Observers for Nonlinear Stochastic Systems

TZYH-JONG TARN, MEMBER, IEEE, AND YONA RASIS

***Abstract*—Motivated by the complexity and the large quantity of on-line operations required for nonlinear filtering problems, observers for nonlinear stochastic systems are constructed based on a Lyapunov-like method. Sufficient conditions on the structure of a nonlinear stochastic system for the existence of an exponentially bounded observer are given. These conditions can be applied off-line. The stabilization of unstable nonlinear stochastic systems using observer feedback are investigated. Sufficient conditions to stabilize cascaded control and observer in a feedback arrangement are given.**

I. INTRODUCTION

THE PROBLEM of estimating the state variables of a dynamical system given observations of the output variables is of fundamental importance in control theory since many feedback control designs require the availability of the state of the controlled plant.

If one considers the class of linear systems, then there are two approaches available. If the output variables can be measured exactly and if there are no stochastic disturbances acting on the system, then one can use a Luenberger observer to reconstruct the state [1]–[3]. On the other hand, if all the output variables are corrupted by additive white noise then one can use a Kalman filter [4]–[7] for a state estimation. The relation between the two for linear systems was clarified by Tse and Athans [8]–[10].

There is extensive past work on nonlinear estimation on stochastic continuous systems using various approaches by Stratonovich [11], Kushner [12], Wonham [13], Bucy [14], [15], Frost and Kailath [16], and others; however, unlike the linear case the solution to the nonlinear estimation problem is given as a nonexplicit representation.

To obtain a practical solution to the nonlinear estimation problem, that is, a suitable filter, one can find algorithms for the evaluation of the representation [17], [18], (these have severe limitations due to the extensive computational work required) or use suboptimal filters [19], [20] based on the assumption that the conditional density can be adequately characterized by its low-order moments.

For stochastic discrete nonlinear systems, the situation is even more unsatisfactory. Only scattered results based on approximation theory exist. Several authors have suggested filtering algorithms based on the Gram–Charlier expansion of the conditional density [21], [22], or on the related Edgeworth expansion [23]. Kizner [24] has suggested basing an approximation scheme on the Hermite expansion. Alspach and Sorenson [25] have suggested using a weighted sum of Gaussian densities to approximate the conditional density. Center [26] suggested a new class of filtering algorithms based on a generalization of standard least square approximation. Bucy and Senne [27] have used a recursive algorithm which is related to generalized least square approximation by step functions.

Summarizing, the usefulness of nonlinear filters is limited by the complexity and the quantity of on-line operations required. As a result, nonlinear filters have not yet become practical.

In this paper we will approach nonlinear estimation through the concepts of observer for stochastic nonlinear systems. Potentially, this route may prove more advantageous for implementation than the approach through estimation theory, as is now advocated. We use a Lyapunov-like method to construct observers for both continuous and discrete nonlinear stochastic systems driven by noises with bounded covariances. Sufficient conditions on the structure of the observed system for the existence of such observers are established. These conditions can be applied off-line and can be used in the design of observers of known characteristics. The sufficient conditions on the stability of cascaded control and observer in a feedback arrangement is presented.

II. LYAPUNOV-LIKE CONDITIONS FOR STOCHASTIC STABILITY

The stability of systems modeled by ordinary differential or difference equations has been studied by many mathematicians. The most powerful tool in this field seems to be Lyapunov's second method.

Lyapunov conditions for the stability of stochastic differential and difference equations were developed by Wonham [28], [29], Kushner [30], [31], Zakai [32], [33], Miyahara [34], and others.

There are several ways to define stability (convergence) of stochastic systems (see, for example, [35]). To avoid confusion we would like to first state what definitions on stochastic stability will be used. Motivated by the criterion frequently used in estimation theory, i.e., minimum mean square, we define stability as follows.

Definition 1: The origin of a continuous stochastic process x_t is said to be asymptotically stable in mean square if there exist constants $\alpha > 0$, $k_1 \geqslant 0$ and $k_2 > 0$ such that

$$E\|x_t\|^2 \leqslant k_1 + k_2 e^{-\alpha t},$$

and then the process x_t is said to be exponentially bounded in mean square with exponent α.

Manuscript received March 10, 1975; revised February 18, 1976. Paper recommended by E. Tse, Past Chairman of the IEEE S-CS Stochastic Control Committee. This work was supported in part by the National Science Foundation under Grants ENG72-04160A02, ENG75-09755, and GK22905A #2 and #3.

T.-J. Tarn is with the Department of Systems Science and Mathematics, Washington University, St. Louis, MO 63130.

Y. Rasis is with the Sachs Electric Company, St. Louis, MO 63110.

Definition 2: The origin of a discrete stochastic process x_n is said to be asymptotically stable in mean square if there exist constants $0<\alpha\leqslant 1$, $k_1\geqslant 0$ and $k_2>0$ such that

$$E\|x_n\|^2\leqslant k_1+k_2(1-\alpha)^n,$$

and then x_n is said to be exponentially bounded in mean square with exponent α.

Remark: Definition 1 (or 2) does not necessarily imply that $E\|x_t\|^2$ (or $E\|x_n\|^2$) decreases for all t (or n), only the bound decreases exponentially and as $t\to\infty$ (or $n\to\infty$) the mean square of the process is bounded by

$$E\|x_\infty\|^2\leqslant k_1$$

where k_1 depends on the noise disturbing the system.

If $k_1=0$, that is, if

$$E\|x_t\|^2\leqslant ke^{-\alpha t}\quad\left(\text{or } E\|x_n\|^2\leqslant k(1-\alpha)^n\right),$$

then x_t (or x_n) converges to zero in mean square (which also implies convergence in probability).

