MEAN-SQUARE JOINT STATE AND NOISE INTENSITY ESTIMATION FOR LINEAR STOCHASTIC SYSTEMS

MICHAEL BASIN\textsuperscript{1,2}, ALEXANDER LOUKIANOV\textsuperscript{2} AND MIGUEL HERNANDEZ-GONZALEZ\textsuperscript{2}

\textsuperscript{1}Department of Physical and Mathematical Sciences
Autonomous University of Nuevo Leon
mbasin@fcfm.uanl.mx; mbasin2007@gmail.com

\textsuperscript{2}Center for Research and Graduate Studies (CINVESTAV)
Campus Guadalajara
Jalisco, Mexico
{mbasin; louk; mglez}@gdl.cinvestav.mx

Received August 2009; revised March 2010

Abstract. This paper presents the mean-square joint state and diffusion coefficient (noise intensity) estimator for linear stochastic systems with unknown noise intensity over linear observations, where unknown parameters are considered Wiener processes. Designing the mean-square joint state and noise intensity estimator presents a significant advantage in the filtering theory and practice, since it enables one to address the mean-square state estimation problems for linear systems influenced by stochastic disturbances with an unknown variable noise level and, in addition, to construct the mean-square estimate for the noise intensity. The original problem is reduced to the filtering problem for an extended state vector that incorporates parameters as additional states. Since the noise intensities cannot be observable in the original linear system, the new quadratic vector variable formed by the diagonal of the matrix square of the system state is introduced. The obtained mean-square filter for the extended state vector also serves as the optimal identifier for the unknown parameters. Performance of the designed mean-square state filter and parameter identifier is verified in an illustrative example.

Keywords: Filtering, Parameter identification, Linear stochastic system

1. Introduction. The problem of the optimal simultaneous state estimation and parameter identification for stochastic systems with unknown parameters has been receiving systematic treatment beginning from the seminal paper [2]. The optimal result was obtained in [2] for a linear discrete-time system with constant unknown parameters within a finite filtering horizon, using the maximum likelihood principle (see, for example, [18]), in view of a finite set of the state and parameter values at time instants. The application of the maximum likelihood concept was continued for linear discrete-time systems in [11] and linear continuous-time systems in [10]. Nonetheless, the use of the maximum likelihood principle reveals certain limitations in the final result: a. the unknown parameters are assumed constant to avoid complications in the generated optimization problem and b. no direct dynamical (difference) equations can be obtained to track the optimal state and parameter estimates dynamics in the “general situation”, without imposing special assumptions on the system structure. Other approaches are presented by the optimal parameter identification methods without simultaneous state estimation, such as designed in [9, 7, 26], which are also applicable to nonlinear stochastic systems. Robust approximate identification in nonlinear systems using various approaches, such as $H_\infty$ filtering, is studied in a variety of papers [4, 5, 8, 12, 13, 14, 15, 16, 19, 20, 21, 22, 23, 24, 25] for
linear stochastic systems with bounded uncertainties in coefficients. The overall comment is that the optimal state filter and diffusion coefficient (noise intensity) identifier in the form of a closed finite-dimensional system of stochastic ODEs has not yet been obtained even in case of uncertain linear stochastic systems with unknown noise intensity in the state equation.

This paper presents the mean-square joint state and diffusion coefficient (noise intensity) estimator for linear stochastic systems with unknown noise intensity over linear observations, where unknown parameters are considered Wiener processes. Designing the mean-square joint state and noise intensity estimator presents a significant advantage in the filtering theory and practice, since it enables one to address the mean-square state estimation problems for linear systems influenced by stochastic disturbances with an unknown variable noise level and, in addition, to construct the mean-square estimate for the noise intensity. Various practical estimation problems, such as the position/velocity estimation problem for a re-entry vehicle landing in an atmosphere with variable unknown drag coefficient (see [1]), can be directly addressed using the designed estimator.

The filtering problem is formalized considering the unknown parameters as additional system states satisfying linear stochastic Ito equations with zero drift and unit diffusion. Since the noise intensities cannot be observable in the original linear system, the new quadratic vector variable formed by the diagonal of the matrix square of the system state is introduced. In view of the Ito formula, the stochastic differential equation for the new variable contains the diffusion coefficient also as a part of the drift term, which makes it observable. Thus, the problem is reduced to the filtering problem for polynomial (bilinear) system states over linear observations, whose solution is obtained in [3, 6]. This paper presents the mean-square algorithm for simultaneous state estimation and parameter identification in linear systems with unknown noise intensity over linear observations. Note that since the original identification problem is reduced to the filtering problem for the extended system state including both state and parameters, the identifiability condition for the original system coincides with the observability condition for the extended system.

