-- --

PARAMETER CONVERGENCE AND UNIQUENESS IN NONMINIMAL PARAMETERIZATIONS FOR MULTIVARIABLE ADAPTIVE CONTROL

Yosef Willner

Department of Electrical Engineering, Technion, Haifa 32000, Israel

Michael Heymann

Department of Computer Science, Technion, Haifa 32000, Israel

and Marc Bodson

Department of Electrical and Computer Engineering,
Carnegie-Mellon University, Pittsburgh, PA 15213, U.S.A.

ABSTRACT

The issue of parameter convergence in multivariable adaptive control is addressed in a general framework. Parameter convergence is proved to be guaranteed if a certain design identity has a unique solution and if the inputs satisfy persistency of excitation conditions. The uniqueness of the solution of the design identity can be obtained in general by using parameterizations which, although nonminimal, are structured so as to guarantee uniqueness. This concept is illustrated with a direct adaptive pole placement algorithm which is modified to guarantee uniqueness and it is shown how the results can be used to establish stability and convergence properties of the algorithm.

Key words: Adaptive control, multivariable systems, pole placement, parameter convergence, persistency of excitation, parameterizations.

1. Introduction

The issue of parameter convergence in adaptive control has received some attention in recent years (see, among others, [1] and [2] in discrete–time and [3] and [4] in continuous–time). It was found that several single–input single–output schemes possessed exponential parameter convergence properties, provided that persistency of excitation (or sufficient richness) conditions were satisfied.

Although it is often argued that parameter convergence is not necessary in adaptive control (boundedness and tracking being the only objectives of model reference adaptive control, for example), there are important reasons to study this problem. First, exponential stability guarantees a certain degree of robustness (*cf.* [4], [5]). In the presence of noise, adaptive schemes exhibit parameter drift and a burst phenomenon (*cf.* [6], [17]) which can be avoided if persistency of excitation conditions are met. The problem can also be avoided using deadzones and projections, but only at the cost of additional prior information.

Another advantage of parameter convergence is that the closed–loop system actually has the asymptotic properties for which the controller was designed. Indeed, consider the case of a model reference adaptive scheme with an input signal that is constant over a long period of time. While the tracking error converges to zero, the closed-loop poles may converge to arbitrary locations. This may result in large transients when the reference input later varies.

It is important to note that we address ourselves here to a strong form of parameter convergence, namely uniform *exponential* parameter convergence to the *nominal* values of the parameters (also called correct, or true values). This form of convergence requires conditions of persistency of excitation. Weaker conditions on the input signals result in weaker forms of parameter convergence. For example, it is known that the parameters of the recursive least–squares algorithm converge without further conditions than those needed for stability ([18]). In that case however, the parameters do not necessarily converge to their nominal values and the convergence is not exponential in general. The advantages of the strong form of parameter convergence mentioned above, as far as robustness and asymptotic performance are concerned, are lost in such cases, where the persistency of excitation conditions are relaxed.

Very few rigorous proofs of stability have been published for multivariable adaptive control algorithms and parameter convergence has not been established. In fact, parameter convergence can often *not* be guaranteed for the existing schemes, even with sufficiently rich inputs. This happens because the parameterizations are not *unique*. In a model reference adaptive control algorithm for example, this means that an infinite number of values of the parameters exist such that model matching is achieved.

In the context of recursive identification, it was shown ([7]) that, using unique parameterizations, frequency–domain conditions on the inputs could be specified under which parameter convergence was guaranteed. The parameterization used there had the additional advantage of being *minimal*, *i.e.*, of requiring the minimal number of parameters necessary to describe the class of systems under consideration. In direct adaptive control, minimality is rarely achieved (even in the SISO case), but it was shown in [1] that it is only necessary for the parameterization to be unique (rather than minimal) to guarantee parameter convergence under suitable persistency of excitation conditions.

Contributions of the Paper

The first contribution in this paper is to extend the results of [1] to the multivariable case (sections 2, 3 and 4), thereby establishing a general framework for the convergence analysis of a large class of adaptive control algorithms. Specifically, the results show that parameter convergence is directly related to the *uniqueness* of a certain design identity. This result is important because it provides a criterion to guarantee parameter convergence and, furthermore, indicates that minimality is not itself necessary. For example, a single–input single–output linear time–invariant system can be described by $2n$ parameters, where n is the order of the system. However, parameter convergence can be achieved with a pole placement algorithm with $4n$ parameters. This is obtained by giving sufficient *structure* to the non–minimal model, so that uniqueness is guaranteed. This paper proves that the same principle holds true for multivariable systems, and gives a general framework in which to test the requirements for making exponential convergence of the parameters to their nominal values possible.

The second contribution of the paper is to show how uniqueness can be guaranteed in a specific adaptive pole placement scheme and to prove the stability and convergence properties of this scheme by applying the general results (sections 5 and 6). As noted above, existing direct adaptive control schemes do not guarantee uniqueness of the design identity. However, we show how the adaptive pole placement scheme of [8] can be modified to achieve this result.

Two nontrivial modifications are incorporated in the scheme of [8]: the first consists in restricting the column degrees of the elements of some polynomial matrices, taking into account the knowledge of the observability indices of the plant. It is interesting to observe the similarity to the situation that arises in recursive parametric identification. There, the knowledge of the observability indices can be used to constrain the column degrees of a left matrix fraction description of the plant, thereby leading to a canonical representation and to the uniqueness of the parameterization. In fact, an interesting feature of a proof presented in this paper is to show how known results on canonical forms can be used to guarantee the uniqueness of a direct adaptive control parameterization (that is, the uniqueness of the solution of the corresponding design identity).

As opposed to the situation in identification, the constraint on the column degrees is insufficient to guarantee uniqueness for the adaptive pole placement algorithm. It is found that the order of a certain observer polynomial matrix in [8] also has to be modified to guarantee uniqueness of the solution of the design identity. This second modification is not obvious and rather technical, but is given in this paper.

We follow, as much as possible, the notation and terminology of [1] and [8], to which this work is most closely related. The reader may wish to consult [1] in particular, for motivation and for additional details on some of the techniques used in this paper.

2. A General Parameter Estimation Problem

We consider discrete–time, linear time–invariant systems modeled by the state equations

$$x(t+1) = A_s\, x(t) + B_s\, u(t)$$
$$y(t) = C_s\, x(t) + E_s\, u(t) \tag{2.1}$$

where $u(t)$ is the $(m \times 1)$ input vector, $y(t)$ the $(p \times 1)$ output vector and $x(t)$ the $(n \times 1)$ state vector. We assume that the system (2.1) is minimal. Let the controllability indices $\mu_i,\ \ 1 \le i \le m$ and the observability indices $\nu_i,\ \ 1 \le i \le p$ be defined as usual (*cf.* [9]). Let $\mu = \max\limits_{1 \le i \le m} (\mu_i)$ the maximal controllability index, simply called the controllability index, and $\nu = \max\limits_{1 \le i \le p} (\nu_i)$ called the observability index. Such a system can also be represented by the right matrix fraction description

$$P(D) \cdot \xi(t) = u(t)$$
$$y(t) = R(D) \cdot \xi(t) \tag{2.2}$$

where $R(D)$ and $P(D)$ are $(p \times m)$ and $(m \times m)$ real polynomial matrices in the unit delay operator (i.e. $D^k x(t) = x(t-k)$). Matrices $R(D)$ and $P(D)$ exist that have the following properties (*cf.* [9])

a) $\partial_{cj} R(D) \le \mu_j$ and $\partial_{cj} P(D) = \mu_j$, where $\partial_{cj}[\,\cdot\,]$ denotes the maximal polynomial degree in the j–th column,

b) $P(0)$ is nonsingular,

c) The matrices $R(D)$ and $P(D)$ are right coprime.

The matrix $P(D)$ can be further constrained, in particular so that it is in some canonical form (*cf.* [10]). This will be discussed in section 5.

