EXPLICIT SOLUTIONS TO A CLASS OF NONLINEAR FILTERING PROBLEMS

by

Eugene Wong

Memorandum No. UCB/ERL M80/16

12 February 1980

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Explicit Solutions to a Class of Nonlinear Filtering Problems

Eugene Wong

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

1. Introduction

Let Z_t be a stochastic process and let X_t be a process of the form

$$X_t = \int_0^t Z_s ds + W_t, \qquad t \geq 0$$

where W_t is a standard Wiener process independent of Z_t. The general filtering problem is to find effective ways of computing the conditional expectation

$$E[f(Z_t)|X_s, 0 \leq s \leq t]$$

for some function f.

Except when Z is of finite state, the Gaussian case and some recently discovered example [4] comprise the entire collection of cases where solutions, in some explicitly computable form, to the nonlinear filtering problem are known. The object of this paper is to add a small but possibly useful class of examples to this collection.

2. A Wiener Series Representation

Let $(\Omega,\mathcal{F},\mathcal{P})$ be a probability space. Let $\{Z_t,W_t, 0 \leq t \leq T\}$ be a pair of independent processes defined on $(\Omega,\mathcal{F},\mathcal{P})$ such that W is a standard Wiener process, and Z is a strong Markov process that is almost surely sample square-integrable. Consider an observation process

(2.1) $$X_t = \int_0^t Z_s ds + W_t, \quad 0 \le t \le T,$$

and denote $F_{xt} = \sigma(X_s, s \le t)$. It is well known (see e.g. [6]) that if we define a probability measure P_0 by

(2.2) $$\frac{dP_0}{dP} = \exp\{-\int_0^T Z_s dW_s - \frac{1}{2}\int_0^T Z_s^2 \, ds\}$$

then (Z,X) has the same distribution under P_0 as (Z,W) under P.

For a bounded f define the unnormalized estimator

(2.3) $$\pi_t f = E_0\{f(Z_t) \frac{dP}{dP_0} \Big| F_{xt}\}$$

To normalize, one would only need to write

(2.4) $$E[f(Z_t) | F_{xt}] = \frac{\pi_t f}{\pi_t 1}$$

where

(2.5) $$\pi_t 1 = L_t = E_0\{\frac{dP}{dP_0} \Big| F_{xt}\}$$

is simply the likelihood ratio.

Now, from (2.2) we have

(2.6) $$\frac{dP}{dP_0} = \exp\{\int_T Z_s dX_s - \frac{1}{2}\int_T Z_s^2 ds\}$$

and the exponential formula for multiple Wiener integrals yields [3]

(2.7) $$\frac{dP}{dP_0} = \sum_{n=0}^{\infty} Z_n \circ X^n$$

where $Z_0 \circ X^0 \equiv 1$ and for $n \geq 1$

$$(2.8) \qquad Z_n \circ X^n = \int_{0<t_1<..<t_n<T} Z_{t_1} Z_{t_2} \cdots Z_{t_n} \; X(dt_1) \ldots X(dt_n)$$

are desymmetrized multiple Wiener integrals. It now follows that

$$(2.9) \qquad \pi_t f = \sum_{n=0}^{\infty} \int_{0<t_1<\ldots<t_n<t} E_0(Z_{t_1} Z_{t_2} \cdots Z_{t_n} f(Z_t)) X(dt_1) \ldots . X(dt_n)$$

The process Z being identically distributed under either measures, E_0 in (2.9) can also be replaced by E.

Now, let Z be a diffusion process, with the density of Z_t being $P(z,t)$. Introduce an unnormalized conditional density $V(z,t)$ of Z_t given the observation by the relationship [6]

$$(2.10) \qquad \pi_t f = \int_{-\infty}^{\infty} V(z,t) \; f(z) dz$$

Then (2.9) reduces to [c.f. 5]

$$(2.11) \qquad V(z,t) = p(z,t) \sum_{n=0}^{\infty} m_n(z,\cdot,t) \circ X^n$$

with

$$(2.12) \qquad m_n(z,t_1,t_2,\ldots,t_n,t) = E(Z_{t_1} Z_{t_2} \cdots Z_{t_n} \big| Z_t = z)$$

and

$$(2.13) \qquad m_n(z,\cdot,t) \circ X^n = \int_{0<t_1<\ldots<t_n<t} m_n(z,t_1,\ldots,t_n,t) \; X(dt_1) \ldots . X(dt_n)$$

From the Markov property of Z, the functions m_n satisfy the recurrence relationships

$$(2.14) \qquad m_n(z,t_k,..,t_n,t) = E[Z_{t_n} m_{n-1}(Z_{t_n},t_1,...,t_n) \big| Z_t = z]$$

The main result of this paper is an explicit evaluation of these functions for a class of stationary Z.

