Neurocomputing 135 (2014) 163–170

Contents lists available at ScienceDirect

Neurocomputing

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

Model-free optimal control design for a class of linear discrete-time systems with multiple delays using adaptive dynamic programming

Jilie Zhang [a], Huaguang Zhang [a,b,*], Yanhong Luo [a], Tao Feng [a]

[a] School of Information Science and Engineering, Northeastern University, Shenyang, Liaoning 110819, P.R. China

[b] State Key Laboratory of Synthetical Automation for Process Industries (Northeastern University), Shenyang, Liaoning 110819, P.R. China

ARTICLE INFO

Article history:
Received 9 July 2013
Received in revised form
4 November 2013
Accepted 16 December 2013
Communicated by D. Liu
Available online 24 January 2014

Keywords:
Model-free optimal control
Discrete-time delay system
Optimal control
Adaptive dynamic programming

ABSTRACT

In this paper, a model-free optimal control scheme for a class of linear discrete-time systems with multiple delays in state, control and output vectors is proposed. The optimal control can be obtained using only measured input/output data from systems, by adaptive dynamic programming (ADP) technology. First, we give a class of systems what we want to address. Then, a model-free optimal control is designed to minimize the given cost functional by ADP technology, which combines a similar *Q*-learning method with a value iteration (VI) algorithm, using only the measured input/output data. Finally, several numerical examples are given to illustrate the effectiveness of our approach.

1. Introduction

Since systems with time delay phenomena are ubiquity in various research fields, such as biology, chemistry, economics, mechanics, electrical, physics, as well as engineering sciences [1–4], the optimal control problem is discussed as a key topic for time-delay problems in [5–7] over the past several decades. In fact, the optimal control for time-delay systems is an infinite-dimensional control problem [8], which is hard to be solved. However, because adaptive (approximate) dynamic programming is a powerful tool for solving optimal control problems [9–11], the optimal control based on ADP attracts considerably attention of researchers.

In recent years, the ADP is used to design the optimal control for control systems [12–18,20,32–34]. The optimal control problem for continuous-time systems is studied in [12–15,17,20,32,34]. While Refs. [16,18,33] design the optimal control for discrete-time systems. However, to the best of our knowledge, the optimal control results based on ADP for time-delay systems are rare. There exist only some relevant results, such as [19,21,22]. An optimal control scheme for nonlinear systems with delays is proposed by using a new iterative ADP algorithm in [21]. In [19], a new iterative heuristic dynamic programming (HDP) algorithm is proposed to solve the optimal control problem for a class of nonlinear discrete time-delay systems with saturating actuators. The local and global optimization searching processes are developed to solve the optimal control problem in the iterative HDP algorithm. Later, Ref. [22] designs the optimal control for tracking control systems by a novel HDP iteration algorithm which contains state updating, control policy iteration and performance index iteration. However, most of the above results design the optimal control for time-delay systems with known knowledge of systems.

Although ADP algorithms, which are used to obtain the optimal control for time-delay systems, have made some progress, how to design the model-free optimal control for time-delay systems is still an open problem. For a simple case without delays, Lewis has made a contribution [23] to the model-free optimal control design, but few researches focus on designing the model-free optimal control for systems with multiple time delays. Therefore, a control we present by the method in [23] is used to drive the time-delay systems, rather than the systems without delays. Namely, we design the optimal control for the equivalent systems by the method in [23], then drive the original systems using it. However, the system must satisfy the certain conditions. Although the systems we address are not general, it has been a progress for designing the model-free optimal control for time-delay systems in the ADP field. The other contribution is that we find a class of systems with delays, which can be drove by an optimal control without delays.

In this paper, we not only expand the necessary and sufficient conditions in [24] to linear discrete-time systems with multiple

* Corresponding author.
E-mail addresses: jilie0226@163.com (J. Zhang), hgzhang@ieee.org (H. Zhang), neuluo@gmail.com (Y. Luo), sunnyfengtao@163.com (T. Feng).

0925-2312/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.12.038

delays by using the bicausal change of coordinates method [25], but also design the model-free optimal control for systems with multiple delays by using the design approach of state estimator based on measured input/output data in [26,27].

The following several factors motivate our research on the model-free optimal control for systems with multiple delays: first, since the system model is usually unknown, the model-based optimal control is not available in practical situations. Therefore, the optimal control which is not dependent on the knowledge of the systems is very useful in practice. This point motivates our research on the model-free optimal control based on ADP for systems with multiple delays. Second, the model-free optimal control design based on ADP for time-delay is an open problem. There are not relevant results at present, while our work deals with the optimal control problem for systems with multiple delays for the first time. Finally, seeking a model-free optimal control for systems with delays is a problem that it is difficult to be solved, because the form of optimal cost functional for systems with multiply time delays in state, input and output vectors cannot be predicted as linear systems. It also motivates our interest.

Here, the model-free optimal control for a class of systems with multiple delays is successfully designed by ADP technology, using the measured input/output data from systems.

This paper is organized as follows. Section 3 gives a class of systems what we want to address. In Section 4, a model-free optimal control for systems with multiple delays is designed by ADP technology, using the measured input-output data. Finally, several numerical examples are given to illustrate the effectiveness of our approach.

Note: In this paper, we use $Q>0$ ($Q<0$, $Q\geq 0$, $Q\leq 0$) to denote the Q matrix as positive (negative, positive semidefiniteness and negative semidefiniteness).