Structured Nonminimal Model

To introduce a general framework for the study of direct adaptive control algorithms, we replace the minimal model (2.2) by a *structured nonminimal model* of the form

$$[C(D) + \sum_{j=1}^{m_a} A_j(D)\,\alpha_j]\; y(t) = [E(D) + \sum_{j=0}^{m_b} B_j(D)\,\beta_j]\; u(t) \tag{2.3}$$

where m_a and m_b are positive integers, $\alpha_j \in \mathbb{R}^{r \times p}$, $\beta_j \in \mathbb{R}^{r \times m}$, $C(D)$ and $E(D)$ are $(r \times p)$ and $(r \times m)$ polynomial matrices with maximal degree l. $A_j(D)$ and $B_j(D)$ are $(r \times r)$ polynomial matrices of the form

$$A_j(D) = \mathrm{diag}\,[a_{ij}(D)] \quad , \quad a_{ij}(D) = \sum_{k=1}^{l} a_{ijk} D^k \quad , \quad a_{ijk} \in \mathbb{R} \tag{2.4}$$

$$B_j(D) = \mathrm{diag}\,[b_{ij}(D)] \quad , \quad b_{ij}(D) = \sum_{k=0}^{l} b_{ijk} D^k \quad , \quad b_{ijk} \in \mathbb{R} \tag{2.5}$$

The structured non–minimal model defined by (2.3), (2.4), (2.5) is an extension of the single–input single–output model of [1]. There, it was shown that the simplified model was adequate to describe several adaptive control algorithms. In section 5, we will show that the multivariable adaptive pole placement algorithm fits into the generalized framework. The integer r is equal to m, the number of inputs, in that case. In other cases, it may take different values (for example, (2.3) may represent a left matrix fraction description, with $r = p$).

At this point, we let r and l be arbitrary integers, but it is assumed that, for $i = 1, 2, \ldots, r$, the polynomials $\{ a_{ij}(D) \}_{j=1}^{m_a}$ are linearly independent over the reals. It is assumed similarly that the $\{ b_{ij}(D) \}_{j=0}^{m_b}$ are linearly independent. These assumptions imply that $\max \{ m_a, m_b \} \le l$.

We assume that the plant can be represented by the model (2.3) for some given matrices $C(D)$, $E(D)$, $A_j(D)$ and $B_j(D)$. One wishes to use the model (2.3) to uniquely estimate the elements of the matrices α_j, $1 \le j \le m_a$, and β_j, $0 \le j \le m_b$, from the plant input–output data. Clearly, (2.3) constitutes a model for the plant (2.2) if and only if the following *design identity* is satisfied

$$[C(D) + \sum_{j=1}^{m_a} A_j(D)\,\alpha_j]\,R(D) = [E(D) + \sum_{j=0}^{m_b} B_j(D)\,\beta_j]\,P(D) \qquad (2.6)$$

Clearly, the elements of the matrices α_j and β_j in (2.3) can be uniquely estimated only if (2.6) has a unique solution $\{ \alpha_1, \ldots, \alpha_{m_a}, \beta_0, \ldots, \beta_{m_b} \}$. Conversely, whenever (2.6) has a unique solution, we will show that the solution can be obtained by a direct estimation algorithm which is exponentially convergent. This will be the focus of the ensuing discussion.

Remark 1: We wish to emphasize that the problem of finding conditions which ensure that (2.6) has a unique solution is quite different when the plant (2.2) is SISO (single–input single–output) and when the plant is MIMO (multi–input multi–output):

In the SISO case, for given polynomials $C(D)$ and $E(D)$, a solution to (2.6) exists if the polynomials $R(D)$ and $P(D)$ are coprime and if the degree of the polynomials $A(D) \triangleq \sum A_j(D)\,\alpha_j$ and $B(D) \triangleq \sum B_j(D)\,\beta_j$ are sufficiently large. Among all solutions of (2.6), there is a unique solution $\{ A(D), B(D) \}$ with minimal degree. Therefore, to ensure a unique solution of (2.6), it is sufficient to bound the degrees of the polynomials $A_j(D)$, $1 \le j \le m_a$, and $B_j(D)$, $0 \le j \le m_b$.

In the MIMO case, (2.6) has a solution if the matrices $R(D)$ and $P(D)$ are right coprime and if the degrees of the elements of the matrices $A(D)$ and $B(D)$ are sufficiently large. But, to ensure the uniqueness of the solution, it is necessary to restrict the maximal *and the minimal* powers in D of each element of $A(D)$ and $B(D)$. In other words, we can guarantee uniqueness by restricting the maximal degree of each element in $A_j(D)$ and $B_j(D)$ and by choosing some of the elements of α_j and β_j as zero. This fact will be made clearer in section 5.

For parameter estimation purposes, it is convenient to write every row of (2.3) as an independent equation

$$[C_i(D) + \sum_{j=1}^{m_a} a_{ij}(D)\,\alpha_{ij}]\;y(t) = [E_i(D) + \sum_{j=0}^{m_b} b_{ij}(D)\,\beta_{ij}]\,u(t) \qquad (2.7)$$

for $i = 1, 2, \ldots, r$, where $C_i(D)$, $E_i(D)$, α_{ij} and β_{ij} are the i–th row of the matrices $C(D)$,

$E(D)$, α_j and β_j, and where the polynomials $a_{ij}(D)$ and $b_{ij}(D)$ are defined in (2.4) and (2.5). These equations can be written as regression equations

$$\bar{\phi}_i^{\,T}(t)\,\bar{\theta}_i^* = E_i(D)\,u(t) - C_i(D)\,y(t) \quad , \quad i = 1, 2, \ldots, r \tag{2.8}$$

where

$$\bar{\phi}_i^{\,T}(t) = [\, a_{i1}(D)\,y^T(t), \ldots, a_{im_a}(D)\,y^T(t),$$

$$-\,b_{i0}(D)\,u^T(t), \ldots, -\,b_{im_b}(D)\,u^T(t)] \tag{2.9}$$

$$\bar{\theta}_i^* = [\,\alpha_{i1}, \ldots, \alpha_{im_a}, \beta_{i0}, \ldots, \beta_{im_b}]^T \tag{2.10}$$

As indicated in remark 1, (2.6) has a unique solution provided that the elements of $A(D) \triangleq \sum A_j(D)\,\alpha_j$ and $B(D) \triangleq \sum B_j(D)\,\beta_j$ satisfy some degree conditions. These conditions depend on the adaptive control problem, and we will give specific conditions in the case of the pole placement algorithm in section 5. The degree conditions imply that some elements in the vectors $\bar{\theta}_i^*$, $1 \le i \le r$ are zero. We delete these zero elements from $\bar{\theta}_i^*$ as well as the corresponding elements from $\bar{\phi}_i(t)$, and define the resulting vectors θ_i^* and $\phi_i(t)$, respectively. (2.8) is then equivalent to

$$\phi_i^{\,T}(t)\,\theta_i^* = E_i(D)\,u(t) - C_i(D)\,y(t) \quad , \quad i = 1, 2, \ldots, r \tag{2.11}$$

Standard estimation procedures, such as the recursive least squares (RLS) algorithm, can be used to estimate each of the parameter vectors θ_i^* using input–output data of the plant. It is well–known (*cf.* [2]) that to insure the global convergence of the estimation algorithms, it is necessary to satisfy a persistency of excitation condition. In section 3, we will introduce linear systems called the *associated–signal systems* of (2.11). Through the use of these systems, we will show in section 4 how the persistency of excitation condition can be satisfied.

3. The Associated–Signal System and Its Output Reachability

For each equation in (2.11), we define the associated–signal system, which is a linear system in state-space form. Its input vector is $u(t)$ (the input vector of the plant (2.1) or (2.2)) and $\phi_i(t)$ is its output vector. Let the state of the associated–signal system be defined as the following $(m(l+\mu))$ vector

$$x_a(t) = [\xi^T(t-1),\ \xi^T(t-2), \ldots, \xi^T(t-l-\mu)]^T \tag{3.1}$$

where $\xi(t)$ is defined in (2.2), l is the same as in (2.4) and (2.5), and μ is the controllability index of the plant (2.1). The matrices $R(D)$ and $P(D)$ (in (2.2)) can be written as

$$R(D) = \sum_{k=0}^{\mu} R_k D^k \quad , \quad P(D) = \sum_{k=0}^{\mu} P_k D^k \tag{3.2}$$

It follows from the first equation of (2.2) that $x_a(t)$ satisfies the discrete–time state equation

$$x_a(t+1) = A\, x_a(t) + B\, u(t) \tag{3.3}$$

where

$$A = \begin{bmatrix} -P_0^{-1}P_1 & -P_0^{-1}P_2 & \cdots & -P_0^{-1}P_\mu & 0 & \ldots & 0 \\ & & & & & \cdot & \\ & & & & & \cdot & \\ & & & & & \cdot & \\ & & I_{m(l+\mu-1)} & & & \cdot & \\ & & & & & \cdot & \\ & & & & & \cdot & \\ & & & & & 0 & \end{bmatrix} , \quad B = \begin{bmatrix} P_0^{-1} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \tag{3.4}$$

