Smoothness Properties of the Conditional Expectation in Finitely Additive White Noise Filtering

H. P. HUCKE

Fern University

G. KALLIANPUR*

University of North Carolina

AND

R. L. KARANDIKAR*

Indian Statistical Institute

Communicated by the Editors

It is shown that for a wide class of signal processes and bounded g , the conditional expectation $\pi(g, y)$ in the white noise filtering model is a C^∞ -functional of the observations in the sense that $\pi(g, y)$ and its Fréchet derivatives (which exist) are random variables on the quasicylindrical probability space on which the observation model is defined. © 1988 Academic Press, Inc.

1. INTRODUCTION

In a recent paper, M. Chaleyat-Maurel has shown that the conditional expectations in the nonlinear filtering problem is a C^∞ -functional in Malliavin's sense [1]. A Malliavin calculus for functionals of finitely additive Gaussian white noise has not yet been developed though, in our view, many of the basic ideas of the former theory carry over naturally to the finitely additive situation.

Received January 26, 1988.

AMS 1980 subject classifications: Primary 60A05, 60G35; Secondary 60H15.

Key words and phrases: C^∞ -functional, finitely additive white noise filtering.

* Research supported by the Air Force Office of Scientific Research Contract F49620 85 C 0144.

In this note, we derive a result close in spirit to Malliavin calculus. We will be concerned with the smoothness properties of the conditional expectation regarded as a functional of the observations. In the same sense as in [1] the result obtained by us may be regarded as a robustness property of the nonlinear filter in the white noise theory. Our result cannot be directly compared with Chaleyat-Maurel's. We are throughout in a Hilbert space setting so that in contrast to the Malliavin theory all directional derivatives are admissible for us. Both the statement and the proof of the main theorems are straightforward. The only thing that sets the proof apart from a standard calculation is the need to show that the various functional derivatives of the filter are also random mappings as defined in [2]. The latter fact is established by relying heavily on properties of lifting maps.

It must be noted that our filtering model assumes signal and noise to be independent whereas in [1] a more general model is considered. However, we are able to prove C^∞ -smoothness of the filter under less restrictive conditions.

2. NOTATION AND TERMINOLOGY

For most of the notation, terminology, and definitions used in this paper we refer the reader to [2] since it would take too much space to repeat them here.

H is an infinite dimensional, separable Hilbert space, $\mathcal{C}$ the field of finite dimensional Borel cylinder sets in H , and m the (finitely additive) canonical Gauss measure on H , i.e., the measure with characteristic functional $\exp(-\frac{1}{2}\|h\|^2)$, ($h \in H$). Let $\mathcal{P}$ denote the class of all orthogonal projections on H with finite dimensional ranges. Let $(\Omega, \mathcal{A}, \Pi)$ be a complete (countably additive) probability space. The triple $(E, \mathcal{E}, \alpha)$ is called a quasicylindrical probability space where $E = \Omega \times H$, $\mathcal{E} = \mathcal{A} \times \mathcal{C}$ and $\alpha = \Pi \odot m$. $\mathcal{E}$ is a field and α is the finitely additive probability on $\mathcal{E}$ such that for any $P \in \mathcal{P}$, the restriction of α to the σ -field $\mathcal{A} \times \mathcal{C}_P$ is the countably additive probability measure $\Pi \times m_P$. Here $\mathcal{C}_P$ is the σ -field of cylinder sets with bases on PH and m_P is the restriction of m to $\mathcal{C}_P$.

Let (L_0, Π_0) be a representation of m with an underlying representation space $(\Omega_0, \mathcal{A}_0, \Pi_0)$ and let $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\Pi}) = (\Omega, \mathcal{A}, \Pi) \otimes (\Omega_0, \mathcal{A}_0, \Pi_0)$. Writing $\tilde{\omega} = (\omega, \omega_0) \in \tilde{\Omega}$, defining $\rho(\tilde{\omega}) = \omega$ and $L(h)(\tilde{\omega}) = L_0(h)(\omega_0)$ for all $h \in H$, it is seen that $(\rho, L, \tilde{\Pi})$ is a representation of the quasicylindrical probability α on the space $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\Pi})$. It can, in fact, be shown that $(\rho, L, \tilde{\Pi})$ can be chosen to possess the property that for each $h \in H$, the map $(h, \tilde{\omega}) \rightarrow L(h)(\tilde{\omega})$ is $B(H) \otimes \tilde{\mathcal{A}}$ measurable. It is such a representation that we shall be working with throughout.

