ESTIMATION OF THE GAUSS–MARKOV PROCESS FROM OBSERVATION OF ITS SIGN*

David SLEPIAN
Bell Laboratories, Murray Hill, NJ 07974, U.S.A. and University of Hawaii, Honolulu, HI 96822, U.S.A.

Received 10 February 1981

Let $X(t)$ be the ergodic Gauss–Markov process with mean zero and covariance function $e^{-|t|}$. Let $D(t)$ be +1, 0 or -1 according as $X(t)$ is positive, zero or negative. We determine the non-linear estimator of $X(t_1)$ based solely on $D(t)$, $-T \leq t \leq 0$, that has minimal mean-squared error $\varepsilon^2(t_1, T)$. We present formulae for $\varepsilon^2(t_1, T)$ and compare it numerically for a range of values of $t_1$ and $T$ with the best linear estimator of $X(t_1)$ based on the same data.

1. Introduction

Let $X(t)$ be a separable version of the continuous parameter ergodic Gaussian process specified by

$$
E X(t) = 0, \quad \rho(\tau) = E X(t)X(t+\tau) = e^{-|\tau|}.
$$

(1)

It is well known that this process is Markovian, and we shall make frequent use of this property. We are concerned here with estimation of $X(t_1)$ based on observation of the related process

$$
D(t) = \begin{cases} 
1, & X(t) > 0, \\
0, & X(t) = 0, \\
-1, & X(t) < 0
\end{cases}
$$

(2)

throughout the interval

$$
I = \{t | -T \leq t \leq 0\}, \quad T > 0.
$$

(3)

We treat both the case $t_1 \in I$ and $t_1 \notin I$. We refer to $D(t)$ as the signum of $X(t)$. For any estimate $\hat{X}(t_1)$ of $X(t_1)$, we adopt the mean-squared error

$$
\varepsilon^2 = E[X(t_1) - \hat{X}(t_1)]^2
$$

(4)

* This work was supported in part by the National Science Foundation under Grant ENG 74-19788.
as a measure of the merit of the estimate. We seek estimates of small mean-squared error.

It is well known that no estimate of $X(t_i)$ based solely on observation of $D(t)$ for $t \in I$ can have smaller mean-squared error than

$$\hat{X}(t_i)_{\text{opt}} = \mathbb{E}[X(t_i) | D(t), t \in I].$$

(5)

In this paper we exhibit explicit expressions for $\hat{X}(t_i)_{\text{opt}}$ and also

$$\varepsilon^2_{\text{opt}} = \mathbb{E}[X(t_i) - \hat{X}(t_i)_{\text{opt}}]^2.$$

(6)

$\hat{X}(t_i)_{\text{opt}}$ is a nonlinear functional of the observed signum of $X(t)$. For comparison we also derive the best linear estimator of $X(t_i)$, i.e., one of form

$$\hat{X}(t_i)_{\text{lin opt}} = \int_{-T}^{0} f(t) D(t) \, dt$$

(7)

having smallest mean-squared error, and we present some numerical data that contrast the performance of these two estimators.

2. Some probabilities associated with $X(t)$

In this section we derive in a simple manner the probabilities of a number of events associated with the excursions of $X(t)$ away from the value zero. These quantities will be of use to us in determining $\hat{X}(t_i)_{\text{opt}}$. Some of the quantities have been derived elsewhere (see [1] and [2]), but by more complicated methods.

Let $p_{X(t_1), X(t_2), \ldots, X(t_n)}(x_1, x_2, \ldots, x_n)$ denote the joint density of $X(t_1), X(t_2), \ldots, X(t_n)$. We shall establish that

$$Q(T, x_1, x_2) \, dx_1 \, dx_2 = \mathbb{P}[0 < x_1 \leq X(0) \leq x_1 + dx_1, \ 0 < x_2 \leq X(T) \leq x_2 + dx_2,$$

$$X(t) > 0 \text{ for } 0 < t < T]$$

$$= [p_{X(0), X(T)}(x_1, x_2) - p_{X(0), X(T)}(x_1, -x_2)] \, dx_1 \, dx_2.$$ 

(8)

From this we shall derive the following:

$$R(T, x) \, dx \, dT = \mathbb{P}[0 < x < X(0) < x + dx, X(t) > 0 \text{ for } 0 < t < T,$$

$$X(t) = 0 \text{ for some } t \text{ such that } T \leq t \leq T + dT]$$

$$= \frac{e^{-T} x e^{-x^2/2h(T)}}{\pi h(T)^{3/2}} \, dx \, dT, \quad x \geq 0$$

(9)

where

$$h(T) = 1 - e^{-|2T|};$$

(10)
\[ S(T, x) \, dx = P[0 < x \leq X(T) \leq x + dx \mid X(t) > 0 \text{ for } 0 < t < T \text{ and } X(0) = 0] = \frac{x}{h(T)} e^{-x^2/2h(T)} \, dx, \quad x > 0. \] (11)

To establish (8) we consider the collection \( C \) of all sample functions \( X(t) \) that satisfy \( 0 < x_1 \leq X(0) \leq x_1 + dx_1 \) and \( 0 < x_2 \leq X(T) \leq x_2 + dx_2 \). We have

\[ P[C] = P[X(0), X(T) \in (x_1, x_2)] \, dx_1 \, dx_2. \] (12)

Now divide \( C \) into two disjoint sets: \( C_1 \), the sample functions that are positive for all \( t \) such that \( 0 \leq t \leq T \), and \( C_2 \), the sample functions that are zero for some \( t \) such that \( 0 < t < T \). We have

\[ P[C] = P[C_1] + P[C_2] = \int Q(T, x_1, x_2) \, dx_1 \, dx_2 + P[C_2]. \] (13)

We shall soon show that

\[ P[C_2] = P[C_1] = P[X(0), X(T) \in (x_1, x_2)] \, dx_1 \, dx_2 \] (14)

where \( C_3 \) is the collection of sample functions \( X(t) \) satisfying \( 0 < x_1 \leq X(0) \leq x_1 + dx_1, -x_2 - dx_2 \leq X(T) \leq -x_2 < 0 \). Eq. (8) will then follow from (12), (13) and (14).

