PERSISTENCY OF EXCITATION, SUFFICIENT RICHNESS AND PARAMETER CONVERGENCE IN DISCRETE TIME ADAPTIVE CONTROL

by

E. W. Bai, and S. S. Sastry

Memorandum No. UCB/ERL M84/91

5 November 1984

PERSISTENCY OF EXCITATION, SUFFICIENT RICHNESS AND PARAMETER CONVERGENCE IN DISCRETE TIME ADAPTIVE CONTROL

by

E. W. Bai, and S. S. Sastry

Memorandum No. UCB/ERL M84/91

5 November 1984

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Persistency of Excitation, Sufficient Richness and Parameter Convergence in Discrete Time Adaptive Control

E.W.Bai[*] and *S.S.Sastry*[*]

Department of Electrical Engineering and Computer Science
Electronics Research Laboratory
University of California, Berkeley,CA 94720

ABSTRACT

The main result of this paper is to give new techniques for deriving explicit frequency domain condition on the exogenously reference trajectory to guarantee parameter convergence in a class of discrete time adaptive control schemes. Other interesting results are relations between persistently exciting vector signal and sufficiently rich scalar signal.

Oct. 29,1984

[*] Research supported in part by IBM Corporation under a Faculty Development Award 83-85. Authors would also like to thank Profs B.D.O.Anderson, K.J.Astrom, M.Bodson and T.Salcudean for several useful discussions.

Persistency of Excitation, Sufficient Richness and Parameter Convergence in Discrete Time Adaptive Control

*E.W.Bai** and *S.S.Sastry**

Department of Electrical Engineering and Computer Science
Electronics Research Laboratory
University of California, Berkeley,CA 94720

1. Introduction

In recent years there has been widespread concern about the sensitivity of adaptive schemes to unmodeled dynamics, output disturbance and the like (see for eg [1]). We feel that one of the important reason for this concern has been a lack of understanding of condition under which the adaptive schemes are genuinely convergent (that is, their parameter estimators converge in addition to the equation error.), it has been shown by us for the continuous time case in [9]. Motivated by results of Anderson and Johnson for the discrete time case in [4], that adaptive schemes that are exponentially convergent (both parameter error and equation error) are robust to the presence of unmodeled dynamics and output disturbance. In the continuous time model reference adaptive system, we have characterized the necessary and sufficient conditions for parameter convergence in terms of spectral content (sufficient richness) of the exogenous reference input. In this paper, we pursue this idea further and establish condition (similar to those of [4]) on the exogenous specified reference trajectory which guarantees parameter convergence for discrete time adaptive schemes of the type proposed by Goodwin,Ramadge and Caines [5,6].(henceforth abbreviated G.R.C.)

During the course of developing the results, we develop an interesting technique for studying the spectral content of possibly unbounded discrete time signal (arising from unstable plant) which yields a sharper result than a corresponding one in [6,lemma 3.4.9 pp 78, and 10]. Our results are similar to those given in Anderson and Johnson [4], however, we feel that our proof and techniques are independently useful and insightful. Our main result is easily stated: If the support of the spectral measure of the specified reference trajectory has more than N points, where N is the number of unknown parameters in the G.R.C. scheme, then the parameters are exponentially convergent. While we have chosen in this paper to focus on G.R.C. scheme with the modified projection parameter update law is for brevity alone. The arguments of this paper are easily modified for several other common discrete time adaptive schemes. An example of such schemes is the discrete time model reference scheme [6, pp 199]. For other parameter update laws, such as the least square scheme, it is also easy to verify that conditions such as those derived in this paper guarantee parameter convergence (though it may not be exponential).

* Research supported in part by IBM Corporation under a Faculty Development Award 83-85. Authors would also like to thank Profs B.D.O.Anderson, K.J.Astrom, M.Bodson and T.Salcudean for several useful discussions.

