Edgeworth Expansions in State Perturbation Estimation

W. W. WILLIAMS, MEMBER, IEEE

Abstract—A formal asymptotic Hermite series expansion is developed for the conditional state density in a scalar estimation problem which differs from the standard Kalman filtering context only by the addition of a small quadratic term in the dynamics. This case is one that might result from a higher order description of perturbations in a more general state estimation situation. The parameters of the expansion here are generated recursively from the measurement data by a system of initial value ordinary differential equations for an expansion of nth order accuracy in the coefficient of the above quadratic term. To at least third order, this expansion has the special form of the standard Edgeworth expansion of corresponding order for the conditional state density, and its parameters can be expressed in terms of this density's cumulants, which are generated by a simpler set of recursive equations.

Introduction

There is considerable interest in state estimation problems where a state variable evolves according to linear dynamics with additive white Gaussian noise, and where linear state measurements are available which are also corrupted by the same type of noise. One reason this case is regarded as important is that it is the form assumed by the first-order description of noise-induced state and measurement perturbations from nominal values in a more general class of situations. If the description of such perturbations is carried out to higher order accuracy, the effect is typically to introduce quadratic perturbation terms in the dynamics and measurements. Some higher order expansions of this sort are considered in more detail in Gustafson and Speyer [1]. The resulting estimation problem can often be scaled so that the state and measurement perturbations are of order unity and the coefficients of the quadratic terms become the relatively small quantities.

A formal analysis has been made of estimation problems with this latter type of structure [2]. The results had the formal appearance of generating an approximation to the conditional probability density of the current state, given the currently available measurements, which was accurate to first order in the small coefficients. Similar formal results have also been obtained for the corresponding “smoothing problem” [3], where the density is conditioned on future measurements as well. In each case this first-order approximation to the conditional state density had the interesting property of being, at least in formal appearance, the standard first-order Edgeworth expansion [4] of that density. Also, the parameters of this expansion (first, second, and third central moments) could be generated recursively from the measurement data by a system of initial value equations. This formalism has been verified [5] under fairly general conditions for the special case of a scalar discrete-time problem in which the only nonlinearity is a small product of the state and noise variables in the dynamics, and in a sense which is strong enough to imply that the optimizing control for a quadratic performance criterion which is generated according to this Edgeworth approximation is indeed accurate to first order in the small coefficients.

Higher order Edgeworth expansions of the conditional state density, with parameters generated from the measurement data by a recursive system of equations, have been proposed by Sorensen and Stubberud [6], [7] for approximate state estimation. This idea is pursued here by the development of higher order approximations to the conditional state density in a scalar estimation problem in which the only nonlinearity is a small quadratic term in the noise and state variables of the dynamics, as in the case of [5]. The development here is only formal, however, i.e., it is

Fig. 9. Estimation error of \( V_y \), I proposed algorithm. II pseudolinear filter \( (m=4^\circ, \sigma=6^\circ) \). Example 3.

Fig. 10. Estimation error of \( V_y \), I proposed algorithm. II pseudolinear filter \( (m=4^\circ, \sigma=6^\circ) \). Example 3.

No attempt has been made to optimize the discretization of the parameter space. This is an area for further investigation.

References


U.S. Government work not protected by U.S. copyright
assumed without rigorous justification that all quantities that appear to be of high enough order to be negligible are so in some appropriate sense. As a result, this problem is considered in its continuous-time version here, or rather time discretizations thereof, because the formalism is actually simpler in the absence of any precise error analysis. Also, the approximations of the conditional state density are in terms of a more general series of Hermite polynomials, although these can be shown by computation to reduce to the special Edgeworth form out to at least third order. The hope is, of course, that this pattern continues to hold for more general types of higher order approximations, in terms of perturbation variables, of nonlinear state estimation problems.

In other related work, Kuznetsov, Stratonovich, and Tikhonov [8] have developed methods for using the more general type of Hermite series expansions just mentioned to characterize multidimensional distributions of random processes under more general conditions for purposes including nonlinear filtering. They call the parameters of these expansions quasi-moments, although these can be shown by computation to reduce to the special Edgeworth form out to at least third order. The hope is, of course, that this pattern continues to hold for more general types of higher order approximations, in terms of perturbation variables, of nonlinear state estimation problems.