Let S be a Polish space, i.e., a complete separable metric space. We shall define classes of S -valued maps on E which form important subclasses of random variables on the finitely additive probability space $(E, \mathcal{E}, \alpha)$.

Let $\mathcal{L}^0(E, \mathcal{E}, \alpha; S)$ be the class of maps f from E to S such that for all $P \in \mathcal{P}$, f_P defined by $f_P(\omega, \eta) = f(\omega, P\eta)$ is $\mathcal{E}_P/B(S)$ -measurable and for all sequences $\{P_j\} \subset \mathcal{P}$ converging strongly to the identity ($P_j \rightarrow^s I$), $R_\alpha(f_{P_j})$ is Cauchy in $\tilde{I}$ -probability. Elements of $\mathcal{L}^0(E, \mathcal{E}, \alpha; S)$ are called S -valued accessible random variables. For $1 \leq q < \infty$, define $\mathcal{L}^q(E, \mathcal{E}, \alpha; S)$ as the class of maps f as above with the additional property that

$$\int_{\Omega} |R_\alpha(f_{P_j}) - R_\alpha(f_{P_l})|^q d\tilde{I} \rightarrow 0.$$

In this case

$$\int_{\Omega} |R_\alpha(f)|^q d\tilde{I} < \infty.$$

The notation here is somewhat different from that adopted in [2] where the class $\mathcal{L}^0$ is denoted by $\mathcal{L}^*$ and $\mathcal{L}^1$ by $\mathcal{L}^{1*}$. Wider classes of random variables are also considered in [2]. The symbol S will be suppressed whenever $S = \mathbb{R}^1$.

Let $\xi: \Omega \rightarrow H$ be a random variable, i.e., a $B(H)/\mathcal{A}$ -measurable map, $B(H)$ being the σ -field of Borel sets in H . The nonlinear filtering model in its abstract form is defined on $(E, \mathcal{E}, \alpha)$ by

$$y = \xi + e, \quad (1)$$

where for $(\omega, \eta) \in E$, $\xi(\omega, \eta) = \xi(\omega)$ and $e(\omega, \eta) = e(\eta) = \eta$. The identity map e on H is called Gaussian white noise, ξ is the signal and y the observation.

Let Q be an arbitrary orthogonal projection on H . If g is a Π -integrable, real random variable on Ω , then the conditional expectation (in the finitely additive theory) $E_\alpha(g|Qy)$ exists and is given by the Bayes formula

$$E_\alpha(g|Qy) = \frac{\sigma_Q(g, y)}{\sigma_Q(1, y)}, \quad (2)$$

where

$$\sigma_Q(g, y) = \int_{\Omega} g(\omega) \exp\{(y, Q\xi(\omega)) - \frac{1}{2} \|Q\xi(\omega)\|^2\} d\Pi(\omega) \quad (3)$$

is called the unnormalized conditional expectation of g . The model (1) covers most of the filtering problems met with in practice including those in

which the observation process takes values in a Hilbert space. In applications, the true signal process is denoted by an S -valued process (X_t) , $(0 \leq t \leq T)$ defined on Ω , and (1) takes the form

$$y_t = \xi_t + e_t, \quad 0 \leq t \leq T, \quad (4)$$

where

e is K -valued Gaussian white noise. Here $H = L^2([0, T], K)$ and K is a possibly infinite dimensional separable Hilbert space; (5a)

$\xi_t(\omega) = h_t(X_t(\omega))$ where $h: [0, T] \times S \rightarrow K$ is measurable and satisfies the condition $\int_0^T \|h_t(X_t(\omega))\|_K^2 dt < \infty$ for each ω (or a.a. ω). (5b)

If Q_t is the orthogonal projection on H with range $H_t := \{f \in H: \int_t^T \|f_s\|_K^2 ds = 0\}$ then the filter one is interested in is the conditional expectation $E_\alpha(g | Q_t y)$ which is given by (2) with $Q = Q_t$. For the sake of notational convenience we shall derive all our results for the abstract model (1) rather than (4).

