



Persistence of Excitation Conditions for Partially Known Systems*

SOURA DASGUPTA,† BRIAN D. O. ANDERSON‡ and R. JOHN KAYE§

Key Words—Persistent excitation, sufficient richness, adaptive identification, linear systems, identifiability.

Abstract—This note derives persistence of excitation conditions on the input for systems with partially known transfer functions. The assumed transfer function model is $[p_0(s) + K^T P(s)]/[q_0(s) + K^T Q(s)]$, where K is the unknown parameter vector $p_0(s)$, $q_0(s)$ are known scalar polynomials and $P(s)$ and $Q(s)$ are known vector polynomials.

1. Introduction

THIS NOTE DERIVES persistence of excitation (p.e.) conditions for the uniform asymptotic convergence of a class of adaptive algorithms presented by Dasgupta *et al.* (1983, 1985, 1988) and Dasgupta (1988). The algorithms deal with systems having proper n th order asymptotically stable transfer functions of the form

$$T(s) = \frac{p_0(s) + K^T P(s)}{q_0(s) + K^T Q(s)} \quad (1)$$

where $K \triangleq [K_1, \dots, K_N]^T$, is the unknown parameter vector, $p_0(s)$ and $q_0(s)$ are known scalar polynomials and $P(s) \triangleq [p_1(s), \dots, p_N(s)]^T$ and $Q(s) \triangleq [q_1(s), \dots, q_N(s)]^T$ are known vectors of polynomials. In this note, no explicit assumptions are made on the degree of any of the polynomials appearing in (1). Nor are any requirements placed on whether or not any of them need to be monic. However, later in this note are stated certain assumptions which are necessary to ensure the existence of input functions which render K uniquely identifiable. These assumptions do in some cases translate to certain degree and monicity requirements.

Similar models also arise in the work of Bai and Sastry (1986). The conditions for uniform asymptotic convergence stated in the above references have been phrased jointly in terms of the input $u(t)$ and the output $y(t)$ of the unknown systems. The problem addressed here is that of translating these to ones involving the system output only.

The results are similar to those of Boyd and Sastry (1983), Narendra and Annaswamy (1988) and Dasgupta *et al.* (1990), which deal with transfer function models of the form

$$\frac{b_0 s^n + b_1 s^{n-1} + \dots + b_n}{s^n + a_1 s^{n-1} + \dots + a_n} \quad (2)$$

when all of the a_i and b_i are unknown. As opposed to (2), models such as (1), reflect greater *a priori* knowledge and are shown in this note to require simpler p.e. conditions.

The format of this note is the following. Section 2 states two excitation conditions involving both $u(t)$ and $y(t)$, stated in Dasgupta *et al.* (1985, 1988) and shows how they are equivalent. Section 3 translates these to input-only conditions. Section 4 discusses certain implications of the results in Section 3. Section 4 concludes.

2. Convergence conditions involving both input and output

For the model in (1), with $\alpha, \gamma > 0$, define

$$h_i(t) = \frac{p_i(s)u(t) - q_i(s)y(t)}{(s + \gamma)^n} \quad \forall i \in \{0, 1, \dots, N\}, \quad (3)$$

$$H(t) \triangleq [h_1(t), \dots, h_N(t)]^T, \quad (4)$$

$$R(t) \triangleq \int_0^t e^{-\alpha(t-\tau)} H(\tau) H^T(\tau) d\tau. \quad (5)$$

In this note s is treated as the differential operator and $h_i(t)$ are assumed generated with arbitrary but finite initial conditions. Then the two conditions required in Dasgupta *et al.* (1985, 1988) and Dasgupta (1988), for the uniform asymptotic convergence of certain parameter identification algorithms, are now stated.

