Acta Automatica Sinica
Vol. 36, No. 1 January, 2010

Available online at www.sciencedirect.com
ScienceDirect

An Optimal Control Scheme for a Class of Discrete-time Nonlinear Systems with Time Delays Using Adaptive Dynamic Programming

WEI Qing-Lai[1] ZHANG Hua-Guang[2] LIU De-Rong[1] ZHAO Yan[3]

Abstract In this paper, an optimal control scheme for a class of nonlinear systems with time delays in both state and control variables with respect to a quadratic performance index function is proposed using a new iterative adaptive dynamic programming (ADP) algorithm. By introducing a delay matrix function, the explicit expression of the optimal control is obtained using the dynamic programming theory and the optimal control can iteratively be obtained using the adaptive critic technique. Convergence analysis is presented to prove that the performance index function can reach the optimum by the proposed method. Neural networks are used to approximate the performance index function, compute the optimal control policy, solve delay matrix function, and model the nonlinear system, respectively, for facilitating the implementation of the iterative ADP algorithm. Two examples are given to demonstrate the validity of the proposed optimal control scheme.

Key words Adaptive dynamic programming (ADP), approximate dynamic programming, time delay, optimal control, nonlinear system, neural networks

The optimal control problem of nonlinear systems has always been a key focus in the control field in the last several decades. Coupled with this is the fact that nothing can happen instantaneously, as is so often presumed in many mathematical models. So strictly speaking, time delays exist in the most practical control systems. Time delays may result in degradation in the control efficiency even instability of the control systems. So there have been many studies on the control systems with time delay in various research fields, such as electrical, chemical engineering, and networked control[1−2]. The optimal control problem for the time-delay systems always attracts considerable attention of the researchers and many results have been obtained[3−5]. In general, the optimal control for the time-delay systems is an infinite-dimensional control problem[3], which is very difficult to solve. So many analysis and applications are limited to a very simple case: the linear systems with only state delays[6]. For nonlinear case with state delays, the traditional method is to adopt fuzzy method and robust method, which transforms the nonlinear time-delay systems to linear systems[7]. For systems with time delays both in states and controls, it is still an open problem[4−5]. The main difficulty lies in the formulation of the optimal controller which must use the information of the delayed control term so as to obtain an efficient control. This makes the analysis of the system much more difficult, and there is no method strictly facing this problem even in the linear cases. This motivates our research.

Adaptive dynamic programming (ADP) is a powerful tool in solving optimal control problems[8−9] and has attached considerable attention from many researchers in recent years, such as [10−16]. However, most of the results focus on the optimal control problems without delays. To the best of our knowledge, there are no results discussing how to use ADP to solve the time-delay optimal control problems. In this paper, the time-delay optimal control problem is solved by the iterative ADP algorithm for the first time. By introducing a delay matrix function, the explicit expression of the optimal control function is obtained. The optimal control can iteratively be obtained using the proposed iterative ADP algorithm which avoids the infinite-dimensional computation. Also, it is proved that the performance index function converges to the optimum using the proposed iterative ADP algorithm.

This paper is organized as follows. Section 1 presents the preliminaries. In Section 2, the time-delay optimal control scheme is proposed based on iterative ADP algorithm. In Section 3, the neural network implementation for the control scheme is discussed. In Section 4, two examples are given to demonstrate the effectiveness of the proposed control scheme. The conclusion is drawn in Section 5.

1 Preliminaries

Basically, we consider the following discrete-time affine nonlinear system with time delays in state and control variables as follows:

$$\begin{aligned}\boldsymbol{x}(k+1) =& f(\boldsymbol{x}(k),\boldsymbol{x}(k-\sigma)) + g_0(\boldsymbol{x}(k),\boldsymbol{x}(k-\sigma))\boldsymbol{u}(k) + \\ & g_1(\boldsymbol{x}(k),\boldsymbol{x}(k-\sigma))\boldsymbol{u}(k-\tau)\end{aligned} \tag{1}$$

