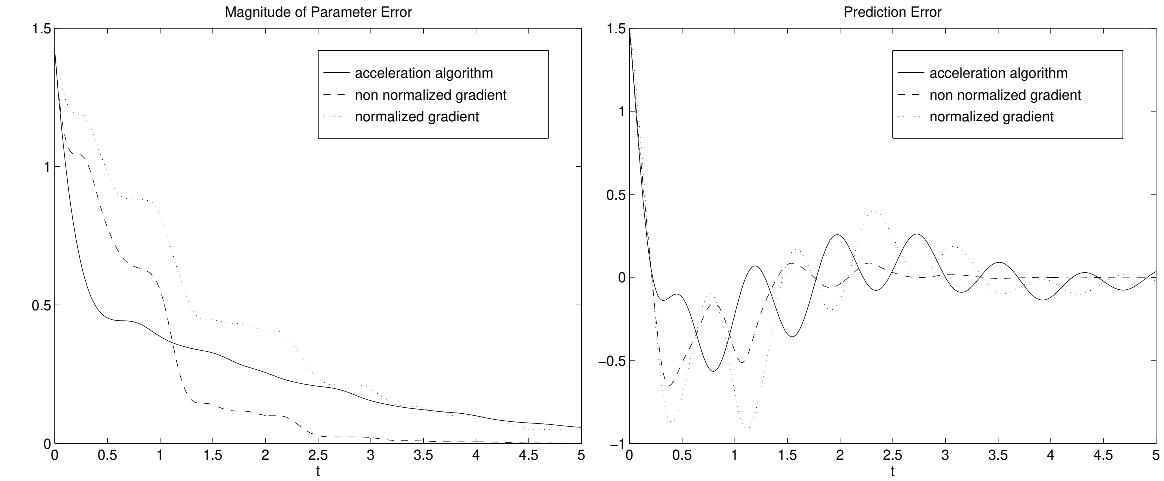

Figure 2: Response to sinusoidal data

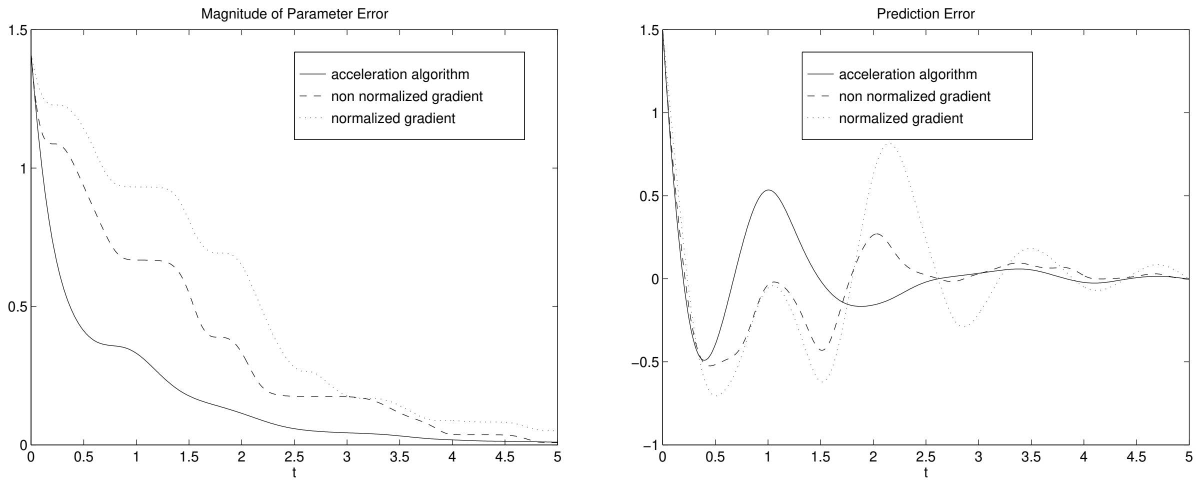

Figure 3: Response to exponentially increasing sinusoid

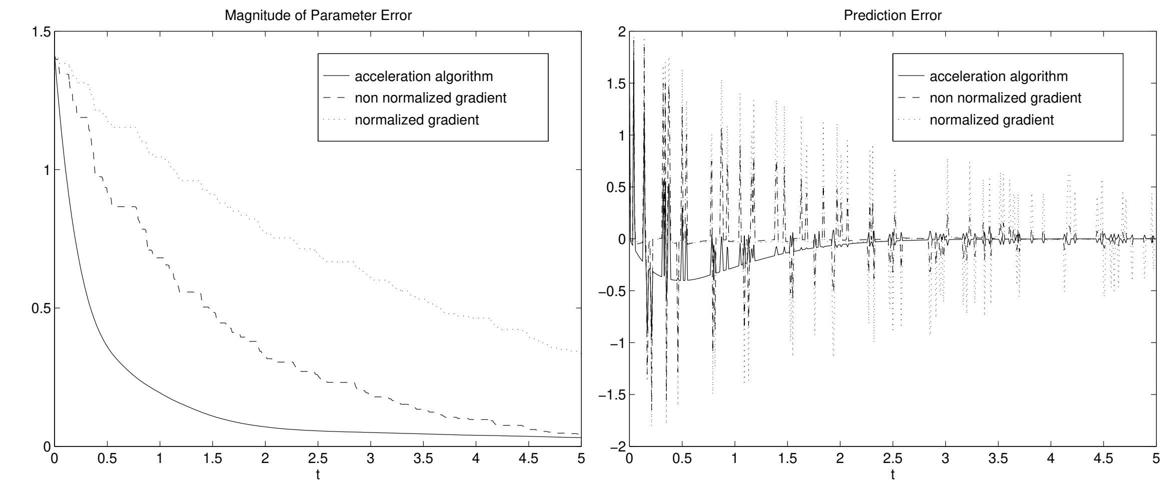

