C43

WP7 - 2:15

Proceedings of the 28th Conference
on Decision and Control
Tampa, Florida • December 1989

IDENTIFICATION OF LINEARLY OVERPARAMETRIZED NON LINEAR SYSTEMS

G. Bastin*, R.R. Bitmead+, G. Campion*[1], M. Gevers*

*Laboratoire d'Automatique, Dynamique et Analyse des Systèmes. Catholic University of Louvain, 1348 Louvain-la-Neuve, Belgium.
+ Department of Systems Engineering - Australian National University - GPO Box 4 Canberra - Australia
[1] Chercheur qualifié FNRS.

Abstract.

Often, a dynamical model is non linear in the unknown parameters, but it can be transformed into an overparametrized linear regression model, where the components of the overparametrization vector are non linear functions of the smaller number of unknown parameters. We present an algorithm that directly identifies the unknown parameters, and we characterize the convergence domain.

1. INTRODUCTION - STATEMENT OF THE PROBLEM

In many practical modelling 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 non linearly 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 by [CF],[BS] and [C]. In this paper we consider the more complex situation where the unknown parameters enter non linearly in the model but can be embedded in a linear over-reparametrization to be made explicit shortly in (1.1). This issue has been previously discussed in a series of papers by [DAK 1,2,3] 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 non linear systems, where the reparametrization is any non linear function of the unknown parameters. The systems under consideration are assumed to be expressed in the following non linear regression form

$$y(t) = \varphi^T(t)\beta(\theta) \qquad (1.1)$$

where $t \in R_+$, $y \in R^m$ is a vector observation sequence, $\varphi \in R^k x R^m$ is a regression matrix made up of known signals, $\theta \in R^n$ is the unknown parameter vector and $\beta(.)$ is a non linear 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.* The field of applications of this problem is illustrated by two typical engineering examples.

An electrical example

Consider the following electrical circuit :

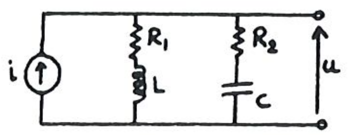

Fig. 1.

The dynamics of this circuit are easily shown to be described by the following differential equation :

$$[R_2LCD^2+(L+R_1R_2C)+R_1]\, i = [LCD^2+(R_1+R_2)\,CD+1]\, u \qquad (1.2)$$

where D stands for the derivative operator (D = (d/dt)). We assume that the values of the resistances R_1 and R_2 are known and, hence, that the only model unknowns are the inductance L and the capacitance C: $\theta_1 = L$ and $\theta_2 = C$. Operating on both sides of (1.2) with the following stable low pass filter: $p(D) = [D^2+\omega_1 D+\omega_2]^{-1}$ with ω_1 and ω_2 arbitrary positive constants, we get the following non linear regression form:

$$y = \sum_{i=1}^{3} \varphi_i(t)\beta_i(\theta)$$
$$y = p(D)D^2(u - R_2 i)$$
$$\varphi_1(t) = p(D)Di$$
$$\varphi_2(t) = p(D)(R_1R_2Di - R_1 i + R_2Du)$$
$$\varphi_3(t) = p(D)(u - R_1 i)$$

and the over-reparametrization :

$$\beta(\theta) = [\, \beta_1(\theta), \beta_2(\theta), \beta_3(\theta)\,]^T$$
$$\beta_1(\theta) = \theta_1^{-1} = C^{-1} \qquad \beta_2(\theta) = \theta_2^{-1} = L^{-1} \qquad \beta_3(\theta) = \theta_1^{-1}\theta_2^{-1} = C^{-1}L^{-1}$$

A mechanical example

The dynamics of a 3-degrees-of-freedom omnidirectionnal symmetric mobile trolley which has been built by the Mechanical Engineering Department of Louvain University is as follows (see also fig. 2 and [CBRR] for further details) :

CH2642-7/89/0000-0618$1.00© 1989 IEEE

$$F(q)\ddot{q} + f(q,\dot{q}) = G(q)u \qquad (1.3)$$

with:

$q = (q_1, q_2, q_3)^T$ the vector of generalized coordinates

$F(q)$ the inertia matrix

$f(q,\dot{q})$ the vector of coriolis and centripetal torques

$G(q)$ the kinematics matrix.

