FIXED-POINT SMOOTHING AS A FILTERING PROBLEM

John B. Moore and Brian D.O. Anderson
Department of Electrical Engineering
University of Newcastle
New South Wales, 2308
Australia

ABSTRACT

The Fixed-Point smoothing problem for nonlinear continuous time systems is shown to be a filtering problem. The smoothing problem is solved using filtering theory, and properties of the solution are examined.

1. INTRODUCTION

We shall be concerned with the fixed-point smoothing problem for nonlinear continuous time systems governed by the standard type of equations:

$$dx_t = f(x_t,t)dt + g(x_t,t)d\beta$$
$$dz_t = h(x_t,t)dt + dw \qquad (1)$$

Here, $d\beta$ is an increment in a vector Wiener process with a derivative of covariance $I\delta(t-\tau)$, and dw is an increment in a vector Wiener process with a derivative of covariance $R_t\delta(t-\tau)$. The two processes are assumed independent. It is well known that, with suitable conditions on $f(\cdot,\cdot)$, etc., there exists a partial differential equation, the conditional Fokker Planck equation, for the filtering density $p(x_t/Z_t)$, where Z_t denotes the process z over the interval $[t_0,t]$. Our aim here is to derive, by a simple application of these known filtering results, an equation for the evolution with t of the fixed-point smoothing density $p(x_{t_0}/Z_t)$. We thereby obtain a result of [3], but by a far simpler procedure. We shall also indicate some consequences of this result.

2. AN AUGMENTED FILTERING PROBLEM

Suppose x_t is an n-vector. Let x_{at} denote a 2n-dimensional vector, satisfying

$$dx_{at} = \begin{bmatrix} dx_t \\ d\bar{x}_t \end{bmatrix} = \begin{bmatrix} f(x_t,t) \\ 0 \end{bmatrix} dt + \begin{bmatrix} g(x_t,t) \\ 0 \end{bmatrix} d\beta$$
$$= f_a(x_{at},t) + g_a(x_{at},t)d\beta \qquad (2)$$

$$dz = h(x_t,t)dt + dw$$
$$= h_a(x_{at},t)dt + dw \qquad (3)$$

The functions f_a, g_a and h_a are defined in an obvious way. The initial density of x_{at} is defined by $\bar{x}_{t_0} = x_{t_0}$ and an initial density $p(x_{t_0})$ for x_{t_0}. The first n entries of x_{at} are the entries of x_t, so that a filtering density of x_{at} will yield a filtering density of x_t. The second n entries of x_{at}, grouped into the vector $\bar{x}_t$, satisfy $d\bar{x}_t = 0$, so that $\bar{x}_t = x_{t_0}$. Therefore, a filtering density of $\bar{x}_t$, derivable from a filtering density of x_{at}, will yield a smoothed density for x_{t_0}. Solution of

---

Work supported by the Australian Research Grants Committee.

the 2n-dimensional filtering problem associated with x_{at} will therefore solve the filtering and smoothing problems associated with x_t. This idea was first used in [4], but in the context of linear smoothing only.

3. CONDITIONAL FOKKER-PLANCK EQUATION

Let us set $p_a = p(x_{at}/Z_t)$, $p_f = p_f(x_t/Z_t)$, $p_s = p(x_{t_0}/Z_t)$, $h_t = h_a(x_{at},t) = h(x_t,t)$, $\hat{h}_t = E\{h_a(x_{at},t)/Z_t\} = E\{h(x_t,t/Z_t)\}$ and $\hat{\hat{h}}_t = E\{h(x_t,t)/Z_t,x_{t_0}\}$. The conditional Fokker-Planck equation for p_a is [1, 2],

$$dp_a = -\sum_{i=1}^{m} \frac{\partial(f_{ai}p_a)}{\partial(x_{at})_i} + \frac{1}{2}\sum_{i,j=1}^{2n} \frac{\partial^2(g_a g_a' p_a)}{\partial(x_{at})_i \partial(x_{at})_j} + (h_t' - \hat{h}_t')R_t^{-1}(dz_t - \hat{h}_t dt)p_a \qquad (4)$$

Using the definitions of f_a and g_a implicit in (2), we obtain easily

$$dp_a = -\sum_{i=1}^{n} \frac{\partial(f_{ai}p_a)}{\partial(x_t)_i} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2(gg'p_a)}{\partial(x_t)_i \partial(x_t)_j} + (h_t' - \hat{h}_t')R_t^{-1}(dz_t - \hat{h}_t dt)p_a \qquad (5)$$