Condition 1. There exists $\delta, \alpha_1, \alpha_2 > 0$ such that for all $t \geq 0$

$$\alpha_1 I \leq \int_t^{t+\delta} H(\tau) H^T(\tau) d\tau \leq \alpha_2 I. \quad (6)$$

Condition 2. There exist $\alpha_3, \alpha_4, \bar{t} > 0$ such that for some \bar{t} , and all $t > \bar{t}$

$$\alpha_3 I \leq R(t) \leq \alpha_4 I. \quad (7)$$

In the sequel $u(t)$ will be required to belong to the set $\Omega_\Delta[0, \infty)$ defined as follows: $u(t) \in \Omega_\Delta[0, \infty)$ if there exists a countable, possibly empty set $C_\Delta = \{t_j\}$, with $t_{j+1} - t_j \geq \Delta$, such that $u(t)$ and its time derivative are continuous and bounded on $R_+ - C_\Delta$, and possess finite 'one sided' limits at each t_j .

We then have the following theorem, a discrete time version of which is available in Johnstone and Anderson (1982).

Theorem 2.1. For the asymptotically stable system (1) with $u(t) \in \Omega_\Delta[0, \infty)$. Conditions 1 and 2 are equivalent.

Proof. Observe that by Lemma A.3 in the appendix all elements of $H(t)$ belong to $\Omega_\Delta[0, \infty)$. Thus, the equivalence of the upper bounds is immediate, and the sequel focusses on the lower bounds only. We first show that Condition 1 implies Condition 2. Notice, Condition 1 implies that for the given δ and α_1 and all $t \geq \delta$ and unit θ ,

$$\int_{t-\delta}^t \{\theta^T H(\tau)\}^2 d\tau \geq \alpha_1.$$

Thus,

$$\theta^T R \theta \geq \int_{t-\delta}^t e^{-\alpha(t-\tau)} \{\theta^T H(\tau)\}^2 d\tau \geq e^{-\alpha\delta} \alpha_1.$$

* Received 1 December 1992; received in final form 28 October 1993. Recommended for publication in revised form by Editor W. S. Levine.

† Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242, U.S.A.

‡ Department of Systems Engineering, Australian National University, ACT 2601, Australia.

§ School of Electrical Engineering and Computer Science, University of New South Wales, P.O. Box 1, Kensington, NSW 2033, Australia.

Thus, Condition 2 holds. Now, suppose Condition 1 is violated. Then for arbitrary $\delta, \epsilon > 0$ there exist unit θ and some t such that

$$\int_t^{t+\delta} (\theta^T \mathbf{H}(\tau))^2 d\tau \leq \epsilon^2. \tag{8}$$

Since $\theta^T \mathbf{H}(\tau)$ is in $\Omega_\Delta[0, \infty)$, by differentiating (8) with respect to t , Lemma A.1 yields,

$$|\theta^T \mathbf{H}(t)|^2 \leq O(\epsilon) \quad \forall \sigma \in [t, t + \delta]. \tag{9}$$

Choose δ to be such that $e^{-\alpha\delta} \leq \epsilon$. Then

$$\begin{aligned} \theta^T \mathbf{R}(t + \delta)\theta &= \int_0^{t+\delta} e^{-\alpha(t+\delta-\tau)} (\theta^T \mathbf{H}(\tau))^2 d\tau \\ &= \int_0^t e^{-\alpha(t+\delta-\tau)} (\theta^T \mathbf{H}(\tau))^2 d\tau \\ &\quad + \int_t^{t+\delta} e^{-\alpha(t+\delta-\tau)} (\theta^T \mathbf{H}(\tau))^2 d\tau \\ &\leq e^{-\alpha\delta} \int_0^t e^{-\alpha(t-\tau)} (\theta^T \mathbf{H}(\tau))^2 d\tau + O(\epsilon) \\ &\leq O(\epsilon), \end{aligned}$$

by the boundedness of $\mathbf{H}(t)$ and the fact that $e^{-\alpha\delta} \leq \epsilon$. Thus Condition 2 is violated, whence Condition 2 implies Condition 1.

▽▽▽

Remark. Notice the lower bound of (6) implies that of (7), regardless of whether $\mathbf{H}(t)$ is bounded.

3. Condition of the input

In this section we derive conditions on the system input which guarantee the satisfaction of (6) and hence, by Theorem 2.1, of (7). The proof of the main theorem relies on lemmas in the Appendix and on techniques developed in Dasgupta *et al.* (1990).

At the outset the following assumptions will be required.

