Acta Automatica Sinica
Vol. 39, No. 2 February, 2013

Available online at www.sciencedirect.com
ScienceDirect

Nearly Optimal Control Scheme Using Adaptive Dynamic Programming Based on Generalized Fuzzy Hyperbolic Model

Ji-Lie ZHANG[1] Hua-Guang ZHANG[1, 2] Yan-Hong LUO[1] Hong-Jing LIANG[1]

Abstract: An effective scheme is presented to design the nearly optimal control for continuous-time (C-T) nonlinear systems. The generalized fuzzy hyperbolic model (GFHM) is used to approximate the solution of the Hamilton-Jacobi-Bellman (HJB) equation (i.e., the value function) for the first time. Further, the approximate solution is utilized to obtain the nearly optimal control. The value function is estimated by only using single GFHM, which captures the mapping between the state and value function. First, we illustrate the design process for the nearly optimal control involving nonlinear systems. Then stability conditions and conservatism analysis are given, and the approximate errors are proven to be uniformly ultimately bounded (UUB). Finally, a numerical example illustrates the effectiveness of our method and an example compared with the adaptive method based on dual neural-network models is used to demonstrate the advantages of our method.

Keywords: Generalized fuzzy hyperbolic model (GFHM), optimal control, adaptive dynamic programming (ADP), approximation optimization, adaptive control

For the last decades, the adaptive dynamic programming (ADP) has played an important role in seeking solutions for the Hamilton-Jacobi-Bellman (HJB) equation[1−9]. Some recent papers[10−12] on ADP techniques present excellent overview of the state-of-the-art developments. In recent years, the nearly optimal control for C-T systems has been the focus of many researchers[13−17]. The main problem is how to solve the HJB equation[11, 18]. Now, there exist three popular approaches to address the problem, that is, Galerkin successive approximation approach, policy iteration (PI)[19] approach based on reinforcement learning algorithm[12, 19−20], and the adaptive approach based on neural networks.

How to obtain the solution of the HJB equation always troubles many researchers. Fortunately, a method which can obtain the closed form solutions of the HJB equation was developed by Beard[21−22]. In the method, the HJB equation is approximated via two steps. The first step is to reduce the HJB equation, which is a nonlinear partial differential equations, to a sequence of linear partial differential equations. The second step is to use the Galerkin spectral method to approximate the GHJB equation. The successive approximation method improves continually the control law and makes it close to the optimal control. However, the calculation process is offline.

Vrabie and Lewis gave the on-line PI method consisting of least squares and reinforcement learning algorithm[13−14], with partially unknown knowledges. The method starts by evaluating the cost of a given admissible initial policy, then obtains a new control policy, which is improved by a smaller value function compared with the previous policy. The policy and value function are repeatedly updated until the policy is no longer changed. It means that the optimal control behavior is obtained. The least squares solutions of the weights can be obtained by weighted residuals method, to approximate the solution of the HJB equation. Abu-Khalaf also designed the nearly optimal control for constrained nonlinear systems[23] by this method. However, the method needs to discretize the HJB equation and utilizes sample data to compute weights.

Recently, the adaptive approach based on dual neural-network models has been used to design the nearly optimal control for nonlinear systems. An online adaptive algorithm which effectively extends adaptive control techniques to successively tune actor and critic neural networks was presented in [15−16, 24]. In 2007, Tamimi used the method to solve the HJB equation for discrete-time nonlinear systems[25]. The method brings in informal style of the Weierstrass higher-order approximation theorem[26], which states that there exists a complete independent basis set approximating the solution for HJB equation.

Theoretically, the Weierstrass higher-order approximation theorem converges for N complete basis, where $N \rightarrow \infty$. For finite values of N, however, the algorithm will be sensitive to the chosen basis. If a smooth mapping is not spanned by finite N independent basis functions, then the neural network model will not be able to strictly approximate the mapping. Choosing a set of appropriate basis, therefore, is not very easy. A bad choice may result in poor precision and even not approximating the function. It motivates us to find a new estimator to overcome these disadvantages. For solving these problems, we select GFHM as an estimator for a mapping. Because GFHM is a strong nonlinear model, we do not consider their independence, and GFHM has universal approximate properties[27], i.e., it can approximate any nonlinear mappings in the compact set. Meanwhile, GFHM also can be seen as a neural network model, so the model weights can be optimized by powerful learning function. On the other hand, the single-network GFHM reduces the number of weight parameters. Computational burden is also reduced greatly, as compared with the dual neural-network models base on Weierstrass higher-order approximation theorem. When dealing with

