PERGAMON

Automatica 36 (2000) 1009–1015

www.elsevier.com/locate/automatica

Brief Paper

Identification of linear time-varying systems using a modified least-squares algorithm☆

Rogelio Lozano*, Dimitrios Dimogianopoulos, Robert Mahony

Université de Technologie de Compiègne, UMR-CNRS 6599, Heudiasyc, Centre de Recherches de Royallieu, BP 20529, 60205 Compiegne Cedex, France

Received 9 February 1998; revised 2 November 1998; received in final form 3 November 1999

Abstract

In this paper a modified version of the standard least-squares algorithm is presented. The aim is to use the proposed modified LS algorithm in linear time-varying systems. The proposed modification involves the addition of extra terms to both the parameter estimates' and the covariance's update laws. We establish a series of properties on the identification error, the parameter estimates and the covariance matrix. These properties are important in an adaptive control context, because LS algorithms provide an easy way of modifying the parameter estimates in order to avoid singular points in the control scheme. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Least-squares estimation; Time-varying plants; Covariance matrix

1. Introduction

Most of the literature on adaptive control deals with time-invariant systems. A few papers have dealt with time-varying systems, mainly Salgado, Goodwin and Middleton (1988), Middleton et al. (1988), Landau et al. (1997), Slotine and Li (1991), Tsakalis and Ioannou (1993) among others. Goodwin and Middleton (1988) and Tsakalis and Ioannou (1993) use only gradient-type parameter estimation algorithms. Least-squares identification algorithms have been proposed by Salgado et al. (1988) and Slotine and Li (1991). In Salgado et al. (1988) a LS identification algorithm was proposed for linear time-invariant (LTI) systems. It is suggested that the proposed algorithm would work in the time-varying case but the analysis is not included. Slotine and Li (1991) proposed two LS identification algorithms with exponential forgetting and presented some convergence properties. Further properties are however required to design an adaptive control algorithm.

In the LTI case least-squares parameter estimation algorithms have been shown to have several advantages over gradient-type algorithms. They minimize a performance index on the square identification error, the "covariance" matrix contains information on the system that can be used to modify the parameters so that they remain in the admissible region, etc. Nevertheless, standard LS identification algorithm are not suitable to estimate time-varying parameters. None of the properties obtained in the LTI case will hold in general in the time-varying case. There is a need of high-performance algorithms that are robust with respect to small time parameter variations. These algorithms should be able to operate for arbitrarily long periods of time without requiring any re-initialization. It is clear that it would be very useful to extend the LS algorithm to obtain such identification algorithms for LTV parameter systems.

The new idea in the proposed algorithm is the combination in the LTV context of the following two elements:

(a) the σ-modification in the update law of the estimate of the parameter vector θ, and
(b) the inclusion of extra terms multiplied by μ in the update law of the covariance matrix $P(t)$.

☆The original version of this paper was presented at the IFAC Workshop on Adaptive Control and Signal Processing which was held in Glasgow, Scotland during August 26–28, 1998. This paper was recommended for publication in revised form by Associate Editor J.W. Polderman under the direction of Editor F.L. Lewis.

* Corresponding author. Tel.: + (0)3-44-23-44-95; fax: + (0)3- 44-23-44-77.

E-mail address: rlozano@hds.utc.fr (R. Lozano).

PII: S0005-1098(00)00010-8

Element (a) has been proposed by Tsakalis and Ioannou (1993) mainly as part of a gradient-type algorithm to assure convergence of the estimated parameters to a given parameter set. Element (b) has been proposed in a slightly more complicated form by Salgado et al. (1988) for use with LTI systems. See also De Mathelin and Lozano (1999) for a direct least squares estimation. The goal of this paper is to use these two elements together and provide the complete analysis for the case of LTV systems. Hence, we are able to establish μ-small in the mean properties on the identification error and various important signals. We must finally point out that the LTI LS algorithm is obtained as a special case when $\mu = 0$. Properties obtained here are useful for designing an adaptive controller. In that case stability will be ensured for small values of μ and an explicit upper bound could then be computed.

2. Identification of time-varying systems

We will be dealing with systems for which we can obtain, after filtering with a *Hurwitz* filter, the following compact form:

$$y(t) = \phi(t)^{\mathrm{T}}\theta^*(t) + w(t), \tag{1}$$

where $y(t)$ is the output, $\phi(t)$ the normalized regressor i.e. $\|\phi(t)\| \le 1$, $\theta^*(t)$ the time-varying parameters and $w(t)$ is a term depending on the error due to swapping terms or unmodeled dynamics (or eventually noise corrupting the system). In Appendix A we show how an expression as (1) can be obtained for linear time-varying systems.

2.1. Assumptions

We make the following assumptions:

(1) The change ratio of the true parameters is known and sufficiently small, i.e.

$$\|\dot{\theta}^*(t)\| \le \varepsilon. \tag{2}$$

The properties of the identification scheme to be proposed will be meaningful only if the upper bound ε is small as will be seen later.

(2) An upper bound M for $\|\theta^*(t)\|$ exists (but may be unknown), i.e.

$$\|\theta^*(t)\| \le M. \tag{3}$$

(3) The disturbance $w(t)$ verifies

$$|w(t)| \le c\varepsilon \sup_{0\le\tau\le t} (\mathrm{e}^{-0.5\xi(t-\tau)}\|\phi(\tau)\|), \tag{4}$$

where ξ, c are constants depending on the chosen filter.