2. System description and preliminaries

2.1. Problem description

Now, we concentrate our attention on linear discrete-time systems with multiple delays in state, input and output vectors. The system model is given by

$$x(k+1)=\sum_{i=0}^{\tau_a}A_ix(k-i)+\sum_{j=0}^{\tau_b}B_ju(k-j) \tag{1a}$$

$$y(k)=\sum_{h=0}^{\tau_c}C_hx(k-h) \tag{1b}$$

where the state $x(k)\in\mathbb{R}^n$, the input $u(k)\in\mathbb{R}^m$ and the measurement output vector $y(k)\in\mathbb{R}^q$; $A_i\in\mathbb{R}^{n\times n}(i=0,\dots,\tau_a)$, $B_j\in\mathbb{R}^{n\times m}$ $(j=0,\dots,\tau_b)$ and $C_h\in\mathbb{R}^{q\times n}(h=0,\dots,\tau_c)$, τ_a,τ_b and $\tau_c\in\mathbb{N}$.

Assumption 1. The system (1) is controllable and observable.

Problem 1. The control object is how to find a control $u(k)$ to minimize the following cost functional subject to the system (1):

$$J(x(k))=\sum_{i=k}^{\infty}\ell(x(i),u(i)), \tag{2}$$

$\ell(x(i),u(i))=y(i)^TQy(i)+u(i)^TRu(i)$ is a utility function, where Q and R are constant weight matrices such that $Q=Q^T>0$ and $R=R^T>0$.

2.2. Preliminaries

We define ∇ as the delay operator, i.e., $\nabla^ix(k)=x(k-i)$, with $i\in\mathbb{N}$. Let $\mathbb{R}[\nabla]$ be the ring of polynomials in ∇ with coefficients in $\mathbb{R}$. Polynomial matrices with ∇ may be written as

$M[\nabla]=M_0+M_1\nabla+\cdots+M_{r_a}\nabla^{r_a}$,

where $M_i(i=0,\dots,r_a)$ are constant matrices, $r_a\in\mathbb{N}$. The addition and the multiplication are defined as usual

$$M[\nabla]+N[\nabla]=\sum_{i=0}^{\sup[r_M,r_N]}(M_i+N_i)\nabla^i,$$

$$M[\nabla]N[\nabla]=N[\nabla]M[\nabla]=\sum_{i=0}^{r_M}\sum_{j=0}^{r_N}M_iN_j\nabla^{i+j}$$

Definition 1 (*Unimodular matrix*). *A* polynomial matrix $A\in\mathbb{R}[\nabla]^{n\times n}$ is said to be unimodular if it has a polynomial inverse on the same ring.

Definition 2 (*Smith invariant factors [24]*). Every $m\times n$ matrix polynomial $P(\lambda)$ of rank r is equivalent to the matrix

$S(\lambda):P(\lambda)=U_1(\lambda)S(\lambda)U_2(\lambda)$

with

$$S(\lambda)=\begin{bmatrix}\Delta(\lambda) & 0\\ 0 & 0\end{bmatrix},$$

$\Delta(\lambda)=\mathrm{diag}[d_1(\lambda),\dots,d_r(\lambda)]$ such that $d_i(\lambda)$ is divisible by $d_{i-1}(\lambda)$ for $i=2,\dots,r$ for some unimodular matrices $U_1(\lambda)$ and $U_2(\lambda)$. The matrix polynomial $S(\lambda)$ is called the Smith form of $P(\lambda)$ and the diagonal elements $d_i(\lambda)$ are the invariant factors.

Definition 3 (*Change of coordinates [25]*). Considering the time-delay system (1), with state coordinates $x(k)$, then

$$z(k)=T[\nabla]x(k)\quad\text{with } T[\nabla]\in\mathbb{R}[\nabla]^{n\times n} \tag{3}$$

is a causal change of coordinates if the Smith Invariant Factors of $T[\nabla]$ have the form ∇^{τ_i} for some $\tau_i\in\mathbb{N}$. If $\tau_i=0,\ \forall i=1,\dots,n$, then the change of coordinates is said to be bicausal.

Definition 4 (*Delay-free equivalence system model*). The linear discrete systems (1) and the delay-free system model

$$z(k+1)=\overline{A}_0z(k)+\overline{B}_0u(k) \tag{4a}$$

$$y=\overline{C}_0z(k) \tag{4b}$$

are equivalent if there exists a unimodular matrix $T[\nabla]$ such that (3).

3. Necessary and sufficient conditions

According to [24], we expand the results of continuous time systems with only state delays and input delays (that in [24]) to that of discrete-time systems with multiple delays in state, input and output vectors. The following lemma shows the results.

Lemma 1. *The original system (1) is equivalent to the delay-free system (4), if and only if there exist $T_i\in\mathbb{R}^{n\times n}$ for $i=0,1,\dots,\tau_m$, with $\tau_c\leq\tau_m=\sup(\tau_a,(n-1)\tau_b)$, such that:*

(a) $\sum_{i=0}^{\kappa}T_iA_{\kappa-i}=\overline{A}_0T_\kappa$ *for* $\kappa=0,\dots,\tau_m+\tau_a$, *with* $\overline{A}_0=A_0$;
(b) $\sum_{i=0}^{\kappa}T_iB_{\kappa-i}=\overline{B}_\kappa$ *for* $\kappa=0,1,\dots,\tau_m+\tau_b$, *with* $\overline{B}_0=B_0;\overline{B}_\kappa=0$, $\forall\kappa>0$;
(c) $\sum_{i=0}^{\kappa}\overline{C}_{\kappa-i}T_i=C_\kappa$ *for* $\kappa=0,1,\dots,\tau_m+\tau_c$, *with* $\overline{C}_0=C_0$ *and* $\overline{C}_\kappa=0,\forall\kappa>0$;
(d) $det(\sum_{i=0}^{\tau_m}T_i\nabla^i)\in\mathbb{R}\backslash\{0\}$.

Proof. If Lemma 1 holds, the system (1) is equivalent to the system (4). Next, the equivalency is proved.