Stochastic observers for nonlinear systems

PRODIP SEN

School of Automation, Indian Institute of Science, Bangalore 560 012, India

Received on July 30, 1977; Revised on April 10, 1978

Abstract

Some results on asymptotic state estimation for a class of continuous time nonlinear stochastic systems are presented. The main property of the estimators developed is that the estimation error goes to zero with probability one and in mean-square as time tends to infinity. The proofs are based on Ito calculus and the martingale convergence theorems.

Key words: Stochastic systems, Asymptotic estimators, Ito calculus, Martingales.

Introduction

The theory of asymptotic state estimation or observer theory (references 7, 4, 5, and the references therein) has been concerned mainly with deterministic dynamic systems. The corresponding theory for stochastic systems seems to have been neglected. Some results in this direction have however been obtained for discrete-time systems[2]. The need for such a theory can be appreciated in the nonlinear estimation case where finite-dimensional realisations of the optimal (minimum error-variance) filters are unavailable barring a few cases, and the behaviour of sub-optimal filters is largely unknown.

In this paper we will be concerned with developing asymptotic state estimators for a class of continuous-time nonlinear probabilistic systems. These estimators will not be optimal in the usual sense but will have the property that the estimate will converge (with probability-one) to the true state as time tends to infinity. We will call such estimators observers (section 3 gives a precise definition), as an extension of the concept of observers from the deterministic theory.

Recently some work has appeared on "observers" for stochastic continuous-time systems[10, 11, 12]. It should be noted however that these "observers" are *not* asymptotic state estimators either with probability-one or in mean-square, nor do they have any optimal property—only the mean-square error remains bounded as $t \to \infty$. Our use of the words "stochastic observer" is *strictly in the sense of with-probability-one (and also possibly in mean-square) asymptotic state estimation* (see the definition in section 3).

We use the framework of Ito calculus. The basic motivation for the approach used here, is from some results obtained elsewhere[8] in continuous-time stochastic approximation, and we use essentially similar arguments.*

2. Preliminaries

In this section we describe the system to be considered and state some associated assumptions. These assumptions will hold throughout the paper unless otherwise mentioned.

The state and observation equations are

$$(\mathrm{NL})\quad \begin{cases} dx(w,t) = f(t, x(w,t))\,dt + g(t, u(w,t))\,dt + A_1(w,t)\,db(w,t) \\ dy(w,t) = h(t, x(w,t))\,dt + A_2(w,t)\,db(w,t) \end{cases}$$

where x, y are n-dimensional vectors, b is an R^p-valued Wiener process, A_1 and A_2 are appropriately dimensioned matrices whose elements are measurable stochastic processes. All stochastic processes are defined on a probability space $(\mathcal{W}, \mathcal{B}, P)$, with $w \varepsilon \mathcal{W}$, the probability space variable. There is an increasing family, $\{\mathcal{A}_t\}$, of sub σ-algebras of $\mathcal{B}$, such that $b(w,t)$ and the elements of $A_1(w,t)$, $A_2(w,t)$ are $\mathcal{A}_t$ measurable for all $t \geq 0$, and the σ-algebra generated by $\{b(w,t) - b(w,\tau), t \geq \tau \geq s\}$ is independent of $\mathcal{A}_s$ for all $s \geq 0$. Only separable versions[6] of all processes will be considered.

$f(t,x)$ satisfies the usual conditions for the existence and uniqueness of the solution of the state equation in (NL), *i.e.*, $f: [0, \infty) x R^n \rightarrow R^n$ is Borel measurable and

$$\| f(t, x_1) - f(t, x_2) \| \leq l_1(t) \| x_1 - x_2 \|, \text{ for } t \geq 0, x_1, x_2 \in R^n$$

and

$$\| f(t,x) \|^2 \leq l_2(t)(1 + \| x \|^2), \text{ for } t \geq 0, x \in R^n,$$

where $l_1(t)$ and $l_2(t)$ are locally bounded. $h(t,x)$ and $g(t,u)$ satisfy the same type of conditions. $u(w,t)$ is a $\{y(w,s), 0 \leq s \leq t\}$ measurable q-vector control, such that the solution of (NL) exists—*e.g.*, u could be a function of y which is locally Lipschitz. Further restrictions on f and h will be stated later wherever necessary.