Remark. Assumption (3) means that the *total filtered and normalized* noise $w(t)$ should (roughly) be bounded by both:

(a) the maximum value of the filtered state-norm signal $\|\phi\|$ over a given time interval and
(b) the maximum variation ε of the system's parameter vector $\theta(t)$.

Condition (a) is quite usual in the sense that it is often found even in the LTI case. Condition (b) is reasonable enough, since the data set (u, y) should contain enough information to assure successful tracking of the TV parameters. Thus any persistent bounded noise can be tolerated by the identification scheme provided that its contribution to the set (u, y) is less significant than the effects due to parameter variation.

3. Description of the modified LS algorithm

There are basically two kinds of identification algorithms that are extensively used in adaptive control: least-squares-type algorithms and gradient-type algorithms. The properties of these algorithms have been thoroughly studied in the LTI case. Our aim is to derive similar properties for LS in the LTV case.

We propose the following least-squares modified algorithm for LTV systems as in (1):

$$\dot{\hat{\theta}}(t) = P(t)\phi(t)e_r(t) - \sigma P(t)\hat{\theta}(t), \tag{5}$$

$$\dot{P}(t) = -P(t)^{\mathrm{T}}\phi(t)\phi(t)^{\mathrm{T}}P(t) + \mu(I + \tfrac{1}{2}(P(t)^{\mathrm{T}} + P(t)) - P(t)^{\mathrm{T}}P(t)) \tag{6}$$

with $P(0) = P(0)^{\mathrm{T}} = P_0 > 0$ being the initial value of $P(t)$. In (5), (6) $\hat{\theta}(t)$ is the estimate of the system parameters, $P(t)^{-1}$ the "covariance" matrix and e_r the normalized identification error defined as

$$e_r(s) = y(s) - \phi(s)^{\mathrm{T}}\hat{\theta}(t), \quad \forall s \in [0, t]. \tag{7}$$

In the previous equation the error $e_r(s)$ is obtained using the most recent estimate vector $\hat{\theta}(t)$. This means that in order to show the dependence of the identification error on t it would be more appropriate to write $e_r(s, t)$. We have chosen to use the simplified term $e_r(s)$ for the sake of simplicity. The properties of this algorithm are established in the following theorem.

Theorem 1. *Let $P(t)$ be the solution of Eq.* (6) *for some initial condition $P(0) = P(0)^{\mathrm{T}} = P_0 > 0$. Then the LS-type algorithm in Eqs.* (5) *and* (7) *when applied to the system in* (1) *verifies the following properties*:

(a) $P(t) = P(t)^{\mathrm{T}}$.
(b) $P(t) > 0$, $R(t) = P(t)^{-1} > 0$ *and* $3\mu \le tr(P) \le K_1$ *for some positive constant K_1 independent of μ.*
(c) *The parametric distance $\tilde{\theta}(t)$, where $\tilde{\theta}(t) = \hat{\theta}(t) - \theta^*(t)$, is bounded.*
(d) *The identification error $e_r(t)$ satisfies*

$$\frac{1}{2}\int_t^{t+T} e_r(\tau)^2\,\mathrm{d}\tau \le q_0 + \mu r_0 T \quad \forall \tau, T > 0, \tag{8}$$

where q_0, r_0 are positive constants.

(e)
$$\int_t^{t+T} (P(\tau)\phi(\tau))^{\mathrm{T}} P(\tau)\phi(\tau)\,\mathrm{d}\tau \le q_1 + \mu r_1 T \quad \forall \tau, T > 0, \tag{9}$$

where q_1, r_1 are positive constants.

(f) *The rate of change of $P(t)$ satisfies*
$$\int_t^{t+T} \|\dot{P}(\tau)\|\,\mathrm{d}\tau \le q_2 + \mu r_2 T \quad \forall \tau, T > 0, \tag{10}$$

where q_2, r_2 are positive constants.

(g) *The rate of change of $\hat{\theta}(t)$ satisfies*
$$\int_t^{t+T} \|\dot{\hat{\theta}}(\tau)\|\,\mathrm{d}\tau \le q_3 + \mu r_3 T \quad \forall \tau, T > 0, \tag{11}$$

where q_3, r_3 are positive constants.

Proof. (a) From (6), it is clear that $\dot{P}(t) = \dot{P}(t)^{\mathrm{T}}$. Given that $P(0) = P(0)^{\mathrm{T}} > 0$, by integration we obtain $P(t) = P(t)^{\mathrm{T}}$. Now that we have established that $P(t)$ is symmetric we may rewrite (6) in a more convenient way:

$$\dot{P}(t) = -P(t)\phi(t)\phi(t)^{\mathrm{T}}P(t) + \mu(I + P(t) - P(t)^2). \tag{12}$$

This form of the update law of $P(t)$ will be used throughout the rest of the paper.

(b) Consider the trace of $\dot{P}(t)$

$$tr(\dot{P}(t)) = \frac{\mathrm{d}}{\mathrm{d}t} tr(P(t)) = -\phi(t)^{\mathrm{T}}P(t)^2\phi(t) + \mu\bar{n} + \mu\, tr(P(t)) - \mu\, tr(P(t)^2), \tag{13}$$

where $\bar{n} = 2n$ where n is the order of the plant. Recall that

$$tr(P(t)) = \sum_{i=1}^{\bar{n}} \lambda_i, \qquad tr(P(t)^2) = \sum_{i=1}^{\bar{n}} \lambda_i^2,$$

where λ_i is an eigenvalue of $P(t)$.