Oct. 29,1984

The paper is organized as follows: we discuss in detail two important concepts -persistent excitation of vector signals and sufficient richness of scalar signals in section 2. In section 2, we also show that sufficient richness of the input to a discrete time (not necessarily stable) linear system implies persistent excitation of the state vector. In section 3, we apply this machinery to prove parameter convergence for the discrete time adaptive scheme of G.R.C. with modified projection parameter update law.

2. Persistency of excitation (PE) and sufficient richness (SR)

The terms of persistency of excitation and sufficient richness have been widely used interchangably in the literature. We propose to make a distinction between them.

Definition 1. A scalar sequence u(t) is said to be *sufficiently rich* (SR) of order n, if there exists $N \in Z_+, \alpha_1 > \alpha_2 > 0$ such that

$$\alpha_1 I \geq \sum_{t=t_0+1}^{t_0+N} \begin{bmatrix} u(t+1) \\ u(t+2) \\ \cdot \\ \cdot \\ u(t+n) \end{bmatrix} [u(t+1), u(t+2), \ldots, u(t+n)] \geq \alpha_2 I$$

uniformly in t_0 .

Definition 2. A vector sequence x(t) $\in R^n$ is said to be *persistently exciting* (PE), if there exists $N \in Z_+$, $\alpha_1 > \alpha_2 > 0$ such that

$$\alpha_1 I \geq \sum_{t=t_0+1}^{t_0+N} x(t) x^T(t) \geq \alpha_2 I$$

uniformly in t_0.

Remarks:

1. The SR condition has an interesting interpretation in the frequency domain. Roughly speaking, if u(t) has at least n spectral lines, then u(t) is SR of order n. We will discuss this further in lemma 2.

2. let U(t)=(u(t+1),.......u(t+n) $)^T$, then u(t) is SR of order n if and only if U(t) is PE.

For the continuous case, it has been shown [3] that the PE condition is very closely related to the autocovariance of the signal, a concept reminiscent of the theory of stationary stochastic process.

Consider the following definition and lemma:

Definition 3. A sequence x(t) $\in R^n$ is said to have *autocovariance* $R_x(k) \in R^{n \times n}$ if and only if

$$\lim_{M \to \infty} \frac{1}{M} \sum_{t=t_0+1}^{t_0+M} x(t) x^T(t+k) = R_x(k)$$

uniformly in t_0 .

Lemma 1.(Characterisation of PE)

Suppose a vector sequence x(t) $\in R^n$ is bounded and has autocovariance $R_x(k)$, then x(t) is PE if and only if $R_x(0) > 0$.

We will not give the proof, since it is the straightforward modification of a similar continuous time result in [3].

Remarks:
The PE condition makes precisely the intuition that R^n can be spanned by x(t) uniformly in N steps when x(t) is PE.

We also have a frequency domain interpretation for PE. Assume that the following limit exists uniformly in t_0

$$\lim_{M\to\infty}\frac{1}{M}\sum_{t=t_0+1}^{t_0+M} x(t)e^{-j\nu t}=X(j\nu) \qquad ,\nu\in(-\pi,\pi)$$

Then, we say x(t) has a *spectral line* at ν with amplitude $X(j\nu)\in C^n$. , it may be proven using techniques similar to [3], if x(t) has spectral lines at frequencies $\nu_1\, \nu_2,\ldots,\nu_n$ with amplitude $X(j\nu_1),X(j\nu_2),\ldots\ldots,X(j\nu_n)$ which are linearly independent, then x(t) is PE.

Lemma 2.(Characterisation of SR in frequency domain)
If a scalar sequence u(t) is bounded and has an autocovariance ,then u(t) is SR of order n if and only if the spectral measure of u(t) is not concentrated on k<n points.

