Identification of Linearly Overparametrized Nonlinear Systems

G. Bastin, R. R. Bitmead, G. Campion, and M. Gevers

Abstract—**Often, a dynamical model is nonlinear in the unknown parameters, but it can be transformed into an overparametrized linear regression model, where the components of the overparametrization vector are nonlinear functions of the smaller number of unknown parameters. We present an algorithm that directly identifies the unknown parameters, we characterize the convergence domains under two different sets of assumptions on the excitation of the signals, and we compute the corresponding convergence rates.**

I. Introduction—Statement of the Problem

In many practical modeling and control applications, a partial prior knowledge of the structure and the parametrization of the system is available. A typical situation is where the only unknowns of the system are the values of a few physical parameters which enter linearly and/or nonlinearly in the model. In such a situation, it is clear that an approach to the parameter estimation problem which ignores the prior knowledge is questionable since it would necessarily result in an attempt to estimate more parameters than necessary. This is the reason why the issue of incorporating prior knowledge on the parametrization in the parameter estimation problem has recently received some attention.

In the case where the unknown parameters enter linearly in the process model, the solution is obviously to reformulate the problem in the form of a linear regression limited to those parameters. However, the practical implementation is not trivial and is discussed in [1], [2], and [3].

In this note we consider the more complex situation where the unknown parameters enter nonlinearly in the model but can be embedded in a linear over-reparametrization to be made explicit short in (1.1). This issue has been previously discussed in a series of papers by Dasgupta, Anderson, and Kay [4]–[6] for single-input single-output (SISO) systems where the reparametrization is a polynomial function of the unknown parameters. Here we shall be concerned with multivariable nonlinear systems, where the reparametrization is any nonlinear function of the unknown parameters.

The systems under consideration are assumed to be expressed in the following nonlinear regression form:

$$y(t) = \varphi^T(t)\beta(\theta) \tag{1.1}$$

where $t \in R_+$, $y \in R^m$ is a vector observation sequence, $\varphi \in R^k \times R^m$ is a regression matrix made up of known signals, $\theta \in R^n$ is the unknown parameter vector, and $\beta(\cdot)$ is a nonlinear mapping from R^n onto a subset of R^k, with $k \geq n$.

It is to be noticed that the vector β constitutes an "over-reparametrization" of the system which enters linearly in the model (1.1).

The problem is to estimate θ from measurements of y and φ.

Manuscript received September 28, 1989; revised September 9, 1990 and February 28, 1991. Paper recommended by Past Associate Editor, G. Verghese.

G. Bastin, G. Campion, and M. Gevers are with the Laboratoire d'Automatique, Dynamique et Analyse des Systèmes, Catholic University of Louvain, Louvain-la-Neuve, Belgium.

R. R. Bitmead is with the Department of Systems Engineering, Australian National University, Canberra, Australia.

IEEE Log Number 9107974.

Outline of the Paper

The paper is organized as follows. In Section II, we state the technical assumptions on the problem structure which will be used subsequently in the analysis. These assumptions concern the structure of the overparametrization mapping $\beta(\theta)$ on the one hand, and the excitation content of the regressor $\varphi(t)$ on the other hand. On this basis, the difference between our approach and that of Dasgupta *et al.* [4]–[6] is emphasized. A gradient algorithm for the estimation of the parameters is presented in Section III, and a Lipshitz condition relative to the dynamics of the estimation error is established. The main convergence results are demonstrated under two different assumptions on the excitation content of φ, in Sections IV and V, respectively. In each case an upper bound for the adaptation gain and a lower bound on the size of the convergence domain are calculated, and their connection with the structure of the overparametrization mapping $\beta(\theta)$ is discussed.

II. Assumptions

In this section, we formulate a set of technical assumptions on the structure of the nonlinear reparametrization $\beta(\theta)$ and on the excitation content of the regressor $\varphi(t)$. These assumptions will be used later in the analysis.

A. Assumption on the Structure of $\beta(\cdot)$

A.1: The function $\beta(\cdot)$ maps an open ball $B_\theta \in R^n$ of radius r, centered on θ^*, onto a set $B_\beta \in R^k$, with $k \geq n$, such that:

- $\beta(\theta)$ is a C^2 function, i.e., its derivatives w.r.t. θ up to order 2 exist and are continuous;
- $\partial\beta/\partial\theta$ has full rank n on B_θ.

In particular, there exist finite constants $k_1 > 0$ and $k_2 > 0$ such that (unless otherwise indicated, all norms are 2-norms throughout the note)

$$\left\| \frac{\partial \beta_i}{\partial \theta} \right\| \leq k_1 \quad \text{and} \quad \left\| \frac{\partial}{\partial \theta}\left(\frac{\partial \beta_i}{\partial \theta_j}\right) \right\| \leq k_2$$
$$i = 1, k \quad j = 1, n \quad \forall \theta \in B_\theta. \tag{2.1}$$

B. Notation

For vector functions $\beta: R^n \to R^k$, we denote by $\partial\beta/\partial\theta$ the $k \times n$ matrix whose (i, j)th element is

$$\left(\frac{\partial \beta}{\partial \theta}\right)_{i,j} \triangleq \frac{\partial \beta_i}{\partial \theta_j}.$$

We also use the notations of Monsieur Dieudonné for the partial derivatives of order 1 and 2:

$$D_i\beta(\theta) \equiv \left(\frac{\partial \beta}{\partial \theta_i}\right)_\theta \qquad i = 1, \cdots, n$$

$$D_{ij}^2\beta(\theta) \equiv D_i(D_j\beta)(\theta) \qquad i, j = 1, \cdots, n.$$

C. Assumptions on the Regressor $\varphi(t)$

We shall make a uniform boundedness and an excitation assumption about the regressor φ. The boundedness assumption is simply as follows.

