Stochastic iterative dynamic programming: a Monte Carlo approach to dual control

Adrian M. Thompson, William R. Cluett

Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, Ont., Canada M5S 3E5

Received 5 July 2002; received in revised form 15 November 2004; accepted 7 December 2004

Abstract

Practical exploitation of optimal dual control (ODC) theory continues to be hindered by the difficulties involved in numerically solving the associated stochastic dynamic programming (SDP) problems. In particular, high-dimensional hyper-states coupled with the nesting of optimizations and integrations within these SDP problems render their exact numerical solution computationally prohibitive. This paper presents a new stochastic dynamic programming algorithm that uses a Monte Carlo approach to circumvent the need for numerical integration, thereby dramatically reducing computational requirements. Also, being a generalization of iterative dynamic programming (IDP) to the stochastic domain, the new algorithm exhibits reduced sensitivity to the hyper-state dimension and, consequently, is particularly well suited to solution of ODC problems. A convergence analysis of the new algorithm is provided, and its benefits are illustrated on the problem of ODC of an integrator with unknown gain, originally presented by Åström and Helmersson (Computers and Mathematics with Applications 12A (1986) 653–662).

Keywords: Dual control; Adaptive control; Stochastic systems; Dynamic programming; Optimal control; Uncertainty

1. Introduction

The optimal dual control (ODC) idea is to predict and make use of future model estimates during the control calculation. Instead of using a fixed model for computation of the entire control policy, in ODC, later controls are computed using models predicted to result from those controls applied earlier. The resulting optimal policy balances system excitation, to improve future model accuracy, against system regulation, to achieve good immediate control. The name “dual control” arises from the need to simultaneously satisfy these two competing objectives.

ODC has proven extremely difficult to implement in practice due to several computational issues and as a result most researchers in the dual control field have been discouraged from attempting to solve the ODC problem. Instead, the recent trend has been towards development of sub-optimal approaches to dual control. Although proving simpler than ODC to implement in practice, these approaches have several disadvantages (Lindoff, Holst, & Wittenmark, 1999; Filatov & Unbehauen, 2000).

Numerically, ODC involves computation of a state- and time-dependent control policy defined over a finite control horizon to minimize a stochastic cost function subject to the dynamics of a non-linear time-varying hyper-system with continuous hyper-state1 and control spaces. Such problems are extremely difficult to solve because (i) existing numerical approaches encounter the curse of dimensionality, which is amplified in ODC due to the presence of model parameters and their uncertainty descriptions within the hyper-state, (ii)
2. Optimal dual control

2.1. Problem description

The ODC problem of interest in this paper takes the following form:

(1) Given a single-input single-output (SISO), linear time invariant (LTI), uncertain model of a dynamical system, of the form

\[ y_{t+1} = \varphi_t^T \theta_t + e_{t+1}, \]

\[ \theta_t \sim N(\hat{\theta}_t, P_t), \]

\[ e_t \sim N(0, \sigma^2), \]

\[ \text{Cov}(e_t, \theta_t;i) = 0 \quad \forall t, i, \]

in which

\[ \varphi_t = \begin{bmatrix} u_t \\ u_{t-1} \\ \vdots \\ -y_{t-n_y+1} \\ -y_{t-1} \\ \vdots \\ -y_{t-n_y+1} \end{bmatrix} \quad \text{and} \quad \theta_t = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n_u} \\ a_1 \\ \vdots \\ a_{n_y} \end{bmatrix}, \]

(2) with model adaptation governed by the equations

\[ K_t = P_t \varphi_t (\varphi_t^T P_t \varphi_t + \sigma^2)^{-1}, \]

\[ \hat{\theta}_{t+1} = \hat{\theta}_t + K_t (y_{t+1} - \varphi_t^T \hat{\theta}_t), \]

\[ P_{t+1} = (I - K_t \varphi_t^T) P_t, \]

(3) compute the feedback control policy

\[ u_t^* = u_t^r (y_t, y_{t-1}, \ldots, y_0, u_{t-1}, u_{t-2}, \ldots, u_0), \]

\[ t = 0, 1, 2, \ldots, N - 1, \]

that minimizes the expected tracking error

\[ J_0 = E \left\{ \sum_{i=0}^{N-1} (y_{t+1} - r_{t+1})^2 \right\}, \]

(4) using a priori knowledge of \( \varphi_0, \hat{\theta}_0, P_0, \sigma^2 \) and \( N \).
دریافت فوری
متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات