INT. J. CONTROL, 1986, VOL. 43, NO. 1, 177–191

Output-error identification methods for partially known systems

SOURA DASGUPTA†, BRIAN D. O. ANDERSON‡
and R. JOHN KAYE§

Output-error methods are applied to the problem of identifying partially known linear time-invariant finite-dimensional systems. The systems considered have a limited number of unknown parameters, which are assumed to enter the numerator and denominator of the transfer function multilinearly. This is representative of mechanical and electrical systems with unknown inertias, frictional coefficients, inductances or resistances. Two output-error schemes are developed and shown to be uniformly asymptotically convergent under certain conditions.

1. Introduction

In this paper we formulate output-error algorithms for the identification of partially known systems. The difference between output- and equation-error identification is illustrated in the Figure. In equation-error schemes the system output y enters the adjustable model twice, directly and through the output error Δy. Consequently terms like $y^2(t)$ appear in the error model. Thus unbiased measurement noise could result in biased parameter estimates. In output error this difficulty is avoided by allowing the output to enter the adjustable model as Δy only.

Exponential convergence of output-error algorithms, however, requires that a certain transfer function be strictly positive real (SPR). Unfortunately this transfer function depends on the unknown system parameters, and the condition for convergence cannot always be checked.

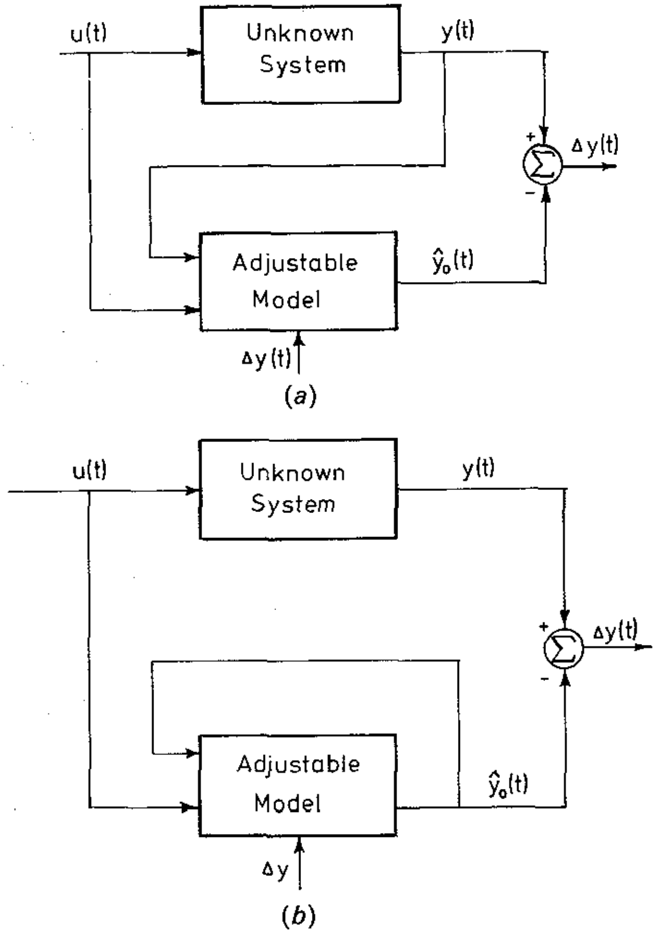
Conventional output-error algorithms (Landau 1979) presuppose a complete lack of knowledge about the unknown system (aside from its degree and relative degree) and ignore all additional knowledge that may well be available. These algorithms thus estimate the numerator and denominator coefficients of the transfer function, having first assumed them all to be unknown. In practice, a great deal of additional knowledge is often available, which if exploited should give rise to parametrizations involving fewer unknowns and better algorithms. Usually, the lack of knowledge in a system relates to certain physical parameter values. Thus all parts of a mechanical system may be known, except perhaps the values of some moments of inertia, frictional coefficients, or the like. Accordingly, in the parametrization considered here the unknown parameters have direct physical significance. Such a parametrization also has the following added attraction: the physics of the system can justify assumptions on the parameter magnitude bounds. Indeed, the algorithms formulated in this paper exploit the knowledge of these assumed magnitude bounds.

Received 6 March 1985.

† Department of Electrical and Computer Engineering, University of Iowa, Iowa City, Iowa 52242, U.S.A.

‡ Department of Systems Engineering, Research School of Physical Sciences, Australian National University, G.P.O. Box 4, Canberra, ACT 2601, Australia.

§ School of Electrical Engineering and Computer Science, University of New South Wales, P. O. Box 1, Kensington NSW 2033, Australia.



(*a*) Equation-error configuration. (*b*) Output-error configuration.

In many linear time-invariant finite-dimensional systems, when most parts are known, but certain parameters associated with physical component values are unknown, the transfer functions can be expressed as the ratios of polynomials having coefficients multilinear in the unknown parameters. For example, when only two parameters are unknown, the transfer function will often have the following form (Dasgupta *et al.* 1984):

$$T(s) = \frac{p_0(s) + k_1 p_1(s) + k_2 p_2(s) + k_1 k_2 p_{12}(s)}{q_0(s) + k_1 q_1(s) + k_2 q_2(s) + k_1 k_2 q_{12}(s)} \tag{1.1}$$

where k_1 and k_2 are the unknown parameters, the polynomials $p_\sigma(s)$ and $q_\sigma(s)$ are known and $\delta[q_0(s)] > \delta[q_\sigma(s)]\ \forall\sigma \neq 0$.

Formulated here are two output-error algorithms which exploit the knowledge of the p_σ and q_σ. As in Dasgupta *et al.* (1984) (which considers equation-error formulations) the algorithms conform to a two-step structure. We illustrate this structure through the two-parameter example of (1.1). The object is to estimate the vector $K = [k_1\ \ k_2\ \ k_1k_2]^{\mathrm{T}}$. The first step obtains an unconstrained estimate $K_u \triangleq [K_{u1}\ \ K_{u2}\ \ K_{u12}]^{\mathrm{T}}$ of K. The estimate of K_u is unconstrained in the sense that K_{u12} need not be the product of K_{u1} and K_{u2}. The second step obtains a constrained estimate

$\hat{K} \triangleq [\hat{k}_1 \ \ \hat{k}_2 \ \ \hat{k}_1\hat{k}_2]^T$ which is the 'closest' to K_u. Simulation results in Dasgupta *et al.* (1984) show that the second step improves upon the accuracy of the parameter estimates.