A.2:

$$\|\varphi(t)\| \leq \varphi_{\max} \qquad \forall t \in R_+.$$

As for the excitation, we shall state here two alternative assumptions, a strong assumption A.3 and a weaker assumption A.3′. Our convergence proof will follow two different routes and will lead to

0018-9286/92$03.00

two different convergence domains, depending on whether the stronger or the weaker assumption is used.

A.3: There exists $\delta_1 > 0$, $T > 0$, and $t_0 \geq 0$ such that

$$P(\theta, t) \equiv \left(\frac{\partial\beta}{\partial\theta}\right)_\theta^T \left[\int_t^{t+T} \varphi(\tau)\varphi^T(\tau)\, d\tau\right]\left(\frac{\partial\beta}{\partial\theta}\right)_\theta \geq \delta_1 I \qquad \forall t \geq t_0, \quad \forall \theta \in B_\theta.$$

A.3′: There exists $\delta_2 > 0$, $T > 0$, and $t_0 \geq 0$ such that

$$\bar{P}(t) \equiv \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta^*}^T \left[\int_t^{t+T} \varphi(\tau)\varphi^T(\tau)\, d\tau\right]\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta^*} \geq \delta_2 I \qquad \forall t \geq t_0$$

where θ^* is the true value of θ.

The problem described by (1.1) could simply be viewed as a nonlinear regression problem, and handled by standard nonlinear regression techniques; see, e.g., [7]. However, with a general nonlinear regression model, not much can be said about the convergence domain and the rate of convergence. Here we have the added assumption that the problem has been reformulated as a linear regression problem, albeit with a larger number of linearly appearing β_i that are nonlinear functions of the smaller number of θ_j^*. This will allow us to make precise statements about domain and rate of convergence. This setup has been studied extensively by Dasgupta, Anderson, and Kaye in a series of papers [4]–[6] for the special case where the β_i are polynomial functions of the θ_j. A simple example would be $\theta = (\theta_1, \theta_2)$ and $\beta(\theta) = (\theta_1, \theta_2, \theta_1^2\theta_2)$. Our algorithm estimates θ directly, whereas in [4]–[6] β is estimated first as an unconstrained estimate and is subsequently modified using a least squares criterion so that the constraints imposed by the polynomial functions $\beta(\theta)$ are satisfied (e.g., $\beta_3 = \beta_1^2\beta_2$ in the example above). Our results extend those of [4]–[6] in two ways: first, $\beta(\theta)$ is not restricted to polynomial functions of θ; second, because we do not estimate β, but the lower dimensional θ, our persistence of excitation (PE) conditions A.3 or A.3′ are much weaker than those of [4]–[6], where the whole vector $\varphi(t)$ was required to be persistently exciting. Here we only require $P(\theta, t)$ [respectively, $\bar{P}(t)$] to be positive definite: its size, $n \times n$, is typically much smaller than the dimension $k \times k$ of $\varphi(t)\varphi^T(t)$. The penalty we pay for these extensions is that our results will be local, rather than global, but such is the nature of life.

III. The Estimation Algorithm

We consider the following estimation algorithm for θ, the estimate of θ^* (we drop the time index for simplicity):

$$\dot{\theta} = \omega\left(\frac{\partial\beta}{\partial\theta}\right)_\theta^T \varphi\left[y - \varphi^T\beta(\theta)\right] = \omega\psi\left[y - \varphi^T\beta(\theta)\right] \tag{3.1}$$

where $\omega > 0$ is the adaptation gain, and ψ denotes

$$\psi(\theta, t) \equiv \left(\frac{\partial\beta}{\partial\theta}\right)_\theta^T \varphi(t) \qquad \theta \in B_\theta. \tag{3.2}$$

This is a gradient algorithm for the minimization of $(y(t) - \varphi^T(t)\beta(\theta))^2$. In the next two sections, we shall analyze the convergence properties of θ under assumptions A.1–A.3 (respectively, A.1–A.3′). Before we embark on this, we derive some useful bounds and expressions for the error equation, that will be valid under both sets of assumptions.

Denoting $\tilde{\theta} \equiv \theta^* - \theta$, and replacing ψ by its expression (3.2), we have

$$\dot{\tilde{\theta}} = -\omega\left(\frac{\partial\beta}{\partial\theta}\right)_\theta^T \varphi\varphi^T\left[\beta(\theta^*) - \beta(\theta)\right]. \tag{3.3}$$

Let θ_1, θ_2 be any two points in B_θ. Then

$$\beta(\theta_2) = \beta(\theta_1) + \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta_1}(\theta_2 - \theta_1) + R(\theta_2 - \theta_1) \tag{3.4}$$

where $R(\theta_2 - \theta_1)$ contains all higher order terms. Using (3.2), (3.3), and (3.4) with $\theta_2 = \theta^*$ and $\theta_1 = \theta$, we can rewrite the error equation as

$$\dot{\tilde{\theta}} = -\omega\psi(\theta, t)\psi^T(\theta, t)\tilde{\theta} - \omega\left(\frac{\partial\beta}{\partial\theta}\right)_\theta^T \varphi\varphi^T R(\tilde{\theta}). \tag{3.5}$$

We denote

$$f(t, \tilde{\theta}) \equiv \omega\left(\frac{\partial\beta}{\partial\theta}\right)_\theta^T \varphi\varphi^T R(\tilde{\theta}) \tag{3.6}$$

and we now derive a Lipschitz bound for $f(t, \tilde{\theta})$.