Assumption 3.1. There exists no nonzero vector θ such that

$$T(s) = \frac{\theta^T \mathbf{P}(s)}{\theta^T \mathbf{Q}(s)}. \tag{10}$$

▽▽▽

Assumption 3.2. If for some vector θ

$$\theta^T \mathbf{P}(s) \equiv \theta^T \mathbf{Q}(s) \equiv 0, \tag{11}$$

then $\theta = 0$.

▽▽▽

As we now argue, Assumptions (10) and (11) ensure that only one choice of \mathbf{K} will satisfy (1). To see this for Assumption 3.1, suppose (10) holds for some nonzero θ . Then for all nonzero scalar d ,

$$\begin{aligned} T(s) &= \frac{p_0(s) + \mathbf{K}^T \mathbf{P}(s)}{q_0(s) + \mathbf{K}^T \mathbf{Q}(s)} = \frac{\theta^T \mathbf{P}(s)}{\theta^T \mathbf{Q}(s)} \\ &= \frac{p_0(s) + (\mathbf{K} + d\theta)^T \mathbf{P}(s)}{q_0(s) + (\mathbf{K} + d\theta)^T \mathbf{Q}(s)}. \end{aligned}$$

Thus an infinite number of parameter vectors describe the input output relationship inherent in (1) with equal precision. The satisfaction of (11) for some non zero θ similarly leads to the existence of multiple parameter values that describe (1). Further, it is easy to see that if two distinct choices of \mathbf{K} satisfy (1), then either Assumption 3.1 or Assumption 3.2 is violated. A direct consequence of these assumptions is a restriction on the dimension of \mathbf{K} , in particular that $\dim(\mathbf{K}) \leq n$, the system order.

Subject to the above assumptions, the following theorem can be proved.

Theorem 3.1. Consider the asymptotically stable n th order system with proper transfer function given in (1). Assume the input $u(t) \in \Omega_\Delta[0, \infty)$ and the Assumptions 3.1 and 3.2

hold. Define m as the maximum degree among the polynomials

$$\{p_i q_j - q_i p_j \mid i \in \{0, 1, \dots, N\}, j \in \{1, \dots, N\}\}.$$

Suppose for any $\bar{\gamma} > 0$

$$\mathbf{U}(t) \triangleq \left[u(t), \frac{u(t)}{s + \bar{\gamma}}, \dots, \frac{u(t)}{(s + \bar{\gamma})^m} \right]^T \tag{12}$$

and there exist $\alpha_s, \alpha_n, \delta' > 0$ such that for all $\sigma \in R_+$,

$$\alpha_s \mathbf{I} \leq \int_\sigma^{\sigma+\delta'} \mathbf{U}(t) \mathbf{U}(t)^T dt \leq \alpha_n \mathbf{I}. \tag{13}$$

Then with $\mathbf{H}(t)$ defined as in Section 2, there exist $\alpha_1, \alpha_2 > 0$ and $\delta > \delta'$ such that for all $\sigma \in R_+$,

$$\alpha_1 \mathbf{I} \leq \int_\sigma^{\sigma+\delta} \mathbf{H}(t) \mathbf{H}(t)^T dt \leq \alpha_2 \mathbf{I}. \tag{14}$$

Proof. Notice the boundedness of $u(t)$ and the asymptotic stability of (1) ensure that the upper bound of (13) implies that of (14). Suppose the lower bound of (14) is violated to obtain a proof by contradiction. Then as in Theorem 2.1 for arbitrary $\delta, \epsilon > 0$ there exists a unit θ and a σ such that

$$|\theta^T \mathbf{H}(t)| \leq \epsilon \quad \forall t \in [\sigma, \sigma + \delta].$$

Then, using (1), (3) and (4) we have that

$$\left| \sum_{i=1}^N \theta_i \frac{p_i(s)u(t) - q_i(s)y(t)}{(s + \gamma)^n} \right| \leq O(\epsilon) \quad \forall t \in [\sigma, \sigma + \delta].$$