Proof: Let

$$U(t)=(u(t+1),\ldots\ldots,u(t+n))^T$$

Now since u(t) has autocovariance, we may define

$$R_U(0)=\lim_{M\to\infty}\frac{1}{M}\sum_{t=t_0+1}^{t_0+M} U(t)U^T(t)=\begin{bmatrix} R_u(0) & R_u(n-1) \\ \cdot & \cdot \\ R_u(-n+1) & R_u(0)\end{bmatrix}$$

with

$$R_u(k)=\lim_{M\to\infty}\frac{1}{M}\sum_{t=t_0+1}^{t_0+M} u(t)u(t+k)$$

It is well-known [7,8] that R_u (k) is a positive semidefinite function. Further from the Herglotz theorem a function $R_u(k)$ defined on the integers is positive semidefinite if and only if

$$R_u(k)=\int_{-\pi}^{\pi} e^{jkw}S_u(dw)$$

where $S_u(w)$ is positive bounded measure.
From lemma 1., u(t) is SR of order n if and only if $R_U(0)$ >0. with

$$R_U(0)=\int_{-\pi}^{\pi}\begin{bmatrix} 1 \\ e^{-jw} \\ \cdot \\ e^{-j(n-1)w}\end{bmatrix}[1,e^{jw},\ldots\ldots,e^{j(n-1)w}]S_u(dw)$$

That is, for any nonzero λ, the following inequality holds

$$\lambda^T R_y(0)\lambda=\int_{-\pi}^{\pi}|\lambda^T W|^2 S_u(dw) >0 \qquad (2.1)$$

where

$$W=(1,e^{-jw},.....,e^{-j(n-1)w})^T$$

Now it is easy to show that the inequality (2.1) holds , if and only if $S_u(w)$ is nonzero at least at n points, since

$$\lambda^T W=\lambda_1+\lambda_2 e^{-jw}+,.....,\lambda_n e^{-j(n-1)w}=0$$

has arbitrary but at most (n-1) roots.

Q.E.D.

In the adaptive tracking problem, the design objective is to make the output of the unknown system track a given reference trajectory. It seems to be reasonable to guess that if reference trajectory is SR, so is the system output. This is precisely what the following lemma does.

Lemma 3. Suppose there are two scalar sequences $y_1(t)$ and $y_2(t)$ satisfying

$$|y_i(t)|<M \qquad i=1,2$$

$$\sum_{t=1}^{\infty}|y_1(t)-y_2(t)|^2<\infty$$

then $y_1(t)$ is SR of order n if and only if $y_2(t)$ is SR of order n.
The proof proceeds by arguments similar to the continuous case [3].

Theorem 1.(Reachability and PE)*
Consider the discrete time system

$$X(t+1)=AX(t)+bu(t) \qquad (2.2)$$

$$t\in Z_+ \;\; X(t)\in R^n \;\; A\in R^{n\times n} \;\; b\in R^{n\times 1} \;\; u(t)\in R$$

Assume that (2.2) is completely reachable, and the input u(t) is SR of order n, then there exists N ∈ Z_+ and α >0 such that

$$\sum_{t=t_0+1}^{t_0+N} X(t)X^T(t) \geq \alpha I$$

uniformly in t_0.(α is independent of the initial condition of (2.2).)

Proof: Note that

$$X(t+j)=A^j X(t)+\sum_{i=1}^{j} A^{j-i} bu(t+i-1) \qquad (2.3)$$

Since the eigenvalues of A are not inside the unit disc ,it is easy to see that the difficulty is the initial condition term. We deal with this as follows, define the characteristic polynomial of A.

$$p(z)=z^n+a_1 z^{n-1}+.....+a_n$$

*The idea for the proof of this theorem was given to us by T.Salcudean. (unpublished notes-to appear)