By using (2.2), (2.4) and (2.5), it can be shown that the vectors $\bar{\phi}_i(t)$ (in (2.9)) satisfy the equations

$$\bar{\phi}_i(t) = \bar{C}_i x_a(t) + \bar{E}_i u(t) \quad , \quad i = 1, 2, \ldots, r \tag{3.5}$$

where $\bar{C}_i$ is the following $((pm_a + m(m_b + 1)) \times (m(l + \mu)))$ matrix

$$\bar{C}_i = \begin{bmatrix} a_{i11} R_0 & a_{i12} R_0 + a_{i11} R_1 & \cdots & a_{i1l} R_\mu \\ \vdots & \vdots & & \vdots \\ a_{im_a 1} R_0 & a_{im_a 2} R_0 + a_{im_a 1} R_1 & \cdots & a_{im_a l} R_\mu \\ -b_{i01} P_0 & -b_{i02} P_0 - b_{i01} P_1 & \cdots & -b_{i0l} P_\mu \\ \vdots & \vdots & & \vdots \\ -b_{im_b 1} P_0 & -b_{im_b 2} P_0 - b_{im_b 1} P_1 & \cdots & -b_{im_b l} P_\mu \end{bmatrix} \tag{3.6}$$

The element of $\bar{C}_i$ in the j–th row and k–th column is the coefficient of D^k in $a_{ij}(D)\, R(D)$ for $j=1,\ldots, m_a$. Similarly, for $j=0,\ldots, m_b$ the element in the $(j + m_a + 1)$–th row and k–th column is the coefficient of D^k in $-(b_{ij}(D) - b_{ij}(0))\, P(D)$. $\bar{E}_i$ is then the $((p\, m_a + m(m_b+1)) \times m)$ matrix

$$\bar{E}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -b_{i00} \cdot I \\ -b_{i10} \cdot I \\ \vdots \\ -b_{im_b 0} \cdot I \end{bmatrix} \tag{3.7}$$

where the elements a_{ijk} and b_{ijk} are defined in (2.4) and (2.5).

By definition, the *associated–signal system* of the *i*–th equation of (2.11) is the system

$$x_a(t+1) = A\, x_a(t) + B\, u(t)$$

$$\phi_i(t) = C_i\, x_a(t) + E_i\, u(t) \qquad (3.8)$$

where the matrices C_i and E_i are obtained from the matrix $\bar{C}_i$ and $\bar{E}_i$ by choosing the rows that correspond to the columns of $\bar{\phi}_i^T(t)$ that were selected to form $\phi_i^T(t)$. It might be pointed out that the state vector $x_a(t)$ and the matrices A and B are the same in all the associated–signal systems.

Now, recall that a linear system is called *output–reachable* if and only if every vector in its output space can be generated (reached) using a suitable input sequence. The following theorem relates the uniqueness of the solution of the design identity (2.6) to the output–reachability of the associated–signal systems.

Theorem 3.1

Assume that (2.6) is solvable.

The solution is unique *if and only if* all r associated–signal systems of (2.11) are output–reachable.

Proof of Theorem 3.1: the proof follows the lines of the proof of theorem 4.1 for the scalar case presented in [1] and is omitted here. □

Our original problem was to ensure that the estimation processes, which are based on (2.11), converge. It is known ([2]) that algorithms such as the RLS algorithm yield a sequence of estimates that converge exponentially fast to θ_i^* provided the sequence $\{ \phi_i(t) \}$ of regression vectors is persistently exciting. The question is to find input sequences $\{ u(t) \}$ such that all output sequences $\{ \phi_i(t) \}$ of the (output-reachable) associated-signal systems will be persistently exciting. This problem is solved in the next section.

4. Uniform Persistent Spanning Of Output–Reachable MIMO Plants

Consider the discrete–time, linear time–invariant plant

$$x_a(t+1) = A\, x_a(t) + B\, u(t)$$

$$y_a(t) = C\, x_a(t) + E\, u(t) \tag{4.1}$$

where $u(t) \in \mathbb{R}^m$, $y_a(t) \in \mathbb{R}^{p_a}$ and $x_a(t) \in \mathbb{R}^{n_a}$. In particular, this plant represents the associated–signal systems of section 3, with $y_a = \phi_i$, $n_a = m(l + \mu)$, etc. The system (4.1) can also be represented by the difference equation

$$d_{n_a}\, y_a(t+1) + d_{n_a-1}\, y_a(t+2) + \ldots d_1\, y_a(t+n_a) + y_a(t+n_a+1)$$

$$= G_{n_a}\, u(t+1) + G_{n_a-1}\, u(t+2) + \ldots + G_0\, u(t+n_a+1) \tag{4.2}$$

where d_i are the coefficients of a monic minimal polynomial for A, i.e. $a(z) = d_{n_a} + d_{n_a-1} z + \ldots + d_1 z^{n_a-1} + z^{n_a}$. The matrices $G_i \in \mathbb{R}^{p_a \times m}$ are defined by

$$G_i = \sum_{k=0}^{i} d_{i-k}\, M_k \quad , \qquad d_0 = 1 \tag{4.3}$$

where the matrices M_k are the Markov parameters, *i.e.*

$$[M_0\, , \, . \, . \, . \, , \, M_{n_a}] = [E,\ CB,\ CAB\, , \, . \, . \, . \, , \, CA^{n_a-1}B] \tag{4.4}$$

We use the following definitions

$$y_{a,\ j}(t) = [y_a(t+1)\, , \, . \, . \, . \, , \, y_a(t+j)] \tag{4.5}$$

$$\bar{u}_k(t) = [u^T(t+1)\, , \, . \, . \, . \, , \, u^T(t+k)]^T \tag{4.6}$$

$$U_{k,j}(t) = [\bar{u}_k(t+1)\, , \, . \, . \, . \, , \, \bar{u}_k(t+j)] \tag{4.7}$$

$$G = [G_{n_a}\, , \, . \, . \, . \, , \, G_0] \tag{4.8}$$

$$d = [d_{n_a}\, , \, . \, . \, . \, , \, d_1,\ 1\,]^T \tag{4.9}$$

Equation (4.2) can then be written as

$$G \cdot \bar{u}_{n_a+1}(t) = y_{a,\ n_a+1}(t) \cdot d \tag{4.10}$$

Definition: The sequence $\{\, y_a(t)\, \}$ is called *persistently exciting* if there exist $\varepsilon > 0$ and integers t_0 and N such that, for all integers $i \geq 0$

$$\lambda_{\min}\, [\; y_{a,\, N}(t_0 + i\, N)\; y_{a,\, N}^T(t_0 + i\, N)\,] \geq \varepsilon > 0 \tag{4.11}$$

Since adaptation algorithms are known to be exponentially convergent provided that the outputs of their associated–signal systems are persistently exciting, a natural question is to find conditions on the inputs which result in this property. The following theorem addresses this issue.

Theorem 4.1

Assume that the plant (4.1) is output–reachable.

If there exist $\varepsilon_1 > 0$ and integers t_1 and $N \geq n_a(m+1) + m$ such that, for all integers $i \geq 0$

$$\lambda_{\min}\,[U_{n_a+1,\,N-n_a}(t_1 + iN)\,U^T_{n_a+1,\,N-n_a}(t_1 + iN)] \geq \varepsilon_1 > 0 \tag{4.12}$$

Then, the sequence $\{ y_a(t) \}$ is persistently exciting for every initial state $x_a(0)$.

Remark 2: Note that (4.11) and (4.12) are equivalent to

$$\lambda_{\min}\left[\sum_{t=t_0+iN+1}^{t_0+iN+N} y_a(t)\,y_a^T(t)\right] \geq \varepsilon > 0 \tag{4.13}$$

and

$$\lambda_{\min}\left[\sum_{t=t_1+iN+1}^{t_1+iN+N-n_a} \bar{u}_{n_a+1}(t)\,\bar{u}^T_{n_a+1}(t)\right] \geq \varepsilon_1 > 0 \tag{4.14}$$

Therefore, theorem 4.1 shows that the persistency of excitation on y_a can be transformed into a similar condition on $\bar{u}_{n_a+1}$, which depends only on the input vector u. The vector $\bar{u}_{n_a+1}(t)$ is obtained by stacking the vectors $u(t+1), \ldots, u(t+n_a+1)$ on top of each other in a long vector. Since the dimension of $\bar{u}_{n_a+1}$ is $(n_a+1)\,m$, the "span" of the sum, $N-n_a$, must be greater than or equal to $(n_a+1)\,m$, and therefore the condition of theorem 4.1.