The conditions on A_1 and A_2 are

(A) $\sup_{t \geq 0} E(\| A_1(w,t) \|^2 + \| A_2(w,t) \|^2) < \infty$, and

(B) $\int_0^\infty E(\| A_1(w,t) \|^2)\, dt < \infty$.

In the above and all that follows, the vector and matrix norms are $\| x \|^2 = x^T x$ and $\| B \|^2 = T_r(B^T B)$ respectively, where T_r = trace and superscript 'T' denotes transpose.

* It has been brought to the attention of the author, by a referee, that related results are available in a book by R. Z. Hasminsku: Stability of systems of Differential equations under random perturbations of parameters, in Russian, an English translation of which has recently been published by the American Mathematical Society.

E denotes the expectation operator on $(\mathcal{W}, \mathcal{B}, P)$ and $E(\cdot \mid \mathcal{A}_t)$ the conditional expectation given the σ-algebra a_t. I_n is the n-dimensional identity matrix.

All stochastic differential equations in this paper, such as (NL), are to be interpreted as corresponding integral equations. The integrals w.r.t. Wiener processes such as $\int A_1(w, s)\, db(w, s)$ are the usual Ito integrals[3]. In our development we will need integrals w.r.t. the observation process $y(w, t)$. Integrals of the form $\int_0^t B(s)\, dy(w, s)$ can be defined so that they are $\{y(w, s), 0 \leq s \leq t\}$ measurable and the following equation holds

$$\int_0^t B(s)\, dy(w, s) = \int_0^t B(s)\, h(s, x(w, s))\, ds + \int_0^t B(s) A_2(w, s)\, db(w, s), \qquad (1)$$

where B is a locally square-integrable matrix of compatible dimensions. The details of the definition and proofs of the above can be found in references 8, 9, 1. We will use equation (1) in our development wherever necessary without further comment.

We will use the abbreviation w.p. 1 for 'with probability one'.

3. Statement of the problem

We wish to develop under suitable conditions, a stochastic dynamic system whose state converges to the state of (NL) (w.p. 1 or in mean square) as time tends to infinity, based on the knowledge of the control u and the observation y. We have the following:

Definition : A stochastic dynamic system is said to be a *with probability one stochastic observer* (WP1SO) for the system (NL), if its state $\hat{x}(w, t)$ is measurable w.r.t. the σ-algebra generated by $\{y(w, s), 0 \leq s \leq t\}$, and w.p. 1 $\|\hat{x}(w, t) - x(w, t)\| \to 0$ as $t \to \infty$. It is said to be a *mean square stochastic observer* (*MSSO*) if $\hat{x}$ is y measurable and $E(\|\hat{x}(w, t) - x(w, t)\|^2) \to 0$ as $t \to \infty$.

The general form of the WP1SO and MSSO which we shall consider for (NL), is

$$\text{(ASE)} \quad \hat{x}(w, t) = x_0 + \int_0^t f(s, \hat{x}(w, s))\, ds + \int_0^t g(s, u(w, s))\, ds + \int_0^t a(s)\,[dy(w, s) - h(s, \hat{x}(w, s))\, ds]$$

where x_0 is an arbitrary n-vector and $a(t) \geq 0$ is a weighting function satisfying

$$\text{(i)} \int_0^\infty a^2(t)\, dt < \infty, \quad \text{(ii)} \int_0^\infty a(t)\, dt = \infty.$$

The estimate $\hat{x}$ is defined by eqn. (ASE), the existence of whose solution will be discussed below, and has no relation to the conditional mean or any other optimal or sub-optimal estimator.

Under the assumptions stated in Section 2, the existence and uniqueness of a w.p.1 continuous solution of (ASE) can be shown by a standard Picard iteration technique and we do not go into the details. Using the $\{y(w,s), 0 \leq s \leq t\}$ measurability of $\int_0^t a(s)\,dy(w,s)$ and $u(w,t)$ in the existence proof for (ASE), the y measurability of $\hat{x}$ can readily be shown. The local boundedness of $E(\|\hat{x}(w,t)\|^2)$ and $E(\|x(w,t)\|^2)$ are also obtained during the existence proofs. We will draw on these facts later as required.

4. Main results

In this section we state and prove the main results of this paper.

