Proceedings of the 35th
Conference on Decision and Control
Kobe, Japan • December 1996

Optimal Adaptive Control of an LQG System

M.C. Campi† and P.R. Kumar‡
†Dept. of Electrical Engineering and Automation
University of Brescia
Via Branze 38
25123 Brescia, Italy
‡Dept. of Electrical and Computer Engineering, and the
Coordinated Science Laboratory
University of Illinois
1308 West Main Street
Urbana, IL 61801,USA

Abstract

We consider the problem of adaptively controlling a linear system so as to minimize a long-term average quadratic cost criterion. It is well known that certainty equivalent controllers based on standard parameter estimators run into an identifiability problem which leads to a strictly suboptimal performance. In this contribution, a cost-biased parameter estimator is introduced to overcome this difficulty. The corresponding adaptive scheme is proven to be stable and optimal when the unknown system parameter lies in an infinite, yet compact, parameter set.

1 Introduction

Consider a linear system

$$x_{t+1} = A^\circ x_t + B^\circ u_t + w_{t+1} \tag{1}$$

where $x_t \in R^n$ is the state, $u_t \in R^m$ the control variable and $\{w_t\}$ is a noise process assumed to be independent and identically distributed with a $N(0, I)$ distribution. The state x_t is observed, and based on this, the goal is to choose the input u_t in such a way as to minimize the long-term average quadratic cost criterion,

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{s=0}^{t-1} \left[x_s^T Q x_s + u_s^T T u_s\right]. \tag{2}$$

(Q and T are symmetric matrices with $Q \geq 0$ and $T > 0$).

The system matrices A° and B° are, however, unknown. Thus we have in our hands an adaptive control problem. All we know is that (A°, B°) belong to a certain compact set Θ as precisely stated in the following assumptions:

(A.i) There is a known compact set Θ such that $(A^\circ, B^\circ) \in$ Interior (Θ).

(A.ii) (A, B) is reachable and $(A, Q^{\frac{1}{2}})$ is observable, for all $(A, B) \in \Theta$.

It is well known that certainty equivalent adaptive control can suffer from an identifiability problem; see Borkar and Varaiya [3], Becker, Kumar and Wei [1], Lin, Kumar and Seidman [12], and other references. In particular, for the case where Θ is a finite set, it is shown in Kumar [5] that a certainty equivalent adaptive controller using least squares parameter estimates can converge with positive probability to a wrong parameter estimate, which then leads to a strictly nonoptimal value of the long-term average cost criterion. For the case of general controlled Markov chains, such a counterexample had earlier been exhibited in Borkar and Varaiya [3]. For the case of linear ARMAX systems with a minimum output variance cost criterion, it has been shown in Becker, Kumar and Wei [1] that the parameter estimates can converge with positive probability to false values; however, due to the special property of the minimum output variance cost criterion, the resulting certainty equivalent control law is still optimal.

Motivated by this general problem of identifiability in closed-loop, a new certainty equivalent adaptive controller was proposed in Kumar and Becker [9] for the class of controlled Markov chains. The novelty of this adaptive controller was the employment of a "cost biased maximum likelihood" parameter estimator, rather than the usual maximum likelihood parameter estimator. This cost biasing modifies the log-likelihood criterion by incorporating an additional term which favors parameter estimates with smaller optimal costs. Such a cost biasing was shown to yield an optimal adaptive controller for controlled Markov chains when the parameter set Θ is finite. This result was extended in Lin and Kumar [11] to the case of general Θ, for controlled Markov chains with finite state spaces. Another extension to the case of finite parameter set Θ, but allowing for a general state space was provided in Kumar [7].

In the reference most pertinent to this paper, in Kumar [6] it was shown that the cost-biased maximum likelihood based certainty equivalent controller yielded an optimal cost for linear systems with quadratic costs, as in (1) and (2), provided that the set Θ is finite.

The case of infinite parameter set Θ introduces several difficulties, and the problem has till now remained unsolved. It is the purpose of the present paper to establish the optimality of the cost biased maximum likelihood parameter estimator together with a certainty equivalent controller, for linear quadratic Gaussian systems, as in (1) and (2), for the more general case of a compact parameter uncertainty set Θ, as in (A.i) and (A.ii).