Figure 4: Response to pseudorandom signal

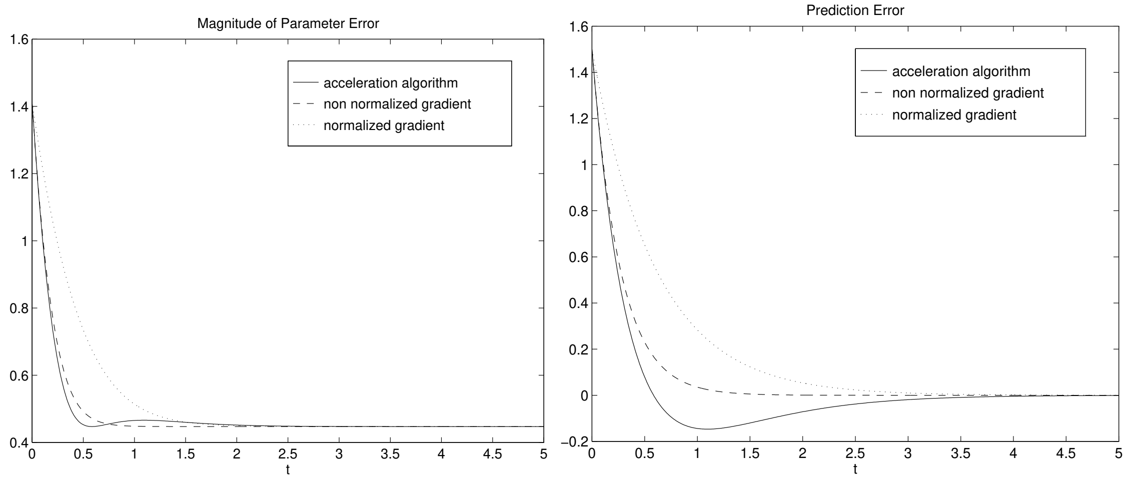

Figure 1: Response to step data

lence and implicit tuning. IEEE *Trans. Automatic Control*, 37(1):15–29, January 1992.