To see that (14) is true, recall that (almost surely) each member \( X(t) \), say, of \( C_3 \) must have a first zero in \((0, T)\), i.e., a smallest number \( t_0, 0 < t_0 < T \), such that \( X(t_0) = 0 \). From this \( X(t) \in C_3 \) form a new function \( \tilde{X}(t) \)

\[ \tilde{X}(t) = \begin{cases} X(t), & 0 \leq t \leq t_0, \\ -X(t), & t_0 < t \leq T. \end{cases} \]

Now \( \tilde{X}(t) \in C_2 \). Furthermore for every \( X(t) \in C_2 \) we can in an analogous manner form a sample function contained in \( C_3 \) by reflecting \( X(t) \) in the \( t \)-axis after its first zero in \((0, T)\). Because the process \( X(t) \) is symmetric about its mean value zero, and because it is Markovian, it follows that \( P[C_2] = P[C_3] \).

To establish (9), we note that

\[ dx_1 \int_{0}^{\infty} \, dx_2 \, Q(T, x_1, x_2) = P[0 < x_1 \leq X(0) \leq x_1 + dx_1, X(t) > 0 \text{ for } 0 \leq t \leq T] \]

\[ = dx_1 \int_{0}^{\infty} R(T', x_1) \, dT', \quad x_1 > 0. \]

Thus

\[ R(T, x) = -\frac{d}{dT} \int_{0}^{\infty} \, dx_2 \, Q(T, x, x_2) \]

\[ = -\frac{d}{dT} \int_{0}^{\infty} \, dx_2 \left[ p_{X(0), X(T)}(x, x_2) - p_{X(0), X(T)}(x, -x_2) \right] \]

\[ = -\frac{d}{dT} \left[ \int_{0}^{\infty} \, dx_2 \, p_{X(0), X(T)}(x, x_2) - \int_{-\infty}^{0} \, dx_2 \, p_{X(0), X(T)}(x, x_2) \right] \] (15)
where we have used (8). Now \( p_{X(0),X(T)}(x_1, x_2) \) is a Gaussian density formed from the covariance matrix

\[
\lambda = \begin{pmatrix} 1 & \lambda_{12} \\ \lambda_{12} & 1 \end{pmatrix}
\]

where \( \lambda_{12} = e^{-|T|} \). It is well known (see for example [6, p. 481]) that \( \partial p / \partial \lambda_{12} = \partial^2 p / \partial x_1 \partial x_2 \). On writing \( d/dT = (d \lambda_{12} / dT)(d/d \lambda_{12}) \), (15) becomes

\[
R(T, x) = e^{-|T|} \left[ \int_0^\infty dx_2 \frac{\partial^2}{\partial x \partial x_2} p_{X(0),X(T)}(x, x_2) \right.
\]

\[
- \int_\infty^0 dx_2 \frac{\partial^2}{\partial x \partial x_2} p_{X(0),X(T)}(x, x_2) \left. \right] 
\]

\[
e^{-|T|} \frac{\partial}{\partial x} \left[ -2 p_{X(0),X(T)}(x, 0) \right] = -e^{-|T|} 2 \frac{\partial}{\partial x} \frac{e^{-x^2/2(1-\lambda_{12}^2)}}{2\pi\sqrt{1-\lambda_{12}^2}}
\]

whence (9) follows.

Since the process \( X(t) \) is invariant under a time reversal, (9) also has the interpretation

\[
R(T, x) \, dx \, dT = P[0 < x < X(T) < x + dx, X(t) > 0 \text{ for } 0 < t < T, \]

\[
X(t) = 0 \text{ for some } t \text{ such that } -dT \leq t \leq 0].
\]

Thus

\[
dT \int_0^\infty dx \, R(T, x) = P[X(t) > 0 \text{ for } 0 < t < T, \]

\[
X(t) = 0 \text{ for some } t \text{ such that } -dT \leq t \leq 0]
\]

\[
e^{-|T|} \frac{dT}{\pi h(T)^{3/2}} \int_0^\infty dx \, x \, e^{-x^2/2h(T)} \]

\[
e^{-|T|} \frac{dT}{\pi h(T)^{1/2}}.
\]

Dividing (9) by (17) we find the conditional density function

\[
S(T, x) = \frac{R(T, x)}{\int_0^\infty R(T, x') \, dx'} = \frac{x}{h(T)} e^{-x^2/2h(T)}, \quad x > 0,
\]

as reported in (11). The conditional probability density displayed there is thus seen to be defined by

\[
P[0 < x \leq X(T) < x + dx \mid X(t) > 0 \text{ for } 0 < t < T \text{ and } X(0) = 0] =
\]

\[
= \lim_{\Delta \to 0} P[0 < x \leq X(T) < x + dx, X(t) > 0 \text{ for } 0 < t \leq T, \]

\[
X(t) = 0 \text{ for some } -\Delta \leq t \leq 0]
\]

\[
\times (P[X(t) > 0 \text{ for } 0 < t \leq T, X(t) = 0 \text{ for some } -\Delta \leq t \leq 0])^{-1}. \quad (18)
\]
A few more probabilities will also be needed in later sections. We note first that
\[
P_{+}(T) = P[X(t) > 0 \text{ for } 0 \leq t \leq T] = \int_{0}^{\infty} dx_{1} \int_{0}^{\infty} dx_{2} P_{X(0),X(T)}(x_{1}, x_{2})
\]
\[
= \int_{0}^{\infty} dx_{1} \int_{0}^{\infty} dx_{2} P_{X(0),X(T)}(x_{1}, x_{2})
\]
\[
- \int_{0}^{\infty} dx_{1} \int_{0}^{\infty} dx_{2} P_{X(0),X(T)}(x_{1}, -x_{2})
\]
\[
= \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin e^{-T} \right] - \left[ \frac{1}{4} - \frac{1}{2\pi} \arcsin e^{-T} \right]
\]
\[
= \left( \frac{1}{\pi} \right) \arcsin e^{-T}.
\]

The two integrals evaluated here can be found in Appendix A as (A6) and the variant discussed above (A11).