To serve basically as an induction hypothesis on the time index, it is assumed temporarily that the conditional probability density of the current state at a generic epoch $i$, given the measurements $z_0, \cdots, z_i$, is of the form

$$p(x) = \left[\frac{e^{-(x-a)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}\right] \left[1 + \sum_{k=1}^{\infty} \frac{a_k H_k(x/a)}{\sqrt{k} \sigma^k} + \delta(x/a) \right]$$

where the epoch subscript $i$ and the conditioning on $z_0, \cdots, z_i$ are now suppressed in the notation, where $n$ is the integer specified in the problem statement, and where

$$H_k(u) = (-1)^k e^{u^2/2} \frac{d^k}{du^k} \left(e^{-u^2/2}\right)$$

and the following product rule holds for $i < k$:

$$H_i H_k = \sum_{m=0}^{i} \binom{i}{m} H_m H_{i-m}$$

The $H_k$ are the modified Hermite polynomials, so that $H_0(u) = 1$, $H_1(u) = u$, $H_2(u) = -u^2 - 1$, etc. Also,

$$\frac{d}{du} H_k(u) = k H_{k-1}(u) \quad k = 1, 2, \ldots$$

and the following product rule holds for $i < k$:

$$H_i H_k = \sum_{m=0}^{i} \frac{i!}{m!(i-m)!} H_m H_{i-m}$$

Hence, the integral, mean and variance of the expansion (5) are unity, $\hat{x}$ and $\sigma$, except for the contribution of the error term $\delta$. For this assumed form of the conditional density at epoch $i$, it is shown in the next two sections that except for an additional error which is formally of order $h^{n-\Delta}$, the conditional density at epoch $i+1$ is also of this form with the same error function $\delta$ if the corresponding parameters are certain functions of $\hat{x}$, $P$, $z_{i+1}$, and the $a_k$.

**Density Propagation**

In this section the “propagated” density $p[x_{i+1}/z_0, \cdots, z_i]$ is computed to an accuracy of order $h^{n-\Delta}$ under the assumption of (5) for epoch $i$. The result is a series of this same form, together with equations for its parameters in terms of $\hat{x}$, $P$, and the $a_k$, such that the residual error function corresponding to $\delta$ of (5) is unchanged to this order of accuracy.

By the standard formulas for probability density transformations, if $s = (1+\Delta)$, then

$$p(s) = \left[\frac{e^{-(s-x)^2/2\sigma^2}}{\sqrt{2\pi\sigma}}\right] \left[1 + \sum_{k=1}^{\infty} \frac{a_k H_k(x/a)}{\sqrt{k} \sigma^k} + \delta(x/a) \right]$$

where
\[ x = (1 + \beta A) \delta \]

\[ v = (1 + 2\beta A) P \]

except for terms of formal order \( A^2 \). Also, to this same accuracy

\[ y = x + \Delta^{1/2} u \]

where \( y \) now denotes \( x_{i+1} \) and \( u = \Delta^{-1/2}(1 + \psi s) w \). Hence,

\[ p(u) = \frac{e^{-u^2/2}(1 + \psi s)^2}{\sqrt{2\pi s}} \]

\[ E(y) = \bar{x} \]

\[ \text{var}(y) = \frac{m = v + qA \left[ (1 + \psi s)^2 + \psi^2 \right]}{} \]

Formally,

\[ p(y) = \int_{-\infty}^{\infty} p_s(y-u\Delta^{1/2}) p_{u/v}(u, y-u\Delta^{1/2}) \, du \]

except for terms of order \( \Delta^{1/2} \) (which are negligible to order \( h^\Delta \) by construction), where

\[ b = q(1 + \psi s)^2 \]

Substituting (17) into the integral of (16) and completing the square in the result in exponential factors gives, except for terms of formal order \( \Delta^{1/2} \),

\[ p(y) = \frac{e^{-(y-\bar{x})^2/2c}}{\sqrt{2\pi c}} \left[ 1 + c + \frac{y-\bar{x}}{c} \right] \]

where

\[ \gamma_n(x) \text{ denotes } \sum_{k=3}^{n} a_{n} H_k(x) \]

and

\[ \bar{u} = \frac{h^{3/2}}{v} \frac{y-\bar{x}}{\psi} \]

\[ c = b - \frac{h^{3/2}}{v} \]

Except for terms which are formally small compared to \( h^\Delta \), the density \( p(y) \) can now be obtained by the following steps.