$$F(q) = \begin{pmatrix} M+1.5.J_r & 0 & -Mb(q) \\ 0 & M+1.5.J_r & -Ma(q) \\ -Mb(q) & -Ma(q) & I+M(l_1^2+l_2^2) + 3J_rL^2 \end{pmatrix}$$

$$f(q,\dot{q}) = \begin{pmatrix} Ma(q)\,\dot{q}_3^2 + 1.5\,J_r\dot{q}_2\dot{q}_3 \\ -Mb(q)\,\dot{q}_3^2 - 1.5\,J_r\dot{q}_1\dot{q}_3 \\ 0 \end{pmatrix}$$

$$a(q) = l_2 \sin q_3 - l_1 \cos q_3$$

$$b(q) = l_2 \cos q_3 - l_1 \sin q_3$$

We assume that all the geometrical and inertial parameters relative to the trolley and to the load, considered separately, are known. This means that, in the model, the parameters M, J_r, I, L are known. The only unknown parameters are those relative to the location of the load on the trolley. This means that the coordinates l_1 and l_2 of the center of mass of the load with respect to the vertical symmetry axis of the trolley are unknown : $\theta_1 = l_1 \quad \theta_2 = l_2$. Hence, it is easy to check, from the definition of $F(q)$ and $f(q,\dot{q})$ that the model (1.3) can be linearly reparametrized as follows:

$$Y(q,\dot{q},\ddot{q},u) = \sum_{i=1}^{3} \varphi_i(q,\dot{q},\ddot{q})\,\beta_i(\theta) \qquad (1.4)$$

$$Y(q,\dot{q},\ddot{q},u) = G(q)u - F_0(q)\,\ddot{q} - f_0(q,\dot{q})$$

$$\varphi_i(q,\dot{q},\ddot{q}) = F_i(q)\,\ddot{q} + f_i(q)\,\ddot{q} \qquad i = 1, 2, 3$$

$$\beta_1 = \theta_1 \qquad \beta_2 = \theta_2 \qquad \beta_3 = \theta_1^2 + \theta_2^2$$

and appropriate definitions of F_i and f_i. The model (1.4) is then in the desired non linear regression form which depends on $q, \dot{q}, \ddot{q}$. In case where, as usual in robotics, the accelerations $\ddot{q}$ are not measured, the model (1.4) can be easily transformed in an equivalent non linear regression model (with the same parametrization) which depends only on q and $\dot{q}$, by appropriate filtering (see e.g. [MG]).

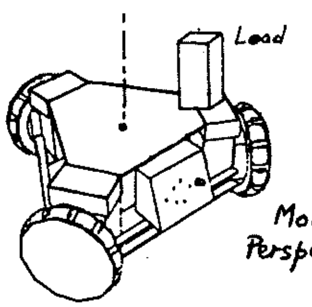

Fig. 2
Mobile Robot
Perspective View

Outline of the paper

The paper is organized as follows. In section 2, 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 [DAK,1,2,3] is emphasized. The estimation algorithm is presented in section 3 and a Lipshitz condition relative to the dynamics of the estimation error is established. The main convergence results are demontrated under two different assumptions on the excitation content of φ, in sections 4 and 5 respectively. In each case a lower bound on the size of the convergence domain is calculated and its connection with the structure of the overparametrization mapping $\beta(\theta)$ is discussed. For reasons of space limitations, the proofs of some of the lemmas have been deleted from this conference version; full proofs can be found in [BBCG].

2. ASSUMPTIONS.

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

Assumption on the structure of $\beta(.)$

A.1. The function $\beta(.)$ maps an open ball $D_\theta \in R^n$ of radius r, centered on θ^*, onto a set $D_\beta \in R^k$, 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 $D_\theta^{(+)}$.