[4] Romeo Ortega. On Morse's new adaptive controller: Parameter convergence and transient performance. IEEE *Trans. Automatic Control*, 38(8):1191–1202, August 1993.

$$\frac{\partial}{\partial \tau}\begin{bmatrix}\phi_{11} & \phi_{12}\\ \phi_{21} & \phi_{22}\end{bmatrix}\bar{M}_1\begin{bmatrix}\phi_{11}^\top & \phi_{21}^\top\\ \phi_{12}^\top & \phi_{22}^\top\end{bmatrix} = \begin{bmatrix}\phi_{11} & \phi_{12}\\ \phi_{21} & \phi_{22}\end{bmatrix}\Bigg(\bar{M}_2$$
$$+\begin{bmatrix}0 & 0\\ 0 & -M_1xx^\top M_1 + 2(1+\mu_2)(M_1M_2M_1M_3M_1xx^\top M_1 + M_1xx^\top M_1M_3M_1M_2M_1)\end{bmatrix}\Bigg)\begin{bmatrix}\phi_{11}^\top & \phi_{21}^\top\\ \phi_{12}^\top & \phi_{22}^\top\end{bmatrix},$$
$$\text{where}\quad \bar{M}_1 = \begin{bmatrix}M_1M_3M_1 + \mu_1M_1 & -M_1/2\\ -M_1/2 & (1+\mu_2)M_1M_2M_1\end{bmatrix},$$
$$\bar{M}_2 = \begin{bmatrix}M_1 & \mu_1M_1xx^\top M_1 - (2+\mu_2)M_1M_2M_1\\ \mu_1M_1xx^\top M_1 - (2+\mu_2)M_1M_2M_1 & 4(1+\mu_2)M_1M_2M_1M_2M_1\end{bmatrix} \tag{7}$$

$$\int_0^t d\tau\,(4(1+\mu_2)m_1^2m_2m_3 - 1)\phi_{12}(t,\tau)M_1xx^\top(\tau)M_1\phi_{12}^\top(t,\tau)$$
$$= \begin{bmatrix}\phi_{11}(t,\tau) & \phi_{12}(t,\tau)\end{bmatrix}\bar{M}_1\begin{bmatrix}\phi_{11}^\top(t,\tau)\\ \phi_{12}^\top(t,\tau)\end{bmatrix}\Bigg|_{\tau=0}^{\tau=t} - \int_0^t d\tau\begin{bmatrix}\phi_{11} & \phi_{12}\end{bmatrix}\bar{M}_2\begin{bmatrix}\phi_{11}^\top\\ \phi_{12}^\top\end{bmatrix} \tag{8}$$

---

right-hand side of (10) tends to $I/4m_1m_3r$, which indicates that the steady-state covariance of the parameter error decreases when the signal x increases in magnitude, and that it can be made smaller via appropriate choices of the gains m_1 and m_3. The situation for the accelerating tuner is hence much more favorable than for the conventional one.

4 Simulations

The simulations in this section compare the behavior of the accelerating tuner (4) with those of the gradient tuner (3) and of a normalized gradient one. All simulations were done in open-loop, with the regressor a two-dimensional signal, and without measurement noise. Figure 1 shows the values of $\sqrt{q^\top q}$ and $x^\top q$ respectively when x is a two-dimensional step signal. In Figure 2 the regressor is a sinusoid, in Figure 3 an exponentially increasing sinusoid, and in Figure 4 a pseudo-random signal generated using MATLAB. No effort was made to optimize the choice of gain matrices (M_1, M_2, and M_3 were all chosen equal to the identity), and the effect of measurement noise was not considered. The performance of the accelerating tuner is comparable, and sometimes superior, to that of the other tuners.

5 Concluding Remarks

Other ideas related to the present one are replacing the integrator in (3) with a positive-real transfer function [1, page 89], and using high-order tuning ([2, 4]). High-order tuning generates as outputs $\hat{p}$ as well as its derivatives up to a given order (in this sense we might consider the present algorithm a second-order tuner), but unlike the accelerating tuner requires derivatives of x up to that same order. We expect that accelerating tuners will find application in adaptive control of nonlinear systems and maybe in dealing with the topological incompatibility known as the "loss of stabilizability problem" in the adaptive control literature.

The stochastic analysis in §3 indicates that the performance and convergence properties of the accelerating tuner, together with its moderate computational complexity, may indeed make it a desirable tool for adaptive filtering applications. It seems that a better transient × steady-state performance compromise is achievable with the accelerating tuner than with the velocity tuner. To verify this conjecture, a study of convergence properties of the accelerating tuner and their relation with the persistence of excitation conditions is in order, as well as more extensive simulations in the presence of measurement noise.

Acknowledgements The authors are grateful to Oswaldo L V Costa and Max Gerken for useful suggestions related to this work.

References

[1] Yoan D Landau. *Adaptive Control: The Model Reference Approach*. Marcel Dekker, New York, 1979.