1) Use (6) to develop a Taylor series expansion giving \( \gamma_n(y-\bar{x} - u\Delta^{1/2})/v \psi \) as a series of Hermite polynomials evaluated at \( (y-\bar{x})/\sqrt{m} \).

2) Substitute the result into (20) and integrate, adding the contribution of the error term \( \delta \) separately because it is not of a "pathological" nature, its contribution to this integral is just \( \delta((y-\bar{x})/\sqrt{m}) \) except for terms of order \( h^{1/2} \Delta \).

3) Use a Taylor series expansion to express the factor outside the integral in (20) as

\[ \frac{e^{-(y-\bar{x})^2/2c}}{\sqrt{2\pi c}} \left[ 1 + \frac{\psi^2 A}{v} \right] \]

\[ \int_{-\infty}^{\infty} \left[ \frac{e^{-\psi^2 A^2/2c}}{\sqrt{2\pi c}} \right] \]

\[ \left[ 1 + \sum_{k=3}^{n} b_{k} H_k \left( \frac{y-\bar{x}}{v} \right) \right] \]

where \( H_k \) are evaluated at \( (y-\bar{x})/\sqrt{m} \).

4) Multiply the results of steps 2) and 3) using (8).

The details of this procedure are too complicated to be presented here, but the eventual result after collecting terms and using the fact that \( a_{3k} \), \( a_{3k-1} \), and \( a_{3k-2} \) are of order \( h^k \) for all \( k = 1, \ldots, n \), is that (formally)

\[ p(y) = \frac{e^{-(y-\bar{x})^2/2m}}{\sqrt{2\pi m}} \left[ 1 + \sum_{k=3}^{n} b_{k} H_k \left( \frac{y-\bar{x}}{v} \right) \right] \]

where

\[ \int_{-\infty}^{\infty} \left[ \frac{e^{-(y-\bar{x})^2/2c}}{\sqrt{2\pi c}} \right] \]

\[ \left[ 1 + \sum_{k=3}^{n} b_{k} H_k \left( \frac{y-\bar{x}}{v} \right) \right] \]

\[ \left[ 1 + \frac{\psi^2 A}{v} \right] \]

\[ \int_{-\infty}^{\infty} \left[ \frac{e^{-\psi^2 A^2/2c}}{\sqrt{2\pi c}} \right] \]

\[ \left[ 1 + \sum_{k=3}^{n} b_{k} H_k \left( \frac{y-\bar{x}}{v} \right) \right] \]

As a consequence, \( b_{3k}, b_{3k-1}, \text{ and } b_{3k-2} \) are also formally of order \( h^k \) for all \( k = 1, \ldots, n \), and differ from \( a_{3k}, a_{3k-1}, \text{ and } a_{3k-2} \) only by order \( h^{1/2} \).

**UPDATING DENSITY WITH NEW MEASUREMENT**

This section consists of a computation of \( p(x_{i+1}/2\theta, \ldots, x_{i+1}) \) to order \( h^\Delta \) by using the Bayes rule and the results of the preceding section. The result is again a series of the form of (5), together with equations for its parameters in terms of \( x, m, z_{i-1}, \text{ and the } \theta, \) such that the residual error function is unchanged to order \( h^\Delta \).

Deleting the conditioning on \( z_{i-1}, \ldots, z_{i} \) and the epoch subscripts of \( z_{i-1} \) and \( r_{i-1} \) in the notation, and defining

\[ v = \Delta^{1/2} \]

and

\[ \bar{u} = \frac{h^{3/2}}{v} \frac{y-\bar{x}}{\psi} \]

\[ c = b - \frac{h^{3/2}}{v} \]

\[ \psi \]

Gives

\[ z = y + \Delta^{-1/2} \psi \]

treated as a term of order \( \Delta^{-1/2} \)

and

\[ p(z/y) = \frac{e^{-(z-y)^2/2\theta}}{\sqrt{2\pi \theta}} \]

(31)
Since the noise variables and initial state are all statistically independent
\[ p(y/z) = gp(y)p(z/y) \]  
(32)

by the Bayes Rule, where \( g \) is a proportionality factor independent of \( y \).

