Neurocomputing 149 (2015) 106–115

Contents lists available at ScienceDirect

Neurocomputing

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

Neural-network-based adaptive optimal tracking control scheme for discrete-time nonlinear systems with approximation errors

Qinglai Wei *, Derong Liu

The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 15 June 2013
Received in revised form
30 August 2013
Accepted 17 September 2013
Available online 1 August 2014

Keywords:
Adaptive dynamic programming
Adaptive critic designs
Approximate dynamic programming
Value iteration
Approximation errors
Optimal tracking control

ABSTRACT

In this paper, a new infinite horizon neural-network-based adaptive optimal tracking control scheme for discrete-time nonlinear systems is developed. The idea is to use iterative adaptive dynamic programming (ADP) algorithm to obtain the iterative tracking control law which makes the iterative performance index function reach the optimum. When the iterative tracking control law and iterative performance index function in each iteration cannot be accurately obtained, the convergence criteria of the iterative ADP algorithm are established according to the properties with finite approximation errors. If the convergence conditions are satisfied, it shows that the iterative performance index functions can converge to a finite neighborhood of the lowest bound of all performance index functions. Properties of the finite approximation errors for the iterative ADP algorithm are also analyzed. Neural networks are used to approximate the performance index function and compute the optimal control policy, respectively, for facilitating the implementation of the iterative ADP algorithm. Convergence properties of the neural network weights are proven. Finally, simulation results are given to illustrate the performance of the developed method.

1. Introduction

Optimal tracking control problems of nonlinear systems have always been the key focus in the control field in the latest several decades. Traditional optimal tracking control is mostly implemented by feedback linearization [1]. However, the controller designed by feedback linearization technique is only effective in the neighborhood of the equilibrium point. When the required operating range is large, the nonlinearities in the system cannot be properly compensated by using a linear model. Hence the control performance of feedback linearization technique is usually unsatisfied and the nonlinear controller design of the optimal tracking control is necessary. The difficulty for nonlinear optimal feedback control lies in solving the time-varying HJB equation which is usually too difficult to solve analytically. To overcome the difficulty, many approximation methods are proposed to obtain optimal tracking control law [2–5]. Among these approximate approaches, adaptive dynamic programming (ADP) algorithm, proposed by Werbos [6,7], has played an important role in seeking approximate solutions of dynamic programming problems as a way to solve the computational issue forward-in-time [8–14]. There are several synonyms used for ADP including "adaptive critic designs" [15], "adaptive dynamic programming" [16,17], "approximate dynamic programming" [18,19], "neural dynamic programming" [20], "neuro-dynamic programming" [21], and "reinforcement learning" [22]. In Werbos [19], ADP approaches were classified into four main schemes: Heuristic Dynamic Programming (HDP), Dual Heuristic Programming (DHP), Action Dependent HDP (ADHDP) (also known as Q-learning [23]), and Action Dependent DHP (ADDHP). In [15], two more ADP that are Globalized-DHP (GDHP) and ADGDHP were proposed. Iterative methods are also used in ADP to obtain the solution of HJB equation indirectly and have received lots of attentions [24–32].There are two main iterative ADP algorithms which are based on policy iteration and value iteration [33].

Policy iteration algorithm for optimal control of continuous-time systems with continuous state and action spaces was given in [34]. In [16], Murray et al. studied the deterministic continuous-time stabilizable systems where an iterative process was proposed to find the optimal control law by starting from an arbitrary admissible control law. In the policy iteration algorithms of ADP, to obtain the iterative performance index functions and iterative control laws, an initial admissible control law of the system is required. But, unfortunately, the admissible control law for nonlinear systems is also difficult to obtain. Thus, the initial conditions for the controller greatly limit the applications of the policy

* Corresponding author.
E-mail addresses: qinglai.wei@ia.ac.cn (Q. Wei), derong.liu@ia.ac.cn (D. Liu).

http://dx.doi.org/10.1016/j.neucom.2013.09.069
0925-2312/

iteration algorithms. Value iteration algorithm of optimal control for discrete-time nonlinear systems was given in [35]. In [18], Al-Tamimi et al. studied the deterministic discrete-time affine nonlinear systems

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

where x_k is the system state and u_k is the system control. Functions $f(x_k)$ and $g(x_k)$ denote system functions. In [18], the performance index function is defined by

$$J(x_k)=\sum_{j=k}^{\infty}(x_j^TQx_j+u_j^TRu_j), \tag{2}$$

where $Q\in\mathbb{R}^{n\times n}$ and $R\in\mathbb{R}^{m\times m}$ are positive definite matrices. In [18], a value iteration algorithm, which was referred to as HDP, was proposed for finding the optimal control law. It starts from $V_0(x_k)\equiv 0$, and then the iteration

$$\begin{cases} u_i(x_k)=-\dfrac{1}{2}R^{-1}g(x_k)^T\dfrac{\partial V_i(x_{k+1})}{\partial x_{k+1}}, \\ V_{i+1}(x_k)=x_k^TQx_k+u_i^T(x_k)Ru_i(x_k)+V_i(x_{k+1}), \end{cases} \tag{3}$$

is introduced for $i=0,1,2,\ldots$, where $x_{k+1}=f(x_k)+g(x_k)u_i(x_k)$. It was proven that $V_i(x_k)$ is nondecreasing and upper bounded, and hence converges to $J^*(x_k)$ as i increases to infinity. In 2008, Zhang et al. [36] applied value iteration of ADP to solve optimal tracking problems for nonlinear systems. Liu et al. [37] realized the value iteration of ADP by GDHP. Although iterative ADP algorithms attract more and more attentions [38–47], for most of the iterative ADP algorithms, the iterative control of each iteration is required to be accurately obtained. These iterative ADP algorithms can be called "accurate iterative ADP algorithms".