Proof of Theorem 4.1

Let $\alpha \in \mathbb{R}^{p_a}$ be a nonzero vector. Using (4.10)

$$\alpha^T G\, U_{n_a+1,\,N-n_a}(t_1)\, U^T_{n_a+1,\,N-n_a}(t_1)\, G^T \alpha = \alpha^T G \left[\sum_{t=t_1+1}^{t_1+N-n_a} \bar{u}_{n_{a+1}}(t)\,\bar{u}^T_{n_{a+1}}(t)\right] G^T \alpha$$

$$= \left[\sum_{t=t_1+1}^{t_1+N-n_a} (\alpha^T y_{a,\,n_a+1}(t)\, d\,)\,(d^T y^T_{a,\,n_a+1}(t)\,\alpha)\right]$$

$$\leq \| d \|^2 \left[\sum_{t=t_1+1}^{t_1+N-n_a} \| \alpha^T y_{a,\, n_a+1}(t) \|^2 \right] = \| d \|^2 \alpha^T \left[\sum_{t=t_1+1}^{t_1+N-n_a} y_{a,\, n_a+1}(t)\, y_{a,\, n_a+1}^T(t) \right] \alpha$$

$$\leq \| d \|^2 \alpha^T \left[\sum_{t=t_1+1}^{t_1+N-n_a} \sum_{j=t+1}^{t+n_a+1} y_a(j)\, y_a^T(j) \right] \alpha \leq \| d \|^2 (n_a+1)\, \alpha^T \left[\sum_{j=t_1+2}^{t_1+N+1} y_a(j)\, y_a^T(j) \right] \alpha$$

$$\leq \| d \|^2 (n_a+1)\, \alpha^T [y_{a,\, N}(t_1+1)\, y_{a,\, N}^T(t_1+1)]\, \alpha \tag{4.15}$$

Since the system (4.1) is output–reachable, the matrix G has full row rank, and using (4.12)

$$\lambda_{\min} [y_{a,\, N}(t_1+1)\, y_{a,\, N}^T(t_1+1)] \geq \frac{1}{\| d \|^2 (n_a+1)} \lambda_{\min} \left[G\, U_{n_a+1,\, N-n_a}(t_1)\, U_{n_a+1,\, N-n_a}^T(t_1)\, G^T \right]$$

$$\geq \frac{\lambda_{\min}(G\, G^T)}{\| d \|^2 (n_a+1)}\, U_{n_a+1,\, N-n_a}^T(t_1)$$

$$\geq \varepsilon_1 \frac{\lambda_{\min}(G\, G^T)}{\| d \|^2 (n_a+1)} > 0 \tag{4.16}$$

If we repeat the proof with $t_1 + iN$ instead of t_1, and let $t_0 = t_1 + 1$

$$\lambda_{\min} \left[y_{a,\, N}(t_0 + iN)\, y_{a,\, N}^T(t_0 + iN) \right] \geq \varepsilon_1 \frac{\lambda_{\min}(G\, G^T)}{\| d \|^2 (n_a+1)} = \varepsilon > 0 \tag{4.17}$$

and the proof is completed. □

The following theorem shows that the results of theorem 4.1 can be extended to cover situations where the input of the plant is calculated from an external reference input and state feedback. Input conditions are transferred to the reference input, assuming that the feedback gain matrix is held constant for sufficiently long periods between updates.

Theorem 4.2

Consider an output–reachable linear plant (4.1). Let the input sequence $u(t)$ be defined by the control law $u(t) = F_N(t)\, x_a(t) + \mathrm{v}(t)$, where $\mathrm{v}(t)$ is an external input and where $F_N(t)$ is a feedback gain matrix.

If the matrix $F_N(t)$ is bounded and changes value only at times $t_i = t_0 + i\,N$, $i = 0, 1, 2,...$, with $N \geq n_a(m+1)+m$, and if the external input $\mathrm{v}(t)$ satisfies the condition that

$$\lambda_{\min}\left[V_{n_a+1,\, N-n_a}(t_i)\, V^T_{n_a+1,\, N-n_a}(t_i) \right] \geq \varepsilon > 0 \tag{4.18}$$

Then, the sequence $\{\, y_a(t) \,\}$ is persistently exciting for every initial state $x_a(0)$.

Proof of Theorem 4.2: The proof is similar to the proof of theorem 5.3 presented in [1] and is omitted here. □

In the next sections, we show how theorems 3.1 and 4.1 can be used to prove the global convergence of an adaptive control algorithm.

5. Adaptive Pole Placement for Linear Multivariable Systems

It is assumed that the following parameters are known: n, m, p, μ_i, $i = 1, 2, \ldots, m$ and ν_i, $i = 1, 2, \ldots, p$. This is all the prior information required by the algorithm. The output of the plant (2.2) is given by

$$y(t) = R(D)\, P^{-1}(D)\, u(t) \tag{5.1}$$

The desired closed-loop dynamics are given by

$$y(t) = R(D)\, P^{*^{-1}}(D)\, \mathrm{v}(t) \tag{5.2}$$

where $\mathrm{v}(t)$ is the external input and $P^*(D)$ is a polynomial matrix that characterizes the desired closed–loop pole locations. The control algorithm is an adaptive version of the control law

$$u(t) = Q^{-1}(D)\, [\, H(D)\, y(t) + K(D)\, u(t)] + \mathrm{v}(t) \tag{5.3}$$

where $Q(D)$ is a fixed $(m \times m)$ polynomial matrix (all zeros of $\det(Q(D))$ are outside the unit circle) and where $H(D)$ and $K(D)$ are $(m \times p)$ and $(m \times m)$ controller matrices. The design equation of the controller is

$$H(D)\, R(D) + K(D)\, P(D) = Q(D)\, [\, P(D) - P^*(D)\,] \tag{5.4}$$

Since $R(D)$ and $P(D)$ are right coprime, there exist $(m \times p)$ and $(m \times m)$ matrices $J(D)$ and $I + S(D)$ that satisfy the Bezout identity

$$J(D)\,R(D) + [I + S(D)]\,P(D) = I \tag{5.5}$$

Using (5.1), (5.4) and (5.5) we get the following nonminimal model of the plant

$$[H(D) + Q(D)\,P^*(D)\,J(D)]\,y(t) = [-K(D) - Q(D)\,P^*(D)\,S(D) + Q(D)(I - P^*(D))]\,u(t) \tag{5.6}$$

This model is of the form (2.3). In this example, the general design identity (2.6) has the form

$$[H(D) + Q(D)\,P^*(D)\,J(D)]\,R(D) = [-K(D) - Q(D)\,P^*(D)\,S(D) + Q(D)(I - P^*(D))]\,P(D) \tag{5.7}$$

In the following theorem, we give conditions that ensure that equation (5.7) has a unique solution $\{ H(D), K(D), J(D), S(D) \}$ such that $\{ H(D), K(D) \}$ satisfy the design equation (5.4).

Theorem 5.1

Consider a plant of the form (2.1). Let $R(D)$ and $P(D)$ be the matrices in model (2.2). Let $Q(D)$ and $P^*(D)$ be $(m \times m)$ matrices of the form

$$Q(D) = \mathrm{diag}\,[q_j(D)] \quad , \quad \deg\,[q_j(D)] = \nu + \mu - \mu_j \quad , \quad q_j(0) = 1$$

$$P^*(D) = \mathrm{diag}\,[p_j^*(D)] \quad , \quad \deg\,[p_j^*(D)] = \mu_j \quad , \quad p_j^*(0) = 1$$

for $j = 1, 2, \ldots, m$, where $q_j(D)$ and $p_j^*(D)$ are polynomials which have zeros outside the unit circle.

Then, (5.7) has a unique solution $\{ H(D), K(D), J(D), S(D) \}$ of the form

$$H_i^{\,j}(D) = \sum_{k = 1+\nu+\mu-\mu_i-\nu_j}^{\nu+\mu-\mu_i} H_{ik}^{\,j}\,D^k \quad , \quad K_i^{\,j}(D) = \sum_{k = k_0(i,\,j)}^{\nu+\mu-\mu_i} K_{ik}^{\,j}\,D^k$$

$$J_i^{\,j}(D) = \sum_{k = 1+\nu+\mu-\mu_i-\nu_j}^{\nu+\mu-\mu_i} J_{ik}^{\,j}\,D^k \quad , \quad S_i^{\,j}(D) = \sum_{k = k_0(i,\,j)}^{\nu+\mu-\mu_i} S_{ik}^{\,j}\,D^k \tag{5.8}$$

where $H_i^{\,j}(D)$ denotes the ij–th element of $H(D)$, $H_{ik}^{\,j}, K_{ik}^{\,j}, J_{ik}^{\,j}, S_{ik}^{\,j} \in \mathbb{R}$ and where $k_0(i,\,j)$ is given by

$$k_0(i,\ j) = \begin{cases} 1 + \mu_j - \mu_i & \text{for } \mu_j \geq \mu_i \\ 0 & \text{for } \mu_j < \mu_i \end{cases} \tag{5.9}$$

The solution $\{H(D), K(D), J(D), S(D)\}$ is also the unique solution of (5.4) (5.5) under the conditions (5.8).

