PERSISTENCE OF EXCITATION PROPERTIES FOR THE IDENTIFICATION OF TIME-VARYING SYSTEMS [1]

Sergio Bittanti
Dipartimento di Elettronica ed Informazione
Politecnico di Milano
piazza Leonardo da Vinci 32
20133 Milano
Italy

M.C. Campi
Dipartimento di Elettronica per l'Automazione
University of Brescia
Via Branze 38
25123 Brescia
Italy

Lei Guo
Institute of Systems Science
Chinese Academy of Science
Beijing 100080
P.R. China

Abstract

The identification of time varying parameters requires that a certain level of information is present in the data through time. Only in this case it is in fact possible to track the parameter variability and form a reliable estimate. This consideration has led to the introduction in the literature of a variety of persistence of excitation notions ranging from the deterministic ones (in the '80's) to more sophisticated stochastic definitions proposed in the last decade.

This paper presents an overview of the existing stochastic excitation notions and discusses important issues like their necessity for tracking and their applicability in different contexts. It appears that the present state of the art is not completely satisfying in terms of completeness and generality of the available results.

[1] Research supported by MURST under the project "Identification and Control of Industrial Systems"

1 Introduction

In the identification of systems subject to *time variability*, it is crucial that the data collected from the system convey continual information on the parameters to be estimated. As a matter of fact, only in this case the identification algorithm can rely on fresh information in forming a reliable estimate of the current value of the system parameters. This concept has been formalized in the system identification literature under the name of *persistence of excitation*, see [1], [2] for classical references.

The condition can take a variety of forms. Letting ϕ_t be the observation vector, a prototype characterization takes the form

$$pr\left(\lambda_{min}\left\{\sum_{i=kh+1}^{(k+1)h}\frac{\phi_i\phi_i^T}{1+\|\phi_i\|^2}\geq k_1|\mathcal{F}_{kh}\right\}\right)\geq k_2,\quad \forall k, \tag{1}$$

for some h, where λ_{min} denotes the minimum eigenvalue, $\mathcal{F}_{kh}$ is the so-called σ-algebra of the past, that is the σ-algebra generated by all system processes up to time kh, and k_1, k_2 are two positive constants. Roughly, this condition requires that, whatever the past evolution of the system might have been, the information carried by data over the next h time points spans the entire parameter space with a finite nonzero probability.

Condition (1) goes back to the early '90's. In a form that is slightly different from but equivalent to (1) it was stated in [3], whereas in the above form it can be found in [4]. Henceforth, such a characterization has been used in many different papers.

In particular, in [3] the convergence and stability of a Kalman filter based algorithm are studied under (1) and in [5] it is proven that a forgetting factor least squares identification algorithm provides bounded estimates if condition (1) is met with and the forgetting factor is large enough. Another contribution using such notion of persistence of excitation is [6], where an expression for the asymptotic estimation error for a forgetting factor least squares algorithm is worked out.

Many more contributions prove results on system identification properties under conditions related to (1). Among others, we quote [7, 8, 9, 10, 11, 12, 13, 14].

In this paper, we critically review the problem of securing excitation in time-varying system identification. Our purpose is twofold:

i) providing an overview of the different approaches that have been adopted in the literature;
ii) focusing on open problems in this context.

2 A conditional persistence of excitation condition

The persistence of excitation condition introduced in [3] takes the form:

$$\lambda_{min}\left\{E\left[\sum_{i=kh+1}^{(k+1)h}\frac{\phi_i\phi_i^T}{1+\|\phi_i\|^2}|\mathcal{F}_{kh}\right]\right\}\geq\delta,\quad\forall k,\qquad(2)$$

for some $\delta > 0$, and is easily seen to be equivalent to (1).

In the case where $\{\phi_t\}$ is a ϕ-mixing process, condition (2) can be seen to be equivalent to the much simpler condition

$$\lambda_{min}\left\{\sum_{i=kh+1}^{(k+1)h}E\left[\frac{\phi_i\phi_i^T}{1+\|\phi_i\|^2}\right]\right\}\geq\delta'>0,\quad\forall k,\qquad(3)$$

which can be verified in a relatively easy way in many situations. The proof of the equivalence between (2) and (3) can be deduced by inspecting the proof of Theorem 2.3 in [15] or as a simple elaboration of the arguments in [11].