4. EQUATIONS FOR FILTERING AND SMOOTHING

Equation (5) will yield equations for the filtering and smoothing densities, p_f and p_s, by integration over the $x_{t_0} = \bar{x}_t$ space and x_t space respectively. Integration over the $\bar{x}_t$ space yields, in a straightforward way, the usual equation

$$dp_f = -\sum_{i=1}^{n} \frac{\partial(f_i p_f)}{\partial(x_t)_i} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2(gg'p_f)}{\partial(x_t)_i \partial(x_t)_j} + (h_t' - \hat{h}_t')R_t^{-1}(dz_t - \hat{h}_t dt)p_f \qquad (6)$$

Now consider integration over the x_t space. The first two terms on the right hand side integrate to zero, assuming p_a and $\frac{\partial p_a}{\partial(x_t)_i}$ approach zero as $||x_t|| \to \infty$. (This is a reasonable assumption, and a standard one used in deriving an equation for the evolution of the mean of an arbitrary function $\phi(x_t,t)$, see [1]). In integrating the third term on the right side, we use the following:

$$\int h'(x_t,t)R_t^{-1}(dz_t - \hat{h}_t dt)p(x_t,\bar{x}_t/Z_t)dx_t$$
$$= \int h'(x_t,t)p(x_t/\bar{x}_t,Z_t)dx_t[p(\bar{x}_t/Z_t)R_t^{-1}(dz_t - \hat{h}_t dt)]$$
$$= \hat{\hat{h}}_t' R_t^{-1}(dz - \hat{h}_t dt)p(\bar{x}_t/Z_t)$$

The result is

$$dp_s = (\hat{\hat{h}}_t' - \hat{h}_t')R_t^{-1}(dz_t - \hat{h}_t dt)p_s \qquad (7)$$

This equation is of course initialized by $p_s(x_{t_0}) = p_f(x_{t_0}) = p(x_{t_0})$.

5. SMOOTHED MEANS

Let $\phi(x_{t_0})$ be an arbitrary function of x_{t_0}, including perhaps x_t or $x_{t_0}x_{t_0}'$. Let $\phi_s = E[\phi x_{t_0}/Z_t]$. Then multiplication of (7) by ϕx_{t_0} and integration with respect to x_{t_0} leads easily to

$$d\hat{\phi}_s = \{E[\phi(x_{t_0})h'(x_t,t)/Z_t] - \hat{\phi}_s\hat{h}'\}R_t^{-1}(dz_t - \hat{h}_t dt) \qquad (8)$$

By setting $\phi(x_{t_0}) = x_{t_0}$, we get an equation for the smoothed mean. In obvious notation,

$$d(\hat{x}_{t_0})_s = [(\widehat{x_{t_0}h_t'})_s - (\hat{x}_{t_0})_s\hat{h}_t']R_t^{-1}(dz_t - \hat{h}_t dt) \qquad (9)$$

By setting $\phi(x_{t_0}) = x_{t_0}x_{t_0}'$, we can obtain an equation for $(\widehat{x_{t_0}x_{t_0}'})_s$. The Ito differential rule will yield an equation for $(\hat{x}_{t_0})_s(\hat{x}_{t_0})_s'$ in terms of the equation for $(\hat{x}_{t_0})_s$, and thence an equation for the error variance associated with $(\hat{x}_{t_0})_s$ follows. With x_i used for the i-th entry of x_{t_0}, the i-j entry of the error variance matrix $E\{[x(t_0) - (\hat{x}_{t_0})_s][x(t_0) - (\hat{x}_{t_0})_s]'\}$, call it P_{ij}, satisfies, with arguments omitted for clarity,

$$dP_{sij} = (\widehat{x_i x_j h'} - \widehat{x_i x_j}\hat{h}' - \hat{x}_i\widehat{x_j h'} - \hat{x}_j\widehat{x_i h'} + 2\hat{x}_i\hat{x}_j\hat{h}')R_t^{-1}(dz - \hat{h}dt) - (\widehat{x_i h'} - \hat{x}_i\hat{h}')R^{-1}(\widehat{hx_j} - \hat{h}\hat{x}_j) \qquad (10)$$

This calculation parallels very closely the corresponding calculation for the error variance associated with filtering, see [1].