The density product in (32) can be expressed as
\[
\frac{e^{-\frac{1}{2}(y-z)^2/\Delta^2}}{\sqrt{2\pi\Delta}} \left( \frac{r+\Delta}{2} \right)^m \frac{e^{-\frac{1}{2}(z-x)^2/\Delta^2}}{\sqrt{2\pi\Delta}} \left( \frac{r+\Delta}{2} \right)^m
\]
(33)

where
\[ j = x + \frac{m\Delta}{r+\Delta}(z-x) = x + \frac{m\Delta}{r}(z-x) + \text{terms of order } \Delta^{1/2}. \]  
(34)

Now let
\[ s = m - \frac{m\Delta}{r} + \frac{6\Delta}{r} \frac{b_2m^{1/2}(z-x)}{r} \]  
(35)

Then the product of the last two factors in expression (33) can be expanded in a formal Taylor series as
\[
\frac{e^{-\frac{1}{2}(y-z)^2/\Delta^2}}{\sqrt{2\pi\Delta}} \left( \frac{r+\Delta}{2} \right)^m \left( 1 + \frac{9b_3m^3}{2r^3} (z-x)^2 \Delta^2 \right) \frac{e^{-\frac{1}{2}(z-x)^2/\Delta^2}}{\sqrt{2\pi\Delta}} \left( \frac{r+\Delta}{2} \right)^m \left[ 1 + \epsilon \left( \frac{z-x}{\Delta} \right) \right]
\]
(36)

except for terms that are of order higher than unity in \( \Delta \), which by construction are arbitrarily small compared to \( h^0 \).

Using (6) to develop a formal Taylor series expansion to first order in \( \Delta \) for \( H_i \) about \( (y-z)/\delta \) in (36), substituting the result in expression (33), and using the product formula (8) and the orders of magnitude (in \( h \)) of the \( b_k \) gives (32) as a series of the form
\[ p(y/z) = g \left[ 1 + \sum_{k=3}^{3n} c_k H_k \left( \frac{y-x}{\Delta} \right) + \epsilon \left( \frac{z-x}{\delta m} \right) \right] \]  
(37)

except for terms formally of order \( h^{n+1} \Delta \) or smaller. From the orthogonality property (7) and the fact that \( p(y/z) \) is a probability density
\[ 1 = \int_{-\infty}^{\infty} p(y/z) \, dy = g \left[ 1 + \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-x)^2/\Delta^2} \epsilon \left( \frac{y-x}{\delta m} \right) \, dy \right]. \]  
(38)

Using the fact that
\[ \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-x)^2/\Delta^2} \epsilon \left( \frac{y-x}{\delta m} \right) \, dy = 0 \]
by construction and a formal Taylor series expansion of \( \epsilon \) about this argument shows that
\[ g = \frac{1}{1+\alpha} \]  
(39)

where
\[ \alpha = \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-x)^2/\Delta^2} \epsilon \left( \frac{y-x}{\delta m} \right) \, dy \]  
(40)

Assuming that \( \epsilon \) is a "well-behaved" function so that \( \epsilon \) is also of order \( h^{n+1} \), it follows from (34) and (35) that
\[ \alpha = - \frac{6b_1}{r} (z-x) \Delta \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-x)^2/\Delta^2} \epsilon \left( \frac{y-x}{\delta m} \right) \, dy \]  
(41)

except for terms which are negligible to order \( h^n \). Since the integral in (41) is of order \( h^n \) and \( b_2 \) is of order \( h^0 \),
\[ a = \theta(z-x) \]  
(42)