Theorem 1 : If f and h satisfy the following conditions

(C) $(x_1 - x_2)^T [f(t, x_1) - f(t, x_2)] \leq 0$, for all $x_1, x_2 \in R^n$ and $t \geq 0$,

(D) for every k > 0, there is a $b(k) > 0$ such that for all $t \geq 0$

$$\inf_{k \leq \|x_1 - x_2\| \leq k^{-1}} (x_1 - x_2)^T [h(t, x_1) - h(t, x_2)] \geq b(k)$$

(interchange k and k^{-1} if $k > 1$), then *(ASE) is a WP1SO for (NL).*

Theorem 2: If f satisfies condition (C) and h satisfies

(E) there is a $d > 0$ such that for all $t \geq 0$, $x_1, x_2 \in R^n$

$(x_1 - x_2)^T [h(t, x_1) - h(t, x_2)] \geq d \, \|x_1 - x_2\|^2$,

then (ASE) is a WP1SO and MSSO for (NL) (Note : (E) $\Rightarrow$ (D)).

To prove the theorems, we have to prove the y- measurability of $\hat{x}$ and the convergence of the estimation error $\hat{x}(w,t) - x(w,t)$ to zero w.p. 1 (theorems 1, 2) and in mean square (theorem 2), as $t \to \infty$. The remarks at the end of the previous section establish the measurability part. Hence the proofs given below will deal with the convergence assertions only.

If A_1 and A_2 are non-random matrices, conditions (A), (B) reduce to

$$\sup_{t \geq 0} (\|A_1(t)\| + \|A_2(t)\|) < \infty, \quad \int_0^\infty \|A_1(t)\|^2 dt < \infty.$$

In order to preserve the clarity of the proofs and avoid making them too long, we will *prove all our results in this context.*

Proof of Theorem 1 : Let $e(w,t) = \hat{x}(w,t) - x(w,t)$. To prove the theorem we have to show that w.p. 1 $\| e(w,t) \| \to 0$ as $t \to \infty$. We will show this by proving w.p. 1 $\| e(w,t) \|^2 \to Z(w)$ as $t \to \infty$, where Z is a finite random variable, and then showing $Z(w) = 0$ w.p. 1. We prove the convergence of $\| e(w,t) \|^2$ first.

Subtracting (NL) from (ASE)

$$\begin{aligned} de(w,t) &= [f(t, e(w,t) + x(w,t)) - f(t, x(w,t))]\,dt \\ &\quad + a(t)[h(t, x(w,t)) - h(t, x(w,t) + e(w,t))]\,dt \\ &\quad + [a(t) A_2(t) - A_1(t)]\,db(w,t). \end{aligned} \tag{2}$$

Applying Ito's differentiation rule[3] we obtain for all $t \geq T_1 \geq 0$

$$\begin{aligned} \| e(w,t) \|^2 &= \| e(w,T_1) \|^2 + 2 \int_{T_1}^{t} e^T(w,s)[f(s, e(w,s) + x(w,s)) \\ &\quad - f(s, x(w,s))]\,ds \\ &\quad + 2 \int_{T_1}^{t} a(s) e^T(w,s)[h(s, x(w,s)) - h(s, x(w,s) + e(w,s))]\,ds \\ &\quad + 2 \int_{T_1}^{t} [a(s) e^T(w,s) A_2(s) - e^T(w,s) A_1(s)]\,db(w,s) \\ &\quad + \int_{T_1}^{t} T_r\{(a(s) A_2(s) - A_1(s))(a(s) A_2(s) - A_1(s))^T\}\,ds. \end{aligned} \tag{3}$$

As discussed in section 3, $\hat{x}$ and x have locally bounded second moments, which implies that $E(\| e(w.t) \|^2)$ is locally bounded. Together with conditions (i), (A), (B) on a, A_1, A_2, (with A_1, A_2 non-random) this implies that the second moment of the integrand of the Ito integral term in eqn. (3) is locally integrable. This yields that the Ito integral term is a Martingale[3] *i.e.*, for all $t \geq T_1 \geq 0$

$$E\left\{ \int_{T_1}^{t} [a(s) e^T(w,s) A_2(s) - e^T(w,s) A_1(s)]\,db(w,s) \,\middle|\, \mathcal{A}_{T_1} \right\} = 0 \text{ w.p. 1.} \tag{4}$$

Taking conditional expectations of eqn. (3) and using (C), (D), eqn. (4) we have for $t \geq T_1 \geq 0$

$$E(\| e(w,t) \|^2 \mid \mathcal{A}_{T_1}) \leq \| e(w,T_1) \|^2 + \int_{T_1}^{t} R(s)\,ds \text{ w.p. 1,} \tag{5}$$

where $R(s)$ is a non-negative function depending on a, A_1, A_2, which is integrable on $[0, \infty)$ due to conditions (i), (A), (B).