Lemma 3.1: Let $f(t, \tilde{\theta})$ be defined by (3.6) and let $\tilde{\theta}_1 \equiv \theta^* - \theta_1$, $\tilde{\theta}_2 \equiv \theta^* - \theta_2$, with $\theta_1, \theta_2 \in B_\theta$. Then, under assumptions A.1, A.2, $f(t, \tilde{\theta})$ satisfies the following Lipschitz condition (we drop the dependence on t for simplicity):

$$\|f(\tilde{\theta}_1) - f(\tilde{\theta}_2)\| \leq \omega\varphi_{\max}^2 k_3 \|\tilde{\theta}_1 - \tilde{\theta}_2\| \tag{3.7}$$

with

$$k_3 = k_2 r\left[2k_1 k\sqrt{n} + k_2\sqrt{k}\, nr\right]. \tag{3.8}$$

Proof:

a)

$$R(\tilde{\theta}_1) - R(\tilde{\theta}_2) = \beta(\theta_2) - \beta(\theta_1) + \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta_2}(\tilde{\theta}_2 - \tilde{\theta}_1) + \left[\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta_2} - \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta_1}\right]\tilde{\theta}_1. \tag{3.9}$$

Consider first the sum of the first three terms of (3.9). The ith component of that vector is

$$\beta_i(\theta_2) - \beta_i(\theta_1) + \left(\frac{\partial\beta_i}{\partial\theta}\right)_{\theta_2}(\theta_1 - \theta_2) = -\frac{1}{2}(\theta_1 - \theta_2)^T\left(\frac{\partial^2\beta_i}{\partial\theta\partial\theta^T}\right)_{\theta_2 + \eta_i(\theta_1 - \theta_2)}(\theta_1 - \theta_2)$$

with $\eta_i \in [0, 1]$. Using A.1 and $\|(\theta_1 - \theta_2)\| \leq 2r$, it follows that the 2 norm of that vector is bounded by $k_2 r\sqrt{kn}\,\|(\theta_1 - \theta_2)\|$. As for the last term of (3.9), we have the following.

$$\left(\frac{\partial\beta_i}{\partial\theta_j}\right)_{\theta_2} - \left(\frac{\partial\beta_i}{\partial\theta_j}\right)_{\theta_1} = \frac{\partial}{\partial\theta}\left(\frac{\partial\beta_i}{\partial\theta_j}\right)_{\theta_1 + \gamma_{ij}(\theta_2 - \theta_1)}(\theta_2 - \theta_1)$$

with $\gamma_{ij} \in [0, 1]$ for $i = 1, \cdots, k$ and $j = 1, \cdots, n$.

Therefore, using A.1

$$\left\|\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta_2} - \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta_1}\right\|_F \leq k_2\sqrt{kn}\,\|(\theta_1 - \theta_2)\| \tag{3.10}$$

where the subscript F denotes the Frobenius norm. Hence,

$$\|R(\tilde{\theta}_1) - R(\tilde{\theta}_2)\| \leq 2k_2 r\sqrt{kn}\,\|(\theta_1 - \theta_2)\|. \tag{3.11}$$

b) It now follows from (3.6) that

$$f(\tilde{\theta}_1) - f(\tilde{\theta}_2) = \omega\left\{ \left(\frac{\partial\beta}{\partial\theta}\right)^T_{\theta_1} \varphi\varphi^T[R(\tilde{\theta}_1) - R(\tilde{\theta}_2)] + \left[\left(\frac{\partial\beta}{\partial\theta}\right)^T_{\theta_1} - \left(\frac{\partial\beta}{\partial\theta}\right)^T_{\theta_2}\right]\varphi\varphi^T R(\tilde{\theta}_2)\right\}.$$

Now the ith element of $R(\tilde{\theta})$ is

$$\tilde{\theta}^T\left(\frac{\partial^2\beta_i}{\partial\theta\partial\theta^T}\right)_{\bar{\theta}_j}\tilde{\theta} \qquad \text{for } \bar{\theta} \in [\theta, \theta^*]_i \quad \forall j.$$

Its norm is bounded by $k_2 r^2\sqrt{n}$. Therefore, using A.1 and A.2, (3.10) and (3.11) give the desired result. ♦

IV. Convergence Results Under A.1 to A.3

In this section we shall derive a bound on the initial error $\tilde{\theta}(0)$ for which asymptotic convergence of $\theta(t)$ to θ^* will be established under the assumptions A.1–A.3 with an additional constraint of slow adaptation. The slow adaptation is required to replace the PE condition of assumption A.3 by the stronger condition that $\psi(\theta, t)$ is persistently exciting for all θ in B_θ. We first establish that preliminary result.