In the illustrative example, performance of the designed mean-square filter is verified for a linear stochastic system with unknown noise intensity in the state equation. The simulation results demonstrate reliable performance of the filter: the state estimate converges to the real state and the parameter estimate converges to the real parameter value rapidly. Note, however, that only local convergence of the state and parameter estimates to their real values can be assured, since the received measurement signal is insufficiently rich to provide global convergence of all state and parameter estimates.

The paper is organized as follows. Section 2 presents the mean-square state filtering and noise intensity identification problem statement for linear systems with unknown noise intensities. In Section 3, the stated problem is reduced to the filtering problem for an extended state vector that incorporates parameters as additional states. The mean-square filtering equations are then obtained. In Section 4, performance of the designed mean-square filter is verified for a linear system with an unknown noise intensity in the example. Section 5 presents conclusions to this study.

2. Filtering Problem for Linear Systems with Unknown State Diffusion Coefficients. Let \((\Omega, F, P)\) be a complete probability space with an increasing right-continuous family of \(\sigma\)-algebras \(F_t; t \geq t_0\), and let \((W_1(t), F_t, t \geq t_0)\) and \((W_2(t), F_t, t \geq t_0)\) be independent standard Wiener processes. The \(F_t\)-measurable random process \((x(t), y(t))\) is described by a linear differential equation for the system state with an unknown diffusion
coefficient (noise intensity) $\theta(t)$

$$dx(t) = (a_0(t) + a(t)x(t))dt + (\theta(t) + b_1(t))dW_1(t), \quad x(t_0) = x_0,$$

and a linear differential equation for the observation process

$$dy(t) = (A_0(t) + A(t)x(t))dt + B(t)dW_2(t).$$

(2)

Here, $x(t) \in \mathbb{R}^n$ is the state vector, $y(t) \in \mathbb{R}^m$, $m \leq n$, is the linear observation vector, and $\theta(t) \in \mathbb{R}^{nxn}$ is the diagonal matrix of unknown diffusion coefficients (noise intensities), whereas the matrix $b_1(t) \in \mathbb{R}^{nxn}$ represents known diffusion coefficients. The initial condition $x_0 \in \mathbb{R}^n$ is a Gaussian vector such that $x_0, W_1(t)$, and $W_2(t)$ are independent. It is assumed that $B(t)B^T(t)$ is a positive definite matrix. All coefficients in (1)–(2) are deterministic functions of time of appropriate dimensions.

It is considered that there is no useful information on values of the unknown parameters $\theta_k(t), k = 1, \ldots, p$, and this uncertainty even grows as time tends to infinity. In other words, the unknown parameters can be modeled as $F_t$-measurable Wiener processes

$$d\theta(t) = \beta(t)dW_3(t),$$

(3)

with unknown initial conditions $\theta(t_0) = \theta_0 \in \mathbb{R}^p$, where $(W_3(t), F_t, t \geq t_0)$ is a standard Wiener process independent of $x_0, W_1(t)$, and $W_2(t)$, $\beta(t) \in \mathbb{R}^{nxp}$ is the intensity matrix.

The estimation problem is to find the mean-square estimate $\hat{z}(t) = [\hat{x}(t), \hat{\theta}(t)]$ of the combined vector of the system states and unknown parameters $z(t) = [x(t), \theta(t)]$, based on the observation process $Y(t) = \{y(s), t_0 \leq s \leq t\}$. As known [17], the mean-square estimate is given by the conditional expectation

$$\hat{z}(t) = m(t) = E(z(t) \mid F_t^Y)$$

of the system state $z(t) = [x(t), \theta(t)]$ with respect to the $\sigma$-algebra $F_t^Y$ generated by the observation process $Y(t)$ in the interval $[t_0, t]$. As usual, the matrix function

$$P(t) = E[(z(t) - m(t))(z(t) - m(t))^T \mid F_t^Y]$$

is the conditional variance of the estimation error with respect to the observations.

The proposed solution to this optimal filtering problem is based on the mean-square filtering equations for polynomial states over linear observations obtained in [3]. Note that direct application of the mean-square filter from [3] to the system (1), (3) does not allow to obtain the mean-square estimate for the parameter $\theta(t)$, since the system (1), (3) cannot be completely observable over the observations (2). To overcome this difficulty, the mean-square filter is applied to the quadratic state vector $\zeta(t) = [x_1^2(t), \ldots, x_n^2(t)]$, where $x(t) = [x_1(t), \ldots, x_n(t)]$ satisfies (1), and the linear state (3) over the observations (2).