Because of the stationarity of \( X(t) \), (19) can also be written as
\[
P[X(t) > 0, t_{1} \leq t \leq t_{2}] = \left( \frac{1}{\pi} \right) \arcsin e^{-(t_{2} - t_{1})}.
\]

Now let
\[
\text{Int}(t_{1}, t_{2}) dt_{1} dt_{2} = P[X(t) > 0 \text{ for } t_{1} \leq t \leq t_{2},
\]
\[
X(t) = 0 \text{ for some } t_{1} - dt_{1} \leq t < t_{1},
\]
\[
X(t) = 0 \text{ for some } t_{2} < t \leq t_{2} + dt_{2}].
\]

From (20) we have
\[
\frac{1}{\pi} \arcsin e^{-(t_{2} - t_{1})} = \int_{t_{1}}^{t_{2}} dr_{1} \int_{r_{1}}^{\infty} dr_{2} \text{Int}(r_{1}, r_{2}),
\]
so that by differentiation we find
\[
\text{Int}(t_{1}, t_{2}) = \frac{1}{\pi} \frac{e^{-(t_{2} - t_{1})}}{h(t_{2} - t_{1})^{3/2}}.
\]

3. \( \hat{X}(t_{1})_{\text{opt}} \) for \( t_{1} \notin I \)

We return now to compute \( \hat{X}(t_{1})_{\text{opt}} \) as given by (5). We treat first the case \( t_{1} \geq 0 \). The case \( t_{1} \leq -T \) will follow at once by the invariance of \( X(t) \) under time reversal.

As concerns the observed data \( D(t), t \in I \), two cases are evident: Case (i) \( D(t) \) is constant for \( -T \leq t \leq 0 \); Case (ii) \( D(t) \) is not constant for \( -T \leq t \leq 0 \). We first consider Case (i).
From (16) and the stationarity of $X(t)$ it follows that

$$P[0 < x_0 \leq X(0) \leq x_0 + dx_0, X(t) > 0 \text{ for } t \in I, \quad X(t) = 0 \text{ for some } -dT - T \leq t < -T] =$$

$$= R(T, x_0) \, dx_0 \, dT.$$

From the Markovian nature of $X(t)$, we have that

$$dx_1 \, dT \int_0^\infty dx_0 \, R(T, x_0) \, p_{X(t_1) \mid X(0)}(x_1 \mid x_0) =$$

$$= P[x_1 \leq X(t_1) \leq x_1 + dx_1, D(t) = 1 \text{ for } t \in I, \quad X(t) = 0 \text{ for some } -dT - T \leq t < -T]$$

where

$$p_{X(t_1) \mid X(t_0)}(x_2 \mid x_1) = \frac{\exp\left\{ -\frac{(x_2 - x_1 e^{-|t_2-t_1|})^2}{2h(t_2-t_1)} \right\}}{\sqrt{2\pi h(t_2-t_1)}}$$

is the conditional density of $X(t_2)$ given that $X(t_1) = x_1$. From (23) we have that

$$P[x_1 \leq X(t_1) \leq x_1 + dx_1, D(t) = 1 \text{ for } t \in I] =$$

$$= dx_1 \int_T^\infty dT' \int_0^\infty dx_0 \, R(T', x_0) \, p_{X(t_1) \mid X(0)}(x_1 \mid x_0).$$

Finally, on using (19) we see that the conditional density of $X(t_1)$ given that $D(t) = 1$ for $t \in I$ is

$$p_{X(t_1) \mid D(t)=1}(x_1) = \frac{\pi}{\arcsin e^{-T}} \int_T^\infty dT' \int_0^\infty dx_0 \, R(T', x_0) \, p_{X(t_1) \mid X(0)}(x_1 \mid x_0)$$

so that

$$E[X(t_1) \mid D(t) = 1 \text{ for } t \in I] =$$

$$= \frac{\pi}{\arcsin e^{-T}} \int_T^\infty dT' \int_0^\infty dx_0 \, dx_1 \, R(T', x_0) \, p_{X(t_1) \mid X(0)}(x_1 \mid x_0)$$

$$= \frac{\pi}{\arcsin e^{-T}} \int_T^\infty dT' \int_0^\infty dx_0 \, e^{-T'} \frac{x_0 e^{-x_0^2/2h(T')}}{\pi h(T')^{3/2}} x_0 \, e^{-t},$$

$$= \frac{\pi}{\arcsin e^{-T}} \int_T^\infty dT' \frac{e^{-T'}}{\pi h(T')^{3/2} \sqrt{2\pi}} = \frac{\sqrt{\frac{\pi}{2}} e^{-(T+T')}}{\arcsin e^{-T}}. \quad (25)$$

By symmetry

$$E[X(t_1) \mid D(t) = -1 \text{ for } t \in I] = \frac{\sqrt{\frac{\pi}{2}} e^{-(T+T')}}{\arcsin e^{-T}}. \quad (26)$$
Suppose now that \( D(0) = 1 \) and that \( D(t) \) is not constant for \( t \in I \). Then there is a largest number \( t_0, -T < t_0 < 0, \) such that \( D(t_0) = 0 \). By observing \( D(t) \) for \( t \in I \) one then knows that \( X(t_0) = 0 \). By the Markov property events before \( t_0 \) need not be considered in calculating \( \mathbb{E}[X(t_1) | D(t) \in I] \) for this case. For the conditional density, \( p_{X(t_1) | D(t)}(x_1) \), of \( X(t_1) \) given \( D(t_0) = 0, D(t) = 1 \) for \( t_0 < t \leq 0 \) we have from (11)

\[
p_{X(t_1) | D(t)}(x_1) = \int_0^\infty dx S(t_0, x)p_{X(t_1), X(t_0)}(x_1 | x),
\]

where we have used the stationarity of \( X(t) \) and the Markov property. Thus

\[
\mathbb{E}[X(t_1) | D(t_0) = 0, D(t) = 1, t_0 < t \leq 0] =
\]

\[
= \int_{-\infty}^\infty dx_1 x_1 \int_0^\infty dx S(t_0, x)p_{X(t_1), X(t_0)}(x_1 | x)
\]

\[
= \int_0^\infty dx e^{-t_1}xS(t_0, x) = \frac{e^{-t_1}}{h(t_0)} \int_0^\infty x^2 e^{-x^2/2h(t_0)} dx
\]

\[
= \sqrt{\frac{3}{2}} e^{-t_1} \sqrt{h(t_0)}.
\]

