Neurocomputing 78 (2012) 14–22

Contents lists available at SciVerse ScienceDirect

Neurocomputing

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

Finite-horizon neuro-optimal tracking control for a class of discrete-time nonlinear systems using adaptive dynamic programming approach ☆

Ding Wang [a], Derong Liu [a,b,*], Qinglai Wei [a]

[a] State Key Laboratory of Intelligent Control and Management of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, PR China
[b] Department of Electrical and Computer Engineering, University of Illinois, Chicago, IL 60607, USA

ARTICLE INFO

Available online 3 September 2011

Keywords:
Adaptive critic designs
Adaptive dynamic programming
Approximate dynamic programming
Finite-horizon optimal tracking control
Learning control
Neural networks
Reinforcement learning

ABSTRACT

In this paper, a finite-horizon neuro-optimal tracking control strategy for a class of discrete-time nonlinear systems is proposed. Through system transformation, the optimal tracking problem is converted into designing a finite-horizon optimal regulator for the tracking error dynamics. Then, with convergence analysis in terms of cost function and control law, the iterative adaptive dynamic programming (ADP) algorithm via heuristic dynamic programming (HDP) technique is introduced to obtain the finite-horizon optimal tracking controller which makes the cost function close to its optimal value within an ε-error bound. Three neural networks are used as parametric structures to implement the algorithm, which aims at approximating the cost function, the control law, and the error dynamics, respectively. Two simulation examples are included to complement the theoretical discussions.

1. Introduction

It is well known that the optimal tracking control problem has been the focus of control systems community for several decades since it is usually encountered in real world systems [1–3]. In the case of infinite-horizon optimal tracking control, the system will not be tracked until the time reaches infinity, while for the finite case, the system must be tracked to a reference trajectory in a finite duration of time. Since many limitations exist in traditional optimal tracking control approaches, such as plant inversion [2] and linearization [3], it is necessary to design direct optimal tracking control schemes for nonlinear systems. In this paper, we will study how to solve this problem through the framework of Hamilton–Jacobi–Bellman (HJB) [4] equation from optimal control theory. Unlike the open-loop optimal controller design for nonlinear systems, however, for closed-loop optimal feedback control, it is difficult to solve directly the time-varying HJB equation which involves solving either nonlinear partial difference or differential equations. Though dynamic programming (DP) has been an useful computational technique in solving optimal control problems for many years, it is often computationally untenable to run it to obtain the optimal solution due to the "curse of dimensionality" [5].

As Poggio and Girosi [6] stated, the problem of learning between input and output spaces is equivalent to that of synthesizing an associative memory that retrieves appropriate output when the input is present and generalizes when a new input is applied. With strong capabilities of self-learning and adaptivity, artificial neural networks (ANN or NN) are an effective tool for implementing intelligent control [7–10]. Besides, it has been used for universal function approximation in adaptive/approximate dynamic programming (ADP) algorithms, which were proposed in [9–11] as a method for solving optimal control problems forward-in-time. There are several synonyms used for ADP including "adaptive dynamic programming" [12–14], "approximate dynamic programming" [9,15,16], "neuro-dynamic programming" [17], "neural dynamic programming" (NDP) [18], "adaptive critic designs" [19], and "reinforcement learning" [15,20]. As an effective intelligent control method, ADP and the related research have gained much attention from researchers [9–19,21–35]. Very good surveys were given in Wang et al. [13], Lewis and Vrabie [14], and Balakrishnan et al. [25]. According to [9,19], ADP approaches were classified into several main schemes: heuristic dynamic programming (HDP), action-dependent HDP (ADHDP), also known as Q-learning [20], dual heuristic dynamic programming (DHP), ADDHP, globalized DHP (GDHP), and ADGDHP. Al-Tamimi et al. [16] proposed a greedy HDP algorithm to solve the discrete-time HJB (DTHJB) equation for optimal control of nonlinear systems. Wang et al. [23] developed an ε-ADP algorithm for studying finite-horizon optimal control of discrete-time nonlinear systems.

☆ This work was supported in part by the NSFC under Grants 60874043, 60904037, 60921061, and 61034002, by Beijing Natural Science Foundation under Grant 4102061, and by the NSF under Grant ECCS-1027602.