In what follows we may take, without loss of generality, g to be non-negative and such that $\int g d\Pi = 1$. Let $d\Pi_1 = g d\Pi$ and $\nu = \Pi_1 \circ (Q\xi)^{-1}$. Then ν is a probability measure on H and

$$\sigma_Q(g, \eta) = \int_H \exp\{(\eta, k) - \frac{1}{2} \|k\|^2\} d\nu(k), \quad \eta \in H. \quad (6)$$

Since, throughout this work, g and Q will remain fixed, it is convenient to suppress g and write $\sigma(\eta)$ for $\sigma_Q(g, \eta)$.

For a Banach space B with norm $\|\cdot\|_B$ let $L(H, B)$ denote the class of all bounded linear transformation $A: H \rightarrow B$, which is itself a Banach space with operator norm. A mapping $f: H \rightarrow B$ is said to be Fréchet differentiable if for every $h \in H$ there exists $f_1(h) \in L(H, B)$ such that

$$\lim_{\|h'\| \rightarrow 0} \frac{1}{\|h'\|} \|f(h+h') - f(h) - f_1(h)[h']\|_B = 0,$$

$f_1(h)$ is called the Fréchet derivative of f at h and is written as $(Df)(h)$.

Let $L^0(H) = \mathbb{R}$, $L^1(H) = L(H, \mathbb{R})$, and for $r \geq 1$, $L^{r+1}(H) = L(H, L^r(H))$. It is well known that the Banach space $L^r(H)$ can be identified with the class of all linear mappings from the r -fold product $H \times \dots \times H$ into $\mathbb{R}$. The norm $\|\cdot\|_r$ on $L^r(H)$ under this identification is given by

$$\|f\|_r = \sup\{|f[h_1, \dots, h_r]|: h_i \in H, \|h_i\| \leq 1\}.$$

A function $f: H \rightarrow \mathbb{R}$ is said to be $(r+1)$ times Fréchet-differentiable, if it is r -times Fréchet differentiable, and $D'f: H \rightarrow L'(H)$ is Fréchet-differentiable and then $D'^{+1}f := D(D'f)$.

Let $L'_{(2)}(H)$ be the subclass of $L'(H)$ consisting of $g \in L'(H)$ for which

$$\|g\|_{r,2}^2 := \sum_{j_1 \cdots j_r} |g[\varphi_{j_1}, \dots, \varphi_{j_r}]|^2 < \infty,$$

where $\{\varphi_j\}$ is any CONS in H . It is well known that $\|g\|_{r,2}$ does not depend on the choice of CONS and that $L'_{(2)}(H)$ is a Hilbert space with norm $\|\cdot\|_{r,2}$ and that $\|g\|_r \leq \|g\|_{r,2}$.

3. MAIN RESULTS

LEMMA 1. *Let the function $\sigma(\eta)$ be defined by (6). Then*

(a) *for every $r \geq 1$, $\sigma(\eta)$ is r -times Fréchet differentiable and the derivative $D'\sigma(\eta)$ is given by*

$$\begin{aligned} D'\sigma(\eta)[h_1, \dots, h_r] \\ = \int [\exp\{(\eta, k) - \tfrac{1}{2}\|k\|^2\}](h_1, k) \cdots (h_r, k) dv(k). \end{aligned} \quad (7)$$

(b) $D'\sigma(\eta) \in L'_{(2)}(H)$.