6. RELATION WITH OTHER RESULTS

Suppose one is interested in a smoothed density $p(x_\tau/z_\sigma,\ \sigma \in [t_0,t))$. Then one could obtain this by computing the filtered density $p(x_\sigma/Z_\sigma)$ forward in time to τ, and then using (7), initialized at τ with $p(x_\tau/Z_\tau)$. This extends the fixed-point smoothing idea marginally. With this extension in mind, one can also obtain a formula of Lo [5] easily. Integrate (7) from τ to t with initial condition $p(x_\tau/Z_\tau)$. One obtains

$$p(x_\tau/Z_t)-p(x_\tau/Z_\tau) = \int_\tau^t p(x_\tau/Z_\sigma)(\hat{\hat{h}}_\sigma-\hat{h}_\sigma)R_\sigma^{-1}(dz_\sigma-\hat{h}_\sigma d\sigma) \tag{11}$$

Here, $\hat{\hat{h}}_\sigma$ is $E[h(x_\sigma,\sigma)/x_\tau,Z_\sigma]$.

It is now well-known that the process with increment $dI_t = dz_t-\hat{h}_t dt$ is a Wiener process whose derivative has covariance $R_t\delta(t-\tau)$. From (7), we have, using the Ito differential rule,

$$d(\ln p_s) = (\hat{\hat{h}}_t'-\hat{h}_t')R_t^{-1}dI_t - \frac{1}{2}(\hat{\hat{h}}_t'-\hat{h}_t')R_t^{-1}(\hat{\hat{h}}_t-\hat{h}_t)dt$$

from which

$$p(x_\tau/Z_t) = p(x_\tau/Z_\tau)\exp\int_\tau^t [\hat{\hat{h}}_\sigma'-\hat{h}_\sigma']R_\sigma^{-1} [dI_\sigma - \frac{1}{2}(\hat{\hat{h}}_\sigma-\hat{h}_\sigma)d\sigma] \tag{12}$$

This formula also appears in [5].

If one multiplies (11) by x_τ and integrates with respect to x_τ, one obtains a relation between the smoothed and filtered means of x_τ as follows:

$$\hat{x}_{\tau/t} - \hat{x}_{\tau/\tau} = \int_\tau^t E\{x_\tau(h_\sigma'-\hat{h}_\sigma')/Z_\sigma\}R_\sigma^{-1}(dz-\hat{h}_\sigma d\sigma) \tag{13}$$

In deriving this equation, one uses the fact that

$$\int_\tau^t x_\tau p(x_\tau/Z_\sigma)\hat{\hat{h}}_\sigma' dx_\tau = \int_\tau^t x_\tau p(x_\tau/Z_\sigma)E[h_\sigma'/x_\tau,Z_\sigma]dx_\tau$$

$$= \int_\tau^t E[x_\tau h_\sigma'/x_\tau,Z_\sigma]p(x_\tau/Z_\sigma)dx_\tau$$

$$= E[x_\tau h_\sigma'/Z_\sigma]$$

Equation (13) appears to be originally due to Frost [6]. It is also an immediate consequence of integrating (9).

7. APPROXIMATE FORMULAS

It is of interest to derive approximate formulas for the smoothed mean and error variance. Though these could be derived from (9) and (10), they follow more directly from known approximate filtering formulas applied to the x_{at} process. Once one has perceived the fact that the smoothing problem can be tackled as a filtering problem, there is no conceptual difficulty in obtaining the formulas.

In order to illustrate the derivation of approximate fixed-point smoothing results, we choose just one approximate nonlinear filter - the truncated second order filter [1] - and carry out the procedure for this case.

With P_{aij} the i-j entry of the error variance matrix $E[x_{at}-\hat{x}_{at})(x_{at}-\hat{x}_{at})']$ we define the following quantities

$$(P_a\partial^2 f_a)_i \triangleq \sum_{j,k=1}^{2n} P_{ajk}\frac{\partial^2 f_{ai}(\hat{x}_a,t)}{\partial x_{aj}\partial x_{ak}}$$

$$(P_a\partial^2 f_a) \triangleq [(P_a\partial^2 f_a)_i, \ldots (P_a\partial^2 f_a)_{2n}]'$$

$$F_a \triangleq \left(\frac{\partial f_{ai}(\hat{x}_a,t)}{\partial x_{aj}}\right)$$

With P_{ij} the i-j entry of $E[(x_t-\hat{x}_t)(x_t-\hat{x}_t)]$ corresponding expressions for $(P\partial^2 f)$ and F may be written down. Likewise the quantities $(P_a\partial^2 h_a)$, H_a, $(P\partial^2 h)$ and H may be defined.