Lemma 4.1: Consider the estimation algorithm (3.1) with the assumptions A.1–A.3. If $\theta(t) \in B_\theta$ $\forall t \in R_+$, and if

$$\omega < \frac{\delta_1}{k_1^3 k_2 k^2\sqrt{n}\,\varphi_{\max}^4 r T^2} \equiv \omega_1 \tag{4.1}$$

then

$$\alpha_1(\omega) I \le \int_t^{t+T}\psi(\theta(\tau), \tau)\psi^T(\theta(\tau), \tau)\, d\tau \le \alpha_2 I \tag{4.2}$$

with

$$\alpha_1(\omega) = \delta_1 - \omega k_1^3 k_2 k^2\sqrt{n}\,\varphi_{\max}^4 r T^2 > 0 \tag{4.3}$$

$$\alpha_2 = k k_1^2\varphi_{\max}^2 T > 0. \tag{4.4}$$

Proof: The upper bound $\alpha_2 I$ follows immediately by A.1 and A.2. Integrating by parts twice successively we can write

$$\begin{aligned}\int_t^{t+T}\left(\frac{\partial\beta}{\partial\theta}\right)^T_\theta \varphi\varphi^T\left(\frac{\partial\beta}{\partial\theta}\right) d\tau &= \left(\frac{\partial\beta}{\partial\theta}\right)^T_{\theta(t+T)}\left(\int_t^{t+T}\varphi\varphi^T d\tau\right)\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta(t+T)} \\ &\quad - \left[\int_t^{t+T}\frac{d}{dt}\left(\frac{\partial\beta}{\partial\theta}\right)^T\left(\int_t^{t+T}\varphi\varphi^T d\sigma\right) d\tau\right]\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta(t+T)} \\ &\quad - \int_t^{t+T}\left(\int_t^{\tau}\left(\frac{\partial\beta}{\partial\theta}\right)^T\varphi\varphi^T d\sigma\right)\frac{d}{dt}\left(\frac{\partial\beta}{\partial\theta}\right) d\tau.\end{aligned} \tag{4.5}$$

The time derivative of $\partial\beta/\partial\theta$ can be expressed as follows:

$$\frac{d}{dt}\left(\frac{\partial\beta}{\partial\theta}\right) = \sum_{i=1}^n [D_{i1}^2\beta(\theta) \cdots D_{in}^2\beta(\theta)]\dot{\theta}_i$$

and is therefore bounded as follows

$$\left\|\frac{d}{dt}\left(\frac{\partial\beta}{\partial\theta}\right)\right\| \le \left\|\frac{d}{dt}\left(\frac{\partial\beta}{\partial\theta}\right)\right\|_F \le k_2\sqrt{nk}\,\|\dot{\theta}\|. \tag{4.6}$$

By assumption A.1 we have

$$\|\beta(\theta_1) - \beta(\theta_2)\| \le k_1\sqrt{k}\,\|\theta_1 - \theta_2\|$$

and, therefore, by (3.3)

$$\|\dot{\theta}\| = \|\dot{\tilde{\theta}}\| \le \omega k_1^2 k\varphi_{\max}^2 r. \tag{4.7}$$

Hence, the 2-norm of each of the last two terms of (4.5) is bounded above by

$$1/2\,\omega k^2 k_1^3 k_2\varphi_{\max}^4\sqrt{n}\, r T^2.$$

Since $\theta(t+T) \in B_\theta$, it follows from assumption A.3 that the first matrix is bounded below by $\delta_1 I$. The result then follows from (4.1). ♦

Before stating our main result, we need the following technical lemma which has been proved in [8].

Lemma 4.2: Consider the linear time-varying system

$$\dot{x} = -\omega\psi\psi^T x \qquad x(0) = x_0 \tag{4.8}$$

with $\omega > 0$, $x \in R^n$, and where ψ satisfies the PE condition (4.2), then $|x(t)| \le Ke^{-at}|x_0|$, where

$$K(\omega) = \sqrt{\frac{1}{1-\gamma(\omega)}}, \quad a(\omega) = -\frac{1}{2T}\log(1-\gamma(\omega)),$$

$$\gamma(\omega) = \frac{2\alpha_1\omega}{(1+n\alpha_2\omega)^2}. \tag{4.9}$$

♦

Consider now the function

$$W(\omega) = \frac{a(\omega)}{\omega K(\omega)} = -\frac{\sqrt{1-\gamma(\omega)}}{2\omega T}\log[1-\gamma(\omega)]$$

for $\omega \ge 0$, with $\gamma(\omega)$ defined by (4.9) and $\alpha_1 = \alpha_1(\omega)$ defined by (4.3), i.e.,

$$\alpha_1(\omega) = \delta_1\left(1 - \frac{\omega}{\omega_1}\right), \quad \gamma(\omega) = \frac{2\,\delta_1\omega}{(1+n\alpha_2\omega)^2}\left(1 - \frac{\omega}{\omega_1}\right). \tag{4.10}$$

It is fairly easy to see that $W(\omega)$ has the form depicted in Fig. 1. With k_3 as defined in (3.8) and assuming that $k_3\varphi_{\max}^2 \le \delta_1/T$, we define for later use ω_2 as the unique value of ω for which $W(\omega_2) = k_3\varphi_{\max}^2$.

Our main result under assumptions A.1–A.3 is now as follows.

