Automatica 72 (2016) 37–45

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Output-feedback adaptive optimal control of interconnected systems based on robust adaptive dynamic programming ☆

Weinan Gao [a], Yu Jiang [a], Zhong-Ping Jiang [a,b], Tianyou Chai [b]

[a] Department of Electrical and Computer Engineering, Tandon School of Engineering, New York University, Brooklyn, NY 11201, USA

[b] State Key Lab of Synthetical Automation for Process Industries, Northeastern University, Shenyang, Liaoning 110819, China

ARTICLE INFO

Article history:
Received 6 August 2014
Received in revised form
29 December 2015
Accepted 5 May 2016

Keywords:
Approximate/adaptive dynamic programming (ADP)
Output-feedback control
Nonlinear dynamic uncertainty
Robust optimal control

ABSTRACT

This paper studies the adaptive and optimal output-feedback problem for continuous-time uncertain systems with nonlinear dynamic uncertainties. Data-driven output-feedback control policies are developed by approximate/adaptive dynamic programming (ADP) based on both policy iteration and value iteration methods. The obtained adaptive and optimal output-feedback controllers differ from the existing literature on the ADP in that they are derived from sampled-data systems theory and are guaranteed to be robust to dynamic uncertainties. A small-gain condition is given under which the overall system is globally asymptotically stable at the origin. An application to power systems is given to test the effectiveness of the proposed approaches.

1. Introduction

As an important subfield of modern control theory, optimal control aims to develop controllers that optimize certain performance (see Lewis, Vrabie, & Syrmosc, 2012). Traditional optimal control methods need to solve the Hamilton–Jacobi–Bellman (HJB) equation (or algebraic Riccati equation (ARE) for linear systems) via the perfect knowledge of the system dynamics. Unfortunately, it is often difficult to obtain an accurate mathematical model for real-world, modern engineering, and natural systems.

Approximate/adaptive dynamic programming (ADP) (see, e.g., Lewis & Liu, 2013; Ni, He, & Wen, 2013; Si, Barto, Powell, & Wunsch, 2004; Vamvoudakis, 2014; Werbos, 1974, 1990; Zhang, Liu, Luo, & Wang, 2013) is a non-model-based approach which gives rise to online approximation of optimal solutions via some recursive numerical methods. The related research has been studied for discrete-time Markov decision processes (Bertsekas & Tsitsiklis, 1996; Powell, 2007; Sutton & Barto, 1998) and discrete-time feedback control systems (Lewis & Vrabie, 2009; Lewis, Vrabie, & Vamvoudakis, 2012; Liu & Wei, 2014; Wang, Zhang, & Liu, 2009). For continuous-time systems, related work can be found in Bhasin, Sharma, Parte, and Dixon (2011), Jiang and Jiang (2012a, 2014), Luo, Wu, Huang, and Liu (2014), Vrabie, Pastravanu, Abu-Khalaf, and Lewis (2009), and Yang, Liu, Ma, and Xu (2016). ADP and output regulation theory is firstly integrated by Gao and Jiang (2016) to solve the problem of asymptotic tracking with disturbance rejection.

Recently, extending these solutions to output-feedback problems has received attention from Lewis and Vamvoudakis (2011), Gao, Huang, Jiang, and Chai (2016); Gao, Jiang, Jiang, and Chai (2014) and Zhu, Modares, Peen, Lewis, and Yue (2015) for linear systems and He and Jagannathan (2005) and Liu, Huang, Wang, and Wei (2013) for nonlinear systems based on neural networks (Ge, Lee, & Harris, 1998). A common feature of these papers is that no dynamic uncertainty (Jiang & Mareels, 1997) is addressed. However, there are numerous practical examples of continuous-time systems arising from engineering and biology for which dynamic uncertainty is unavoidable.

The contribution of this paper is threefold. First, different from existing output-feedback ADP for discrete-time linear systems (Lewis & Vamvoudakis, 2011) or continuous-time static output-feedback ADP design which requires the accurate knowledge of the input matrix (Zhu et al., 2015), a dynamic output-feedback ADP approach is proposed for continuous-time linear systems without the exact knowledge of any system matrices. By employing

☆ This work has been supported in part by the U.S. National Science Foundation grant ECCS-1501044, and in part by the National Natural Science Foundation of China grant 61374042, and the Project 111 (No. B08015). The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Raul Ordonez under the direction of Editor Miroslav Krstic.

