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:

\[ dx_t = f(x_t,t)dt + g(x_t,t)d\theta \]
\[ dz_t = h(x_t,t)dt + dw \]

Here, \( d\theta \) is an increment in a vector Wiener process with a derivative of covariance \( I(t-t) \), and \( dw \) is an increment in a vector Wiener process with a derivative of covariance \( R(t-t) \). 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

\[ dx_{at} = \begin{bmatrix} dx_t \\ dz_t \end{bmatrix} = \begin{bmatrix} f(x_t,t)dt + g(x_t,t)d\theta \\ h(x_t,t)dt + dw \end{bmatrix} \]

\[ = f_a(x_{at},t)dt + g_a(x_{at},t)d\theta \]  (2)

\[ dz = h(x_t,t)dt + dw \]

The functions \( f_a, g_a \), and \( h_a \) are defined in an obvious way. The initial density of \( x_{at} \) is defined by \( x_{t_0} = x_{t_0} \) 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 \( \tilde{x}_t \), satisfy \( d\tilde{x}_t = 0 \), so that \( \tilde{x}_t = x_{t_0} \). Therefore, a filtering density of \( \tilde{x}_t \), 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 \( x_{t} \) 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 
\[
p_a = p(x_{at}/Z_{t}), \quad p_f = p_f(x_{t}/Z_{t}), \quad p_s = p(x_{t_0}/Z_{t}),
\]
\( h_{t} = h(x_{at}, t) = h(x_{t}, t), \)
\( \hat{h}_{t} = E(h_{a}(x_{at}/t)/Z_{t}) = E(h(x_{t}/t/Z_{t})) \) and
\( \hat{h}_{t} = E(h(x_{t}/t)/Z_{t}/x_{t_0}). \) 
The conditional Fokker-Planck equation for \( p_a \) is [1, 2],
\[
dp_a = - \sum_{i=1}^{2n} \frac{\partial (f_{a} p_{a})}{\partial (x_{at})_i} + \frac{1}{2} \sum_{i,j=1}^{2n} \frac{\partial^{2} (g_{a} g_{a} p_{a})}{\partial (x_{at})_i \partial (x_{at})_j} + (h_{t} - \hat{h}_{t}) R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt) p_{a}
\]

Using the definitions of \( f_{a} \) and \( g_{a} \) implicit in (2), we obtain easily
\[
dp_a = - \sum_{i=1}^{n} \frac{\partial (f_{a} p_{a})}{\partial (x_{at})_i} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^{2} (g_{a} g_{a} p_{a})}{\partial (x_{at})_i \partial (x_{at})_j} + (h_{t} - \hat{h}_{t}) R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt) p_{a}
\]

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 \( x_{t_0} = x_{t} \) space and \( x_{t} \) space respectively. Integration over the \( x_{t_0} \) space yields, in a straightforward way, the usual equation
\[
dp_{f} = - \sum_{i=1}^{n} \frac{\partial (f_{a} p_{a})}{\partial (x_{t})_i} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^{2} (g_{a} g_{a} p_{a})}{\partial (x_{t})_i \partial (x_{t})_j} + (h_{t} - \hat{h}_{t}) R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt) p_{f}
\]

Now consider integration over the \( x_{t} \) space. The first two terms on the right hand side integrate to zero, assuming \( p_{a} \) and \( \frac{\partial p_{a}}{\partial (x_{t})_j} \) approach zero as \( ||x_{t}|| \rightarrow \infty \). 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:
\[
\int h'(x_{t}, t) R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt) p(x_{t}, x_{t}/Z_{t}) dx_{t}
\]
\[
= \int h'(x_{t}, t) p(x_{t}/x_{t}, Z_{t}) dx_{t} \int [p(x_{t}/Z_{t}) R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt)]
\]
\[
= h_{t} R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt) p(x_{t}/Z_{t})
\]
The result is
\[
dp_{s} = (\hat{h}_{t} - h_{t}) R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt) p_{s}
\]