Then for any $k > 0$

$$\left(k - \frac{1}{k}\lambda_i\right)^2 = 2\left(\frac{k^2}{2} + \frac{1}{2k^2}\lambda_i^2 - \lambda_i\right) \ge 0$$

and $k^2/2 + (1/2k^2)\lambda_i^2 \ge \lambda_i$. Hence summing from $i = 1, \ldots, \bar{n}$ yields

$$tr(P(t)) \le \bar{n}\,\frac{k^2}{2} + \frac{1}{2k^2} tr(P(t)^2). \tag{14}$$

From (13) and the above

$$\begin{aligned} tr(\dot{P}(t)) &= \frac{\mathrm{d}}{\mathrm{d}t} tr(P(t)) \\ &\le \mu\left[\bar{n} + \bar{n}\frac{k^2}{2} + \frac{1}{2k^2} tr(P(t)^2) - tr(P(t)^2)\right] \\ &\le \mu\left[\bar{n} + \bar{n}\frac{k^2}{2} - \bar{k}\, tr(P(t)^2)\right], \end{aligned} \tag{15}$$

where $\bar{k} = (1/2k^2 - 1)$. Choosing $k > \frac{1}{\sqrt{2}}$ it follows that $\bar{k} > 0$. From (14) it follows that *if* $P(t)$ is such that

$$tr(P(t)) > \bar{n}\,\frac{k^2}{2} + \frac{1}{2k^2}K^2 = K_P, \quad \text{with } K^2 = \frac{\bar{n}(1 + k^2/2)}{\bar{k}}, \tag{16}$$

then $tr(P(t)^2) > K^2$ which from (15) implies that $tr(\dot{P}(t)) < 0$. We therefore conclude that

$$tr(P(t)) < \max(tr(P_0), K_P) = K_1, \tag{17}$$

Let us now obtain an upper bound for $R(t) = P(t)^{-1}$. Differentiating the identity $P(t)R(t) = I$, we have

$$\dot{R}(t) = -R(t)\dot{P}(t)R(t), \tag{18}$$

$$\dot{R}(t) = \phi(t)\phi(t)^{\mathrm{T}} + \mu(I - R(t) - R(t)^2). \tag{19}$$

Note that, as $(\lambda_i(R) - 1)^2 \ge 0$ where $\lambda_i(R)$ is an eigenvalue of $R(t)$, we can obtain

$$\sum_{i=1}^{\bar{n}} (\lambda_i(R)^2 - 2\lambda_i(R) + 1) \ge 0$$

or

$$tr(R(t)^2) \ge 2\, tr(R(t)) - \bar{n}. \tag{20}$$

From (19) and (20) we get

$$\begin{aligned} tr(\dot{R}(t)) &= \|\phi(t)\|^2 + \mu\bar{n} - \mu\, tr(R(t)) - \mu\, tr(R(t)^2) \\ &\le 1 + 2\mu\bar{n} - 3\mu\, tr(R(t)). \end{aligned} \tag{21}$$

Thus if $R(t)$ is such that $tr(R(t)) > (1 + 2\mu\bar{n})/3\mu$ then $tr(\dot{R}(t)) < 0$ and thus

$$tr(R(t)) < \max\left(tr(R_0), \frac{1 + 2\mu\bar{n}}{3\mu}\right). \tag{22}$$

Thus $P(t)$ and $R(t)$ exist and are bounded. Note that P can become singular only when $t \to \infty$ and $\mu \to 0$.

(c) In order to study the properties of the proposed identification scheme the following function V is considered:

$$V(t) = \tfrac{1}{2}\tilde{\theta}(t)^{\mathrm{T}}R(t)\tilde{\theta}(t), \quad \text{where } \tilde{\theta}(t) = \hat{\theta}(t) - \theta^*(t). \tag{23}$$

Differentiating and using (5), (19)

$$\begin{aligned} \dot{V} &= \dot{\tilde{\theta}}^{\mathrm{T}}R\tilde{\theta} + \tfrac{1}{2}\tilde{\theta}^{\mathrm{T}}\dot{R}\tilde{\theta} \\ &= \dot{\tilde{\theta}}^{\mathrm{T}}R\tilde{\theta} + \tfrac{1}{2}\tilde{\theta}^{\mathrm{T}}(\phi\phi^{\mathrm{T}} + \mu I - \mu R - \mu R^2)\tilde{\theta} \\ &= (\phi^{\mathrm{T}}Pe_r - \sigma\hat{\theta}^{\mathrm{T}}P - \dot{\theta}^{*\mathrm{T}})R\tilde{\theta} \\ &\quad + \tfrac{1}{2}\tilde{\theta}^{\mathrm{T}}(\phi\phi^{\mathrm{T}} + \mu I - \mu R - \mu R^2)\tilde{\theta} \\ &= e_r\phi^{\mathrm{T}}\tilde{\theta} - \sigma\hat{\theta}^{\mathrm{T}}\tilde{\theta} - \dot{\theta}^{*\mathrm{T}}R\tilde{\theta} + \tfrac{1}{2}(\phi^{\mathrm{T}}\tilde{\theta})^2 \\ &\quad + \tfrac{1}{2}\mu\|\tilde{\theta}\|^2 - \mu V - \tfrac{1}{2}\mu\|R\tilde{\theta}\|^2. \end{aligned} \tag{24}$$