In both the algorithms presented, the second step is the same and constitutes a steepest-descent algorithm, $(K_u - \hat{K})^T\Lambda(K_u - \hat{K})$ being the minimizing function with Λ a diagonal matrix whose first two diagonal elements are positive and third non-negative. The first step in the first algorithm is gradient descent, while that in the second algorithm is least-squares.

Both the algorithms are shown to be globally uniformly asymptotically stable, whenever the inputs are persistently exciting, a certain transfer function is SPR and the parameter magnitude bounds are known. The convergence proofs require a key lemma, given in the Appendix, which extends a result of Boyd and Sastry (1983).

Sections 2 and 3 present the gradient-descent and least-squares algorithms respectively and analyse them for convergence. Section 4 makes the concluding remarks.

2. A two-step gradient-descent output-error algorithm

Consider the asymptotically stable unknown system

$$Q(s, k)y(t) = P(s, k)u(t) \tag{2.1 a}$$

where k is the N-dimensional unknown parameter vector, $n = \delta[Q(s, k)] \geqslant \delta[P(s, k)]$ and

$$Q(s, k) = q_0(s) + \sum_{\sigma \subset S} \left(\prod_{i \in \sigma} k_i\right) q_\sigma(s), \tag{2.1 b}$$

$$P(s, k) = p_0(s) + \sum_{\sigma \subset S} \left(\sum_{i \in \sigma} k_i\right) p_\sigma(s) \tag{2.1 c}$$

In (2.1), $S \triangleq \{1, 2, \ldots N\}$; $\delta[q_0(s)] \geqslant \delta[q_\sigma(s)]\ \forall \sigma \subset S$ and $p_\sigma(s)$, $q_\sigma(s)$ are known. Usually, N will be small.

Throughout this paper we shall abuse notation by referring, for example, to $y(s)$ as the Laplace transform of $y(t)$. Quantities like $y_i(t) \triangleq (a(s)/b(s))u(t)$ will refer to the solutions of the differential equation $b(p)y_i(t) = a(p)u(t)$ with $p \triangleq d/dt$, with arbitrary finite initial conditions. For vectors such as

$$V(t) \triangleq \left[\frac{p_1(s)y(t) - q_1(s)u(t)}{\beta(s)} \quad \frac{p_2(s)y(t) - q_2(s)u(t)}{\beta(s)} \quad \cdots \right.$$
$$\left. \cdots \quad \frac{p_{12\ldots N}(s)y(t) - q_{12\ldots N}(s)u(t)}{\beta(s)}\right]^T$$

the initial conditions are assumed to be zero. Also, for $\sigma = \{1, 2\}$, for example, $p_\sigma(s)$ will refer to $p_{12}(s)$.

Let K be a vector containing the multilinear combinations of the k_i and let K_u be an unconstrained estimate of K, in the sense described in §1. Thus, for $N = 2$, $K = [k_1 \ \ k_2 \ \ k_1k_2]^T$ and $K_u = [K_{u1} \ \ K_{u2} \ \ K_{u3}]^T$.

Let $\beta(s)$ be a Hurwitz polynomial of degree n. Then for the following adjustable model:

$$\left\{\frac{q_0(s)}{\beta(s)} + \sum_{\sigma \subset S} K_{u\sigma}(t)\frac{q_\sigma(s)}{\beta(s)}\right\} y(t) = \left\{\frac{p_0(s)}{\beta(s)} + \sum_{\sigma \subset S} K_{u\sigma}(t)\frac{p_\sigma(s)}{\beta(s)}\right\} u(t) \tag{2.2}$$

the lemma below relates the output error $\Delta y(t) = \hat{y}(t) - y(t)$ to the parameter error $\Delta K_u(t) = K_u(t) - K$. Note that $K_{u\sigma}(t)$ are elements of $K_u(t)$.

Lemma 2.1
Define $\hat{V}(t)$ as the vector whose elements are

$$\frac{p_\sigma(s)}{\beta(s)}u(t) - \frac{q_\sigma(s)}{\beta(s)}\hat{y}(t) \quad \forall\sigma \subset S$$

then

$$\Delta y(t) = \frac{\beta(s)}{Q(s,k)}\{\hat{V}^{\mathrm{T}}(t)\Delta K_u(t)\} \tag{2.3}$$

Proof
Equation (2.3) follows directly from dividing (2.1) by $\beta(s)$, subtracting (2.2) and rearranging terms. □

In the sequel the following two assumptions are made, the first of which will remain in force throughout the paper, while the second will hold for this section only.

Assumption 2.1
There exist known positive m_i and M_i such that

$$m_i \leqslant k_i \leqslant M_i \quad \forall i \in S \tag{2.4}$$

Assumption 2.2
The transfer function $\beta(s)/Q(s, k)$ is strictly positive real.

Remark 2.1
In Assumption 2.1 m_i can be made positive by introducing a suitable translation in the parameters. The problem of selecting a $\beta(s)$ to ensure the satisfaction of Assumption 2.2 for all k is an open question. However, given the knowledge of the parameter magnitude bounds and of the power p_σ and q_σ, its selection should be considerably more simple in this case than in the algorithms designed for the more conventional parametrization.