3. Mean-Square Filter for Linear Systems with Unknown State Diffusion Coefficients. The identification algorithm for the parameter $\theta(t)$ satisfying (3) is based on considering the variable $\zeta(t) = A(x(t)x^T(t)) \in \mathbb{R}^n$, where $A$ is the 3D tensor defined by $A_{ijk} = 1$, if $i = j = k$, and $A_{ijk} = 0$, otherwise. In other words, $\zeta(t) = [x_1^2(t), \ldots, x_n^2(t)]$ is the vector of diagonal entries of the matrix $x(t)x^T(t)$. In view of the Ito formula (see [17]), the stochastic Ito differential equation for $\zeta(t)$ is given by

$$d\zeta(t) = A(x(t)(a_0(t) + a(t)x(t))^T)dt$$

$$+ A(x(t)((\theta(t) + b_1(t))dW_1(t))^T) + A((a_0(t) + a(t)x(t))x^T(t))dt$$

$$+ A((\theta(t) + b_1(t))dW_1(t))x^T(t)) + A((\theta(t) + b_1(t))(\theta(t) + b_1(t))^T)dt$$

$$= 2(A(x(t)(a_0(t) + a(t)x(t))^T))dt + A(x(t)((\theta(t) + b_1(t))dW_1(t))^T)$$

$$+ A((\theta(t) + b_1(t))dW_1(t))x^T(t)) + A((\theta(t) + b_1(t))(\theta(t) + b_1(t))^T)dt,$$

(4)
with the initial condition \( \tilde{z}(t_0) = E(A(x(t_0)\xi(t_0)) | F^Y_{t_0}) \). The Equation (4) is bilinear with respect to the extended state vector \( z(t) = [x(t), \theta(t)] \).

Thus, the estimation problem is now reformulated as to find the optimal estimate \( \tilde{z}_1(t) = [\tilde{z}(t), \tilde{x}(t), \tilde{\theta}(t)] \) for the state vector \( z_1(t) = [z(t), x(t), \theta(t)] \), governed by the Equations (4), (1), (3), based on the observation process \( Y(t) = \{y(s), t_0 \leq s \leq t \} \), satisfying the Equation (2). The solution of this problem is obtained using the mean-square filtering equations for polynomial states over linear observations [3] and is given by

\[
d\tilde{z}(t) = A(\tilde{x}(t))a_G(t)dt + A(x(t))a_1(t)dt + A(P_{xx}(t))a_1(t)dt
\]

\[
d\tilde{x}(t) = (a_0(t) + a(t)\tilde{x}(t))dt + P_{xx}(t)[dy(t) - (A_0(t) + A(t)\tilde{x}(t))dt],
\]

\[
d\tilde{\theta}(t) = P_{y\theta}(t)[dy(t) - (A_0(t) + A(t)\tilde{x}(t))dt],
\]

with the initial conditions \( \tilde{z}(t_0) = E(A(x(t_0)\xi(t_0)) | F^Y_{t_0}), \tilde{x}(t_0) = E(x_0 | F^Y_{t_0}), \) and \( \tilde{\theta}(t_0) = E(\theta_0 | F^Y_{t_0}) \).

\[
dP_{zz}(t) = 2(a_0(t) + a(t)\tilde{x}(t))dt + A((a(t)P_{xx}(t))\tilde{x}(t))dt + A(P_{y\theta}(t))\tilde{x}(t)dt
\]

\[
dP_{xx}(t) = (a(t)P_{xx}(t)) + P_{xx}(t)a_1(t)dt + \tilde{\theta}(t)\tilde{\theta}(t)
\]

\[
dP_{y\theta}(t) = (P_{y\theta}(t)a_1(t)) - P_{y\theta}(t)A(t)(G(t)G^T(t))^{-1}A(t)P_{xx}(t)dt,
\]

with the initial condition \( P(t_0) = E((z_1(t_0) - \tilde{z}_1(t_0))(z_1(t_0) - \tilde{z}_1(t_0)^T) | F^Y_{t_0}) \), where the equations for the other entries of the conditional variance matrix \( P(t) \) of the estimation error \( z_1(t) - \tilde{z}_1(t) \) are not required. Here, the \(*\)-product of a vector \( a \in R^n \) by a matrix \( A \in R^{n \times n} \) results in a matrix \( (a * A) \in R^{n \times n} \) obtained by multiplication of each row of the matrix \( A \) by the corresponding component of the vector \( a \): \( (a * A)_{ij} = a_iA_{ij}, \) \( i,j = 1, \ldots, n \).

