Report 99-029

Bayes Formula for Optimal Filter with n-ple Markov Gaussian Errors

Pranab K. Mandal
V. Mandrekar

ISSN: 1389-2355

Bayes Formula for Optimal Filter with n-ple Markov Gaussian Errors

Pranab K. Mandal, EURANDOM, NL,

V. Mandrekar, Michigan State University, USA

Abstract

We consider the nonlinear filtering problem where the observation noise process is n-ple Markov Gaussian. A Kallianpur-Striebel type Bayes formula for the optimal filter is obtained.

Keywords and Phrases: Nonlinear Filtering, Bayes Formula, Kallianpur-Striebel Formula, n-ple Markov Gaussian process

Introduction

Professor Hida has been interested in non-linear problems and the study of n-ple Markov processes. The purpose of this work is to obtain a Bayes formula for the optimal filter in the case where the noise process is n-ple Markov in the sense of Levy - Hida ([6], [3]). We apply the general result obtained by us in [7]. The main effort here is to compute tractable form of the reproducing kernel Hilbert space (RKHS) using Goursat representation of n-ple Markov processes obtained by one of the authors [8] under active encouragement of T. Hida. As a consequence of our work we are able to obtain a proper form of the Bayes formula given by Kunita [5] for the noise process being a solution of stochastic differential equation of order n. It is a great pleasure to dedicate this work to Professor T. Hida who has been a friend and a mentor for one of the authors (Mandrekar) for more than thirty years.

The article is organized as follows. In section 1, we start with the main result from [7] that we will be using. We also introduce the notations and definitions we will follow throughout this article. We identify the suitable form of the RKHS of an n-ple Markov Gaussian process in Section 2. Finally, in section 3 we obtain the Bayes formula for n-ple Markov Gaussian process.

1. Preliminaries and Notations

The general filtering problem can be described as follows. The signal or system process $\{X_t, 0 \leq t \leq T\}$ is unobservable. Information about $\{X_t\}$ is obtained by observing another process Y which is a function of X corrupted by noise i.e.,

$$Y_t = \beta(t, X) + N_t \quad 0 \leq t \leq T, \tag{1}$$

where β_t is measurable with respect to $\mathcal{F}_t^X$, the σ-field generated by the "past" of the signal i.e., $\sigma\{X_u, 0 \leq u \leq t\}$ (augmented by the inclusion of zero probability sets) and N_t is some noise process. The observation σ-field $\mathcal{F}_t^Y = \sigma\{Y_u, 0 \leq u \leq t\}$. The aim of filtering theory is to get an estimate of X_t based on $\mathcal{F}_t^Y$. This is given by the conditional distribution of X_t given $\mathcal{F}_t^Y$ or equivalently, the conditional expectations $E(f(X_t)|\mathcal{F}_t^Y)$ for a large enough class of functions f. Kallianpur and Striebel [4] provided an explicit Bayes formula for the conditional expectation in case N_t is a Brownian Motion W_t. In [7], this formula was extended to general N_t, a Gaussian process, with certain restriction on $\beta(t, X)$. In order to explain it we need to introduce the idea of Reproducing Kernel Hilbert Space (RKHS).

Given a Gaussian process $\{N_t, 0 \leq t \leq T\}$ (WLOG assume $EN_t = 0$) let $R_N(s,t) := EN(t)N(s)$ be its covariance function. It determines uniquely the distribution of N. Once the Gaussian process is clear in the context we will drop the subscript N. Given a symmetric, non-negative definite function R (covariance) we can associate with it a unique Hilbert space of functions called RKHS $H(R)$ of R satisfying

$$\begin{cases} \text{(a)} \;\; R(\cdot, t) \in H(R) \;\; \text{for} \;\; 0 \leq t \leq T \\ \text{(b)} \;\; \text{For all} \;\; f \in H(R), \;\; (f, R(\cdot, t)) = f(t), \;\; 0 \leq t \leq T. \end{cases} \tag{2}$$