In particular, there exist finite constants $k_1>0$ and $k_2>0$ such that:

$$\left\|\frac{\partial\beta_i}{\partial\theta}\right\|_2 \leq k_1 \text{ and } \left\|\frac{\partial}{\partial\theta}\left(\frac{\partial\beta_i}{\partial\theta_j}\right)\right\|_2 \leq k_2 \quad i = 1,k \quad j = 1,n \qquad \forall\theta\in D_\theta$$

$^+$*Notation :* For vectors $\beta \in \mathbf{R}^k$ and $\theta \in R^n$, we denote by $\partial\beta/\partial\theta$ the k x n matrix whose (i,j)-th element is

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

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 :

A.2. $\quad \|\varphi(t)\| \leq \varphi_{max} \quad \forall t \in R_+$

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 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

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

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. However, 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 set up has been studied extensively by Dasgupta, Anderson and Kaye in a series of papers [DAK,1,2,3] 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 results extend those of DAK in three ways : first, $\beta(\theta)$ is not restricted to polynomial functions of θ; second, our algorithm estimates θ directly, whereas in DAK 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); third, 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 [DAK 1,2,3], where the whole vector $\varphi(t)$ was required to be persistently exciting. Here we only require $P(\theta,t)$ (resp. $\bar{P}(t)$) to be positive definite : its size, n x n, is typically much smaller than the dimension k x 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.

3. THE ESTIMATION ALGORITHM

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

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

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

$$\psi(\theta,t) \stackrel{\Delta}{=} \left(\frac{\partial\beta}{\partial\theta}\right)_\theta^T \varphi(t) \qquad \theta\in D_\theta \qquad (3.2)$$

In the next two sections, we shall analyze the convergence properties of θ under assumptions A.1 to A.3 (resp. A.1 to 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 (2.1), we have

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

Let θ_1, θ_2 be any two points in D_θ. Then:

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

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

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

We denote: $f(t,\tilde{\theta}) \equiv \omega\left(\frac{\partial\beta}{\partial\theta}\right)_\theta^T \varphi\varphi^T R(\tilde{\theta})$ (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 $\theta_1, \theta_2 \in D_\theta$. Then, under assumptions A.1, A.2, $f(t,\theta)$ satisfies the following Lipschitz condition (we drop the dependence on t for simplicity):

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

$$k_3 = k_2 r\,[2k_1 k\sqrt{n} + k_2\sqrt{k}\; nr] \qquad (3.8)$$

Proof : see [BBCG]

4. 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 to 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 D_θ. We first establish two preliminary results.

Lemma 4.1

Consider the estimation algorithm (3.1) with the assumptions A.1 to A.3. If $\theta(t) \in D_\theta$ $\forall t\in \mathbb{R}_+$, and if:

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

Then: $\alpha_1(\omega) I \leq \int_t^{t+T} \psi(\theta(\tau),\tau)\, \psi^T(\theta(\tau),\tau) d\tau \leq \alpha_2 I$ (4.1)

with: $\alpha_1(\omega) = \delta_1 - \omega k_1^3\, k_2 k^2 \sqrt{n}\varphi_{max}^4\, r\, T^2 > 0$ (4.2)

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

Proof : see [BBCG].

Lemma 4.2.
Consider the linear time-varying system:

$$\dot{x} = -\omega\psi\psi^T x \qquad x(0) = x_0 \tag{4.3}$$

with $\omega > 0$, $x \in \mathbb{R}^n$, and where ψ satisfies the PE condition (4.1). Then $|x(t)| \le Ke^{-at}|x_0|$, where:

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

$$\gamma(\omega) = \frac{2\alpha_1\omega}{(1+n\alpha_2\omega)^2} \tag{4.4}$$

Proof:
Let $V(t) = x^T(t)x(t)$. Extending the derivations of [K] to the multivariable case, it can be shown that: $\dot{V}(t) \le 0 \quad \forall t \ge 0$ and:

$$\frac{V(t)-V(t+T)}{V(t)} \ge \frac{2\alpha_1\omega}{(1+n\alpha_2\omega)^2} \equiv \gamma(\omega)$$

Hence: $V(t+T) \le (1-\gamma(\omega))V(t)$.

This, and: $V(t+\tau) \le V(t)\ \forall\tau \ge 0$ implies:

$$V(t) \le K_1 \exp\left[\frac{t}{T}\log(1-\gamma(\omega))\right]V(0)$$

with $K_1 = (1-\gamma(\omega))^{-1}$. Taking the square root gives the desired result.
QED.

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 (4.4) and $\alpha_1 = \alpha_1(\omega)$ defined by (4.2), i.e.