Theorem 4.1: Consider the estimation algorithm (3.1) with the assumptions A.1–A.3, and the additional assumption.

A.4: r is chosen small enough so that, with k_3 defined by (3.8),

$$k_3\varphi_{\max}^2 \le \frac{\delta_1}{T}. \tag{4.11}$$

Let the adaptation gain ω be chosen such that $\omega < \omega_2$, with ω_2 defined by $W(\omega_2) = k_3\varphi_{\max}^2$ (see Fig. 1), and let

$$\|\tilde{\theta}(0)\| < r\sqrt{1-\gamma(\omega)}. \tag{4.12}$$

Then

$$\|\tilde{\theta}(t)\| \le \frac{1}{\sqrt{1-\gamma(\omega)}}\exp(-\lambda(\omega)t)\|\tilde{\theta}(0)\| \le r \qquad \forall t \ge 0 \tag{4.13}$$

where

$$\lambda(\omega) \equiv -\frac{1}{2T}\ln(1-\gamma(\omega)) - \frac{\omega\varphi_{\max}^2 k_3}{\sqrt{1-\gamma(\omega)}}.$$

Proof: Equation (3.5) can be rewritten as

$$\dot{\tilde{\theta}} = -\omega\psi\psi^T\tilde{\theta} + f(t, \tilde{\theta}) \tag{4.14}$$

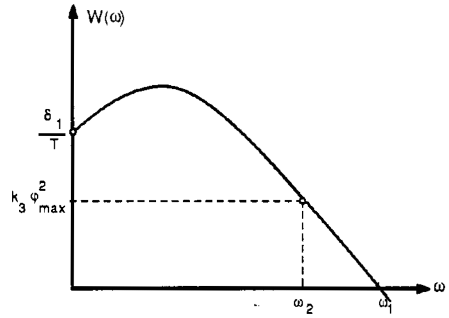

Fig. 1.

where $f(t, 0) = 0$ and $f(t, \tilde{\theta})$ satisfies the Lipschitz condition (3.7). It follows from (4.12) that there exists a positive constant $\epsilon \geq 0$ such that

$$\|\tilde{\theta}(0)\| < (r - \epsilon)\sqrt{1 - \gamma(\omega)}.$$

We demonstrate by contradiction that

$$\|\tilde{\theta}(t)\| < (r - \epsilon) \qquad \forall t.$$

Suppose there exists a finite $t_1 > 0$ such that

$$\|\tilde{\theta}(t)\| < r - \epsilon \qquad 0 \leq t < t_1, \qquad \|\tilde{\theta}(t_1)\| = r - \epsilon. \quad (4.15)$$

Then, it is clear that

$$\|\tilde{\theta}(\sigma)\| < r, \qquad \forall \sigma, \quad 0 \leq \sigma \leq t_1.$$

Hence, since $\omega < \omega_2 \leq \omega_1$, ψ satisfies the PE condition (4.2) with $\alpha_1(\omega)$ defined by (4.3). Therefore, the homogeneous equation

$$\dot{\tilde{\theta}} = -\omega\psi\psi^T\tilde{\theta} \qquad (4.16)$$

is exponentially asymptotically stable, and

$$\|\tilde{\theta}(\sigma)\| \leq K(\omega)e^{-a(\omega)\sigma}\|\tilde{\theta}(0)\| \qquad \sigma \in [0, t_1]$$

with $K(\omega)$ and $a(\omega)$ defined by (4.9). Since $\omega < \omega_2$, it also follows that

$$\frac{\omega\varphi_{\max}^2 k_3 K(\omega)}{a(\omega)} < 1$$

where $\omega\varphi_{\max}^2 k_3$ is the Lipschitz constant of the perturbation term $f(t, \tilde{\theta})$ (see Lemma 3.1). It then follows from the total stability theorem (see, e.g., [9]) that, for $\sigma \in [0, t_1]$

$$\|\tilde{\theta}(\sigma)\| \leq \frac{1}{\sqrt{1 - \gamma(\omega)}} \exp(-\lambda(\omega)\sigma)\|\tilde{\theta}(0)\|$$

$$\leq \frac{\|\tilde{\theta}(0)\|}{\sqrt{1 - \gamma(\omega)}} < r - \epsilon$$

where

$$\lambda(\omega) \equiv -\frac{1}{2T}\log(1 - \gamma(\omega)) - \frac{\omega\varphi_{\max}^2 k_3}{\sqrt{1 - \gamma(\omega)}}.$$

This is in contradiction with (4.15). Hence

$$\|\tilde{\theta}(t)\| < (r - \epsilon) \qquad \forall t \geq 0 \qquad (4.17)$$

and the theorem follows. ♦

Comments:

1) The total stability theorem essentially says that if the perturbation term $f(t, \tilde{\theta})$ is Lipschitz and if the homogeneous equation (4.15) is exponentially stable, then the perturbed $\theta(t)$ remains within a ball of radius r, and its norm decreases with a slower rate [hence, the second term in (4.13)] provided the initial condition is within a ball of smaller radius $r/K(\omega)$. The effect of ω on the radius of the initial condition ball and on the speed of convergence λ can be seen from Fig. 2.

2) The condition (4.11) can always be satisfied by choosing r small enough, i.e., which implies that $\theta(0)$ must be closer to the θ^*. However, it is interesting to note that the richer φ is (i.e., the larger δ_1/T is; see the PE condition A.3), the larger the convergence radius r is allowed to be.

