An Actively Adaptive Control Policy for Linear Models

L. Pronzato, C. Kulcsár and E. Walter

Abstract

We consider the quadratic optimal control of a discrete-time linear system with unknown parameters. The control action affects both the future behavior of the system and the precision of the future state and parameter estimates (dual effect).

Using the information matrix, the posterior uncertainty on the parameters can be predicted, in order to design an *actively adaptive* control policy. The case of an autoregressive model is considered. Illustrative examples are presented, together with possible extensions and remaining open questions.

AMS 1991 Subject Classification—93E35, 49L20, 62K05

Key words—Adaptive control, dual control, stochastic control, experiment design, active learning, autoregressive models, information matrix.

I. Introduction

We consider the quadratic optimal control of a linear system with unknown parameters. The present control affects both the future behavior of the system and the precision of the future state and parameter estimates. This is the well known dual effect [1], and optimal control should take into account two conflicting objectives: (i) to follow the prescribed state trajectory, (ii) to help estimate the state and/or parameters in order to control the system better. When the model parameters θ are known and only the state has to be estimated, the Certainty Equivalence Principle (CEP) leads to classical LQ feedback control laws, see, e.g., [2], [3]. This principle does not hold true in the case of unknown parameters, and the corresponding stochastic dynamic programming problem can usually not be solved.

The most widely used suboptimal policy corresponds to Forced (or heuristic) Certainty Equivalence Principle (FCEP) [4], for which each new control action is computed on the basis of the current state and parameter estimates. Possible improvements correspond to open-loop feedback optimal [5] and cautious [6] control.

All these suboptimal policies are *passively adaptive*, i.e. they do not anticipate their consequences on the estimation, and do not take the presence of future measurements into account. Conversely, *active learning* means taking into account the fact that the model parameters will be estimated from the data to be collected. The design of an active-control policy requires the evaluation of the so-called *optimal cost-to-go*, which, in general, can only be approximated. For this purpose, we shall use ideas from *optimal experiment design* for parameter estimation, which is based on the prediction of the consequences of a present action on the future precision of the model parameters, and is thus intrinsically *active*.

Section II states the problem as one of dynamic programming and defines the notations to be used. Section III presents the actively adaptive algorithm that is proposed. Two simple illustrative examples are given in Section IV, and some possible extensions as well as remaining open questions are stated

Author for correspondence. Laboratoire I3S, bât. 4, CNRS-URA1376, Sophia Antipolis, 06560, Valbonne, France. E-mail: pronzato@unice.fr, Fax: (33) 92 94 28 98.

Laboratoire des Signaux et Systèmes, CNRS-SUPELEC, 91192 Gif-sur-Yvette, France.