Robust Controller Design of Continuous-time Nonlinear System Using Neural Network

Xiangnan Zhong, Haibo He, and Danil V. Prokhorov

Abstract—In this paper, we propose an optimal control method based on the solution of Hamilton-Jacobi-Bellman (HJB) equation for the continuous-time nonlinear system with bounded unknown perturbation. The robust control system is converted into the corresponding optimal control system with appropriate performance index and the equivalence of the transformation is proved, i.e., the solution of the optimal control problem can globally asymptotically stabilize the robust control system. Adaptive dynamic programming (ADP) based approach is presented to iteratively approximate the optimal performance index and obtain the optimal control policy. A neural network with adaptive weights is applied to implement this approach. An example is given to illustrate the proposed method.

I. INTRODUCTION

Stability analysis and controller design are among the most important issues in control problem. In practice, the control systems are always subject to exogenous perturbation, model uncertainties or other changes due to the limited knowledge of certain physical phenomena, especially for the nonlinear systems. Since uncertainties are frequently a source of instability, they are necessarily considered during the controller design process in order to obtain the high performance of the systems. Generally, we should find a robust controller which works even if the actual system deviates from its ideal model on which the controller design is based. In recent years, robust control problems have witnessed extensive studies and many approaches have been proposed such as fuzzy approach [1], [2], sliding mode approach [3], Lyapunov approach [4], etc. In [5], the feasibility to convert robust control problem into the corresponding optimal control problem is presented for nonlinear system with uncertainties. However, the detailed implementation process is not discussed.

In terms of optimal control problem, controller design for linear system can be obtained by solving the Riccati equation. However, when comes to the nonlinear control problem, Riccati equation becomes the well-known Hamilton-Jacobi-Bellman (HJB) equation which is difficult to tackle directly [6], [7]. Fortunately, adaptive dynamic programming (ADP), proposed by Werbos[8], [9], has been widely recognized as one of the “core methodologies” to achieve optimal control in stochastic process in a general case to achieve brain-like intelligent control. Extensive efforts and promising results have been achieved over the past decades. Among these achievements, we highlight Al-Tamimi et al [10], Abu-Khalaf et al [11], Wei and Liu [12], [13], [14], Lewis and Vamvoudakis [15], He et al [16], [17], Tyukin and Prokhorov [18], [19] and Zhang et al [20], [21] from the theoretical perspective. These achievements cover a wide variety of problems, including system stability, convergence analysis, controller design, optimal control, state prediction, etc. Furthermore, Lin presented the feasibility of designing the optimal controller for the optimal control system to stabilize the robust control system in [22] and divided the uncertainties into matched uncertainties and unmatched uncertainties according to whether the unknown perturbation is in the range space of the input matrix. In [23], [24], Adhyaru expended this transformation method into bounded input robust control problems.

In this paper, we build on the developments in [10], [22], [23], [24] to design the robust controller for continuous-time nonlinear system with uncertainties by finding the optimal control policy for the corresponding optimal control system with appropriate performance function. In other words, the solution of the optimal control problem can globally asymptotically stabilize the robust control system. The equivalent analysis of the transformation between these two problems has been proved. The ADP neural network based approach is applied to iteratively approximate the performance index and improve the control policy until the near-optimal value is achieved.

The rest of this paper is organized as follows. In Section II, we formulate the robust control system analyzed in this paper and the corresponding optimal control system with appropriate performance index. Then the HJB equation for continuous-time nonlinear system is introduced. The detailed equivalent analysis of the transformation between these two systems is presented in Section III. Section IV first proposes the ADP approach to approximate the performance index and obtain the control policy iteratively. Then, neural network, actually the adaptive critic network, is introduced to implement this ADP scheme. A numerical example with four cases is given in Section V to demonstrate the effectiveness of the proposed approach, followed by conclusion in Section VI.
II. PROBLEM STATEMENT

Consider the continuous-time nonlinear system with unknown perturbation

\[ \dot{x}(t) = f(x(t)) + g(x(t))u(t) + d(x(t))W(x(t)) \]  

(1)

where \( x(t) \in \mathbb{R}^n \) is the state vector with the initial state \( x(0) = x_0 \) and \( u(t) \in \mathbb{R}^m \) is the input vector, \( f(x(t)), g(x(t)) \) and \( d(x(t)) \) are the differentiable system functions. \( W(x(t)) \) is the unknown perturbation which is bounded by a non-negative known function \( W_m(x(t)) \), i.e.

\[ \|W(x(t))\| \leq W_m(x(t)) \]  

(2)

In this paper, we assume \( f(0) = 0 \) and \( W(0) = 0 \), which means \( x = 0 \) is an equilibrium point.

For system (1), in order to deal with the robust control problem, we should find a feedback control policy \( u(x(t)) = \mu \) so that the close-loop system

\[ \dot{x}(t) = f(x(t)) + g(x(t))\mu + d(x(t))W(x(t)) \]  

(3)

is globally asymptotically stable for all uncertainties \( W(x(t)) \) satisfied equation (2). Note that if \( d(x(t)) = g(x(t)) \), the unknown perturbation \( W(x(t)) \) is in the range space of \( g(x(t)) \). We call this condition as matched condition. This paper deals with the robust control problem with unmatched condition, i.e., \( d(x(t)) \neq g(x(t)) \).