**Theorem 3.1.** The mean-square finite-dimensional filter for the extended state vector \( z_1(t) = [z(t), x(t), \theta(t)] \), governed by the Equations (4), (1), (3) over the linear observations (2) is given by the Equations (5) for the mean-square estimate \( \tilde{z}_1(t) = [\tilde{z}(t), \tilde{x}(t), \tilde{\theta}(t)] \) = \( E([\tilde{z}(t), x(t), \theta(t)] | F^Y_{t}) \) and the Equations (6) for certain components of the estimation error variance \( P(t) = E((z_1(t) - \tilde{z}_1(t))(z_1(t) - \tilde{z}_1(t)^T) | F^Y_{t}) \). This filter, applied to the subvector \( \theta(t) \), also serves as the mean-square identifier for the vector of unknown parameters \( \theta(t) \) in the Equation (1), yielding the estimate subvector \( \tilde{\theta}(t) \) as the mean-square parameter estimate.

**Proof:** The proof directly follows from the optimal filtering equations for incompletely measured polynomial states over linear observations, obtained in [3].

**Remark 3.1.** Note that in view of the physical sense of a noise intensity and the structure of a Wiener process, only the absolute values of the noise intensity diagonal matrix entries, \( \| \theta_{11}(t) \|, \ldots, \| \theta_{nn}(t) \| \), are significant. Those absolute values can be unambiguously determined if the diagonal entries of the matrix \( \theta(t) \theta(t)^T, [\theta_1^2(t), \ldots, \theta_n^2(t)] \), appearing in the right-side of (4), are identified.
Thus, based on the general mean-square filtering equations for polynomial systems, the mean-square state filter and parameter identifier is obtained for the linear system state (1) with unknown parameters, modeled by (3), over the linear observations (2). Since the original identification problem is reduced to the filtering problem for the extended system state including both state and parameters, the identifiability condition for the original system (1) coincides with the observability condition for the extended system (4), (1), (3).

4. Example. This section presents an example of designing the optimal filter and identifier for a linear system state with an unknown noise intensity, based on linear state measurements.

Let the scalar system state \( x(t) \in R \) satisfy the linear equation with an unknown noise intensity \( \theta(t) \in R \)

\[
dx(t) = x(t)dt + \theta(t)W_1(t), \quad x(0) = x_0, \quad (7)
\]

and the scalar observation process \( y(t) \in R \) be given by the linear equation

\[
dy(t) = x(t)dt + dW_2(t), \quad (8)
\]

where \( W_1(t) \) and \( W_2(t) \) are standard Wiener processes. The parameter \( \theta(t) \) is modeled as a standard Wiener process, i.e., satisfies the equation

\[
d\theta(t) = dW_3(t), \quad \theta(0) = \theta_0. \quad (9)
\]

The considered Wiener processes are assumed independent.

In view of the Ito formula (see [17]), the equation for the variable \( \zeta(t) = x^2(t) \) takes the form

\[
d\zeta(t) = 2x^2(t)dt + \theta^2(t)dt + 2x(t)\theta(t)dW_1(t). \quad (10)
\]

The filtering problem is to find the optimal estimate for the bilinear-linear state (10), (7) and (9) using the linear observation Equation (8). The simulation time is set to \( T = 1 \).

The filtering Equations (5), (6) take the following particular form for the system (10), (7), (9)

\[
d\hat{\zeta}(t) = (2\hat{x}^2(t) + 2P_{xx}(t) + \hat{\theta}^2(t) + P_{\theta\theta}(t))dt + P_{\zeta x}(t)[dy(t) - \hat{x}(t)dt],
\]

\[
d\hat{x}(t) = \hat{x}(t)dt + P_{xx}(t)[dy(t) - \hat{x}(t)dt],
\]

\[
d\hat{\theta}(t) = P_{\theta x}(t)[dy(t) - \hat{x}(t)dt],
\]

with the initial conditions \( \hat{\zeta}(0) = E(x_0^2 \mid y(0)), \hat{x}(0) = E(x_0 \mid y(0)), \) and \( \hat{\theta}(0) = E(\theta_0 \mid y(0)) \).

\[
\hat{P}_{\zeta x}(t) = P_{\zeta x}(t) + 4\hat{x}(t)P_{xx}(t) + 2\hat{\theta}(t)P_{\theta x}(t) + 4\hat{\theta}(t)P_{\theta\theta}(t) + 2\hat{x}(t)\hat{\theta}^2(t) + 2\hat{x}(t)P_{\theta\theta}(t) - P_{\zeta x}(t)P_{xx}(t) \quad (11)
\]

\[
\hat{P}_{xx}(t) = 2P_{xx}(t) + P_{\theta\theta}(t) + \hat{\theta}^2(t) - P_{xx}(t),
\]

\[
\hat{P}_{\theta x}(t) = P_{\theta x}(t) - P_{\theta x}(t)P_{xx}(t),
\]

\[
\hat{P}_{\theta\theta}(t) = 1 - P_{\theta\theta}(t),
\]

with the initial condition \( P(0) = E((z_1(0) - \hat{z}_1(0))(z_1(0) - \hat{z}_1(0)^T) \mid y(0)) \).