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 \( x_{t_0} \) or \( x_{t_0} x_{t_{0}} \). Let
\[
\phi(x_{t_0}) = E[\Phi(x_{t_0}/Z_{t})].
\]
Then multiplication of (7) by \( \phi(x_{t_0}) \) and integration with respect to \( x_{t_0} \) leads easily to
\[
d\hat{\phi}_{s} = (\phi(x_{t_0}) h'_{t} x_{t_0}/Z_{t} - \hat{\phi}_{s} R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt) p_{s}
\]

By setting \( \phi(x_{t_0}) = x_{t_0} \), we get an equation for the smoothed mean. In obvious notation,
\[
d(\hat{x}_{t_0}) = (\hat{x}_{t_0} - \hat{x}_{t_0}) R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt)
\]

By setting \( \phi(x_{t_0}) = x_{t_0} x_{t_{0}} \), we can obtain an equation for \( (x_{t_0} x_{t_{0}}) \). The Ito differential rule will yield an equation for \( (x_{t_0} x_{t_{0}}) \) in terms of the equation for \( (x_{t_0}) \), and thence an equation for the error variance associated with \( (x_{t_0}) \) follows. With \( x_{i} \) used for the i-th entry of \( x_{t_0} \), the i-j entry of the error variance matrix \( E[(x_{t_0} - (\hat{x}_{t_0}))^{T}(x_{t_0} - (\hat{x}_{t_0}))] \), call it \( P_{ij} \), satisfies, with arguments omitted for clarity,
\[
dP_{ij} = (x_{i} x_{j} h'_{i} x_{j} - x_{i} x_{j} h'_{i} x_{j} + x_{i} x_{j} h'_{i} x_{j} + 2 x_{i} x_{j} h'_{i} x_{j})
\]
\[
- (x_{i} x_{j} h'_{i} x_{j}) R_{t}^{-1}(dz_{t} - \hat{h}_{t} dt)
\]

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 \( p(x_T/\mathbf{z}_\tau, c \in (t_0, t)) \). Then one could obtain this by computing the filtered density \( p(x_T/\mathbf{z}_\tau) \) forward in time to \( \tau \), and then using (7), initialized at \( \tau \) with \( p(x_T/\mathbf{z}_\tau) \). 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 \( \tau \) to \( t \) with initial condition \( p(x_T/\mathbf{z}_\tau) \). One obtains

\[
p(x_T/\mathbf{z}_\tau) - p(x_T/\mathbf{z}_t) = \int_\tau^t p(x_T/\mathbf{z}_\tau)(\hat{h}_0 - \hat{h}_\tau)R_\tau^{-1}(dz_\tau - \hat{h}_\tau dt)
\]

Here, \( \hat{h}_0 \) is \( E[h(x_{0,0})/x_{0,T}] \).

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

\[
d\ln p_s = (\hat{h}_t - \hat{h}_\tau)R_t^{-1}d\mathbf{I}_t - \frac{1}{2}(\hat{h}_t - \hat{h}_\tau)R_t^{-1}(\hat{h}_t - \hat{h}_\tau)dt
\]

from which

\[
p(x_T/\mathbf{z}_t) = p(x_T/\mathbf{z}_\tau)\exp\left[\int_\tau^t E[\hat{h}_t - \hat{h}_\tau]/\mathbf{z}_\tau]R_t^{-1}(dz_\tau - \hat{h}_\tau dt)\right]
\]

This formula also appears in [5].