Similarly

\[
\mathbb{E}[X(t_1) | D(t_0) = 0, D(t) = -1, t_0 < t \leq 0] = -\sqrt{\frac{1}{2}} e^{-t_1} \sqrt{h(t_0)}.
\]

We summarize: If \( t_1 \neq 0 \),

\[
\hat{X}(t_1)_{\text{opt}} = \begin{cases} 
\alpha, & D(t) = 1 \text{ for all } t \in I, \\
-\alpha, & D(t) = -1 \text{ for all } t \in I, \\
\beta(t_0), & D(t_0) = 0, D(t) = 1 \text{ for } -T < t_0 < t \leq 0, \\
-\beta(t_0), & D(t_0) = 0, D(t) = -1 \text{ for } -T < t_0 < t \leq 0,
\end{cases}
\]

\[
\alpha = \sqrt{\frac{1}{2}} \pi e^{-2T} \arcsin e^{-T}, \quad \beta(t_0) = \sqrt{\frac{1}{2}} \pi e^{-t_1} \sqrt{1 - e^{-2|t_0|}}^{1/2}.
\] (27)

As for the optimum mean-squared error, we have

\[
e_{\text{opt}}^2 = \mathbb{E}[X(t_1) - \mathbb{E}[X(t_1) | D(t), t \in I]^2 = \mathbb{E}X(t_1)^2 - \mathbb{E}\mathbb{E}[X(t_1) | D(t), t \in I]^2
\]

\[
= 1 - \mathbb{E} \hat{X}(t_1)_{\text{opt}}^2.
\]

\[
= 1 - 2\alpha^2 \mathbb{P}[D(t) = 1 \text{ for all } t \in I]
\]

\[
-2 \int_{-T}^0 dt_0 \beta^2(t_0) \mathbb{P}[D(t) = 1 \text{ for } t_0 < t \leq 0, D(t) = 0
\]

for some \( t \) such that \( -dt_0 - t_0 < t \leq t_0 \)

\[
= 1 - 2\alpha^2 \frac{1}{\pi} \arcsin e^{-T} - 2 \int_{-T}^0 dt_0 \beta^2(t_0) \frac{e^{-|t_0|}}{\pi \sqrt{h(t_0)}}
\]
by (19) and (17). Using (27), we find
\[ \varepsilon_{opt}^2 = 1 - \frac{1}{2} e^{-2t_1} \left[ \frac{1}{2} \pi - \arcsin e^{-T} - e^{-T} \sqrt{1 - e^{-2T}} + \frac{2 e^{-2T}}{\arcsin e^{-T}} \right]. \] (28)

We note the limiting values
\[ \varepsilon_{opt}^2 = \begin{cases} 1 - (2/\pi) e^{-2t_1}, & T = 0, \\ \left(1 - \frac{1}{4} \pi e^{-2t_1} + O(e^{-T}) \right), & T \to \infty. \end{cases} \] (29)

For \( t_1 = 0 \) these values are respectively 0.36338 and 0.21460. Thus the mean-squared error of \( \hat{X}(0)_{opt} \) based on observing the sign of \( X(t) \) for all \( t < 0 \) is 59% of the mean-squared error obtained on observing just the sign of \( X(0) \). The value 0.2146 is approached exponentially fast, so a small amount of past sign data gives considerable aid in estimating \( X(0) \) accurately.

4. \( \hat{X}(t_1)_{opt} \) for \( t_1 \in I \)

Now let \( t_1 \in I \) and define
\[ I_L = \{ t \mid -T < t \leq t_1 \}, \quad I_R = \{ t \mid t_1 \leq t < 0 \}. \] (30)

If \( D(t_1) = 0 \), then \( X(t_1) = 0 \) and, of course, \( \hat{X}(t_1) = 0 \). Suppose now that \( D(t_1) = 1 \). Four cases present themselves: \textit{Case (i)} \( D(t) \) constant for \( t \in I \); \textit{Case (ii)} \( D(t) \) constant for \( t \in I_L \), but not constant for \( t \in I_R \); \textit{Case (iii)} \( D(t) \) constant for \( t \in I_R \), but not constant for \( t \in I_I \); \textit{Case (iv)} \( D(t) \) not constant for \( t \in I_L \) and \( D(t) \) not constant for \( t \in I_R \). We shall treat them separately, but first we evaluate an integral that will be useful to us.

Suppose \( t_0 < t_1 < t_2 \) and define
\[ u = t_1 - t_0, \quad v = t_2 - t_1. \] (31)

Then from (8) and the Markov nature of \( X(t) \),
\[ W(u, v, x_1) \, dx_1 = P[0 < x_1 \leq X(t_1) \leq x_1 + dx_1, X(t) > 0 \text{ for } t_0 \leq t \leq t_2] \]
\[ = dx_1 \int_0^\infty dx_0 \int_0^\infty dx_2 \, Q(u, x_0, x_1) \frac{Q(v, x_1, x_2)}{p_{X(t_1)}(x_1)}. \] (32)