with the initial condition given by $\boldsymbol{x}(s) = \boldsymbol{\phi}(s)$, $s = -\sigma, -\sigma+1, \cdots, 0$, where $\boldsymbol{x}(k) \in \mathbf{R}^n$ is the state vector, $f\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$ and $g_0, g_1\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^{n\times m}$ are differentiable functions and the control $\boldsymbol{u}(k) \in \mathbf{R}^m$. The state and control delays σ and τ are both nonnegative integral numbers. Assume that $f(\boldsymbol{x}(k), \boldsymbol{x}(k-\sigma)) + g_0(\boldsymbol{x}(k), \boldsymbol{x}(k-\sigma))\boldsymbol{u}(k) + g_1(\boldsymbol{x}(k), \boldsymbol{x}(k-\sigma))\boldsymbol{u}(k-\tau)$ is Lipschitz continuous on a set Ω in $\mathbf{R}^n$ containing the origin, and that system (1) is controllable in the sense that there exists a bounded control on Ω that asymptotically stabilizes the system. In this paper, how to design an optimal state feedback controller for this class of delayed discrete-time systems is mainly discussed. Therefore, it is desired to find the optimal control $\boldsymbol{u}(\boldsymbol{x})$ satisfying $\boldsymbol{u}(\boldsymbol{x}(k)) = \boldsymbol{u}(k)$ to minimize the generalized performance functional as follows:

Manuscript received September 5, 2008; accepted March 3, 2009
Supported by National High Technology Research and Development Program of China (863 Program) (2006AA04Z183), National Natural Science Foundation of China (60621001, 60534010, 60572070, 60774048, 60728307), and the Program for Changjiang Scholars and Innovative Research Groups of China (60728307, 4031002)
1. Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, P. R. China 2. School of Information Science and Engineering, Northeastern University, Shenyang 110004, P. R. China 3. Department of Automatic Control Engineering, Shenyang Institute of Engineering, Shenyang 110136, P. R. China
DOI: 10.1016/S1874-1029(09)60008-2

$$V(\boldsymbol{x}(0),\boldsymbol{u})=\sum_{k=0}^{\infty}\Big(\boldsymbol{x}^{\mathrm{T}}(k)Q_0\boldsymbol{x}(k)+2\boldsymbol{x}^{\mathrm{T}}(k)Q_1\boldsymbol{x}(k-\sigma)+\boldsymbol{x}^{\mathrm{T}}(k-\sigma)Q_2\boldsymbol{x}(k-\sigma)+\boldsymbol{u}^{\mathrm{T}}(k)R_0\boldsymbol{u}(k)+2\boldsymbol{u}^{\mathrm{T}}(k)R_1\boldsymbol{u}(k-\tau)+\boldsymbol{u}^{\mathrm{T}}(k-\tau)R_2\boldsymbol{u}(k-\tau)\Big) \tag{2}$$

where $\begin{bmatrix} Q_0 & Q_1 \\ Q_1^{\mathrm{T}} & Q_2 \end{bmatrix} \geq 0, \begin{bmatrix} R_0 & R_1 \\ R_1^{\mathrm{T}} & R_2 \end{bmatrix} > 0$, and $l(\boldsymbol{x}(k), \boldsymbol{x}(k-\sigma), \boldsymbol{u}(k), \boldsymbol{u}(k-\tau)) = \boldsymbol{x}^{\mathrm{T}}(k)Q_0\boldsymbol{x}(k) + 2\boldsymbol{x}^{\mathrm{T}}(k)Q_1\boldsymbol{x}(k-\sigma) + \boldsymbol{x}^{\mathrm{T}}(k-\sigma)Q_2\boldsymbol{x}(k-\sigma) + \boldsymbol{u}^{\mathrm{T}}(k)R_0\boldsymbol{u}(i) + 2\boldsymbol{u}^{\mathrm{T}}(k)R_1\boldsymbol{u}(k-\tau) + \boldsymbol{u}^{\mathrm{T}}(k-\tau)R_2\boldsymbol{u}(k-\tau)$ is the utility function. Let $V^*(\boldsymbol{x})$ denote the optimal performance index function which satisfies

$$V^*(\boldsymbol{x}) = \min_{\boldsymbol{u}} V(\boldsymbol{x},\boldsymbol{u}) \tag{3}$$