3) Finally, we note that if $\beta(\theta)$ is linear, $k_2 = k_3 = 0$, $\gamma(\omega) > 0$ for all ω, $\lambda = -1/2T\ln(1 - \gamma(\omega))$, (4.11) is always satisfied, and the classical exponential convergence results of the linear regression case are recovered, without any constraint on $\|\theta(0)\|$ or ω.

V. Convergence Results Under A1, A2, and A3′

In this section, an analysis, parallel to that of Section IV, will be carried out under the weaker assumption A3′ on the persistency of excitation of the regressor. Roughly speaking, assumption A3′ requires that the regressor $\varphi(t)$ must be sufficiently rich only for the true system, that is if the parameter is exact ($\theta = \theta^*$), while assumption A3 requires a sufficient richness for all the models corresponding to all the admissible parameter values (i.e., $\forall\theta \in B_\theta$). Clearly, A.3′ is a weaker requirement on $\varphi(t)$ than A.3, and A.3 implies A.3′.

From assumptions A.2 and A.3′, it follows directly that:

$$\alpha_1 I \leq \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta^*}^T \left[\int_t^{t+T} \varphi(\tau)\varphi^T(\tau)\, d\tau\right] \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta^*} \leq \alpha_2 I \quad (5.1)$$

with $\alpha_1 = \delta_2$ and $\alpha_2 = kk_1^2\varphi_{\max}^2 T$.

The error equation (3.3) is rewritten as follows:

$$\dot{\tilde{\theta}} = -\omega\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta^*}^T \varphi\varphi^T \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta^*} \tilde{\theta} + f_1(t, \tilde{\theta}) \qquad (5.2)$$

where

$$f_1(t, \tilde{\theta}) = -\omega\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta}^T \varphi\varphi^T \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta^*} \tilde{\theta} + \omega\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta^*}^T \varphi\varphi^T \left(\frac{\partial\beta}{\partial\theta}\right)_{\theta^*} \tilde{\theta}$$
$$+ \omega\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta}^T \varphi\varphi^T R(-\tilde{\theta}). \quad (5.3)$$

Let $\tilde{\theta}_1 \equiv (\theta^* - \theta_1)$, $\tilde{\theta}_2 \equiv (\theta^* - \theta_2)$ with $\theta_1, \theta_2 \in B_\theta$. Then, under assumptions A.1 and A.2, it can be shown that $f_1(t, \tilde{\theta})$ satisfies the following Lipschitz condition:

$$\|f_1(\tilde{\theta}_1) - f_1(\tilde{\theta}_2)\| \leq \omega\varphi_{\max}^2 k_4 \|\tilde{\theta}_1 - \tilde{\theta}_2\| \qquad (5.4)$$

with

$$k_4 = k_2 r\left[3k_1 k\sqrt{n} + k_2\sqrt{k}\, nr\right]. \qquad (5.5)$$

According to Lemma 4.2, we define the following quantities:

$$\gamma_1(\omega) = \frac{2\alpha_1\omega}{(1 + n\alpha_2\omega)^2}, \quad K_1(\omega) = \sqrt{\frac{1}{1 - \gamma_1(\omega)}},$$
$$a_1(\omega) = -\frac{1}{2T}\log(1 - \gamma_1(\omega)). \qquad (5.6)$$

Consider now the function

$$W_1(\omega) = \frac{a_1(\omega)}{\omega K_1(\omega)} = -\frac{\sqrt{1 - \gamma_1(\omega)}}{2\omega T}\log[1 - \gamma_1(\omega)]. \qquad (5.7)$$

$W_1(\omega)$ has the form depicted in Fig. 3. With k_4 as defined in (5.5)

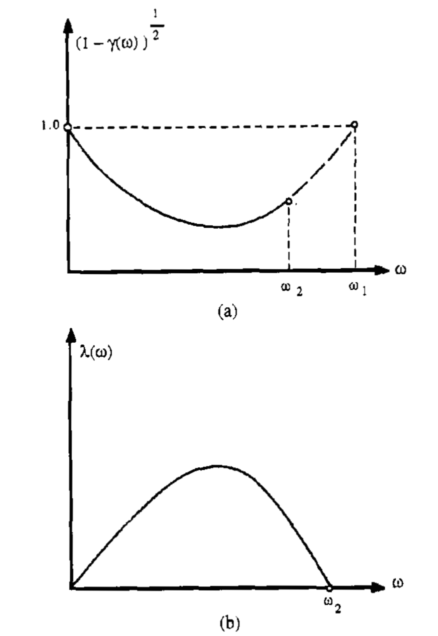

Fig. 2. (a) Size of allowable initial condition versus adaptation gain. (b) Convergence rate versus adaptation gain.

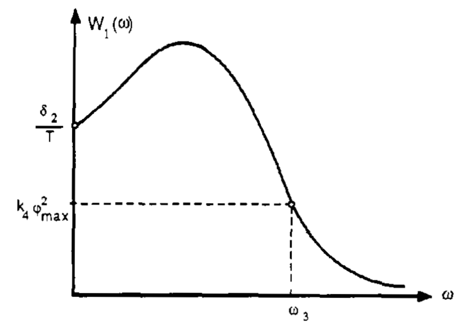

Fig. 3.

and assuming that

$$k_4\varphi^2_{\max} \le \frac{\alpha_1}{T} = \frac{\delta_2}{T}$$

we define ω_3 as the unique value of ω for which

$$W_1(\omega_3) = k_4\varphi^2_{\max}.$$

Theorem 5.1: Consider the estimation algorithm (3.1) with the assumptions A1–A3', and the additional assumption A4'.

