ON THE CONVERGENCE OF LION'S IDENTIFICATION

METHOD WITH RANDOM INPUTS.

by

Harold J. Kushner*

* This research was supported in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR Grant No. AF-AFSOR-693-66, in part by the National Science Foundation under Grant No. GK-967, and in part by the National Aeronautics and Space Administration under Grant No. NGR-40-002-015. Brown University

ABSTRACT

An interesting identification scheme for scalar input, scalar output linear systems, proposed by Lion in [1], is investigated for a wide class of random inputs. The inputs include the class of functions $u(t) = \Sigma k_i \bar{u}_i(t)$, where $\{\bar{u}_1(t), \ldots \bar{u}_k(t)\}$ is a Markov process which is asymptotically stationary. An invariant set theorem for random systems [2] is used to prove convergence (in probability) of the identification algorithm proposed in [1].

1. Introduction

Consider the scalar, asymptotically stable, reduced form, input-output representation

$$(1)\quad \left(\frac{d^n}{dt^n} + \sum_{i=o}^{n-1} a_i \frac{d^i}{dt^i}\right)y(t) = \left(\sum_{i=o}^{m} b_i \frac{d^i}{dt^i}\right)u(t),$$

where $m < n$, and u is the scalar input.

More precisely, there is a least dimension, asymptotically stable, system

$$(1a)\quad \dot{x} = Ax + Bu$$

$$y = Hx$$

which represents (1). In Laplace transform form

$$y(s) = \frac{N(s)}{D(s)}\, u(s) + \sum_{o}^{n-1} \frac{Q_i(s)}{D(s)}\; x_i(0)$$

where N,D are the obvious polynomials and N/D is in reduced form. Also

$$y(t) = y^o(t) + \hat{\delta}\,(t)$$

where $y^o(t)$ is the solution for $x(0) = 0$ and $\hat{\delta}(t) \to 0$ exponentially fast.

In an interesting paper [1], Lion investigated a versatile method for identifying the constant coefficients $\{a_i, b_i\}$. In a sense, the method is a type of model reference system, and Lion proved that the continually adjusted 'model' parameters $\{\alpha_i(t), \beta_i(t)\}$ converged to values $\{\alpha_i^o, \beta_i^o\}$ from which the $\{a_i, b_i\}$ could usually be computed, provided that the input was periodic and contained sufficiently many frequencies.

The periodicity requirement appeared, since a Liapunov function technique was used and an invariant set theorem appealed to. The latter theorem required periodicity. From the results of [1], one cannot assert convergence of the algorithm when the inputs are random. In this paper, an invariant set theorem for random systems [2] is applied to yield that Lion's algorithm converges for a very wide class of asymptotically stationary inputs. In fact, it seems likely that the class includes all inputs which are of practical interest, which are of the form $u(t) = \sum_1^s k_i \bar{u}_i(t)$, where† $\bar{u}(t) = (u_1(t),\ldots,u_s(t))'$ is an asymptotically stationary Markov process.

The result is interesting because the class of inputs is quite realistic, and important and because it illustrates the applicability of little known stochastic stability results to a very practical problem. Next a brief summary of Lion's method is given. In Section 3 the random inputs are described. Section 4 discusses stochastic invariant sets, and Section 5 states and proves the convergence theorem for the identification algorithm.

---

†The prime' denotes transpose.

2. Lion's Identification Scheme

The description of Lion's method is brief. It is included partially for purposes of self containment of the paper but we also require a more exact treatment, especially in regard to the effects of initial conditions, and transient terms. For more motivation and detail see [1].