3. Processes of the Pearson Class

We shall restrict our attention to a class of stationary diffusion processes Z_t that have a transition density of the forms

$$(3.1) \qquad p(z,t|z_0,t_0) = p(z) \sum_{k=0}^{\infty} e^{-\lambda_k(t-t_0)} \phi_k(z)\, \phi_k(z_0)$$

where p(z) is the stationary density and ϕ_k are orthonormal polynomials of degree k. Densities of the form (3.1) were introduced by Barrett and Lampard [1]. In [7] diffusion processes with such transition densities were exhaustively studied subject to the additional condition that p(z) is of the Pearson type [2]. It was found that such processes fall into three categories, corresponding to the classical Hermite, Laguerre and Jacobi polynomials respectively. In terms of the Fokker Planck equation for the trnasition density p

$$(3.2) \qquad \frac{1}{2}\frac{\partial^2}{\partial z^2}[\sigma^2(z)p] - \frac{\partial}{\partial z}[m(z)p] = \frac{\partial}{\partial t} p$$

these cases can be summarized as follows:

$$(3.3a) \qquad \sigma^2(z) = 2, \quad m(z) = -z$$

$\phi_k(z)$ are Hermite polynomials

$$(3.3b) \qquad z > 0,\ \sigma^2(z) = 2z,\ m(z) = (\alpha+1)-z,\ z \geq 0$$

$\phi_k(z)$ are Laguerre polynomials

(3.3c) $|z| < 1,\ \ \sigma^2(z) = 2(1-z^2),\ \ m(z) = (\alpha-\beta) - (\alpha+\beta+2)z \quad \alpha,\beta > -1$

$\phi_k(z)$ are Jacobi polynomials

Observe that $z\phi_k(z)$ is a polynomial of degree k+1. Furthermore, for any $j \leq k-2$ $z\phi_j(z)$ is a polynomial of degrees k-1 or less and hence is orthonormal to ϕ_k, i.e.,

$$\int p(z)\ z\phi_k(z)\ \phi_j(z)dz = 0 \qquad j \leq k-2.$$

It follows that $z\phi_k(z)$ is at most a linear combination of ϕ_k and $\phi_{k\pm 1}$. We shall write

$$(3.4) \qquad z\phi_k(z) = a_{k+1}\phi_{k+1}(z) + b_k\phi_k(z) + c_{k-1}\phi_{k-1}(z)$$

for the general 3-term recurrence relationship, and use this to evaluate the conditional moments $m_n(z,\cdot)$ explicitly.

We note that for any of these cases we have

$\lambda_0 = 0$ and $\phi_0(z) = 1$.

4. An Explicit Solution

We begin with the following observation:

Theorem 4.1. If Z is a stationary Markov process with a transition function of the form (3.1). Then, $m_n(z,\cdot)$ are of the form

$$(4.1) \qquad m_n(z,t_1,\ldots,t_n,t) = \sum_{p=0}^{n} \alpha_{np}(t_2-t_1,t_3-t_2,\ldots,t-t_n)\ \phi_p(z)$$

where α_{np} satisfy the recurrence relationship

(4.2) $$\alpha_{np}(t_2-t_1,\ldots,t-t_n)$$

$$= e^{-\lambda_p(t-t_n)} \; a_p\, \alpha_{n-1,p-1}(t_2-t_1,\ldots,t_n-t_{n-1})$$