Proof. Denote the right-hand side of (7) by $g_r[h_1, \dots, h_r]$. The integral appearing in (7) is finite since

$$|(\eta, k)| \leq \tfrac{1}{2} \{ \|2\eta\|^2 + \tfrac{1}{2}\|k\|^2 \} = 2\|\eta\|^2 + \tfrac{1}{8}\|k\|^2 \quad (8)$$

and

$$|(h, k)| \leq \|h\| \cdot \|k\|. \quad (9)$$

Let $\{\varphi_j\}$ be a CONS in H . Note that

$$\begin{aligned} \|g_r\|_{r,2}^2 &= \sum_{j_1 \cdots j_r} \left[\int \exp\{(\eta, k) - \tfrac{1}{2}\|k\|^2\} \cdot (\varphi_{j_1}, k) \cdots (\varphi_{j_r}, k) dv(k) \right]^2 \\ &\leq \sum_{j_1 \cdots j_r} \int \exp\{2(\eta, k) - \|k\|^2\} \cdot (\varphi_{j_1}, k)^2 \cdots (\varphi_{j_r}, k)^2 dv(k) \\ &= \int \exp\{2(\eta, k) - \|k\|^2\} \cdot \|k\|^{2r} dv(k) \\ &< \infty, \end{aligned} \quad (10)$$

in view of (8) and (9). Hence $g_r(\eta) \in L'_{(2)}(H)$. Let us write $\sigma(\eta) = g_0(\eta)$. To complete the proof, we will show that g_r is Fréchet differentiable for $r \geq 0$ and that $Dg_r = g_{r+1}$. This will show that $D'\sigma(\eta) = g_r(\eta)$. For $\eta, h \in H$, let

$$v(\eta, h) := \|g_r(\eta + h) - g_r(\eta) - g_{r+1}(\eta)(h)\|_{r,2}.$$

Recall that $g_{r+1}(\eta)(h)[h_1, \dots, h_r] = g_r(\eta)[h_1, \dots, h_r, h]$ under the identification of $L(H, L'(H))$ with $L^{r+1}(H)$. Thus

$$\begin{aligned} v^2(\eta, h) &= \sum_{j_1 \dots j_r} |\{g_r(\eta + h) - g_r(\eta)\}[\varphi_{j_1}, \dots, \varphi_{j_r}] \\ &\quad - g_{r+1}(\eta)[\varphi_{j_1}, \dots, \varphi_{j_r}, h]|^2 \\ &= \sum_{j_1 \dots j_r} \left[\int \exp\{\langle \eta, k \rangle - \frac{1}{2} \|k\|^2\} \right. \\ &\quad \cdot (\varphi_{j_1}, k) \cdots (\varphi_{j_r}, k) \cdot v_1(h, k) dv(k) \left. \right]^2, \end{aligned} \quad (11)$$

where

$$v_1(h, k) := \exp\{\langle h, k \rangle\} - 1 - \langle h, k \rangle.$$

Since

$$|v_1(h, k)| \leq |\langle h, k \rangle|^2 \exp(\langle h, k \rangle) \leq \|h\|^2 \|k\|^2 \exp(\langle h, k \rangle),$$

we get from (1)

$$\begin{aligned} v^2(\eta, h) &\leq \int \exp\{2\langle \eta, k \rangle - \|k\|^2\} \cdot \|k\|^{2r} \\ &\quad \cdot \|h\|^4 \|k\|^4 \exp(2\langle h, k \rangle) dv(k) \\ &\leq \|h\|^4 \cdot \exp\{4\|h\|^2 + 4\|\eta\|^2\} \\ &\quad \cdot \int \exp\{-\frac{1}{2}\|k\|^2\} \|k\|^{2r+4} dv(k) \end{aligned}$$

using (8) and (9). This shows that

$$\lim_{\|h\| \rightarrow 0} \frac{1}{\|h\|} \cdot v(\eta, h) = 0$$

and hence that $g_r: H \rightarrow L'_{(2)}(H)$ is Fréchet differentiable with $Dg_r = g_{r+1}$. Since $\|\cdot\|_{r,2} \leq \|\cdot\|_r$, this implies that $g_r: H \rightarrow L'(H)$ is Fréchet differentiable. ■

Our next step is to show that $D'\sigma(y)$ is a random variable.

LEMMA 2. Suppose that ν satisfies

$$\int \|k\|^{2r} d\nu(k) < \infty. \quad (12)$$

Then $D'\sigma(y) \in \mathcal{L}(E, \mathcal{E}, \alpha; L_2^{(r)}(H))$.