Let H(s) be an asymptotically stable, finite order, rational transfer function, where order of denomenator minus order of numerator $\geq$ n. Define the functions† $y_k(t)$, $u_k(t)$, via the Laplace transform relations $y_k(s) = H(s)(s+c)^k y(s)$, $u_k(s) = H(s)(s+c)^k u(s)$. Intuitively, the ratios of the $y_k(s)$ (or $u_k(s)$) yield estimates of the derivatives. Introduce the system error functions (in complex domain and time domain, resp.)

$$\text{(2a)} \qquad \epsilon(s) \equiv y_n(s) + \sum_{0}^{n-1} \alpha_i y_i(s) + \sum_{0}^{m} \beta_i u_i(s)$$

$$\text{(2b)} \qquad \epsilon(t) \equiv y_n(t) + \sum_{0}^{n-1} \alpha_i y_i(t) + \sum_{0}^{m} \beta_i u_i(t).$$

It is easy to see that if all 'initial condition' are zero, there are $\alpha_i = \alpha_i = \alpha_i^o$, $\beta_i = \beta_i^o$, which do not depend on the input, and for which $\epsilon(t) \equiv 0$. In any case, for these coefficients $\{\alpha_i^o, \beta_i^o\}$ in (2b), $\epsilon(t) \to 0$ exponentially. The $\{a_i, b_i\}$ can usually be calculated from $\{\alpha_i^o, \beta_i^o\}$. See Lion [1] for a

---

† s always denotes the 'Laplace Transform' variable.

discussion and examples of this point. <u>Thus the identification problem is simply reduced to the calculation of</u> $\{\alpha_i^o, \beta_i^o\}$. In fact, placing the problem in this form was one of the nice aspects of [1]. The concern in this paper is solely with the calculation of $\{\alpha_i^o\ \beta_i^o\}$.

Next, choose any constant $k > 0$, and define the algorithm

$$
(3) \quad
\begin{aligned}
\dot{\alpha}_j &= -\frac{k}{2}\,[\partial \epsilon^2(t)/\partial \alpha_j] = -k\,\epsilon(t) y_j(t) \\
\dot{\beta}_j &= -\frac{k}{2}\,[\partial \epsilon^2(t)/\partial \beta_j] = -k\,\epsilon(t) u_j(t).
\end{aligned}
$$

Define the column rectors $z, w(t)$, and matrix $A(t)$,

$$z = (\alpha_o - \alpha_o^o, \ldots, \alpha_n - \alpha_n^o, \beta_o - \beta_o^o, \ldots, \beta_m - \beta_m^o)'$$

$$w(t) = (y_o(t), \ldots, y_{n-1}(t), u_o(t), \ldots, u_m(t))'.$$

$$A(t) = -kw(t)w'(t).$$

Let $w^o(t)$, $y_k^o(t)$, $u_k^o(t)$, etc., denote the quantities $w(t)$, $y_k(t)$, $u_k(t)$, etc., when the initial condition $X(0)$ in (1a) is zero. For the remainder of this section $u(t)$ is a periodic input. Then $A(t)$ is uniformly bounded.

(3) yields (4a)

$$
(4a) \quad
\begin{aligned}
\dot{z} &= -kw(t)w'(t)z - kw(t)[y_n(t) + \sum_0^{n-1} \alpha_i^o y_i(t) + \sum_0^m \beta_i^o u_i(t)] \\
&= A^o(t)z + \tilde{\delta}_t = A_p^o(t)z + [A_T^o(t)z + \tilde{\delta}_t]
\end{aligned}
$$

where $A_T^o(t)$ is the transient part of $A^o(t)$, and tends to zero exponentially. $A_p^o(t)$ is periodic.

Alternatively,

$$\text{(4b)} \qquad \dot{z} = -kw^o(t)\in(t) + \delta_t$$

$$\tilde{\delta}_t = [A(t) - A^o(t)]z - kw(t)[y_n(t) + \sum_o^{n-1} \alpha_i^o y_i(t) + \sum_o^m \beta_i^o u_i(t)]$$

$$\delta_t = -k\in(t)[w(t) - w^o(t)].$$

With Liapunov function $V(z) = z'z \equiv |z|^2$, we have[†]

$$\text{(5)} \qquad \frac{d}{dt}(z'z) = -2k(z'w)^2 - k(z'w)[y_n + w'\alpha^o]$$

$$\equiv -2k\ (z'w)^2 + \rho_t,$$

where $\alpha^o = (\alpha_o^o, \ldots \alpha_{n-1}^o, \beta_o^o, \ldots, \beta_m^o)'$. Then since, the $\{y_i(t), u_i(t)\}$ are uniformly bounded,

$$\frac{d}{dt}|z|^2 \leq K_2|z|$$

for some real K_2, which implies that $|z(t)| \leq K_o + K_1 t$. Since $[y_n(t) + w'(t)\alpha^o] \to 0$ exponentially, it is also true that ρ_t is uniformly bounded and is integrable. This, and the form (5), imply that $|z(t)|^2$ is uniformly bounded and that $z'(t)w(t) \to 0$. Also, by (5) and the non-negativity of $z'z$, $z'(t)w(t)$ is square