One interesting fact is that, in a ϕ-mixing context, condition (3) is, in a well-defined sense, see [15], also a necessary condition for guaranteeing tracking properties of the LMS identification algorithm, as well as for RLS with forgetting factor. Since such a condition is also sufficient for tracking, it appears to be the right condition in this context. The related discussions may be found in [15] for LMS and in [12] for RLS with forgetting factor. For signals $\{\phi_t\}$ which are more general than ϕ-mixing processes, it was also shown that the above condition is both necessary and sufficient for the stability of the standard LMS in the paper [16].

While the ϕ-mixing condition includes a number of important stochastic processes, such as M-dependent processes or signal generated by filters fed by bounded white noise, it is also true that a process generated by a dynamical system is not ϕ-mixing in general.

3 Assessing condition (2)

In Section 1, a number of papers providing results which hold true under the persistence of excitation condition (2) have been referenced. Clearly, in order to apply these results, one has first to verify that the persistence of excitation condition is satisfied in the setting under study.

It is advisable to distinguish two situations. In the case of systems without autoregressive parts, it turns out that condition (2) can be easily related to the properties of the inputs to the system. In this way, explicit assumptions can be derived for the persistence of excitation condition to hold. For example, in [6] condition (2) has been discussed in connection with Hammerstein models and linear combiners. On the other hand, assessing (2) is a difficult task when the system contains an autoregressive component and we mainly focus on this case in the present paper.

To be specific, let us consider the time-varying state variable system described by the equation

$$\phi_{t+1}=A_t\phi_t+w_{t+1}.\qquad(4)$$

In (4), w_t is a remote unmeasurable noise signal that plays the role of a latent variable in the generation of ϕ_t.

A typical example of system (4) is the time-varying scalar autoregressive model

$$y_{t+1} = \theta_t^T \phi_t + v_{t+1}, \tag{5}$$

where

$$\phi_t = [y_t, y_{t-1}, \dots, y_{t-p+1}],$$
$$\theta_t = [a_1(t), \dots, a_p(t)].$$

In this case, by letting

$$A_t \triangleq \begin{bmatrix} a_1(t) & \cdots & & a_p(t) \\ 1 & \ddots & & 0 \\ & \ddots & & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \quad w_{t+1} \triangleq \begin{bmatrix} v_{t+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{6}$$

system (4) is immediately recovered. Clearly, (4) can accommodate many other specific situations than the autoregressive system (5).

In [17], the problem of securing the persistence of excitation condition (2) for system (4) has been studied in a stationary context. There, it is shown that if w_t is i.i.d, the regressors are bounded, and the system is deterministically stable according to the following assumption

$$\limsup_{t\to\infty} \rho^{-t} \|A_t A_{t-1} \cdots A_0\| = 0, \text{ almost surely},$$

for some $\rho > 1$, then the persistence of excitation condition is satisfied, provided that

$$E[\phi_i \phi_i^T | \mathcal{F}_t], \forall i \geq t+n, \text{ for some integer n.}$$

The latter condition is satisfied e.g. in the case of the autoregressive system (5).

The restrictive condition of this result is that the system is required to be deterministically stable, a condition which is not satisfied in many cases. Extending this result to milder stability conditions of stochastic type such as the one proposed in [18] and reported in [17] as Assumption 2' is in general impossible. This is shown by a counterexample provided in [17].

In conclusion, the persistence of excitation condition (2) appears to be too stiff for systems with an autoregressive part in general.

4 A generalized excitation condition

As pointed out above, the persistence of excitation condition (2) cannot be verified by signals generated from autoregressive stochastic models in general. One way of relaxing the excitation condition has been presented in [15], where the key idea is to replace the constant lower bound δ by a time-varying random process satisfying certain "excitation" properties. To be precise, the generalized excitation condition introduced in [15] is as follows:

Generalized Excitation Condition:

There exists an integer $h > 0$ such that $\{\lambda_k\} \in S^0(\lambda)$ for some $\lambda \in (0,1)$, where

$$\lambda_k \triangleq \lambda_{min} \left\{ E\left[\sum_{i=kh+1}^{(k+1)h} \frac{\phi_i \phi_i^T}{1+\|\phi_i\|^2} |\mathcal{F}_{kh} \right] \right\},$$

and

$$\begin{aligned} S^0(\lambda) \quad \triangleq \{ \quad & a = \{a_k, k \geq 0\} |\ a_k \in [0,1], \\ & E\Pi_{j=i+1}^k (1-a_j) \leq M\lambda^{k-i}, \\ & \forall k \geq i \geq 0, \text{for some } M > 0\}. \end{aligned}$$

