EXPONENTIAL CONVERGENCE OF ADAPTIVE IDENTIFICATION AND CONTROL ALGORITHMS

B. D. O. Anderson* and C. R. Johnson, Jr.**

**Department of Electrical and Computer Engineering, University of Newcastle, New South Wales, Australia*
***Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va 24061, USA.*

Abstract Output and equation error adaptive identification algorithms are shown to be exponentially convergent under a persistently exciting (or spanning) condition on the system inputs together with several other standard conditions. An adaptive control algorithm is shown to be exponentially convergent under a persistently exciting condition on the reference trajectory together with some standard conditions.

Keywords Identification; adaptive control; self-adjusting systems.

INTRODUCTION

This paper describes results on the exponential convergence of certain algorithms used in adaptive identification and control. At the outset, one can reasonably address the question, why bother establishing exponential convergence, as opposed simply to convergence? The answer has at least two facets. First, exponential convergence is a faster form of convergence than say a $1/k$ convergence as $k \to \infty$, and there are time constants associated with it which allow the stating of finite times in which convergence is practically complete; minimization of the time constants can even be attempted, and the concept of a time constant is readily related to by engineers. Second, systems which are exponentially convergent are generally systems which are tolerant of modelling errors, noise, and various inadequacies of realization. On the other hand, systems with a $1/k$ rate of convergence may when "unidealized" become unstable. This kind of difficulty is well understood in Kalman filtering, being considered under the notion of filter divergence, see e.g. Anderson and Moore (1979). It can be simply grasped by noting that where the iterations $x_{k+1} = 1/2\ x_k$ and $x_{k+1} = 1/k\ x_k$ both lead to $x_k \to 0$, the iterations $x_{k+1} = 1/2\ x_k + u_k$ and $x_{k+1} = 1/k\ x_k + u_k$ have very different properties for u_k say a constant sequence; the former gives a bounded x_k, the latter an unbounded x_k. Again if u_k is a white noise process, a dichotomy of behaviour occurs: in the first case, $E[x_k^2]$ is bounded, in the second case, it is not. It is in fact doubtful that a proof of convergence without exponential convergence of an adaptive algorithm should be regarded as providing justification in itself for use of the algorithm; moreover, if the real reason for using adaptive identification or control is to cope with a slowly-varying plant, it is highly doubtful that a rate of convergence slower than exponential in the stationary plant case will translate to satisfactory behaviour in the nonstationary case. This point is covered in some of the references on exponential convergence in adaptive schemes, the contributions of which are now described in more detail.

Among the earliest contributions, we note that of Lion, (1967); this author considered a continuous-time equivalent of what one calls, at least in discrete time, equation error identification. The condition for exponential convergence required signals to be periodic, so that Lyapunov theorems on periodic linear systems could be appealed to. Through misunderstanding of these Lyapunov theorems, a number of authors later asserted that when the system signals comprised linear combinations of sinusoids, exponential stability could be claimed. This was wrong, since a linear combination of sinusoids, though almost periodic, need not be periodic, and certain Lyapunov theorems valid in the periodic case are known to be invalid in the almost periodic case, for the first time it is believed, in Anderson (1974a, 1974b). Reference, Anderson (1974a) applied a new and powerful tool to the stability analysis which had been developed in Anderson and Moore (1969) - a time-varying Lyapunov lemma which can be used to deduce an exponential stability result precisely when a uniform observability property holds. The proof in Anderson (1974a) contained a proof of stability under even more general conditions than apply in the almost periodic case, and the result with these more general conditions was rederived in Morgan and Narendra (1977a). The derivation in Morgan and Narendra (1977a) on the one hand has the great advantage of offering many insightful examples and providing necessity and sufficiency of the stability condition rather than just sufficiency; on the other hand, it is lengthy, due to the fact that the general tools of Anderson (1974b) were not appealed to, but rather sometimes rederived in highly specific form. Reference, Anderson (1977a) is a development of the results of Anderson (1974a), and includes the main results of Morgan and Narendra (1977a) and also Morgan and Narendra (1977b). The application of Anderson (1977a) to problems of adaptive identification is described in Anderson (1977b). Several other contributions should be mentioned: those of Sondhi and Mitra (1976) and Weiss and Mitra (1979) establish exponential convergence results; reference G. Kreisselmeier (1977) obtains independently and in another way a number of the results of Anderson (1974a, 1974b, 1977a, 1977b), Morgan and Narendra (1977a, 1977b); reference Yuan and Wonham (1977) also contains important results on the design of system input signals to ensure exponential convergence of the identification scheme.