Continuous Systems

Consider a dynamical system modeled by the Itô stochastic differential equation

$$dx_t=f(x_t)\,dt+\sigma(x_t)\,dw_t \tag{1}$$

where $x_t'=(x_t^{1'},x_t^{2'})$ and $f(x_t)$ are vectors in R^{2n}, x_t^1 and x_t^2 are vectors in R^n, $\sigma(x_t)$ is a $2n\times q$ matrix valued function, w_t is the standard q-dimensional Brownian motion process, and $f(x_t)$ and $\sigma(x_t)$ satisfy a global Lipschitz condition, namely, for all $x,y\in R^{2n}$

(C1) $\|f(x)-f(y)\|+\|\sigma(x)-\sigma(y)\|\leqslant c\|x-y\|$

and linear growth condition

(C2) $\|f(x)\|^2+\|\sigma(x)\|^2\leqslant k(1+\|x\|^2)$

for vectors $\|f\|=(\Sigma_i f_i^2)^{1/2}$ and for matrices $\|\sigma\|=(\Sigma_{i,j}\sigma_{i,j}^2)^{1/2}$.

Associated with the processes defined by (1) is the Kolmogorov (backward) differential operator

$$\mathcal{L}(\cdot)=\sum_i f_i(x)\frac{\partial(\cdot)}{\partial x_i}+\frac{1}{2}\operatorname{tr}(\sigma(x)\sigma'(x)\nabla\nabla'(\cdot))$$

where

$$\nabla(\cdot)=\operatorname{col}\frac{\partial(\cdot)}{\partial x_j}.$$

Also, we denote

$$E_af(x_t)\triangleq E[f(x_t)|x_0=a].$$

We will consider functions $V(x)$ with the following properties [32].

1) $V(x)$ is real-valued, nonnegative, and twice continuously differentiable.

2) Let $F(a,t)$ stand for any of the functions

$$E_aV(x_t),E_a|\mathcal{L}V(x_t)|\quad\text{or}\quad E_a\left|\frac{\partial V(x_t)}{\partial x_i}\sigma_{ij}(x_t)\right|^2.$$

Then $F(a,t)$ is, for each a, bounded in t in any bounded t interval.

Theorem 1: If x_t is generated by (1) and if $V(\epsilon_t)$ (with $\epsilon_t=x_t^1-x_t^2,\epsilon_0=a$) satisfies 1) and 2) such that

a) $V(\epsilon_t)\geqslant c\|\epsilon_t\|^2$, for all $\epsilon_t\in R^n$, $c>0$, $V(0)=0$,

b) $\mathcal{L}V(\epsilon_t)\geqslant k-\alpha V(\epsilon_t)$, for all $\epsilon_t\in R^n$, $k\geqslant 0$, $\alpha>0$,

then

$$E_af(x_t)\triangleq E[f(x_t)|x_0=a].$$

Proof: The proof follows essentially from Zakai [32]. By properties 1) and 2), the expectation of the integrated form of Itô's rule gives

$$E_aV(\epsilon_t)=V(\epsilon_0=a)+\int_0^t E_a\mathcal{L}V(\epsilon_s)\,dt.$$

It follows that $E_aV(\epsilon_t)$ is absolutely continuous in t (with respect to Lebesgue measure). Therefore, for almost all s, $s\geqslant 0$

$$\frac{dE_aV(\epsilon_s)}{ds}=E_a\mathcal{L}V(\epsilon_s)\leqslant k-\alpha E_aV(\epsilon_s),$$

so we have

$$\frac{de^{\alpha s}E_aV(\epsilon_s)}{ds}\leqslant ke^{\alpha s}.$$

Integrating the last inequality we get

$$e^{\alpha t}E_aV(\epsilon_t)-V(\epsilon_0=a)\leqslant\frac{k}{\alpha}(e^{\alpha t}-1),$$

from which the result of the theorem follows.

Remarks: 1) A scalar function $V(\epsilon_t)=V(x_t^1-x_t^2)$ satisfying conditions a) and b) of Theorem 1 will be called a Lyapunov-like function for system (1). Since V is not positive definite on R^{2n}, the existence of the Lyapunov-like function does not guarantee any Lyapunov stability for system (1).

2) Theorem 1 gives sufficient conditions so that the origin of the stochastic process ϵ_t will be asymptotically stable in mean square in the large (i.e., for all initial states $\epsilon_0=a$). Since as $t\to\infty$,

$$E\|\epsilon_\infty\|^2\leqslant\frac{k}{c\alpha}$$

regardless of the initial state ϵ_0.

Corollary 1: If x_t is generated by (1) and if $V(x_t^1,\epsilon_t)$ (with $\epsilon_t=x_t^1-x_t^2$, $(x_0^1,\epsilon_0)=a$) satisfies 1) and 2) such that

a) $V(x_t^1,\epsilon_t)\geqslant c\|(x_t^{1'},\epsilon_t')\|^2$, for all $x_t\in R^{2n}$, $c>0$, $V(0,0)=0$,

b) $\mathcal{L}V(x_t^1,\epsilon_t)\leqslant k-\alpha V(x_t^1,\epsilon_t)$, for all $x_t\in R^{2n}$, $k\geqslant 0$, $a>0$,

then

$$cE_a\|(x_t^{1'},\epsilon_t')\|^2\leqslant V(a)e^{-\alpha t}+\frac{k}{\alpha}(1-e^{-\alpha t}).$$

Proof: Similar to Theorem 1.

Discrete Systems

Consider a dynamical system described by the stochastic difference equation

$$x_{k+1}=f(x_k)+\sigma(x_k)w_k \tag{2}$$

where $x_k'=(x_k^{1'},x_k^{2'})$ and $f(x_k)$ are vectors in R^{2n}, x_k^1 and x_k^2 are vectors in R^n, $\sigma(x_k)$ is a $2n\times q$ matrix valued function and w_k is a q-dimensional sequence of uncorrelated normalized Gaussian random variables.