Let $\sqrt{\mu/2} = \delta$. Then

$$-\dot{\theta}^{*\mathrm{T}}R\tilde{\theta} = -\left\|\delta R\tilde{\theta} + \frac{\dot{\theta}^*}{2\delta}\right\|^2 + \delta^2\|R\tilde{\theta}\|^2 + \left(\frac{\|\dot{\theta}^*\|}{2\delta}\right)^2.$$

Introducing the above into (24) and using (7) we obtain

$$\dot{V} = \frac{1}{2}w^2 - \frac{1}{2}e_r^2 - \sigma\hat{\theta}^{\mathrm{T}}\tilde{\theta} + \left(\frac{\|\dot{\theta}^*\|}{2\delta}\right)^2 - \left\|\delta R\tilde{\theta} + \frac{\dot{\theta}^*}{2\delta}\right\|^2 + \frac{1}{2}\mu\|\tilde{\theta}\|^2 - \mu V. \tag{25}$$

Consider the following equation resulting from the perfect square $\|(\sigma/2\delta)\hat{\theta} - \delta\tilde{\theta}\|^2$:

$$-\sigma\hat{\theta}^{\mathrm{T}}\tilde{\theta} = \left\|\frac{\sigma}{2\delta}\hat{\theta} - \delta\tilde{\theta}\right\|^2 - \delta^2\|\tilde{\theta}\|^2 - \left(\frac{\sigma}{2\delta}\right)^2\|\hat{\theta}\|^2, \tag{26}$$

choosing $\sigma = 2\delta^2$ we get

$$-\sigma\hat{\theta}^{\mathrm{T}}\tilde{\theta} = \delta^2\|\theta^*\|^2 - \delta^2\|\tilde{\theta}\|^2 - \left(\frac{\sigma}{2\delta}\right)^2\|\hat{\theta}\|^2 \tag{27}$$

Introducing (27) into (25) and using (2) and (3) we obtain

$$\begin{aligned}\dot{V} &\le \frac{1}{2}w^2 - \frac{1}{2}e_r^2 + \left(\frac{\varepsilon}{2\delta}\right)^2 + \delta^2\|\theta^*\|^2 - \left(\frac{\sigma}{2\delta}\right)^2\|\hat{\theta}\|^2 - \mu V \\ &\le \frac{1}{2}w^2 - \frac{1}{2}e_r^2 + \left(\frac{\varepsilon}{2\delta}\right)^2 + \delta^2 M^2 - \mu V. \end{aligned}\tag{28}$$

From (28) we observe that the term $V_1 = (\varepsilon/2\delta)^2 + \delta^2 M^2$ should be as small as possible. A reasonable choice for δ is the one which minimizes V_1. Thus, we consider the constraint $\partial V_1/\partial\delta = 0$. Its solution is $\delta = \delta_{\mathrm{opt}} = \sqrt{\varepsilon/2M}(=\sqrt{\mu/2})$. From assumption (3) and the fact that $\|\phi\| \le 1$ it follows that

$$\tfrac{1}{2}w^2 \le \tfrac{1}{2}c^2\varepsilon^2 \tag{29}$$

Introducing (29) in (28) we obtain

$$\begin{aligned}\dot{V} &\le -\frac{1}{2}e_r^2 + \frac{1}{2}c^2\varepsilon^2 + \varepsilon M - \mu V \\ &\le -\frac{1}{2}e_r^2 + \frac{\varepsilon}{M}\frac{l}{2} - \frac{\varepsilon}{M}V \end{aligned}\tag{30}$$

or

$$\dot{V} \le -\frac{1}{2}e_r^2 - \frac{\varepsilon}{M}\left(V - \frac{l}{2}\right),$$

where $l = 2(M^2 + \frac{1}{2}c^2\varepsilon M)$. If $V \ge \frac{l}{2}$ then $\dot{V} \le 0$. Consequently, $l/2$ is the upper bound of V and, consequently, we have

$$\begin{aligned} V &= \frac{1}{2}\tilde{\theta}^{\mathrm{T}}R\tilde{\theta} \le \frac{l}{2}, \\ \|\tilde{\theta}\| &\le \sqrt{\frac{l}{(\lambda_{\min}R)}}, \end{aligned}\tag{31}$$

where $(\lambda_{\min}R)$ denotes the smallest of the minimum eigenvalues of $R(t)$ for all t. This lower bound exists (see part (b) of theorem) and is independent of μ. In fact,

$$\begin{aligned}(\lambda_{\min}R) &= \left(\frac{1}{(\lambda_{\max}P)}\right) \\ &\ge \max\left(\frac{1}{tr(P)}\right) \quad \text{(using (17))} \\ &\ge K_1^{-1}. \end{aligned}\tag{32}$$

Then from (31) and (32) we conclude

$$\|\tilde{\theta}\| \le \sqrt{lK_1}. \tag{33}$$

Moreover $\|\tilde{\theta}(t)\| = \|\hat{\theta}(t) - \theta^*(t)\|$ and

$$\big|\|\hat{\theta}(t)\| - \|\theta^*(t)\|\big| \le \|\hat{\theta}(t) - \theta^*(t)\| \le \|\hat{\theta}(t)\| + \|\theta^*(t)\|. \tag{34}$$