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

It is fairly easy to see that $W(\omega)$ has the form depicted in Fig.3

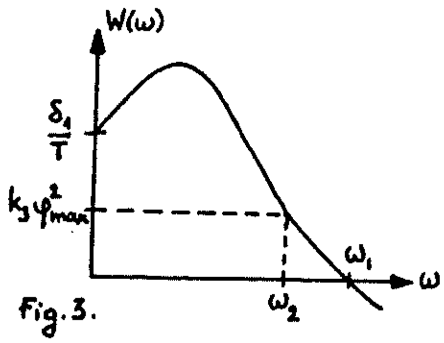

Fig.3.

With k_3 as defined in (3.8) and assuming that $k_3\varphi^2_{max} \le \delta_1/T$, we define for later use ω_2 as the unique value of ω for which $W(\omega_2) = k_3\varphi^2_{max}$. Our main result under Assumptions A.1 to A.3 is now as follows.

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

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

$$k_3\varphi^2_{max} \le \frac{\delta_1}{T} \tag{4.5}$$

Let the adaptation gain ω be chosen such that $\omega < \omega_2$, with ω_2 defined by $W(\omega_2) = k_3\varphi^2_{max}$ (see Fig.3), and let:

$$\|\tilde{\theta}(0)\| < r\sqrt{1-\gamma(\omega)} \tag{4.6}$$

Then: $\|\tilde{\theta}(t)\| < r \quad \forall t \ge 0$

Proof:
Equation (3.5) can be rewritten as

$$\dot{\tilde{\theta}} = -\omega\psi\psi^T\tilde{\theta} + f(t,\tilde{\theta}) \tag{4.7}$$

where $f(t,0) = 0$ and $f(t,\tilde{\theta})$ satisfies the Lipschitz condition (3.7). It follows from (4.6) that there exists a positive constant $\varepsilon > 0$ such that

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

We demonstrate by contradiction that $\|\tilde{\theta}(t)\| < (r-\varepsilon) \quad \forall t$.
Suppose there exists a finite $t_1 > 0$ such that:

$$\|\tilde{\theta}(t)\| < r-\varepsilon \quad 0 \le t < t_1,\ \|\tilde{\theta}(t_1)\| = r-\varepsilon \tag{4.8}$$

Then, it is clear that $\|\tilde{\theta}(\sigma)\| < r,\ \forall\sigma,\ 0 \le \sigma \le t_1$

Hence, since $\omega < \omega_2 \le \omega_1$, ψ satisfies the PE condition (4.1) with $\alpha_1(\omega)$ defined by (4.2). Therefore the homogeneous equation $\dot{\tilde{\theta}} = -\omega\psi\psi^T\tilde{\theta}$ is exponentially asymptotically stable, and:

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

with $K(\omega)$ and $a(\omega)$ defined by (4.4). Since $\omega < \omega_2$, it also follows that: $\frac{\omega\varphi^2_{max}k_3K(\omega)}{a(\omega)} < 1$ where $\omega\varphi^2_{max}k_3$ 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. [ABJKKMPR]) that, for $\sigma \in [0,t_1]$

$$\|\tilde{\theta}(\sigma)\| \le \frac{1}{\sqrt{1-\gamma(\omega)}} \exp(-\lambda(\omega)\sigma) \|\tilde{\theta}(0)\| \le \frac{\|\tilde{\theta}(0)\|}{\sqrt{1-\gamma(\omega)}} < r - \varepsilon$$

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 (4.8). Hence: $\|\tilde{\theta}(t)\| < (r-\varepsilon) \quad \forall t \ge 0$ and the theorem follows. QED.

Comments

1) The Total Stability Theorem essentially says that if the perturbation term $f(t, \tilde{\theta})$ is Lipschitz and if the homogeneous equation (4.3) 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 (4.7)) 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.4.

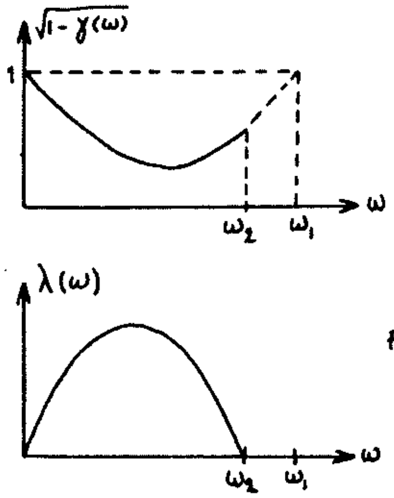

Fig.4.