[2] A Stephen Morse. High-order parameter tuners for the adaptive control of nonlinear systems. In *US-Italy Joint Seminar on Systems, Models, and Feedback: Theory and Applications*, Capri, 1992.

[3] A Stephen Morse. Towards a unified theory of parameter adaptive control – Part 2: Certainty equiva-

This equation can be given the following interpretation: for small t, when $\phi(t,0)$ is close to the identity, the covariance of $q(t)$ remains close to $qq^\top(0)$, the outer product of the error in the initial guess of the parameters with itself. As $\phi(t,0) \to 0$, which will happen if $x(\cdot)$ is persistently exciting, $E_w\{qq^\top(t)\}$ tends to $\sigma_w^2 M/2$. This points to a compromise between higher convergence speeds and lower steady-state parameter error, which require respectively larger and smaller values of the gain M. Algorithms that try for the best of both worlds — parameter convergence in the mean-square sense — often utilize time-varying, decreasing gains; an example is the least-squares algorithm.

We shall now attempt a similar analysis for the acceleration tuner (5) applied to (1), which results in the differential equation

$$\begin{bmatrix} \dot q \\ \ddot q \end{bmatrix} = \begin{bmatrix} 0 & I \\ -M_1 xx^\top & -2M_1(M_2 + xx^\top M_1 M_3) \end{bmatrix} \begin{bmatrix} q \\ \dot q \end{bmatrix} + \begin{bmatrix} 0 \\ M_1 x w \end{bmatrix}.$$

Let

$$\begin{bmatrix} \dot\phi_{11} & \dot\phi_{12} \\ \dot\phi_{21} & \dot\phi_{22} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -M_1 xx^\top & -2M_1(M_2 + xx^\top M_1 M_3) \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix},$$

where $\phi_{11}(t,t) = \phi_{22}(t,t) = I$, $\phi_{12}(t,t) = \phi_{21}(t,t) = 0$, each ϕ_{ii} is a function of (t,τ) unless otherwise noted, and the dot signifies derivative with respect to the first argument. If $\dot q(0) = 0$,

$$q(t) = \phi_{11}(t,0)q(0) + \int_0^t d\tau\, \phi_{12}(t,\tau) M_1 x(\tau) w(\tau).$$

Following the same reasoning used for the velocity tuner, one concludes that $E_w\{q(t)\} = \phi_{11}(t,0)q(0)$ and that

$$E_w\{qq^\top(t)\} = \phi_{11}(t,0)qq^\top(0)\phi_{11}^\top(t,0) + \sigma_w^2 \int_0^t d\tau\, \phi_{12}(t,\tau) M_1 xx^\top(\tau) M_1 \phi_{12}^\top(t,\tau), \quad (6)$$

however the properties of the acceleration and velocity tuners are not yet directly comparable because the right-hand side of (6) does not lend itself to immediate integration. To obtain comparable results, we employ the ungainly but easily verifiable formula (7), valid for arbitrary scalars μ_1 and μ_2, and make the

Simplifying Assumption: For $i = 1, 2$, and 3, $M_i = m_i I$, where m_i are scalars and I is the $n \times n$ identity matrix.

Premultiplying (7) by $[I\ 0]$, postmultiplying by $[I\ 0]^\top$, integrating from 0 to t, and using the simplifying assumption gives formula (8). Taking $\mu_1 = \mu_2 = 0$ in (8), $\bar M_2$ results positive-semidefinite, therefore

$$\int_0^t d\tau\, \phi_{12}(t,\tau) M_1 xx^\top(\tau) M_1 \phi_{12}^\top(t,\tau) \le \frac{1}{4m_1^2 m_2 m_3 - 1}\left(m_1^2 m_3 I - \begin{bmatrix} \phi_{11}(t,0) & \phi_{12}(t,0) \end{bmatrix} \bar M_1 \begin{bmatrix} \phi_{11}^\top(t,0) \\ \phi_{12}^\top(t,0) \end{bmatrix} \right). \quad (9)$$

The combination of (6) and (9) shows that m_1 can be increased without affecting q's steady-state covariance. On the other hand, to decrease the covariance we need to increase m_2, which roughly speaking means increasing damping in (4). Since m_1 and m_2 can be increased without affecting the stability properties shown in §2, a better transient $\times$ steady-state performance compromise might be achievable with the acceleration tuner than with the velocity tuner, at least in the case when M_1, M_2, and M_3 are "scalars." Notice that $4m_1^2 m_2 m_3 > 1$ by construction.