where \( \theta \) is of order \( h^n \). Because the values of \( (z-x) \) are statistically independent innovations for distinct epochs, however, this term too can be ignored, since its cumulative effect in (39) for \( g \) over a time interval of order unity (basically a sum of \( 1/9 \) such terms) is a zero-mean random variable with variance of order \( h^{n+1} \) (because \( z-x \) is of order \( \Delta^{1/2} \)), which is negligible to order \( h^n \) with arbitrarily high probability by the Chebychev inequality. Hence, \( g \) in (37) can be taken as unity here. Deleting other similar zero-mean uncorrelated random terms leads eventually to the following equations determining the \( c_k \) of (37):
\[ \frac{1}{\Delta} (c_1 - b_1) = \left( 4b_4 - 9b_5 \right) \frac{\tilde{b}_1 m}{r} (z-x) + \frac{m}{2r} (15b_6 - 7b_7 b_2 - 3b_3) \]  
(42)

\[ \frac{1}{\Delta} (c_2 - b_2) = \left( 5b_6 - 12b_2 b_3 \right) \frac{\tilde{b}_1 m}{r} (z-x) \]  
(43)

\[ \frac{1}{\Delta} (c_3 - b_3) = \left( 6b_5 - 15b_4 b_2 - 3b_3 \right) \frac{\tilde{b}_1 m}{r} (z-x) + \frac{m}{2r} (315b_2^2 + 90b_2^2) \]  
(44)

\[ \frac{1}{\Delta} (c_4 - b_4) = \left( 7b_6 - 18b_5 b_2 - 3b_4 \right) \frac{\tilde{b}_1 m}{r} (z-x) + \frac{m}{2r} (216b_3 b_2 - 3b_3 - 3b_5) \]  
(45)

\[ \frac{1}{\Delta} (c_5 - b_5) = \left[ \left( k+1 \right) b_{k+1} + 3b_k \left( b_k + b_{k-2} \right) \right] \frac{\tilde{b}_1 m}{r} (z-x) \]  
(46)

except for terms which are formally negligible compared to \( h^n \). Furthermore, \( \epsilon(\tilde{y}-x)/\tilde{b}_1 \) is the same as \( \epsilon(\tilde{y}-x)/\tilde{b}_1 \) except for terms which can be neglected to this order of accuracy. From the form of (42)-(46), \( c_k, c_{k+1}, \) and \( c_{k-2} \) are formally of order \( h^k \), and differ from \( b_k, b_{k+1}, \) and \( b_{k-2} \) by order \( h^k \Delta \), for all \( k = 1, \ldots, n \). Also, for \( k = 3n, b_{k+1} = 0 \) in (46).

**Composite Result**

Since (34), (35), (37), and (42)-(46) describe the situation at epoch \( i + 1 \) corresponding to (5) for epoch \( i \), they, (10), (11), (15), and (24)-(28) form a recursion for generating the conditional state density to a formal accuracy of \( h^n \) from the incoming measurements. This follows by what is essentially an induction argument on \( i \). Equation (5) is clearly the form of the Gaussian prior density for \( x_0 \) with \( \theta_0 \) as specified. \( \delta \) and the \( a_k \) zero,
and an identically zero error \( \delta \). The previous two sections constitute the basis of an induction step showing that the conditional density remains of the form within an error small compared to \( h^n \), if the parameters are generated according to this recursion and the number of epochs is small compared to \((h A)^{-1}\), since the error accumulated at each step is of order \( h^{n+1} A \). In terms of the continuous-time problem being approximated [[(3) and (4)] this latter condition corresponds to a time interval small compared to \( 1/h \).

If the coefficients \( a_k \) are defined (for generic epoch \( i \)) as

\[
a_k = P^{k/2} a_k; \quad k = 3, \ldots, 3n
\]

(47)
it is a straightforward but tedious matter of substitution to convert this form to within an error small compared to \( h^n \) and \( \epsilon \). Expressing the composite recursion for \( k = 3 \) (proper) quasi-moment of Kuznetsov, Stratonovich, and Tikhonov [8], [10].