Manuscript received April 10, 2012; accepted August 20, 2012
Supported by National Basic Research Program of China (973 Program) (2009CB320601), National High Technology Research and Development Program of China (863 Program) (2012AA040104), National Natural Science Foundation of China (50977008, 61034005), and Science and Technology Research Program of the Education Department of Liaoning Province (LT2010040)
Recommended by Associate Editor De-Rong LIU
Citation: Ji-Lie Zhang, Hua-Guang Zhang, Yan-Hong Luo, Hong-Jing Liang. Nearly optimal control scheme using adaptive, dynamic programming based on generalized fuzzy hyperbolic model. *Acta Automatica Sinica*, 2013, **39**(2): 142−149
1. School of Information Science and Engineering, Northeastern University, Shenyang 110819, China 2. State Key Laboratory of Synthetical Automation for Process Industries (Northeastern University), Shenyang 110819, China

high-dimension systems, our method, the single-network GFHM, has more obvious advantage for reducing computational burden. The more important point is that the norm of derivative for hyperbolic tangent function is less than one, so we can find that using GFHM has less conservatism than using the general neural network for stability conditions. The proof will be illustrated in the paper.

In this paper, we present a new method to design the nearly optimal control making use of generalized fuzzy hyperbolic model as an estimator, because of GFHM being an universal approximator. Only a GFHM is used to estimate the value function, so it saves half of storage space, as compared with that using the dual neural-network models as estimator[16]. Then, we can obtain the temporal difference residual error (i.e. the error of the approximate HJB equation) for C-T systems. To minimize the error, the gradient descent algorithm is utilized to search the nearly optimal value. Furthermore, we analyze the stability conditions and prove the weight, optimal control input and state errors are uniformly ultimately bounded (UUB).

The contributions of our method include:

1) By the single GFHM, reducing the number of the networks architecture and the approximation error.

2) It is a simple method that only a GFHM is enough to design the nearly optimal control rather than dual neural-network models including actor NN and critic NN. It reduces half of storage space.

3) Because the norm of derivative for hyperbolic tangent function is less than one, our method has less conservatism for stability conditions.

The rest of this paper is organized as follows. In Section 1, some definitions and notions are given. The nearly optimal control design process for nonlinear systems is stated in Sections 2 and 3. In Section 4, we give stability and conservatism analysis for our method and prove the state, optimal control input and weight approximate errors are uniformly ultimately bounded (UUB). Finally, a numerical example is given to illustrate the effectiveness of our method and demonstrate the advantages of our method by comparing with the adaptive method based on the dual neural-network models.

1 Preliminaries

Definition 1. Given a plant with n input variables $\boldsymbol{x} = [x_1(t), \cdots, x_n(t)]^{\mathrm{T}}$ and a single output variable y, if the output variable corresponds to a group of fuzzy rules which satisfy the following conditions:

1) For output variable y, the corresponding group of fuzzy rules has the following form:

Rule.

If x_1 is F_{x_1} and x_2 is $F_{x_2}, \cdots$, and x_n is F_{x_n}

Then $y = \theta^{\pm}_{F_{x_1}} + \theta^{\pm}_{F_{x_2}} + \cdots + \theta^{\pm}_{F_{x_n}}$,

where F_{x_i} $(i = 1, \cdots, n)$ are fuzzy sets of x_i, which include P_{x_i} (positive) and N_{x_i} (negative), and $\theta^{\pm}_{F_{x_i}}$ $(i = 1, \cdots, n)$ are $2n$ real constants corresponding to F_{x_i}.

2) The constant terms $\theta^{\pm}_{F_{x_i}}$ in the Then-part correspond to F_{x_i} in the If-part; that is, if the language value of F_{x_i} term in the If-part is P_{x_i}, $\theta^{+}_{F_{x_i}}$ must appear in the Then-part; if the language value of F_{x_i} term in the If-part is N_{x_i}, $\theta^{-}_{F_{x_i}}$ must appear in the Then-part; if there is no F_{x_i} in the If-part, $\theta^{\pm}_{F_{x_i}}$ is zero in the Then-part.

3) There are 2^n fuzzy rules in the rule base; that is, there are a total of 2^n input variable combinations of all the possible P_{x_i} and N_{x_i} in the If-part.

Then we call this group of fuzzy rules as the fuzzy hyperbolic rule base (FHRB). To describe a plant with a single output variable, we will need an FHRB.