It is quite obvious that the excitation condition (2) is a special case of the above generalized excitation condition, because for any constant $\delta \in (0,1)$, we have $\{\delta\} \in S^0(\lambda)$ with $\lambda = 1-\delta$. Furthermore, it was shown in Section 4 of [15] that the above generalized excitation condition is a unified stability condition for all the three standard tracking algorithms, i.e. Kalman filtering algorithm, least-mean-squares algorithm and the forgetting factor RLS algorithm. Tracking error bounds of these three algorithms have also been established under the same condition there.

Now, a natural question is: can we verify the generalized excitation condition for a nontrivial class of time-varying autoregressive processes? To give a concrete answer, let us consider again the autoregressive process (5) and assume that $\{v_t\}$ is a zero mean independent sequence that is independent of ϕ_0 and satisfies

$$\inf_t Ev_t^2 > 0, \quad \sup_t E|v_t|^9 < \infty.$$

It goes without saying that the random coefficient process $\{\theta_t\}$ determines whether or not the regressor process $\{\phi_t\}$ satisfies the generalized excitation condition.

The following two examples were provided and analyzed in [15].

Example 1.

Let $\{A_t\}$ in (6) be an independent random sequence which is independent of $\{v_t\}$. If there exists $\delta \in (0,1)$ such that

$$\sup_k E\|\Pi_{i=kp}^{(k+1)p-1} A_i\|^4 < \delta$$

and

$$\sup_k E\|A_k\|^q < \infty,$$

where $q \triangleq \max\{4, 2(p+1)\}$, then the regressor $\{\phi_t\}$ generated by the autoregressive model satisfies the generalized excitation condition.

Next, we present an example where the random coefficient vector $\{\theta_t\}$ is not an independent sequence.

Example 2.

Let θ_t be decomposed as

$$\theta_t = \theta + e_t,$$

where θ corresponds to a stable AR(p) model, and e_t is a random vector process which is generated by a stable ARMA model:

$$e_t + F_1 e_{t-1} + \cdots + F_{t-q} e_{t-q} = w_t + G_1 w_{t-1} + \cdots + G_r w_{t-r},$$

where $\{w_t\}$ is a Gaussian white noise sequence which is independent of $\{v_t\}$ with small variance. Then $\{\phi_t\}$ generated by (5) satisfies the generalized excitation condition.

Of course, the generalized excitation condition can also include many other interesting cases which cannot be verified by the excitation condition (1).

5 Discussion

Apparently, the overview of the persistence of excitation literature described in the previous sections does not provide a complete answer to a number of questions that can be posed in this context. And, indeed, the existing literature is not complete and many problems remain open at the present stage of knowledge. The purpose of this section is to better focus on such problems.

As we have seen, condition (2) is in some well-defined sense a sufficient as well as necessary condition for tracking properties for LMS and RLS with forgetting factor in the case of ϕ-mixing processes. On the other hand, ϕ-mixing is not general enough for many applications (in particular, in relation to autoregressive systems). The obvious question then is:

i) how far is (2) from being necessary in other, more general, contexts?

All identification settings that have been discussed in previous sections and for which a persistence of excitation condition has been proven require a stability assumption of some sort on the true system. These conditions entail that the observation vector ϕ_t remains bounded in some sense.

ii) is it necessary to impose a stability-like condition in order to prove (2)?

An intuitive reasoning would suggest that if the system to be identified is unstable, than more information is available and, consequently, its parameters should be better estimated by an optimal algorithm. On the other hand, apart from the Kalman filter algorithm in a specific Gaussian context, for all other algorithms no claim of optimality is possible.

Technically speaking, in the unstable case the term $\|\phi_i\|^2$ appearing at the denominator of (2) (and also of the generalized excitation condition) grows unbounded. In order to prove the excitation condition, also the numerator has to grow unbounded, and this requires that ϕ_t becomes large in all the directions of the parameter space. If the observation vector is large in certain directions only, then the excitation condition cannot be expected to be satisfied.