Here $(\cdot, \cdot)$ denote the inner product in $H(R)$. Thus there is one-one correspondence between Gaussian processes and RKHS's generated by covariances ([1]). In fact, the simplest way to describe $H(R_N)$ is to consider functions $f(t) = EN(t)U$, where $U \in \overline{sp}\{N(t), 0 \leq t \leq T\}$ and $(f_1, f_2) = EU_1U_2$. We denote the unique U by $< N, f >$. Under the assumption that $\beta(\cdot, X) \in H(R_N)$ a.s., the following extension of Kallianpur-Striebel formula was given in [7]. We use the notations $< N, f >_t$ and $|| \cdot ||_t$ to denote $< N, f >$ and the norm, respectively, corresponding to the RKHS of $R_N|_{[0,t]\times[0,t]}$.

Theorem 1. *Suppose that the observation process is given by (1) with noise N a Gaussian process with covariance R and $\beta(\cdot, X(w)) \in H(R)$ a.s. w.*

Then for an $\mathcal{F}_T^X$ measurable and integrable function $g(T, X)$

$$E(g(T,X)|\mathcal{F}_t^Y) = \frac{\int g(T,x) e^{<Y,\beta(\cdot,x)>_t - \frac{1}{2}||\beta(\cdot,x)||_t^2} dP_X(x)}{\int e^{<Y,\beta(\cdot,x)>_t - \frac{1}{2}||\beta(\cdot,x)||_t^2} dP_X(x)}$$

where $< Y, \beta(\cdot, x) >_t = < N, \beta(\cdot, x) >_t + ||\beta(\cdot, x)||_t^2$.

As stated in the introduction, we want to compute explicitly the above formula in case of n-ple Markov processes, which involves computing the RKHS of such processes. For this we need the form of $L^2(M)$, the space of vector valued functions, square integrable with respect to matrix valued measure M (Rozanov [10]). Let M be a non-negative $n \times n$ matrix valued measure M on a measurable space $(S, \mathcal{S})$. For two n-dimensional (row) vectors, ψ_1 and ψ_2 we define the integral

$$\int \psi_1 dM \psi_2^* = \int (M' \psi_2, \psi_1) d\mu$$

for any non-negative σ-finite measure μ which dominates the measures $m_{i,j}$, $(i, j = 1, 2, \ldots, n)$, the entries of M. Here $M' = ((dm_{ij}/d\mu))$. It can be easily seen that the definition above does not depend on μ.

In particular, $tr M = \sum_{i=1}^n m_{ii}$ is a dominating measure. We denote by

$$L^2(M) = \left\{ \psi : S \to \mathcal{R}^n(\text{row}) \text{ such that } \int \psi dM \psi^* < \infty \right\}.$$

Then $L^2(M)$ is a Hilbert space under the inner product $(\psi_1, \psi_2)_M = \int \psi_1 dM \psi_2^*$ for $\psi_1, \psi_2 \in L^2(M)$.

Let $\{\underline{Z}(t), 0 \le t \le T\}$ be an n-dimensional (row) Gaussian Martingale, then

$$E(\underline{Z}(t) - \underline{Z}(s))^*(\underline{Z}(t) - \underline{Z}(s)) = F_z(t) - F_z(s) \quad (3)$$

where F is a non-negative matrix valued "increasing" function (that is, $F_z(t) - F_z(s)$ is non-negative definite matrix if $s \leq t$). Let M_z be the matrix-valued measure associated with F_z. Then it is easy to see that we can define

$$\int (\psi(u), d\underline{Z}(u)) = \sum_{i=1}^{n} \int \psi_i(u) dZ_i(u) \tag{4}$$

for all $\psi \in L^2(M_z)$ and $\overline{sp}\{Z_i(u), i = 1, 2, \ldots, n, u \leq t\}$ equals

$$\left\{ \int_0^t (\psi(u), d\underline{Z}(u)), \;\; \psi \in L^2(M_z) \right\}.$$