A4': r is chosen small enough so that, with k_4 defined by (5.5),

$$k_4\varphi^2_{\max} \le \frac{\delta_2}{T}. \tag{5.8}$$

Let the adaptation gain ω be chosen such that $\omega < \omega_3$, and let

$$\|\tilde{\theta}(0)\| < r\sqrt{1-\gamma_1(\omega)}. \tag{5.9}$$

Then

$$\|\tilde{\theta}(t)\| \to 0 \text{ exponentially fast, i.e.,}$$

$$\|\tilde{\theta}(t)\| \le \frac{1}{\sqrt{1-\gamma_1(\omega)}}\exp\left[-\lambda_1(\omega)t\right]\|\tilde{\theta}(0)\| \le r \quad \forall t \ge 0 \tag{5.10}$$

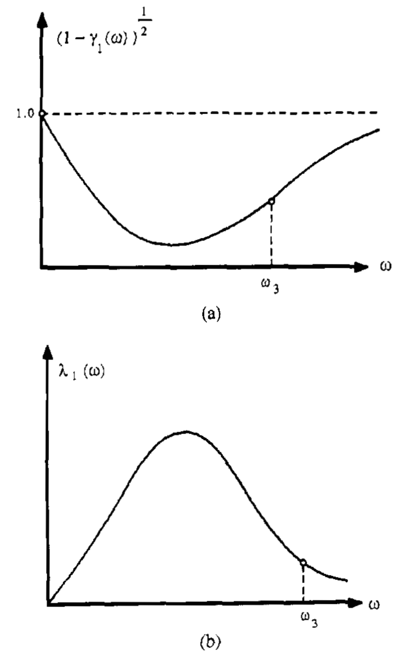

Fig. 4. (a) Size of allowable initial condition versus adaptation gain. (b) Convergence rate versus adaptation gain.

where

$$\lambda_1(\omega) = -\frac{1}{2T}\log\left[1-\gamma_1(\omega)\right] - \frac{\omega\varphi^2_{\max}k_4}{\sqrt{1-\gamma_1(\omega)}}. \tag{5.11}$$

Proof: Follows straightforwardly from the total stability theorem. ♦

Comment: In this case the effect of ω on the radius of the initial condition ball and on the speed of convergence λ_1 is seen from Fig. 4.

VI. Discussion and Conclusion

We have followed two different (but fairly parallel) ways for the analysis of a parameter estimator for a class of nonlinear regression problems. The reader might believe that this is redundant and that one way is better than the other. This is actually not the case, as is shown by the following argumentation.

Suppose that the regressor $\varphi(t)$ is given (from an experiment on the system) and that it is sufficiently rich in the sense of both A3 and A3'. Then it follows from the analysis that the radius r of the admissible domain B_θ for the parameter estimates must be chosen such that

$$\text{first analysis (A3)}: \delta_1(r) \ge k_3(r)\varphi^2_{\max}T \tag{6.1}$$

$$\text{second analysis (A3')}: \delta_2 \ge k_4(r)\varphi^2_{\max}T \tag{6.2}$$

with $\delta_1(r) \le \delta_2$ and $k_3(r) \le k_4(r)$.

$k_3(r)$ and $k_4(r)$ can be viewed as a measure of the degree of nonlinearity in the parameterization ($k_3 = k_4 = 0$ when $\beta(\theta)$ is linear function of θ). They are both monotonically increasing with r. $\delta_1(r)$ and δ_2 are a measure of the regressor richness. $\delta_1(r)$ is monotonically decreasing with r.

It is clear that no definite conclusion can be drawn from (6.1) and (6.2) regarding the respective sizes of B_θ arising from the first and the second analysis. Either way could yield a larger B_θ depending on the particular structure of the nonlinearity in specific applications.

Acknowledgment

The results presented in this note have been obtained within the framework of the Belgian Program on Concerted Research Actions and on Interuniversity Attraction Poles initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility rests with its authors.

References

[1] E. W. Bai and S. S. Sastry, "Parameter identification using prior information," *Int. J. Contr.*, vol. 44, no. 2, pp. 455-473, 1986.

[2] J. P. Clary and G. F. Franklin, "Self tuning control with *a priori* plant knowledge," in *Proc. 23th IEEE CDC*, 1984, pp. 1510-1513.

[3] C. Canudas de Wit, "Adaptive control for partially known systems: Theory and applications," Ph.D. dissertation, Laboratoire d'Automatique de Grenoble, 1987.

[4] S. Dasgupta, B. D. O. Anderson, and R. J. Kaye, "Robust identification of partially known systems," in *Proc. 22th IEEE CDC*, 1983, p. 1510.

[5] —, "Identification of physical parameters in structured systems," *Automatica*, vol. 24, pp. 217-225, Mar. 1988.

[6] S. Dasgupta, "Adaptive identification of systems with polynomial parametrizations," *IEEE Trans. Circuits Syst.*, vol. 35, no. 5, pp. 599-603, May 1988.