Now define

$$V(t)=X(t+n)+a_1X(t+n-1)+.....+a_nX(t)$$

We have

$$V(t)=A^nX(t)+\sum_{i=1}^{n}A^{n-i}bu(t+i-1)$$

$$+a_1A^{n-1}X(t)+\sum_{i=1}^{n-1}a_1A^{n-1-i}bu(t+i-1)+........$$

$$+a_{n-1}AX(t)+a_{n-1}bu(t)+a_nX(t)$$

From the Cayley-Hamilton theorem, it may readily be seen that the contribution of the initial condition term to V(t) is identically zero. Thus

$$V(t)=[A^{n-1}b+a_1A^{n-2}b+.....+a_{n-1}b,..........,Ab+a_1b,b]\begin{bmatrix} u(t) \\ u(t+1) \\ \cdot \\ \cdot \\ u(t+n-1) \end{bmatrix}$$

$$=G[u(t),........u(t+n-1)]^T$$

where

$$G=[A^{n-1}b+a_1A^{n-2}b+...+a_{n-1}b,....,Ab+\alpha_1b,b]$$

Since (A,b) is reachable, G is a full rank matrix, hence it follows from the definition 1. and the hypothesis that

$$\sum_{t=t_0+1}^{t_0+N}V(t)V^T(t)=G\sum_{t=t_0+1}^{t_0+N}\begin{bmatrix} u(t) \\ \cdot \\ \cdot \\ u(t+n-1) \end{bmatrix}[u(t),.........,u(t+n-1)]G^T$$

$$\geq\alpha GG^T\geq\gamma I \qquad (2.4)$$

uniformly in t_0, where $\gamma=\alpha\sigma_{\min}(GG^T)$, let

$$Y(t_0)=[V(t_0+1),V(t_0+2),........,V(t_0+N)]$$

Then from (2.4)

$$Y(t_0)Y^T(t_0)\geq\gamma I$$

Now observe that

$$Y(t_0)=[X(t_0+1),X(t_0+2),.......,X(t_0+n+N)]M \qquad (2.5)$$

with $M\in R^{(N+n)\times N}$ given by

$$M=\begin{bmatrix} a_n & 0 & & 0 \\ a_{n-1} & a_n & & \cdot \\ \cdot & a_{n-1} & & \cdot \\ \cdot & \cdot & & 0 \\ \cdot & \cdot & & a_n \\ a_1 & \cdot & & a_{n-1} \\ 1 & a_1 & & \cdot \\ 0 & 1 & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & a_1 \\ 0 & 0 & & 1 \end{bmatrix}$$

Using (2.5)

$$Y(t_0)Y^T(t_0)=WMM^TW^T$$

where

$$W=[X(t_0+1),X(t_0+2),\ldots\ldots,X(t_0+n+N)]$$

the following equation then follows immediately

$$WMM^TW^T \le \sigma_{\max}(MM^T)WW^T$$

or

$$WW^T \ge \frac{1}{\sigma_{\max}(MM^t)}WMM^TW^T = \frac{Y(t_0)Y^T(t_0)}{\sigma_{\max}(MM^t)}$$

Since M is a constant matrix, and $\sigma_{\max}(MM^T)$ is nonzero and independent of t_0. We have

$$\sum_{t_0=1}^{t_0+n+N} X(t)X^T(t) \ge \frac{\gamma}{\sigma_{\max}(MM^T)}I$$

This completes the proof.

Q.E.D.

Following is the dual of the theorem 1.

Theorem 2.(Observability and SR)
Consider the discrete time system

$$X(t+1)=AX(t)$$
$$y(t)=CX(t)$$
$$t\in Z_+, X(t)\in R^n, A\in R^{n\times n}, C\in R^{1\times n}, y(t)\in R$$

Suppose system is completely observable and X(t) is PE, then y(t) is SR of order n.

It is essential to note that the proof of the theorem 1. uses only the reachability. It allows one to deal with unstable systems. From a practical point of view, this is important, since the stability of the model may not be known prior to applying adaptive control.

From the lemma 2., it is easy to see that the input containing a linear combination of n/2 sinusoids is enough to produce PE states. A similar result in [5] needs 2n sinusoids for the same purpose. (Note that a sinusoid of nonzero frequency ν has symmetric spectrum at both ν and $-\nu$.)

In the next section, we will use the results obtained here to establish the global exponential convergence of the G.R.C. scheme.