2. n-ple Markov Gaussian Processes

Following Hida [3] and Levy [6], we define a Gaussian process N to be n-ple Markov if for each s, the set

$$\{E(N(t)|\mathcal{F}_s^N), t \geq s\}$$

in $L^2(\Omega, \mathcal{F}_T^N, P)$ has exactly n linearly independent elements. It was shown in [8] (see also [9]) that such processes have Goursat representation, i.e.,

$$N(t) = \sum_{i=1}^{n} \varphi_i(t) Z_i(t) \tag{5}$$

where $\underline{Z} = (Z_1, \ldots, Z_n)$ is a non-singular vector valued martingale and $det((\varphi_i(t_j))) \neq 0$ for any $0 \leq t_1 < t_2 < \cdots < t_n \leq T$. Here by nonsingular we mean $F_z(t)$ of (3) is a non-singular $n \times n$ matrix for each t. In particular, N has representation (5) with

$$H(N : t) := \overline{sp}\{N(s), s \leq t\} = H(\underline{Z} : t) := \overline{sp}\{Z_i(s), s \leq t, 1 \leq i \leq n\}.$$

We now derive the RKHS generated by N of the form (5).

Lemma 1. *Let* $N(t) = \sum_{i=1}^{n} \varphi_i(t) Z_i(t)$, *where* $\underline{Z}(t)$ *is an n-dimensional Gaussian martingale. Then*

$$H(R_N) = H := \left\{ f : f(t) = \int_0^t \varphi(t) dM_z(u) \psi^*(u), \psi \in L^2(M_z) \right\}$$

with

$$(f_1, f_2)_H = \int_0^T \psi_1(u) dM_z(u) \psi_2^*(u). \qquad (6)$$

Proof : Note that

$$\begin{aligned} R_N(\cdot, t) &= EN(\cdot)N(t) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_i(\cdot) \varphi_j(t) m_{ij}(t, \cdot). \end{aligned}$$

Define

$$\psi_t(u) = (\varphi_1(t) 1_{[0,t]}(u), \varphi_2(t) 1_{[0,t]}(u), \cdots, \varphi_n(t) 1_{[0,t]}(u)).$$

Then

$$\psi_t \in L^2(M_z) \quad \text{and} \quad R_N(\cdot, t) = \int_0^{\cdot} \varphi(\cdot) dM_z(u) \psi_t^*(u),$$

i.e., $R_N(\cdot, t) \in H$. Also, if $f(t) = \int_0^t \varphi(t) dM_z(u) \psi^*(u)$ then

$$\begin{aligned} (f, R(\cdot, t)) &= \int_0^T \psi_t(u) dM_z(u) \psi^*(u) \\ &= \int_0^t \varphi(t) dM_z(u) \psi^*(u) = f(t). \end{aligned}$$

Thus H and R_N satisfy the conditions (a) and (b) described in equation (2) and hence the lemma is proved. ∎

From the definition of $< N, f >$ and lemma 1 it is easy to check that for

$f \in H(R_N)$ such that $f(t) := \int_0^t \varphi(t) dM_z(u)\psi^*(u)$, we have

$$< N, f >_t = \int_0^t (\psi(u), dZ(u)), \qquad (7)$$

where the integral is defined as in (4).

Now, suppose $N(t)$ is a solution of the stochastic differential equation of order n given by

$$L_t N(t) = \dot{W}(t) \quad \text{with} \quad \dot{W}(t) \ \text{“white noise”} \qquad (8)$$

and where

$$L(t) = a_n(t)\left(\frac{d}{dt}\right)^n + a_{n-1}\left(\frac{d}{dt}\right)^{n-1} + \cdots + a_0(t), \quad (\text{with } a_n(t) > 0\ ).$$

More precisely, $N(t)$ is $(n-1)$ times differentiable Gaussian process and satisfies

$$a_n(t) dN^{(n-1)}(t) + \left(a_{n-1}(t)N^{(n-1)}(t) + \ldots + a_0(t)N(t)\right) dt = dW(t).$$