2) The condition (4.5) can always be satisfied by choosing r small enough, i.e. which implies that $\theta(0)$ must be closer to the true 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, $\lambda = -1/2T \ln(1-\gamma(\omega))$, (4.5) is always satisfied, and the classical exponential convergence results of the linear regression case are recovered, without any constraint on $\|\theta(0)\|$ or ω.

5. CONVERGENCE RESULTS UNDER A1, A2 AND A3'

In this section, an analysis, parallel to that of section 4, 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. $\forall\theta \in D_\theta$). 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 :

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

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

The error equation (3.3) is rewritten as follows

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

where

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

$$+ \omega\left(\frac{\partial\beta}{\partial\theta}\right)_{\theta}^T \varphi\varphi^T R(-\tilde{\theta}) \qquad (5.1)$$

We now derive a Lipschitz bound for $f_1(t,\tilde{\theta})$

Lemma 5.1.

Let $f_1(t,\tilde{\theta})$ be defined by (5.1) and let $\tilde{\theta}_1 \triangleq \theta^*-\theta_1$, $\tilde{\theta}_2 \triangleq \theta^*-\theta_2$ with $\theta_1, \theta_2 \in D_\theta$. Then, under assumption A.1 and A.2, $f(t,\tilde{\theta})$ satisfies the Lipschitz condition: $\|f(\tilde{\theta}_1) - f(\tilde{\theta}_2)\|_2 \le \omega\varphi_{max}^2 k_4 \|\tilde{\theta}_1 - \tilde{\theta}_2\|_2$
with $k_4 = k_2 r\,[\,3\,k_1\,k\,\sqrt{n} + k_2\sqrt{k}\,n\,r]$ (5.2)

Proof : see [BBCG].

According to lemma 4.2, we define the following quantities:

$$\gamma_1(\omega) = \frac{2\alpha_1\omega}{(1+n\alpha_2\omega)^2}, \quad K_1(\omega) = \sqrt{\frac{1}{1-\gamma_1(\omega)}}$$

$$a_1(\omega) = -\frac{1}{2T}\log(1-\gamma_1(\omega))$$

Define the function: $W_1(\omega) = \frac{a_1(\omega)}{\omega K_1(\omega)} = -\frac{\sqrt{1-\gamma_1(\omega)}}{2\omega_T}\log\{1-\gamma_1(\omega)\}$

which has the form depicted on fig 5.

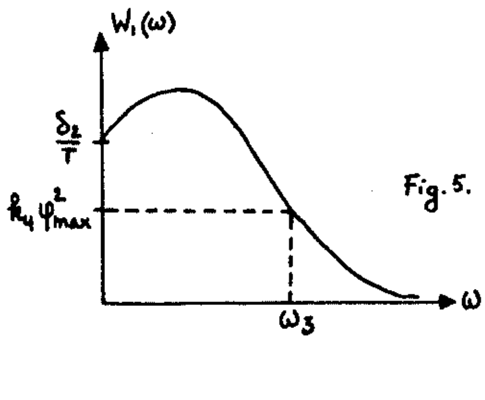

Fig.5.

With k_4 as defined in (5.2) and assuming that: $k_4\varphi_{max}^2 \leq \frac{\alpha_1}{T} = \frac{\delta_2}{T}$
we define ω_3 as the unique value of ω for which: $W_4(\omega_3)=k_4\varphi_{max}^2$

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

A4'. r is chosen small enough so that, with k_4 defined by (5.2),

$$k_4\varphi_{max}^2 \leq \frac{\delta_2}{T}$$

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

$$\| \tilde{\theta}(0) \|_2 < r\sqrt{1-\gamma_1(\omega)}$$

Then:

1) $\| \tilde{\theta}(t) \| \leq r \quad \forall t \in R_+$

2) $\| \tilde{\theta}(t) \|_2 \to 0$ exponentially fast i.e.

$$\| \tilde{\theta}(t) \|_2 \leq \frac{1}{\sqrt{1-\gamma_1(\omega)}} \exp[-\lambda_1 t] \ \| \tilde{\theta}(0) \|_2 \leq r \quad \forall t \geq 0$$

where: $\lambda_1 \stackrel{\Delta}{=} -\frac{1}{2T} \log[1-\gamma_1(\omega)] - \frac{\omega\varphi_{max}^2 k_4}{\sqrt{1-\gamma_1(\omega)}}$

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.6.