---

[†] Sometimes the t argument is dropped, for simplicity of writing.

integrable. In turn, this implies that† $\epsilon(t) = y_n(t) + \sum_0^{n-1} \alpha_i(t)y_i(t) + \sum_0^m \beta_i(t)u_i(t) \to 0$ and is square integrable (but not necessarily integrable). Finally, $\tilde{\delta}_t$ and δ_t and $A_T^o(t)z(t)$ are all $O(e^{-\lambda t})$ for some $\lambda > 0$, and are uniformly bounded.

Using this, the periodicity of $A_p^o(t)$, and the invariant set Theorem [3], we conclude that z(t) tends to the largest invariant set consistent with $\epsilon(t) \equiv 0$ and $\tilde{\delta}_t \equiv 0$. These†† conditions, put into (4a) and (4b), give $\dot{z} = 0$ or $z(t) \to \text{constant} \equiv z$, or $\alpha_i(t) \to \tilde{\alpha}_i$, $\beta_i(t) \to \tilde{\beta}_i$. Then we need only determine the constants satisfying

$$0 \equiv \epsilon(t) \equiv y_n^o(t) + \sum_0^{n-1} \tilde{\alpha}_i y_i^o(t) + \sum_0^{m} \tilde{\beta}_i u_i^o(t),$$

and the identification is complete.

---

† To see this, write $z'w + y_n - y_n = \epsilon(t) - [y_n + \Sigma\alpha_i^o y_i + \Sigma\beta_i^o u_i] \to 0$. Since the bracketed term $\to 0$ exponentially, $\epsilon(t) \to 0$.

†† The argument here is a little careless. The random case, of greater interest here, is more careful at this point. In fact, it is not required to use the invariant set theorems for periodic inputs. If the inputs u, and the $\{u_k, y_k\}$ are 'state variabilized' (as they can be since u is periodic), the ordinary invariant set theorem can be used. In fact, the result is true for any input u(t) which is the uniformly bounded and uniformly differentiable solution to a differential equation. This 'dynamical' aspect of the result also appears in the random case, where the inputs are related to a Markov process.

3. Identification with Random Inputs

In order to use state variable - or dynamical methods (such as invariant set theorems or Liapunov theorems) it is required that the input have a 'dynamical' structure. In particular, let $\bar{u}(t) = (\bar{u}_1(t),\ldots,\bar{u}_s(t))'$ be a vector Markov process. The random case analysis is almost the same as the deterministic case except for one crucial point. Assume

(A1) $\bar{u}(t)$ is continuous on the right w.p.1. (with probability one), has a stationary transition function and

$$E|\bar{u}(t)|^2 = M_o < \infty, \qquad \text{for some real } M_o.$$

The input is $\Sigma k_i \bar{u}_i(t) \equiv u(t)$.

(A2) $P\{|\bar{u}(t)| \leqq M_1 e^{rt} \text{ for all } t\} \to 1$ as $r \to \infty$, for some random M.

(A3) $P\{\bar{u}(t) \in A, \bar{u}(0) \in B\} \to P\{A\}P_o\{B\}$, as $t \to \infty$ for all A,B, and some ergodic measure $P\{.\}$.

(A4) If $f(x)$ is bounded and continuous, then so is $E_x f(\bar{u}(t))$ where E_x is the expectation given $\bar{u}(0) = x$.

(A5) $P_x\{|\bar{u}(t+h) - \bar{u}(t)| > \epsilon\} \overset{h}{\to} 0$ uniformly in x,t (for any fixed $\epsilon > 0$) for x,t in any finite region $(x = \bar{u}(0) =$ initial condition).

(A6) Let $S_u(\omega)$, the spectral density of $u(t)$, exist and be non-zero over some interval.

(A7) Both (1) and H(s) are asymptotically stable, (1) is completely controllable and observable.