Virtually all the above references deal with adaptive situations where the plant is excited by deterministic signals. New techniques are required to handle the presence of random exciting signals. Some of the most general results for coping with such signals can be found in Bitmead and Anderson (1978, 1980a, 1980b), Farden, Goding and Sayood (1979), Jones (1973).

Where then lies the novelty of this paper? In our view, the novelty of the paper arises on two distinct grounds. First, the paper presents results for the exponential stability of an output error identification algorithm and of an adaptive control algorithm. Second, the conditions are free of a deficiency of the conditions in most but not all of the references cited above. Consider the discrete-time identification problem. Let $\{u_k\}$ and $\{y_k\}$ be the input and output sequence of the unknown plant and $\{z_k\}$ the output of the adaptive identifier. Then typical conditions for exponential stability are given, e.g. Bitmead and Anderson (1980b), in terms of the behaviour of a vector $x_k = [u_{k-1}\ u_{k-2} \cdots u_{k-m}\ z_{k-1} \cdots z_{k-n}]$ for some m, n. Of course, a condition in terms of u_k only is desired. This paper explains how such a condition can be obtained. Previous contributions have obtained such a condition on the input alone only when equation error identification has been considered, for which x_k is replaced by

$$x_k = [u_{k-1}\ u_{k-2} \cdots u_{k-m}\ y_{k-1} \cdots y_{k-n}]$$

and $\{u_k\}$ has been almost periodic. For the adaptive control problem, a desirable condition is one on the reference trajectory which the plant output is supposed to track.

The next section sets up the basic apparatus for passing from a condition involving inputs and outputs to one involving inputs only (for identification) or outputs only (for control). This apparatus is powerful enough to apply for both deterministic and stochastic signals with very little adjustment but because of space restrictions, we consider deterministic signals only. For almost all of Section 3, we consider output error identification schemes; the last part of this section is able to treat very efficiently equation error schemes. In Section 4, we consider adaptive control.

For adaptive identification or control, the main requirement for exponential convergence is a persistently exciting or spanning condition on the input or reference trajectory respectively.

REFORMULATING PERSISTENT EXCITATION CONDITIONS

The purpose of this section is to provide the tools which allow conversion of "persistent excitation" conditions involving input and output quantities to ones involving, for the identification problem, input quantities alone, and for the adaptive control problem, output quantities alone. Proofs may be found in a more extended version of the paper.

Theorem 1. Consider the single-input, single-output system

$$y_k + \alpha_1 y_{k-1} + \ldots + \alpha_n y_{k-n} = \beta_1 u_{k-1} + \ldots + \beta_m u_{k-m} \quad (1)$$

assume that $\alpha(z) = z^n + \alpha_1 z^{n-1} + \ldots + \alpha_n$ and $\beta(z) = \beta_1 z^{n-1} + \ldots + \beta_m z^{n-m}$ are coprime and that $\{y_k\}$ is a bounded sequence Then with

$$x_k = [y_k\ y_{k-1} \cdots y_{k-n+1}\ u_k\ u_{k-1} \cdots u_{k-m+1}]' \quad (2)$$

and S any integer with $S \geq 2n+m-2$,

$$\rho_2 I > \sum_{j}^{j+S} x_k x_k' > \rho_1 I > 0 \quad (3)$$

for all j and some ρ_1, ρ_2 if

$$\rho_4 I > \sum_{j}^{j+S-n+1} \begin{bmatrix} u_{k+n} \\ \vdots \\ u_{k-m+1} \end{bmatrix} [u_{k+n} \cdots u_{k-m+1}] > \rho_3 I > 0 \quad (4)$$

for some ρ_3, ρ_4 and all j.

The upper bound in (4) ensures $\{u_k\}$ is bounded, and if $\alpha(z)$ has all roots in $|z| < 1$, $\{y_k\}$ is bounded. However, it is conceivable that the system be unstable, but located in a stable closed-loop.