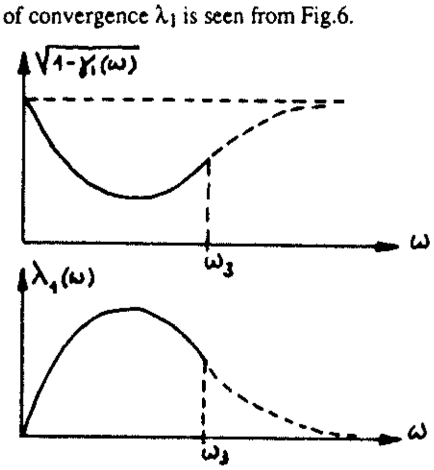

Fig.6.

6. DISCUSSION AND CONCLUSION

We have followed two different (but fairly parallel) ways for the analysis of a parameter estimator for a class of non linear 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 D_θ for the parameter estimates must be chosen such that :

first analysis (A3) : $\delta_1(r) \geq k_3(r)\,\varphi_{max}^2$

second analysis (A3') : $\delta_2 \geq k_4(r)\,\varphi_{max}^2$

with : $\delta_1(r) \leq \delta_2$ and $k_3(r) \leq k_4(r)$

$k_3(r)$ and $k_4(r)$ can be viewed as a measure of the degree of non linearity in the parametrization ($k_3 = k_4 = 0$ when $\beta(\theta)$ is linear function of θ). They are both monotically increasing with r. $\delta_1(r)$ and δ_2 are a measure of the regressor richness. $\delta_1(r)$ is monotically decreasing with r. It is clear that no definite conclusion can be drawn regarding the respective sizes of D_θ arising from the first and the second analysis. Either way could yield a larger D_θ depending on the particular structure of the non linearity in specific applications.

ACKNOWLEDGEMENTS.

The results presented in this paper 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

[BBCG] BASTIN G., BITMEAD R.R., CAMPION G. and GEVERS M. (1989) : "Identification of linearly overparametrized non linear systems", submitted to IEEE Trans. Automatic Control.

[BS] BAI E.W. and SASTRY S.S. (1986) : "Parameter Identification using prior information", Int. Journal of Control, vol. 13

[CF] CLARY J.P., and FRANKLIN G.F. (1984) : "Self tuning control with a priori plant knowledge", Proceedings 23th IEEE CDC, pp. 1510-1513.

[C] CANUDAS de WIT C. (1987) : "Adaptive control for partially known systems : Theory and Applications", Ph. D. thesis, Laboratoire d'Automatique de Grenoble.

[DAK1] DASGUPTA S., ANDERSON B.D.O, KAYE R.J. (1983) : "Robust identification of partially known systems", 22th IEEE CDC, pp. 1510

[DAK2] DASGUPTA S., ANDERSON B.D.O, KAYE R.J. (1988) : "Identification of physical parameters in sructured systems", Automatica, March 1988.

[DAK3] DASGUPTA S. (1988) : "Adaptive identification of systems with polynomial parametrizations", IEEE Trans. Circuits systems, vol. 35, n° 5, May 1988.

[CBRC] CAMPION G., BASTIN G., ROLIN D, RAUCENT B. (1988), "External linearization control for an omnidirectionnal mobile robot", IFAC SYROCO 88, p. 71.1.

[K] KREISSELMEIER G. (1977): "Adaptive observers with exponential convergence rate", IEEE Trans. Aut. Cont. AC 22(1), pp.2-8.

[ABJKKMPR] ANDERSON B.D.O. et al. (1986): "Stability of adaptive systems", MIT Press, Cambridge, Mass.

[MG] MIDDLETON R.H. and GOODWIN G.C. (1986):"Adaptive computed torque control for rigid link manipulators", Proc. 25th IEEE-CDC, Athens, pp.68-73.