E-mail addresses: weinan.gao@nyu.edu (W. Gao), yu.jiang@nyu.edu (Y. Jiang), zjiang@nyu.edu (Z.-P. Jiang), tychai@mail.neu.edu.cn (T. Chai).

http://dx.doi.org/10.1016/j.automatica.2016.05.008
0005-1098/

the sampled-data system theory (Chen & Francis, 1995), the unmeasurable state can be reconstructed in terms of input/output data, whereby we can iteratively solve the ARE. Second, we, for the first time, develop an output-feedback value iteration (VI) ADP algorithm for continuous-time linear systems by using the sampled-data approach.

As the third contribution, we study the robust optimal redesign issue for a class of interconnected systems with dynamic uncertainties. The state and order of dynamic uncertainties are unknown. Because of the implementation of sampled-data output-feedback robust optimal controllers, the closed-loop system is a hybrid system that involves both continuous-time and discrete-time dynamics. Thus, robustness analysis cannot be conducted directly by our previous work on state-feedback robust ADP (Jiang & Jiang, 2013, 2014). Instead, we derive the global asymptotic stability of the closed-loop interconnected system based on a combined application of Lyapunov theory, sampled-data systems theory, and nonlinear small-gain method. To the best of the authors' knowledge, this paper represents the first step towards the ADP design of output-feedback adaptive optimal controllers for continuous-time nonlinear systems with both static and dynamic uncertainties.

The remainder of this paper is organized as follows. In Section 2, we formulate the control problem, and briefly review the linear–quadratic regulator (LQR) theory. In Section 3, we develop adaptive optimal output-feedback strategies by using both policy iteration (PI) and VI based ADP methods. Robustness and suboptimality of the closed-loop system are analyzed in Section 4. An application to a practical example on the power systems is presented in Section 5. Finally, conclusions are contained in Section 6.

Notations. Throughout this paper, $\mathbb{R}_+$ denotes the set of nonnegative real numbers. $|\cdot|$ represents the Euclidean norm for vectors and the induced norm for matrices. A continuous function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ belongs to class $\mathcal{K}$ if it is increasing and $\alpha(0) = 0$. It belongs to class $\mathcal{K}_\infty$ if, in addition, it is proper. A continuous function $\beta : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ belongs to class $\mathcal{KL}$ if for each fixed t, the function $\beta(\cdot, t)$ is of class $\mathcal{K}$ and, for each fixed s, the function $\beta(s, \cdot)$ is non-increasing and tends to 0 at infinity. $\otimes$ indicates the Kronecker product operator and $\mathrm{vec}(A) = [a_1^T, a_2^T, \ldots, a_m^T]^T$, where $a_i \in \mathbb{R}^n$ are the columns of $A \in \mathbb{R}^{n\times m}$. For a symmetric matrix $P \in \mathbb{R}^{m\times m}$, $\mathrm{vecs}(P) = [p_{11}, 2p_{12}, \ldots, 2p_{1m}, p_{22}, 2p_{23}, \ldots, 2p_{m-1,m}, p_{mm}]^T \in \mathbb{R}^{\frac{1}{2}m(m+1)}$. For an arbitrary column vector $v \in \mathbb{R}^n$, $\mathrm{vecv}(v) = [v_1^2, v_1v_2, \ldots, v_1v_n, v_2^2, v_2v_3, \ldots, v_{n-1}v_n, v_n^2]^T \in \mathbb{R}^{\frac{1}{2}n(n+1)}$. $\lambda_M(P)$ and $\lambda_m(P)$ denote the maximum and the minimum eigenvalue of the real symmetric matrix P. For any piecewise continuous function $u : \mathbb{R}_+ \to \mathbb{R}^m$, $\|u\|$ stands for $\sup_{t\geq 0} |u(t)|$.

2. Problem formulation and preliminaries

Consider a linear subsystem interacting with a nonlinear subsystem known as the dynamic uncertainty, characterized by the ζ-system:

$$\dot{x} = Ax + B(u + \Delta(\zeta, y)), \tag{1}$$

$$\dot{\zeta} = g(\zeta, y), \tag{2}$$

$$y = Cx \tag{3}$$

where $x \in \mathbb{R}^n$, $\zeta \in \mathbb{R}^p$ are unmeasurable states with an unknown integer $p > 0$, $u \in \mathbb{R}^m$ the input, $y \in \mathbb{R}^r$ the output. $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $C \in \mathbb{R}^{r\times n}$ are unknown system matrices with (A, B) controllable, (A, C) observable satisfying $|A| \leq A_M$, $|B| \leq B_M$, and $|C| \leq C_M$. $g : \mathbb{R}^p \times \mathbb{R}^r \to \mathbb{R}^p$ and $\Delta : \mathbb{R}^p \times \mathbb{R}^r \to \mathbb{R}^m$ are two unknown, locally Lipschitz functions with $g(0, 0) = 0$ and $\Delta(0, 0) = 0$.