The proposed two-step algorithm is

$$\dot{K}_u(t) = -\hat{V}(t)\Delta y(t); \quad \prod_{i\in\sigma} m_i \leqslant K_{u\sigma}(s) \leqslant \prod_{i\in\sigma} M_i \quad \forall\sigma \subset S \tag{2.5}$$

$$\dot{\hat{k}}(t) = -\left[\frac{\partial \hat{K}(t)}{\partial \hat{k}(t)}\right]^{\mathrm{T}} \Lambda(\hat{K}(t) - K_u(t)) - \Gamma\Psi(\hat{k}(t))$$

$$m_i \leqslant \hat{k}_i(0) \leqslant M_i, \quad \forall i \in S \tag{2.6}$$

where $\hat{K}(t)$ is a $2^N - 1$ vector with σth element $\prod_{i\in\sigma}\hat{k}_i(t)$. Thus $\hat{K}(t)$ is a constrained estimate of K. The diagonal matrix $\Lambda = \mathrm{diag}\,[\lambda_1, \ldots, \lambda_N, \ldots, \lambda_\sigma, \ldots, \lambda_S]$ has $\lambda_\sigma \geqslant 0\ \forall\sigma \subset S$, and $\lambda_i > 0\ \forall i \in S$. The term $\Psi(\hat{k}(t))$, in the fashion of Kreisselmeier (1985), is introduced to prevent the $\hat{k}_i(t)$ from straying far outside the intervals $[m_i, M_i]$, and in particular from becoming negative. The ith element of $\Psi(\hat{k}(t))$ is given by, for $i \in S$,

$$\Psi_i(\hat{k}(t)) = \begin{Bmatrix} \hat{k}_i - M_i & \text{when } \hat{k}_i > M_i \\ 0 & \text{when } m_i \leqslant k_i \leqslant M_i \\ \hat{k}_i - m_i & \text{when } \hat{k}_i < m_i \end{Bmatrix} \tag{2.7}$$

For $i = 1, \ldots, N$, let $\tilde{M}_i$ be such that $M_i \leqslant \tilde{M}_i$. The choice of Γ, which is positive-definite diagonal matrix, will be described following Theorem 2.1, when we shall prove that

$$0 \leqslant \hat{k}_i(t) \leqslant \tilde{M}_i \quad \forall t \in R_+, \mathrm{i} \in S \tag{2.8}$$

Finally, observe that (2.5) represents a standard gradient-descent output-error algorithm (Landau 1979), while the first term in (2.6) is obtained as $\frac{1}{2}\partial\{(\hat{K}(t) - K_u(t))^T\Lambda(\hat{K}(t) - K_u(t))\}/\partial\hat{k}(t)$. Thus (2.6) attempts to project the unconstrained estimate K_u onto the appropriate constraint surface.

We now prove the uniform asymptotic convergence of $\hat{k}$ to k in two steps. Theorem 2.1 states that even without persistence of excitation $\|\Delta K_u\| = \|K_u - K\|$ is bounded and the output error is in $\mathscr{L}^2$. We then show via Theorem 2.2 that, under persistance of excitation, $\hat{k}$ converges to k, uniformly asymptotically.

For the proof, we shall require that the input $u(\cdot)$ lie in the set $\Omega_\Delta[0, \infty)$, defined as follows: there exists a countable, possibly empty, ordered set $C_\Delta = \{t_1, t_2, \ldots\}$ with $t_{i+1} - t_i \geqslant \Delta$ such that u, $\dot{u}$ are continuous and bounded on $R_+ - C_\Delta$, and u, $\dot{u}$ possess finite limits from the right and left at each t_i.

Theorem 2.1

Consider the unknown system (2.1), adjustable system (2.2) and the adjustment law (2.5). Suppose (2.1) is asymptotically stable, $u(t) \in \Omega_\Delta[0, \infty)$ and $Q(s, k)/\beta(s)$ is SPR. Then

$$\text{(i) } K_u(t) \text{ is bounded } \quad \forall t \geqslant 0 \tag{2.9}$$

$$\text{(ii) } \int_0^\infty \Delta y^2(t)\,dt < \infty \tag{2.10}$$

$$\text{(iii) } \int_0^\infty \|\hat{V}(t) - V(t)\|_2^2\,dt < \infty \tag{2.10}$$

where

$$V(t) \triangleq \left[\frac{p_1(s)u(t) - q_1(s)y(t)}{\beta(s)} \quad \frac{p_2(s)u(t) - q_2(s)y(t)}{\beta(s)} \quad \ldots\right]^T$$

Proof

Note first of all that (2.3) needs adjustment:

$$\Delta y(t) = \frac{\beta(s)}{Q(s, k)}\{\hat{V}^T(t)\Delta K_u(t) + \varepsilon_1(t)\}$$

where $\varepsilon_1(t)$ arises owing to initial condition effects and decays exponentially to zero owing to the stability of $Q(s, k)$.

However, its exponentially decaying nature implies that $\varepsilon_1(t)$ can be ignored. Since $\beta(s)/Q(s, k)$ is SPR there exist $x(t)$, A, b, c, d, L and r such that (Anderson 1977)

$$\left.\begin{aligned} \dot{x}(t) &= Ax(t) + b(\hat{V}^T(t)\Delta K_u(t)) \\ \Delta y(t) &= C^T x + d(\hat{V}^T(t)\Delta K_u(t)) \end{aligned}\right\} \tag{2.12}$$

where

$$A + A^{\mathrm{T}} = -LL^{\mathrm{T}} - 2rI$$

$$b = c - (2d)^{1/2}L$$

and $\{A, b, c, d\}$ is time-invariant.