According to the Bellman′s optimal principle, we can get the following Hamilton-Jacobi-Bellman (HJB) equation

$$V^*(\boldsymbol{x}(k))=\min_{\boldsymbol{u}(k)}\Big\{\boldsymbol{x}^{\mathrm{T}}(k)Q_0\boldsymbol{x}(k)+2\boldsymbol{x}^{\mathrm{T}}(k)Q_1\boldsymbol{x}(k-\sigma)+\boldsymbol{x}^{\mathrm{T}}(k-\sigma)Q_2\boldsymbol{x}(k-\sigma)+\boldsymbol{u}^{\mathrm{T}}(k)R_0\boldsymbol{u}(k)+2\boldsymbol{u}^{\mathrm{T}}(k)R_1\boldsymbol{u}(k-\tau)+\boldsymbol{u}^{\mathrm{T}}(k-\tau)R_2\boldsymbol{u}(k-\tau)+V^*(\boldsymbol{x}(k+1))\} \tag{4}$$

For the optimal control problem, the state feedback control $\boldsymbol{u}(\boldsymbol{x})$ must not only stabilize the system on Ω but also guarantee that (2) is finite, i.e., $\boldsymbol{u}(\boldsymbol{x})$ must be admissible[17].

Definition 1. A control $\boldsymbol{u}(\boldsymbol{x})$ is defined to be admissible with respect to (3) on Ω if $\boldsymbol{u}(\boldsymbol{x})$ is continuous on Ω, $\boldsymbol{u}(0) = \boldsymbol{0}$, $\boldsymbol{u}(\boldsymbol{x})$ stabilizes (1) on Ω, and $\forall \boldsymbol{x}(0) \in \Omega$, $V(\boldsymbol{x}(0))$ is finite.

2 Properties of the iterative ADP approach

Since the nonlinear delayed system (1) is infinite-dimensional[3] and the control variable $\boldsymbol{u}(k)$ couples with $\boldsymbol{u}(k-\tau)$, it is nearly impossible to obtain the expression of the optimal control by solving the HJB equation (1). To overcome the difficulty, a new iterative algorithm is proposed in this paper. The following lemma is necessary to apply the algorithm.

Lemma 1. For the delayed nonlinear system (1) with respect to the performance index function (2), if there exists a control $\boldsymbol{u}(k) \neq \boldsymbol{0}$ at time point k, then there exists a bounded matrix function $M(k)$ that makes

$$\boldsymbol{u}(k-\tau) = M(k)\boldsymbol{u}(k) \tag{5}$$

hold for $j = 0, 1, \cdots, n$.

Proof. As $\boldsymbol{u}(k)$ and $\boldsymbol{u}(k-\tau_j)$, $j = 0, 1, \cdots, n$ are bounded real vectors, can construct a function that satisfies

$$\boldsymbol{u}(k-\tau) = h(\boldsymbol{u}(k)) \tag{6}$$

where $j = 0, 1, \cdots, n$. Then, using the method of undetermined coefficients, let $M(\boldsymbol{u}(k))$ satisfy

$$h(\boldsymbol{u}(k)) = M(\boldsymbol{u}(k))\boldsymbol{u}(k) \tag{7}$$

Then, we can obtain $M(\boldsymbol{u}(k))$ expressed as

$$M(\boldsymbol{u}(k)) = h(\boldsymbol{u}(k))\boldsymbol{u}^{\mathrm{T}}(k)\left(\boldsymbol{u}(k)\boldsymbol{u}^{\mathrm{T}}(k)\right)^{-1} \tag{8}$$

where $\left(\boldsymbol{u}(k)\boldsymbol{u}^{\mathrm{T}}(k)\right)^{-1}$ means the generalized inverse matrix of $\left(\boldsymbol{u}(k)\boldsymbol{u}^{\mathrm{T}}(k)\right)$. On the other side, $\boldsymbol{u}(k)$ and $\boldsymbol{u}(k-\tau)$ are both bounded real vectors, then we have $h(\boldsymbol{u}(k))$ and $\left(\boldsymbol{u}(k)\boldsymbol{u}^{\mathrm{T}}(k)\right)^{-1}$ are bounded. So $M(k) = M(\boldsymbol{u}(k))$ is the solution. □