By Lemma A.2 and the fact that $q_0(s) + \mathbf{K}^T \mathbf{Q}(s)$ has degree n , there exists $\delta_1 < \delta$ such that

$$\begin{aligned} &\left| \frac{(s + \gamma)^n (q_0(s) + \mathbf{K}^T \mathbf{Q}(s))}{(s + \bar{\gamma})^{2n}} \right. \\ &\quad \left. \times \sum_{i=1}^N \theta_i \frac{p_i(s)u(t) - q_i(s)T(s)u(t)}{(s + \gamma)^n} \right| \\ &\leq O(\epsilon) \quad \forall t \in [\sigma + \delta_1, \sigma + \delta], \end{aligned}$$

whence, from (1)

$$\begin{aligned} &\left| \sum_{i=1}^N \{ \theta_i [p_i(s) \{q_0(s) + \mathbf{K}^T \mathbf{Q}(s)\} \right. \\ &\quad \left. - q_i(s) \{p_0(s) + \mathbf{K}^T \mathbf{P}(s)\}] / (s + \gamma)^{2n} \} u(t) \right| \\ &\leq O(\epsilon) \quad \forall t \in [\sigma + \delta_1, \sigma + \delta], \end{aligned}$$

thus

$$\begin{aligned} &\left| \left\{ \left(\sum_{i=1}^N \theta_i p_i(s) \right) (q_0(s) + \mathbf{K}^T \mathbf{Q}(s)) \right. \right. \\ &\quad \left. \left. - \left(\sum_{i=1}^N \theta_i q_i(s) \right) (p_0(s) + \mathbf{K}^T \mathbf{P}(s)) \right\} / (s + \bar{\gamma})^{2n} \right\} u(t) \right| \\ &\leq O(\epsilon) \quad \forall t \in [\sigma + \delta_1, \sigma + \delta] \end{aligned}$$

and

$$\begin{aligned} &| \{ (\theta^T \mathbf{P}(s)) (q_0(s) + \mathbf{K}^T \mathbf{Q}(s)) \\ &\quad - (\theta^T \mathbf{Q}(s)) (p_0(s) + \mathbf{K}^T \mathbf{P}(s)) \} / (s + \gamma)^{2n} \} u(t) | \\ &\leq O(\epsilon) \quad \forall t \in [\sigma + \delta_1, \sigma + \delta]. \end{aligned}$$

Now

$$\theta^T \mathbf{P}(s) (q_0(s) + \mathbf{K}^T \mathbf{Q}(s)) \neq \theta^T \mathbf{Q}(s) (p_0(s) + \mathbf{K}^T \mathbf{P}(s))$$

as otherwise one of Assumption 3.1 or 3.2 will be violated. Thus by the definition of m

$$\theta^T \mathbf{P}(s) (q_0(s) + \mathbf{K}^T \mathbf{Q}(s)) - \theta^T \mathbf{Q}(s) (p_0(s) + \mathbf{K}^T \mathbf{P}(s))$$

is a nonzero polynomial. The degree can be checked to be at most m . Hence, there exists a nonzero vector $\xi \triangleq [\xi_0, \dots, \xi_m]^T$ such that

$$\left| \frac{\sum_{i=0}^m \xi_i (s + \bar{\gamma})^i}{(s + \bar{\gamma})^{2n}} u(t) \right| \leq O(\epsilon) \quad \forall t \in [\sigma + \delta_1, \sigma + \delta].$$

Then by Lemmas A.1 and A.3 and some positive ν

$$\left| \frac{(s + \bar{\nu})^{2n-m} \sum_{i=1}^m \xi_i (s + \bar{\nu})^i}{(s + \bar{\nu})^{2n}} u(t) \right| \leq O(\epsilon^\nu) \forall t \in [\sigma + \delta_1, \sigma + \delta],$$

whence

$$|\xi^T U(t)| \leq O(\epsilon^\nu) \forall t \in [\sigma + \delta_1, \sigma + \delta].$$

Then it can be seen that (13) is violated, whence (13) implies (14).

▽▽▽

Note that the definition of m ensures that $m \leq 2n$. For $u(t)$ a linear combination of sinusoids, (13) requires that $u(t)$ have at least $(m + 1)/2$ distinct frequency components. As opposed to this, for (2) the corresponding number is $(2n + 1)/2$.