Thus
\[ Y(u, v) = \int_0^\infty dx_1 \, x_1 W(u, v, x_1) \]
\[ = \int_0^\infty dx_0 \int_0^\infty dx_1 \int_0^\infty dx_2 \, x_1 [p_{01}(x_0, x_1) - p_{01}(x_0, -x_1)] \]
\[ \times [p_{21}(x_2 | x_1) - p_{21}(-x_2 | x_1)] \]
\[ = \int_0^\infty dx_0 \int_0^\infty dx_1 \int_0^\infty dx_2 \, x_1 [p_{012}(x_0, x_1, x_2) - p_{012}(x_0, -x_1, -x_2) \]
\[ - p_{012}(x_0, x_1, -x_2) + p_{012}(x_0, -x_1, x_2)] \] (33)
on substituting expression (8) for \( Q \) and where we have written \( p_{01} \) for \( p_{X(t_0), X(t_1)} \), etc. The integrals in (33) can now be evaluated using (A4) of Appendix A and the rule discussed in connection with (A11). One finds

\[
Y(u, v) = B_2(\lambda_{12}, \lambda_{13}, \lambda_{23}) - B_2(-\lambda_{12}, -\lambda_{13}, \lambda_{23})
\]

\[
- B_2(\lambda_{12}, -\lambda_{13}, -\lambda_{23}) + B_2(-\lambda_{12}, \lambda_{13}, -\lambda_{13}),
\]

\[
\lambda_{12} = e^{-u}, \quad \lambda_{23} = e^{-v}, \quad \lambda_{13} = e^{-(u+v)}
\]

in the notation of Appendix A. Eq. (A4) and some algebra yield the result

\[
Y(u, v) = \frac{1}{\pi} \sqrt{\frac{2}{\pi}} \left[ e^{-u} \arcsin \sqrt{1 - \frac{h(v)}{h(u+v)}} + e^{-v} \arcsin \sqrt{1 - \frac{h(u)}{h(u+v)}} \right].
\] (34)

We now return to the four alternatives for \( D(t) \), \(-T \leq t \leq 0\). Eq. (19) gives the probability that \( D(t) = 1 \) for all \( t \in I \) so that for Case (i) we have

\[
E[X(t_1) | D(t) = 1 \text{ for } -T \leq t < t_2, D(t_2) = 0] = \frac{\pi Y(t_0, t_2)}{\arcsin e^{-t}}
\] (35)

where

\[
t_r = -t_1, \quad t_e = T + t_1
\]

are nonnegative quantities giving the distance from \( t_1 \) to the right and left ends of the observation interval \( I \).

To treat Case (ii), let \( t_2 \) be the smallest number greater than \( t_1 \) such that \( D(t_2) = 0 \). By hypothesis, \( t_2 < 0 \). Then, using (9) for \( R \),

\[
P[0 < x_1 < X(t_1) \leq x_1 + dx_1, D(t) = 1 \text{ for } -T \leq t < t_2, D(t_2) = 0 \text{ for some } t_2 \leq t \leq t_2 + dT] =
\]

\[
= dx_1 \int_0^\infty dx_0 Q(t_2-t_1, x_1) \frac{R(t_2-t_1, x_1)}{p_{X(t_1)(x_1)}}
\]

\[
= -dx_1 \int_0^\infty dx_0 \int_0^\infty dx_2 Q(t_2-t_1, x_1) \frac{Q(t_2-t_1, x_1, x_2)}{p_{X(t_1)(x_1)}}
\]

where we have used the first equality of (15) to express \( R \) in terms of \( Q \). The probability that \( D(t) \) has its first zero in \( I \) between \( t_2 \) and \( t_2 + dT \) is, by (17),

\[
e^{-(T+t_2)} dT/\left(\pi h(T+t_2)^{1/2}\right).
\]

On dividing (36) by that quantity, we obtain the conditional density for \( X(t_1) \) knowing \( D(t) = 1 \) for \(-T \leq t < t_2\) and \( D(t_2) = 0 \). Thus

\[
E[X(t_1) | D(t) = 1 \text{ for } -T \leq t < t_2, D(t_2) = 0] =
\]

\[
= \frac{-\pi \sqrt{h(T+t_2)}}{e^{-(T+t_2)}} \int_0^\infty dx_1 x_1 W(t_2-t_1, x_1)
\]

\[
= \frac{-\pi \sqrt{h(t_e+t_0)}}{e^{-t_e+t_0}} \int_0^\infty \frac{dx_1 \ x_1 W(t_2-t_1, x_1)}{Y(t_0, t_0)}
\] (37)
by definition (33) of $Y$. Here $t_0 = t_2 - t_1$ is the positive distance from $t_1$ to the first zero of $D(t)$ to the right. Since, from (34),

$$\frac{-d}{dv} Y(u, v) = \frac{1}{\pi} \sqrt{\frac{2}{\pi}} \left[ e^{-v} \arcsin \sqrt{1 - \frac{h(u)}{h(u + v)}} + \frac{e^{-(u+v)} \sqrt{h(u)h(v)}}{h(u + v)} \right],$$

(38)

$
\hat{X}(t_1)_{opt}$ is now specified for Case (ii). By time reversal, the formula for $
\hat{X}(t_1)_{opt}$ for Case (iii) is obtained from these results by replacing $t_0$ by $t_0$ and by replacing $t_0$ by $t_0$. Here $t_0 = t_2 - t_1$ where $t_2$ is the largest number less than $t_1$ for which $D(t_2) = 0$. Case (iv) follows in a similar way. We have

$$P[0 < x_1 \leq X(t_1) \leq x_1 + dx_1, D(t) = 1 \text{ for } t_1 - t_0 < t < t_1 + t_0, \nonumber$$

$$D(t) = 0 \text{ for some } t_1 - t_0 < t < t_1 - t_0$$

and also for some $t_1 + t_0 < t < t_1 + t_0 + dT] = \nonumber$$

$$= dx_1 \ dT \ dT \ R(t_0, x_1) \frac{R(t_0, x_1)}{p_{X(t_1)}(x_1)} \nonumber$$

$$= dx_1 \ dT \ dT \ \frac{\partial^2}{\partial t_0 \partial t_0} W(t_0, t_0, x_1).$$

From (21) we see that the conditional density for $X(t_1)$ given $D(t_1 - t_0) = D(t_1 + t_0) = 0$ and $D(t) = 1$ for $t_1 - t_0 < t < t_1 + t_0$ is given by

$$\frac{\pi h(t_0 + t_0)^{3/2}}{e^{-(t_0 + t_0)} dx_1 \ \frac{\partial^2}{\partial t_0 \partial t_0} W(t_0, t_0, x_1).$$

Thus

$$E[X(t_1) \mid D(t) = 1 \text{ for } t_1 - t_0 < t < t_1 + t_0, D(t_1 - t_0) = D(t_1 + t_0) = 0] = \nonumber$$

$$= \frac{\pi h(t_0 + t_0)^{3/2}}{e^{-(t_0 + t_0)} \ \frac{\partial^2}{\partial t_0 \partial t_0} Y(t_0, t_0)$$

(39)

where we have used definition (33). From (38) one finds

$$\frac{\partial^2}{\partial u \partial v} Y(u, v) = (2/\pi)^{3/2} e^{-(u+v)} \sqrt{h(u)h(v)}$$

(40)

so that $\hat{X}(t_1)_{opt}$ is now specified for this case as well.