Theorem 2: If x_n is generated by (2) and if there exists a function $V(\epsilon_n)$ (with $\epsilon_n=x_n^1-x_n^2, \epsilon_0=a$) such that

a) $V(\epsilon_n)\geqslant c\|\epsilon_n\|^2$, for all $\epsilon_n\in R^n$, $c>0$, $V(0)=0$

b) $E_{\epsilon_n} V(\epsilon_{n+1})-V(\epsilon_n)\leqslant k-\alpha V(\epsilon_n)$, for all $\epsilon_n\in R^n$, $k>0$, $0<\alpha\leqslant 1$,

then

$$cE_a\|\epsilon_n\|^2\leqslant(1-\alpha)^nV(a)+k\sum_{i=0}^{n-1}(1-\alpha)^i.$$

Proof: From condition b)

$$E_{\epsilon_{n-1}}V(\epsilon_n)\leqslant k+(1-\alpha)V(\epsilon_{n-1}).$$

Using the property

$$E_{\epsilon_{n-2}}V(\epsilon_n)=E_{\epsilon_{n-2}}E_{\epsilon_{n-1}}V(\epsilon_n)$$

we have

$$E_{\epsilon_{n-2}}V(\epsilon_n)\leqslant E_{\epsilon_{n-2}}[k+(1-\alpha)V(\epsilon_{n-1})]$$
$$=k+(1-\alpha)E_{\epsilon_{n-2}}V(\epsilon_{n-1})$$

and by using condition b) again, we get

$$E_{\epsilon_{n-2}}V(\epsilon_n)\leqslant k+(1-\alpha)[k+(1-\alpha)V(\epsilon_{n-2})]$$
$$=k+k(1-\alpha)+(1-\alpha)^2V(\epsilon_{n-2}).$$

Continuing this way and applying condition a) we obtain

$$cE_a\|\epsilon_n\|^2\leqslant(1-\alpha)^nV(a)+k\sum_{i=0}^{n-1}(1-\alpha)^i.$$

Remark: Theorem 2 gives sufficient conditions so that the origin of the stochastic process ϵ_n will be asymptotically stable in mean square in the large. Since as $n\to\infty$

$$E\|\epsilon_\infty\|^2\leqslant\frac{k}{c\alpha}$$

regardless of the initial state ϵ_0.

Corollary 2: If x_n is generated by (2) and if there exists a function $V(x_n^1,\epsilon_n)$ (with $\epsilon_n=x_n^1-x_n^2$, $(x_0^1,\epsilon_0)=a$) such that

a) $V(x_n^1,\epsilon_n)\geqslant c\|(x_n^{1'},\epsilon_n')\|^2$, for all $x_n\in R^{2n}$, $c>0$, $V(0)=0$;

b) $E_{(x_n^1,\epsilon_n)}V(x_{n+1}^1,\epsilon_{n+1})-V(x_n^1,\epsilon_n)\leqslant k-\alpha V(x_n^1,\epsilon_n)$, for all $x_n\in R^{2n}$, $k>0$, $0<\alpha\leqslant 1$,

then

$$cE_\alpha\|(x_n^{1'},\epsilon_n')\|^2\leqslant(1-\alpha)^nV(a)+k\sum_{i=1}^{n-1}(1-\alpha)^i.$$

Proof: Similar to Theorem 2.

III. Requirements of the Observer Design

Consider a nonlinear system modeled by the Itô stochastic differential equation

$$dx_t=f(x_t)\,dt+\sigma_1(x_t)\,dw_t \tag{3}$$

with observation equation

$$dy_t=h(x_t)\,dt+\sigma_2(x_t)\,dv_t \tag{4}$$

where x_t and $f(x_t)$ are n-dimensional vectors, y_t and $h(x_t)$ are m-dimensional vectors, $f(\cdot)$ and $h(\cdot)$ are continuously differentiable, w_t and v_t are, respectively, q- and r-dimensional uncorrelated standard Brownian motion processes, and $\sigma_1(x_t)$ and $\sigma_2(x_t)$ are matrix functions of appropriate dimensions. Also assume that both system (3) and the observation equation (4) satisfy (C1) and (C2).

The problem is to design a dynamic system (observer, state estimator) for system (3), using the output, y_t from (4) as the input, such that the difference between the observer output z_t and the system state x_t is exponentially bounded in mean square. Such an observer output can then be used as an estimate of the system state.

In choosing the observer structure, two possible approaches to the design should be distinguished. One is to seek the best possible observer without any additional constraints. On the other hand, additional constraints (motivated by practical, mathematical, or computational considerations) may be imposed.

In general, the observer can be described by any nonlinear stochastic differential equation of the form

$$dz_t=g_1(z_t,y_t)dt+g_2(z_t,y_t)dy_t. \tag{5}$$

In this paper we limit ourselves to observers that satisfy the following constraints.

a) The observer is an identity observer, that is, the observer state, z_t, has the same dimension as the system state x_t.

b) The observer is linear with respect to the observed data (but is not a linear observer), this constraint limits us to observers of the form

$$dz_t=g(z_t)\,dt+B\,dy_t$$

where B is an $n\times m$ constant matrix.

c) The observer is an asymptotic observer for the given system in the absence of system noise w_t and observation noise v_t. The motivation for this constraint is that we would like the observer to be an asymptotic observer even when some, or all, the components of the system state or the observation are noise free.

Under the above constraints the observer equation (5) reduces to [36]

$$dz_t=f(z_t)\,dt+B[dy_t-h(z_t)\,dt] \tag{6}$$

where B is an $n\times m$ constant matrix.

To solve (6) we have to choose an appropriate initial state z_0. The best possible choice would be $z_0=x_0$, but since in general x_0 is unknown we have to guess, or assume, some value for z_0. Because of this the initial error

$x_0 - z_0$ can be big and unknown and therefore it is important to require that the error be exponentially bounded in mean square in the large (i.e., for all values of an initial error).

IV. Observers for Nonlinear Systems

In this section we apply the Lyapunov-like method presented in Section II to obtain sufficient conditions on the structure of a nonlinear stochastic system for the existence of an observer.