Thus with $z(t) = [x^{\mathrm{T}}(t), \Delta K_{\mathrm{u}}^{\mathrm{T}}(t)]^{\mathrm{T}}$ we have that

$$\dot{z}(t) = \begin{bmatrix} A & b\hat{V}^{\mathrm{T}}(t) \\ -\hat{V}(t)c^{\mathrm{T}} & -d\hat{V}(t)\hat{V}^{\mathrm{T}}(t) \end{bmatrix} z(t) \tag{2.13}$$

Selecting the Lyapunov function

$$L_1(t) = z^{\mathrm{T}}(t)z(t)$$

we observe that, with some manipulation,

$$\begin{aligned} \dot{L}_1(t) &= -2rx^{\mathrm{T}}(t)x(t) - (x^{\mathrm{T}}(t)L + (2d)^{1/2}\Delta K_{\mathrm{u}}^{\mathrm{T}}(t)\hat{V}(t))^2 \\ &\leqslant 0 \end{aligned} \tag{2.14}$$

Thus $\|z(t)\|$ is decrescent, and in particular

$$\Delta K_{\mathrm{u}}^{\mathrm{T}}(t)\Delta K_{\mathrm{u}}(t) \leqslant \Delta K_{\mathrm{u}}^{\mathrm{T}}(0)\Delta K_{\mathrm{u}}(0) + x^{\mathrm{T}}(0)x(0) \tag{2.15}$$

whence (i) follows. Then $\exists \bar{M}_1$ and $\bar{M}_2$ such that

$$\left.\begin{aligned} &\int_0^\infty x^{\mathrm{T}}(t)x(t)\,dt < \bar{M}_1 \\ &\text{and} \\ &\int_0^\infty (x^{\mathrm{T}}(t)L + (2d)^{1/2}\Delta K_{\mathrm{u}}^{\mathrm{T}}(t)\hat{V}(t))^2\,dt < \bar{M}_2 \end{aligned}\right\} \tag{2.16}$$

Thus

$$\int_0^\infty 2d(\Delta K_{\mathrm{u}}^{\mathrm{T}}(t)\hat{V}(t))^2\,dt < \bar{M}_2 + \int_0^\infty \|x(t)\|_2^2\,dt\,\|L\|_2^2 < \bar{M}_2 + \bar{M}_1\,\|L\|_2^2 \tag{2.17}$$

From (2.12)

$$\int_0^\infty \Delta y^2(t)\,dt \leqslant \|c\|_2^2 \int_0^\infty \|x\|_2^2\,dt + d^2 \int_0^\infty |\hat{V}^{\mathrm{T}}K_{\mathrm{u}}|^2\,dt \leqslant \bar{M}_4 < \infty \tag{2.18}$$

Thus (2.10) is proved. Equation (2.11) is proved by noting that

$$\hat{V}(t) - V(t) = G(s)\Delta y(t)$$

where $G(s)$ is an asymptotically stable proper transfer function. □

Remark 2.2

From (2.15) we see that

$$\{K_u(t) - K\}^T\{K_u(t) - K\} \leqslant \Delta K_u^T(0)\Delta K_u(0) + x^T(0)x(0)$$

Suppose a bound on the magnitude of the initial state vector in any minimal realization of (2.1) is known. Then an *a priori* bound $\bar{M}_5$ on the magnitude of $x(0)$ will also be available. Now if the initial conditions in (2.5) are satisfied then

$$\begin{aligned}\|K_u(t)\|^2 &\leqslant \|K\|^2 + \sum_{\sigma \subset S}\left\{\prod_{i\in\sigma} M_i - \prod_{i\in\sigma} m_i\right\}^2 + \tilde{M}_5 \\ &\leqslant \sum_{\sigma \subset S}\left[\left(\prod_{i\in\sigma} M_i\right)^2 + \left\{\prod_{i\in\sigma} M_i - \prod_{i\in\sigma} m_i\right\}^2\right] + \tilde{M}_5 \\ &= M^2 \end{aligned} \tag{2.19}$$

for some M. The diagonal matrix $\Gamma = \operatorname{diag}[\gamma_1, \ldots, \gamma_N]$ in (2.6) is defined in terms of M and earlier defined quantities by

$$\gamma_i = \max\left\{\frac{M}{m_j}\sum_{\substack{\sigma\subset S\\ j\in\sigma}} \lambda_\sigma\left(\prod_{\substack{i\in\sigma\\ i\neq j}} \tilde{M}_i\right), \frac{1}{M_j}\sum_{\substack{\sigma\subset S\\ j\in\sigma}}\left(\prod_{\substack{i\in\sigma\\ i\neq j}} \tilde{M}_i\right)\left[\left(\prod_{i\in\sigma} \tilde{M}_i\right) + M\right]\right\} \tag{2.20}$$

and ensures satisfaction of (2.8). The calculation, which involves showing that $\dot{\hat{k}} > 0$ if $\hat{k}_i = 0$ and $\dot{\hat{k}}_i < 0$ if $\hat{k}_i = \tilde{M}_i$, is an easy one.

Remark 2.3

Equation (2.10) does *not* ensure that $\lim_{t\to\infty} \Delta y(t) = 0$, as $\hat{y}(t)$ may not be bounded.

To prove our main result, Theorem 2.2, we need the following proposition, which can be proved by a simple extension of a result in Dasgupta *et al.* (1984).

Proposition 2.1

Suppose all the elements of the vectors $\hat{k}(t)$ and k are always positive. Then

$$\frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)}\Lambda(\hat{K}(t) - K) + \Gamma\Psi(\hat{k}(t)) \equiv 0 \quad \text{iff } \hat{k}(t) \equiv k.$$

□

Theorem 2.2

For (2.1), (2.2), (2.5) and (2.6) let $K_{u\sigma} \geqslant 0\ \forall\sigma \subset S$. Assume that (2.1) is asymptotically stable, Assumption 2.2 holds, $u(t) \in \Omega_\Delta[0, \infty)$ and there exist no θ such that $\|\theta\| = 1$ and

$$\sum_{\sigma\subset S}\theta_\sigma p_\sigma(s) \equiv \sum_{\sigma\subset S}\theta_\sigma q_\sigma(s) \equiv 0 \tag{2.21}$$

Then $[x^T(t)\ \ K_u^T(t) - K^T\ \ \hat{k}^T(t) - k^T]$ converges exponentially fast to zero if there exist $\alpha_1, \alpha_2, \delta > 0$ such that $\forall v \in R_+$

$$\alpha_1 I \leqslant \int_v^{v+\delta} W(t)W^T(t)\,dt \leqslant \alpha_2 I \tag{2.22}$$

Here

$$W^{\mathrm{T}}(t) \triangleq \left[u(t) \quad \frac{1}{s+\gamma} u(t) \quad \dots \quad \frac{1}{(s+\gamma)^m} u(t) \right]$$

for any $\gamma < 0$ and m the highest degree among the polynomials $p_0 q_\sigma - q_0 p_\sigma \; \forall \sigma \subset S$ and $p_\sigma q_{\bar\sigma} - q_\sigma p_{\bar\sigma} \; \forall \sigma, \bar\sigma \subset S$ and $\sigma \neq \bar\sigma$.