Suppose $\{u_k\}$ consists of a sum of p distinct sinusoidal terms and possibly a constant component; then $\{u_k\}$ also satisfies a difference equation, of order $q = 2p$ or $2p+1$, according as whether the constant term is present or not. If $q \geq n+m$, it is not hard to check that (4) must hold. So if there is an appropriate number of sinusoids present in $\{u_k\}$, (5) holds with $S = 2n+m-2$. This type of condition arises frequently in the literature, see e.g. Lion (1967), Kreisselmeier (1977), Yuan and Wonham (1977).

Secondly, we state a result relevant to adaptive control.

Theorem 2. Consider the single-input, single-output system (1) and assume that $\alpha(z) = z^n + \alpha_1 z^{n-1} + \ldots + \alpha_n$ and $\beta(z) = \beta_1 z^{n-1} + \ldots + \beta_m z^{n-m}$ are coprime and that $\{u_k\}$ is a bounded sequence. Let x_k be as in (2) and let S be any integer $\geq 2m+n-2$. Then (3) holds for all j and some ρ_1, ρ_2 if

$$\rho_4 I > \sum_{j}^{j+S-m+1} \begin{bmatrix} y_{k+m} \\ \vdots \\ y_{k-n+1} \end{bmatrix} [y_{k+m} \cdots y_{k-n+1}] > \rho_3 I > 0 \tag{5}$$

for some ρ_3, ρ_4 and all j. In case it is known that $\beta_1 = \beta_2 = \ldots = \beta_{d-1} = 0$ and $\beta_d \neq 0$, the result holds for $S \geq 2m+n-1-d$.

OUTPUT AND EQUATION ERROR IDENTIFICATION

We shall begin by studying output error identification. We state the basic problem, which requires the postulation of a particular form of identifier, and then summarise the solution. We then look at equation error identification, which is simpler.

The basic problem; plant and identifier structures.

We follow the notation of Johnson (1979), and postulate a model

$$y_k + \alpha_1 y_{k-1} + \ldots + \alpha_n y_{k-n} = \beta_1 u_{k-1} + \ldots + \beta_m u_{k-m} \tag{6}$$

(with n, m assumed known). We assume that n is minimal. We also assume that coefficients γ_i are known such that the transfer function

$$H(z) = \frac{1 + \sum_{i=1}^{n} \gamma_i z^{-i}}{1 + \sum_{i=1}^{n} \alpha_i z^{-1}} \tag{7}$$

is strictly positive real. This means that the model is necessarily stable. An adaptive identifier is constructed which at time k is parameterised by $\hat{\alpha}_i(k)$, $\hat{\beta}_i(k)$. It is considered as having two outputs, an a priori output $\hat{y}_k$ and an a posteriori output z_k:

$$\hat{y}_k = -\sum_{i=1}^{n} \hat{\alpha}_i(k) z_{k-i} + \sum_{j=1}^{m} \hat{\beta}_j(k) u_{k-j} \tag{8}$$

$$z_k = -\sum_{i=1}^{n} \hat{\alpha}_i(k+1) z_{k-i} + \sum_{j=1}^{m} \hat{\beta}_j(k+1) u_{k-j} \tag{9}$$

The parameters are updated with the aid of a smoothed error defined as

$$v_k = y_k - z_k + \sum_{i=1}^{n} \gamma_i [y_{k-i} - z_{k-i}] \tag{10}$$

with the update equations

$$\hat{\alpha}_i(k+1) = \hat{\alpha}_i(k) - \mu_i z_{k-i} v_k \qquad \mu_i > 0 \tag{11}$$

$$\hat{\beta}_j(k+1) = \hat{\beta}_j(k) + \rho_j u_{k-j} v_k \qquad \rho_j > 0 \tag{12}$$

Calculation of $\hat{\alpha}_i(k+1)$, according to (11), needs v_k which according to (10) needs $\hat{\alpha}_i(k+1)$; to avoid this circularity, one uses the following equivalent expression for v_k:

$$v_k = \frac{y_k - \hat{y}_k + \sum_{i=1}^{n} \gamma_i [y_{k-i} - z_{k-i}]}{1 + \sum_{i=1}^{n} \mu_i z_{k-i}^2 + \sum_{j=1}^{m} \rho_j u_{k-j}^2} \tag{13}$$

With bounded inputs, it can be shown by a simple Lyapunov analysis (Johnson, submitted for publication) or by appeal to hyperstability arguments, Johnson (1979) that $\hat{y}_k \to y_k$, i.e., that the identifier output converges to the plant output.