Continuous Systems

Consider a dynamic system modeled by (3) with observation equation (4). In the following we present sufficient conditions between the matrix B and the functions f and h of the system, such that the error between the system state and the state of an observer modeled by (6) be exponentially bounded in mean square with given exponent α.

Theorem 3: If the covariances $\sigma_1(x_t)$, $\sigma_2(x_t)$ are bounded and if there exists an $n\times n$ positive definite symmetric matrix Q and a constant $n\times m$ matrix B such that $Q(\nabla f - B\nabla h + (\alpha/2)I)$ is uniformly negative semidefinite, then the error $\epsilon_t \triangleq x_t - z_t$ is exponentially bounded in mean square with given exponent $\alpha > 0$.

Proof: The proof is to show that there exists a Lyapunov-like function having properties 1) and 2) and satisfying the conditions of Theorem 1.

Define

$$V(\epsilon_t) = V(x_t, z_t) = (x_t - z_t)' Q (x_t - z_t)$$

where Q is an $n\times n$ positive definite symmetric matrix. $V(\epsilon_t)$ satisfies condition 1) trivially. Equation (1) satisfies (C1) and (C2) implies that $E_a\|x_t\|^p$, $p>0$ is bounded in any bounded t interval; therefore, $V(\epsilon_t)$ satisfies condition 2), because from assumptions V, $|\mathcal{L}V|$ and $|(\partial V/\partial x_i)\,\sigma_{ij}|$ are dominated by polynomials. Since

$$V(\epsilon_t) \geq \lambda_{\min}(Q)\|\epsilon_t\|^2, \qquad \lambda_{\min}(Q) > 0$$

condition a) of Theorem 1 is satisfied. The error will be exponentially bounded in mean square if

$$\mathcal{L}V(x_t, z_t) \leq k - \alpha V(x_t, z_t). \tag{7}$$

When x_t and z_t are given by (3) and (6), respectively, then by setting

$$\sup_{x_t \in R^n} \operatorname{tr}\left(\sigma_1(x_t)\sigma_1'(x_t) + B\sigma_2(x_t)\sigma_2'(x_t)B'\right)Q = k < \infty$$

(7) can be written as

$$2(x_t - z_t)'Q\left\{f(x_t) - Bh(x_t) - \left[f(z_t) - Bh(z_t)\right]\right\} \leq -\alpha(x_t - z_t)'Q(x_t - z_t). \tag{8}$$

By the fundamental theorem of integral calculus for vector-valued functions of several variables [37], we have

$$\left[f(x_t) - Bh(x_t)\right] - \left[f(z_t) - Bh(z_t)\right] = \int_0^1 (\nabla f - B\nabla h)(w_s)(x_t - z_t)\,ds \tag{9}$$

where

$$w_s \triangleq s x_t + (1-s) z_t.$$

Substituting (9) into (8) and recalling that $(x_t - z_t) \triangleq \epsilon_t$ is independent of s we get

$$2\epsilon_t' \int_0^1 Q\left[\nabla f - B\nabla h + \frac{\alpha}{2} I\right](w_s)\,ds \cdot \epsilon_t \leq 0. \tag{10}$$

Thus, we see that (7) will be satisfied if the matrix $Q[\nabla f - B\nabla h + \frac{1}{2}\alpha I]$ is uniformly negative semidefinite.

Remark: When the nonlinear system is modeled by the stochastic differential equation

$$dx_t = f(x_t)\,dt + g(y_t)\,dt + \sigma_1(x_t)\,dw_t$$

that is, when we have an observation dependent feedback, the observer is

$$dz_t = f(z_t)\,dt + g(y_t)\,dt + B\left[dy_t - h(z_t)\,dt\right].$$

For this observer we can show that the condition that the matrix $Q(\nabla f - B\nabla h + \frac{1}{2}\alpha I)$ be uniformly negative semidefinite is sufficient for the error ϵ_t to be exponentially bounded in mean square with exponent α.

Example 1: Consider the continuous nonlinear system

$$dx_1 = x_1\,dt + dw_t$$
$$dx_2 = (x_1 - 3x_2 + e^{-x_2})\,dt$$

with observation

$$dy = x_1\,dt + x_2\,dt + \sin x_1\,dv_t.$$

Let us choose $Q = I$, $B' = [2, 1]$; then it is easy to check that this system satisfies the condition of Theorem 3.

Nonlinear Observers for Discrete Systems

Now let us shift our attention to the discrete systems. We consider a dynamic system modeled by the stochastic difference equation

$$x_{k+1} = f(x_k) + \sigma_1(x_k) w_k \tag{11}$$

with observation equation

$$y_k = h(x_k) + \sigma_2(x_k) v_k \tag{12}$$

where x_k and $f(x_k)$ are n-dimensional vectors, y_k and $h(x_k)$ are m-dimensional vectors, $f(\cdot)$ and $h(\cdot)$ are continuously differentiable, w_k and v_k are, respectively, q- and r-dimensional sequences of zero mean uncorrelated normalized Gaussian random variables, w_k and v_k are independent of x_k and $\sigma_1(x_k)$ and $\sigma_2(x_k)$ are matrix functions with appropriate dimensions.

As in the continuous case, we look for sufficient conditions on the system structure (the functions f and h) such that the error, $\epsilon_k \triangleq x_k - z_k$, between the system state and the state of an observer modeled by the stochastic difference equation

$$z_{k+1} = f(z_k) + B[y_k - h(z_k)] \tag{13}$$

will be exponentially bounded in mean square with given exponent $0 < \alpha \leqslant 1$.

Theorem 4: If the covariances $\sigma_1(x_k)$, $\sigma_2(x_k)$ are bounded and if there exist an $n \times m$ matrix B such that

$$\|(\nabla f - B \nabla h)(w_s)\| \leqslant \sqrt{1-\alpha}, \qquad \text{for all } w_s \in R^n$$

where

$$\|(\nabla f - B \nabla h)(w_s)\| = \sup_{\|x_k\|=1} \|(\nabla f - B \nabla h)(w_s) \cdot x_k\|;$$

then the error, $\epsilon_k \triangleq x_k - z_k$, is exponentially bounded in mean square with given exponent α.