Consequently from assumption (2) and the LHS of (34), we have

$$\|\theta^*(t)\| - \|\hat{\theta}(t) - \theta^*(t)\| \le \|\hat{\theta}(t)\| \le \|\theta^*(t)\| + \|\hat{\theta}(t) - \theta^*(t)\|$$

or (using (33))

$$\|\hat{\theta}(t)\| \le M + \sqrt{lK_1}. \tag{35}$$

(d) From (30) it follows

$$\dot{V} \le -\frac{1}{2}e_r^2 - \mu\left(V - \frac{l}{2}\right) \le -\frac{1}{2}e_r^2 + \mu\frac{l}{2}. \tag{36}$$

Integrating the above yields:

$$\begin{aligned}\frac{1}{2}\int_t^{t+T} e_r(\tau)^2\,\mathrm{d}\tau &\le -V(t+T) + V(t) + \mu\tfrac{l}{2}T \\ &\le V(t) + \mu\tfrac{l}{2}T \\ &\le q_0 + \mu r_0 T \quad \forall\tau, T > 0. \end{aligned}\tag{37}$$

(e) Integrating (12) we get

$$\begin{aligned}\int_t^{t+T} P(\tau)\phi(\tau)\phi(\tau)^{\mathrm{T}}P(\tau)\mathrm{d}\tau &= P(t) - P(t+T) + \mu\int_t^{t+T}(I + P(\tau) - P(\tau)^2)\,\mathrm{d}\tau \\ &\le P(t) + \mu T I + \mu\int_t^{t+T} P(\tau)\,\mathrm{d}\tau. \end{aligned}\tag{38}$$

Consider the trace of the above expression:

$$\begin{aligned} tr&\left(\int_t^{t+T} P(\tau)\phi(\tau)\phi(\tau)^{\mathrm{T}}P(\tau)\,\mathrm{d}\tau\right) \\ &\le tr(P) + \bar{n}\mu T + \mu\int_t^{t+T} tr(P(\tau))\,\mathrm{d}\tau \\ &\le tr(P) + \bar{n}\mu T + \mu K_1 T, \end{aligned}\tag{39}$$

where K_1 is defined in (17). Thus,

$$\int_t^{t+T} tr(P(\tau)\phi(\tau)\phi(\tau)^{\mathrm{T}}P(\tau))\,\mathrm{d}\tau = \int_t^{t+T} \phi(\tau)^{\mathrm{T}}P(\tau)^2\phi(\tau)\,\mathrm{d}\tau \le tr(P) + \mu(\bar{n} + K_1)T \le q_1 + \mu r_1 T \quad \forall \tau, T > 0. \tag{40}$$

(f) From (12) one has

$$\|\dot{P}(t)\| \le \|P(t)\phi(t)\phi(t)^{\mathrm{T}}P(t)\| + \mu\|I + P(t) - P(t)^2\|. \tag{41}$$

Integrating we get

$$\int_t^{t+T} \|\dot{P}(\tau)\|\,\mathrm{d}\tau \le \int_t^{t+T} \|\phi(\tau)^{\mathrm{T}}P(\tau)^2\phi(\tau)\|\,\mathrm{d}\tau + \mu\int_t^{t+T} [\|I\| + \|P(\tau)\| + \|P(\tau)^2\|]\,\mathrm{d}\tau \tag{42}$$

and from (40)

$$\int_t^{t+T} \|\dot{P}(\tau)\|\,\mathrm{d}\tau \le q_1 + \mu r_1 T + \mu(\|I\| + K_1 + K_1^2)T \le q_1 + \mu(r_1 + \|I\| + K_1 + K_1^2)T \le q_2 + \mu r_2 T. \tag{43}$$

(g) Taking norms in (5), we have

$$\|\dot{\hat{\theta}}(t)\| = \|P(t)\phi(t)e_r(t)\| + \|\sigma P(t)\hat{\theta}(t)\| \le \|P(t)\phi(t)\|\ |e_r(t)| + \sigma\|P(t)\hat{\theta}(t)\|. \tag{44}$$

Notice that

$$|e_r|\ \|P\phi\| = \|P\phi\|^2 + \tfrac{1}{4}|e_r|^2 - (\|P\phi\| - \tfrac{1}{2}|e_r|)^2 \le \|P\phi\|^2 + \tfrac{1}{4}|e_r|^2. \tag{45}$$

Thus, from (44) and (45)

$$\|\dot{\hat{\theta}}(t)\| \le \|P\phi\|^2 + \tfrac{1}{4}|e_r|^2 + \sigma\|P(t)\hat{\theta}(t)\|. \tag{46}$$

Recall that $\sigma = 2\delta^2 = \mu$. Integrating (46) and using (35) we obtain

$$\int_t^{t+T} \|\dot{\hat{\theta}}(\tau)\|\,\mathrm{d}\tau \le \int_t^{t+T} \|P(\tau)\phi(\tau)\|^2\,\mathrm{d}\tau + \frac{1}{4}\int_t^{t+T} |e_r(\tau)|^2\,\mathrm{d}\tau + \mu K_1(\sqrt{lK_1} + M)\int_t^{t+T}\mathrm{d}\tau \tag{47}$$

or from (40), (37)

$$\int_t^{t+T} \|\dot{\hat{\theta}}(\tau)\|\,\mathrm{d}\tau \le q_1 + \mu r_1 T + \tfrac{1}{2}(q_0 + \mu r_0 T) + \mu K_1(\sqrt{lK_1} + M)T \le q_3 + \mu r_3 T. \qquad \square \tag{48}$$

Remark. (1) The μ-small in the mean properties are meaningful only if ε (and as a consequence μ) is small, i.e. if the plant parameters vary *slowly* with time.