$$+ b_p\, \alpha_{n-1,p}(t_2-t_1,\ldots,t_n-t_{n-1})$$

$$+ c_p\, \alpha_{n-1,p+1}(t_2-t_1,\ldots,t_n-t_{n-1})$$

$$n \geq p \geq 0$$

Proof: We note from (3.1) that

(4.3) $$E[\phi_k(Z_s)|Z_t = z]$$

$$= e^{-\lambda_k(t-s)} \phi_k(z), \qquad t \geq s$$

Hence, from (3.4) we have

$$m_1(z,t_1,t) = E[Z_{t_1}\big|Z_t = z]$$

$$= E[a_1\phi_1(Z_{t_1}) + b_0\phi_0(Z_{t_1})\big|Z_t = z]$$

$$= a_1 e^{-\lambda_1(t-t_1)}\phi_1(z) + b_0 e^{-\lambda_0(t-t_1)}\phi_0(z)$$

so that (4.1) holds for $n = 1$, and we have $\alpha_{10} = b_0 e^{-\lambda_0(t-t_1)} = b_0$, $\alpha_{11} = a_1 e^{-\lambda_1(t-t_1)}$.

Suppose that (4.1) holds for $k \leq n-1$. Then, from (2.14) we have

(4.4) $$m_n(z,t_1,\ldots,t_n,t)$$

$$= \sum_{p=0}^{n-1} \alpha_{n-1,p}(t_2-t_1,\ldots,t_n-t_{n-1})$$

$$E[Z_{t_n}\phi_p(Z_{t_n})\big|Z_t = z]$$

$$= \sum_{p=0}^{n-1} \alpha_{n-1,p}(t_2-t_1,\ldots,t_n-t_{n-1})$$

$$\{a_{p+1}\phi_{p+1}(z)e^{-\lambda_{p+1}(t-t_n)}$$

$$+ b_p\phi_p(z)e^{-\lambda_p(t-t_n)}$$

$$+ c_{p-1}\phi_{p-1}(z)e^{-\lambda_{p-1}(t-t_n)}\}$$

which is again of the form (4.1).

If we rearrange terms in (4.3), we get (4.2). ¤

In (4.2) let's adopt the convention that $\alpha_{np} = 0$ whenever $p > n$ or $n < 0$. Then the equation holds for any n and p. Observe that when $n = p$, we have

$$\alpha_{nn} = e^{-\lambda_n(t-t_n)} a_n \alpha_{n-1,n-1}$$

which can be solved immediately to yield

$$\alpha_{nn}(\tau_1,\tau_2,\ldots,\tau_n) = \prod_{k=1}^{n} a_k e^{-\lambda_k\tau_k}$$

and that in turn can be used to solve for $\alpha_{n\ n-1}$, etc. It is convenient to work with Laplace transforms and make a change in notation as follows:

$$(4.5)\qquad \hat{\alpha}_p^{(\nu)}(s_1,s_2,\ldots,s_{p+\nu}) = \int_0^\infty \cdots \int_0^\infty e^{-(s_1\tau_1+\ldots+s_{p+\nu}\tau_{p+\nu})}$$

$$\alpha_{p+\nu,p}(\tau_1,\tau_2,\ldots,\tau_{p+\nu})d\tau_1 \ldots d\tau_{p+\nu}$$

Then, (4.2) becomes

$$(4.6) \qquad \hat{\alpha}_p^{(\nu)}(s_1,s_2,\ldots,s_{p+\nu}) = \frac{1}{(s_{p+\nu}+\lambda_p)} \{a_p \hat{\alpha}_{p-1}^{(\nu)}(s_1,\ldots,s_{p+\nu-1})$$

$$+ b_p \hat{\alpha}_p^{(\nu-1)}(s_1,s_2,\ldots,s_{p+\nu-1})$$

$$+ c_p \, \hat{\alpha}_{p+1}^{(\nu-2)}(s_1,s_2,\ldots,s_{p+\nu-1})\}$$

which can be solved immediately to yield

$$(4.7) \qquad \hat{\alpha}_p^{(0)} = \prod_{k=1}^{p} \frac{a_k}{(s_k+\lambda_k)} \, , \quad \hat{\alpha}_0^{(0)} = 1$$

verifying the result that we obtained earlier for α_{nn}.

The general solution for $\hat{\alpha}_p^{(\nu)}$ is given as follows.