We summarize these results for the case $t_1 \in I$: Let

$t_0 = \text{distance from } t_1 \text{ to the right endpoint of } I,$

$t_0 = \text{distance from } t_1 \text{ to the left endpoint of } I,$

$t_0 = \text{distance from } t_1 \text{ to the first zero of } D(t) \text{ to the right of } t_1,$

$t_0 = \text{distance from } t_1 \text{ to the first zero of } D(t) \text{ to the left of } t_1.$
Then

\[
\hat{X}(t_{1})_{_{\text{opt}}} = \begin{cases} 
0, & D(t_{1}) = 0, \\
\pm \alpha(t_{\ell}, t_{r}), & D(t_{1}) = \pm 1, t_{\ell_{0}} > t_{\ell}, t_{r_{0}} > t_{r} \\
\pm \beta(t_{\ell}, t_{r_{0}}), & D(t_{1}) = \pm 1, t_{\ell_{0}} > t_{\ell}, t_{r_{0}} < t_{r} \\
\pm \beta(t_{r_{0}}, t_{\ell}), & D(t_{1}) = \pm 1, t_{\ell_{0}} < t_{\ell}, t_{r_{0}} > t_{r} \\
\pm \gamma(t_{\ell_{0}}, t_{r_{0}}), & D(t_{1}) = \pm 1, t_{\ell_{0}} < t_{\ell}, t_{r_{0}} < t_{r} 
\end{cases}
\]  

(41)

Here

\[
\alpha(u, v) = \sqrt{\frac{2}{\pi}} \frac{1}{\arcsin e^{-u + v}} \left[ e^{-u} \arcsin s(u, v) + e^{-v} \arcsin s(v, u) \right],
\]

\[
s(u, v) = \sqrt{1 - h(v)/h(u + v)},
\]

\[
\beta(u, v) = \sqrt{\frac{2}{\pi}} \frac{\sqrt{h(u + v)}}{e^{-u + v}} \left[ e^{-v} \arcsin s(v, u) + \frac{e^{-(u + v)} \sqrt{h(u)h(v)}}{h(u + v)} \right],
\]

\[
\gamma(u, v) = 2\sqrt{2/\pi} \frac{\sqrt{h(u)h(v)}}{h(u + v)}.
\]

(42)

For the optimum mean-squared error we find

\[
\varepsilon_{\text{opt}}^{2} = 1 - \mathbb{E} \hat{X}(t_{1})^{2}_{\text{opt}}
\]

\[
= 1 - 2\alpha(t_{\ell}, t_{r})\mathbb{P}[D(t) = 1 \text{ for all } t \in I]
\]

\[-2 \int_{0}^{t_{\ell}} dt_{r_{0}} \beta^{2}(t_{\ell}, t_{r_{0}})\mathbb{P}[D(t) = 1 \text{ for } t_{1} - t_{\ell} \leq t < t_{1} + t_{r_{0}} \text{ and } D(t) = 0 \text{ for some } t]
\]

\such{t_{1} + t_{r_{0}} \leq t \leq t_{1} + t_{r_{0}} + dt_{r_{0}}}

\[-2 \int_{0}^{t_{r}} dt_{\ell_{0}} \beta^{2}(t_{r}, t_{\ell_{0}})\mathbb{P}[D(t) = 1 \text{ for } t_{1} - t_{r} < t \leq t_{1} + t_{r} \text{ and } D(t) = 0 \text{ for some } t \text{ such that}]
\]

\[t_{1} - t_{\ell_{0}} - dt_{\ell_{0}} \leq t \leq t_{1} - t_{\ell_{0}}]

\[-2 \int_{0}^{t_{r}} dt_{r_{0}} \int_{0}^{t_{\ell_{0}}} dt_{\ell_{0}} \gamma^{2}(t_{\ell_{0}}, t_{r_{0}})\mathbb{P}[D(t) = 1 \text{ for } t_{1} - t_{\ell_{0}} < t < t_{1} + t_{r_{0}} \text{ and } D(t) = 0 \text{ for some } t \text{ such that}]
\]

\[t_{1} - t_{\ell_{0}} - dt_{\ell_{0}} \leq t \leq t_{1} - t_{\ell_{0}}]

\[t_{1} + t_{r_{0}} \leq t \leq t_{1} + t_{r_{0}} + dt_{r_{0}}
\]
60 D. Slepian / Estimation from sign

\[ 1 - 2\alpha \left( t_\epsilon, t_\epsilon \right)^2 \frac{1}{\pi} \arcsin e^{-\left( t_\epsilon + t_0 \right)} - 2 \int_0^{t_\epsilon} \beta^2(t_\epsilon, t_0) \frac{e^{-\left( t_\epsilon + t_0 \right)}}{\pi \sqrt{h(t_\epsilon + t_0)}} \, dt_0 \]

\[ -2 \int_0^{t_\epsilon} \beta^2(t_\epsilon, t_0) \frac{e^{-\left( t_\epsilon + t_0 \right)}}{\pi \sqrt{h(t_\epsilon + t_0)}} \, dt_0 \]

\[ -2 \int_0^{t_\epsilon} dt_0 \int_0^{t_\epsilon} dt_\epsilon \gamma^2(t_\epsilon, t_0) \frac{e^{-\left( t_\epsilon + t_0 \right)}}{\pi h(t_\epsilon + t_0)^{3/2}} \] (43)

by (20), (17) and (22).