Define $z(w,t) = \| e(w,t) \|^2 + \int_{t}^{\infty} R(s)\,ds$ (well defined and finite since R is integrable). Then using eqn. (5) and the non-negativity of R, it is easy to show that z is a non-negative supermartingale, and hence $z(w, t)$ converges w.p.1 to a finite-valued random variable as $t \to \infty$ (reference 6, p. 526 and noting that a supermartingale is a negative submartingale). Also $\int_0^\infty R(s)\,ds \to 0$ as $t \to \infty$ since R is integrable. The convergence of this term and of $z(w, t)$ immediately yields, from the definition of z:

$$\| e(w, t) \|^2 \to Z(w) \text{ w.p. 1 as } t \to \infty, \tag{6}$$

where $Z(w)$ is a finite valued random variable.

We now show $Z(w) = 0$ w.p. 1. Taking expectations of eqn. (3) with $T_1 = 0$, transposing the 2nd integral to the left, using condition (C), eqn. (4), we have

$$2E\{ \int_0^t a(s)\, e^T(w, s) [h(s, x(w, s) + e(w, s)) - h(s, x(w, s))]\, ds\}$$

$$\leq E(\| e(w, 0) \|^2) + \int_0^\infty R(s)\, ds = K_1 < \infty,$$

by the integrability of R and the finiteness of $E(\| x(w, 0) \|^2)$. Condition (D) shows the non-negativity of the integrand of the L.H.S. of the above inequality, which permits the application of Fubini's theorem to the above to yield

$$0 \leq \int_0^t a(s)\, E\{e^T(w, s) [h(s, x(w, s) + e(w, s))$$

$$- h(s, x(w, s))]\}\, ds \leq K, \text{ for } t \geq 0, \text{ where } K \text{ is finite.} \tag{7}$$

Since $\int_0^\infty a(t)\, dt = \infty$ by (ii), we immediately obtain from (7)

$$\liminf_{t \to \infty} E\{e^T(w, t) [h(t, x(w, t) + e(w, t)) - h(t, x(w, t))]\} = 0.$$

Using the definition of lim inf and the non-negativity of the term in braces [condition (D)] we have

$$e^T(w, t_i) [h(t_i, x(w, t_i) + e(w, t_i)) - h(t_i, x(w, t_i))] \to 0$$

in 1st mean as $i \uparrow \infty$, where $\{t_i\}$ is a sequence of real numbers increasing to infinity. Then (ref. 6 p. 164) a subsequence $\{t_{i_n}\}$ of $\{t_i\}$ can be extracted such that

$$\lim_{n \to \infty} e^T(w, t_{i_n}) [h(t_{i_n}, x(w, t_{i_n}) + e(w, t_{i_n}))$$

$$- h(t_{i_n}, x(w, t_{i_n}))] = 0, \text{ w.p. 1.} \tag{8}$$

Let $S = \{$intersection of the probability-one sets of (6) and (8)$\}$. Obviously $P(S) = 1$. Let $w_0 \in S$. Let if possible $Z(w_0) > 0$. Fix k with $0 < k^2 < \min(1, Z(w_0))$

Then using the convergence of $\| e(w_0, t) \|^2$ to $Z(w_0)$, there is a $T' < \infty$ such that for all $t \geq T'$, $k^{-1} \geq \| e(w_0, t) \| \geq k$. Condition (D) then yields

$$e^T(w_0, t)\,[h(t, x(w_0, t) + e(w_0, t)) - h(t, x(w_0, t))] \geq b(k) > 0,$$

for all $t \geq T'$, which contradicts eqn. (8) above. Hence $Z(w_0)$ is not greater than zero, which implies since Z is non-negative, that $Z(w_0) = 0$. Since w_0 was chosen to be an arbitrary element of a probability-one set we have $Z(w) = 0$ w.p. 1, which was to be shown.