Proof. Fix $\{P_i\} \subseteq \mathcal{P}$, $P_i \rightarrow^s I$. Let $Z(k, \tilde{\omega})$, $Z_i(k, \tilde{\omega})$ be defined by

$$Z(k, \tilde{\omega}) := \exp\{(k, \xi(\omega)) + L_0(k)(\omega_0) - \frac{1}{2} \|k\|^2\}$$

and

$$Z_i(k, \tilde{\omega}) := \exp\{(k, \xi(\omega)) + L_0(P_i k)(\omega_0) - \frac{1}{2} \|k\|^2\}, \tilde{\omega} = (\omega, \omega_0).$$

Then from the definition of lifting for cylinder functions (see [2]) it follows that

$$R_\alpha([D'\sigma \circ y]_{P_i} [\varphi_{j_1} \cdots \varphi_{j_r}](\tilde{\omega})) = \int Z_i(k, \tilde{\omega}) f_j(k) d\nu(k),$$

where $j = (j_1, \dots, j_r)$ and $f_j(k) = (\varphi_{j_1}, k) \cdots (\varphi_{j_r}, k)$. Let $U_{it}(\tilde{\omega}) := \|R_\alpha(\{D'\sigma \circ y\}_{P_i}) - R_\alpha([D'\sigma \circ y]_{P_i})\|_{r,2}^2(\tilde{\omega})$. To complete the proof we will show that $U_{it} \rightarrow 0$ in $\tilde{H}$ -probability:

$$\begin{aligned} U_{it}(\tilde{\omega}) &\leq \sum_j \left[\int |Z_i(k, \tilde{\omega}) - Z_i(k, \tilde{\omega})| |f_j(k)| d\nu(k) \right]^2 \\ &\leq \sum_j \int |Z_i(k, \tilde{\omega}) - Z_i(k, \tilde{\omega})| d\nu(k) \\ &\quad \cdot \int |Z_i(k, \tilde{\omega}) - Z_i(k, \tilde{\omega})| \cdot f_j^2(k) d\nu(k) \\ &= \int |Z_i(k, \tilde{\omega}) - Z_i(k, \tilde{\omega})| d\nu(k) \\ &\quad \cdot \int |Z_i(k, \omega) - Z_i(k, \tilde{\omega})| \|k\|^{2r} d\nu(k). \end{aligned} \quad (13)$$

We have used Hölder's inequality above.

Define a probability measure Π' on $\tilde{\mathcal{Q}}$ by

$$\begin{aligned} \frac{d\Pi'}{d\tilde{\Pi}}(\tilde{\omega}) &= C \cdot \exp(-\frac{1}{2} \|\xi(\omega)\|^2) \\ &\quad \cdot \exp(-L_0(\xi(\omega))(\omega_0) - \frac{1}{2} \|\xi(\omega)\|^2), \end{aligned}$$

where, $\tilde{\omega} = (\omega, \omega_0) \in \tilde{\Omega}$. The constant C is chosen such that $\Pi'(\tilde{\Omega}) = 1$. It is easy to see that

$$\int \exp(\tfrac{1}{2} \|\xi(\omega)\|^2) d\Pi' < \infty.$$

Let μ be any countably additive finite measure on H . Then, it is easy to check that

$$Z_i \rightarrow Z \quad \text{in } \mu \otimes \Pi' \text{ measure}$$

and

$$\iint Z_i(k, \tilde{\omega}) d\Pi'(\tilde{\omega}) d\mu(k) \rightarrow \iint Z(k, \tilde{\omega}) d\Pi'(\tilde{\omega}) d\mu(k),$$

where the integrals appearing above are finite. By arguments similar to the proof of Scheffé's theorem, it follows that

$$\iint |Z_i - Z| d\Pi' d\mu \rightarrow 0$$

and as a consequence

$$\int |Z_i(k, \tilde{\omega}) - Z(k, \tilde{\omega})| d\mu(k) \rightarrow 0 \quad (14)$$

in Π' -probability. Since $\tilde{\Pi} \ll \Pi'$, (14) also holds in $\tilde{\Pi}$ -probability. The assumption (12) implies that ν_1 defined by

$$\frac{d\nu_1}{d\nu}(k) = \|k\|^{2r}$$

is a finite measure. Thus (14) for $\mu = \nu$ and $\mu = \nu_1$ implies that $U_{ii} \rightarrow 0$ in $\tilde{\Pi}$ -probability. ■