3. Exponential convergence

To study the exponential convergence of all existing discrete time adaptive control schemes individually is tedious, and many schemes are related. Therefore, for brevity, we discuss the G.R.C. scheme with the modified projection algorithm for parameter update law.

Consider the plant which is modeled by

$$y(t)+\alpha_1 y(t-1)+.....+\alpha_n y(t-n)=\beta_0 u(t-d)+......\beta_m u(t-d-m) \qquad (3.1)$$

That is

$$\alpha(z^{-1})y(t)=z^{-d}\beta(z^{-1})u(t)$$

where

$$\alpha(z^{-1})=1+\alpha_1 z^{-1}+.......+\alpha_n z^{-n}$$
$$\beta(z^{-1})=\beta_0+\beta_1 z^{-1}+.......+\beta_m z^{-m}$$

Assume $\beta_0 \neq 0$ and that d, m n are known, but not the values α_i and β_i. Further, assume that $\alpha(z^{-1})$ and $\beta(z^{-1})$ are coprime and that the zeros of $\beta(z^{-1})$ lie inside the unit disk.

The objective of the adaptive control is to get y(t) to track a given reference trajectory.

First convert the model to a d-step-ahead predictor form [5].

$$y(t+d)=a_0 y(t)+a_1 y(t-1)+......a_{n-1} y(t-n+1)$$
$$+b_0 u(t)+.......+b_{m+d+1} u(t-m-d+1) \qquad (3.2)$$

where $b_0 \neq 0$ Consider the control law and update law by (3.3) and (3.4) respectively.

$$\varphi^T(t)\vartheta(t)=y^*(t+d) \qquad (3.3)$$

$$\vartheta(t)=\vartheta(t-1)+a(t)\varphi(t-d)[c+\varphi^T(t-d)\varphi(t-d)]^{-1}[y(t)-\varphi^T(t-d)\vartheta(t-1)] \qquad (3.4)$$

where $\vartheta(t)$ is parameter estimate, $y^*(t)$ is a reference output and

$$\varphi(t)=[y(t),y(t-1),...,y(t-1+n),u(t),......,u(t-m-d+1)]^T$$

For this scheme, Goodwin and co-workers [5,6] have shown the following facts:

(1) y(t) and u(t) are bounded.

(2) $\lim_{t\to\infty}[y(t)-y^*(t)]=0$

(3) $\lim_{n\to\infty}\sum_{t=1}^{n}[y(t)-y^*(t)]^2<\infty$

Thus the scheme above has attractive global convergence properties, however, the rate of convergence is not guaranteed to be exponential. Moreover, nothing can be said about the parameter errors. It is known [4] that if $\varphi(t)$ is PE, then both the output error and paramcter error will converge to zero exponentially. We will find condition on the reference trajectory which will guarantee that $\varphi(t)$ is PE.

Lemma 4. Consider the system (3.1), described by the predictor form (3.2), then φ (t) is PE if either of the following conditions holds:

1. u(t) is SR of order (n+m+d), or the spectral measure of the u(t) is not concertrated on k < (n+m+d) points and y(t) is bounded.

2. y(t) is SR of order (n+m+d), or the spectral measure of the y(t) is not concertrated on k < (n+m+d) points and u(t) is bounded.

Proof:
The proof of 1. and 2. proceed by similar arguments, hence we only prove 2. here.

From theorem 1., we know that if the system is reachable, then the SR of input implies that state variable is PE. What we do next is to use exactly this idea to build a system, whose input is y(t) and state is φ (t).

From the predictor form (3.2)

$$u(t)=\frac{1}{b_0}y(t+d)-\frac{a_0}{b_0}y(t)-.....-\frac{a_{n-1}}{b_0}y(t-n+1)$$
$$-\frac{b_1}{b_0}u(t-1)-...-\frac{b_{m+d-1}}{b_0}u(t-m-d+1)$$

Using (3.1), we rewrite the above equation as

$$u(t)=\frac{1}{b_0}y(t+d)+c_1y(t-1)+......+c_ny(t-n)$$
$$+c_{n+1}u(t-1)+.......+c_{n+m+d}u(t-d-m)$$

where c_i (i=1,2,....n+m+d) are constants.