Proof: The proof is to show that there exists a Lyapunov-like function satisfying the conditions of Theorem 2.

Define

$$V(\epsilon_k) = V(x_k, z_k) = (x_k - z_k)' Q (x_k - z_k)$$

where Q is an $n \times n$ positive definite symmetric matrix. Since

$$V(\epsilon_k) \geqslant \lambda_{\min}(Q) \|\epsilon_k\|^2 \quad \text{and} \quad \lambda_{\min}(Q) > 0$$

condition a) of Theorem 2 is satisfied. The error will be exponentially bounded in mean square if

$$E_{\epsilon_k} V(\epsilon_{k+1}) - V(\epsilon_k) \leqslant k - \alpha V(\epsilon_k) \tag{14}$$

when $\epsilon_k = x_k - z_k$, and x_k and z_k are given by (11) and (13), respectively,

$$\begin{aligned} E_{\epsilon_k} V(\epsilon_{k+1}) = E_{\epsilon_k} \{ & f(x_k) - Bh(x_k) - [f(z_k) - Bh(z_k)] \\ & + \sigma_1(x_k) w_k - B\sigma_2(x_k) v_k \}' \\ & \cdot Q \{ f(x_k) - Bh(x_k) - [f(z_k) - Bh(z_k)] \\ & + \sigma_1(x_k) w_k - B\sigma_2(x_k) v_k \}. \end{aligned} \tag{15}$$

By the assumption on w_k and v_k we have

$$\begin{aligned} & E_{\epsilon_k} \{ \sigma_1(x_k) w_k - B\sigma_2(x_k) v_k \}' \\ & \quad \cdot Q \{ f(x_k) - Bh(x_k) - [f(z_k) - Bh(z_k)] \} = 0 \end{aligned} \tag{16}$$

and

$$\begin{aligned} & E_{\epsilon_k} \{ f(x_k) - Bh(x_k) - [f(z_k) - Bh(z_k)] \}' \\ & \quad \cdot Q \{ \sigma_1(x_k) w_k - B\sigma_2(x_k) v_k \} = 0. \end{aligned} \tag{17}$$

Since

$$\begin{aligned} & \{ \sigma_1(x_k) w_k - B\sigma_2(x_k) v_k \}' Q \{ \sigma_1(x_k) w_k - \sigma_2(x_k) v_k \} \\ & \quad = \operatorname{tr} Q \{ \sigma_1(x_k) w_k - B\sigma_2(x_k) v_k \} \{ \sigma_1(x_k) w_k - B\sigma_2(x_k) v_k \}', \end{aligned}$$

and w_k and v_k are uncorrelated normalized random variables, we obtain

$$\begin{aligned} & E_{\epsilon_k} \{ \sigma_1(x_k) w_k - B\sigma_2(x_k) v_k \}' Q \{ \sigma_1(x_k) w_k - B\sigma_2(x_k) v_k \} \\ & \quad = \operatorname{tr}(\sigma_1(x_k)\sigma_1'(x_k) + B\sigma_2(x_k)\sigma_2'(x_k)B')Q. \end{aligned} \tag{18}$$

Substituting (16), (17), and (18) into (15) and setting

$$\sup_{x_k \in R^n} \operatorname{tr}(\sigma_1(x_k)\sigma_1'(x_k) + B\sigma_2(x_k)\sigma_2'(x_k)B')Q = k < \infty,$$

then (14) can be written as

$$\begin{aligned} & \{ f(x_k) - Bh(x_k) - [f(z_k) - Bh(z_k)] \}' Q \{ f(x_k) - Bh(x_k) \\ & \quad - [f(z_k) - Bh(z_k)] \} \leqslant (1-\alpha)(x_k - z_k)' Q (x_k - z_k). \end{aligned} \tag{19}$$

By the fundamental theorem of integral calculus for vector-valued functions of several variables, we have

$$\begin{aligned} & [f(x_k) - Bh(x_k)] - [f(z_k) - Bh(z_k)] \\ & \qquad = \int_0^1 (\nabla f - B \nabla h)(w_s)(x_k - z_k)\, ds \end{aligned} \tag{20}$$

where $w_s = s x_k + (1-s) z_k$.

Substituting (20) into (19), recalling that $\epsilon_k = x_k - z_k$ is independent of s, we get

$$\begin{aligned} & \epsilon_k' \int_0^1 (\nabla f - B \nabla h)'(w_s)\, ds\, Q \cdot \int_0^1 (\nabla f - B \nabla h)(w_s)\, ds\, \epsilon_k \\ & \qquad \leqslant (1-\alpha) \epsilon_k' Q \epsilon_k. \end{aligned} \tag{21}$$

Thus, (14) will be satisfied if

$$\|(\nabla f - B \nabla h)(w_s)\| \leqslant \sqrt{1-\alpha}, \qquad \text{for all } w_s \in R^n. \tag{22}$$

Remark: When the nonlinear system is modeled by the stochastic difference equation

$$x_{k+1} = f(x_k) + g(y_k) + \sigma_1(x_k) w_k,$$

the observer will be

$$z_{k+1} = f(z_k) + g(y_k) + B[y_k - h(z_k)].$$

For this observer we can show that the condition

$$\|(\nabla f - B \nabla h)(w_s)\| \leqslant \sqrt{1-\alpha}$$

for all $w_s \in R^n$ is sufficient for the error ϵ_k to be exponentially bounded in mean square.

Example 2: Consider the discrete nonlinear system

$$\begin{aligned} x_1(k+1) &= a \sin x_1(k) + 2x_2(k) \\ x_2(k+1) &= x_2(k) + w_k \end{aligned}$$

with observation

$$y(k) = c \sin x_1(k) + x_2(k) + 0.2 v_k.$$

Let $Q = I$ and choose $B' = [2, 1]$. It is easy to see that if for any given α, $(a - 2c^2) + c^2 \leqslant 1 - \alpha$, then this system satisfies the conditions of Theorem 4.