For most real-world control systems, however, the accurate iterative control laws in the iterative ADP algorithms cannot be obtained. As approximation structures are used to achieve the optimal control law and the performance index function, there must exist approximation errors between the approximation functions and the expected ones. This shows that the convergence properties in the accurate iterative ADP algorithms may be invalid for the iterative ADP with approximation errors. Till now, the discussion on the convergence properties of the iterative ADP algorithms with approximation errors is very little. Only in [41], based on iterative θ-ADP algorithm, an optimal regulation control scheme with approximation errors was proposed for discrete-time nonlinear systems, while in [41], the convergence of neural network weights is not analyzed. To the best of our knowledge, there are no discussions on the ADP algorithm for optimal tracking control problems with approximation errors.

In this paper, we will develop a new iterative ADP scheme for infinite horizon optimal tracking control problems. The main contribution of this paper is that the optimal tracking control problems with finite approximation errors are solved effectively using the present iterative ADP algorithms. A convergence analysis of the performance index function is developed and the least upper bound of the converged iterative performance index function is also presented. The convergence criteria are obtained. In order to facilitate the implementation of the iterative ADP algorithms, we use neural networks to obtain the iterative performance index function and the optimal tracking control policy, respectively. The convergence properties of the neural network weights are proven to guarantee the effectiveness of the neural network applications. Finally, simulation results are given to show the effectiveness of the developed iterative ADP algorithm.

The rest of this paper is organized as follows. In Section 2, the problem formulation is presented. In Section 3, the iterative ADP algorithm for the optimal tracking control problem is derived. The convergence criteria for the iterative ADP algorithm is also analyzed in this section. In Section 4, the neural network implementation with convergence proof for the optimal control scheme is discussed. In Section 5, numerical results and analysis are presented to demonstrate the effectiveness of the developed optimal control scheme. Finally, in Section 6, the conclusion is drawn and our future work will be declared.

2. Problem formulation

Consider a class of affine nonlinear systems of the form:

$$x(k+1)=f(x(k))+g(x(k))u(x(k)) \tag{4}$$

where $x(k)\in\Re^n, f(x(k))\in\Re^n, g(x(k))\in\Re^{n\times m}$, the input $u(k)\in\Re^m$ and $g(\cdot)$ has a generalized inverse. Here, assume that the system is controllable on $\Omega\subset\Re^n$. For infinite-time optimal tracking problem, the control objective is to design optimal control $u(x(k))$ for system (4) such that the state $x(k)$ track the specified desired trajectory $\eta(k)\in\Re^n$, $k=0,1,\ldots$. Define the tracking error as

$$z(k)=x(k)-\eta(k). \tag{5}$$

Define the following quadratic performance index:

$$J(z(0),\underline{u}_0)=\sum_{k=0}^{\infty}\{z^T(k)Qz(k)+(u(k)-u_e(k))^TR(u(k)-u_e(k))\} \tag{6}$$

where $Q\in\Re^{n\times n}$ and $R\in\Re^{m\times m}$ are positive definite matrices and $\underline{u}_0=(u(0), u(1),\ldots)$. Let

$$U(z(k),v(k))=z^T(k)Qz(k)+v^T(k)Rv(k)$$

be the utility function, where $v(k)=u(k)-u_e(k)$. Let $u_e(k)$ denote the expected control introduced for analytical purpose, which can be given as

$$u_e(k)=g^{-1}(\eta(k))(\eta(k+1)-f(\eta(k))) \tag{7}$$

where $g^{-1}(\eta(k))g(\eta(k))=I$ and $I\in\Re^{m\times m}$ is the identity matrix. Combining (4) and (5), we can get

$$\begin{aligned} z(k+1)&=f(z(k)+\eta(k))-\eta(k+1)+g(z(k)+\eta(k)) \\ &\quad\times(v(k)+g^{-1}(\eta(k))(f(\eta(k))-\eta(k+1))). \end{aligned} \tag{8}$$

We will study optimal tracking control problems for (4). The goal of this paper is to find an optimal tracking control scheme which tracks the desired trajectory $\eta(k)$ and simultaneously minimizes the performance index function (6). The optimal performance index function is defined as

$$J^*(z(k))=\inf_{\underline{v}_k}\{J(z(k),\underline{v}_k)\}, \tag{9}$$

where $\underline{v}_k=(v(k),v(k+1),\ldots)$. According to Bellman's principle of optimality, $J^*(z(k))$ satisfies the discrete-time HJB equation:

$$J^*(z(k))=\inf_{v(k)}\{U(z(k),v(k))+J^*(F(z(k),v(k)))\}. \tag{10}$$

Then, the law of optimal single control vector can be expressed as

$$v^*(z(k))=\arg\inf_{v(k)}\{U(z(k),v(k))+J^*(z(k+1))\}. \tag{11}$$

Hence, the HJB equation (10) can be written as

$$J^*(z(k))=U(z(k),v^*(z(k)))+J^*(z(k+1)). \tag{12}$$

In [36], based on the greedy HDP iteration technique, the performance index and control policy are updated by iterations, with the iteration number i increasing from 0 to ∞. First, the initial performance index $V_0(z(k))\equiv 0$. Then, for $i=0,1,\ldots$, the control $v_i(k)$ and $V_{i+1}(z(k))$ are computed by the following two equations:

$$v_i(k)=\arg\min_{v(k)}\{z^T(k)Qz(k)+v^T(k)Rv(k)+V_i(z(k+1))\} \tag{13}$$