(2) Property (e) is required to obtain properties (f) and (g). These two properties together with (d) are essential to design an adaptive control system, in the sense that the signals $\|\dot{\hat{\theta}}(t)\|$ and $\|\dot{P}(t)\|$ should be as smooth as possible to prevent from having abrupt changes in the controller parameters. In this case stability will be ensured for small values of μ and an explicit upper bound could be computed.

(3) If we use the proposed algorithm for a LTI system we will not obtain better results than if we had used a standard LS algorithm. However a standard LS algorithm will not be able to handle LTV systems while the proposed algorithm will.

(4) As we have mentioned in the introduction a simpler modified LS algorithm was proposed in Slotine and Li (1991). In order to make a comparison, let us observe the algorithm proposed in this book together with the one proposed in this paper:

$$\dot{R} = \phi\phi^{\mathrm{T}} - \mu(R - K^{-1}) \quad \text{or}$$

$$\dot{P} = -P\phi\phi^{\mathrm{T}}P + \mu(P - K^{-1}P^2) \quad \text{(Salgado et al.)}$$

$$\dot{R} = \phi\phi^{\mathrm{T}} + \mu(I - R - R^2) \quad \text{or}$$

$$\dot{P} = -P\phi\phi^{\mathrm{T}}P + \mu(I + P - P^2) \quad \text{(Lozano et al.)}$$

The reason to include the extra term R^2 in our algorithm can be illustrated if we consider the following experiment:

We make sure that the system's input signal is persistently exciting during a long time. Consequently P becomes very small. Then if the excitation is less significant, P will grow more slowly when using the algorithm of Slotine and Li (1991) than when using the algorithm proposed in this paper. This can be observed by studying the two previous equations of P and can be checked by carrying out simulations. During the time that P is small no tracking of the TV parameters can be assured. Thus, in certain circumstances, the inclusion of the extra term R^2 can enhance the tracking ability of the estimation algorithm.

(5) Although Eq. (12) features common points with a *Riccati* equation, a thorough examination reveals that this is not true. In order to prove this, we will compare the corresponding terms of the two equations. The form of a *Riccati* equation is (Kailath, 1980):

$$\dot{P}(t) = -P(t)F(t) - F(t)^{\mathrm{T}}P(t) - C^{\mathrm{T}}C + P(t)(G(t)R^{-1}G(t))P(t)$$

for the system

$$\dot{x}(t) = F(t)x(t) + G(t)u(t),$$

$$y(t) = Cx(t).$$

Consequently, we have $F = \mu/2(-I)$ which is acceptable, $C^{\mathrm{T}}C = \mu(-I)$ which is not acceptable and $G(t)R^{-1}G(t) = -(\mu I + \phi(t)\phi(t)^{\mathrm{T}})$. The last term means that matrix

R should always be negative definite, which is impossible. Hence (12) cannot be treated as a *Riccati* equation.

4. Conclusion

We have proposed a modified version of the standard LS algorithm that can be applied to time-varying systems. We have proved that the estimates are bounded. We have shown that the different system signals are μ-small in the mean. Current research is underway to apply the proposed algorithm in the adaptive control context.

Acknowledgements

The authors would like to thank the reviewers for their comments to improve the paper.

Appendix A. Obtaining expression (1) for a LTV system

Consider the following linear time-varying system:

$$A(D,t)\bar{y}(t) = B(D,t)\bar{u}(t) \tag{A.1}$$

with

$$\begin{aligned} A(D,t) &= D^n + a_1(t)D^{n-1} + \cdots + a_n(t), \\ B(D,t) &= b_0(t)D^m + b_1(t)D^{m-1} + \cdots + b_m(t), \end{aligned} \tag{A.2}$$

where $\bar{y}(t)$ is the output and $\bar{u}(t)$ the input.

The system's state is not supposed to be measurable. Thus, we will apply a *Hurwitz* filter $F(D) = (D^n + f_1 D^{n-1} + \cdots + f_n)$ in order to obtain a filtered model similar to (A.1) but with a measurable state. This procedure involves:

(a) the filtering of the input and the output,
(b) the derivation of a model of the form $y(t) = \theta(t)^{\mathrm{T}}\phi_f(t) + w(t)$ similar to the original $y(t) = \theta(t)^{\mathrm{T}}\phi(t)$ but with $\phi_f(t)$ being a measurable signal and
(c) the demonstration that the resulting 'noise' term $w(t)$ is small under certain conditions.