Remark 2.4

Just as the non-negativity of the entries of $\hat{k}$ can be assured by inclusion of an additional term that pulls $\hat{k}$ inwards from its boundaries (see $-\Gamma\Psi(\hat{k}(t))$ in (2.6)), in a similar manner we can adjust (2.5) to ensure that $K_{u\sigma}(t) \geqslant 0 \; \forall \sigma \subset S, \forall t$.

Remark 2.5

In case there exists $\|\theta\| = 1$ such that (2.21) holds, there can exist a $\bar{K} \neq K$ and not necessarily of constrained form (i.e. if $N = 2$, $\bar{K} = [\bar{k}_1 \;\; \bar{k}_2 \;\; \bar{k}_3]$ without necessarily $\bar{k}_3 = \bar{k}_1 \bar{k}_2$), such that

$$\frac{p_0(s) + \sum k_\sigma p_\sigma(s)}{q_0(s) + \sum k_\sigma q_\sigma(s)} = \frac{p_0(s) + \sum \bar{k}_\sigma p_\sigma(s)}{q_0(s) + \sum \bar{k}_\sigma q_\sigma(s)}$$

In other words, if we are not concerned about the constraints ($k_3 = k_1 k_2$ for example), we have a non-identifiable situation.

Proof

From a result in Dasgupta *et al.* (1984), (2.21), (2.22) and the asymptotic stability of (2.1) imply $\exists \; \beta_3, \beta_4, \delta_2 > 0$ such that $\forall \, v \in R_+$

$$\beta_3 I \leqslant \int_v^{v+\delta_2} V(t) \, V^{\mathrm{T}}(t) \, dt \leqslant \beta_4 I \tag{2.23}$$

where V has the elements

$$\frac{p_\sigma(s)}{\beta(s)} u(t) - \frac{q_\sigma(s)}{\beta(s)} y(t) \quad \forall \sigma \subset S$$

Then by (2.11) and Lemma A.1, $\exists \beta_5, \beta_6, \delta_3 > 0$ such that $\forall v \in R_+$

$$\beta_5 I \leqslant \int_v^{v+\delta_3} \hat{V}(t) \hat{V}^{\mathrm{T}}(t) \, dt \leqslant \beta_6 I \tag{2.24}$$

Because $u \in \Omega_\Delta[0, \infty)$ and because of the way in which $\hat{V}(\cdot)$ is derived from $u(\cdot)$, $\hat{V} \in \Omega_\Delta[0, \infty)$ also. By a result in Anderson (1977), (2.24) implies that (2.13) is exponentially stable, and there exists an $\bar{L}_1$ such that

$$c_1 \{ \|K_u(t) - K\|^2 + \|x\|^2 \} \leqslant \bar{L}_1 \leqslant c_2 \{ \|K_u(t) - K\|^2 + \|x\|^2 \}$$

and

$$\dot{\bar{L}}_1 \leqslant -c_3 \{ \|K_u(t) - K\|^2 + \|x(t)\|^2 \}$$

for some c_1, c_2 and $c_3 > 0$ (see Krasovskii 1963, p. 86). It is not difficult to see that $\hat{K}(t)$ is bounded, whence $\exists c_4$ such that

$$\left\| \frac{\partial \hat{K}^{\mathrm{T}}(t)}{\partial \hat{k}(t)} \Lambda \right\| < c_4$$

Consider the Lyapunov function

$$\bar{L} = \tfrac{1}{2}(\hat{K}(t) - K)^{\mathrm{T}}\Lambda(\hat{K}(t) - K) + \tfrac{1}{2}\Psi^{\mathrm{T}}(\hat{k}(t))\Gamma\Psi(\hat{k}(t)) + \left(1 + \frac{c_4^2}{2c_3}\right)\bar{L}_1$$

It is easy to that, with respect to $[\hat{K}(t) - K \;\; K_u(t) - K \;\; x(t)]$, $\bar{L}$ is positive-definite and $\dot{\bar{L}}$ is negative-definite. The result follows directly. □

Remark 2.6

In the ideal situation of no noise, no time-variation of parameters and exact modellability of the plant by the assumed model, Theorem 2.2 shows that the second step of the algorithm (projection of K_u onto the constraint surface) is unnecessary, since K_u converges to K. But in the presence of non-ideality, the theorem, which guarantees exponential stability in the ideal case, implies, together with concepts of total stability (Krasovskii 1963), that the estimates will be indicative of correct values, at least if the non-idealities are not great, and then the advantage of the projection on to the constraint surface becomes clear.

3. A recursive least-squares formulation

Consider the unknown system and the adjustable model defined by (2.1) and (2.2) respectively and the update scheme

$$\dot{\hat{R}}(t) = -\alpha\hat{R}(t) + \hat{V}(t)\hat{V}^{\mathrm{T}}(t) \quad \forall t \geqslant 0; \quad \hat{R}(0) = 0 \tag{3.1}$$

$$\dot{X}(t) = \alpha X(t) - X(t)\hat{V}(t)\hat{V}^{\mathrm{T}}(t)X(t), \quad X(t_0) = \hat{R}^{-1}(t_0) \quad \forall t \geqslant t_0 \tag{3.2}$$

$$\dot{K}_u(t) = -X(t)\hat{V}(t)\Delta y(t) \quad \forall t \geqslant t_0 \tag{3.3}$$

$$\dot{\hat{k}}(t) = \frac{\partial \hat{K}^{\mathrm{T}}(t)}{\partial \hat{k}(t)}\Lambda[\hat{K}(t) - K_u(t)] - \Gamma\Psi(\hat{k}(t)) \quad \forall t > t_0 \tag{3.4}$$

$$m_i \leqslant \hat{k}_i(t_0) \leqslant M_i, \; \forall i \in S$$

where t_0 is the first time instant at which $\hat{R}(t)$ becomes well conditioned, and Ψ and Γ are defined in (2.7) and (2.9). The choice of the bound M on $K_u(t)$ will be explained at a later stage, but assume for the time being that M is such that (2.8) is always satisfied. Here (3.3) is an unconstrained least-squares output-error algorithm similar to that in Landau (1979).