Lemma 1[27]. Given an FHRB, if we define the membership function of P_{x_i} and N_{x_i} as:

$$\mu_{P_{x_i}}(x_i) = \mathrm{e}^{-\frac{1}{2}(x_i-\phi_i)^2},$$
$$\mu_{N_{x_i}}(x_i) = \mathrm{e}^{-\frac{1}{2}(x_i+\phi_i)^2},$$

where $i = 1, \cdots, n$ and ϕ_i are positive constants, if we denote $\theta^{+}_{F_{x_i}}$ by $\theta_{P_{x_i}}$ and $\theta^{-}_{F_{x_i}}$ by $\theta_{N_{x_i}}$, and after applying singleton fuzzifier, product inference, and the center of gravity defuzzifier to fuzzy rule, then we can derive

$$y = \boldsymbol{\theta}^{\mathrm{T}} \tanh(\Phi \boldsymbol{x}) + \zeta,$$

where $\boldsymbol{\theta} = [\theta_1\ \theta_2\ \cdots\ \theta_n]^{\mathrm{T}}$ is an ideal vector, ζ is a constant scalar, $\tanh(\Phi\boldsymbol{x}) = [\tanh(\phi_1 x_1)\ \ \cdots\ \ \tanh(\phi_n x_n)]^{\mathrm{T}}$, and $\Phi = \mathrm{diag}\{\phi_i\}$ $(i = 1, \cdots, n)$. We call it a generalized fuzzy hyperbolic model (GFHM), as $n = m (m = 1, ..., \infty)$ (the generalized input variables), detailed see [34].

Lemma 2[28]. For any given real continuous $f(\boldsymbol{x})$ on the compact set $U \subset \mathbf{R}^n$ and any arbitrary $\delta > 0$, there exists an $h(\boldsymbol{x}) \in F$ (F is the set of all the fuzzy basis function expansions) such that

$$\sup_{\boldsymbol{x} \in U} |f(\boldsymbol{x}) - h(\boldsymbol{x})| < \delta.$$

Remark 1. Lemma 2 indicates that GFHM can be seen as a function approximator.

Definition 2[29]. The Frobenius norm of matrix $A \in \mathbf{R}^{n\times n}$ is defined as $\|A\|_F^2 = \mathrm{tr}(A^{\mathrm{T}}A) = \sum a_{ij}^2$ with $\mathrm{tr}(\cdot)$ being the trace operator. There exists the following inequality

$$\|A\boldsymbol{x}\| \leq \|A\|_F \|\boldsymbol{x}\|.$$

2 Optimal control design

Consider the following known nonlinear C-T systems:

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}) + g(\boldsymbol{x})\boldsymbol{u}, \tag{1}$$

where $\boldsymbol{x} \in \mathbf{R}^n$ is the state vector, $f(\boldsymbol{x}) \in \mathbf{R}^n$, $g(\boldsymbol{x}) \in \mathbf{R}^{n\times p}$ and $\boldsymbol{u} \in \mathbf{R}^p$ is the control input vector. $f(\boldsymbol{x}) + g(\boldsymbol{x})\boldsymbol{u}$ is Lipschitz continuous nonlinear function vector with $f(0) = 0$, on a set $\boldsymbol{\Omega} \subseteq \mathbf{R}^n$ which contains the origin.

Assumption 1. System (1) is stabilizable on $\boldsymbol{\Omega}$, i.e., there exists a continuous control function $\boldsymbol{u}$ such that the system is asymptotically stable on $\boldsymbol{\Omega}$.

We define the infinite horizon cost functional:

$$J(\boldsymbol{x_0}) = \int_0^{\infty} r(\boldsymbol{x}, \boldsymbol{u})\mathrm{d}t, \tag{2}$$

where $r(\boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{x}^{\mathrm{T}}Q\boldsymbol{x} + \boldsymbol{u}^{\mathrm{T}}R\boldsymbol{u}$ is the utility function, $Q = Q^{\mathrm{T}} > 0$ and $R = R^{\mathrm{T}} > 0$.

Definition 3 (Admissible policy)[15]. A control policy $\boldsymbol{\mu}(\boldsymbol{x})$ is defined as admissible such that it must not only stabilize the systems on $\boldsymbol{\Omega}$ but also make the integral of cost functional finite.

Definition 4 (Uniformly ultimately bounded (UUB))[30]. The equilibrium point $\boldsymbol{x_e} = 0$ of (1) is said to be uniformly ultimately bounded (UUB) if there exists a compact set $S \subset \mathbf{R}^n$ that for all $\boldsymbol{x_0} \in S$ there exist a bound