No doubt one can conceive unstable systems with an autoregressive part such that the observation vector grows unbounded in certain directions only, so that the excitation condition is violated. Now, is it true that the identification algorithms have no tracking capabilities in theses cases? The LMS algorithm is not direction selective. This means that if the observation vector is large, then it looses responsiveness in all directions in the parameter space and may become unable to track variations in the parameters. This same rationale does not apply however to other algorithms, in particular the Kalman filter. For this algorithm, even the generalized excitation condition may be too demanding in order to guarantee tracking properties.

iii) is it possible to conceive different persistence of excitation conditions for the Kalman filter algorithm that allow to treat unstable systems?

Much work in the literature on time varying system identification has been devoted to establish stability results and the boundedness of the tracking error.

iv) is it possible to work out approximate (though tight) expressions relating the variance of the tracking error to the excitation level?

A quantitative bound of the tracking error is provided in [6] under general conditions. In particular, no weak dependence conditions on the regressors are required. While this result is of great theoretical interest because it explicitely shows the dependence of the tracking error on fundamental estimation variables (such as the memory length of the algorithm, the parameter variability and the noise variance), admittedly the provided bound is not tight.

In [14], the problem of quantifying the tracking error has been studied under some weak dependence conditions on the regressors. There, it is shown that the tracking error variance is approximated by means of the solution to a deterministic equation (where certain stochastic variables are replaced by their expected value). This result provides quantitative reliable bounds in the case when the regressor dependence vanishes fast enough.

References

[1] K. Astrom. *Introduction to stochastic control theory.* Academic Press, 1970

[2] K. Astrom and P. Eykhoff. System identification - a survey *Automatica*, 7:123–167, 1971.

[3] L. Guo. Estimating time-varying parameters by the Kalman filter based algorithm: stability and convergence. *IEEE Trans. on Automatic Control*, AC-35:141–147, 1990.

[4] S. Bittanti S., M.. Campi: Tuning the forgetting factor in RLS identification algorithms *30th IEEE Conference on Decision and Control*, Brighton (UK), 1991.

[5] S. Bittanti and M. C. Campi. Bounded error identification of time-varying parameters by RLS techniques. *IEEE Trans. on Automatic Control*, AC-39:1106–1110, 1994.

[6] M. C. Campi. Exponentially weighted least squares identification of ime-varying systems with white disturbances. *IEEE Trans. on Signal Processing*, SP-42:2906–2914, 1994.

[7] E. Eweda and O. Macchi. Tracking error bounds of adaptive nonstationary filtering. *Automatica*, 21:293–302, 1985.

[8] L. Ljung and S. Gunnarson. Adaptive tracking in system identification - a survey. *Automatica*, 26:7–22, 1990.

[9] S. Bittanti and M. C. Campi. Adaptive RLS algorithms under stochastic excitation - l^2 convergence analysis. *IEEE Trans. on Automatic Control*, AC-36:963–967, 1991.

[10] S. Bittanti and M. C. Campi. Adaptive RLS algorithms under stochastic excitation - strong consistency analysis. *Syst. Contr. Lett.*, 17:3–8, 1991.

[11] M. C. Campi. Performance of RLS identification algoritms with forgetting factor: a ϕ-mixing approach. *J. of Mathematical Systems, Estimation and Control*, :29–53, 1997.

[12] L. Guo, L. Ljung, and P. Priouret. Performance analysis of the forgetting factor RLS algorithm. *Int. J. Adapt. Control and Signal Process.*, pages 525–537, 1993.

[13] L. Guo and L. Ljung. Exponential stability of general tracking algorithms. *IEEE Trans. on Automatic Control*, AC-39:1376–1387, 1995.

[14] L. Guo and L. Ljung. Performance analysis of general tracking algorithms. *IEEE Trans. on Automatic Control*, AC-39:1388–1401, 1995.

[15] L.Guo. Stability of recursive stochastic tracking algorithms. *SIAM J. Control and Optimization*, Vol.32, No.5 : 1195-1225, 1994

[16] L.Guo, L.Ljung and G-J. Wang. Necessary and sufficient conditions for stability of LMS. *IEEE Trans. on Automatic Control*, Vol.42, No.6 : 761-770, 1997.

[17] S. Bittanti and M. C. Campi. Persistence of excitation properties for time-varying utoregressive systems. *SIAM J. Control and Optimization*, 2000, to appear.

[18] V. Solo. On the stability of slowly time-varying linear systems. *Mathematics of Control Signals and Systems*, pages 331–350, 1994.