Let us note that the conjunction of positive real transfer functions with adaptive problems has proceeded for some time, see e.g., Butchart and Shackloth (1965), Ionescu and Monopoli (1977), Johnson (1980b), Landau (1976), Lin and Narendra (1978), Parks (1966).

It is easy to see that the convergence of $\hat{y}_k$ to y_k need not be exponentially fast - this can be checked for the example $y_k = \beta_1 u_{k-1}$, where $u_{k-1} = 0$ except when $k-1 = 2^\alpha$, $= 1, 2, 3, \ldots$ when $\rho u_{k-1}^2 = 1$; a simple calculation shows that $\hat{y}_k - y_k \to 0$ at a $1/k$ rate. Slower rates again are possible. It is also to see that even though u_k might be such as to apparently cause $\hat{y}_k$ to approach y_k exponentially fast, the convergence might not be robust, in that after convergence has apparently occurred, variation of u_k may mean that $\hat{y}_k$ no longer tracks y_k - consider for example a plant with non-constant transfer function, and suppose that u_k is constant over $[0, K]$ for some very large, arbitrary but fixed K. Then in $[0, K]$, $\hat{y}_k \to y_k$ exponentially fast. But if at time $k+1$, u_k becomes sinusoidally varying, $\hat{y}_k$ no longer tracks y_k, at least until the identifier has learnt the value of the plant transfer functions at the relevant frequency. To avoid both slow and nonrobust convergence we want the identifier parameters to approach the plant parameters exponentially fast. Accordingly, we are led to the basic problem: find conditions on $\{u_k\}$ such that $\hat{y}_k \to y_k$, $\hat{\alpha}_i(k) \to \alpha_i$ and $\hat{\beta}_i(k) \to \beta_i$ exponentially fast.

Solution to Basic Problem

Theorem 3. Consider the plant (6) where n, m are assumed known, $\alpha(z) = z^n + \alpha_1 z^{n-1} + \ldots + \alpha_n$, $\beta(z) = \beta_1 z^{n-1} + \ldots + \beta_m z^{n-m}$ are coprime but otherwise unknown. Suppose also that γ_i are known such that $H(z)$ in (7) is strictly positive real, and that the adaptive identification scheme of (8)-(13) is used. Then $y_k - \hat{y}_k$, $\alpha_i - \hat{\alpha}_i(k)$ and $\beta_j - \hat{\beta}_j(k)$ converge exponentially to zero provided that for S some integer $\geq 2n+m-2$, for all j and some ρ_3, ρ_4,

$$\rho_4 I > \sum_{j}^{j+S-n+1} \begin{bmatrix} u_{k+n} \\ \vdots \\ u_{k-m+1} \end{bmatrix} [u_{k+n} \cdots u_{k-m+1}] > \rho_3 I > 0$$

Major Steps in Proof[1]

1. Define

$$e_k = [y_{k-1} - z_{k-1} \cdots y_{k-n} - z_{k-n}]' \tag{14}$$

$$\phi_k = [-\alpha_1 + \hat{\alpha}_1(k) \ldots -\alpha_n + \hat{\alpha}_n(k) \;\; \beta_1 - \hat{\beta}_1(k) \cdots \beta_m - \hat{\beta}_m(k)]' \tag{15}$$

and $q_k = [e_k' \; \phi_k']'$. Then one can determine F_k such that

$$q_{k+1} = F_{k+1} q_k \tag{16}$$

2. Argue that (16) converges exponentially fast if the vector

$$x_k = [z_{k-1} \; z_{k-2} \cdots z_{k-n} \; u_{k-1} \cdots u_{k-m}] \tag{17}$$

satisfies a condition like that of (3)