In the next section we specify the details of the cost biased least squares parameter estimator, and the certainty equivalent adaptive control law. In the subsequent section, we provide the main result, and a statement of the intermediate results through which it is established. The final section provides some concluding remarks and suggestions for future research.

2 The Adaptive Control Scheme

Given a pair of matrices $(A, B) \in \Theta$, let $J(A, B)$ denote the optimal long term average cost (2) for the system with $x_{t+1} = Ax_t + Bu_t + w_{t+1}$, where $\{w_t\}$ is i.i.d. with $N(0, I)$ distribution. For such a system, the optimal control law is

$$u_t = K(A, B)x_t,$$

where

$$K(A, B) = -[B^T P(A, B)B + T]^{-1} B^T P(A, B)A,$$

with $P(A, B)$ the unique positive semidefinite solution of the Riccati equation,

$$P = A^T PA - A^T PB(B^T PB + T)^{-1} B^T PA + Q.$$

In an adaptive control problem, the "true" matrices (A°, B°) are not known and therefore we follow a certainty equivlent methodology by replacing them by an estimate. The heart of our adaptive scheme lies in the cost-biased least squares parameter estimator described below.

We choose a deterministic time sequence $\{\mu_t\}$ such that $\mu_t \to +\infty$ and $\mu_t = o(\log t)$ as $t \to +\infty$.

The cost-biased least squares parameter estimates sequence $\{(\widehat{A}_t, \widehat{B}_t)\}$ is give by

$$(\widehat{A}_t, \widehat{B}_t) = \begin{cases} \arg\min_{(A,B)\in\Theta} \{\sum_{s=1}^{t} \|x_s - Ax_{s-1} - Bu_{s-1}\|^2 \\ \qquad + \mu_t J(A, B)\} \ \textit{for } t \textit{ even} \\ (\widehat{A}_{t-1}, \widehat{B}_{t-1}) \ \textit{for } t \textit{ odd.} \end{cases} \tag{3}$$

(when there is more than one minimizer above, any of them can be chosen).

The distinguishing feature of the criterion (3) is the term $\mu_t J(A, B)$, which introduces a mild bias in favor of parameters (A, B) with lower optimal costs. The biasing is "mild" because $\mu_t = o(\log t)$. On the other hand it is non-negligible because $\mu_t \to +\infty$. Without this term one would just have the usual least squares parameter estimator, with its attendant difficulty in identifying systems in closed-loop.

The control action is determined according to the certainty equivalent methodology as

$$u_t = K(\widehat{A}_t, \widehat{B}_t)x_t.$$

The intuitive rationale for the cost biasing in the least squares criterion is as follows. Suppose that one simply employs a straightforward least squares parameter estimator. Then, generically, it can be shown that the parameter estimator sequence $(\widehat{A}_t^{LS}, \widehat{B}_t^{LS})$ converges to a limiting random variable $(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS})$, see Kumar [8]. Such a limiting estimate results in a limiting controller $u_t = K(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS})x_t$. It is natural to expect that the least squares estimator will identify, at a minimum, the closed-loop behavior of the system. Thus, one expects that the behavior of the true system (A°, B°) with the loop closed by $K(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS})$ will be the same as the closed-loop imagined system:

$$A^\circ + B^\circ K(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS}) = \widehat{A}_\infty^{LS} + \widehat{B}_\infty^{LS} K(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS}).$$

This implies that the cost of running the true system (A°, B°) with the feedback gain $K(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS})$ is the same as the cost of running the imagined system $(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS})$ with the feedback $K(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS})$. The latter is, however, the optimal configuration for the system $(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS})$, while the former is not necessarily an optimal configuration for the system (A°, B°). Thus one has

$$J(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS}) \geq J(A^\circ, B^\circ),$$

that is, the least squares estimator has a natural tendency to return estimates with larger optimal cost than the optimal cost associated with the true system. This motivates the idea of somehow introducing a bias into the parameter estimator so that it favors parameters (A, B) with smaller values of $J(A, B)$. Thus one conceives of adding a term such as $\mu_t J(A,B)$ to the squared error. However, one needs to choose μ_t with care. One does not want to destroy the ability of the least squares estimator to identify closed loop dynamics. This is achieved by choosing $\mu_t = o(logt)$. On the other hand, one definitely does want the $\mu_t J(A,B)$ to assert itself, and this is achieved by choosing $\mu_t \to +\infty$.