Approximate Analysis: The derivation of inequality (9) does not involve any approximations, and therefore provides an upper bound on $E_w\{qq^\top(t)\}$, valid independently of $x(\cdot)$. A less conservative estimate of the integral in (6) can be obtained by replacing $xx^\top$ by its average value R in the definition of $\bar M_2$ in (8). This approximation seems reasonable because $\bar M_2$ appears inside an integral, but calls for more extensive simulation studies.

To obtain a useful inequality, we require $\bar M_2 \ge 0$; namely, using the Schur complement

$$4(1+\mu_2) M_1 M_2 M_1 M_2 M_1 \ge M_1(\mu_1 xx^\top - (2+\mu_2) M_2) M_1 (\mu_1 xx^\top - (2+\mu_2) M_2) M_1,$$

or, using the simplifying assumption and substituting $xx^\top$ by its approximation R

$$4(1+\mu_2) I \ge \left((2+\mu_2) I - \frac{\mu_1}{m_2} R \right)^2.$$

Suppose further that $R = rI$. Looking for the least conservative estimate, we pick $\mu_1 = \frac{m_2}{r}(2 + \mu_2 - 2\sqrt{1+\mu_2})$, the least value of μ_1 that keeps $\bar M_2 \ge 0$. Thus

$$\int_0^t d\tau\, \phi_{12}(t,\tau) M_1 xx^\top(\tau) M_1 \phi_{12}^\top(t,\tau) \le \frac{m_1^2 m_3 r + m_1 m_2 (\mu_2 + 2 - 2\sqrt{1+\mu_2})}{4m_1^2 m_2 m_3 r (1+\mu_2) - r} I - \Xi(t) \quad (10)$$

with $\Xi(t) = \frac{[\,\phi_{11}(t,0)\ \ \phi_{12}(t,0)\,] \bar M_1 \begin{bmatrix} \phi_{11}^\top(t,0) \\ \phi_{12}^\top(t,0) \end{bmatrix}}{4m_1^2 m_2 m_3 r (1+\mu_2) - r}$.

Taking $\mu_2 = 0$ we repeat the previous, exact result. For large positive values of μ_2 the first term of the because the unknown q appears only in scalar product with x. Choose another function of Lyapunovian inspiration:

$$\begin{aligned} V &= q^\top q + q^\top \dot{q} + \dot{q}^\top \dot{q} \\ &= \begin{bmatrix} q^\top & \dot{q}^\top \end{bmatrix} \begin{bmatrix} I & I/2 \\ I/2 & I \end{bmatrix} \begin{bmatrix} q \\ \dot{q} \end{bmatrix} \geq 0. \end{aligned}$$

Taking derivatives and using the accelerating tuner (4) gives

$$\begin{aligned} \dot{V} &= 2q^\top \dot{q} + \dot{q}^\top \dot{q} - (q^\top + 2\dot{q}^\top)(2\dot{q} + xx^\top q + 2xx^\top \dot{q}) \\ &= -3\dot{q}^\top \dot{q} - (q + 2\dot{q})^\top xx^\top (q + 2\dot{q}) \leq 0. \end{aligned}$$

Integrating $\dot{V}$ we obtain

$$V(t) = V(0) - \int_0^t \dot{q}^\top \dot{q} - \int_0^t \Big(x^\top (q + 2\dot{q}) \Big)^2,$$

which leads immediately to the desired properties:

$$q, \dot{q} \in \mathcal{L}^\infty; \quad \dot{q} \in \mathcal{L}^2; \quad x^\top (q + 2\dot{q}) \in \mathcal{L}^2.$$