Observers for a Class of Nonlinear Systems

Theorems 3 and 4 provide sufficient conditions for the existence of explicit nonlinear observers as represented by (6) and (13). These conditions can be checked off-line.

In general, it is difficult to check whether a given matrix is uniformly negative semidefinite. We would like to find more efficient sufficient conditions on the system structure so that the error between the system state and the observer be exponentially bounded in mean square with given exponent α.

Along this direction we are considering a special class of nonlinear systems modeled by the stochastic differential equation

$$dx_t = Ax_t\,dt + f(x_t)\,dt + g(y_t)\,dt + \sigma_1(x_t)\,dw_t \quad (23)$$

with observation equation

$$dy_t = Cx_t\,dt + \sigma_2(x_t)\,dv_t \quad (24)$$

where x_t, $f(x_t)$ and $g(y_t)$ are n-dimensional vectors, y_t an m-dimensional vector, $f(\cdot)$ is continuously differentiable, w_t and v_t are, respectively, q- and r-dimensional uncorrelated standard Brownian motion processes, A and C are constant $n\times n$ and $m\times n$ matrices, respectively, and $\sigma_1(x_t)$ and $\sigma_2(x_t)$ are matrix functions of appropriate dimensions. Also, we assume that $f(x_t)$, $g(y_t)$, $\sigma_1(x_t)$ and $\sigma_2(x_t)$ satisfy (C1) and (C2).

We are seeking a bound on the nonlinear function $f(x_t)$ such that the existence of an asymptotic observer for the linear system, obtained by assuming $f(x_t)=0$, will imply the existence of an asymptotic observer for the whole system.

Theorem 5: Given a dynamic system modeled by (23) with observation equation modeled by (24). If the covariances $\sigma_1(x_t)$, $\sigma_2(x_t)$ are bounded and if there exist two constant $n\times n$ positive definite symmetric matrices P and Q, and a constant $n\times m$ matrix B such that

$$Q\left(A - BC + \frac{1}{2}\alpha I\right) + \left(A - BC + \frac{1}{2}\alpha I\right)'Q = -P \quad (25)$$

$$\{\lambda_{\min}(P)/2\lambda_{\max}(Q)\} > \|\nabla f\|_\infty \quad (26)$$

where $\|\nabla f\|_\infty = \sup_{x\in R^n}\|\nabla f\|$. Then with the observer given by

$$dz_t = Az_t\,dt + f(z_t)\,dt + g(y_t)\,dt + B\left[dy_t - Cz_t\,dt\right], \quad (27)$$

the error ϵ_t is exponentially bounded in mean square with exponent α.

Proof: The proof is to show that when (25) and (26) are satisfied there exists a Lyapunov-type function having properties 1) and 2) and satisfying the conditions of Theorem 1.

Define

$$V(\epsilon_t) = V(x_t, z_t) = (x_t - z_t)'Q(x_t - z_t)$$

where Q is an $n\times n$ positive definite symmetric matrix. Clearly, $V(\epsilon_t)$ has properties 1) and 2) and condition a) of Theorem 1 is satisfied. The error will be exponentially bounded in mean square if

$$\mathcal{L}V(\epsilon_t) = \mathcal{L}V(x_t, z_t) \leqslant k - \alpha V(x_t, z_t). \quad (28)$$

From this point, we follow the same line of reasoning used in going from (7) to (10) in the proof of Theorem 3. Assuming that (25) is satisfied, we arrive at the following equation:

$$-\epsilon_t' P\epsilon_t + 2\epsilon_t' \int_0^1 Q\nabla f(w_s)\,ds\cdot\epsilon_t \leqslant 0. \quad (29)$$

Finally, for real symmetric matrices P, we have

$$\lambda_{\min}(P)\|\epsilon_t\|^2 \leqslant \epsilon_t' P\epsilon_t \leqslant \lambda_{\max}(P)\|\epsilon_t\|^2. \quad (30)$$

By the Schwartz inequality we obtain

$$|\epsilon_t' Q\,\nabla f(w_s)\epsilon_t| \leqslant \|Q\|\cdot\|\nabla f\|_\infty\cdot\|\epsilon_t\|^2 \quad (31)$$

where for constant matrices $\|Q\| \triangleq [\lambda_{\max}(Q'Q)]^{1/2}$.

Substituting inequalities (30) and (31) into (29) we see that (28) will be satisfied if $\{\lambda_{\min}(P)/2\lambda_{\max}(Q)\} > \|\nabla f\|_\infty$, which is (26). Thus, according to Theorem 1 the error is exponentially bounded in mean square with the given exponent α.

Remark: It should be pointed out that this class of systems is important since many systems consist of a linear plant with nonlinear output feedback, bounded nonlinear perturbations, and linear observation.

V. Stabilization of Stochastic Nonlinear Systems

Methods of using feedback control to stabilize otherwise unstable systems were developed early in the Laplace transform based control theory of the 1950's. Later state space based approaches were proposed by several researchers, among them Langenhop [38] and Wolovich [39] for deterministic linear systems, and Haussman [40] for linear stochastic systems.

In the previous section, we proposed to design the observers for unforced nonlinear systems which are either stable or unstable. One of the most important roles of observers is to provide input for feedback control design and thus to facilitate system stabilization. So the problem of stability of the overall feedback controlled system with the observer is important.

A cascade of a nonlinear observer and a controller would not be expected to produce the optimal stochastic controller necessarily, but it should produce a good, efficient, implementable, and stable control which may not be significantly inferior in expected cost to the theoretically ultimate optimal stochastic control.