Numerical simulation results are obtained solving the system of filtering Equations (11)–(12). The obtained values of \( \hat{x}(t), \) estimate for \( x(t), \) and \( \hat{\theta}(t), \) estimate for \( \theta(t), \) are compared to the real values of the state variable \( x(t) \) and parameter \( \theta(t) \) in (7) and (9).

For the filter Equations (11), (12) and the reference system (7)–(9) involved in simulation, the following initial values are assigned: \( x(0) = 3, \hat{x}(0) = 1, \hat{\zeta}(0) = 1, \theta(0) = 0, P_{xx}(0) = 87, P_{\theta\theta}(0) = 23, P_{\zeta}(0) = 12, P_{\theta x}(0) = 46. \) The unknown parameter \( \theta \) is assigned as \( \theta = 1. \) Gaussian disturbances \( dW_1(t), dW_2(t), \) and \( dW_3(t) \) are realized using the
built-in MatLab white noise function. High-frequency oscillations in the output graphs are smoothed by a Butterworth low-pass built-in MatLab filter, with low-pass threshold \( \nu = 30 \).

The following graphs are obtained: graphs of the reference state \( x(t) \) and its mean-square estimate \( \hat{x}(t) \) (Figure 1), graphs of the state \( \zeta(t) \) and its mean-square estimate \( \hat{\zeta}(t) \) (Figure 2), and graphs of the unknown parameter \( \theta \) and its mean-square estimate \( \hat{\theta}(t) \) (Figure 3). The graphs of all those variables are shown in the entire simulation interval from \( t_0 = 0 \) to \( T = 1 \).

It can be observed that the state estimate \( \hat{x}(t) \) converges to the real state \( x(t) \), the state estimate \( \hat{\zeta}(t) \) converges to the real state \( \zeta(t) \), and the parameter estimate \( \hat{\theta}(t) \) converges to the real value of the unknown parameter \( \theta(t) \) rapidly, in less than one time unit. This behavior can be classified as very reliable, especially taking into account considerable deviations in the initial values for the real state and its estimate and large values of the initial error variances. Note, however, that only local convergence of the state and parameter estimates to their real values can be assured, since the received measurement signal is insufficiently rich to provide global convergence of all state and parameter estimates.

![Graphs of the reference state and its mean-square estimate](image)

**Figure 1.** Graphs of the reference state \( x(t) \) (thin line) and its mean-square estimate \( \hat{x}(t) \) (thick line) in the simulation interval \([0, 1]\).

5. **Conclusions.** This paper presents the mean-square state filter and diffusion coefficient (noise intensity) identifier for linear stochastic systems with unknown noise intensity over linear observations, where unknown parameters are considered Wiener processes. Since the noise intensities cannot be observable in the original linear system, the new quadratic vector variable formed by the diagonal of the matrix square of the system state is introduced. In view of the Ito formula, the stochastic differential equation for the new variable contains the diffusion coefficient also as a part of the drift term, which makes it observable. The mean-square state filter and parameter identifier is designed in the form of a closed finite-dimensional system of stochastic ODEs. This result is theoretically proved based on the previously obtained mean-square filter for polynomial system states over linear observations and numerically verified. The simulation results show reliable behavior of the designed mean-square state filter and parameter identifier. The proposed approach would be applied in the future to designing joint state and noise intensity estimators for nonlinear polynomial systems and systems with other than Gaussian noises.
Figure 2. Graphs of the state $\zeta(t)$ (thin line) and its mean-square estimate $\hat{\zeta}(t)$ (thick line) in the simulation interval $[0, 1]$

Figure 3. Graphs of the unknown parameter $\theta$ (thin line) and its mean-square estimate $\hat{\theta}(t)$ (thick line) in the simulation interval $[0, 1]$

Acknowledgment. The authors thank the Mexican National Science and Technology Council (CONACyT) for financial support under Grants 55584 and 46069 and Sabbatical Fellowship for the first author.

REFERENCES

M. V. Basin, P. Shi and D. Calderon-Alvarez, Optimal filtering for polynomial states over polynomial observations, Proc. of the 47th IEEE Conf. on Decision and Control, Cancun, Mexico, pp.5128-5133, 2008.


