



An iterative adaptive dynamic programming algorithm for optimal control of unknown discrete-time nonlinear systems with constrained inputs

Derong Liu*, Ding Wang, Xiong Yang

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

ARTICLE INFO

Article history:

Received 2 November 2011
 Received in revised form 3 April 2012
 Accepted 15 July 2012
 Available online 31 July 2012

Keywords:

Adaptive dynamic programming
 Approximate dynamic programming
 Control constraints
 Globalized dual heuristic programming
 Neural networks
 Optimal control

ABSTRACT

In this paper, the adaptive dynamic programming (ADP) approach is employed for designing an optimal controller of unknown discrete-time nonlinear systems with control constraints. A neural network is constructed for identifying the unknown dynamical system with stability proof. Then, the iterative ADP algorithm is developed to solve the optimal control problem with convergence analysis. Two other neural networks are introduced for approximating the cost function and its derivatives and the control law, under the framework of globalized dual heuristic programming technique. Furthermore, two simulation examples are included to verify the theoretical results.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

The nonlinear optimal control has been the focus of control fields for many decades [8,16]. It often needs to solve the nonlinear Hamilton–Jacobi–Bellman (HJB) equation. For instance, the discrete-time HJB (DTHJB) equation is more difficult to work with than the Riccati equation because it involves solving nonlinear partial difference equations. Although dynamic programming has been a useful technique in handling optimal control problems for many years, it is often computationally untenable to perform it to obtain the optimal solutions [4].

Effective techniques have been employed to construct learning systems [22,20,37,19,35,3,12,11]. Characterized by strong abilities of self-learning and adaptivity, artificial neural networks (ANN or NN) are also a functional tool to implement learning control [33,15,13,34]. Additionally, they are often used to carry out universal function approximation in adaptive/approximate dynamic programming (ADP) algorithms. The ADP method was proposed by Werbos [33,34] to deal with optimal control problems forward-in-time. There were several synonyms used for ADP, including “adaptive critic designs” [21], “adaptive dynamic programming” [30,17], “approximate dynamic programming” [34,24,2], “neuro-dynamic programming” [5], “neural dynamic programming” [23], and “reinforcement learning” [6].

In recent years, ADP and related research have gained much attention from researchers [1,2,5,6,9,10,14,17,18,21,23–32,34,36]. According to [21] and [34], ADP approaches were classified into several main schemes: heuristic dynamic programming (HDP), action-dependent HDP (ADHDP), also known as Q-learning, dual heuristic dynamic programming (DHP), ADDHP, globalized DHP (GDHP), and ADGDHP. Al-Tamimi et al. [2] proposed a greedy HDP iteration algorithm to

* Corresponding author. Tel.: +86 10 62557379; fax: +86 10 62650912.

E-mail addresses: derong.liu@ia.ac.cn (D. Liu), ding.wang@ia.ac.cn (D. Wang), xiong.yang@ia.ac.cn (X. Yang).

solve the DTHJB equation of optimal control of discrete-time affine nonlinear systems. Abu-Khalaf and Lewis [1], Vrabie and Lewis [27], and Vamvoudakis and Lewis [25] investigated the continuous-time nonlinear optimal control problems based on the idea of ADP.

With the increasing complexity of industry processes, the data-based method has achieved great interest among control engineers. It does not need to build accurate mathematical models of controlled plants and thus has significant practical value. Kim and Lewis [14] presented a model-free H_∞ control design scheme for unknown linear discrete-time systems via Q-learning, which was expressed in the form of linear matrix inequality. Campi and Savaresi [7] proposed a virtual reference feedback tuning approach which was in fact a data-based method. In this paper, we solve the constrained optimal control problem of unknown discrete-time nonlinear systems based on the iterative ADP algorithm via GDHP technique (i.e., iterative GDHP algorithm). An NN model is constructed as an identifier to learn the unknown controlled plant. Then, the iterative ADP algorithm is introduced to solve the DTHJB equation with convergence proof. Next, the optimal controller can be designed by employing the GDHP technique.

This paper is organized as follows: In Section 2, the optimal control problem and the DTHJB equation are recalled for discrete-time nonlinear systems. In Section 3, we first design an NN identifier for unknown controlled system with stability proof. Then, the optimal control scheme based on the iterative ADP algorithm is developed with convergence analysis. In Section 4, the implementation of iterative ADP algorithm is presented through NN-based GDHP technique. In Section 5, two numerical examples are given to demonstrate the effectiveness of the proposed optimal control scheme. In Section 6, concluding remarks are given.

2. Preliminaries

In this paper, we study the nonlinear discrete-time systems described by

$$x_{k+1} = F(x_k, u_k), \quad k = 0, 1, 2, \dots \quad (1)$$

where $x_k \in \mathbb{R}^n$ is the state vector and $u_k = u(x_k) \in \mathbb{R}^m$ is the control vector. Let the initial state be denoted by x_0 . The system function $F(x_k, u_k)$ is continuous for $\forall x_k, u_k$ and $F(0,0) = 0$. Hence, $x = 0$ is an equilibrium state of system (1) under control $u = 0$. We define $\Omega_u = \{u_k | u_k = [u_{1k}, u_{2k}, \dots, u_{mk}]^T \in \mathbb{R}^m, |u_{ik}| \leq \bar{u}_i, i = 1, 2, \dots, m\}$, where \bar{u}_i is the saturating bound for the i th actuator. Let $\bar{U} = \text{diag}\{\bar{u}_1, \bar{u}_2, \dots, \bar{u}_m\}$ be a constant diagonal matrix.

The objective for general optimal control problems is to find the control law $u(x)$ which minimizes the infinite horizon cost function given by

$$J(x_k) = \sum_{i=k}^{\infty} U(x_i, u_i),$$

where U is the utility function, $U(0, 0) = 0$, and $U(x_i, u_i) \geq 0$ for $\forall x_i, u_i$. According to Bellman's optimality principle, the optimal cost function

$$J^*(x_k) = \min_{u_k, u_{k+1}, \dots, u_\infty} \sum_{i=k}^{\infty} U(x_i, u_i)$$

can be rewritten as

$$J^*(x_k) = \min_{u_k} \left\{ U(x_k, u_k) + \min_{u_{k+1}, \dots, u_\infty} \sum_{i=k+1}^{\infty} U(x_i, u_i) \right\}.$$

In other words, $J^*(x_k)$ satisfies the DTHJB equation

$$J^*(x_k) = \min_{u_k} \{ U(x_k, u_k) + J^*(x_{k+1}) \}. \quad (2)$$

The corresponding optimal control u^* is

$$u^*(x_k) = \arg \min_{u_k} \{ U(x_k, u_k) + J^*(x_{k+1}) \}. \quad (3)$$

In many literatures [2,9,28], the utility function is chosen as

$$U(x_i, u_i) = x_i^T Q x_i + u_i^T R u_i, \quad (4)$$

where Q and R are positive definite matrices with suitable dimensions. However, when dealing with constrained optimal control problems, it is not the case any more. Inspired by the work of [1,36], we can employ a generalized non-quadratic functional

$$Y(u_i) = 2 \int_0^{u_i} \psi^{-T}(\bar{U}^{-1}s) \bar{U} R ds \quad (5)$$

Download English Version:

<https://daneshyari.com/en/article/393971>

Download Persian Version:

<https://daneshyari.com/article/393971>

[Daneshyari.com](https://daneshyari.com)