Theorem 4.2. Let u_k, $b_k^{(\nu)}$ and $c_k^{(\nu)}$ be defined as follows: $(k \geq 1, \ \nu \geq 1)$

$$(4.8) \qquad u_k = \prod_{j=1}^{k} (\frac{b_0}{s_j})$$

$$(4.9) \qquad b_k^{(\nu)} (\frac{b_k}{s_{k+\nu}+\lambda_k}) \prod_{j=1}^{k} (\frac{s_{j+\nu}+\lambda_j}{s_{j+\nu-1}+\lambda_j})$$

$$(4.10) \qquad c_k^{(\nu)} = 0 \qquad \nu = 1$$

$$= \frac{c_{k-1} a_k}{(s_{k+\nu-1}+\lambda_{k-1})(s_{k+\nu}+\lambda_k)} \prod_{j=1}^{k} (\frac{s_{j+\nu}+\lambda_j}{s_{j+\nu-2}+\lambda_j}) \qquad \nu \geq 2$$

For $\nu \geq 1$, $p \geq 0$ and $1 \leq k \leq p+1$, define a ν-dimensional row vector $a_{pk}^{(\nu)}$ as follows:

$$(4.11)\qquad a_{p1}^{(\nu)} = \left(b_1^{(\nu)},\ u_\nu\left(\frac{s_\nu c_1^{(\nu)}}{u_{\nu-1}}\right),\ u_\nu\left(\frac{s_{\nu-1} c_1^{(\nu-1)}}{u_{\nu-2}}\right),\ \ldots,\ u_\nu\left(\frac{s_2 c_1^{(\nu)}}{u_1}\right)\right)$$

$$a_{pk}^{(\nu)} = (b_k^{(\nu)},\ c_k^{(\nu)},\ 0 \ldots\ldots 0)\ ,\ 2 \leq k \leq p$$

$$a_{pp+1}^{(\nu)} = (0,\ c_{p+1}^{(\nu)},\ 0 \ldots\ldots 0)$$

Finally, define $\nu+1$ by ν matrices

$$(4.12)\qquad A_{pk}^{(\nu)} = \begin{bmatrix} a_{pk}^{(\nu)} \\ \delta_{pk} I_\nu \end{bmatrix}$$

where I_ν is the $\nu\times\nu$ identity matrix.

Then, $\hat{\alpha}_p^{(\nu)}$ are given as follows:

$$(4.13)\qquad \begin{bmatrix} \hat{\alpha}_p^{(\nu)} \\ \vdots \\ \hat{\alpha}_p^{(0)} \end{bmatrix} = \prod_{j=1}^{p} \frac{a_j}{(s_{j+\nu}+\lambda_j)} \Bigg[\sum_{k=0}^{\nu} u_k \left\{ \sum_{m_\nu=1}^{p+1} \sum_{m_{\nu-1}=1}^{m_\nu+1} \cdots \sum_{m_{k+1}=1}^{m_{k+2}+1} A_{pm_\nu}^{(\nu)} A_{m_\nu m_{\nu-1}}^{(\nu-1)} \cdots A_{m_{k+2}m_{k+1}}^{(k+1)} 1_{k+1} \right\} \Bigg]$$

when 1_k is the k-dimensional unit column vector.

<u>Proof</u>: We begin by iterating (4.6) in p and get

$$(4.14)\qquad \hat{\alpha}_p^{(\nu)} = \prod_{j=1}^{p} \frac{a_j}{(s_{j+\nu}+\lambda_j)}\ \hat{\alpha}_0^{(\nu)} + \sum_{m=1}^{p} \frac{1}{\prod_{j=1}^{m} \frac{a_j}{(s_{j+\nu}+\lambda_j)}} \left(\frac{b_m}{s_{m+\nu}+\eta_m}\right) \hat{\alpha}_m^{(\nu-1)} + \left(\frac{c_m}{s_{m+\nu}+\lambda_m}\right) \hat{\alpha}_{m+1}^{(\nu-2)}$$

for $p \geq 1$ and

$$(4.15) \qquad \hat{\alpha}_0^{(\nu)} = \frac{b_0}{(s_\nu+\lambda_0)} \hat{\alpha}_0^{(\nu-1)} + \frac{c_0}{(s_\nu+\lambda_0)} \hat{\alpha}_1^{(\nu-a)}$$