The following result, stated without proof, shows that the infinite memory associated with (3.1) and (3.2) ensures that $X(t)$ is the inverse of $\hat{R}(t)$ $\forall t \geqslant t_0$.

Lemma 3.1

Suppose $\hat{R}(t)$ and $X(t)$ are defined by (3.1) and (3.2) and that $t_0 > 0$ such that $\hat{R}^{-1}(t_0)$ exists. Then

$$\hat{R}(t)X(t) = X(t)\hat{R}(t) = I \quad \forall t \geqslant t_0 \tag{3.5}$$

The convergence analysis proceeds on similar lines to that in §2. The SPR condition, however, needs adjustment. Assumption 2.2 must now be replaced by the following.

Assumption 3.1
The transfer function $\beta(s)/Q(s, k) - \frac{1}{2}$ is strictly positive real.

Theorem 3.1 shows that the output error is in $\mathscr{L}^2$ as long as $\hat{R}(t_0)$ is invertible for some $t_0 > 0$.

Theorem 3.1
For the unknown system (2.1), assumed asymptotically stable, the adjustable system (2.2) and adaptive law (3.3), the following are true as long as $\beta(s)/Q(s, k) - \frac{1}{2}$ is SPR, $u(t) \subset \Omega_\Delta[0, \infty)$ and $t_0 > 0$ such that $\hat{R}(t_0) \geqslant \alpha_1 I$:

(i) $\Delta K_u^T(t)\hat{R}(t)\Delta K_u(t)$ is bounded (3.6)

(ii) $$\int_0^\infty \Delta y^2(t)\, dt < \infty \tag{3.7}$$

(iii) $$\int_0^\infty \| \hat{V}(t) - V(t)\|^2\, dt < \infty$$

(iv) $\hat{R}(t) \leqslant \alpha_2 I$ for some finite α_2 and all $t \geqslant 0$ (3.9)

Proof
By Lemma 3.1 and the SPR nature of $\beta(s)/Q(s, k) - \frac{1}{2}$ there exist $x(t)$, A, b, c, d, L, r such that

$$\left.\begin{aligned} \dot{x}(t) &= Ax(t) + b(\hat{V}^T(t)\Delta K_u(t)) \\ \Delta y(t) &= c^T x(t) + d(\hat{V}^T(t)\Delta K_u(t)) \end{aligned}\right\} \tag{3.10}$$

$$\left.\begin{aligned} A + A^T &= -LL^T - 2rI \\ b &= c - (2d - 1)^{1/2} L \end{aligned}\right\} \tag{3.11}$$

Choosing a state variable

$$z^T(t) = [x^T(t) \quad \Delta K_u^T(t)] \tag{3.12}$$

and a Lyapunov-like function

$$L_2(z, t) = z^T \begin{bmatrix} I & 0 \\ 0 & \hat{R}(t) \end{bmatrix} z \tag{3.13}$$

it can be shown that $\dot{L}_2 \leqslant 0$ so that

$$\Delta K_u^T(t)\hat{R}(t)\Delta K_u(t) \leqslant \Delta K_u^T(t_0)\hat{R}(t_0)\Delta K_u(t_0) + x^T(t_0)x(t_0) \tag{3.14}$$

which proves (3.6).

Moreover, (3.7) and (3.8) follow from arguments as in Theorem 2.1. Finally, let $\bar{M}_1$ be such that

$$\int_0^\infty \|\hat{V}(t) - V(t)\|^2 dt < \bar{M}_1$$

As (2.1) is asymptotically stable and $u(t) \in \Omega_\Delta[0, \infty)$,

$$\int_0^\tau \exp\ [-\alpha(t - \tau)] \| V(\tau)\|^2 d\tau \leqslant \bar{M}_2 < \infty \quad \forall t \in R_+$$

Thus

$$\left[\int_0^t \exp\left[-\alpha(t-\tau)\,\|\hat{V}(\tau)\|^2\,d\tau\right]^{1/2} \leqslant \bar{M}_1^{1/2} + \bar{M}_2^{1/2}$$

whence for any unit vector θ of appropriate dimension,

$$\theta^{\mathrm{T}}\hat{R}(t)\theta = \int \exp\left[-\alpha(t-\tau)\right](\theta^{\mathrm{T}}\hat{V}(\tau))^2\,d\tau$$

$$\leqslant (\bar{M}_1^{1/2} + \bar{M}_2^{1/2})^2$$

Thus (3.9) is proved. □

Remark 3.1

The condition under which the non-singularity of $R(t_0)$ at some time t_0 can be guaranteed is identical with the input conditions given in Theorem 3.1. This is because, up to $t = t_0$, the adjustable system is constant. Thus, by considerations similar to results in Dasgupta *et al.* (1983, 1984), it can be shown that $\exists\alpha_3, \delta > 0$ such that $\forall v \leqslant t_0 - \delta$

$$\alpha_3 I \leqslant \int_v^{v+\delta} \hat{V}(t)\hat{V}^{\mathrm{T}}(t)\,dt \tag{3.15}$$

Moreover, as the following theorem shows, this is enough to ensure the non-singularity of $R(t_0)$.