3. Argue that (16) converges exponentially fast if the condition on x_k is replaced by the same condition on x_k, defined by

$$x_k = [y_{k-1} \; y_{k-2} \cdots y_{k-n} \; u_{k-1} \cdots u_{k-n}] \tag{18}$$

4. Apply Theorem 1 to deduce the result.

Equation Error Identification

Results for equation error identification, Mendel (1973), Johnson (1979), are much easier to achieve. In the simplest approach to equation error identification, we consider the plant (6) as before, i.e. n, m known, n minimal, and we explicitly assume the plant is stable. We also consider the identifier

$$\hat{y}_k = -\sum_{i=1}^{n} \hat{\alpha}_i(k) y_{k-i} + \sum_{j=1}^{m} \hat{\beta}_j(k) u_{k-j} \tag{19}$$

where we set

$$\hat{\alpha}_i(k+1) = \hat{\alpha}_i(k) - \frac{\mu_i}{1 + \sum_{i=1}^{n} y_{k-i}^2 + \sum_{j=1}^{m} u_{k-j}^2} y_{k-i}(y_k - \hat{y}_k)$$

$$\hat{\beta}_j(k+1) = \hat{\beta}_j(k) + \frac{\rho_j}{1 + \sum_{i=1}^{n} \mu_i y_{k-i}^2 + \sum_{j=1}^{m} \rho_j u_{k-j}^2} u_{k-j}(y_k - \hat{y}_k)$$

With ϕ_k the parameter error vector of (15) it is possible to find $\tilde{F}_k$ such that

$$\phi_{k+1} = \tilde{F}_k \phi_k \tag{20}$$

A simplified version of the argument applying for output error leads to the same conclusion as that of Theorem 2, without the requirement for the strictly positive real $H(z)$. The argument revolves around establishing exponential stability of (20). Of course, exponential convergence of ϕ_k to zero assures exponential convergence of $\hat{y}_k$ to y_k.

ADAPTIVE CONTROL

To keep the formulation simple, we shall consider a single-input, single-output system. We shall follow the first projection algorithm of Goodwin Ramadge and Caines (1980) shown in Johnson (1980a) to underlie several approaches to adaptive control. We shall show that when the output of the unknown plant is required to follow a sufficiently rich reference trajectory, then the estimate of the plant parameters will converge exponentially fast to the true value, as will the plant output to the reference trajectory.

The basic problem; plant and controller structures.

We assume the plant is modelled by

$$y_k + \alpha_1 y_{k-1} + \ldots + y_{k-n} = \beta_d u_{k-d} + \ldots + \beta_m u_m$$
$$\beta_d \neq 0 \tag{21}$$

where the values of d, m, n are assumed known, but not the values of the α_i, β_j. We assume that $\alpha(z) = z^n + \alpha_1 z^{n-1} + \ldots + \alpha_n$ and $\beta(z) = \beta_d z^{n-d} + \ldots + \beta_m z^{n-m}$ are

---

1 See extended version of paper for full detail.

ntial convergence.

What do these results lead on to? The question was in part answered in the introduction, where we referred to papers discussing the robustness of adaptive algorithms, e.g. to noise or slowly-varying parameters, in the event that exponential convergence has been established. Otherwise, we shall limit ourselves to two observations. First, one should be on guard when we assume a plant order larger than is really the case, perhaps more so for adaptive control. For then we cannot expect to identify the plant, and indeed one simulation study reports to us (Hannan, private communication) has resulted in the adaptive model predicting a pole-zero cancellation which migrates to +1; this hardly augurs well for adaptive control, and is worrisome for identification too. For similar reasons, we can be concerned about the absence of persistent excitation conditions. In equation error identification, it is not hard to argue that with absence of persistence excitation and with noise in the calculations, one or more linear functionals of the parameter estimation vector will diverge. As against these conclusions though, one could well argue of any real plant that it is impossible to overmodel its order, so that the difficulties are largely illusory.

The second observation is as positive as the first was negative. Proving exponential stability opens up the possibility of proving results on approximate identification and adaptive control, and in particular results applicable when the unknown plant order is underestimated (as may always be the case in practice). The relevant equations are normally perturbations of those equations for which exponential stability has been established; certain perturbations are always possible, in the sense that at worst bounded errors will result. We are at present applying this idea to the study of the adaptive control of nonminimum phase plants, where the plant output is supposed to follow a reference trajectory. Similar calculations to those justifying approximate identification should also justify the robustness of various algorithms when the plant is slowly time-varying.

