Neurocomputing 156 (2015) 166–175

Contents lists available at ScienceDirect

Neurocomputing

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

Nearly finite-horizon optimal control for a class of nonaffine time-delay nonlinear systems based on adaptive dynamic programming

Ruizhuo Song [a], Qinglai Wei [b,*], Qiuye Sun [c]

[a] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China
[b] The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[c] School of Information Science and Engineering, Northeastern University, Shenyang, Liaoning 110004, China

ARTICLE INFO

Article history:
Received 15 September 2014
Received in revised form
5 December 2014
Accepted 24 December 2014
Communicated by H. Zhang
Available online 5 January 2015

Keywords:
Adaptive dynamic programming
Approximate dynamic programming
Adaptive critic designs
Nonlinear systems
Optimal control
Time-delay

ABSTRACT

In this paper, a novel adaptive dynamic programming (ADP) algorithm is developed to solve the nearly optimal finite-horizon control problem for a class of deterministic nonaffine nonlinear time-delay systems. The idea is to use ADP technique to obtain the nearly optimal control which makes the optimal performance index function close to the greatest lower bound of all performance index functions within finite time. The proposed algorithm contains two cases with respective different initial iterations. In the first case, there exists control policy which makes arbitrary state of the system reach to zero in one time step. In the second case, there exists a control sequence which makes the system reach to zero in multiple time steps. The state updating is used to determine the optimal state. Convergence analysis of the performance index function is given. Furthermore, the relationship between the iteration steps and the length of the control sequence is presented. Two neural networks are used to approximate the performance index function and compute the optimal control policy for facilitating the implementation of ADP iteration algorithm. At last, two examples are used to demonstrate the effectiveness of the proposed ADP iteration algorithm.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Time-delay phenomenons are often encountered in physical and biological systems, and require special attention in engineering applications [1]. Transportation systems, communication systems, chemical processing systems, metallurgical processing systems and power systems are examples of time-delay systems. Delays may result in degradation in the control efficiency even instability of the control systems [2]. So there have been many works about systems with time delays in various research areas such as electrical, chemical engineering and networked control [3]. In the past few decades, the stabilization and the control of time-delay systems have always been the key focus in the control field [4–7]. Furthermore, there are many researchers who studied the controllability of linear time-delay systems [8–10]. They proposed some related theorems to judge the controllability of the linear time-delay systems. In addition, the optimal control problem is often encountered in industrial production. So the investigation of the optimal control for time-delay systems is significant. In [11] Chyung has pointed out the disadvantages of discrete time-delay system written as an extended system by increasing the dimension method to deal with the optimal control problem. So some direct methods for linear time-delay systems were presented in [11,12]. While for nonlinear time-delay system, due to the complexity of systems, the optimal control problem is rarely researched. This motivated our research interest.

As is well known, dynamic programming is very useful in solving the optimal control problems [13–15]. But it is often computationally untenable to run dynamic programming [16]. In the early 1970s, Werbos set up the basic strategy for ADP [17] to overcome the "curse of dimensionality" of dynamic programming. In [18], Werbos classified ADP approaches into four main schemes: heuristic dynamic programming (HDP), dual heuristic dynamic programming (DHP), action dependent heuristic dynamic programming (ADHDP), and action dependent dual heuristic dynamic programming (ADDHP). In recent years, ADP algorithms have made great progress [19–24]. In [25], an iteration ADP scheme with convergence proof was proposed for solving the optimal control problem of nonlinear discrete-time systems. In [26], an optimal tracking controller was proposed for a class of nonlinear discrete-time systems with time delays based on a novel HDP algorithm. In [27], a ADP-based optimal control is developed for complex-valued systems. Note that most of the results of the present study are about the infinite-horizon optimal control. The system cannot be really stabilized or tracked until the time

* Corresponding author.

http://dx.doi.org/10.1016/j.neucom.2014.12.066
0925-2312/© 2014 Elsevier B.V. All rights reserved.

reaches infinity. While for finite-horizon control problems, the system must be stabilized to zero or tracked to a desired trajectory within finite time. The controller design of finite-horizon problems still presents a challenge to control engineers as the lack of methodology and the control step is difficult to determine. Few results relate to the finite-horizon optimal control based on ADP algorithm. As we know that [28] solved the finite-horizon optimal control problem for a class of discrete-time nonlinear systems using ADP algorithm. But the method in [28] cannot be used in nonlinear time-delay systems. As the delay states in time-delay systems are coupling with each other. The state of current time k is decided by the states before k and the control law, while the control law is not known before it is obtained. So based on the research results in [28], we proposed a new ADP algorithm to solve the nearly finite-horizon optimal control problem for discrete time-delay systems through the framework of Hamilton–Jacobi–Bellman (HJB) equation.

In this paper the optimal controller is designed based on the original time-delay systems, directly. The state updating method is proposed to determine the optimal state of the time-delay system. For finite-horizon optimal control, the system can reach to zero when the final running step N is finite. But it is impossible in practice. So the results in this paper are in the sense of an error bound. The main contributions of this paper can be summarized as follows:

1. The finite-horizon optimal control for deterministic discrete nonaffine time-delay systems is studied based on the ADP algorithm.
2. The state updating is used to determine the optimal states of HJB equation.
3. The relationship between the iteration steps and the length of the control sequence is given.

This paper is organized as follows. In Section 2, the problem formulation is presented. In Section 3, the nearly finite-horizon optimal control scheme is developed based on the iteration ADP algorithm and the convergence proof is given. In Section 4, two examples are given to demonstrate the effectiveness of the proposed control scheme. In Section 5, the conclusion is drawn.