[7] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*. Cambridge, MA: M.I.T. Press, 1983.

[8] G. Kreisselmeier, "Adaptive observers with exponential rate of convergence," *IEEE Trans Automat. Contr.*, vol. AC-22, pp. 2-8, 1977.

[9] B. D. O. Anderson *et al.*, *Stability of Adaptive Systems.* Cambridge, MA: M.I.T. Press, 1986.

Lower Summation Bounds for the Discrete Riccati and Lyapunov Equations

Nicholas Komaroff and Bahram Shahian

Abstract—**Lower eigenvalue summation (including trace) bounds for the solution of the discrete algebraic Riccati and Lyapunov matrix equations are presented. These are tighter than or supplement existing results.**

I. Introduction

Consider the discrete algebraic Riccati equation (DARE)

$$P = A'PA - A'PB(I + B'PB)^{-1}B'PA + Q, \qquad Q = Q' \geq 0 \tag{1}$$

where A, P, $Q \in R^{n \times n}$, $B \in R^{n \times r}$, $(')$, I and (≥ 0) denote the transpose, the unit matrix, and positive semidefiniteness, respectively. When $B = 0$, (1) becomes the discrete algebraic Lyapunov equation (DALE).

$$P = A'PA + Q. \tag{2}$$

Manuscript received July 28, 1989; revised February 9, 1990.

N. Komaroff was with the Department of Electrical Engineering, University of California, Los Angeles, CA 90025. He is now with Queensland University of Technology, School of Electrical and Electronic Systems Engineering, Brisbane QLD 4001, Australia.

B. Shahian is with the California State University, Long Beach, CA 90840.

IEEE Log Number 9108140.

The above equations are of central importance in signal processing and control theory [1]. Knowledge of ranges of the magnitudes of the solutions to (1) and (2) gives design information about systems governed by these equations, and provides a starting point for numerical solution algorithms. Such ranges are given by bounds on eigenvalues of the solution P, and on their summations and products—see [2] for a summary and some applications.

Lower bounds for $\operatorname{tr}(P)$, the trace of P, have recently been obtained in [3]-[6], and for summations of eigenvalues in [5]. In this note we derive summation lower bounds, that include the trace, for the eigenvalues of P in (1) and (2) that are tighter than, or supplement those in [3]-[6]. Our results are expressed by Theorems 2.1 and 2.2, and corollaries.

In what follows, $\lambda_i(X)$ denotes the ith eigenvalue of a matrix X, $i = 1, 2, \cdots, n$. All eigenvalues are ordered such that their real parts are nonincreasing

$$\operatorname{Re} \lambda_1(X) \geq \operatorname{Re} \lambda_2(X) \geq \cdots \geq \operatorname{Re} \lambda_n(X).$$

The abbreviation RHS (LHS) means right- (left)-hand side.

The following theorems and lemmas shall be used.

Theorem 1.1 [7, p. 246]: Let symmetric matrices X, Y be positive semidefinite. Then

$$\prod_1^k \lambda_i(XY) \leq \prod_1^k \lambda_i(X)\lambda_i(Y) \qquad k = 1, 2, \cdots, n \tag{3}$$

with equality when $k = n$. This theorem is due to Horn, 1950.

Lemma 1.1 [8]: Let a_i, b_i be nonnegative real numbers such that $a_1 \leq b_1, a_1 a_2 \leq b_1 b_2, \cdots, a_1 \cdots a_n \leq b_1 \cdots b_n$. Then for any real exponent $s > 0$

$$\sum_1^k a_i^s \leq \sum_1^k b_i^s, \qquad k = 1, 2, \cdots, n. \tag{4}$$

Theorem 1.2 [9]: Let matrices $X, Y \geq 0$ and $1 \leq i, j \leq n$. Then

$$\lambda_{i+j-n}(XY) \geq \lambda_j(X)\lambda_i(Y), \qquad i + j \geq n + 1 \tag{5}$$

$$\lambda_{i+j-1}(XY) \leq \lambda_j(X)\lambda_i(Y), \qquad i + j \leq n + 1. \tag{6}$$

Theorem 1.3 [8]: Let X be any arbitrary n by n matrix. Then

$$\sum_1^k |\lambda_i(X)|^2 \leq \sum_1^k \lambda_i(X'X), \qquad k = 1, 2, \cdots, n. \tag{7}$$

Theorem 1.4 [8]: Let A be an arbitrary n by n matrix. Then

$$\prod_1^k |\lambda_i(A)|^2 \leq \prod_1^k \lambda_i(A'A), \qquad k = 1, 2, \cdots, n \tag{8}$$

with equality when $k = n$. Because of the equality when $k = n$

$$\prod_1^k |\lambda_{n-i+1}(A)|^2 \geq \prod_1^k \lambda_{n-i+1}(A'A). \tag{9}$$

Theorem 1.5 [7, p. 241]: Let X, Y be symmetric n by n matrices. Then

$$\sum_1^k \lambda_i(X + Y) \leq \sum_1^k [\lambda_i(X) + \lambda_i(Y)], \qquad k = 1, 2, \cdots, n. \tag{10}$$

This theorem is due to Fan, 1949.

Theorem 1.6 [7, p. 245]: Let X, Y be symmetric n by n

0018-9286/92$03.00