Let us define the filtered signals:

$$F\bar{y}_F = \bar{y}, \qquad F\bar{u}_F = \bar{u}. \tag{A.3}$$

The original system (A.1) can be written as

$$\begin{aligned} F\bar{y} &= (F-A)\bar{y} + B\bar{u} \\ &= (F-A)F\bar{y}_F + BF\bar{u}_F \text{ (using (A.3))} \\ &= (F-A)F\bar{y}_F + BF\bar{u}_F \pm F(F-A)\bar{y}_F \pm FB\bar{u}_F \\ &= F(F-A)\bar{y}_F + FB\bar{u}_F \\ &\quad + [(F-A)F - F(F-A)]\bar{y}_F \\ &\quad + [BF - FB]\bar{u}_F. \end{aligned} \tag{A.4}$$

Let us denote

$$\begin{aligned} x(t)^{\mathrm{T}}\theta^*(t) &\doteq (F-A)\bar{y}_F + B\bar{u}_F \\ F\eta_F &\doteq [(F-A)F - F(F-A)]\bar{y}_F + [BF - FB]\bar{u}_F. \end{aligned} \tag{A.5}$$

Then from (A.4) we obtain

$$F[\bar{y}(t) - x(t)^{\mathrm{T}}\theta^*(t) - \eta_F(t)] = F[e_x(t)] = 0, \tag{A.6}$$

where e_x is an exponentially decaying term. Thus,

$$\bar{y}(t) = x(t)^{\mathrm{T}}\theta^*(t) + \eta_F(t) + e_x(t), \tag{A.7}$$

where

$$x(t) = \left[\frac{D^{n-1}}{F}\bar{y}(t), \ldots, \frac{1}{F}\bar{y}(t), \frac{D^m}{F}\bar{u}(t), \ldots, \frac{1}{F}\bar{u}(t)\right]^{\mathrm{T}}.$$

The signal $\eta_F(t)$ arises from the filtering procedure. Let us obtain the complete expression of $\eta_F(t)$. From the second equation in (A.5) it follows:

$$\begin{aligned} F\eta_F &= (FA - AF)\bar{y}_F + (BF - FB)\bar{u}_F, \\ \eta_F &= \sum_{i=1}^{n} (-1)^i \frac{F_i}{F}[B^{(i)}\bar{u}_F(t) - A^{(i)}\bar{y}_F(t)] \end{aligned} \tag{A.8}$$

with

$$\begin{aligned} F_i(D) &= \left(\frac{\partial}{\partial D}\right)^i F(D)\frac{1}{i!}, \\ A^{(i)}(D,t) &= a_1^{(i)}(t)D^{n-1} + \cdots + a_n^{(i)}(t), \\ B^{(i)}(D,t) &= b_0^{(i)}(t)D^m + b_1^{(i)}(t)D^{m-1} + \cdots + b_m^{(i)}(t). \end{aligned} \tag{A.9}$$

Consequently η_F depends on:

(a) the initial values of $\bar{y}(t)$, $\bar{u}(t)$,
(b) the derivatives of $\bar{y}(t)$, $\bar{u}(t)$ and
(c) the time derivatives $a_1^{(i)}(t), \ldots, a_n^{(i)}(t)$ and $b_0^{(i)}(t), \ldots, b_m^{(i)}(t)$, $i = 1, \ldots, n$ in (A.9).

Finally, we normalize (A.7) using the normalization signal $m(t) = 1 + \sup_{0 \le \tau \le t} \|x(\tau)\|$. Thus we obtain (1) with $y(t) = \bar{y}(t)/m(t)$, $\phi(t) = x(t)/m(t)$, $w(t) = \eta_F(t)/m(t)$.

We also have to make an *additional assumption*. The magnitude of the derivatives $a_1^{(i)}(t), \ldots, a_n^{(i)}(t)$ and $b_0^{(i)}(t), \ldots, b_m^{(i)}(t)$, $i = 1, \ldots, n$ in (A.9) should be of order ε as in Assumption 1. This means that $a_{(.)}(t), b_{(.)}(t)$ should be smooth signals, thus ensuring that $w(t)$ is a small bounded quantity. Consequently from (A.8)

$$\begin{aligned} Fw(t) &= \frac{1}{m(t)} \sum_{i=1}^{n} (-1)^i F_i [B^{(i)}\bar{u}_F(t) - A^{(i)}\bar{y}_F(t)] \\ &= \frac{1}{m(t)} \sum_{i=1}^{n} (-1)^i F_i x^{\mathrm{T}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^i \theta^*(t). \end{aligned} \tag{A.10}$$

Let $W(t)=[w^{(n-1)}(t)\dots w^{(0)}(t)]^{\mathrm{T}}$. Then (A.10) can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t}W(t)=G_F W(t)+H_F\left\{\frac{1}{m(t)}\sum_{i=1}^{n}(-1)^i F_i x^{\mathrm{T}}(t)\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^i\theta^*(t)\right\}, \quad \text{(A.11)}$$

where G_F is the state space matrix of the filter F and $H_F=[1,0,\dots,0]^{\mathrm{T}}$. The solution of (A.11) is

$$W(t)=\mathrm{e}^{G_F t}W(0)+\int_0^t \mathrm{e}^{G_F(t-s)}H_F \times\left\{\frac{1}{m(t)}\sum_{i=1}^{n}(-1)^i F_i x^{\mathrm{T}}(s)\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^i\theta^*(s)\right\}\mathrm{d}s.$$