Conditions (A1) - (A5) (except for (A3)) are satisified by any physical Markov process - of which some component may form an input to the system (1) - as far as the author is aware. In particular, any (Itô) system $\dot{x} = f(x) + \sigma(x)\xi$, with white noise ξ and f,σ Lipschitz satisfies (A1), (A2), (A4), (A5). (A3) is partially a technical condition - but it is not unreasonable. It says partly that the effects of the initial condition wears off - which is reasonable (in fact, convergence may not always take place without it). The asymptotic invariance part can be weakened at the expense of analytic difficulty , but such ergodic assumptions are quite common.

Next, we state variabilize the relevant quantities. There are constant, asymptotically stable, systems for which

$$\dot{\tilde{X}} = B\tilde{X} + C\bar{u},\ y(t) = \tilde{H}\tilde{X}(t)$$

$$\dot{X}^{yi} = B^{yi}X^{yi} + C^{yi}y(t),\ y_i(t) = H^{yi}X^{yi}(t) \tag{6}$$

$$\dot{X}^{ui} = B^{ui}X^{ui} + C^{ui}\bar{u},\ u_i(t) = H^{ui}X^{ui}(t).$$

Obviously $(\bar{u}(t), \tilde{X}(t), X^{yi}(t), X^{ui}(t), i = 0,\dots) \equiv \hat{X}(t)$ is Markov and satisfies (A1) - (A6). So does the Markov process $(X(t), X^{o}(t))$, where again, the superscript ^{o} implies that all initial conditions in (6) are zero.

There are random variables $\{\alpha_i^o, \beta_i^o\}$ for which (2) holds, and the random $\epsilon(t) \to 0$ exponentially (w.p.1.). (The exponent may depend on the random realization, but this is unimportant.) (4) also is meaningful. The superscript o implies again that all initial conditons are zero. In fact, the entire paragraph below (4) also holds w.p.1. with no change - except that λ is random and $\lambda > 0$ w.p.1.

We may write (4) as $\dot{z} = F(\hat{X})$, and observe, owing to (4) and (5) and the statements following (4), that (A1), (A2), (A4), (A5) are satisfied by the z(t) process also - and, in fact by the joint Markov $(\hat{X}(t), z(t)) \equiv X(t)$ process.

To complete the analysis, the following results on 'stochastic' invariant sets is required.

4. Invariant Sets for Stochastic Systems.

Deterministic Case

Let $\dot{x} = f(x)$ be a deterministic system and V a Liapunov function with $\dot{V} = -K(x) \leq 0$. Then, under other boundedness and smoothness conditions which we need not mention here, $x_t \to M \equiv \{x: K(x) = 0\}$. Furthermore, (invariant set theorem) let S = {collection of all points on all paths satisfying $K(y_t) \equiv 0$ <u>and</u> $\dot{y}_t = f(y_t)$, $\infty > t > -\infty$}. S is the largest invariant set contained in M and $x_t \to S$ as $t \to \infty$. The advantage of the invariant set theorem is that one can substitute the $K(y_t) \equiv 0$ <u>directly</u>

into the dynamics - and test the consequences. Otherwise, all the Liapunov theorem yields is that $K(x_t)$ tends to zero - and we must go through a (sometimes very difficult) limiting operation to determine the location of the asymptotic part of the path x_t. See [3], [4] for more details.

Stochastic Case

An analogous situation holds in the stochastic case. This article is not the place to dwell in detail on the stochastic invariant set theorem (see [2], or [5] for a more elementary proof for the case of discrete time processes with countably many states). (Note, also that the result in [2] is more general than indicated by the following argument). However, the following description should be helpful. Let X_t be a homogeneous Markov process which satisfies (A1), (A4) and (A5) for X_t replacing $\bar{u}(t)$. First, some definitions are given. Let ϕ_t be the probability measure of X_t; i.e., $\phi_t(A) = P\{X_t \in A\}$.