Hence we arrive at the cost-biased least squares parameter estimator (3).

The analysis of this scheme is however quite intricate. Due to space limitations, details are omitted in this paper. In the next section we will indicate some of the principal intermediary steps through which the optimality of this scheme is established.

Before turning to the analysis it should be noted that the parameter estimation sequence defined in (3) is not of a recursive nature. However, one can conceive of other similarly motivated modifications to recursively specified schemes, whose goal still is to mildly favor parameter estimates with lower optimal costs; this is a topic for future research.

3 The Main Results

Let $(\widehat{A}_t^{LS}, \widehat{B}_t^{LS})$ denote the usual least square parameter estimates.

Defining $v_s^T := [x_s^T, u_s^T]$, the subspace

$$\bar{E} := \left\{ x \in \mathbf{R}^{n+m} : x^T \sum_{s=1}^{\infty} v_s v_s^T x < +\infty \right\} \quad (4)$$

is called the *unexcited* subspace and its orthogonal complement E is the *excited* subspace, see Bittanti, Bolzern and Campi [2].

Given (A,B), let $(A,B)_E$ and $(A,B)_{\bar{E}}$ denote the matrices in $\mathbf{R}^{n\times(n+m)}$ formed by projecting the rows of (A,B) onto E and $\bar{E}$, respectively.

By employing the Bayesian embedding procedure, see Kumar [8], it can be shown that there exists a random limit $(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS})$ such that,

$$\lim_{t\to\infty} (\widehat{A}_t^{LS}, \widehat{B}_t^{LS}) = (\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS}) \quad a.s.,$$

for all $(A^\circ, B^\circ) \in \Theta$, except for a Lebesgue null set in Θ.

Henceforth, we will assume that (A°, B°) does not lie in the Lebesgue null set.

Furthermore, by employing the properties of the excitation subspace, it can be shown, see Campi [4], that the true system is identified up to its unexcited components,

$$(\widehat{A}_\infty^{LS}, \widehat{B}_\infty^{LS})_E = (A^\circ, B^\circ)_E \quad a.s.$$

This last equation reveals that the asymptotic error of the least squares estimate is confined to the unexcited subspace. Taking into account that the total amount of information in this subspace is finite (see definition (4)), it is possible to show that the biasing term $\mu_t J(A,B)$ will eventually succeed in pushing the estimate away from the set where the value of the optimal cost is higher than the optimal cost for the true system.

Theorem 1

$$\limsup_{t\to\infty} J(\widehat{A}_t, \widehat{B}_t) \leq J(A^\circ, B^\circ) \quad a.s.$$

Next one turns to showing that the biasing sequence does not damage the valuable ability of the least squares parameter estimator to identify at least the closed loop dynamics of the control system. This is done by showing that for all $\delta > 0$,

$$\sum_{s=1}^{t} 1((\widehat{A}_s, \widehat{B}_s) \in \mathcal{B}_\delta) = O(\mu_t) \quad a.s., \quad (5)$$

where $1(A)$ is the indicator function of a set A, and

$$\mathcal{B}_\delta := \{(A,B) \in \Theta : \|A^\circ + B^\circ K(A,B) - A - BK(A,B)\| \geq \delta\}.$$