4. Consolidation

In this section we make certain observations which provide insights into the nature of the result in Theorem 3.1. The fact that $u(t)$ must have $(m + 1)/2$ distinct frequencies to guarantee p.e. requires further explanation. Essentially, a p.e. input is one which ensures that the unknown parameters are uniquely identifiable even after the decay of initial condition induced transients. Now, for an N parameter system to be identifiable, the input must generate N linearly independent equations in the unknown parameters. Given that there are two pieces of information contained in each distinct frequency component, namely its magnitude and phase, on the face of it, a linear combination of $N/2$ sinusoids should suffice for the parameters to be uniquely identifiable ($1/2$ a frequency is a constant component). Yet, m is in general greater than N , and thus our result appears to demand more than what intuition suggests it should.

To reconcile this apparent conservatism in the result of Theorem 3.1, suppose in (1) the number of unknown parameters is four, that the input is a linear combination of sinusoids and thus infinitely differentiable, but that it is such that two different parameter vectors K_1 and K_2 produce precisely the same output. Then simple calculations reveal that for this input and choice of K_1 and K_2 , with $G(s, K_1, K_2)$ defined as the operator below,

$$(p_0(s) + K_1^T P(s))(q_0(s) + K_2^T Q(s)) - (q_0(s) + K_1^T Q(s))(p_0(s) + K_2^T P(s))$$

the following holds:

$$G(s, K_1, K_2)u(t) = 0. \tag{15}$$

It turns out that for every K_1 , and every pair of conjugate imaginary numbers there exists a K_2 such that the operator multiplying $u(t)$ in (15) has zeros at this pair. Equally, although for every K_1 , there may exist K_2 different from K_1 such that this operator has as many as m zeros on the imaginary axis, only one arbitrary conjugate pair of imaginary numbers can be simultaneously chosen as the zeros of this operator by the suitable selection of K_2 . Thus, as long as the sinusoidal input $u(t)$ has two sinusoidal frequency components, then for generic values of these frequencies, the signal in (15) will not be zero, whence no K_2 different from K_1 will describe the input output relationship. On the other hand, without a priori knowledge of the parameter values, one cannot be certain that the imaginary axis zeros have been avoided and that the two frequencies in $u(t)$ will not coincide with the imaginary axis zeros of the operator. Thus, one cannot preclude the possibility that the two frequency sinusoidal component $u(t)$ will not be annulled completely by the operator in (15) for some value of K_2 distinct from K_1 . On the other hand, should the input have $(m + 1)/2$ distinct frequencies, and be real, one can be certain that for arbitrary K_2 distinct from K_1 , the signal in (15) is nonzero. Thus, the only way to guarantee p.e. without further prior knowledge is to choose the input signal in the manner dictated by Theorem 3.1.

Now suppose the parameter vector in (1) is known to have nonlinear interdependence between its elements. For

example, as in Dasgupta *et al.* (1983, 1985, 1988), suppose the dimension of K_1 is three and that its third element is the product of the first two. Then local, though not global, identifiability of the parameters would require the input to be capable of generating just two, rather than three independent equations. (By local identifiability is implied the fact that, though more than one parameter value may describe the input-output relationship equally well, these values are isolated and do not lie on a continuum.) Then pursuing an argument similar to that set out in the foregoing, one can show that one generic frequency components would result in local identification of (1). However, the same argument also reveals, that as many as $m/2$ nongeneric frequency components may yet be annulled by the operator in (15) for suitable K_2 , distinct from K_1 , despite K_2 being constrained to be on the manifold where the third element equals the product of the first two. Thus, even for local identifiability, without further prior knowledge, one requires as complex an input as needed to achieve global identifiability.

5. Conclusion

We have established persistence of excitation conditions for adaptive algorithms for parameterizations which utilize partial knowledge about the unknown system.