We first consider the stochastic continuous system

$$dx_t = f(x_t)\,dt + u_t\,dt + \sigma_1(x_t)\,dw_t \quad (32)$$

$$dy_t = h(x_t)\,dt + \sigma_2(x_t)\,dv_t \quad (33)$$

where (32) and (33) satisfy the assumption of (3) and (4) and the unforced system of (32) is unstable, u_t is an additive feedback control which is designed to stabilize the unstable system, that is, to steer its state to the origin. However, we cannot design u_t to be a function of x_t since the state of the system is not available for direct measurement. In order to implement the feedback control law we need the estimate of the state. Assume $u_t = g(z_t)$, so the observer for (32) and (33) will be

$$dz_t = f(z_t)\,dt + g(z_t)\,dt + B[dy_t - h(z_t)\,dt]. \tag{34}$$

Now the design problem is to choose a proper continuously differentiable function $g(z_t)$ and a constant matrix B such that both the state of the system (32) and the error $\epsilon_t = x_t - z_t$ will be exponentially bounded in mean square. The following theorem provides sufficient condition.

Theorem 6: If $f(0) = g(0) = 0$ and if there exists a $2n \times 2n$ positive definite symmetric matrix Q, a constant $n \times m$ matrix B such that the $2n \times 2n$ matrix QP is uniformly negative semidefinite, where

$$P \triangleq \begin{bmatrix} \left(\nabla f + \nabla g + \frac{\alpha}{2} I\right)(w_s) & -\nabla g(w_s) \\ 0 & \left(\nabla f - B \nabla h + \frac{\alpha}{2} I\right)(w_s) \end{bmatrix}.$$

Then both the state x_t of (32) and the error $\epsilon_t = x_t - z_t$ are exponentially bounded in mean square with exponent α.

Next we consider the discrete system

$$x_{k+1} = f(x_k) + u_k + \sigma_1(x_k) w_k \tag{35}$$

$$y_k = h(x_k) + \sigma_2(x_k) v_k \tag{36}$$

where the unforced system of (35) is unstable and (35) and (36) satisfy the assumptions of (11) and (12).

To stabilize the system we assume that $u_k = g(z_k)$ where $g(\cdot)$ is continuously differentiable, so the observer for system (35) and (36) is given by

$$z_{k+1} = f(z_k) + g(z_k) + B[y_k - h(z_k)]. \tag{37}$$

The following theorem gives a sufficient condition so that both the state of (35) and the error $\epsilon_k = x_k - z_k$ will be exponentially bounded in mean square.

Theorem 7: If $f(0) = g(0) = 0$ and if there exists a constant $n \times m$ matrix B such that $\|P\| \leqslant \sqrt{1-\alpha}$ for all $w_s \in R^n$, where

$$P = \begin{bmatrix} \left(\nabla f + \nabla g + \frac{\alpha}{2} I\right)(w_s) & -\nabla g(w_s) \\ 0 & \left(\nabla f - B \nabla h + \frac{\alpha}{2} I\right)(w_s) \end{bmatrix},$$

then both the state x_k of (35) and the error $\epsilon_k = x_k - z_k$ are exponentially bounded in mean square with given exponent α.

Remark: The proofs of Theorem 6 and Theorem 7 are the applications of Corollary 1 and Corollary 2. They are essentially the same as given in the proofs of Theorem 3 and Theorem 4.

VI. Conclusions

In this study we used Lyapunov-like methods to obtain sufficient conditions required to design exponential observers for nonlinear stochastic systems (continuous and discrete). An observer is a dynamic system driven by the observed output of the system and designed in such a way that the error between the observer output and the system state is exponentially bounded in mean square. Therefore, we could use the output of the observer as an estimate of the state. It should be noted that there is no optimality requirement for designing the observer in this paper.

We also developed sufficient conditions so that a feedback control that is a function of the observer output (the estimated system state) will stabilize an otherwise unstable system.

There are several possible directions in which this research can be continued. Some of them are as follows.

1) Remove the constraint that the observer be linear with respect to the observed data and look for an observer of the form

$$dz_t = f(z_t)\,dt + B(z_t, y_t)[dy_t - h(z_t)\,dt]$$

for continuous systems, or

$$z_{k+1} = f(z_k) + B(z_k, y_k)[y_k - h(z_k)]$$

for discrete systems, where here B is an $n \times m$ matrix function instead of a constant matrix.

2) Develop an observer whose structure depends on the noise covariances $\sigma_1(x)\sigma_1'(x)$ and $\sigma_2(x)\sigma_2'(x)$ in addition to the system structure so that a lower mean-square error may be obtained.

Acknowledgment

The authors would like to thank Prof. D. L. Elliott for his suggestions and criticism. Many thanks are also due to one of the referees for his careful review and helpful comments.

References

[1] D. G. Luenberger, "Observing the state of a linear system," *IEEE Trans. Mil. Electron.*, vol. MIL-18, pp. 74–80, Apr. 1964.

[2] ——, "Observers for multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 190–197, Apr. 1966.

[3] J. J. Bongiarno, Jr., and D. C. Youla, "On observer in multivariable control systems," *Int. J. Contr.*, vol. 8, no. 3, pp. 221–243, 1968.

[4] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME* (J. Basic Eng.), ser. D, vol. 82, pp. 35–45, Mar. 1960.

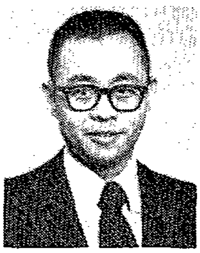
[5] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME* (J. Basic Eng.), ser. D, vol. 83, pp. 95–107, Dec. 1961.

[6] M. Aoki and J. R. Huddle, "Estimation of the state vector of a linear stochastic system with a constrained estimator," *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-12, pp. 432–433, Aug. 1967.

[7] M. Athans and E. Tse, "A direct derivation of the optimal linear filter using the maximum principle," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 690–698, Dec. 1967.

[8] E. Tse and M. Athans, "Optimal minimal-order observer-estimator for discrete linear time-varying systems," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 416–426, Aug. 1970.