In terms of the continuous-time problem being approximated (i.e., to \( n = 3 \) in the notation here) that the expansion of (48) has the special form of a standard \( n \)th order Edgeworth expansion of the conditional density \( p(x/Y) \). In the third-order case, for example, the result can be expressed as

\[
p(x/Y) = \frac{e^{-(x-2)/2P}}{\sqrt{2\pi P}} \left\{ 1 + \sum_{k=3}^{3n} a_k P^{-k/2} H_k \left( \frac{x-\bar{x}}{\sqrt{P}} \right) \right\}
\]

(48)
to order \( h^n \), where (since \( \sqrt{\Delta} \approx h^n \))

\[
\dot{x} = \beta + \frac{P}{r} (z - \bar{x}); \quad \dot{\bar{x}} = 0
\]

(49)

\[P = 2fP + q \left[ (1 + \psi^2) + \frac{P^2}{r} \right] - \frac{P^2}{r} + 6a_3 (z - \bar{x}); \quad P(o) = P_o \]

(50)

\[
\dot{a}_3 = 3 \left( \frac{f - P}{r} + q \psi^2 \right) a_3 + P q (1 + \psi^2) + \frac{a_1}{2 P r} (135a_3^2 - 72 P a_4)
\]

\[
+ 4a_4 (z - \bar{x}); \quad a_3(0) = 0
\]

(51)

\[
\dot{a}_4 = 4 \left( \frac{f - P}{r} + \frac{3}{2} q \psi^2 \right) a_4 + 3 a_3 q (1 + \psi^2) + \frac{P^2}{r} q^2 + \frac{a_1}{2 P r} (126 a_4 - 9 P a_3 - 120 P a_5)
\]

\[
+ 5a_5 (z - \bar{x}); \quad a_4(0) = 0
\]

(52)

\[
\dot{a}_5 = 5 \left( \frac{f - P}{r} + \frac{5}{2} q \psi^2 \right) a_5 + 4 a_3 q (1 + \psi^2) + 3 a_1 P q^2 \psi
\]

\[
+ \frac{a_1}{2 P r} (315 a_5 - 24 P a_3 - 90 P a_6 - 33 a_2)
\]

\[
+ \frac{a_1}{2 P r} \left( 432 a_6 - 30 P a_3 - 36 P (7 a_7 - 3 a_2) \right)
\]

\[
+ \left( \frac{7 a_3 - 3 a_4}{2} \right) (z - \bar{x}); \quad a_6(0) = 0
\]

(53)

\[
\dot{a}_6 = 6 \left( \frac{f - P}{r} + \frac{5}{2} q \psi^2 \right) a_6 + q (5 a_3 + P a_5) (1 + \psi^2) \psi
\]

\[
+ 4 a_4 Q P q^2 + \frac{a_1}{2 P r} (432 a_6 - 30 P a_3 - 36 P (7 a_7 - 3 a_2))
\]

\[
+ \left( \frac{7 a_3 - 3 a_4}{2} \right) (z - \bar{x}); \quad a_6(0) = 0
\]

(54)

\[
\dot{a}_k = k \left( \frac{f - P}{r} + \frac{k - 1}{2} q \psi^2 \right) a_k + q (1 + \psi^2) \left[ (k - 1) a_{k-1} + P a_{k-3} \right]
\]

\[
+ P q^2 \left( (k - 2) a_{k-2} + \frac{1}{2} P a_{k-4} \right)
\]

\[
+ \frac{9 a_3}{2 P r} \left[ (k + 1) a_{k+1} + (k - 1) P a_{k-1} \right]
\]

\[
- \frac{3 a_1}{2} \left[ (k + 1) a_{k+1} + (k - 1) P a_{k-1} \right]
\]

\[
+ \left( \frac{(k + 1) a_{k+1} - 3 a_{k-1}}{r} \right) (z - \bar{x}); \quad k \geq 7; \quad a_6(0) = 0
\]

(55)

where \( Y(t) \) in (48) denotes the function \( \{ \{ y(s), s \}; 0 \leq s \leq t \} \), with \( y \) now denoting the data of (4) and \( z \) is the formal time-derivative of \( y \), which, for example, could reasonably be implemented here by taking

\[
z(t) = \frac{1}{\Delta} \left[ y(\Delta) - y((t-1)\Delta) \right]; \quad t \in \{I, (i+1)\Delta \}.
\]

(56)

The structure of the equation system (49)–(55) is such that, with the construction of (56), they can be regarded as ordinary differential equations without the need for correction terms of the sort described by Wong and Zakai [11], [12]. Of course, \( a_k+1 = 0 \) in (55) for \( k = 3n \). Many of the terms in this equation become negligible to order \( h^n \) as \( n \) grows.

