SMOOTHING OF NOISY RANDOM-TELEGRAPH-TYPE SIGNALS

David J. Clements and Brian D.O. Anderson
Department of Electrical Engineering
University of Newcastle, New South Wales, Australia

Abstract

For a class of Markov processes in additive gaussian white noise, both the equations for the optimal fixed-point smoother and those for a sub-optimal fixed-lag smoother are presented. Simulation results for the random telegraph wave are discussed.

1. INTRODUCTION

It has been shown in the literature\cite{1} that fixed-lag smoothers for both the continuous and discrete-time linear-gaussian problems exhibit the following desirable properties: (1) the expected value of the mean square error is a non-increasing function of the lag and (2) for a value of lag of the order of the dominant time constant of the associated (zero-lag) filter, practically all of the possible improvement that smoothing offers over filtering is attained. Further, for stationary time-invariant processes\cite{2}, the ratio of the fixed-lag error to the filtering error decreases as the signal-to-noise ratio increases. Hence, for those applications where a time delay between the estimate of the state of a process and the state itself is acceptable, it is worthwhile looking at the possibility of using fixed-lag smoothing rather than filtering.

Intuitively, it seems plausible that these properties should in fact hold for a larger class of processes than that above. In this paper, we derive and examine properties of a sub-optimal fixed-lag smoothed estimate of the states of a class of Markov processes in the presence of additive gaussian white noise. Simulation of the fixed-lag smoother for the special case of the random telegraph wave suggests that the above properties do in fact hold for the class of processes under study.

2. A NONLINEAR FILTERING RESULT

Suppose \( \{x(t), t \geq 0\} \) is a Markov process with a distinct states \( a_1, \ldots, a_n \) and stationary transition probabilities \( P_{ij}(h) = \Pr(x(t+h) = a_j | x(t) = a_i) \) satisfying

\[
P_{ij}(h) = \begin{cases} 1 - \nu_{ij} h + o(h) & \text{if } i = j \\ \nu_{ij} h + o(h) & \text{if } i \neq j \end{cases}
\]

(1)

where \( h > 0 \) and \( \nu_{ij}, \nu_{jj} \) are nonnegative constants such that \( \sum_{j} \nu_{ij} = \nu_{jj} \). Let a measurement process \( \{z(t), t \geq 0\} \) be defined by

\[
dz(t) = h(x(t)) dt + \beta(t) dw(t)
\]

(2)

where \( dw(t) \) is an increment in a Wiener process independent of \( x \) with a derivative of covariance \( \delta(t-t') \), \( \beta(t) \) is a continuously differentiable function bounded away from zero and \( h \) is a real-
real-valued function with values \( h_1, \ldots, h_n \) not necessarily distinct. Denote the \( z \) process over the interval \([0, t]\) by \( Z_t \) and denote the filtered probabilities \( P_z(x(t) = a_j | Z_t) \) by \( p_j(t) \). Then, following Nonham[3], it can be shown that these probabilities satisfy

\[
dp_j(t) = \left[ -\sum_{j=1}^{n} p_j(t) + \sum_{j=1}^{n} \sum_{i=1}^{n} p_j(t) \right] dt \\
+ \beta(t) \beta(t)^{-1} P_z(t) \left[ h(t)^{-1} \hat{v}(t) \right] \left[ ds(t) - \hat{v}(t) dt \right]
\]

with \( \hat{v}(t) = \beta(t)^{-1} P_z(t) \left[ h(t)^{-1} \hat{v}(t) \right] \), the conditional expectation of \( h(x(t)) \). The initial condition for (3) is just the initial probability distribution for \( x(0) \).

3. FIXED-POINT SMOOTHING

For each fixed \( t \), define the augmented process \( (X(t), \tau) \) by \( X(t) = (x(t), x(t)) \). This process has \( N \times N \) states \( a_1, a_2, \ldots, a_N \); or in component form \( (a_i, a_j) \) with \( i = (1, j) \) and \( i, j = 1, 2, \ldots, N \). From the properties of \( x \), it also follows that this new process in Markov and has stationary transition probabilities \( P_{ij}(h) \) satisfying

\[
P_{ij}(h) = \begin{cases} 
1 - V_j h + o(h) & i = j \\
V_j h + o(h) & i \neq j
\end{cases}
\]

for \( i = (1, j) \) and \( j = (1, j) \). The nonnegative constants \( V_{ij} \), \( V_{ij} \) are defined by \( V_i = V_{i1} \), \( V_{jj} = V_{1j} \), and satisfy \( V_{i1} V_{1j} = V_{i1} \). \( V_{ij} \) and satisfy \( V_{ij} = V_{1j} \).