5. Numerical results

The last term in (43) can be expressed in terms of elementary functions. We find for its value

\[ \frac{16}{3\pi^2} \left[ \frac{1}{2} \pi - a \sqrt{1 - a^2} - b \sqrt{1 - b^2} + \frac{ab(2 - a^2 - b^2)}{\sqrt{1 - a^2 b^2}} \right. \]

\[ - \arcsin a - \arcsin b + \arcsin ab \]

where \( a = e^{-t_\epsilon} \) and \( b = e^{-t_0} \). While we have not been able to express the single integrals in (43) in elementary terms, they are easily evaluated numerically. Fig. 1 shows the results of evaluating (28) and (43) for various values of \( T \) and \( t_1 \). Here we have placed the origin at the center of the observation interval so that the curves shown can be extended to negative values of \( t_1 \) by reflection in the \( \epsilon^2 \)-axis. It is seen that, as \( t_1 \) approaches the end of the observation interval from without, \( \epsilon^2 \) decreases rapidly. The mean-squared estimation error continues to decrease as the point at which \( X(t) \) is to be estimated approaches the center of the observation interval, but it is soon very nearly equal to its minimum value which is attained at the center of the observation interval. This minimum itself is nearly independent of the duration of the observation interval for \( T > 1.5 \). It has the limiting value \( 1 - 8/3\pi = 0.1512 \) which is the mean-squared estimation error at the ‘center’ of an observation interval of infinite duration.

The solid curves on Fig. 2 show how the optimal mean-squared estimation error at the edge and at the center of the observation interval depends on the duration \( T \) of that interval. For \( T > 1 \), optimal estimation of \( X(t) \) at the edge of the observation interval is about 40% larger than the mean-squared error of the estimate of \( X(t) \) at the center of the interval. For \( T < 0.1 \), however, \( X(t) \) can be estimated at the center and the edge of the observation interval with about the same accuracy.

The dashed curves on Fig. 2 show analogous results for the optimal linear estimator. It is easy to show that the weighting function \( f(t) \) of (7) that provides a
Fig. 1. The mean-squared error $\epsilon^2$ for the optimal estimator of $X(t)$ at a point distant $t_1$ from the center of the observation interval of $D(t)$ of duration $T$.

Fig. 2. Mean-squared error at the observation interval edge and center for the optimal estimator and for the optimal linear estimator.
linear estimator with least mean-squared error satisfies the integral equation

\[ \int_{-T}^{0} E[D(t)D(t')] f(t') \, dt' = EX(t)D(t), \quad -T \leq t \leq 0. \]

For the case at hand this becomes

\[ \frac{2}{\pi} \int_{-T}^{0} \text{arcsin}(e^{-|t-t'|}) f(t') \, dt' = e^{-|t-t'|}, \quad -T < t < 0. \quad (44) \]

The mean-squared error for this best linear estimator is then given by

\[ \varepsilon_{\text{lin}}^2 = 1 - \frac{2}{\pi} \int_{-T}^{0} f(t) e^{-|t-t'|} \, dt. \quad (45) \]

We have discussed in [3] how a class of integral equations containing (44) as a special case can be solved numerically. We have applied that technique to (44) and have thereafter evaluated (45) to obtain the dashed curves of Fig. 2. Morrison [4] has treated (44) and (45) in detail when \( T = \infty \) by Wiener–Hopf techniques. Our results for large \( T \) agree with his. For large \( T \) it is seen that the mean-squared error for the best linear estimator is about 52\% larger than the error for the best non-linear estimator.

**Appendix A. Some Gaussian integrals**

Let

\[ p_n(x_1, \ldots, x_n) = \frac{\exp\left(-\frac{1}{2} \sum_{i=1}^{n} \frac{1}{\lambda_{ii}} x_i x_i\right)}{(2\pi)^{n/2} |\lambda_n|^{1/2}} \]

be the \( n \)-dimensional density for Gaussian variates having mean zero, unit variance and non-singular symmetric covariance matrix

\[ (\lambda_n) = \begin{pmatrix} 1 & \lambda_{12} & \lambda_{13} & \cdots & \lambda_{1n} \\
\lambda_{21} & 1 & \lambda_{23} & \cdots & \lambda_{2n} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\lambda_{n1} & \lambda_{n2} & \lambda_{n3} & \cdots & 1 \end{pmatrix}. \quad (A2) \]

The following integrals are needed in this paper:

\[ A(\lambda_{12}, \lambda_{13}, \lambda_{23}) = \int_{-\infty}^{\infty} dx_1 \int_{0}^{\infty} dx_2 \int_{0}^{\infty} dx_3 p_3(x_1, x_2, x_3) \]

\[ = \frac{1}{4\pi} \left[ \frac{1}{2} \pi + \text{arcsin} \lambda_{12} + \text{arcsin} \lambda_{13} + \text{arcsin} \lambda_{23} \right]; \quad (A3) \]
\[
B_i(\lambda_{12}, \lambda_{13}, \lambda_{23}) = \int_0^\infty dx_1 \int_0^\infty dx_2 \int_0^\infty dx_3 \, p_3(x_1, x_2, x_3)
\]
\[
= \frac{1}{\sqrt{32\pi}} \sum_{j=1}^3 \lambda_j \left[ 1 - \frac{2}{\pi} \arcsin \sigma_j \right], \quad i = 1, 2, 3,
\]
(A4)