Remark 3: Note that any solution of (5.4), (5.5) is clearly a solution of (5.9). The reverse, however, is not so obvious. Theorem 5.1 shows that, under the degree constraints, (5.4) and (5.5) have unique solutions $\{H(D), K(D)\}$ and $\{J(D), S(D)\}$ which, together, constitute the unique solution of (5.7). To achieve this objective, we observe that constraints were imposed on the *lowest* as well as the highest degrees of $H(D)$, $K(D)$, $J(D)$, $S(D)$. Furthermore, compared to the scheme of [8], the degrees of the $q_j(D)$'s were increased from $\nu + \mu$ to $\nu + \mu - \mu_j$. This modification is not necessary for the uniqueness of the solutions of (5.4), (5.5), but was found to be needed to prove that any solution of (5.7) is a solution of (5.4), (5.5).

Proof of Theorem 5.1

Preliminaries:

The proof is easier to derive in terms of the forward shift operator, rather than in terms of the backward shift operator or delay D. In this framework, the proof is also similar to the proof for model reference adaptive control given in [4], pp. 288. We define

$$\bar{Q}(z) = \mathrm{diag}\,[z^{\nu+\mu-\mu_i}]\, Q(D)\Big|_{D=z^{-1}} \qquad \bar{P}^*(z) = P^*(D)\Big|_{D=z^{-1}} \mathrm{diag}\,[z^{\mu_j}]$$

$$\bar{H}(z) = \mathrm{diag}\,[z^{\nu+\mu-\mu_i}]\, H(D)\Big|_{D=z^{-1}} \qquad \bar{K}(z) = \mathrm{diag}\,[z^{\nu+\mu-\mu_i}]\, K(D)\Big|_{D=z^{-1}}$$

$$\bar{J}(z) = \mathrm{diag}\,[z^{\nu+\mu-\mu_i}]\, J(D)\Big|_{D=z^{-1}} \qquad \bar{S}(z) = \mathrm{diag}\,[z^{\nu+\mu-\mu_i}]\, S(D)\Big|_{D=z^{-1}}$$

$$\bar{R}(z) = R(D)\Big|_{D=z^{-1}} \mathrm{diag}\,[z^{\mu_j}] \qquad \bar{P}(z) = P(D)\Big|_{D=z^{-1}} \mathrm{diag}\,[z^{\mu_j}] \tag{5.10}$$

Note that all these matrices are polynomial matrices in z. The constraints on the degrees of $H(D)$, $K(D)$, $J(D)$ and $S(D)$ in (5.8) may be shown to be *equivalent* to the following constraints on $\bar{H}(z)$, $\bar{K}(z)$, $\bar{J}(z)$ and $\bar{S}(z)$

$$\partial_{cj}(\bar{H}(z)) \leq \nu_j - 1 \qquad \partial_{cj}(\bar{J}(z)) \leq \nu_j - 1 \tag{5.11}$$

$$\partial_{ri}(\bar{K}(z)) \leq \nu + \mu - \mu_i \qquad \partial_{cj}(\bar{K}(z)) \leq \nu + \mu - \mu_j - 1 \tag{5.12}$$

$$\partial_{ri}(\bar{S}(z)) \le \nu + \mu - \mu_i \qquad \partial_{cj}(\bar{S}(z)) \le \nu + \mu - \mu_j - 1 \tag{5.13}$$

The constraints on $Q(D)$, $P^*(D)$ are equivalent to

$$\begin{aligned} &\bar{Q}(z) = \text{diag}\,[\bar{q}_j(z)] \quad , \quad \deg[\bar{q}_j(z)] = \nu + \mu - \mu_j \quad , \quad \bar{q}_j(0) \ne 0 \\ &\bar{P}^*(z) = \text{diag}\,[\bar{p}_j^*(z)] \quad , \quad \deg[\bar{p}_j^*(z)] = \mu_j \quad , \quad \bar{p}_j^*(0) \ne 0 \end{aligned} \tag{5.14}$$

provided that $\bar{q}_i(z)$, $\bar{p}_i(z)$ are *monic* polynomials with zeros inside the unit circle. With these definitions, (5.4) and (5.5) are equivalent to

$$\bar{H}(z)\,\bar{R}(z) + \bar{K}(z)\,\bar{P}(z) = \bar{Q}(z)\,(\bar{P}(z) - \bar{P}^*(z)) \tag{5.15}$$

$$\bar{J}(z)\,\bar{R}(z) + \bar{S}(z)\,\bar{P}(z) = z^{\nu+\mu} I - \text{diag}\,[z^{\nu+\mu-\mu_i}]\,\bar{P}(z) \tag{5.16}$$

while the design identity (5.7) is

$$\begin{aligned} (z^{\nu+\mu}\bar{H}(z) + \bar{Q}(z)\,\bar{P}^*(z)\,\bar{J}(z))\,\bar{R}(z) = {}& (-z^{\nu+\mu}\bar{K}(z) - \bar{Q}(z)\,\bar{P}^*(z)\,\bar{S}(z) \\ & + z^{\nu+\mu}\bar{Q}(z) - \bar{Q}(z)\,\bar{P}^*(z)\,\text{diag}\,[z^{\nu+\mu-\mu_j}])\,\bar{P}(z) \end{aligned} \tag{5.17}$$

>From the properties of $R(D)$, $P(D)$, it follows that $\bar{R}(z)$, $\bar{P}(z)$ are right coprime, with $\partial_{cj}(\bar{R}(z)) \le \mu_j$, $\partial_{cj}(\bar{P}(z)) = \mu_j$ and $\Gamma_c(\bar{P}(z)) = P(0)$ is non singular (where $\Gamma_c(\bar{P}(z))$ denotes the matrix whose *j*–th column contains the coefficients of z^{μ_j} in the *j*–th column of $\bar{P}(z)$). It is a remarkable fact (*cf.* [10]) that there exists a *canonical* pair $(\bar{R}(z), \bar{P}(z))$, such that $\bar{P}(z)$ satisfies

$$\begin{aligned} &\partial_{cj}(\bar{P}(z)) = \mu_j \qquad \left[\Gamma_c(\bar{P}(z)) - I\right]_{ij} = 0 \qquad i \ge j \\ &\partial_{ri}(\bar{P}(z)) = \mu_i \qquad \Gamma_r(\bar{P}(z)) - I = 0 \end{aligned} \tag{5.18}$$

Note that the canonical $\bar{P}(z)$ is not only column-reduced but also row-reduced. Similarly, there exists a *canonical left matrix fraction description* $(\tilde{P}(z), \tilde{R}(z))$ such that

$$\tilde{P}(z)\,\bar{R}(z) = \tilde{R}(z)\,\bar{P}(z) \tag{5.19}$$

and $\tilde{P}(z)$ satisfies

$$\begin{aligned} &\partial_{ri}(\tilde{P}(z)) = \nu_i \qquad \left[\Gamma_r(\tilde{P}(z)) - I\right]_{ij} = 0 \qquad j \ge i \\ &\partial_{cj}(\tilde{P}(z)) = \nu_j \qquad \Gamma_c(\tilde{P}(z)) - I = 0 \end{aligned} \tag{5.20}$$

with $\partial_{ri}(\tilde{R}(z)) \le \nu_i$. With these preliminaries, we are ready to proceed with the proof of theorem 5.1.