* Corresponding author at: State Key Laboratory of Intelligent Control and Management of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, PR China. Tel.: +86 10 62557379, +1 312 355 4475; fax: +86 10 62650912, +1 312 966 6465.
E-mail addresses: ding.wang@ia.ac.cn (D. Wang), derong.liu@ia.ac.cn, dliu@ece.uic.edu (D. Liu), qinglai.wei@ia.ac.cn (Q. Wei).

0925-2312/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2011.03.058

With the rapid development of NN technology and recently, the ADP method, various new strategies were devised to deal with the optimal tracking control problems. Park et al. [36] used the multilayer NN to design an optimal tracking neuro-controller for discrete-time nonlinear systems with quadratic cost function. Zhang et al. [32] gave a novel infinite-horizon optimal tracking control scheme for discrete-time nonlinear systems via greedy HDP algorithm. Dierks and Jagannathan [31] utilized the NDP technique to solve the HJB equation forward-in-time for optimal tracking control of affine nonlinear systems. However, to the best of our knowledge, there is still no result to solve the finite-horizon optimal tracking control problem for discrete-time nonlinear systems based on iterative ADP algorithm via HDP technique (iterative HDP algorithm for brief). In this paper, for the first time, we will provide an iterative ADP algorithm to design finite-horizon near-optimal tracking controller for a class of discrete-time nonlinear systems.

The rest of this paper is organized as follows. In Section 2, we present the problem statement, transform the finite-horizon optimal tracking control problem into an optimal regulation problem, and introduce the DTHJB equation for nonlinear systems. Section 3 starts by deriving the iterative ADP algorithm with convergence analysis, and then the finite-horizon optimal tracking control scheme is proposed which makes the cost function close to its optimal value within an ε-error bound. In Section 4, the NN implementation of the iterative ADP algorithm is presented. In Section 5, two examples are given to substantiate the theoretical results. Section 6 contains concluding remarks.

2. Problem statement

Consider the discrete-time nonlinear systems given by

$$x_{k+1} = f(x_k) + g(x_k)u_p(x_k), \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state, $u_p(x_k) \in \mathbb{R}^m$ is the control vector, $f(\cdot)$ and $g(\cdot)$ are differentiable in their argument with $f(0)=0$. Assume that $f+gu_p$ is Lipschitz continuous on a set Ω in $\mathbb{R}^n$ containing the origin, and that the system (1) is controllable in the sense that there exists a continuous control on Ω that asymptotically stabilizes the system. In the following part, $u_p(x_k)$ is denoted by u_{pk} for simplicity.

The objective for optimal tracking control problem is to determine optimal control law u_p^*, so as to make the nonlinear system (1) to track a reference (or desired) trajectory r_k in an optimal manner. Here, we assume that the reference trajectory r_k satisfies

$$r_{k+1} = \phi(r_k), \tag{2}$$

where $r_k \in \mathbb{R}^n$ and $\phi(r_k) \in \mathbb{R}^n$. Then, we define the tracking error as

$$e_k = x_k - r_k. \tag{3}$$

Inspired by the work of [31,32,36], we define the steady control corresponding to the reference trajectory r_k as

$$u_{dk} = g^{-1}(r_k)(\phi(r_k) - f(r_k)), \tag{4}$$

where $g^{-1}(r_k)g(r_k) = I_m$ and I_m is an $m \times m$ identity matrix.

By denoting

$$u_k = u_{pk} - u_{dk} \tag{5}$$

and using (1)–(4), we obtain

$$\begin{cases} e_{k+1} = f(e_k+r_k) + g(e_k+r_k)g^{-1}(r_k)(\phi(r_k), \\ \qquad -f(r_k)) - \phi(r_k) + g(e_k+r_k)u_k \\ r_{k+1} = \phi(r_k) \end{cases} \tag{6}$$

as the new system. Note that in system (6), e_k and r_k are regarded as the system variables while u_k is seen as system input. The second equation of system (6) only gives the evolution of the reference trajectory which is not affected by the system input. Therefore, for simplicity, (6) can be rewritten as

