The paper presents some recently obtained results on the nonlinear filtering problem for infinite dimensional processes. The optimal filter is obtained as the unique solution of certain measure valued equations. Robustness properties - both pathwise and statistical - are given and a preliminary result shows consistency with the stochastic calculus theory. Applications to random fields and models of voltage potential in neurophysiology are briefly discussed.
THE FILTERING PROBLEM FOR INFINITE DIMENSIONAL STOCHASTIC PROCESSES

by

G. Kallianpur

and

R.L. Karandikar

Technical Report No. 175

January 1967
THE FILTERING PROBLEM FOR INFINITE DIMENSIONAL STOCHASTIC PROCESSES

G. Kallianpur
University of North Carolina, Chapel Hill

and

R.L. Karandikar
Indian Statistical Institute, New Delhi
THE FILTERING PROBLEM FOR INFINITE DIMENSIONAL STOCHASTIC PROCESSES

G. Kallianpur
University of North Carolina, Chapel Hill
and
R.L. Karandikar
Indian Statistical Institute, New Delhi

Abstract

The paper presents some recently obtained results on the nonlinear filtering problem for infinite dimensional processes. The optimal filter is obtained as the unique solution of certain measure valued equations. Robustness properties — both pathwise and statistical — are given and a preliminary result shows consistency with the stochastic calculus theory. Applications to random fields and models of voltage potential in neurophysiology are briefly discussed.

1. Introduction

The finitely additive white noise approach to filtering, smoothing and prediction developed in Kallianpur and Karandikar [3,4] is particularly useful when the stochastic processes involved take values in infinite dimensional spaces. Since there is no natural measure (such as Lebesgue measure) in infinite dimensional spaces, the optimal filter is given by a measure valued differential equation in which the observed process occurs as a parameter. The existence and uniqueness of solution has been established in [3] for the filtering problem to which we confine ourselves in this paper. However, the result in [3] is not applicable to the most general observation model and, in particular, does not cover linear filtering. We will first state a version of the filtering result that is a significant improvement over the theorems obtained in [3] and then show how it can be applied to the case when the system process is a Banach or Hilbert space valued Markov process,
e.g., an Ornstein-Uhlenbeck process. Certain problems involving random fields (e.g. horizontal or vertical filtering) can also be solved using the white noise formulation. We conclude by discussing the robustness of the measure valued filter including a property which we call statistical robustness (Theorem 3). Our final result (Theorem 4) extends some of our consistency results to the measure valued case.

Corresponding results for prediction and smoothing problems will appear in our forthcoming work [4].

For notation and definitions not given in this paper, the reader is referred to [3].

2. Measure-valued Equations for the Optimal Filter

Let $X = (X_t)$ be a Markov process defined on a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and taking values in a Polish space $S$. Let $K$ be a separable Hilbert space with inner product $(\cdot, \cdot)_K$ and norm $\|\cdot\|_K$ and let

$$H := L^2([0,T], K) := \{ z : [0,T] \to K \text{ with } \|z\|^2 := \int_0^T \|z_s\|_K^2 \, ds < \infty \}.$$ 

Let $h : [0,T] \times S \to K$ be a measurable function such that

$$\int_0^T \|h_s(X_s(\omega))\|_K^2 \, ds < \infty$$

for every $\omega \in \Omega$. 

Denote by $(H, \mathcal{C}, \mathbb{P})$ the finitely additive Gaussian white noise measure and by $\mathbb{I}$, the identity map from $H$ to $H$. Let $\omega \mapsto \mathbb{I}(\omega)$ be the $A\otimes \mathcal{B}(H)$-measurable, $H$-valued process $\mathbb{I}_t(\omega) = h_t(X_t(\omega)), t \in [0,T]$. Both $\mathbb{I}$ and $e$ can be defined in an obvious manner on $(E, \mathcal{E}, \mathbb{P}) = (\Omega, \mathcal{A}, \mathbb{P}) \otimes (H, \mathcal{C}, \mathbb{P})$. Then the abstract filtering model

$$y = \mathbb{I} + e$$

(2)

takes the more concrete form

$$y_t = h_t(X_t) + e_t, \quad 0 \leq t \leq T.$$ 

(3)

In (3) $(y_t)$ is the observed process and $(e_t)$ is $K$-valued white noise.
Although our results also include the case of finite dimensional $K$, as mentioned earlier, we are concerned in this paper with applications where $\dim K = \infty$.