Now it is not difficult to check that the vector φ (t) is generated by the following state equation with the input y(t).

$$\begin{bmatrix} y(t) \\ y(t-1) \\ \cdot \\ y(t-n+1) \\ u(t) \\ \cdot \\ u(t-m-d+1) \end{bmatrix} = \begin{bmatrix} -a_1 & -a_2 & -a_n & 0 & b_0 & b_m \\ 1 & \cdot & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & 0 & \cdot & 0 \\ c_1 & c_2 & c_n & c_{n+1} & c_{n+d} & c_{n+m+d} \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y(t-1) \\ y(t-2) \\ \cdot \\ y(t-n) \\ u(t-1) \\ \cdot \\ u(t-m+d) \end{bmatrix} + \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ 1/b_0 \\ 0 \\ \cdot \\ 0 \end{bmatrix} y(t+d)$$

or

$$\varphi(t)=A\varphi(t-1)+by(t+d)$$

It may be verified from the same transformation as in [5] that coprimeness of $\alpha(z^{-1})$ and $\beta(z^{-1})$ guarantees that

$$rank[\lambda I-A \quad b]=n+m+d \qquad \text{for any } \lambda\in C$$

This implies the system is completely reachable. By theorem 1., the SR of y(t) implies the PE of φ (t).

Q.E.D.

Theorem 2. If reference output y^* (t) is SR of order (n+m+d), or the spectral measure of the y^* (t) is not concertrated on k<(n+m+d) points, then the output error and parameter error converge to zero exponentially.

Proof: This follows directly from theorem 1., lemma 3., and lemma 4,.

4.Conclusion:

We have given a condition in terms of spectral measure for parameter convergence of the adaptive control of not necessarily stable linear system. The paper has dealt only with the exponential convergence of G.R.C. scheme with modified projection parameter update law, but the idea is more general and may be readily extended to several other discrete time adaptive schemes.

Reference:

(1) C.E.Rohrs,et al "Robustness of adaptive control algorithms in the presence of unmodeled dynamics" Proc. 21 st IEEE Conf. on Decision and Control 1982 pp.3-11

(2) S.Boyd and S.Sastry "On parameter convergence in adaptive control" System and Control Letter Vol.3(1983) pp 311-319

(3) S.Boyd and S.Sastry "Necessary and sufficient conditions for parameter convergence in adaptive control" Proc. American Control Conf. 1984 San Diego. pp 1584-1588

(4) B.D.O.Anderson and C.R.Johnson "Exponential convergence of adaptive identification and control algorithm" Automatica Vol.18(1982) pp 1-13

(5) G.C.Goodwin and K.S.Sin "Adaptive Filtering, Prediction, and Control" 1984 Prentice-Hall

(6) G.C.Goodwin, P.J.Ramadge, P.E.Caines "Discrete multivariable adaptive control" IEEE Trans. on Automatic Control Vol.25 1980 pp 449-456

(7) C.W. Burrill "Measure, integration and Probability" McGraw-Hill 1973 New York pp 281

(8) L.Ljung " Characterization of the concept of persistently exciting in frequency domain" Coden-Lutfd2/Tfrt-3038 Rept.7119 Nov.1971

(9) M.Bodson and S.S.Sastry "Small signal I/O stability of nonlinear control system: Application to the robustness of a MRAC scheme" Memorandum No.UCB/ERL M84/70 17 Sep. 1984 University of California, Berkeley

(10) H.Elliott, R.Christi and M.Das "Global stability of a direct hybrid adaptive pole placement algorithm" Tech. Rep.#UMASS-ECE-No.81-1 University of Massachusetts,Amherst