Applying norms and using the *additional assumption* we get

$$\begin{aligned}\|W(t)\| &\le \|W(0)\|\mathrm{e}^{-\xi t}+\int_0^t \mathrm{e}^{-\xi(t-s)}K_1\varepsilon\frac{1}{m(t)}\|x^{\mathrm{T}}(s)\|\,\mathrm{d}s\\ &\le \|W(0)\|\mathrm{e}^{-\xi t}+K_1\varepsilon\int_0^t \mathrm{e}^{-0.5\xi(t-s)}\\ &\quad\times\left(\sup_{0\le s\le t}\mathrm{e}^{-0.5\xi(t-s)}\|\phi(s)\|\right)\mathrm{d}s\\ &\le \|W(0)\|\mathrm{e}^{-\xi t}+K_1\varepsilon\left(\sup_{0\le s\le t}\mathrm{e}^{-0.5\xi(t-s)}\|\phi(s)\|\right)\\ &\quad\times\int_0^t \mathrm{e}^{-0.5\xi(t-s)}\,\mathrm{d}s \qquad \text{(A.12)}\end{aligned}$$

or

$$|w(t)|\le \|W(0)\|\mathrm{e}^{-\xi t}+c\varepsilon\left(\sup_{0\le s\le t}\mathrm{e}^{-0.5\xi(t-s)}\|\phi(s)\|\right)=b_w,$$

where ξ, c depend on the choice of the *Hurwitz* filter. Note that if we consider $W(0)\ne 0$, the term $\|W(0)\|\mathrm{e}^{-\xi t}$ decays exponentially and thus the last of the above inequalities is similar to (4) after a short time. Eq. (A.12) is important, as now $w(t)$ in (1) can be treated as a bounded noise term. The above procedure follows the ideas in Goodwin et al. (1988).

References

De Mathelin, M., & Lozano, R. (1999). Robust adaptive identification of slowly time-varying parameters with bounded disturbances. *Automatica, 35*(7), 1291–1306.

Goodwin, G. C., & Middleton, R. H. (1988). Adaptive control of time-varying linear systems. *IEEE Transactions on Automatic Control*, *33*(2), 150–155.

Landau, I. D., Lozano, R., & M'Saad, M. (1997). *Adaptive Control*. Springer-Verlag CCE Series, 1997.

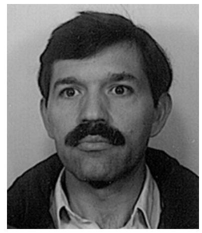
Middleton, R. H., Goodwin, G. C., Hill, D., & Mayne, D. (1988). Design issues in adaptive control. *IEEE Transactions on Automatic Control, 33* (1), 50–58.

Kailath, T. (1980). *Linear systems*. New York: Prentice-Hall.

Salgado, M., Goodwin, G. C., & Middleton, R. (1988). Modified least squares algorithm incorporating exponential resetting and forgetting. *International Journal of Control*, *47*(2), 477–491.

Slotine, J. J. E., & Li, W. (1991). *Nonlinear applied control*. Englewood Cliffs, NJ: Prentice-Hall.

Tsakalis, K. S., & Ioannou, P. A. (1993). *Linear time varying systems: Control and adaptation*. Englewood Cliffs, NJ: Prentice-Hall.



Rogelio Lozano was born in Monterrey Mexico, on July 12, 1954. He received the B.S. degree in electronic engineering from the National Polytechnic Institute of Mexico in 1975, the M.S. degree in electrical engineering from Centro de Investigacion y de Estudios Avanzados (CIEA), Mexico in 1977, and the Ph.D. degree in automatic control from Laboratoire d'Automatique de Grenoble, France, in 1981. He joined the Department of Electrical Engineering at the CIEA, Mexico, in 1981 where he worked until 1989. He was Head of the Section of Automatic Control from June 1985 to August 1987. He has held visiting positions at the University of Newcastle, Australia, from November 1983 to November 1984, NASA Langley Research Center VA, from August 1987 to August 1988, and Laboratoire d'Automatique de Grenoble, France, from February 1989 to July 1990. Since 1990 he is a CNRS Research Director at University of Technology of Compiégne, France. He is head of the Laboratory Heudiasyc, UMR 6599 CNRS since January 1995. Dr Lozano is Associate Editor of Automatica since 1987 and of Int. J. of Adaptive Control and Signal Processing since 1993. His research interest are in adaptive control of linear, nonlinear systems and robot manipulators

Dimitrios Dimogianopoulos was born in Patras Greece, on 14 May 1971. He obtained his degree in Mechanical Engineering from the University of Patras in 1994. He, then, moved to Compiègne in France, where he received his D.E.A (Diplôme d'Études Approfondies) in Systems Control from the University of Technology of Compiègne, in 1995. He obtained his Ph.D. on Adaptive Control of Linear Time Varying Systems from the same University in late 1999. He is currently doing his military service in Greece. His interests are in mechanical and automotive systems, adaptive control theory and applications in motion systems and non linear control.

Robert Mahony obtained a science degree majoring in applied mathematics and geology from the Australian National University (ANU) in 1989. After working for a year as a geophysicist processing marine seismic data he returned to study at ANU and obtained a Ph.D. in systems engineering in 1994. Between 1994 and 1997 he worked as a Research Fellow in the Cooperative Research Centre for Robust and Adaptive Systems based in the Research School of Information Sciences and Engineering, ANU, Australia. From 1997 to 1999 he had a position as a post-doctoral fellow in the CNRS laboratory for Heuristics Diagnostics and complex systems (Heudiasyc), Compiegne University of Technology, France. Having returned to Australia in late 1999, he is now a Logan Fellow in the department of Engineering and Computer Science at Monash University, Melbourne, Australia. His main interests are in non-linear control theory, with applications in mechanical systems and motion systems, mathematical systems theory and optimization techniques with applications in linear algebra and geometry.