tilde{\varphi}^{ij}(t,y,x)$, $i \geq 1$, $1 \leq j \leq m$, can be computed off-line for $t \in [0,T], x, y \in \mathcal{R}^d$ and stored. The 'on-line' computation then requires computing the series

$$p(t)(f) = \int f(x)p(t,x)dx + \sum_{i=1}^{\infty} \sum_{j=1}^{m} \frac{\xi_{ij}}{\sqrt{e-1}} \int f(x)(\psi^{ij}(t,x) - p(t,x))dx,$$

for $t > 0$. Here $\{\xi_{ij}\}$ are the only observation-dependent terms. These depend on the entire trajectory $Y(s)$, $s \in [0,T]$. We may, however, take conditional expectation with respect to $\mathcal{F}_t$ on both sides and use the martingale property of $\Lambda_{ij}(\cdot)$ to obtain

$$p(t)(f) = \int f(x)p(t,x)dx + \sum_{i=1}^{\infty} \sum_{j=1}^{m} \frac{\tilde{\xi}_{ij}(t)}{\sqrt{e-1}} \int f(x)(\varphi^{ij}(t,x) - p(t,x))dx,$$

where

$$\tilde{\xi}_{ij}(t) = E_0[\xi_{ij}/\mathcal{F}_t] = (\Lambda_{ij}(t) - 1)/\sqrt{e-1}, t \geq 0.$$

By Ito's differentiation rule,

$$\begin{aligned} d\tilde{\xi}_{ij}(t) &= (e-1)^{-1/2}\Lambda_{ij}(t)e_i(t)dY_j(t) \\ &= ((e-1)^{-1/2} + \tilde{\xi}_{ij}(t))e_i(t)dY_j(t), t \in [0,T], \end{aligned}$$

with $\tilde{\xi}_{ij}(0) = 0$. The $\tilde{\xi}_{ij}(\cdot)$ can be recursively computed on-line. Alternatively, we may first compute $\Lambda_{ij}(\cdot)$ recursively by

$$d\Lambda_{ij}(t) = \Lambda_{ij}(t)e_i(t)dY_j(t), \Lambda_{ij}(0) = 1,$$

and then $\tilde{\xi}_{ij}(\cdot)$ in terms of it. If $\{e_i(\cdot)\}$ are continuously differentiable, this can be arranged so as to avoid any stochastic integration, because

$$\begin{aligned} \Lambda_{ij}(t) &= \exp\left(\int_0^t e_i(s)dY_j(s) - \frac{1}{2}\int_0^t e_i^2(s)ds\right) \\ &= \exp\left(e_i(t)Y_j(t) - \int_0^t Y_j(s)\dot{e}_i(s)ds - \frac{1}{2}\int_0^t e_i^2(s)ds\right), t \in [0,T]. \end{aligned}$$

This analysis was confined to the time interval $[0,T]$. The same procedure can be repeated on the successive time intervals $[kT,(k+1)T]$, $k \geq 1$, by shifting the time orgin to kT and replacing $\pi_0(q)$ in (4.2′) by $p(kT)(q)$ available from the preceding interval.

A natural approximation for $p(t)(f)$ in practice would be the truncated series

$$\int f(x)p(t,x)dx + \sum_{i=1}^{N} \sum_{j=1}^{m} \frac{\tilde{\xi}_{ij}(t)}{\sqrt{e-1}} \int f(x)(\varphi^{ij}(t,x) - p(t,x))dx.$$

for sufficiently large $N \geq 1$. The truncation error is

$$\varsigma(t) = \sum_{i=N+1}^{\infty} \sum_{j=1}^{m} \frac{\tilde{\xi}_{ij}(t)}{\sqrt{e-1}} \int f(x)(\varphi^{ij}(t,x) - p(t,x))dx.$$

It would be interesting to obtain good analytic bounds on $E[\varsigma(t)^2]$.

References

[1] Davis, M.H.A., Marcus, S.I., An introduction to nonlinear filtering, in 'Stochastic Systems: The Mathematics of Filtering and Identification and Applications' (M. Hazewinkel, J.C. Willems, eds.) *Proc. of the NATO Advanced Study Institute at Les Arcs*, June 1980, D. Reidel, Dordrecht, 1981, pp. 53–76.

[2] Ito, K., Nisio, M., On the convergence of sums of independent Banach space valued random variables, *Osaka Math.* J.5 (1968), pp. 35–48.

[3] Ladyzenskaja, O.A., Ural'seva, N.N.; Solonnikov, V.A., Linear and quasilinear equations of parabolic type, *Trans. Math. Monographs* 23, A.M.S., Providence, 1968.

[4] Mikulevicius, R.; Rozovskii, B.L., Nonlinear filtering revisited: a spectral approach, Center for Applied Math. Sciences, Report No. CAMS 93-13, Univ. of Southern California, June 1993.