[9] ——, "Observer theory for continuous time linear systems," *Inform. Contr.*, vol. 22, pp. 405–434, June 1973.
[10] E. Tse, "Observer-estimators for discrete time systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 10–16, Feb. 1973.
[11] R. L. Stratonovich, "Conditional Markov process theory," *Theory Prob. Appl.*, vol. 5, pp. 156–173, 1960.
[12] H. J. Kushner, "On differential equations satisfied by conditional probability densities of Markov processes," *SIAM J. Contr.*, vol. 2, pp. 106–119, 1964.
[13] W. M. Wonham, "Some applications of stochastic differential equations to optimal nonlinear filtering," *SIAM J. Contr.*, vol. 2, pp. 347–369, 1965.
[14] R. S. Bucy, "Nonlinear filtering," *IEEE Trans. Automat. Contr.*, vol. AC-10, p. 198, Apr. 1965.
[15] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*. New York: Interscience, 1968.
[16] P. A. Frost and T. K. Kailath, "An innovations approach to least-squares estimation—Part III: Nonlinear estimation in white Gaussian noise," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 217–226, June 1971.
[17] C. Hecht, "Digital realization of nonlinear filters," in *Proc. 2nd Symp. Nonlinear Estimation Theory and Its Applications*, 1971, pp. 152–158.
[18] R. S. Bucy and K. D. Senne, "Digital synthesis of nonlinear filters," *Automatica*, vol. 7, no. 3, pp. 287–298, 1971.
[19] J. B. Pearson, "On nonlinear least-squares filtering," *Automatica*, vol. 4, pp. 97–105, 1967.
[20] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970, ch. 9.
[21] J. R. Fisher, "Optimal nonlinear filtering," in *Advances in Control Theory*, vol. 5, C. T. Leondes, Ed. New York: Academic, pp. 197–200, 1967.
[22] K. Srinivasan, "State estimation by orthogonal expansion of probability distributions," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 3–10, Feb. 1970.
[23] H. W. Sorenson and A. R. Stubberud, "Nonlinear filtering by approximation of the a posteriori density," *Int. J. Contr.*, vol. 8, July 1968.
[24] W. Kizner, "Optimal nonlinear estimation based on orthogonal expansions," JPL Tech. Rep., pp. 32-1366, Apr. 1969.
[25] D. L. Alspach and H. W. Sorenson, "Approximation of density functions by a sum of Gaussians for nonlinear Bayesian estimation," in *Proc. 1st Symp. Nonlinear Estimation Theory and its Applications*, pp. 19–31.
[26] J. L. Center, "Practical nonlinear filtering of discrete observation," in *Proc. of 2nd Symp. Nonlinear Estimation Theory and Its Applications*, Sept. 13–15, 1971.
[27] R. S. Bucy and K. D. Senne, "Realization of optimum discrete-time nonlinear estimators," in *Proc. 1st Symp. Nonlinear Estimation Theory and its Applications*, pp. 6–17, 1970.
[28] W. M. Wonham, "Liapunov criteria for weak stochastic stability," *J. Differential Equations*, vol. 2, pp. 195–207, 1966.
[29] ——, "A Liapunov method for the estimation of statistical averages," *J. Differential Equations*, vol. 2, pp. 365–377, 1966.
[30] H. J. Kushner, *Introduction to Stochastic Control*. New York: Holt, Rinehart and Winston, 1971.
[31] ——, *Stochastic Stability and Control*. New York: Academic, 1967.
[32] M. Zakai, "On the ultimate boundedness of moments associated with solutions of stochastic differential equations," *SIAM J. Contr.*, vol. 5, no. 4, pp. 588–593, 1967.
[33] ——, "A Liapunov criterion for the existence of stationary probability distributions for systems perturbed by noise," *SIAM J. Contr.*, vol. 7, no. 3, pp. 390–397, 1969.
[34] Y. Miyahara, "Ultimate boundedness of the systems governed by stochastic differential equations," *Nagoya Math. J.*, vol. 47, pp. 111–144, 1972.
[35] E. Wong, *Stochastic Processes in Information and Dynamical Systems*. New York: McGraw-Hill, 1971.
[36] S. R. Kou, D. L. Elliott, and T. J. Tarn, "Exponential observers for nonlinear dynamic systems," *Inform. Contr.*, vol. 29, pp. 204–216, Nov. 1975.
[37] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*. New York: Academic, 1970.
[38] C. E. Langenhop, "On the stabilization of linear systems," *Proc. Amer. Math. Soc.*, vol. 15, no. 5, pp. 735–742, 1964.
[39] W. A. Wolovich, "On the stabilization of controllable systems," *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-13, pp. 569–572, Oct. 1968.
[40] V. G. Haussman, "Stability of linear systems with control dependent noise," *SIAM J. Contr.*, pp. 382–394, May 1973.



Tzyh-Jong Tarn (M'71) was born in Szechwan, China in 1937. He received the B.S. degree in chemical engineering from Taiwan Provincial Cheng Kung University in 1959, the M.S. degree in chemical engineering from Stevens Institute of Technology, Hoboken, NJ, in 1965, and the D.Sc. degree in control systems science and engineering from Washington University, St. Louis, MO, in 1968.

Since 1968 he has been with the Program of Control Systems Science and Engineering, now the Department of Systems Science and Mathematics, Washington University, where he was a Research Associate until 1969, an Assistant Professor from 1969 to 1972, and an Associate Professor since 1972. His research interests include filtering and estimation, optimal control, structure and realization of bilinear, and nonlinear systems.

Dr. Tarn is a member of the AIChE and Sigma Xi.

Yona Rasis received the B.S. and M.S. degrees in electrical engineering from the Technion—Israel Institute of Technology, Haifa, Israel, and the D.Sc. degree in control systems science and engineering from Washington University, St. Louis, MO, in 1963, 1969, and 1974, respectively.

From 1963 to 1966 he served in the Israeli Air Force and during the academic year 1969–1970 he was an Instructor at the Technion—Israel Institute of Technology. Since 1974 he has been with Sachs Electric Company, St. Louis, MO. His main interests are estimation and applications of modern control theory.