In the proof of equation (5), the fact that the parameter estimate is kept constant at each odd time instant is used. Suppose that at a certain odd time instant s the estimate $(\widehat{A}_s, \widehat{B}_s) = (\widehat{A}_{s-1}, \widehat{B}_{s-1})$ does not fall in the set $\mathcal{B}_\delta$. Then,

$$\|x_{s+1} - Ax_s - Bu_s\|^2$$
$$= \left\| ([A^\circ + B^\circ K(\widehat{A}_{s-1}, \widehat{B}_{s-1})] - [A + BK(\widehat{A}_{s-1}, \widehat{B}_{s-1})])x_s + w_{s+1} \right\|^2$$

is such that $[A^\circ + B^\circ K(\widehat{A}_{s-1}, \widehat{B}_{s-1})] - [A + BK(\widehat{A}_{s-1}, \widehat{B}_{s-1})]$ is away from zero for any (A,B) in a suitable neighborhood of $(\widehat{A}_{s-1}, \widehat{B}_{s-1})$. This discrepancy is emphasized by x_s since it is driven by the noise w_s which is independent of $[A^\circ + B^\circ K(\widehat{A}_{s-1}, \widehat{B}_{s-1})] - [A + BK(\widehat{A}_{s-1}, \widehat{B}_{s-1})]$. A careful use of this argument leads to the conclusion that the estimate $(\widehat{A}_s, \widehat{B}_s)$ cannot visit too often the set $\mathcal{B}_\delta$, as stated in equation (5).

As a consequence of result (5) and Theorem 1, it follows that the adaptive gain $K(\widehat{A}_s, \widehat{B}_s)$ is nearly optimal except for the exceptional time instants (self-tuning property).

Theorem 2

$$\sum_{s=0}^{t-1} 1\left(\left\|K(\widehat{A}_s,\widehat{B}_s)-K(A^\circ,B^\circ)\right\|>\delta\right) = O(\mu_t)$$

a.s., for all $\delta > 0$.

The intuitive rationale in the proof of this theorem is simple. Equation (5) asserts that closed-loop identification holds with the exception of very rare time instants. This in turn implies that the cost obtained by running the true system with the estimated feedback gain is close to the optimal cost for the estimated system. On the other hand, by Theorem 1 the latter is less than or equal to the optimal cost for the true system in the long run. The only way to reconcile these two facts is to conclude that the estimated gain leads to a feedback loop whose cost is close to the optimal cost. The conclusion of Theorem 2 is finally drawn by observing that the optimal gain matrix for the true system is unique.

The closed-loop system is time-varying but, in view of Theorem 2, it is destabilized very rarely. From this, one can establish the stability of the system.

Theorem 3

$$\limsup_{t\to\infty}\frac{1}{t}\sum_{s=0}^{t-1}(\|x_s\|^p+\|u_s\|^p) < +\infty \quad a.s.,$$

for all p.

The last result we want to discuss is the optimality of the proposed scheme.

The dynamic programming equation for the long-term average cost control problem at hand, see Kumar and Varaiya [10], allows one to write,

$$\frac{1}{t}\sum_{s=0}^{t-1}[x_s^TQx_s+u_s^TTu_s]$$

$$=\frac{1}{t}\sum_{s=0}^{t-1}J(\widehat{A}_s,\widehat{B}_s)$$

$$+\frac{1}{t}\sum_{s=0}^{t-1}\Big\{x_s^TP(\widehat{A}_s,\widehat{B}_s)x_s$$

$$-E[x_{s+1}^TP(\widehat{A}_{s+1},\widehat{B}_{s+1})x_{s+1}/\text{past up to time s}]\Big\}$$

$$-\frac{1}{t}\sum_{s=0}^{t-1}E[x_{s+1}^T(P(\widehat{A}_s,\widehat{B}_s)$$

$$-P(\widehat{A}_{s+1},\widehat{B}_{s+1}))x_{s+1}/\text{past up to time s}]$$

$$-\frac{1}{t}\sum_{s=0}^{t-1}\Big\{(\widehat{A}_sx_s+\widehat{B}_su_s)^TP(\widehat{A}_s,\widehat{B}_s)(\widehat{A}_sx_s+\widehat{B}_su_s)$$

$$-(A^\circ x_s+B^\circ u_s)^TP(\widehat{A}_s,\widehat{B}_s)(A^\circ x_s+B^\circ u_s)\Big\}.$$

By Theorem 1, the first term in the right-hand-side is less than or equal to the optimal cost $J(A^\circ,B^\circ)$ as $t\to\infty$, while the second term disappears by standard martingale convergence results. Then, to prove optimality one has only to show that the last two terms in the right-hand-side converge to zero. The convergence of the last but one term is proven by showing that matrix $P(\widehat{A}_s,\widehat{B}_s)$ is nearly invariant, with rare exceptions in time and by making use of the stability established in Theorem 3. As for the last term, its convergence follows from the closed-loop identification result (5).