It has also been verified by direct substitution out to third order in \( h \) and \( \epsilon \), which, for example, could reasonably be implemented here by taking

\[
\dot{x} = \beta + \frac{P}{r} (z - \bar{x}); \quad \dot{\bar{x}} = 0
\]

(58)

\[
\dot{p} = 2fP + q \left[ (1 + \psi^2) + \frac{P^2}{r} \right] - \frac{P^2}{r} + \frac{K_3}{2} (z - \bar{x}); \quad p(0) = p_o
\]

(59)

\[
\dot{a}_3 = 3 \left( \frac{f - P}{r} + \psi^3 - \frac{r}{2P} \right) a_3 + 4 a_3 q (1 + \psi^2) + \frac{P^2}{r} q^2 + \frac{a_1}{2 P r} (315 a_5 - 24 P a_3 - 90 P a_6 - 33 a_2)
\]

\[
+ \frac{a_1}{2 P r} \left( 432 a_6 - 30 P a_3 - 36 P (7 a_7 - 3 a_2) \right)
\]

\[
+ \left( \frac{7 a_3 - 3 a_4}{2} \right) (z - \bar{x}); \quad a_6(0) = 0
\]

(60)

\[
\dot{a}_4 = 4 \left( \frac{f - P}{r} + \frac{3}{2} q \psi^2 \right) a_4 + 3 a_3 q (1 + \psi^2) + \frac{P^2}{r} q^2 + \frac{a_1}{2 P r} (126 a_4 - 9 P a_3 - 120 P a_5)
\]

\[
+ 5a_5 (z - \bar{x}); \quad a_4(0) = 0
\]

(61)

\[
\dot{a}_5 = 5 \left( \frac{f - P}{r} + \frac{5}{2} q \psi^2 \right) a_5 + 4 a_3 q (1 + \psi^2) + 3 a_1 P q^2 \psi
\]

\[
+ \frac{a_1}{2 P r} \left( 315 a_5 - 24 P a_3 - 90 P a_6 - 33 a_2 \right)
\]

\[
+ \frac{a_1}{2 P r} \left( 432 a_6 - 30 P a_3 - 36 P (7 a_7 - 3 a_2) \right)
\]

\[
+ \left( \frac{7 a_3 - 3 a_4}{2} \right) (z - \bar{x}); \quad a_6(0) = 0
\]

(62)

The terms in (57) are grouped by orders of magnitude, since it is apparent from (58)–(62) that \( \kappa_3, \kappa_4, \) and \( \kappa_5 \) are of orders \( h^3 \), \( h^5 \), and \( h^7 \), respectively.

Unfortunately, not even a formal demonstration has been found to show for general \( n \) that the \( n \)th order approximation to \( p(x/Y) \) is its \( n \)th order Edgeworth expansion. Computationally, the special Edgeworth form has the advantage of reducing the number of independent coefficients by a factor of almost three—from \( 3n+2 \) to \( n+2 \). Conceptually, this form is interesting because it suggests some kind of connection with the Central Limit Theorem of statistics, from which Edgeworth expansions are originally derived. Finally, it should be noted that, even formally, the results here only constitute an asymptotic expansion for \( p(x/Y) \) in powers of the parameter \( h \) for some arbitrary but fixed \( n \), not a series converging to this.
density for fixed \( h \) and increasing \( s \). In the context of refinements of the Central Limit Theorem, the associated Edgeworth expansions sometimes converge in this latter sense also, but often do not for interesting classes of probability densities [4].

**References**


The Asymptotic Minimum Variance Estimate of Stationary Linear Single Output Processes

U. SHAKED and B. BOBOVSKY

**Abstract**—The problem of minimum error variance estimation of single output linear stationary processes in the presence of weak measurement noise is considered. By applying a domain analysis to the case of single input systems and white observation noise, explicit and simple expressions are obtained for the error covariance matrix of estimate and the optimal Kalman gains both for minimum- and nonminimum-phase systems. It is found that as the noise intensity approaches zero, the error covariance matrix of estimating the output and its derivatives becomes insensitive to uncertainty in the system parameters. This matrix depends only on the shape of the high frequency tail of the power-density spectrum of the observation, and thus it can be easily determined from the system transfer function. The theory developed is extended to deal with white measurement noise in multi-input systems where an analogy to the single input nonminimum-phase case is established. The results are also applied to colored measurement noise problems and a simple method to derive the minimum error covariance matrices and the optimal filter transfer functions is introduced.