Proof of Theorem 2 : Since condition (E) implies condition (D), theorem 1 holds and w.p. 1 convergence to zero is at hand. For the mean square convergence we proceed as follows. Taking expectations of eqn. (3), using conditions (C), (E), the martingale property of Ito integral term in eqn. (3) (see eqn. (4) of theorem 1), and the definition of R (see eqn. (5)), we have for all $t \geq T_1 \geq 0$

$$E(\| e(w, t) \|^2) \leq E(\| e(w, T_1) \|^2) + \int_{T_1}^{t} R(s)\,ds$$

$$- 2d \int_{T_1}^{t} a(s)\,E(\| e(w, s) \|^2)\,ds.$$

By the integrability of R (theorem 1) and the non-integrability of $a(\cdot)$ (condition (ii)) and $d > 0$ (condition (E)), we see that the proposition of the appendix holds in its entirety with the following identifications

$$m(t) = E(\| e(w, t) \|^2), f(t) = R(t), h(t) = 2d\,a(t),$$

and we therefore have $E(\| e(w, t) \|^2) \to 0$ as $t \to \infty$. Since the proposition of the appendix gives an explicit bound, we have rate of convergence information available in this case.

Remark 1 : The proofs for the general case where A_1, A_2 are random is a bit too long to be given here. The basic problem is that the martingale property of the Ito integral term in eqn. (3) is lost unless restrictions on higher (than 2nd) order moments of A_1, A_2 are imposed. To avoid this a truncated process argument is used. A sequence of stopping times[6] and a related sequence of stopped processes is defined in terms of the process $e(w, t)$. Equations and inequalities concerning these are developed as in the proofs presented in this paper, and later limits are taken. In this approach, general (nonlinear) dependence of A_1, A_2 on the state x can also be included, subject to the usual restrictions on nonlinearities in Ito equations. The details of all this and related results are available in the author's thesis[9].

Remark 2 : It should be noted that the conditions on the observation function h, as given by (D) or (E) can be relaxed somewhat. For example (E) can be replaced by

$$(x_1 - x_2)^T K(t)\,[h(t, x_1) - h(t, x_2)] \geq d \| x_1 - x_2 \|^2$$

for some bounded matrix $K(t)$. Then, in the observer equation (ASE), $a(t)$ has to be replaced by $a(t)\ K(t)$. Similarly for (D). The modifications required in the proof are trivial.

Remark 3 : In the case where f and h are both linear, a variation of the approach of this paper can be used to relax the conditions on the system which would otherwise be needed if the results of Theorems 1 or 2 are applied directly. The conditions on $F(t)$ and $H(t)$, where $f(t, x) = F(t)\,x$ and $h(t, x) = H(t)\,x$, are then

(i) The state transition matrix corresponding to $F(t) : \Phi_F(t, t_0)$ is bounded for $t \geq t_0 \geq 0$.

(ii) $(F(t), H(t))$ is a uniformly completely observable pair.

The conditions on the noise remain essentially the same. The details of the above (too long to be given here) will be presented elsewhere and are also available in the author's thesis[9].

5. Discussion

The restrictions on systems to which the results of this paper apply are of two types: conditions on the deterministic part, and conditions on the noise terms.

The most important restrictions for the deterministic part are those embodied in (C) and (D) ((E) is a special case of (D)). Condition (C) on the state eqn. means that the distance between two solutions of (NL), in the noise-less case, should remain in a bounded region depending on their initial distance. Notice that (C) is considerably weaker than the condition $(x_1 - x_2)^T [f(t, x_1) - f(t, x_2)] \leq -k \| x_1 - x_2 \|^2$, which would imply that all the solutions of (NL) tend to the same solution as $t \to \infty$ (in the noise free case). In fact in this case an observer which does not use the observations can easily be built ! Condition (D) (with the extensions of Remark 2) on the observation function can be interpreted as an observability condition. It implies that in a noise-less situation the observations due to different states are uniformly distinguishable. In the absence of a reasonably well-developed observability theory for nonlinear systems, this condition, which requires that the state and observation have the same dimension, seems unavoidable.

The major restriction on the noise terms is condition (B) which implies that the noise in the state-equation dies out in mean square sufficiently rapidly. This can be also seen to imply that the part of the covariance of $x(w, t)$ due to the noise, has bounded second moment. For example if $f = 0$, $g = 0$, (NL) becomes : $dx(w, t) = A_1(w, t)\, db(w, t)$, which implies $E\{\| x(w, t) \|^2\} = \int_0^t E\{\| A_1(w, s) \|^2\}\, ds$. Thus the boundedness of $E\{\| x(w, t) \|^2\}$ is just condition (B). This seems to be a reasonable assumption for studying the infinite time behaviour of stochastic systems. The other restriction on the noise, given by condition (A), is a mild assumption and is an extension of the usual local boundedness required when finite time behaviour is examined.