Now, denote for $p \geq 1$

$$(4.15) \qquad \hat{\alpha}_p^{(\nu)} = [\prod_{j=1}^{p} \frac{a_j}{(s_{j+\nu}+\lambda_j)} \gamma_p^{(\nu)}$$

Then, we have

$$(4.16) \qquad \gamma_p^{(\nu)} = \hat{\alpha}_0^{(\nu)} + \sum_{m=1}^{p} \frac{1}{\prod_{j=1}^{m} \frac{a_j}{(s_{j+\nu}+\lambda_j)}}$$

$$\left\{ (\frac{b_m}{s_{m+\nu}+\lambda_m}) \prod_{j=1}^{m} \frac{a_j}{(s_{j+\nu-1}+\lambda_j)} \gamma_m^{(\nu-1)} \right.$$

$$\left. + (\frac{c_m}{s_{m+\nu}+\lambda_m}) \prod_{j=1}^{m+1} \frac{a_j}{(s_{j+\nu-2}+\lambda_j)} \gamma_{m+1}^{(\nu-2)} \right\}$$

which simplies to yield

$$(4.17) \qquad \gamma_p^{(\nu)} = \hat{\alpha}_0^{(\nu)} + \sum_{m=1}^{p} b_m^{(\nu)} \gamma_m^{(\nu-1)} + \sum_{m=1}^{p+1} c_m^{(\nu)} \gamma_m^{(\nu-2)}$$

where $b_m^{(\nu)}$ and $c_m^{(\nu)}$ are as defined in (4.9) and (4.10).

Equation (4.15) can be iterated to yield

$$(4.18) \qquad \hat{\alpha}_0^{(\nu)} = \frac{b_0}{\prod_{j=1}^{\nu} (s_j+\lambda_0)} + \sum_{k=0}^{\nu-2} \frac{c_0 a_1 b_0^{\nu-k-2}}{\prod_{j=k+1}^{\nu} (s_j+\lambda_0)} (\frac{s_{k+1}+\lambda_0}{(s_{k+1}+\lambda_1)}) \gamma_1^{(k)}$$

which is of the form

$$(4.19) \qquad \hat{\alpha}_0^{(\nu)} = u_\nu + \sum_{k=0}^{\nu-2} s_{k+2}\, c_1^{(k+2)} (\frac{u_\nu}{u_{k+1}}) \gamma_1^{(k)}$$

With the use of (4.19), we can now rewrite (4.16) in the form of

$$(4.20)\qquad \begin{bmatrix} \gamma_p^{(\nu)} \\ \gamma_p^{(\nu-1)} \\ \vdots \\ \gamma_p^{(0)} \end{bmatrix} = \sum_{m=1}^{p+1} A_{pm}^{(\nu)} \begin{bmatrix} \gamma_m^{(\nu-1)} \\ \vdots \\ \gamma_m^{(0)} \end{bmatrix} + u_\nu \quad 1_{\nu+1}$$

where $A_{pm}^{(\nu)}$ are as defined by (4.12) and (4.11). Equation (4.20) can now be iterated in ν. With $\gamma_p^{(0)} = 1$ we get

$$(4.21)\qquad \begin{bmatrix} \gamma_p^{(\nu)} \\ \vdots \\ \gamma_p^{(0)} \end{bmatrix} = \sum_{k=0}^{\nu} u_k \left[\sum_{m_\nu=1}^{p+1} \sum_{m_{\nu-1}=1}^{m_\nu+1} \cdots \sum_{m_{k+1}=1}^{m_{k+2}+1} \left\{ A_{pm_\nu}^{(\nu)} A_{m_\nu m_{\nu-1}}^{(\nu-1)} \cdots A_{m_{k+2}m_{k+1}}^{(k+1)} \quad 1_{k+1} \right\} \right]$$

whence the desired result (4.13) follows immediately using (4.15). ¤

5. The Symmetric Case

There are some cases for which the polynomials $\phi_n(z)$ contain only even or odd terms according as n is even or odd respectively. This is the situation, for example, for Gegenbauer polynomials (which include both Chebyshev and Legendre polynomials), and most importantly for Hermite polynomials which correspond to Z_t being a Gaussian process. We shall refer to these cases collectively as the symmetric case.