**I. INTRODUCTION**

The general problem of optimal estimation of linear processes in the presence of measurement noise has been vastly dealt with in the past. Special attention has been devoted to the stationary case with measurement noise of very small intensity [1]. The importance of this case stems from the following reasons:

1. In many practical problems in communication the processes are stationary and the signal-to-noise ratio (SNR) is high [2]. In FM communication, for example, the threshold occurs at SNR > 100 [14].
2. The solution for this case, which is relatively easy to obtain, may indicate the maximal accuracy achievable in the estimation process [3].
3. It is shown later that the theory which is developed for cases of weak measurement noise can be readily extended to deal with colored noise problems. The solution obtained by this approach is shown to be identical to results found by the commonly used algorithm of [4], but significantly easier to derive.

This problem is dual to the steady-state "cheap-control" problem of time invariant regulators that has been intensively investigated in the literature [3], [5], [6].

Significant contributions to the solution of this case have been made in the past [2], [5]–[8]. A comprehensive treatment of the single output process case is introduced in the present paper. The asymptotic properties of the closed-loop poles of the optimal filter [9], [10] and the "domain equivalent of the algebraic Riccati equation [11] are first used to derive explicit expressions for \( k \), the optimal Kalman gain, in the case of single stochastic driving input and white measurement noise of small intensity \( \rho \).

Choosing a specific canonical state space representation for the asymptotically stable linear process, it is found that these gains are insensitive to the location of the open-loop poles, assuming, of course, that all these poles stay in the left half plane. By denoting the system order by \( n \) and the excess of poles over zeros in the system open-loop transfer function by \( i \) and \( k! = (k_1, k_2, \ldots, k_i) \), it is found that, in the limit as \( \rho \) tends to zero, \( k! \approx a_{i} \rho^{1-i} \), where \( a_{i} \) is related to the input mapping matrix of the minimum-phase image of the system for \( i = 1 \) and to the coefficient of the \( n \)th order Butterworth polynomial for \( i < 1 \).

The expression for \( k! \) is substituted in the algebraic Riccati equation to provide a Lyapunov type equation for the MECM (minimum error covariance matrix of estimate). From this equation the different powers of \( \rho \) according to which the estimation error covariance of the system output and its first \( i–1 \) derivatives tend to zero are found explicitly. The theory developed is then extended to also deal with systems of several inputs, and the obtained results are used to derive the solution of the colored measurements noise problem.

The paper is organized as follows. In Section II the explicit solution for the single input case is given. Minimum and nonminimum-phase systems are considered. In the minimum-phase case the pattern according to which the elements of the MECM approach asymptotically to zero is found. In the nonminimum-phase case the submatrix of the MECM whose elements always approach zero in the limit as the noise intensity tends to zero is located. The part of the MECM that does not tend to zero, in this limit, is also determined and it is related to the inherent structural difference that exists between the system and its minimum-phase image.

In Section III the theory is extended to cope with multi-input cases. The different asymptotic behavior of the elements of the MECM in this case is discussed and the similarity to the result found in the single-input nonminimum-phase case is demonstrated. The approach of this section is applied in Section IV to derive solutions to the singular colored measurement noise filtering problems. Two problems are treated. In the first, the intensity of the measurement noise is assumed to approach zero and results that are identical to the solution of the white noise problem are obtained. The second problem concerns the case where the colored measurement noise contains a white noise component of intensity that tends to zero. The asymptotic behavior of the MECM is discussed and the optimal filter, which applies finite gains only, is derived.

An example that illustrates different aspects of the proposed way of solution is given in Section II.

**II. THE CASE OF SINGLE INPUT AND WHITE MEASUREMENT NOISE**

Given the linear time-invariant system

\[
\dot{x} = Ax + b w \\
y = c' x
\]

with the observation process


Paper recommended by A. Ephratiades, Past Chairman of the Estimation Committee.

The authors are with the Department of Electronic Communications, Control and Computer Science, Tel-Aviv University, Tel-Aviv, Israel.

0018-9286/81/0400-0498$00.75 © 1981 IEEE