Then ([2]),

$$N(t) = \int_0^t \sum_{i=1}^n \varphi_i(t) g_i(u) dW(u),$$

where $\{g_i(u)\}_{i=1}^n$ are solutions of the adjoint homogeneous equation corresponding to $L_t h = 0$ and $\{\varphi_i(t)\}_{i=1}^n$ are determined so that $F(t,u) := \sum_{i=1}^n \varphi_i(t) g_i(u)$ satisfies the following conditions :

$$\left.\frac{\partial^k F(t,u)}{\partial u^k}\right|_{u=t} = \begin{cases} 0, & k = 0, 1, \ldots, (n-2) \\ (-1)^{(n-1)}/a_n(t), & k = (n-1). \end{cases}$$

Hence with $Z_i(t) = \int_0^t g_i(u) dW(u)$

$$N(t) = \sum_{i=1}^n \varphi_i(t) Z_i(t)$$

is of the form as in Lemma 1. Here $m_{ij}(B) = \int_B g_i(u)g_j(u)d\lambda(u)$ where λ is the Lebesgue measure. The space $L^2(M_z)$ consists of the functions ψ so that

$$\int_0^T \left(\sum_{i=1}^n \psi_i(u)g_i(u) \right)^2 d\lambda(u) < \infty$$

and $f \in H(R_N)$ iff

$$\begin{aligned} f(t) &= \int_0^t \left(\sum_{i=1}^n \varphi_i(t)g_i(u) \right) \left(\sum_{j=1}^n \psi_j(u)g_j(u) \right) d\lambda(u) \\ &= \int_0^t F(t,u)g^*(u)du \end{aligned}$$

for some $g^* \in L^2(\lambda)$, if we assume $H(N:t) = H(W:t)$. This is the case for Levy - Hida n-ple Markov processes. In this case we get as in [7]

$$< N, f >_t = \int_0^t g^*(u)dW(u).$$

This also is the case if $N(t)$ is a solution of (8) as $g \in L^2(\lambda)$ and $g \perp F(t,\cdot)1_{[0,t]}(\cdot)$ for all t, implies

$$\int_0^t F(t,u)g(u)du = 0 \ \ \forall t.$$

Then using the fact that $F(t,u)$ is the Riemann function ([2]) we get

$$g(t) = L_t \int_0^{(\cdot)} F(\cdot,u)g(u)du = 0.$$

3. Bayes Formula for n-ple Markov noise processes

We now obtain a Bayes formula for the filter in case the observation process Y is of the form (1) with $N(t) = \sum_{i=1}^n \varphi_i(t)Z_i(t)$ and $\underline{Z}(t)$ an n-dimensional Gaussian martingale.

Theorem 2. *Let X be a system process and Y be given by (1) with N as above and*

$$\beta(t,X) = \int_0^t \varphi(t)dM_z(u)\psi^*(u,X) \quad \textit{with} \;\; \psi(\cdot,X) \in L^2(M_z) \text{ a.e..}$$

Then for an $\mathcal{F}_T^X$-measurable and integrable function $g(T,X)$,

$$E\left(g(T,X)\,\middle|\mathcal{F}_t^Y\right) = \frac{\int g(T,x)e^{\int_0^t (\psi(u,x),d\hat{Y}(u)) - \frac{1}{2}\int_0^t \psi(u,x)dM_z(u)\psi^*(u,x)}dP_X(x)}{\int e^{\int_0^t (\psi(u,x),d\hat{Y}(u)) - \frac{1}{2}\int_0^t \psi(u,x)dM_z(u)\psi^*(u,x)}dP_X(x)}$$

where

$$\hat{Y}(t) = \underline{Z}(t) + \int_0^t \psi(u,X)dM_z(u).$$

Proof : First notice that, from lemma 1, $\beta(\cdot,X) \in H(R_N)$. Also, from (6), it follows that