Existence:

We first show that there exists a solution that satisfies (5.11). This result is available in the literature [11], but we give here a brief proof for the sake of completeness. Since $\bar{R}(z)$ and $\bar{P}(z)$ are right coprime, there exist matrices $\bar{U}(z)$ and $\bar{V}(z)$ such that

$$\bar{U}(z)\,\bar{R}(z) + \bar{V}(z)\,\bar{P}(z) = I \tag{5.21}$$

The general solution of (5.15) is of the form

$$\begin{aligned} \bar{H}(z) &= \bar{Q}(z)\,(\bar{P}(z) - \bar{P}^{*}(z))\,\bar{U}(z) + \bar{Q}_1(z)\,\tilde{P}(z) \\ \bar{K}(z) &= \bar{Q}(z)\,(\bar{P}(z) - \bar{P}^{*}(z))\,\bar{V}(z) - \bar{Q}_1\,\tilde{R}(z) \end{aligned} \tag{5.22}$$

and the solution of (5.16) is

$$\begin{aligned} \bar{J}(z) &= \left[\, z^{\nu+\mu} I - \mathrm{diag}\,[z^{\nu+\mu-\mu_i}]\,\bar{P}(z) \right] \bar{U}(z) + \bar{Q}_2(z)\,\tilde{P}(z) \\ \bar{S}(z) &= \left[\, z^{\nu+\mu} I - \mathrm{diag}\,[z^{\nu+\mu-\mu_i}]\,\bar{P}(z) \right] \bar{V}(z) - \bar{Q}_2(z)\,\tilde{R}(z) \end{aligned} \tag{5.23}$$

where $\bar{Q}_1(z)$ and $\bar{Q}_2(z)$ are arbitrary $(m \times p)$ polynomial matrices. From the polynomial matrix division theorem (*cf.* [9, p. 389], [4, p. 282]), there exist matrices $\bar{Q}_1(z)$ and $\bar{Q}_2(z)$ such that

$$\partial_{cj}(\bar{H}(z)) \le \nu_j - 1 \qquad \partial_{cj}(\bar{J}(z)) \le \nu_j - 1 \tag{5.24}$$

It follows that $\bar{H}(z)$ and $\bar{J}(z)$ satisfy the degree constraints (5.11). Concerning the degree constraints on $\bar{K}(z)$, we multiply (5.15) on the left by $\mathrm{diag}\,[z^{-(\nu+\mu-\mu_i)}]$ and on the right by $\mathrm{diag}\,[z^{-\mu_j}]$ to obtain

$$\begin{aligned} &\mathrm{diag}\,[z^{-(\nu+\mu-\mu_i)}]\,\bar{H}(z)\cdot\bar{R}(z)\,\mathrm{diag}\,[z^{-\mu_j}] + \mathrm{diag}\,[z^{-(\nu+\mu-\mu_i)}]\,\bar{K}(z)\cdot\bar{P}(z)\,\mathrm{diag}\,[z^{-\mu_j}] \\ &\quad = \mathrm{diag}\,[z^{-(\nu+\mu-\mu_i)}]\,\bar{Q}(z)\cdot(\bar{P}(z) - \bar{P}^{*}(z))\,\mathrm{diag}\,[z^{-\mu_j}] \end{aligned} \tag{5.25}$$

Since $\partial_{ri}(\bar{H}(z)) \le \nu - 1 < \nu + \mu - \mu_i$, and using the properties of $\bar{R}(z)$, $\bar{P}(z)$, $\bar{P}^{*}(z)$, $\bar{Q}(z)$, it follows that

$$\lim_{z\to\infty}\ \mathrm{diag}\,[z^{-(\nu+\mu-\mu_i)}]\,\bar{K}(z) = I - (\,\Gamma_c(\,\bar{P}(z))\,)^{-1} < \infty \tag{5.26}$$

and therefore

$$\partial_{ri}(\overline{K}(z)) \le \nu + \mu - \mu_i \tag{5.27}$$

The other constraint on $\overline{K}(z)$ is obtained by multiplying (5.15) on the left by $z^{-(\nu+\mu)}$

$$z^{-(\nu+\mu)} \cdot \overline{H}(z)\,\overline{R}(z) + \overline{K}(z)\,\mathrm{diag}\,[z^{-(\nu+\mu-\mu_j)}] \cdot \mathrm{diag}\,[z^{-\mu_i}]\,\overline{P}(z)$$
$$= \mathrm{diag}\,[z^{-(\nu+\mu-\mu_i)}]\,\overline{Q}(z) \cdot \mathrm{diag}\,[z^{-\mu_i}]\,(\overline{P}(z) - \overline{P}^{\,*}(z)) \tag{5.28}$$

where we used the fact that $\overline{Q}(z)$ is diagonal and that the product of diagonal matrices commutes. Since $\partial_{ri}(\overline{H}(z)\,\overline{R}(z)) \le \nu + \mu - 1$, it follows that

$$\lim_{z \to \infty} \overline{K}(z)\,\mathrm{diag}\,[z^{-(\nu+\mu-\mu_j)}] = I - (\Gamma_r(\overline{P}(z)))^{-1} = 0 \tag{5.29}$$

and therefore

$$\partial_{cj}(\overline{K}(z)) \le \nu + \mu - \mu_j - 1 \tag{5.30}$$

The proof for the constraints on $\overline{J}(z)$ follows along identical lines.

Uniqueness:

To prove uniqueness, we first establish that (5.17) can be satisfied *only* if (5.15) and (5.16) are satisfied (the converse being obvious). Rewrite (5.17) as

$$z^{\nu+\mu}\left[\overline{H}(z)\,\overline{R}(z) + \overline{K}(z)\,\overline{P}(z) - \overline{Q}(z)\,(\overline{P}(z) - \overline{P}^{\,*}(z))\right]$$
$$= -(\overline{Q}(z)\,\overline{P}^{\,*}(z))\left[\overline{J}(z)\,\overline{R}(z) + \overline{S}(z)\,\overline{P}(z) - z^{\nu+\mu} I + \mathrm{diag}\,[z^{\nu+\mu-\mu_i}]\,\overline{P}(z)\right] \tag{5.31}$$

>From the degree conditions and the properties of $\overline{R}$, $\overline{P}$, note that $\partial\overline{J}_{ik}(z) \le \nu_k - 1 \le \nu - 1$, $\partial\overline{R}_{kj}(z) \le \mu_j \le \mu$, $\partial\overline{S}_{ik}(z) \le \nu + \mu - \mu_k - 1$, $\partial\overline{P}_{kj}(z) \le \mu_k$. Further, $\partial(z^{\nu+\mu-\mu_i}\,\overline{P}_{ik}(z) - z^{\nu+\mu}) \le \nu+\mu-1$. It follows that the maximal degree of any element in the right bracket in (5.31) is $\nu + \mu - 1$. However, the elements on the left side have $\nu + \mu$ zeros at $z = 0$. and $\overline{Q}(z)\,\overline{P}^{\,*}(z)$ is a diagonal matrix with elements that have no zeros at $z = 0$. Therefore, (5.31) can only be valid if both sides are equal to zero, *i.e.*, if both (5.15) and (5.16) are satisfied.

Now, assume that there existed another solution $\overline{H}(z) + \delta\overline{H}(z)$, $\overline{K}(z) + \delta\overline{K}(z)$, $\overline{J}(z) + \delta\overline{J}(z)$, $\overline{S}(z) + \delta\overline{S}(z)$. The following homogeneous equations would have to be satisfied

$$\delta\overline{H}(z)\,\overline{R}(z) + \delta\overline{K}(z)\,\overline{P}(z) = 0$$
$$\delta\overline{J}(z)\,\overline{R}(z) + \delta\overline{S}(z)\,\overline{P}(z) = 0 \tag{5.32}$$

Since $\bar{R}(z)$ and $\bar{P}(z)$ are coprime, and since $\partial_{cj}(\delta \bar{H}(z)) \le \nu_j - 1$ and $\partial_{cj}(\delta \bar{J}(z)) \le \nu_j - 1$, this implies (*cf.* [12]) that $\delta \bar{H}(z) = \delta \bar{K}(z) = \delta \bar{J}(z) = \delta \bar{S}(z) = 0$. □

Expressions for the Structured Nonminimal Model

We now show how the model (5.6) can be put in the form (2.3). The matrices in (5.6) can be written as follows

$$H(D) = \sum_{k=1}^{\nu+\mu-\mu_{\min}} H_k D^k \quad , \quad K(D) = \sum_{k=0}^{\nu+\mu-\mu_{\min}} K_k D^k$$

$$J(D) = \sum_{k=0}^{\nu+\mu-\mu_{\min}} J_k D^k \quad , \quad S(D) = \sum_{k=0}^{\nu+\mu-\mu_{\min}} S_k D^k \tag{5.33}$$

where $H_k, J_k \in \mathbb{R}^{m \times p}$, $K_k, S_k \in \mathbb{R}^{m \times m}$ and $\mu_{\min} = \min_{1 \le j \le m} \{ \mu_j \}$. Let $\gamma = \nu + \mu - \mu_{\min}$. Substitution of (5.33) in (5.6) yields the following parameterization for (2.3)

$$C(D) = 0 \quad , \qquad E(D) = Q(D)\,[I - P^*(D)] \tag{5.34}$$

$$B_0(D) = -I \quad , \quad \beta_0 = K_0 \tag{5.35}$$

while, for $j = 1, 2, \ldots, \gamma$:

$$A_j(D) = D^j \cdot I \;\;,\;\; \alpha_j = H_j \;\;,\;\; B_j(D) = -D^j \cdot I \;\;,\;\; \beta_j = K_j \tag{5.36}$$

and for $j = \gamma + 1, \ldots, 2\gamma + 1$:

$$A_j(D) = Q(D)\,P^*(D)\,D^{j-\gamma-1} \quad , \quad \alpha_j = J_{j-\gamma-1}$$

$$B_j(D) = Q(D)\,P^*(D)\,D^{j-\gamma-1} \quad , \quad \beta_j = S_{j-\gamma-1} \tag{5.37}$$

with $m_a = m_b = 2\gamma + 1$, $l = \nu + \mu + \gamma = 2\nu + 2\mu - \mu_{\min}$, $r = m$.