The truncated second order filter associated with the augmented system (2) - (3) is, [1],

$$d\hat{x}_a = [f_a(\hat{x}_{at},t) + \tfrac{1}{2}(P_a\partial^2 f_a)]dt + P_a H_a' R^{-1}\{dz - [h_a(\hat{x}_{at},t) + \tfrac{1}{2}(P_a\partial^2 h_a)]dt\} \tag{14}$$

$$dP_a = [F_aP_a+P_aF_a' + \widehat{g_ag_a'} - P_aH_a'R^{-1}H_aP_a]dt - \tfrac{1}{2}P_a(P_a\partial^2 h_a)'R^{-1} \{dz-[h_a(\hat{x}_{at},t) + \tfrac{1}{2}(P_a\partial^2 h_a)]dt\} \tag{15}$$

The quantity $\widehat{g_a g_a'}$ may be evaluated as described in [1], where an explicit formula is also given for the scalar state case. The partitioning of (2) and (3) can now be applied to this equation. It is not difficult to establish the following relationships

$$P_a = \begin{bmatrix} P & P_{fs}' \\ P_{fs} & P_s \end{bmatrix} , \quad h_a(x_{at},t) = h(x_t,t)$$

$$f_a(x_{at},t) = \begin{bmatrix} f(x_t,t) \\ 0 \end{bmatrix} , \quad g_a(x_{at},t) = \begin{bmatrix} g(x_t,t) \\ 0 \end{bmatrix}$$

$$(P_a \partial^2 h_a) = (P\partial^2 h) \quad , \quad (P_a \partial^2 f_a) = \begin{bmatrix} (P\partial^2 f) \\ 0 \end{bmatrix}$$

$$H_a = [\, H \quad 0 \,] \quad , \quad F_a = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix}$$

Application of these to (14) and (15) yields firstly, the usual truncated second order filter formulas for (1) as

$$d\hat{x} = [f(\hat{x},t) + \tfrac{1}{2}(P\partial^2 f)]dt + PH'R^{-1}d\tilde{z} \qquad (16)$$

$$d\tilde{z} = dz - [h(\hat{x},t) + \tfrac{1}{2}(P\partial^2 h)]dt \qquad (17)$$

$$dP = [FP+PF'+\widehat{gg'}-PH'R^{-1}HP]dt - \tfrac{1}{2}(P\partial^2 h)'R^{-1}d\tilde{z} \quad (18)$$

and secondly, truncated second order fixed-point smoothing formulas for (1) as

$$d(x_{t_0})_s = P_{fs}H'R^{-1}d\tilde{z} \qquad (19)$$

$$dP_{fs} = [P_{fs}F'-P_{fs}H'R^{-1}HP]dt - \tfrac{1}{2}P_{fs}(P\partial^2 h)'R^{-1}d\tilde{z} \qquad (20)$$

$$dP_s = [-P_{fs}H'R^{-1}HP_{fs}']dt - \; P_s(P\partial^2 h)'R^{-1}d\tilde{z} \qquad (21)$$

Observe that these latter smoothing equations are linear and thus relatively easy to solve.

REFERENCES

[1] A.H. Jazwinski, "Stochastic Processes and Filtering Theory," Academic Press, New York, 1970.

[2] L.E. Zachrisson, "On Optimal Smoothing of Continuous Time Kalman Processes," *Information Sciences*, Vol. 1, 1969, pp. 143-172.

[3] G.M. Lee, "Nonlinear Interpolation," *IEEE Transactions on Information Theory*, Vol. IT-17, No. 1, January 1971, pp. 45-49.

[4] H.J. Kushner, "On the differential equations satisfied by conditional probability densities of Markov processes with applications," *SIAM Journal on Control*, Vol. 2, No. 1, 1964, pp. 106-119.

[5] J.T. Lo, "On Optimal Nonlinear Estimation - Part I: Continuous Observations," *Proceedings of the 8th Annual Allerton Conference on Circuit and System Theory*, October 1970.

[6] P.A. Frost, "Estimation in Continuous-Time Nonlinear Systems," *Ph.D. Dissertation*, Department of Electrical Engineering, Stanford University, California, May 1968.