$\{\phi_t\}$ is said to be weakly bounded if for each $\epsilon > 0$, there is a compact set K_ϵ for which $\phi_t(K_\epsilon) \geq 1 - \epsilon$ for all $t \geq 0$. Clearly $\{\phi_t\}$ is weakly bounded if $E|X_t| \leq M_o < \infty$ for all t. If the measures of $\{Y_t\}$ are weakly bounded, and the matrix A is asymptotically stable, then clearly, the measures of $\{X_t, Y_t\}$, for $\dot{X}_t = AX_t + Y_t$, are weakly bounded. A set of probability measures M is an invariant set (for $\{X_t\}$) if it satisfies the following: Let ψ be a measure in M. Then there is a process X_t, $t \in (-\infty, \infty)$, with measures ψ_t so that $\psi_o = \psi$ and $\psi_t \in M$.

Thus, if $\psi \in M$, so is an entire trajectory of measures over $(-\infty,\infty)$. The stochastic counterpart of the deterministic invariant set is a set of measures, since it is the measures which have a semigroup property. Weak boundedness of the measures (implied by (A1)) replaces the boundedness condition on the state trajectory which is required for the deterministic problem.

Let ψ be a measure and B be the largest (measurable) set on which $\psi(B) = 0$. Then the complement of B is the support of ψ, $S(\psi)$. If $\{\psi_\alpha\} = G$ is a collection of measures, $S(G) \equiv \bigcup_\alpha S(\psi_\alpha)$.

The stochatic invariant set theorems says

Theorem 0. (A composit of results proved in [2].) Assume (A1), (A4), (A5) for X_t. Let $G(X_t) \to 0$ w.p.1. Then $X_t \to L$ in probability, where L is the support set of the largest invariant set of measures M consistent with $G(X_t) \equiv 0$ for $t \in (-\infty,\infty)$ (i.e., if $\psi \in M$, then ψ is concentrated on the set $\{X: G(X) = 0\}$) Let $\dot{V}(X_t) = -K(x) \leq 0$, where $X_t = x$, and $V(x) \geq 0$, for all x. Then $X_t \to L_1$, where L_1 is the support set of the largest invariant set of measures consistent with $K(X_t) \equiv 0$, $t \in (-\infty,\infty)$. Let $\dot{V}(X_t) = -K(x) + \rho(x)$, where $\rho(X_t)$ is bounded and integrable on $[0,\infty]$. Then the same conclusion holds.

5. The Problem Continued

The convergence result is contained in

Theorem 1. *Under* (A1) - (A7), $\alpha_i(t) \to \alpha_i^o$, $\beta_i(t) \to \beta_i^o$ *in probability* as $t \to \infty$, *where* $\{\alpha_i^o, \beta_i^o\}$ *are constant.*

Note. Recall that obtaining $\{\alpha_i^o, \beta_i^o\}$ yields the solution to the identification problem.

Proof. Let $V(x) = z'z$, where $X(t)$ is the composite Markov process and z is as previously defined. Then $\dot{V}(x) = -K(x) + \rho_t$, where ρ_t is given by (5), and, again $K(x) = -2k(z'w)^2$. Now, by the discussion immediately succeeding (5), $K(X(t)) \equiv 0$ implies that $\in(t) \equiv 0$. Next, the set of probability measures $\{\psi_t\}$ for the process $\{X(t)\}$ consistent with $\dot{\in}(t) \equiv 0$ will be calculated. In particular, what are the possible systems (and their probability measures)

$$\hat{X}(t),\ z(t)$$

consistent with

$$(9) \qquad 0 \equiv y_n(t) + \sum_{0}^{n-1} \alpha_i(t) y_i(t) + \sum_{0}^{m} \beta_i(t) u_i(t) = \in(t)$$

Recall that the probability law of the process X(t) is given by the law of $\bar{u}(t)$, and the relations (3), (6) which generate the other components of $\{\hat{X}(t), z(t)\}$. Next, some further implications of $\in(t) \equiv 0$ will be obtained. If (3) is to be consitent with $\in(t) \equiv 0$, then $z(t) = z$, a random variable. Thus, (9) can be replaced by (9')

$$(9') \qquad 0 = y_n(t) + \sum_{o}^{n-1} \tilde{\alpha}_i y_i(t) + \sum_{o}^{m} \tilde{\beta}_i u_i(t) \equiv \tilde{\epsilon}(t),$$

where $\{\tilde{\alpha}_i, \tilde{\beta}_i\}$ are random variables.