With \( h(X(t) | \tau) = x(t) \), the measurements are given by

\[
dx(t) = x(t) dt + \beta(t) dw(t), \quad \tau \geq t
\]

Then denote the filtered probabilities of \( X \), \( P_z(x(t) | \tau) = a_j | Z_t \) by \( p_j(t) \) and the fixed-point smoothing probabilities of \( x \), \( P_z(x(t) = a_j | Z_t) \) by \( p_j(t) \) with \( i = (1, j) \). We have the filtered probabilities of \( x \) completely determining the fixed-point probabilities of \( x \). From Section 2 the former are easily found as the solution of the equations

\[
dp_j(t) = \left[ -\beta(t) p_j(t) + \sum_{j=1}^{n} \sum_{i=1}^{n} p_j(t) \right] dt \\
+ \beta(t) \beta(t)^{-1} P_z(t) \left[ h(t)^{-1} \hat{v}(t) \right] \left[ ds(t) - \hat{v}(t) dt \right]
\]

with the initial condition \( P_z(t) = \delta_{ij} p_j(t) \), \( i = (1, j) \) and with \( \hat{v}(t) = \beta(t)^{-1} P_z(t) \left[ h(t)^{-1} \hat{v}(t) \right] \), the conditional expectation of \( x(t) \).

The above approach to the continuous-time fixed-point smoothing problem has been used previously for both the linear[4] and nonlinear[3] cases.

RANDOM TELEGRAPH WAVE EXAMPLE

For a process with states \( \pm 1 \) and switching parameter \( v \) with \( u_j = v \) and \( u_j = -v \) for \( i, j = 1, 2 \), the fixed-point equations (6) are, with independent variables \( q_1, Z \) and \( D \),

\[
dq_1(t) = -2q_1(t) + \beta^2 \left[ 1 - q_1(t)^2 \right] dt, \quad q(0) = 0 \tag{7}
\]

\[
dZ(t) = \beta^2 \left[ 2D(t) - 1 - Z(t) \right] dt, \quad Z(0) = 0 \tag{8}
\]

\[
dD(t) = \beta^2 \left[ 2D(t) - 1 - Z(t) \right] dt, \quad D(0) = 1 \tag{9}
\]

where \( q \) is the conditional expectation of \( Z \), \( Z \) is the fixed-point estimate of \( x \), \( D \) is the probability that both components of the \( X \)-process are the same; and \( I \) is the innovations process with \( dI(t) = dz(t) - q(t) dt \).

4. FIXED-LAG SMOOTHING (DISCRETE-TIME)

For the fixed-lag smoothing problem we need to know the evolution of the quantities \( P_{z}(t | t+1) \) with \( L \) fixed and \( t \) varying. One approach which suggests itself is to take equation (6) for \( P_z(t | t) \), to express \( P_z(t | t+1) \) as an integral using this equation, and then to compute the differential now letting \( t \) rather than \( t \) vary. Although this method works for the linear-gaussian problem[6] it fails here because of the conditional rather than unconditional expectations in the integrand. We thus treat only the discrete-time case and derive a
suboptimal fixed-lag estimate as a linear combination of the states of a discrete-time nonlinear system driven by the measurements. With obvious notation we have, discretizing (6)

\[ P_z(k|z) = P_z(k|z_k) + \sum_{j=1}^{N} V_{ij}P_z(k|z_{j-1})T + \beta(k)^{-2}P_z(k|z_{1-j})(z_{j-1} - z(k))T(k) \]

where \( k \) is fixed, \( 1 \) takes the values \( k, k+1, k+2, \ldots \), and 1(\( k \)) = \( z(k+1) - z(k) - \alpha(k)T \). Now, fix \( \alpha \) in these equations, and in writing down (10) for each \( k \) in the range \( 1 \) to \( L+1 \), define the new variables \( P_{ij}(k) = P_z(k|z_{j-1}) \) for \( i=1, \ldots, N \) and \( j=1, \ldots, L+1 \). We then obtain

\[ P_{i,j+1}(k) = P_{ij}(k) + \sum_{j=1}^{N} V_{ij}P_{ij}(k)T + \beta(k)^{-2}P_{ij}(k)(z_{j-1} - z(k))T(k) \]

while the quantities \( P_{ij}(k) \) are none other than the filtered probabilities associated with the original \( x \)-process, and hence are updated by the discrete-time version of the filter equations (3). Finally, the suboptimal fixed-lag estimate of \( x(k|L) \) given \( z(k) \) is