The slow variation property $\dot{q} \in \mathcal{L}^2$ follows without the need for normalization, and now we obtain $x^\top(q + 2\dot{q}) \in \mathcal{L}^2$ instead of $x^\top q \in \mathcal{L}^2$ as before. We might regard $x^\top(q + 2\dot{q})$ as a modified error, which can be used in the stability analysis of a detectable or "tunable" adaptive system via an output-injection argument; see [3]. A generalization of (4) is

$$\ddot{q} = -M_1 \Big(x(x^\top \hat{p} - y) + 2(M_2 + xx^\top M_1 M_3)\dot{q} \Big), \tag{5}$$

with M_1, M_2 and M_3 constant, symmetric, positive-definite $n \times n$ matrices such that $M_2^{-1} < 4M_1M_3M_1$ and $M_2M_1M_3 + M_3M_1M_2 > M_1^{-1}/2$. The properties of tuner (5), which can be obtained using the positive-definite function

$$\begin{bmatrix} q^\top & \dot{q}^\top \end{bmatrix} \begin{bmatrix} M_2 & M_1^{-1}/2 \\ M_1^{-1}/2 & M_3 \end{bmatrix} \begin{bmatrix} q \\ \dot{q} \end{bmatrix}$$

in the same manner as before, are

$$q, \dot{q} \in \mathcal{L}^\infty; \quad \dot{q} \in \mathcal{L}^2; \quad x^\top (q + 2M_1M_3\dot{q}) \in \mathcal{L}^2.$$

3 Covariance Analysis

We now consider the effects on the expected value and covariance of $q(t)$ of the presence of a measurement error. The assumptions are that $w(t)$ is a white noise with zero average and covariance $\sigma_w^2 \delta(t - \tau)$ and that $x(t)$ are given, deterministic data. For comparison purposes, first consider what happens when the conventional tuner (2) is applied to (1) in the presence of measurement error w:

$$\dot{q} = -Mxx^\top q + Mxw. \tag{3'}$$

The solution to the equation above can be written in terms of $(-Mxx^\top)$'s state transition matrix $\phi(t, \tau)$ as follows

$$q(t) = \phi(t,0)q(0) + \int_0^t d\tau \, \phi(t,\tau) M x(\tau) w(\tau),$$

hence $E_w\{q(t)\} = \phi(t,0)q(0)$ because $E\{w(\tau)\} = 0$ by assumption. Here the notation $E_w\{\cdot\}$, denoting the expectation with respect to the random variable w, is used to emphasize that the stochastic properties of x are not under consideration. The conclusion is that $q(t)$ will converge to zero in average as fast as $\phi(t,0)$ does. The well-known persistency of excitation conditions on $x(\cdot)$ are sufficient for the latter to happen.

To study the second moment of the parameter error, write

$$\begin{aligned} qq^\top(t) &= \phi(t,0) qq^\top(0) \phi^\top(t,0) \\ &\quad + \phi(t,0)q(0) \int_0^t d\tau \, w(\tau) x^\top(\tau) M \phi^\top(t,\tau) \\ &\quad + \left(\int_0^t d\tau \, \phi(t,\tau) M x(\tau) w(\tau) \right) q^\top(0) \phi^\top(t,0) \\ &\quad + \left(\int_0^t d\tau \, \phi(t,\tau) M x(\tau) w(\tau) \right) \cdot \\ &\qquad \left(\int_0^t d\tau \, w(\tau) x^\top(\tau) M \phi^\top(t,\tau) \right). \end{aligned}$$

The covariance of $q(t)$ can be written as the sum of four terms. The first is deterministic. The second term

$$\begin{aligned} &E_w\Big\{ \phi(t,0)q(0) \int_0^t d\tau \, w(\tau) x^\top(\tau) M \phi^\top(t,\tau) \Big\} \\ &\quad = \phi(t,0)q(0) \int_0^t d\tau \, E\{w(\tau)\} x^\top(\tau) M \phi^\top(t,\tau) = 0 \end{aligned}$$

because $w(t)$ has zero mean, and the third term is likewise zero. The fourth term

$$\begin{aligned} &E_w \Big\{ \int_0^t d\tau_1 \, \phi(t,\tau_1) M x(\tau_1) w(\tau_1) \\ &\qquad \int_0^t d\tau_2 \, w(\tau_2) x^\top(\tau_2) M \phi^\top(t,\tau_2) \Big\} \\ &= \int_0^t d\tau_1 \int_0^t d\tau_2 \, \phi(t,\tau_1) M x(\tau_1) \\ &\qquad E\{w(\tau_1) w(\tau_2)\} x^\top(\tau_2) M \phi^\top(t,\tau_2) \\ &= \sigma_w^2 \int_0^t d\tau_1 \, \phi(t,\tau_1) M x(\tau_1) x^\top(\tau_1) M \phi^\top(t,\tau_1), \end{aligned}$$