The vectors $\bar{\theta}_i^*$, for $1 \le i \le m$, are given by

$$\bar{\theta}_i^* = [H_{i1}, \ldots, H_{i\gamma}, J_{i0}, \ldots, J_{i\gamma},$$

$$K_{i0}, \ldots, K_{i\gamma}, \tag{5.38}$$

where H_{ik}, J_{ik}, K_{ik} and S_{ik} are the i–th row of the matrices H_k, J_k, K_k and S_k (which are defined in (5.33)). The vectors $\bar{\phi}_i(t)$, $1 \le i \le m$ are given by

$$\bar{\phi}_i^T(t) = [y^T(t-1), \ldots, y^T(t-\gamma), q_i(D)\,p_i^*(D)\,y^T(t-\gamma),$$

$$u^T(t), \ldots, u^T(t-\gamma), q_i(D)\,p_i^*(D)\,u^T(t-\gamma)] \tag{5.39}$$

where $q_i(D)$ and $p_i^*(D)$ are the polynomials defined in theorem 5.1. Each of the parameter vectors θ_i^*, $1 \le i \le m$, is obtained from $\bar{\theta}_i^*$ by deleting the elements that, according to the conditions in (5.8), are zero. In the same way, we obtain the vectors $\phi_i(t)$ from $\bar{\phi}_i(t)$, $1 \le i \le m$.

6. Stability and Convergence Properties

We now show that the general theorems of sections 3 and 4 apply to the adaptive pole placement scheme. >From theorem 3.1 and theorem 5.1, it follows that all m associated–signal systems are output–reachable, provided that the degree conditions in (5.8) are satisfied. By using theorem 4.2, we obtain that all sequences $\{ \phi_i(t) \}$, $1 \le i \le m$ (which are the associated–signal systems outputs) are persistently exciting, provided that the external input sequence $\{ \mathrm{v}(t) \}$ satisfies condition (4.18). We only need to show that the adaptive version of the control law (5.3) is of the form of the control law in theorem 4.2.

The adaptive control law is given by

$$\begin{aligned} u(t) &= Q^{-1}(D)\,[\,H(D,\,t)\,y(t) + K(D,\,t)\,u(t)\,] + \mathrm{v}(t) \\ &= H(D,\,t)\,y(t) + K(D,\,t)\,u(t) + Q(D)\,\mathrm{v}(t) - (Q(D) - I)\,u(t) \end{aligned} \tag{6.1}$$

where $H(D,\,t)$ and $K(D,\,t)$ are the estimates of $H(D)$ and $K(D)$ at time t. For analysis purposes, it is useful to express (6.1) row by row, using (2.2)

$$\begin{aligned} u_i(t) = [H_i(D,\,t)\,R(D) + K_i(D,\,t)\,P(D) - (q_i(D) - 1)\,P_i(D)]\,\xi(t) \\ + q_i(D)\,\mathrm{v}_i(t) \quad , \quad i = 1, 2, \ldots, m \end{aligned} \tag{6.2}$$

where $u_i(t)$ and $\mathrm{v}_i(t)$ are the i–th component of $u(t)$ and $\mathrm{v}(t)$, $H_i(D,\,t)$, $K_i(D,\,t)$ and $P_i(D)$ are the i–th row of $H(D,\,t)$, $K(D,\,t)$ and $P(D)$, and where $q_i(D)$ are polynomials defined in theorem 5.1. We will use the following definitions

$$\begin{aligned} H_i(D,\,t)\,R(D) &= \sum_{k=1}^{\nu + 2\mu - \mu_i} \\ K_i(D,\,t)\,P(D) &= \sum_{k=1}^{\nu + 2\mu - \mu_i} M_{ik}(t)\,D^k + K_i(0,\,t)\,P(D) \\ (q_i(D) - 1)\,P_i(D) &= \sum_{k=1}^{\nu + 2\mu - \mu_i} N_{ik}\,D^k \end{aligned} \tag{6.3}$$

where $L_{ik}(t)$, $M_{ik}(t)$, $N_{ik} \in \mathbb{R}^{m \times m}$. Substitution of (6.3) in (6.2) yields

$$u_i(t) = \sum_{k=1}^{\nu + 2\mu - \mu_i} [(L_{ik}(t) + M_{ik}(t) - N_{ik})\, D^k + K_i(0, t)\, P(D)]\, \xi(t) + q_i(D)\, v_i(t)$$

$$= \sum_{k=1}^{\nu + 2\mu - \mu_i} [(L_{ik}(t) + M_{ik}(t) - N_{ik})\, \xi(t-k) + K_i(0, t)\, u(t) + q_i(D)\, v_i(t) \qquad (6.4)$$

for $i = 1, 2, \ldots, m$. Equation (6.4) can be written as follows

$$u_i(t) = \overline{F}_i(t)\, x_a(t) + K_i(0, t)\, u(t) + q_i(D)\, v_i(t) \qquad (6.5)$$

where $x_a(t)$ is the state vector of the associated–signal systems defined in (3.1). In this case, $l = 2\nu + 2\mu - \mu_{\min}$, $n_a = \dim[x_a] = m(l + \mu) = m(2\nu + 3\mu - \mu_{\min})$. $\overline{F}_i(t)$ is the following $(1 \times m(2\nu + 3\mu - \mu_{\min}))$ row vector

$$\overline{F}_i(t) = [L_{i1}(t) + M_{i1}(t) - N_{i1},\ L_{i2}(t) + M_{i2}(t) - N_{i2}, \ldots,$$

$$L_{i\nu + 2\mu - \mu_i}(t) + M_{i\nu + 2\mu - \mu_i}(t) - N_{i\nu + 2\mu - \mu_i},\ 0, \ldots, 0] \qquad (6.6)$$

By writing the m equations (6.5) for $i = 1, 2, \ldots, m$ we obtain

$$u(t) = \overline{F}(t)\, x_a(t) + K(0, t)\, u(t) + Q(D)\, v(t) \qquad (6.7)$$

where

$$\overline{F}(t) = [F_1^{\,T}(t), F_2^{\,T}(t), \ldots, F_m^{\,T}(t)]^T \qquad (6.8)$$

>From the constraints on $K_i^{\,j}(D)$ in (5.8), it follows that the matrix $[I - K(0, t)]$ is upper triangular and has a unit diagonal for all t. Therefore, its inverse always exists and is also upper triangular with unit diagonal. Hence, we can write

$$u(t) = F(t)\, x_a(t) + [I - K(0, t)]^{-1} Q(D)\, v(t) \qquad (6.9)$$

where

$$F(t) = [I - K(0, t)]^{-1}\, \overline{F}(t) \qquad (6.10)$$

In fact, it can be shown that $F(t)$ is given by

$$F(t) = [I - K(0, t)]^{-1} \begin{bmatrix} \theta_1^{\,T}(t) & & & 0 \\ & \theta_2^{\,T}(t) & & \\ & & \ddots & \\ 0 & & & \theta_m^{\,T}(t) \end{bmatrix} \qquad (6.11)$$

bold cdot ˜˜ T sub 1 minus T sub 2 ˜˜ right }

where $\theta_i(t)$ are the parameter vectors estimates of θ_i^* (*cf.* (2.8)–(2.10) and after) and where T_1 and T_2 are *fixed* real matrices, which depend on the elements of $R(D)$, $P(D)$ and $Q(D)$.