[10]

ALTERNATIVE DERIVATION

We now outline an alternative derivation of (11) which is based on a discrete-time approximation [7] to the continuous-time processes involved and the idea [1] that the discrete-time fixed-lag smoothing problem can be posed as one of filtering a related process. Consider first a discrete-time Markov process \( \{x(k), k=0, 1, \ldots \} \) for which the transition probabilities \( P_z(x(k+1) = a_j | x(k) = a_i) \) are denoted by \( P_{i,j} \) with \( P_{i,j} = V_{ij}T \) for \( i,j \) and \( P_{i,j} = \gamma_{ij} \) for \( i \neq j \), \( T \) being the discretization interval; and second, the observation process defines as

\[ (k) = s(k)T + \beta(k)w(k)/T \]

where \( \{w(k), k=0, 1, \ldots \} \) is Gaussian white noise with unit variance.

Suppose we define a new Markov process \( \{s(k), k=L+1, \ldots \} \) with \( S(k) = \{s(k) \ s(k-1) \ldots s(k-L-1)\} \), then the fixed-lag probabilities of the \( s \)-process are simply sums of the filtered probabilities of the \( x \)-process. The latter can be derived using a well-known recursive formula [9,17] and approximating [7] for small \( T \).

5. SIMULATION - DISCUSSION AND RESULTS

For our example, the random telegraph wave, we generated the approximated discrete-time processes mentioned in the final part of section 4. It was found that for reliable results the following points needed to be considered.

(a) Consider the discrete-time filter equation (see (7))

\[ q(k+1) = -2\alpha q(k)T + \beta^{-2}[1-q(k)^{-2}][x(k)-q(k)]T 
+ \beta^{-1}(1-q(k)^{-2})w(k)T \]

\[ q(0) = 0 \]

(b) The definition of \( q \) implies \( |q(k)| \leq 1 \) for all \( k \). However, from (12) it is clear that for a sufficiently large value of \( w \), this bound will be violated. Then, the term \( 1-q(k)^{-2} \) is destabilizing in the sense that its effect on the increment is to cause \( |q(k+1)| > |q(k)| \). Hence, for a sequence of values of \( w \) with suitable sign, \( q \) may become arbitrarily large. An ad hoc solution, which was found to be satisfactory, was to redefine \( q(k) \) as \( \pm 1 \) at each iteration whenever \( |q(k)| > 1 \). An alternative method is to bound the noise samples and choose a sufficiently small sampling interval \( T \).

(b) It is also necessary to ensure that the discrete-time processes are, in fact, approximations to their continuous-time counterparts. Each term on the right side of (12) is small if we choose \( T \) such that \( T < \beta / 2\alpha \sqrt{T} \) and \( VT \) is small. Although ad hoc, this bounding proves to be a successful guideline in the choice of parameters in the simulation. This bound on \( T \) implies that the larger the signal-to-noise ratio the greater is the amount of computation per unit of time.
An example of the results of simulation is shown in Fig. 1, with $v = 50$, $\beta = 0.07$ and $T = 0.0005$. Then $VT = 0.025$ which is "small" and $\beta/20\sqrt{V} \approx 0.0005$. Thus the conditions on $T$ are satisfied. It is apparent from Fig. 1 that there is an initial rapid improvement in estimation with lag, followed by an evening out to a steady-state value. For the filter eq.(12), we can associate a time constant of $1/2VT$ sampling intervals. For our example, $1/2VT = 20$. Compare this with Fig. 1 where the lag for which practically all improvement is obtained is about 15. We found that every example for which $T$, $v$ and $\beta$ were satisfied, this type of behaviour was observed.

Fig. 2 shows a plot of improvement of smoothing over filtering versus a noise measure $\mu = \sqrt{q}$. The similarity with the linear-gaussian case can be seen.

![Figure 1](image1)

![Figure 2](image2)

REFERENCES