There are several directions in which the ideas of this paper could be extended in a straightforward manner. Thus we would expect that multiple-input, multiple-output systems need hold few terrors if appropriate (efficient) parameterizations are used, and that other adaptive control algorithms for trajectory following (which all seem to fall into one family - see Johnson (1980a)) should be equally amenable to analysis. Continuous-time problems raise a slightly more subtle issue. In passing from input-output conditions for exponential stability to input-only conditions, it is necessary to demand that the inputs not become faster and faster (else the contributions to the output becomes smaller and smaller). This can be accomplished by using a technical device due to Yuan and Wonham (1977), and also used in Anderson (1977a, 1977b).

REFERENCES

Anderson, B.D.O. (1974a). Tech. Rept. EE7402 Uni. of Newcastle, Australia.

Anderson, B.D.O. (1974b). Proc. IEEE Conf. Decision and Control, 273-281.

Anderson, B.D.O. (1977a). IEEE Trans. Auto. Control, AC-22, 1, 83-88.

Anderson B.D.O. (1977b). Automatica, 13, 4, 401-408.

Anderson, B.D.O. and J.B. Moore (1969). SIAM J. Cont. 7, 3, 398-414.

Anderson, B.D.O. and J.B. Moore (1979). Optimal Filtering, Prentice Hall Inc. Englewood Cliffs, N.J.

Astrom, K.J. (1970). Introduction to Stochastic Control Theory, Academic Press, N.Y.

Bitmead, R.R. and B.D.O. Anderson (1978). Proc. IEEE Conf. Decision and Cont. 580-585.

Bitmead, R.R. and B.D.O. Anderson (1980a). IEEE Trans. Auto. Control, AC-25, 4, 782-788.

Bitmead, R.R. and B.D.O. Anderson (1980b). IEEE Trans. Auto. Control, AC-25, 4, 788-794.

Butchart, R.L. and B. Shackloth (1965). Proc. IFAC Symp. Teddington, England.

Conley, C.C. and R.K. Miller (1965). J. Diff. Equations, 1, 333-336.

Farden, D.C., J.C. Goding, Jr. and K. Sayood, Proc. IEEE on Acoust. Speech & Sig. Proc.

Goodwin, G.C., P.J. Ramadge and P.E. Caines (1980). IEEE Trans. Auto. Control. AC-25, 3, 449-450.

Hannan, E.J., Private communication.

Ionescu, T. and R. Monopoli, (1977). Automatica, 13, 507-517.

Johnson, C.R., Jr. (1979a). IEEE Trans. Syst. Man, Cybernet. SMC9, 7, 392-397.

Johnson, C.R., Jr. (1979b). IEEE Trans. Info. Theory, IT-25, 6, 745-749.

Johnson, C.R., Jr. (1980a) IEEE Trans. Auto. Control, AC-25, 4.

Johnson, C.R., Jr. (1980b). VPI&SU Tech. Rept. EE8002.

Jones, S.K. (1973). Ph.D. Thesis, Southern Methodist University.

Kreisselmeier, G. (1977). IEEE Trans. Auto. Control, AC-22, 1, 2-8.

Landau, I.D. (1976). IEEE Trans. Auto. Control, AC-21, 2, 194-202.

Lin, Y-H. and K.S. Narendra, Rept. 7802, System & Information Sciences, New Haven.

Mendel, J.M. (1973). Marcel Dekker.

Lion, P.M. (1967). AIAA J. 5, 10, 1835-1842.

Morgan, A.P. and K.S. Narendra, (1977a). SIAM J. Contr. 15, 1. 5-24.

Morgan, A.P. and K.S. Narendra, (1977b). SIAM J. Contr. 15, 1. 163-176.

Parks, P.C. (1966). IEEE Trans. Auto. Contr. AC-11, 3, 362-367.

Sondhi, M.M. and D. Mitra, (1976). Proc. IEEE, 64, 11, 1583-1597.

Weiss, A. and D. Mitra, (1979). IEEE Trans. Info. Theory, IT-25, 6, 637-652.

Yuan, J.S-C. and W.M. Wonham (1977). IEEE Trans. Auto. Control, AC-22, 4, 530-538.

For Discussion see page 930