According to Lemma 1, the HJB equation becomes

$$V^*(\boldsymbol{x}(k))=\boldsymbol{x}^{\mathrm{T}}(k)Q_0\boldsymbol{x}(k)+2\boldsymbol{x}^{\mathrm{T}}(k)Q_1\boldsymbol{x}(k-\sigma)+\boldsymbol{x}^{\mathrm{T}}(k-\sigma)Q_2\boldsymbol{x}(k-\sigma)+\boldsymbol{u}^{*\mathrm{T}}(k)R_0\boldsymbol{u}^*(k)+2\boldsymbol{u}^{*\mathrm{T}}(k)R_1M^*(k)\boldsymbol{u}^*(k)+\boldsymbol{u}^{*\mathrm{T}}(k)M^{*\mathrm{T}}(k)R_2M^*(k)\boldsymbol{u}^*(k)+V^*(\boldsymbol{x}(k+1)) \tag{9}$$

where $\boldsymbol{u}^*(k)$ is the optimal control and $\boldsymbol{u}^*(k-\tau) = M^*(k)\boldsymbol{u}^*(k)$.

2.1 Derivation of the iterative ADP algorithm

According to the Bellman′s principle of optimality, we can obtain the optimal control by differentiating the HJB equation (9) with respect to control $\boldsymbol{u}$.

Then, we can obtain the optimal control $\boldsymbol{u}^*(k)$ formulated as

$$\boldsymbol{u}^*(k)=-\frac{1}{2}\left(R_0+2R_1M^*(k)+M^{*\mathrm{T}}(k)R_2M^*(k)\right)^{-1}\times\left(g_0\left(\boldsymbol{x}(k),\boldsymbol{x}(k-\sigma)\right)+g_1\left(\boldsymbol{x}(k),\boldsymbol{x}(k-\sigma)\right)M^*(k)\right)^{\mathrm{T}}\times\frac{\partial V^*(\boldsymbol{x}(k+1))}{\partial \boldsymbol{x}(k+1)} \tag{10}$$

In (10), the inverse of the term $\left(R_0 + 2R_1M^*(k) + M^{*\mathrm{T}}(k)R_2M^*(k)\right)$ should exist and a proof is presented in Appendix to guarantee the existence of the inverse.

From (10), the explicit optimal control expression $\boldsymbol{u}^*$ is obtained by solving the HJB equation (9). We can see that the optimal control $\boldsymbol{u}^*$ depends on M^* and $V^*(x)$, where $V^*(x)$ is a solution to the HJB equation (9). While how to solve the HJB equation is still open, there is currently no method for rigorously seeking for this performance index function of this delayed optimal control problem. Furthermore, the optimal delay matrix function M^* is also unknown which makes the optimal control $\boldsymbol{u}^*$ more difficult to obtain. So an iterative index i is introduced into the ADP approach to obtain the optimal control iteratively.

Firstly, for $i = 0, 1, \cdots$, let

$$\boldsymbol{u}^{(i+1)}(k-\tau) = M^{(i)}(k)\boldsymbol{u}^{(i+1)}(k) \tag{11}$$

where $M^{(0)}(k) = I$ and $\boldsymbol{u}^{(0)}(k-\tau) = M^{(0)}(k)\boldsymbol{u}^{(0)}(k)$. We start with initial performance index $V^{(0)}(\boldsymbol{x}(k)) = 0$, and the control $\boldsymbol{u}^{(0)}(k)$ can be computed as follows

$$\boldsymbol{u}^{(0)}(\boldsymbol{x}(k)) = \arg\min_{\boldsymbol{u}}\left\{\Gamma^0 + V^{(0)}(\boldsymbol{x}(k+1))\right\} \tag{12}$$

where

$$\Gamma^0=\boldsymbol{x}^{\mathrm{T}}(k)Q_0\boldsymbol{x}(k)+2\boldsymbol{x}^{\mathrm{T}}(k)Q_1\boldsymbol{x}(k-\sigma)+\boldsymbol{x}^{\mathrm{T}}(k-\sigma)Q_2\boldsymbol{x}(k-\sigma)+\boldsymbol{u}^{(0)\mathrm{T}}(k)R_0\boldsymbol{u}^{(0)}(k)+2\boldsymbol{u}^{(0)\mathrm{T}}(k)R_1M^{(0)}(k)\boldsymbol{u}^{(0)}(k)+\boldsymbol{u}^{(0)\mathrm{T}}(k)M^{(0)\mathrm{T}}(k)R_2M^{(0)}(k)\boldsymbol{u}^{(0)}(k)$$

Then, the performance index function is updated as

$$V^{(1)}(\boldsymbol{x}(k)) = \Gamma^0 + V^{(0)}(\boldsymbol{x}(k+1)) \tag{13}$$