References

Bai, E. W. and S. S. Sastry (1986). Parameter identification using prior information. *Int. J. of Control*, **44**, 455-473.
 Boyd, S. and S. S. Sastry (1983). On parameter convergence in adaptive control. *Systems and Control Letters*, **3**, 311-319.
 Dasgupta, S., B. D. O. Anderson and R. J. Kaye (1983). Robust identification of partially known systems. *Proc. 22nd CDC San Antonio, Texas*, pp. 1510-1514.
 Dasgupta, S., B. D. O. Anderson and R. J. Kaye (1985). Output error identification methods for partially known systems. *Int. J. of Control*, **43**, 177-191.
 Dasgupta, S., B. D. O. Anderson and R. J. Kaye (1988). Identification of physical parameters in structured systems. *Automatica*, **24**, 217-225.
 Dasgupta, S., B. D. O. Anderson and A. C. Tsoi (1990). Input only conditions for continuous time adaptive systems problems. *IEEE Trans. on Auto.*, **35**, 78-82.
 Dasgupta, S. (1988). Adaptive identification of systems with polynomial parameterizations. *IEEE Trans. on Circuits and Systems*, **35**, 593-603.
 Johnstone, R. M. and B. D. O. Anderson (1982). Exponential convergence of recursive least squares with exponential forgetting factor—adaptive control. *Systems and Control Letters*, **2**, 69-76.
 Mitrinovic, D. S. (1970). *Analytic Inequalities*. Springer-Verlag, Berlin.
 Narendra, K. S. and A. M. Annaswamy (1988). *Stable Adaptive Systems*. Prentice-Hall, Englewood Cliffs, NJ.

Appendix

The proof of the theorems depend on the following lemmas, the first of which has been obtained from Mitrinovic (1970).

Lemma A.1. If $f(\cdot)$ is an n times differentiable function on an interval I of length Δ and if $|f(x)| \leq M_0$ and $|f^{(n)}(x)| \leq M_n$, then for $x \in I$ and for $0 < k < n$

$$|f^{(k)}(x)| \leq 4e^{2k} ({}^n C_k)^k M_0^{1-k/n} M_n^{k/n}, \tag{A.1}$$

where $M_n = \max(M_n, M_0 n! \Delta^{-n})$.

▽▽▽

Lemma A.2. For any asymptotically stable system with a proper transfer function $W(s)$, if the input $u(t)$ is such that

$$\begin{aligned} |u(t)| &\leq M \text{ on } [0, T] \\ |u(t)| &\leq \epsilon \text{ for all } t > T, \end{aligned} \tag{A.2}$$

and if the initial state lies in some fixed ball B of radius R then there exists a $\delta_1(\epsilon, M, R)$, independent of T , such that for $t > \delta_1 + T$, the output $y(t)$ satisfies $|y(t)| < O(\epsilon)$.

Proof. For any minimal realization F, G, H, J of $W(s)$, asymptotic stability and (A.2) imply the existence of a $K(M, R)$ such that $\|x(t)\| < K, \forall t \leq T$, with $x(t)$ the state vector.

Then for $t > T$

$$\|y(t)\| \leq K \|e^{F(t-T)}\| \|H\| + O(\epsilon).$$

Thus if δ_1 is selected to make $K \|e^{F\delta_1}\| \|H\| < \epsilon$ it follows that for $t > \delta_1 + T$, $\|y(t)\| < O(\epsilon)$.

Lemma A.3. If $u(t) \in \Omega_\Delta[0, \infty)$, then under the assumption of arbitrary finite initial conditions, for any Hurwitz polynomial $D(s)$ and polynomials $N_1(s)$ and $N_2(s)$, such that $\delta[N_1(s)] \leq \delta[D(s)]$ and $\delta[N_2(s)] = 1 + \delta[D(s)]$, the following

properties hold:

$$(i) \left\{ \frac{N_1(s)}{D(s)} \right\} u \in \Omega_\Delta[0, \infty);$$

and

$$(ii) \left\{ \frac{N_2(s)}{D(s)} \right\} \text{ is continuous and bounded on } \{[0, \infty) - C_\Delta\},$$

and has finite limits as $t \downarrow t_i$ and $t \uparrow t_i$, $t_i \in C_\Delta$.

Proof. Trivial. VVV