THEOREM 3. *For any integrable function f , for any orthogonal projection Q , $\sigma_Q(g, y)$ is r -times Fréchet differentiable for all $r \geq 1$. Further, (a) if*

$$\int \|Q\xi(\omega)\|^{2r} \cdot |g(\omega)| d\Pi(\omega) < \infty \quad (15)$$

then

$$D^r \sigma_Q(g, y) \in \mathcal{L}(E, \mathcal{E}, \alpha; L'_{(2)}(H));$$

(b) *if $\|\xi\|, g$ are bounded, then for all $q \geq 1, r \geq 1$,*

$$D^r \sigma_Q(g, y) \in \mathcal{L}^q(E, \mathcal{E}, \alpha; L'_{(2)}(H)).$$

Proof. Part (a) follows from the preceding lemmas. The proof of (b) is based on the easily verifiable fact that

$$\int [U_{il}]^{q/2} d\tilde{H} \rightarrow 0$$

as $(i, l) \rightarrow \infty$. We are now in a position to prove the main result of this paper.

THEOREM 4. *Suppose ξ and g satisfy (15). Then, $\pi_Q(g, y)$ is r -times Fréchet differentiable and*

$$D^r \pi_Q(g, y) \in \mathcal{L}(E, \mathcal{E}, \alpha; L'_{(2)}(H)). \quad (16)$$

Proof. $\pi_Q(g, y) = f_1(y)/f_2(y)$, where

$$f_1(y) = \sigma_Q(g, y), \quad f_2(y) = \sigma_Q(1, y).$$

Now, f_1, f_2 are both r -times F -differentiable and $f_2 > 0$. From this it is easy to check that f_1/f_2 is also r -times F -differentiable. It can be shown that $D^r(f_1/f_2)$ can be expressed as

$$D^r(f_1/f_2) = A_r(f_1, f_2, Df_1, Df_2, \dots, D^r f_1, D^r f_2, 1/f_2)$$

where A_r is a continuous mapping from

$$\mathbb{R} \times \mathbb{R} \times L^1_{(2)}(H) \times L^1_{(2)}(H) \times \dots \times L^r_{(2)}(H) \times L^r_{(2)}(H) \times \mathbb{R}$$

into $L^r_{(2)}(H)$. Since $f_1, f_2, 1/f_2 \in \mathcal{L}(E, \mathcal{E}, \alpha; \mathbb{R})$, and $D^i f_1, D^i f_2 \in \mathcal{L}(E, \mathcal{E}, \alpha; L^i_{(2)}(H))$, $1 \leq i \leq r$, assertion (16) follows from the continuity of A_r . ■

A functional $f(y)$ will be said to be a C^∞ -functional of the observations if $D^r f$ exists for all $r \geq 1$ and $D^r f(y) \in \mathcal{L}(E, \mathcal{E}, \alpha; L^r_{(2)}(H))$.

We have proved above that if

$$E \|\xi\|^r < \infty \quad \text{for all } r \geq 1,$$

then for all g bounded $\pi_Q(g, y)$ is a C^∞ -functional of the observations.

A concrete application of this result to the models considered in [2] shows that the conditional expectations in the filtering, prediction, and smoothing problems (for finite or infinite dimensional signals) are C^∞ -functionals of the observations.

REFERENCES

- [1] CHALEYAT-MAUREL, M. (1986). Robustesse du filtre et calcul des variations stochastique. *J. Funct. Anal.* **18** No. 1 55–71.
- [2] KALLIANPUR, G., AND KARANDIKAR, R. L. (1985). White noise calculus and nonlinear filtering theory. *Ann. Probab.* **13** No. 4 1033–1107.