Letting $Q_t$ be the orthogonal projection on $H$ with range $H_t := \{ \cdot \in H : \int_0^T \| \cdot_S^d \|_K^2 \, ds = 0 \}$ ($0 \leq t \leq T$), it can be shown that the conditional expectation $E_\lambda [g(X_t) | Q_t y]$ exists in the finitely additive theory and is given by the Bayes formula

$$E_\lambda [g(X_t) | Q_t y] = \frac{\sigma_t (g, Q_t y)}{\sigma_t (1, Q_t y)} \quad (4)$$

where

$$\sigma_t (g, Q_t y) = \int g(X_t (\omega)) \exp \{ \int_0^t (y_S^d h_S (X_S (\omega))) \, ds - \frac{1}{2} \int_0^t \| h_S (X_S (\omega)) \|_K^2 \, ds \} \, d\omega. \quad (5)$$

It is assumed, of course, that $g : S \to \mathbb{R}$ is such that $g(X_t)$ is $\mathbb{F}$-integrable.

The main problem is to determine, recursively, the conditional distribution $F_t^y (\cdot)$ of $X_t$ given $y_s$, $0 \leq s \leq t$ or, equivalently, the unnormalized conditional distribution $\gamma_t^y (\cdot)$, where

$$F_t^y (B) = \frac{\gamma_t^y (B)}{\gamma_t^y (S)} \quad \text{and} \quad \sigma_t (g, Q_t y) = \int_S g(x) \gamma_t^y (dx).$$

Let $\mathcal{M} := \mathcal{M}(S, S(S))$ be the class of all finite Borel measures on $S$.

Before stating our result we need to impose the following conditions on $X$: Let $\mathcal{D} := D([0, \infty), S)$ be the Skorokhod space of functions from $[0, \infty)$ to $S$ which are right continuous and have left limits and let $A_t^S := \{ x_u, s \cdot u \cdot t \}$ where $0 \leq s < t < \infty$ and $(x_u)$ are the coordinate maps.

(A) (i) The paths of the process $X$ belong to $\mathcal{D}$;

(ii) The Markov process $X_t$ admits a transition probability function $P(\cdot; \cdot, \cdot)$;

(iii) For all $(s, x) \in [0, \infty) \times S$, there exists a (countably additive) probability measure $P_{s, x}$ on $(\mathcal{D}, A_t^S)$ such that for all $k \geq 1$,
The extended generator of $X_t$ will be denoted by $L$ and its domain by $D$. More precisely, $L$ is the generator of $\hat{X}_t := (t, X_t)$ with augmented state space $[0, \infty) \times S$ and stationary transition probabilities. Let $(V^S_t)$ be the two parameter semigroup of $(X_t)$ and $(T_t)$, the one parameter semigroup of $(\hat{X}_t)$ with extended generator $L$ and domain $D$. See [3].

Theorem 1. Let the $S$-valued Markov process $X_t$ satisfy Condition (A) and

$$\int_0^T \| h_s(X_s) \|^2_K ds < \infty. \quad (6)$$

For each $\gamma \in H$ let $c^\gamma_s(x) : [0, T] \times S \to \mathbb{R}$ be defined by

$$c^\gamma_s(x) = (h_s(x), \eta_s) - \frac{1}{2} \| h_s(x) \|^2_K. \quad (7)$$

Denote by $V^S_t$ the semigroup corresponding to the transition probabilities of $X$ and also write $N_0 := x_0^{-1}$. Then the following conclusions hold for all $\gamma \in H$:

(a) $F^\gamma_t$ is the unique solution of the integral equation

$$F^\gamma_t = f(0, \cdot), N_0 > + \int_0^t (Lf + c^\gamma_s f)(s, \cdot), \gamma_s > ds, f \in D. \quad (8)$$

(b) $F^\gamma_t$ is the unique solution of the equation

$$F^\gamma_t g = (V^S_t g, N_0 > + \int_0^t c^\gamma_s (V^S_s g)(\cdot), \gamma_s > ds \quad (9)$$

for all $g \in \mathcal{B}(S)$, the bounded, Borel measurable functions on $S$.

(c) $F^\gamma_t$ is the unique solution to
\[
\langle f(t, \cdot), F^Y_t \rangle = \langle f(0, \cdot), N_0 \rangle + \int_0^t \langle (Lf + c_s^Y f)(s, \cdot), F^Y_s \rangle ds
\]
\[
- \int_0^t \langle c_s^Y, F^Y_s \rangle < f, F^Y_s > ds, \ f \in \mathcal{D},
\]
(10)

(d) \( F^Y_t \) is the unique solution of the equation

\[
\langle g, F^Y_t \rangle = \int_0^t \langle c_s^Y (v_s^g), F^Y_s \rangle ds - \int_0^t \langle c_s^Y, F^Y_s \rangle < v_s^g, F^Y_s > ds
\]
\] (11)

for all \( g \in \mathcal{B}(S) \).