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

\[
\hat{x}_t/\tau - \hat{x}_t/\tau = \int_\tau^t E[x_t(h_0 - \hat{h}_\tau)/\mathbf{z}_\tau]R_\tau^{-1}(dz_\tau - \hat{h}_\tau dt)
\]

In deriving this equation, one uses the fact that

\[
\int_\tau^t x_T p(x_T/\mathbf{z}_\tau)\hat{h}_\tau dx_T = \int_\tau^t x_T p(x_T/\mathbf{z}_\tau)E[h_0 - h_T/\mathbf{z}_\tau]dx_T
\]

\[
= \int_\tau^t E[x_T h_T/\mathbf{z}_\tau]p(x_T/\mathbf{z}_\tau)dx_T
\]

\[
= E[x_T h_T/\mathbf{z}_\tau]
\]

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 \( x_T \) 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 \( P_{ij} \) 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

\[
(\sum_{a} P_{ij}) \quad \frac{\partial^2 f_a}{\partial x_j \partial x_k} \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j}
\]

\[
(P_{ij}) \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j}
\]

\[
(F_a) \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j} \quad \frac{\partial^2 f_a}{\partial x_j}
\]

With \( P_{ij} \) the i-j entry of \( E[(x_{at} - \hat{x}_{at})(x_{at} - \hat{x}_{at})'] \), corresponding expressions for \( (P_{ij}) \) and \( F_a \) may be written down. Likewise the quantities \( (P_{ii}^2 h_a), H_a, (P_{ii}^2 h) \) and \( H \) may be defined.

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

\[
d\hat{\mathbf{x}}_a = [f_a(\hat{x}_{at}, t) + \frac{1}{2}(P_{ij} \sum_{a} h_a)]dt
\]

\[
+ P_{ij} H_a R_a^{-1}(dz - h_a(\hat{x}_{at}, t) + \frac{1}{2}(P_{ij} \sum_{a} h_a)dt)
\]

\[
dP_a = [F_a P_a + P_a F_a + \sum_{a} F_a R_a^{-1} H_a P_a]dt
\]

\[
- \frac{1}{2} P_{ij} (P_{ij} h_a)^{-1} (dz - h_a(\hat{x}_{at}, t) + \frac{1}{2}(P_{ij} \sum_{a} h_a)dt)
\]

222
The quantity $g_{a}g_{a}^{T}$ 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

\[
P_{a} = \begin{bmatrix} P & P_{a}^{T} \\ P_{a} & P_{a}^{T} \end{bmatrix}, \quad h_{a}(x_{at},t) = h(x_{t},t)
\]

\[
f_{a}(x_{at},t) = \begin{bmatrix} f(x_{t},t) \\ 0 \end{bmatrix}, \quad g_{a}(x_{at},t) = \begin{bmatrix} g(x_{t},t) \\ 0 \end{bmatrix}
\]

\[
(P_{a}P_{a}^{T}) = (P_{a}^{T}P_{a}), \quad (P_{a}P_{a}^{T}) = \begin{bmatrix} P_{a}P_{a}^{T} & 0 \\ 0 & 0 \end{bmatrix}
\]

\[
H_{a} = \begin{bmatrix} H & 0 \end{bmatrix}, \quad F_{a} = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix}
\]

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

\[
\dot{x} = [f(x_{t},t) + \frac{1}{2}(P_{a}^{T}P_{a})]dt + PH^{-1}d\bar{z}
\]

\[
\dot{\bar{z}} = dz - [h(x_{t},t) + \frac{1}{2}(P_{a}^{T}P_{a})]dt
\]

\[
\dot{P} = [FP+PF^{T} - PH^{-1}HP]dt - \frac{1}{2}(P_{a}^{T}P_{a})R^{-1}d\bar{z}
\]

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

\[
\dot{x}_{s} = P_{a}^{T}H^{-1}d\bar{z}
\]

\[
\dot{P}_{s} = [P_{a}^{T}F^{T} - P_{a}^{T}H^{-1}HP]dt - \frac{1}{2}P_{s}(P_{a}^{T}P_{a})R^{-1}d\bar{z}
\]

\[
\dot{P}_{s} = [-P_{a}^{T}H^{-1}HP_{s}]dt - P_{s}(P_{a}^{T}P_{a})R^{-1}d\bar{z}
\]

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

REFERENCES