Theorem 3.2

Suppose $\exists\alpha_3, t_1 > 0$ such that $\forall v$ and some $T < t_1$

$$\int_{g(v-T)}^{v} \hat{V}(\tau)\hat{V}^{\mathrm{T}}(\tau)\,d\tau \geqslant \alpha_3 I \tag{3.16}$$

where

$$g(\tau) = \begin{cases} \tau & \text{for } \tau > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Then $\exists\alpha_2 > 0$ such that $\forall v > v_0$ and some v_0

$$R(v) = \int_0^v \exp\left[-\alpha(v-\tau)\right]\hat{V}(\tau)\hat{V}^{\mathrm{T}}(\tau)\,d\tau \geqslant \alpha_2 I \tag{3.17}$$

Proof

Suppose (3.17) is violated. Then for arbitrary $\varepsilon > 0$ there exists a unit θ such that for some v

$$\theta^{\mathrm{T}}R(v)\theta \leqslant \varepsilon^2$$

whence

$$\int_0^v \exp\left[-\alpha(v-t)\right]\{\theta^{\mathrm{T}}\hat{V}(t)\}^2\,dt < \varepsilon$$

Thus

$$\int_{g(v-t_1)}^{v} \exp\left[-\alpha(v-t)\right]\{\theta^{\mathrm{T}}\hat{V}(t)\}^2\,dt < \varepsilon$$

whence, by the definition of $g(v - t_1)$,

$$\int_{g(v-t_1)}^{v} \{\theta^{\mathrm{T}}\hat{V}(t)\}^2\,dt < \varepsilon \exp(\alpha t_1)$$

and then (3.16) is violated. Thus (3.16) implies (3.17). □

Theorem 3.3 below shows that $\Delta K_u \to 0$ exponentially fast whenever $u(t)$ is persistently exciting in the sense that (2.22) holds. This in turn implies that $\hat{k} - k$ converges to zero with uniform asymptotic stability as long as the $K_{u\sigma}$ are all positive. In proving this theorem we need to appeal to the notion of uniform complete observability (u.c.o.) defined in Anderson (1977)

Theorem 3.3

For (2.1), (2.2) and (3.2)–(3.4) suppose there exists no θ such that (2.2) holds, (2.1) is asymptotically stable and $\beta(s)/Q(s, k) - \frac{1}{2}$ is SPR. Then $\Delta K_u(t)$ converges exponentially to zero if $\exists\alpha_1, \alpha_2, \delta > 0$ such that

$$\alpha_1 I \leqslant \int_{v}^{v+\delta} W(t)W^{\mathrm{T}}(t)\,dt \leqslant \alpha_2 I \quad \forall v \in R_+ \tag{3.18}$$

Here

$$W(t) \triangleq \left[u(t) \quad \frac{1}{s+\gamma}u(t) \quad \cdots \frac{1}{(s+\gamma)}u(t)\right]^{\mathrm{T}}$$

for any positive γ and m = highest degree among the polynomials $p_0 q_\sigma - q_0 p_\sigma$ $\forall\sigma \subset S$ and $p_\sigma q_{\bar{\sigma}} - q_\sigma p_{\bar{\sigma}}$, $\forall\sigma, \bar{\sigma} \subset S$, $\sigma \neq \bar{\sigma}$, $u(t) \in \Omega_\Delta[0, \infty)$. Moreover, if $K_{u\sigma} > 0$ $\forall\sigma \subset S$ then $[\hat{k}^{\mathrm{T}}(t) - k^{\mathrm{T}}\, x^{\mathrm{T}}(t)\, K_u^{\mathrm{T}}(t) - K^{\mathrm{T}}]$ converges uniformly asymptotically to zero.

Proof

As in Theorem 2.2, (3.18) and (3.8) and the non-satisfaction of (2.21) and Lemma A.1 imply the existence of $\bar{\delta}, \alpha_3, \alpha_4 > 0$ such that

$$\int_{v}^{v+\bar{\delta}} \|\hat{V}(t)\|^2\,dt < \alpha_4 \tag{3.19}$$

and

$$\int_{v}^{v+\delta} \hat{V}(t)\hat{V}^{\mathrm{T}}(t)\,dt > \alpha_3 I \quad \forall v \in R_+ \tag{3.20}$$

Thus by Theorem 3.2, $\exists t, \alpha_5 > 0$ such that

$$\hat{R}(t) \geqslant \alpha_5 I \quad \forall t > t_1 \tag{3.21}$$

Thus L_2 given by (3.13) with $z(t)$ defined by (3.12) is a Lyapunov function. Suppose

$$F(t) = \begin{bmatrix} A & b\hat{V}^{\mathrm{T}}(t) \\ -X(t)\hat{V}(t)c^{\mathrm{T}} & -dX(t)\hat{V}(t)\hat{V}^{\mathrm{T}}(t) \end{bmatrix}$$

$$H(t) = \begin{bmatrix} (2r)^{1/2}I & L \\ 0 & (2d-1)^{1/2}\hat{V}(t) \end{bmatrix}$$

Then let $\bar{H}(t)$ be defined by

$$\begin{bmatrix} 0 & 0 \\ 0 & \alpha\hat{R}(t) - \hat{V}(t)\hat{V}^{\mathrm{T}}(t) \end{bmatrix} = F^{\mathrm{T}}(t)\begin{bmatrix} I & 0 \\ 0 & \hat{R}(t) \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & \hat{R}(t) \end{bmatrix} F(t) + \bar{H}(t)\bar{H}^{\mathrm{T}}(t)$$

Using (3.13), this is equivalent to

$$\begin{aligned} \bar{H}(t)\bar{H}^{\mathrm{T}}(t) &= \begin{bmatrix} 0 & 0 \\ 0 & \alpha\hat{R}(t) \end{bmatrix} + H(t)H^{\mathrm{T}}(t) \\ &= \bar{R}(t) + H(t)H^{\mathrm{T}}(t) \end{aligned}$$

Using arguments similar to those in Anderson (1977), it can be shown that (3.18) and (3.19) imply that $[F, H]$ is uniformly completely observable. Using the uniform positive-definiteness of $\hat{R}(t)$ and (3.22), it can be shown that this implies the uniform complete observability of $[F, \bar{H}]$ and, consequently, again following Anderson (1977), $z(\cdot)$ is exponentially stable. Then by arguments similar to Theorem 2.2 we have that $[\hat{k}^{\mathrm{T}} - k^{\mathrm{T}}\ K_{\mathrm{u}}^{\mathrm{T}} - K^{\mathrm{T}}\ x^{\mathrm{T}}]$ is also exponentially stable. □