The weighting function $a(t)$ in the observer equation (ASE) is of free choice as long as it satisfies the general conditions specified in section 3. For example $\frac{k}{(t+1)^{\alpha}}$, $k > 0$, $\frac{1}{2} < \alpha \leq 1$, will do. In the general case of theorem 1, information regarding the rate of convergence is not available and hence selection of $a(t)$ with regard to this is not possible. In theorem 2, however, an estimate of the rate of convergence of the 2nd moment of the error is possible in terms of $a(t)$, as mentioned in the proof. Using this information an optimum choice of k and α in $k/(t+1)^{\alpha}$ is possible. Even here it is difficult to get a comparison of all possible $a(t)$ one may use. It should also be noted that the rate information obtained is only of the 2nd moment of the error—the rate of convergence of the sample functions is not obtained.

In the non-linear case, finite dimensional realisations of the optimal (minimum mean square-error) filter are very rare, and as such the optimal filter is unimplementable in most cases. We have developed here an estimator which is implementable and which, though not optimal in the minimum error-variance sense, has the nice property of convergence to the true state as time increases. Also the estimator does not require the knowledge of the noise coefficient matrices and hence will be useful in situations where these are unknown. The optimal filter, even when realisable, would require such knowledge.

Further we emphasize that we have specified w.p. 1 or sample function behaviour of our estimator—*almost all sample functions converge to the true state.* In the case of the minimum error-variance filter or of the observers of[10-12], sample function behaviour is *not* specified, and we point out that in any practical situation only sample functions of the stochastic processes involved are available.

As a last point, the minimal amount of computation required by the observer of this paper is obvious from eqn. (ASE). This is of course a characteristic property of any scheme originating from stochastic approximation theory whence these results arose.

6. Acknowledgements

The author wishes to thank Prof. M. R. Chidambara for helpful discussions.

Appendix

A result on integral equations, needed in the proof of theorem 2 is presented here. Although this result may be known there does not seem to be a convenient reference for it. We develop it under conditions suitable to our needs though more general formulations are possible.

Proposition : Let $m(t)$, $f(t)$ and $h(t)$ be non-negative locally bounded functions on $[T_1, \infty)$, $T_1 \geq 0$, satisfying

$$m(t) \leq m(s) + \int_s^t f(u)\,du - \int_s^t m(u)\,h(u)\,du, \quad t \geq s \geq T_1.$$

Then

$$m(t) \leq m(T_1)\exp\Big(-\int_{T_1}^t h(s)\,ds\Big) + \int_{T_1}^t f(s)\exp\Big(-\int_s^t h(u)\,du\Big)\,ds.$$

Moreover if $\int_{T_1}^{\infty} f(s)\,ds < \infty$ and $\int_{T_1}^{\infty} h(s)\,ds = \infty$, then $m(t) \to 0$ as $t \to \infty$.

Proof : It can be verified directly that a solution of the equation ($n(t) \geq 0$, locally bounded) :

$$n(t) = n(T_1) + \int_{T_1}^t f(s)\,ds - \int_{T_1}^t n(s)\,h(s)\,ds$$

is

$$n(t) = n(T_1)\exp\Big(-\int_{T_1}^t h(s)\,ds\Big) + \int_{T_1}^t f(s)\exp\Big(-\int_s^t h(u)\,du\Big)\,ds.$$

The local boundedness or non-negativity hypotheses justify the invocation of Fubini's theorem, which is required for this verification.

We take $n(T_1) = m(T_1)$, then to prove the 1st part of the proposition we have to show $m(t) \leq n(t)$ for all $t \geq T_1$. Define $z(t) = m(t) - n(t)$. Then by the hypothesis

$$z(t) \leq z(s) - \int_s^t z(u)\,h(u)\,du, \quad t \geq s \geq T_1. \tag{A.1}$$

Define $A = \{t \geq T_1 : z(t) \leq 0\}$. This is non-empty since $T_1 \in A$. For the 1st part of the proposition we then have to prove that $A = [T_1, \infty)$.