For the symmetric case the coefficient b_k in the recurrence relationship (3.4) is necessarily zero for every k. It follows from (4.9) that $b_k^{(\nu)}$ are identically zero, and the result of theorem 4.2 simplies a great deal as is indicated as follows:

<u>Theorem 5.1</u>. For the symmetric case we have

$$\alpha_p^{(2\nu+1)} = 0$$

$$(5.1) \qquad \alpha_p^{(2\nu)} = \prod_{j=1}^{p} \frac{a_j}{(s_{2\nu+j}+\lambda_j)} \left\{ \sum_{m_\nu=1}^{p+1} \sum_{m_{\nu-1}=1}^{m_\nu+1} \cdots \sum_{m_1=1}^{m_2+1} c_{m_\nu}^{(2\nu)} c_{m_{\nu-1}}^{(2\nu-2)} \cdots c_{m_1}^{(2)} \right\}$$

<u>Proof</u>: Since $b_k^{(\nu)} \equiv 0$, (4.17) becomes

$$(5.2) \qquad \gamma_p^{(\nu)} = \sum_{m=2}^{p+1} c_m^{(\nu)} \gamma_m^{(\nu-2)} + \hat{\alpha}_0^{(\nu)}$$

and (4.18) now takes the form

$$(5.3) \qquad \hat{\alpha}_0^{(\nu)} = \left(\frac{c_0}{s_\nu}\right) \hat{\alpha}_1^{(\nu-2)}$$

with the use of (5.2) for $\hat{\alpha}_0^{(\nu)}$, (5.2) can be rewritten as

$$(5.4) \qquad \gamma_p^{(\nu)} = \sum_{m=1}^{p+1} c_m^{(\nu)} \gamma_m^{(\nu-2)}$$

where $c_m^{(\nu)}$ is given by (4.10). Since $\gamma_p^{(0)} = 1$ and $\gamma_p^{(1)} = 0$, we have $\gamma_p^{(\nu)} = 0$ for all ν odd, and

$$(5.5) \qquad \gamma_p^{(2\nu)} = \sum_{m_\nu=1}^{p+1} \sum_{m_{\nu-1}=1}^{m_\nu+1} \sum_{m_1=1}^{m_2+1} c_{m_\nu}^{(2\nu)} c_{m_{\nu-1}}^{(2\nu-2)} \cdots c_{m_1}^{(2)}$$

whence (5.1) follows. ⌑

It is interesting to note that in the Gaussian case (c.f. 33a) the terms $c_k^{(\nu)}$ are given by

$$(5.6) \qquad c_k^{(\nu)} = \frac{k}{(s_{\nu-1}+1)(s_\nu+2)} \prod_{j=1}^{k-2} \left(\frac{s_{j+\nu}+j}{s_{j+\nu}+j+2} \right)$$

Acknowledgement. Most of the work was done while I was at MIT, as a guest of the Department of Electrical Engineering and Computer Science and the Laboratory for Information and Decision Systems. In particular, the results were inspired by the work of Mitter and Ocone [5] and by many conversations with Professor S. K. Mitter.

Research sponsored by the U.S. Army Research Office Grant DAAG29-78-G-0186.

References

1. J.F. Barrett and D.G. Lampard. "An expansion for some second-order probability distributions and its application to noise problems," "IRE Trans. IT-1 (1955) 10-15.

2. H. Cramer. Mathematical Methods of Statistics, Princeton Univ. Press, Princeton, N.J., 1946, 248-249.

3. K. Ito. "Multiple wiener integrals," J. Math. Soc. Japan 3 (1951) 157-169.

4. S.K. Mitter. "On the analogy between the mathematical problems of nonlinear filtering and quantum physics, "Ricerche di Automatica, in press.

5. S.K. Mitter and D. Ocone. "Multiple integral expansion for non-linear filtering," Proceedings of the 18th IEEE Conference on Decision of Control, Fort Lauderdale, December, 1979.

6. E. Wong. Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York, 1971.

7. E. Wong. "The construction of a class of stationary Markoff processes, "AMS Proceedings of Symposium in Applied Math. 16 (1963) 264-276.