Remark 2.1. The system (1)–(3) belongs to the class of interconnected systems studied by Saberi, Kokotovic, and Summers (1990). If $\Delta(\zeta, y) = \Delta_1(y)$, the system (1) and (3) is a linear system with nonlinear output injection (see Krener & Isidori, 1983).

Assumption 2.1. The ζ-system with y regarded as the input and Δ as the output has the strong unboundedness observability (SUO) property with zero offset (Jiang, Teel, & Praly, 1994), i.e., there exist a function σ_1 of class $\mathcal{KL}$ and a function γ_1 of class $\mathcal{K}$ such that for any measurable essentially bounded control $y(t)$ on $[0, T)$ with $0 < T \leq +\infty$, the solution $\zeta(t)$ of (2) right maximally defined on $[0, T')(0 < T' \leq T)$ satisfies

$$|\zeta(t)| \leq \sigma_1(|\zeta(0)|, t) + \gamma_1(\|[y_{[0,t]}^T, \Delta_{[0,t]}^T]^T\|), \quad \forall t \in [0, T')$$

where $y_{[0,t]}$ and $\Delta_{[0,t]}$ are truncated functions of y and Δ over $[0, t]$, respectively.

Assumption 2.2. The ζ-system is input-to-output stable (IOS) (Sontag, 2007), i.e., there exist a function σ_Δ of class $\mathcal{KL}$ and a function γ_Δ of class $\mathcal{K}$ such that, for any initial state $\zeta(0)$, any measurable essentially bounded input y and any $t \geq 0$,

$$|\Delta(t)| \leq \sigma_\Delta(|\zeta(0)|, t) + \gamma_\Delta(\|y\|). \tag{4}$$

Considering the reduced-order system (1) and (3) in the absence of the dynamic uncertainty

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ y &= Cx, \end{aligned} \tag{5}$$

define the cost as

$$J_c(x(0)) = \int_0^\infty (y^T(\tau)Qy(\tau) + u^T(\tau)Ru(\tau))d\tau \tag{6}$$

where $Q = Q^T \geq 0$ and $R = R^T > 0$ with $(A, \sqrt{Q}C)$ observable. Moreover, a minimal cost $J_c^* = x^T(0)P^*x(0)$ in (6) is obtained by using the following control policy

$$u = -R^{-1}B^TP^*x := -K^*x \tag{7}$$

where $P^* = (P^*)^T > 0$ is the unique solution to the algebraic Riccati equation (ARE):

$$A^TP^* + P^*A + C^TQC - P^*BR^{-1}B^TP^* = 0. \tag{8}$$

A discretized model of (5) is obtained by taking periodic sampling

$$\begin{aligned} x_{k+1} &= A_dx_k + B_du_k, \\ y_k &= Cx_k \end{aligned} \tag{9}$$

where $A_d = e^{Ah}$, $B_d = (\int_0^h e^{A\tau}d\tau)B$, and $h > 0$ is the sampling period. Suppose the sampling frequency $\omega_h = 2\pi/h$ is non-pathological (Chen & Francis, 1995). Then, both (A_d, C) and $(A_d, \sqrt{Q}C)$ are observable with (A_d, B_d) controllable. The cost for (9) is

$$J_d(x(0)) = \sum_{j=0}^\infty (y_j^TQ_dy_j + u_j^TR_du_j) \tag{10}$$

where $Q_d = Qh$, $R_d = Rh$. Notice that (10) can be viewed as a first-order approximation of equation (7) in Melzer and Kuo (1971) which itself is the discretized version of cost (6).

The optimal control law minimizing (10) is

$$u_k = -(R_d + B_d^TP_d^*B_d)^{-1}B_d^TP_d^*A_dx_k := -K_d^*x_k \tag{11}$$

where $P_d^* = P_d^{*T} > 0$ is the unique solution to

$$\begin{aligned} &A_d^TP_d^*A_d - P_d^* + C^TQ_dC \\ &- A_d^TP_d^*B_d(R_d + B_d^TP_d^*B_d)^{-1}B_d^TP_d^*A_d = 0. \end{aligned} \tag{12}$$

The sensitivities of P_d^* and K_d^*, with respect to sampling period h, are discussed in the following lemma.