The uniqueness property asserted above is to be understood to hold in the class of \( \{K_t \} \subseteq \mathcal{M} \) satisfying the following conditions:

For \( A \subseteq \mathcal{B}(S) \), \( K_0(A) = E_A^{-1}(X_0) \) and \( t - K_t(A) \) is a bounded Borel measurable function;

\( K_t \) is absolutely continuous with respect to \( \mathbb{P} \circ X^{-1}_t \) and

\[
\frac{\partial K_t}{\partial \mathbb{P} \circ X^{-1}_t} \leq R, \quad 0 \leq t \leq T,
\]
(13)

for a suitable constant \( R \leq \infty \).

3. Examples

The theorems of the preceding section can be applied to certain problems of filtering of random fields. Suppose that \( X = (X_{sx}) \), \( 0 \leq s \leq T, \ 0 \leq x \leq b \) is a two parameter, real-valued, sample continuous process. We shall refer to \( x \) as the spatial parameter. The observation model is

\[
Y_{tx} = \int_0^t \int_0^x h_s(X_{sv}) dv ds + W_{tx}
\]
(14)

where the following conditions are imposed:

(i) \( W = (W_{tx}) \) is a standard, Yeh-Wiener process.

(ii) \( X \) and \( W \) are independent.

(iii) \( \mathbb{E} \int_0^T \int_0^b \hat{h}_s(X_{sx})^2 dv ds < \infty \).

The assumptions on \( X \) allow us to regard \( X_t \) as a sample continuous pro-
cess taking values in the Banach space $B := C_0([0,b], \mathbb{R})$. So we have $S = B$ and $X \in C([0,T];B)$. We shall make the further assumption that $X_t$ is a (B-valued) Markov process. Consider now the white noise model.

Take $K = H$, the (reproducing kernel) Hilbert space of real valued absolutely continuous functions on $[0,b]$ with square integrable derivatives. Let $h : [0,T] \times B \to H$ be given by

$$h_t(\eta)(x) := \int_0^x \hat{h}_t(\eta_y) \, dy, \quad \eta \in B.$$ 

Then the white noise observation model is given by

$$y_t = h_t(X_t) + e_t, \quad 0 \leq t \leq T,$$

where $e_t$ is $H$-valued Gaussian white noise. For more details and a somewhat different treatment of this problem, see [7].

The conversion of the filtering problem involving a two parameter random field into a filtering problem for an infinite dimensional (in this case B-valued) signal process is useful in cases when observations can be assumed to be available for all values of the spatial variable $x$. The latter situation occurs when $X_{tx}$ represents the (random) voltage potential at time $t$ and "site" $x$ of a spatially extended neuron, the neuron being modeled as a thin cylinder or segment $[0,b]$ (see [5,6]). Another example in which neither parameter has the connotation of time occurs in problems of physical geodesy [1].

An example of a B-valued Markov process $X_t$ is the Malliavin Ornstein-Uhlenbeck (O-U) process which is given by the unique solution of the SDE $dX_{tx} = -\frac{1}{2} X_{tx} \, dt + dB_{tx}$, $X_{0x} \sim B (X_{0x}$ Gaussian) where $B = (B_{tx})$ is a Yeh-Cameron Wiener process. The generator $L$ (the so-called Ornstein-Uhlenbeck operator) is well known: For $0 \leq x_1, \ldots, x_n \leq b$ and for functions $f : B \to \mathbb{R}$ of the form

$$f(\eta) = \tilde{f}((\eta(x_1), \ldots, \eta(x_n)))$$

where $\tilde{f} \in C^2_b(\mathbb{R}^n)$ we have
\[(Lf)(\eta) = \frac{1}{2} \sum_{i,j=1}^{n} \chi_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\eta(x_1), \ldots, \eta(x_n)) \quad \text{and} \quad \sum_{i=1}^{n} \eta(x_i) \frac{\partial f}{\partial x_i}(\eta(x_1), \ldots, \eta(x_n)).\]

Other examples in which the infinite dimensional signal process \(X_t\) does not arise from a multiparameter stochastic process also occur in the study of neuronal behavior when the geometrical shape of the neuron is more complicated. Even in the simplest such models, \(X_t\) is an O-U process (more generally, a non-Gaussian, diffusion process) whose sample paths lie in \(C([0,T], K)\) where \(K\) is some infinite dimensional Hilbert space. A natural filtering problem is given by the linear model

\[Y_t = X_t + e_t\]

which follows from (3) by taking \(B = K\) and \(h_t(\eta) = \eta\).

4. Robustness of the Measure Valued Optimal Filter

We shall now state some typical results on the robustness properties of the measure valued optimal filter. These are extensions of results recently obtained by H.P. Hucke in his thesis [2]. Robustness is used in two senses here: the first in the sense that has become conventional in the theory, viz., the continuous pathwise dependence of the optimal filter on the observations; the second is robustness as commonly used in statistics.

Write

\[M := \mathbb{E} \int_{0}^{T} ||h_s(X_s)||_K^2 ds \leq \infty.\]

Theorem 2. Let \(y, y' \in \mathcal{M}\) and denote the total variation norm

\[||Y_t - y_t' ||_{\text{var}} := \sup_{A \in \mathcal{B}(S)} |Y_t(A) - y_t'(A)|.\]

Then

\[||Y_t - y_t' ||_{\text{var}} \leq M \cdot ||Q_t y - Q_t y' || \cdot \exp \left\{ \frac{1}{2} ||Q_t y||^2 + \frac{1}{2} ||Q_t y'||^2 \right\}.\]

Theorem 3. Let \((X^k) \ (k = 1, 2, \ldots)\) be signal processes satisfying assump-
tion (A) of Theorem 1. Suppose further that the following conditions are satisfied:

(a) \( h : [0,T] \times S \rightarrow K \) is continuous.
(b) The sequence of measures \( \pi_0(K)^{-1} \) converges weakly \((\rightarrow)\) to \( \pi_0^{-1} \) in \((\mathbb{D}, \mathcal{S}(\mathbb{D}))\).
(c) \( \pi_0^{-1}([x \in \mathbb{D} : x_t \neq x_{t-1}]) = 0 \) for each \( t \).

Then for every \( y \in H \), \( \gamma^k, y \Rightarrow \gamma^y \) in \( M(S) \) (in the topology of weak convergence.)

The statistical significance of the above result is clear. It says that small changes in the distribution of the signal process cause only small deviations in the optimal filter. Further implications of statistical robustness in specific problems (such as the Kalman filter with non-Gaussian initial distribution) will be discussed in a later work.

To what extent do the results of this paper imply analogous results for the stochastic calculus theory? Such consistency results when signal and noise are finite dimensional have been presented in [3]. For the infinite dimensional case, results at this level of generality do not seem to have been obtained in the conventional theory. Nevertheless, using the white noise model of this paper one might hope to derive robust versions for the countably additive model of at least some of our results. We conclude this topic by stating a preliminary result of this kind.

Let \((\gamma, K, B)\) be an abstract Wiener space, \( \gamma : K \times B, X := C_0([0,T], B) \) and \( \gamma \), the Wiener measure on \((X, \mathcal{B}(X))\). Recalling that \( H = L^2([0,T], K) \), a representation space for \((H, \mathcal{C}, m)\) is given by \((X, \mathcal{B}(X), \mu)\). Denote by \( W_t \), the coordinate map on \( X \). Letting \( (\hat{\gamma}, \hat{A}, \hat{\Pi}) = (\gamma, A, \Pi) \otimes (X, \mathcal{B}(X), \mu) \) and defining all processes involved on the product space \( \hat{\gamma} \) in the usual manner we obtain a stochastic calculus model corresponding to (15)

\[
\gamma_t = \int_0^T [h_u(X_u)]du + W_t. \tag{16}
\]
Finally let $\hat{Y}_t, \hat{\mathcal{F}}_t$ be the conditional measures for the optimal filter for (16). Let $\mathcal{Y}_t$ stand for either one of these. Similarly let $\mathcal{Y}_t$ be either one of $\mathcal{Y}_t$ or $\mathcal{F}_t$.

Theorem 4.

(a) Let $g \in L^1(\mathcal{F}, \mathbb{F}, \mathbb{P})$ where $g(\omega) = f(X_t(\omega))$. Then, with $\mathcal{R}$ denoting the lifting map we have

$$\mathcal{R} \mathbb{E}_{\mathbb{P}}(g | Q_\mathcal{Y}) = \mathbb{E}_{\mathbb{P}}(g | \mathcal{F}_t).$$

(b) Let $f : \mathcal{S} \rightarrow \mathbb{R}$ be bounded and continuous. Then

$$\mathcal{R} \langle f, \mathcal{G}_\mathcal{Y} \rangle = \langle f, \hat{\mathcal{G}}_\mathcal{Y} \rangle$$

where $\langle f, \nu \rangle := \int f(x) d\nu(x)$ for $\nu \in \mathcal{M}(\mathcal{S})$.

References


END
DATE
FILMED
DEC.
1987