This leads to the following theorem

Theorem 4

$$\limsup_{t\to\infty}\frac{1}{t}\sum_{s=0}^{t-1}(x_s^TQx_s+u_s^TTu_s)=J(A^\circ,B^\circ) \quad a.s.$$

Hence the cost biased adaptive control scheme overcomes the closed-loop identifiability problem, and is indeed self-optimizing.

4 Concluding Remarks

A fundamental difficulty in many minimum variance type adaptive control schemes is the problem of handling non-minimum phase systems, due to the lack of any weighting of the input energy. This can be alleviated by considering a full quadratic cost criterion. Then, however, one runs into the fundamental obstacle to adaptive control posed by the inability of identifying a system when it operates in closed-loop.

A way out of this dilemma is to employ a more fine grained scheme which carefully exploits properties of the set to which the least squares parameter estimates converge, namely that their optimal cost must necessarily be larger than the optimal cost. Such a scheme is presented and analyzed in this paper.

Unfortunately, there are two drawbacks. First, the scheme presented is nonrecursive. However, we feel that this can perhaps be removed by employing a modification to a recursive scheme, using the biasing idea suggested here. The second drawback is that we have only addressed systems where the full state is observed. This needs to be removed.

Both the above problems suggest useful research opportunities.

Acknowledgment

M. C. Campi would like to acknowledge the financial support of MURST under the 60% project "Adaptive

identification, prediction and control".
The research of P. R. Kumar has been supported in part by the U.S. Army Research Office under Contract No. DAAH-04-95-1-00090, an in part by the Joint Service Electronics Program under Contract No. N00014-96-1-0129.

References

[1] A. Becker, P. R. Kumar, and C. Z. Wei. Adaptive control with the stochastic approximation algorithm: Geometry and convergence. *IEEE Transactions on Automatic Control*, AC-30(4):330–338, 1985.

[2] S. Bittanti, P. Bolzern, and M. Campi. Recursive least squares identification algorithms with incomplete excitation: Convergence analysis and application to adaptive control. *TAC*, 35(12):1371–1373, 1990.

[3] V. Borkar and P. P. Varaiya. Adaptive control of Markov chains, I: Finite parameter set. *IEEE Transactions on Automatic Control*, AC-24:953–958, 1979.

[4] M.C. Campi. The Problem of Pole-Zero Cancellation in Transfer Function Identification and Application to Adaptive Stabilization. *Automatica*, 32: 849-857, 1996.

[5] P. R. Kumar. Optimal adaptive control of linear-quadratic-Gaussian systems. *SIAM Journal on Control and Optimization*, 21(2):163–178, 1983.

[6] P. R. Kumar. Optimal adaptive control of linear quadratic Gaussian systems. *SIAM Journal on Control and Optimization*, 163–178, 1983.

[7] P. R. Kumar. Simultaneous identification and adaptive control of unknown systems over finite parameter sets. *IEEE Transactions on Automatic Control*, AC-28:68–76, 1983.

[8] P. R. Kumar. Convergence of least-squares parameter estimate based adaptive control schemes. *IEEE Transactions on Automatic Control*, AC-35(5):416–424, 1990.

[9] P. R. Kumar and A. Becker. A new family of optimal adaptive controllers for Markov chains. *IEEE Transactions on Automatic Control*, AC-27:137–146, 1982.

[10] P. R. Kumar and P. P. Varaiya. *Stochastic Systems: Estimation, Identification and Adaptive Control.* Englewood Cliffs, NJ: Prentice-Hall, 1986.

[11] W. Lin and P. R. Kumar. Optimal adaptive controllers for unknown systems. *IEEE Transactions on Automatic Control*, AC-27:765–774, 1982.

[12] W. Lin, P. R. Kumar, and T. I. Seidman. Will the self-tuning approach work for general cost criteria? *Systems & Control Letters*, 6(2):77–85, 1985.