where Fubini's Theorem and the fact $E\{w(\tau_1)w(\tau_2)\} = \sigma_w^2 \delta(\tau_1 - \tau_2)$ were used. Adding the first and fourth terms and performing the integration results in

$$\begin{aligned} E_w\{qq^\top(t)\} &= \phi(t,0) qq^\top(0) \phi^\top(t,0) \\ &\quad + \frac{\sigma_w^2}{2} \Big(M - \phi(t,0) M \phi^\top(t,0) \Big). \end{aligned}$$

A Tuner that Accelerates Parameters

Felipe M Pait[1] and Paulo A Atkinson[2]
pait,atk@lac.usp.br
Universidade de São Paulo – PEE
Laboratório de Automação e Controle
Av. Prof. Gualberto trav 3–158
São Paulo SP 05508–900 Brasil

Abstract

We propose a tuner, suitable for adaptive control and (in its discrete-time version) adaptive filtering applications, that sets the second derivative of the parameter estimates rather than the first derivative as is done in the overwhelming majority of the literature. Comparative stability and performance analyses are presented.

Key Words: Adaptive Control; Parameter Estimation; Adaptive Filtering; Covariance Analysis.

1 Parameter Adjustment

In adaptive control and recursive parameter estimation one often needs to adjust recursively an estimate $\hat{p}$ of a vector p, which comprises n constant but unknown parameters, using measurements of a quantity

$$y = x^\top p + w. \tag{1}$$

Here $x : [0, \bar{t}) \to \mathbb{R}^n$ is a vector of known data, often called the regressor, and w is a measurement error signal. The goal of tuning is to keep both the estimation error $x^\top \hat{p} - y$ and the parameter error $\hat{p} - p$ as small as possible.

There are several popular methods for dealing with the problem above, for instance least-squares. Maybe the most straightforward involve minimizing the prediction error via gradient-type algorithms of the form:

$$\dot{\hat{p}} = -Mx(x^\top \hat{p} - y), \tag{2}$$

where M is a constant, symmetric, positive-definite gain matrix. Let us define $q = \hat{p} - p$ and analyze differential equations (1) and (2), which under the assumption that $w(t)$ is identically zero read:

$$\dot{q} = -Mxx^\top q. \tag{3}$$

The nonnegative function $V = \frac{1}{2} q^\top M^{-1} q$ has time derivative $\dot{V} = q^\top M^{-1} \dot{q} = -q^\top x x^\top q$, hence

$$\int_0^t \dot{V} = V(t) - V(0) = -\int_0^t (x^\top q)^2.$$

Inspection of the equation above reveals that V is limited in time, thus $q \in \mathcal{L}^\infty$, and also that the error $x^\top q \in \mathcal{L}^2$ (norms are taken on the interval $[0, \bar{t})$ where all signals are defined). These are the main properties an algorithm needs in order to be considered a suitable candidate for the role of a tuner in an adaptive control system. Often $\dot{q} \in \mathcal{L}^2$ or something similar is also a desirable property. To obtain the latter, normalized algorithms can be used; however, the relative merits of normalized versus unnormalized tuners are still somewhat controversial. Another alternative is to use a time-varying M, as is done in least-squares tuning.

In §2 we present a tuner that sets the second derivative of $\hat{p}$, and in §3 the effects of a white noise w on the performance of the two algorithms are compared. Then we show some simulations and make concluding remarks.

2 The Accelerating Algorithm

Classical tuners are such that the *velocity* of adaptation (the first derivative of the parameters) is set proportional to the regressor and to the prediction error $x^\top \hat{p} - y = x^\top q$. We propose to set the *acceleration* of the parameters:

$$\boxed{\ddot{q} = -xx^\top q - 2(I + xx^\top)\dot{q}.} \tag{4}$$

Notice that the the formula above is implementable (using $2n$ integrators) if measurement error is absent,

[1]Research supported by CNPq – Brazilian Research Council, grant 523949/95-2.
[2]Research supported by FAPESP – São Paulo Research Agency, scholarship 97/00582-5.