Let $\{ t_i \}$ be a sequence of integers such that $t_i = i\,N, \quad i = 0, 1, 2, \ldots$ where N is a positive integer to be determined later. The feedback gain matrix is held constant during each period of length N and the adaptive control law is *modified* so that

$$u(t) = F_N(t)\, x_a(t) + w(t) \qquad (6.12)$$

where

$$F_N(t) = F(t_i) \qquad \text{for} \quad t_i \le t < t_{i+1} \qquad (6.13)$$

$$w(t) = [I - K(0, t_i)]^{-1}\, \bar{v}(t) \qquad \text{for} \quad t_i \le t < t_{i+1} \qquad (6.14)$$

$$\bar{v}(t) = Q(D)\, v(t) \qquad (6.15)$$

It is known (see [2]) that, with a RLS algorithm with covariance resetting, the estimates $\theta_i(t)$ remain in a bounded region of the parameter space. By the triangular property of $K(0, t_i)$, it follows that $(I - K(0, t_i))^{-1}$ is bounded. The matrices T_1 and T_2 are fixed so that, by (6.11), $F_N(t)$ remains bounded. The sequence $\{ w(t) \}$ in (6.14) depends on the value of $K(0, t_i)$ but since $[I - K(0, t)]$ is bounded, it follows that if

$$\lambda_{\min}\, [\bar{V}_{n_a+1,\, N-n_a}(t_i)\, \bar{V}^T_{n_a+1,\, N-n_a}(t_i)] > \varepsilon_1 > 0 \qquad (6.16)$$

then

$$\lambda_{\min}\, [W_{n_a+1,\, N-n_a}(t_i)\, W^T_{n_a+1,\, N-n_a}(t_i)] > \varepsilon > 0 \qquad (6.17))$$

$Q(D)$ being fixed, the external input sequence $\{ v(t) \}$ must be chosen so that (6.16) will be satisfied. Using theorem 4.2, we obtain the following proposition that summarizes the results.

Proposition 5.1

Consider a linear minimal system (2.1). Assume that the observability indices $\nu_i, \quad 1 \le i \le p$ and the controllability indices $\mu_i, \quad 1 \le i \le m$ are known. Let $Q(D)$ and $P^*(D)$ be defined as in theorem 5.1. Define m estimation equations of the form (2.11), for matrices $H(D)$, $K(D)$, $J(D)$ and $S(D)$ that satisfy the degree constraints in (5.8). Every parameter vector $\theta_i^*, \quad 1 \le i \le m$ is estimated with a RLS algorithm with covariance resetting. The adaptive control law is given by

$$u(t) = Q^{-1}(D)\,[\,H_N(D,\,t)\,y(t) + K_N(D,\,t)\,u(t) + Q(D)\,\mathrm{v}(t)\,] \tag{6.18}$$

where $H_N(D,\,t)$ and $K_N(D,\,t)$ are the estimates of the matrices $H(D)$ and $K(D)$, updated periodically so that

$$H_N(D,\,t) = H(D,\,t_i)\,,\quad K_N(D,\,t) = K(D,\,t_i)\quad \text{for}\quad t_i \le t < t_{i+1}$$

where $t_i = i\,N$, $i = 0, 1, 2, \ldots$ and $N \ge m(2\nu + 3\mu - \mu_{\min})(m+1) + m$. The external input sequence $\{\,\mathrm{v}(t)\,\}$ satisfies (6.16) (where $\bar{v}(t)$ is defined in (6.15)).

Then, the transfer matrix of the closed–loop system converges exponentially fast to $T_{cl}(D) = R(D)\,P^{*^{-1}}(D)$, for every initial state of the system and for all initial conditions of the estimation algorithm.

7. Conclusions

In this paper, we showed how a multivariable adaptive pole placement algorithm could be designed so that parameter convergence is guaranteed under persistency of excitation conditions. More generally, it was proved that parameter convergence would follow provided that a certain design identity was satisfied, so that the results of this paper are applicable to a wide range of adaptive control algorithms.

An advantage of parameter convergence is that the closed–loop system asymptotically has the properties for which the controller was designed. In particular, the scheme presented here does not have the uncertainty of a matrix $U(D)$ found in [8] (present even with persistently exciting signals). On the other hand, more prior information is needed, that is, the observability indices have to be known in addition to the controllability indices. While the persistency of excitation conditions were used to assess stability, it is known that such conditions are not necessary to prove stability in adaptive control, but only to prove exponential convergence of the parameters to the nominal values (*cf.* [2], [4], and [13], [14] specifically for adaptive pole placement algorithms). The results of this paper may also be related to the work of [15], which discusses the minimum value of N in theorem 4.2 for the SISO case, and to the work of [7], which transforms the persistency of excitation condition into a condition on the number of spectral components of the inputs (sufficient richness condition) in the case of multivariable identification. Special signals such that the persistency of excitation condition is satisfied were also investigated in [16].

8. Acknowledgements

Marc Bodson gratefully acknowledges the support of the National Science Foundation under grant ECS 88–10145 and the support of the Lady Davis Foundation at the Technion, Israel. Michael Heymann's work was supported by the Technion Fund for Promotional Research. The authors would like to thank the anonymous reviewers for their constructive comments.

References

[1] HEYMANN M., *Persistency of Excitation Results for Structured Nonminimal Models,* IEEE Trans. on Automatic Control, 33 (1988), pp. 112–116.

[2] GOODWIN G.C. and SIN K.S., *Adaptive Filtering, Prediction and Control,* Prentice–Hall, Englewood Cliffs, New Jersey, 1984.

[3] BOYD S. and SASTRY S., *Necessary and Sufficient Conditions for Parameter Convergence in Adaptive Control,* Automatica, 22 (1986), pp. 629–639.

[4] SASTRY S. and BODSON M., *Adaptive Control: Stability, Convergence, and Robustness,* Prentice–Hall, Englewood Cliffs, New Jersey, 1989.

[5] ANDERSON B.D.O. and JOHNSTONE R.M., *Adaptive Systems and Time–Varying Plants,* Int. J. Control, 37 (1983), pp. 367–377.

[6] ANDERSON B.D.O., *Adaptive Systems, Lack of Persistency of Excitation and Bursting Phenomena,* Automatica, 21 (1985), pp. 247–258.

[7] DE MATHELIN M. and BODSON M., *Frequency Domain Conditions for Parameter Convergence in Multivariable Recursive Identification,* Automatica, (1990).

[8] ELLIOTT H., WOLOVICH W.A. and DAS M., *Arbitrary Adaptive Pole Placement for Linear Multivariable Systems,* IEEE Trans. on Autom. Control, 29 (1984), pp. 221–229.

[9] KAILATH T., *Linear Systems,* Prentice-Hall, Englewood Cliffs, New Jersey, 1980.

[10] BEGHELLI S. and GUIDORZI R., *A New Input–Output Canonical Form for Multivariable Systems,* IEEE Trans. on Autom. Control, 21 (1976), pp. 692–696.

[11] WOLOVICH W.A. and ANTSAKLIS P.J., *The Canonical Diophantine Equations with Applications,* 22 (1984), pp. 777–787.

[12] FORNEY G.D., *Minimal Bases of Rational Vector Spaces with Applications to Multivariable Linear Systems,* SIAM J. Control, 13 (1975), pp. 493–520.

[13] POLDERMAN J.W., *Adaptive Control & Identification: Conflict or Conflux?* Ph.D. Dissertation, University of Groeningen, 1987.

[14] LOZANO–LEAL R. and GOODWIN G.C., *A Globally Convergent Adaptive Pole Placement Algorithm Without a Persistency of Excitation Requirement,* IEEE Trans. on Autom. Control, 30 (1985), pp. 795–798.

[15] FEUER A. and HEYMANN M., *On Minimum Spanning Blocks in Discrete Linear Systems,* IEEE Trans. on Autom. Control, 31 (1986), pp. 352–355.

[16] WILLNER J., *Convergence of Parameters in Estimation and Adaptive Control of Multivariable Systems,* Master Thesis (in Hebrew), Dept of Electrical Engineering, Technion, Israel, 1986.

[17] HEYMANN M., LEWIS J.H. and MEYER G., ‘‘Remarks on the Adaptive Control of Linear Plants with Unknown High–Frequency Gain,’’ *Systems and Control Letters,* vol. 5, pp. 357–362, 1985.

[18] KUMAR P.R., ‘‘Convergence of Adaptive Control Schemes Using Least–Squares Parameter Estimates,’’ *IEEE Trans. on Automatic Control*, vol. 35, no. 4, pp. 416–424, 1990.