Remark 3.2

Suppose $\exists\alpha_1, \alpha_2 > 0$ such that $\alpha_1 I \leqslant R(t)\ \forall t$ and let $\lambda_{\max}R(t_0) = \alpha_2$. Then (3.14) implies

$$\Delta K^{\mathrm{T}}(t)\Delta K(t) \geqslant \frac{\alpha_2}{\alpha_1}\Delta K^{\mathrm{T}}(t_0)\Delta K(t_0) + \frac{\bar{M}}{\alpha_1}$$

where $\bar{M}$ is the bound on $x^{\mathrm{T}}(0)x(0)$. As $\alpha_1 > 0$ one can choose a sufficiently large $u(t)$ to ensure that $\alpha_1 > 1$, i.e.

$$\Delta K^{\mathrm{T}}(t)\Delta K(t) \leqslant \alpha_2\Delta K^{\mathrm{T}}(t_0)\Delta K(t_0) + \bar{M}$$

Also

$$\alpha_2 \leqslant \frac{1}{\alpha}\max\{\|\hat{V}(t)\|^2 : t \in [0, t_0]\} = \bar{M}_3$$

Thus for large enough $u(t)$

$$\Delta K^{\mathrm{T}}(t)\Delta K(t) \leqslant \bar{M} + \bar{M}_3$$

It is reasonable to expect that conservative estimates of both max $\|V(t)\|^2$ and $\bar{M}$ would be known. Thus this information can be used in employing the technique explained earlier, which ensures that $K_{\mathrm{u}\sigma}(t) > 0\quad \forall t \geqslant 0$.

4. Conclusions

Two new output-error identification schemes have been presented in this paper, Their application to physical systems where there is uncertainty in physical element values only is expected to produce good results. These schemes exploit all the *a priori* information, both in the parametrization of the system and also in the identification algorithms.

Both schemes use two-step procedures. The first step produces a parameter estimate ignoring the inherent relationship between the parameters. In one scheme this is achieved using gradient descent and in the other, recursive least-squares. The second step, common to both schemes, sharpens up the estimate produced by the first using the multilinear constraints.

The analysis of this paper shows that both schemes are globally uniformly asymptotically stable whenever certain strict positive realness conditions and persistence of excitation conditions are satisfied.

Appendix. A key lemma

Lemma A.1 is an extension of a result in Boyd and Sastry (1983) to unbounded signals. All norms considered are Euclidean.

Lemma A.1

Suppose that

$$\int_0^{\infty} \|V - \hat{V}\|^2 \, dt < M_1 \tag{A 1}$$

$$\int_v^{v+\delta} \|V\|^2 \, dt < \alpha_1 \tag{A 2}$$

for some M_1, α_1 and all v. Then

$$\int_v^{v+\delta} \|\hat{V}\|^2 \, dt < \alpha_2 \tag{A 3}$$

for some α_2 and all v. Suppose that (A 2) does not necessarily hold but that

$$\alpha_3 I \leqslant \int_v^{v+\delta} V V^{\mathrm{T}} \, dt \tag{A 4}$$

for some $\alpha_3 > 0$ and all v. Then

$$\alpha_4 I \leqslant \int_v^{v+\bar{\delta}} \hat{V} \hat{V}^{\mathrm{T}} \, dt \tag{A 5}$$

for some α_4, $\bar{\delta} > 0$ and all v.

Proof

Inequality (A 3) follows from (A 1) and (A 2) by Minkowskii's inequality. Next, let $\bar{\delta} = p\delta$, where p is an integer to be specified. Then for arbitrary θ of unit norm

$$\left\{\int_{\nu}^{\nu+\delta}(\theta^{\mathrm{T}}\hat{V})^2dt\right\}^{1/2} = \left\{\int_{\nu}^{\nu+\delta}(\theta^{\mathrm{T}}\hat{V} - \theta^{\mathrm{T}}V + \theta^{\mathrm{T}}V)^2\,dt\right\}^{1/2}$$

$$\geqslant \left\{\int_{\nu}^{\nu+\delta}(\theta^{\mathrm{T}}V)^2dt\right\}^{1/2} - \left\{\int_{\nu}^{\nu+\delta}(\theta^{\mathrm{T}}\hat{V} - \theta^{\mathrm{T}}V + \theta^{\mathrm{T}}V)^2dt\right\}^{1/2}$$

$$\geqslant (\alpha_3 p)^{1/2} - \left\{\int_{\nu}^{\nu+\delta}\|\hat{V} - V\|^2dt\right\}^{1/2}$$

$$\geqslant (\alpha_3 p)^{1/2} - M_1^{1/2}.$$

Thus if $p > M_1/\alpha_3$ then (A 5) holds. □

References

ANDERSON, B. D. O. 1977, *I.E.E.E. Trans. autom. Control*, **22**, 83.

BOYD, S., and SASTRY, S., 1983, *Syst. Control Lett*, **3**, 311.

DASGUPTA, S., ANDERSON, B. D. O., and KAYE, R. J., 1984 Identification of economically parametrized systems, *Proc. IFAC World Congress*, Budapest, Vol. X, edited by J. Gertler and L. Keviczky (Oxford: Pergamon), pp. 96–100.

DASGUPTA, S., ANDERSON, B. D. O., and TSOI, A. C., 1983, Input conditions for continuous-time adaptive system problems, *Proc. 22nd Conf. on Decision and Control*, San Antonio, Texas, U.S.A., December 1983 (New York: I.E.E.E.), pp. 211–216.

KRASOVSKII, N. N., 1963, *Stability of Motion* (Stanford University Press).

KREISSELMEIER, G., 1985, *Automatica*, **21**, 425.

LANDAU, Y. D., 1979, *Adaptive Control: The Model Reference Approach* (New York: Marcel Dekker).