Define $B = \{t \geq T_1$: there is a sequence $\{s_n\}$, $s_n \in A$ such that $s_n \uparrow t\}$. It is obvious that

$$A \cup B \subset [T_1, \infty). \tag{A.2}$$

Let $t_0 \in B$, then by definition there is a sequence $\{s_n\}$ $s_n \in A, s_n \uparrow t_0$. Using these and (A.1) we obtain

$$z(t_0) \leq z(s_n) - \int_{s_n}^{t_0} z(u)\,h(u)\,du \leq -\int_{s_n}^{t_0} z(u)\,h(u)\,du \to 0$$

as $n \to \infty$. Thus $z(t_0) \leq 0$, *i.e.*,

$$B \subset A. \tag{A.3}$$

Let if possible there be $t_1 \in (T_1, \infty)$ and $t_1 \notin B$. Then by definition there is $t_2 \in (T_1, t_1)$ such that $A \cap (t_2, t_1) = \phi$ (null set). Define $t_3 = \sup\ (A \cap [0, t_2])$. Then there is a sequence $\{r_n\} \subset A \cap [0, t_2]$, $r_n \uparrow t_3$. Then from the definition of B we have $t_3 \in B$ and using (A.3) : $t_3 \in A$, *i.e.*,

$$z(t_3) \leq 0. \tag{A.4}$$

By definition of t_2, t_3

$$(t_3, t_1) \cap A = \phi. \tag{A.5}$$

Now let $t \in (t_3, t_1)$. By (A.1)

$$z(t) \leq z(t_3) - \int_{t_3}^{t} z(s)\, h(s)\, ds.$$

Using (A.4) and (A.5) and the non-negativity of h therefore $z(t) \leq 0$. But again by (A.5) since $t \in (t_3, t_1)$, $z(t) > 0$—a contradiction.
Hence t_1 cannot exist, and using $T_1 \in B$,

$$[T_1, \infty) \subset B. \tag{A.6}$$

From (A.2), (A.3), (A.6) we immediately have $A = [T_1, \infty)$, thus proving the 1st part of the proposition.

For the convergence assertion we use the bound of the 1st part of the proposition. The 1st term of the bound goes to zero as $t \to \infty$, since $\int_{T_1}^{\infty} h(t) = \infty$. Since f is integrable, Lebesgue's dominated convergence theorem can be applied to the 2nd term to. yield its convergence to zero.

References

1. BALAKRISHNAN, A. V. *Stochastic Differential Systems I*, Springer-Verlag, Berlin, 1973.
2. FUJITA, S. AND FUKAO, T. Convergence conditions of a dynamic stochastic approximation method for nonlinear stochastic discrete time dynamic systems, *IEEE Trans. Automatic Control*, AC-17, 1972, pp. 715–717.
3. GIKHMAN, I. I. AND SKOROKHOD, A. V. *Stochastic Differential Equations*, Springer-Verlag, Berlin, 1972.
4. IKEDA, M., MAEDA, H. AND KODAMA, S. Estimation and feedback in linear time-varying systems: A deterministic approach, *SIAM J. Control*, 1975, **13**, 304–326.
5. KOU, S. R., ELLIOTT, D. L. AND TARN, T. J. Exponential observers for nonlinear dynamic systems, *Information and Control*, 1975, **29**, 204–216.
6. LOEVE, M. *Probability Theory*, 3rd ed., Van Nostrand, Princeton, N. J., 1963.

7. LUENBERGER, D. G. — An introduction to observers, *IEEE Trans. Automatic Control*, AC-16, 1971, pp. 596-602.

8. SEN, P. AND ATHREYA, K.B. — On stochastic approximation procedures in continuous time, *Journal of the Indian Institute of Science*, 1978, 60 (A), 3, pp. 135-147.

9. SEN, P. — Stochastic observers for continuous time systems, *Ph.D. Dissertation*, Indian Institute of Science, Bangalore, May 1977.

10. SUNAHARA, Y., AIHARA, S. AND KISHINO, K. — Stochastic observability and controllability of nonlinear systems, *Int. J. Control*, 1975, **22** (1), 65-82.

11. SUNAHARA, Y., AIHARA, S. AND SHIRAWA, M. — Stochastic observability for noisy nonlinear stochastic systems, *Int. J. Control*, 1975, **22** (4), 461-480.

12. TARN T. J. AND RASIS Y. — Observers for nonlinear stochastic systems, *IEEE Trans. Automatic Control*, AC-21, Aug. 1976, pp. 441-448.