2. Problem statement

Consider a class of deterministic nonaffine time-delay nonlinear systems

$$x(t+1)=F(x(t),x(t-h_1),x(t-h_2),\ldots,x(t-h_l),u(t)),$$
$$x(t)=\chi(t),\ \ -h_l\le t\le 0 \tag{1}$$

where $x(t)\in\Re^n$ is the state and $x(t-h_1),\ldots,x(t-h_l)\in\Re^n$ are time-delay states. $u(t)\in\Re^m$ is the system input. $\chi(t)$ is the initial state, h_i, $i=1,2,\ldots,l$, is the time delay, set $0<h_1<h_2<\ldots<h_l$, and they are nonnegative integer numbers. $F(x(t),x(t-h_1),x(t-h_2),\ldots,x(t-h_l),u(t))$ is the known function. $F(0,0,\ldots,0)=0$.

For any time step k, the performance index function for state $x(k)$ under the control sequence $U(k,N+k-1)=(u(k),u(k+1),\ldots,u(N+k-1))$ is defined as

$$J(x(k),U(k,N+k-1))=\sum_{j=k}^{N+k-1}\{x^T(j)Qx(j)+u^T(j)Ru(j)\}, \tag{2}$$

where Q and R are positive definite constant matrixes.

In this paper, we focus on solving the nearly finite-horizon optimal control problem for system (1). The feedback control $u(k)$ must not only stabilize the system within finite time steps but also guarantee the performance index function (2) to be finite. So the control sequence $U(k,N+k-1)=(u(k),u(k+1),\ldots,u(N+k-1))$ must be admissible.

Definition 1. N time steps control sequence: for any time step k, we define the N time steps control sequence $U(k,N+k-1)=(u(k),u(k+1),\ldots,u(N+k-1))$. The length of $U(k,N+k-1)$ is N.

Definition 2. Final state: we define final state $x_f=x_f(x(k),U(k,N+k-1))$, i.e., $x_f=x(N+k)$.

Definition 3. Admissible control sequence: an N time steps control sequence is said to be admissible for $x(k)$, if the final state $x_f(x(k),U(k,N+k-1))=0$ and $J(x(k),U(k,N+k-1))$ is finite.

Remark 1. Definitions 1 and 2 are used to state conveniently the admissible control sequence, i.e. Definition 3, which is necessary for the theorems of this paper.

Remark 2. It is important to point out that the length of control sequence N cannot be designated in advance. It is calculated by the proposed algorithm. If we calculate that the length of optimal control sequence is L at time step k, then we consider that the optimal control sequence length at time step k is $N=L$.

According to the theory of dynamics programming [29], the optimal performance index function is defined as

$$J^*(x(k))=\inf_{U(k,N+k-1)}J(x(k),U(k,N+k-1)) \tag{3}$$

$$J^*(x(k))=\inf_{u(k)}\{x^T(k)Qx(k)+u^T(k)Ru(k)+J^*(x(k+1))\}, \tag{4}$$

and the optimal control policy is

$$u^*(k)=\arg\inf_{u(k)}\{x^T(k)Qx(k)+u^T(k)Ru(k)+J^*(x(k+1))\}, \tag{5}$$

so the state under the optimal control policy is

$$x^*(t+1)=F(x^*(t),x^*(t-h_1),\ldots,x^*(t-h_l),u^*(t)),\quad t=0,1,\ldots,k,\ldots, \tag{6}$$

and then, the HJB equation is written as

$$\begin{aligned}J^*(x^*(k))&=J(x^*(k),U^*(k,N+k-1))\\&=(x^*(k))^TQx^*(k)+(u^*(k))^TRu^*(k)+J^*(x^*(k+1)).\end{aligned} \tag{7}$$

Remark 3. From Remark 2, we can see that the length N of the optimal control sequence is unknown finite number and cannot be designated in advance. So we can say that if at time step k, the length of the optimal control sequence is N, then at time step $k+1$, the length of the optimal control sequence is $N-1$. Therefore, the HJB equation (7) is established.

In the following, we will give an explanation about the validity of Eq. (3). First, we define $U^*(k,N+k-1)=(u^*(k),u^*(k+1),\ldots,u^*(N+k-1))$, i.e.

$$U^*(k,N+k-1)=\arg\inf_{U(k,N+k-1)}J(x(k),U(k,N+k-1)). \tag{8}$$

Then we have

$$\begin{aligned}J^*(x(k))&=\inf_{U(k,N+k-1)}J(x(k),U(k,N+k-1))\\&=J(x(k),U^*(k,N+k-1)).\end{aligned} \tag{9}$$

Then according to (2), we can get

$$\begin{aligned}J^*(x(k))&=\sum_{j=k}^{N+k-1}\{x^T(j)Qx(j)+(u^*(j))^TRu^*(j)\}\\&=x^T(k)Qx(k)+(u^*(k))^TRu^*(k)+\cdots+x^T(N+k-1)Qx(N+k-1)\\&\quad+(u^*(N+k-1))^TRu^*(N+k-1).\end{aligned} \tag{10}$$

Eq. (10) can be written as

$$\begin{aligned}J^*(x(k))&=x^T(k)Qx(k)+(u^*(k))^TRu^*(k)+\cdots+x^T(N+k-2)Qx(N+k-2)\\&\quad+(u^*(N+k-2))^TRu^*(N+k-2)\\&\quad+\inf_{u(N+k-1)}\{x^T(N+k-1)Qx(N+k-1)\\&\quad+u^T(N+k-1)Ru(N+k-1)\}.\end{aligned} \tag{11}$$

Download English Version:

https://daneshyari.com/en/article/406248

Download Persian Version:

https://daneshyari.com/article/406248

Daneshyari.com