where

\[
\sigma_1 = \frac{\lambda_{12}\lambda_{13} - \lambda_{23}}{\sqrt{(1 - \lambda_{12}^2)(1 - \lambda_{13}^2)}}, \quad \sigma_2 = \frac{\lambda_{12}\lambda_{23} - \lambda_{13}}{\sqrt{(1 - \lambda_{12}^2)(1 - \lambda_{23}^2)}},
\]
\[
\sigma_3 = \frac{\lambda_{13}\lambda_{23} - \lambda_{12}}{\sqrt{(1 - \lambda_{13}^2)(1 - \lambda_{23}^2)}},
\]
\[
C_i(\lambda_{12}, \lambda_{13}, \lambda_{23}) = \int_0^\infty dx_1 \int_0^\infty dx_2 \int_0^\infty dx_3 \, x_i^2 p_3(x_1, x_2, x_3)
\]
\[
= A(\lambda_{12}, \lambda_{13}, \lambda_{23}) + \frac{1}{4\pi} \left[ \frac{2\lambda_{ij}\lambda_{ik} - \lambda_{ij}^2\lambda_{ik} - \lambda_{ik}^2\lambda_{ij}}{\sqrt{1 - \lambda_{ik}^2}} + \lambda_{ik} \sqrt{1 - \lambda_{ik}^2} + \lambda_{ij} \sqrt{1 - \lambda_{ij}^2} \right]
\]
(A5)

where \(i, j, k\) is some permutation of 1, 2, 3:

\[
\bar{A}(\lambda_{12}) = \int_0^\infty dx_1 \int_0^\infty dx_2 \, p_2(x_1, x_2) = \frac{1}{4} + \frac{1}{2\pi} \arcsin \lambda_{12};
\]
(A6)

\[
\bar{B}_i(\lambda_{12}) = \int_0^\infty dx_1 \int_0^\infty dx_2 \, x_i p_2(x_1, x_2) = \frac{1 + \lambda_{12}}{\sqrt{8\pi}}, \quad i = 1, 2;
\]
(A7)

\[
\bar{C}_i(\lambda_{12}) = \int_0^\infty dx_1 \int_0^\infty dx_2 \, x_i^2 p_2(x_1, x_2)
\]
\[
= \bar{A}_2(\lambda_{12}) + \frac{\lambda_{12} \sqrt{1 - \lambda_{12}^2}}{2\pi}, \quad i = 1, 2,
\]
(A8)

The last three equations immediately follow from (A3)–(A5) by choosing \(\lambda_{13} = \lambda_{23} = 0\). The third Gaussian variate is then independent of the first two and one has \(A = \frac{1}{2} \bar{A}, B_i = \frac{1}{2} \bar{B}_i, C_i = \frac{1}{2} \bar{C}_i, i = 1, 2\), since the probability that the third variate be positive is \(\frac{1}{2}\).

Eqs. (A3) and (A6) appear in many places (see, for example, [5] or [6, Eqs. (6) and (7)]).

Eq. (A4) can be derived as follows:

\[
\sum_{j=1}^3 \lambda_{ij}^{-1} B_i(\lambda_{12}, \lambda_{13}, \lambda_{23}) = L_i = -\int_0^\infty dx_1 \int_0^\infty dx_2 \int_0^\infty dx_3 \frac{\partial}{\partial x_i} p_3(x_1, x_2, x_3),
\]
\[
i = 1, 2, 3, \quad (A9)
\]
as can be seen by differentiating \( p_3 \) given by (A1). The \( \lambda_{ij}^{-1} \) on the left-hand side of (A9) are elements of the inverse of \( \lambda_3 \) of (A2). We have, for example,

\[ L_1 = -\int_0^\infty dx_1 \int_0^\infty dx_2 \int_0^\infty dx_3 \frac{\partial}{\partial x_1} p_3 = \int_0^\infty dx_2 \int_0^\infty dx_3 p_3(0, x_2, x_3). \]

By suitable change of variables of integration the double integral here can be evaluated by using (A6). Permuting indices gives expressions for \( L_2 \) and \( L_3 \). From (A9), \( B_i(\lambda_{12}, \lambda_{13}, \lambda_{23}) = \sum_j \lambda_{ij} L_j \) and the result of the computation is (A4).

To obtain (A5), we observe that

\[ C_i(\lambda_{12}, \lambda_{13}, \lambda_{23}) = -\frac{2}{\sqrt{|\lambda_3|}} \int_0^\infty dx_1 \int_0^\infty dx_2 \int_0^\infty dx_3 \frac{\partial}{\partial \lambda_{ji}^{-1}} \sqrt{|\lambda_3|} p_3(x_1, x_2, x_3) \]

\[ = -\frac{2}{\sqrt{|\lambda_3|}} \frac{\partial}{\partial \lambda_{ji}^{-1}} \sqrt{|\lambda_3|} A(\lambda_{12}, \lambda_{13}, \lambda_{23}) \]  

(A10)

with \( A \) given explicitly by the right-hand side of (A3). The differentiation indicated in (A10) can be carried out using the formulae

\[ \frac{\partial}{\partial \lambda_{ji}^{-1}} \sqrt{|\lambda_3|} = \frac{\partial}{\partial \lambda_{ji}^{-1}} \frac{1}{\sqrt{|\lambda_3|}} = -\frac{1}{2} \lambda_{ij} \sqrt{|\lambda_3|} \]

and

\[ \frac{\partial}{\partial \lambda_{ji}^{-1}} = -\sum_{\mu\nu} \lambda_{i\mu} \lambda_{j\nu} \frac{\partial}{\partial \lambda_{\mu\nu}}. \]

Straight-forward algebra yields (A5).

Integrals differing from (A3)-(A8) by change of sign of an argument of \( p_3 \) or \( p_2 \) can readily be obtained from (A3)-(A8) by appropriate change of sign of the \( \lambda \)'s. The rule is simple: \( \lambda_{ij} \) is to be replaced by \( -\lambda_{ij} \) if either \( x_i \) or \( x_j \) is negated in \( p \), but not both. Thus, for example,

\[ \int_0^\infty dx_1 \int_0^\infty dx_2 \int_0^\infty dx_3 p_3(x_1, -x_2, x_3) = A(-\lambda_{12}, \lambda_{13}, -\lambda_{13}), \]

\[ \int_0^\infty dx_1 \int_0^\infty dx_2 \int_0^\infty dx_3 p(x_1, -x_2, -x_3) = A(-\lambda_{12}, -\lambda_{13}, \lambda_{23}). \]  

(A11)

References