Next, what is an invariant set of measures M for $\{X(t)\}$, which is consistent with (9') for all $t \in (-\infty,\infty)$, and with $\alpha_i(t) = \alpha_i$, $\beta_i(t) = \tilde{\beta}_i$? The measures in M must also be consitent with (A3). The asymptotic stability of (1) and the systems (6) and condition (A3) imply that (A3) also holds if the collection $y_i(t)$, $u_i(t)$ replace $\bar{u}(t)$ in (A3). Using this observation and the fact that (9') and the trajectories of measures in M must be consitent with the observation for all $-\infty < t < \infty$, it is seen that the $\{\tilde{\alpha}_i, \tilde{\beta}_i\}$ may be taken to be independent of the $\{y_i(t), u_i(t)\}$. This is intuitively reasonable; since (9') must hold for all $t \in (-\infty,\infty)$, we can consider that it starts at $-2T < 0$. Then, for all $t \geq -T$, the $\{\tilde{\alpha}_i, \tilde{\beta}_i\}$ and $\{y_i(t), u_i(t)\}$ are 'almost' independent . Letting $T \to \infty$, gives the assertion.

Finally, write $R(\tau) = \lim_{\tau\to\infty} E[\hat{\epsilon}(t)\, \hat{\epsilon}(t+\tau) \mid \tilde{\alpha}_i, \tilde{\beta}_i]$
The Fourier tranform of $R(\tau)$ is

$$(10) \qquad \left|\sum_{o}^{n} \tilde{\alpha}_j H_j(i\omega)\; T(i\omega) + \sum_{o}^{m} \tilde{\beta}_j H_j(i\omega)\right|^2 \; S_u(\omega) = 0$$

where† $\tilde{\alpha}_n = 1$, $H_j(s) = (s+c)^j H(s)$ and $T(s)$ is the transfer function of (1). By (A6), (10) has a unique constant solution $\{\alpha_i^o, \beta_i^o\}$. Thus, the only values of the random variables $\tilde{\alpha}_i, \tilde{\beta}_i$, consistent with $\epsilon(t) \equiv 0$ and the other conditions, are constants. By the invariant set theorem, then, any measure ψ in M must correspond to $z(t)$ = constant = 0. Let $\hat{X}(t)$ in $X(t) = (\hat{X}(t), z(t))$ have dimension q. Then the support of any in measure ψ in M must be contained in $R^q x\{0,0,\ldots,0\}$, where the number of zeros are the dimension of $z(t)$. Thus $z(t) \to 0$ in probability.

Q.E.D.

---

† Let u_t be a second order stationary process, which is input to a system with output y_t and asymptotically stable transfer function $Q(s)$. Let $S_1(\omega)$ be the (asymptotic) spectral density of u_t. Then that of y_t is $|Q(i\omega)|^2 S_1(\omega)$.

CONCLUSIONS

An interesting identification scheme for scalar input, scalar output linear systems, proposed by Lion in [1], is investigated for a wide class of random inputs. The inputs include the class of functions $u(t) = \Sigma k_i \bar{u}_i(t)$, where $\{\bar{u}_1(t), \ldots \bar{u}_k(t)\}$ is a Markov process which is asymptotically stationary. An invariant set theorem for random systems [2] is used to prove convergence (in probability) of the identification algorithm proposed in [1].

REFERENCES

[1] P.M. Lion, "Rapid Identification of Linear and Nonlinear Systems", Proc. 1966 Joint Automatic Control Systems Conf., Univ. of Washington.

[2] H.J. Kushner, "The Concept of Invariant Set for Stochastic Dynamical Systems and Applications to Stochastic Stability", in Stochastic Optimization and Control, (Ed. H.F. Karreman), Wiley, 1968

[3] J.P. LaSalle, "Asymptotic Stability Criteria", in Proceedings of Symposia in Applied Math., Vol 13, Amer. Math. Soc.

[4] J.P. LaSalle, S. Lefschetz, Stability by Liapunovs Direct Method, Academic Press, 1961.

[5] H.J. Kushner, Stochastic Stability and Control, Academic Press, 1967.