$$e_{k+1} = F(e_k, u_k). \tag{7}$$

Now, let e_0 be an initial state of system (7) and define $\underline{u}_0^{N-1} = (u_0, u_1, \ldots, u_{N-1})$ be a control sequence with which the system (7) gives a trajectory starting from e_0: $e_1, e_2, \ldots, e_N$. We call the number of elements in the control sequence $\underline{u}_0^{N-1}$ the length of $\underline{u}_0^{N-1}$ and denote it as $|\underline{u}_0^{N-1}|$. Then, $|\underline{u}_0^{N-1}| = N$. The final state under the control sequence $\underline{u}_0^{N-1}$ is denoted as $e^{(f)}(e_0, \underline{u}_0^{N-1}) = e_N$.

Definition 1. A nonlinear dynamical system is said to be stabilizable on a compact set $\Omega \in \mathbb{R}^n$, if for all initial conditions $e_0 \in \Omega$, there exists a control sequence $\underline{u}_0^{N-1} = (u_0, u_1, \ldots, u_{N-1})$, $u_i \in \mathbb{R}^m$, $i = 0, 1, \ldots, N-1$, such that the state $e^{(f)}(e_0, \underline{u}_0^{N-1}) = 0$.

Let $\underline{u}_k^{N-1} = (u_k, u_{k+1}, \ldots, u_{N-1})$ be the control sequence starting at k with length $N-k$. For finite-horizon optimal tracking control problem, it is desired to find the control sequence which minimizes the following cost function:

$$J(e_k, \underline{u}_k^{N-1}) = \sum_{i=k}^{N-1} U(e_i, u_i), \tag{8}$$

where U is the utility function, $U(0,0)=0$, $U(e_i,u_i) \geq 0$ for $\forall e_i, u_i$. In this paper, the utility function is chosen as the quadratic form as follows:

$$U(e_i, u_i) = e_i^T Q e_i + u_i^T R u_i.$$

This quadratic cost function can not only force the system state to follow the reference trajectory, but also force the system input to be close to the steady value in maintaining the state to its reference value. In fact, it can also be expressed as

$$U(e_i, u_i) = [e_i^T \ r_i^T] \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} e_i \\ r_i \end{bmatrix} + u_i^T R u_i,$$

when considered from the angle of system (6).

In this sense, the nonlinear tracking problem is converted into a regulation problem and the finite-horizon cost function for tracking is written in terms of e_k and u_k. Then, the problem of solving the finite-horizon optimal tracking control law u_p^* for system (1) is transformed into seeking the finite-horizon optimal control law u^* for system (7) with respect to (8). As a result, we will focus on how to design u^* in the following sections.

For finite-horizon optimal control problems, the designed feedback control must be finite-horizon admissible, which means it must not only stabilize the controlled system on Ω within finite number of time steps but also guarantee the cost function to be finite.

Definition 2. A control sequence $\underline{u}_k^{N-1}$ is said to be finite- horizon admissible for a state $e_k \in \mathbb{R}^n$ with respect to (8) on Ω if $\underline{u}_k^{N-1}$ is continuous on a compact set $\Omega_u \in \mathbb{R}^m$, $u(0)=0$, $e^{(f)}(e_k, \underline{u}_k^{N-1}) = 0$ and $J(e_k, \underline{u}_k^{N-1})$ is finite.

Let

$$\mathfrak{A}_{e_k} = \{\underline{u}_k : e^{(f)}(e_k, \underline{u}_k) = 0\}$$

be the set of all finite-horizon admissible control sequences of e_k and

$$\mathfrak{A}_{e_k}^{(i)} = \{\underline{u}_k^{k+i-1} : e^{(f)}(e_k, \underline{u}_k^{k+i-1}) = 0, |\underline{u}_k^{k+i-1}| = i\}$$

be the set of all finite-horizon admissible control sequences of e_k with length i. Define the optimal cost function as

$$J^*(e_k) = \inf_{\underline{u}_k} \{J(e_k, \underline{u}_k) : \underline{u}_k \in \mathfrak{A}_{e_k}\}. \tag{9}$$