$$||\beta(\cdot,x)||_t^2 = \int_0^t \psi(u)dM_z(u)\psi^*(u).$$

Then the theorem follows from Theorem 1 and equation (7). ∎

In case $Z_i(t) = \int_0^t g_i(u)dW(u)$, we get a more appealing one dimensional form because in this case

$$\beta(t,X) = \int_0^t F(t,u)\phi^*(u,X)d\lambda(u)$$

where

$$\phi^*(u,X) = \sum_{i=1}^n g_i(u)\psi_i(u,X)$$

and

$$E\left(g(T,X)\,\middle|\mathcal{F}_t^Y\right) = \frac{\int g(T,x)e^{\int_0^t \phi^*(u,x)d\hat{Y}(u) - \frac{1}{2}\int_0^t \phi^*(u,x)^2 d\lambda(u)}dP_X(x)}{\int e^{\int_0^t \phi^*(u,x)d\hat{Y}(u) - \frac{1}{2}\int_0^t \phi^*(u,x)^2 d\lambda(u)}dP_X(x)}$$

where $\hat{Y}(t) = \int_0^t \phi^*(u, X)du + W(t)$.

In particular, when the noise process (N_t) has derivative n_t satisfying the stochastic differential equation (8) and the observation process (y_t) is defined by

$$y_t = h(X_t) + n_t,$$

where $h(\cdot) \equiv h(X_.)$ is n times continuously differentiable, we get the following (corrected) formula due to Kunita [5]

$$E\left(g(X_t)\left|\mathcal{F}_t^Y\right.\right) = \frac{\int g(x_t) e^{\int_0^t L_s h(x_.)d\hat{Y}_s - \frac{1}{2}\int_0^t (L_s h(x_.))^2 ds} dP_X(x)}{\int e^{\int L_s h(x_.)d\hat{Y}_s - \frac{1}{2}\int_0^t (L_s h(x_.))^2 ds} dP_X(x)}$$

where

$$\begin{aligned} \hat{Y}(t) &= \int_0^t L_s h(X_.)ds + W(t) \\ &= \int_0^t a_n(u)dy^{(n-1)}(u) + \int_0^t \sum_{i=0}^{n-1} a_i(u)y^{(i)}(u)du. \end{aligned}$$

This follows immediately once we note that here

$$\beta(t, X) = h(X_t) = \int_0^t F(t, u)L_u h(X_.)du.$$

References

[1] Chatterji, S. D. and Mandrekar, V. (1978). Equivalence and singularity of Gaussian measures and applications. Probabilistic Analysis and Related Topics 1. (Ed. A. T. Barucha - Ried), Academic Press, New York, 169-197.

[2] Dolph, C. L. and Woodbury, M. A. (1952). On the relation between Green's functions and covariances of certain stochastic processes and its applications to unbiased linear prediction, 72, 519-550.

[3] Hida, T. (1960). Canonical representations of Gaussian processes and their applications. Mem. Coll. Sci. Kyoto, A33, 109-155.

[4] Kallianpur, G. (1980). Stochastic Filtering Theory, Springer - Verlag, New York.

[5] Kunita, H. (1991). Representation and stability of nonlinear filters associated with Gaussian noises, Stochastic Processes (Kallianpur Fest.) (Ed. Stamatis Combanis et al.) Springer - Verlag, New York, 201-210.

[6] Levy, P.(1957). Fonctions linearment Markovian d'ordre n, Math. Japonica, 4, 113-121.

[7] Mandal, Pranab K. and Mandrekar, V. (1998). Bayes formula for Gaussian noise processes and its applications. Preprint.

[8] Mandrekar, V. (1974). On the multiple Markov property of Levy - Hida for Gaussian processes, Nagoya Math. J., 54, 69-78.

[9] Pitt, L. (1975). Hida-Cramer multiplicity theory for multiple Markov processes and Goursat representations, Nagoya Math. J., 57, 99-128.

[10] Rozanov, Yu. (1961). Spectral theory of multidimensional stationary random processes with discrete time, Selected transl. in Math. Stat. and Prob. I, 127-132.