Consider to use ADP to solve this kind of problem. Generally, ADP, however, can not deal with the robust control problem directly. The following shows we can convert the robust control problem into an optimal control problem by designing an optimal controller for the corresponding auxiliary system with appropriate performance index.

In order to simplify the formula, we eliminate the time index \( t \) in the following statement.

Decompose the uncertainty term \( d(x)W(x) \) into a matched component and an unmatched component by projecting \( d(x)W(x) \) onto the range of \( g(x) \):

\[ d(x)W(x) = g(x)g^*(x)d(x)W(x) \]

(4)

\[ + (I - g(x)g^*(x))d(x)W(x) \]

in which \( g^*(x) \) is the Moore-Penrose pseudoinverse of function \( g(x) \).

Thus, consider the following auxiliary system

\[ \dot{x} = f(x) + g(x)u + (I - g(x)g^*(x))d(x)v \]  

(5)

where \([u; v]\) is the continuous-time control policy of system (5). In this new system, an augmented control \( v \in \mathbb{R}^p \) is added to deal with the unmatched uncertainty component. \( v \) plays an important role in finding the optimal control policy \( u^* \) and proving the closed-loop stability of the system (1).

Assume that system (5) is controllable. It is desired to find the optimal control \([u^*; v^*]\) that minimizes the performance index

\[ V(x_0) = \int_0^\infty \left[ g_m^2(x(\tau)) + \eta^2 m^T \|W_m(x(\tau))\|^2 \right] d\tau + U(x(\tau), u(\tau), v(\tau)) \]  

(6)

where \( \eta > 0 \) is the design parameter. \( U(x, u, v) \) is the utility function with \( U(0, 0, 0) = 0 \). In this paper, the utility function is given by

\[ U(x, u, v) = x^T Q x + u^T R u + \eta^2 v^T M v \]  

(7)

in which \( Q, R \) and \( M \) are symmetric and positive definite matrices with appropriate dimensions. And, they can be described by

\[ Q = q^T \cdot q \]  

(8)

\[ R = r^T \cdot r \]  

(9)

\[ M = m^T \cdot m \]  

(10)

In addition, \( g_m(x) \) is the upper bound of the term \( \|x^T g^*(x) d(x) W(x)\| \), which is

\[ \|x^T g^*(x) d(x) W(x)\|^2 \leq g_m^2(x) \]  

(11)

Note that the performance index (6) is a modified form of the ordinary one, in which \( U(x, u, v) \) reflects the auxiliary system described in (5), and \( g_m^2(x) + \eta^2 m^T \|W_m(x)\|^2 \) reflects the uncertainty.

For optimal control problem, the control vector must not only stabilize the system on the compact set \( \Omega \in \mathbb{R}^n \), but also guarantee that the performance index (6) is finite, which called admissible control.

**Definition 1** (Admissible Control) A control law is said to be an admissible control with respect to (6) on \( \Omega \), if \([u; v]\) is continuous on \( \Omega \) and can stabilize system (5) for all \( x_0 \in \Omega \), \([u; v]\in [0; 0] \) as \( x = 0 \). For \( v(x_0), v(x) \) is finite.

Equation (6) can be expended as follows

\[ V(x_0) = \int_0^T \left[ g_m^2(x(\tau)) + \eta^2 m^T \|W_m(x(\tau))\|^2 \right] d\tau + U(x(\tau), u(\tau), v(\tau)) \]

(12)

\[ + \int_0^\infty \left[ g_m^2(x(\tau)) + \eta^2 m^T \|W_m(x(\tau))\|^2 \right] d\tau + U(x(\tau), u(\tau), v(\tau)) \]

\[ + \int_0^T U(x(\tau), u(\tau), v(\tau)) d\tau + V(x(T)) \]

If the associate performance index (6) is continuously differentiable, then after transformation, equation (12) becomes

\[ \lim_{T \to 0} \frac{[V(x(T)) - V(x_0)]}{T} = - \lim_{T \to 0} \int_0^T \left[ g_m^2(x(\tau)) + \eta^2 m^T \|W_m(x(\tau))\|^2 \right] d\tau + U(x(\tau), u(\tau), v(\tau)) \]

(13)

Therefore, we obtain the infinitesimal version of (6) as

\[ V_x^T (f(x) + g(x)u + (I - g(x)g^*(x))d(x)v + g_m^2(x) + \eta^2 m^T \|W_m(x)\|^2 + U(x, u, v) = 0 \]

(14)

with \( V(0) = 0 \).
Equation (14) is a sort of nonlinear Lyapunov equation and the notation $V_x$ denotes the partial derivative of the performance index $V(x)$ with respect to $x$

$$V_x = \frac{\partial V(x)}{\partial x} \quad (15)$$

According to Bellman’s optimality equation, the optimal performance index is given by

$$V^*(x_0) = \min_{u,v} \int_0^\infty [g_m^2(x(\tau)) + \eta^2 ||m^T||^2 W_m^2(x(\tau))]$$

$$+ U(x(\tau), u(\tau), v(\tau))]d\tau \quad (16)$$

This equation becomes the well-known Hamilton-Jacobi-Bellman (HJB) equation on substitution of the optimal control

$$u^* = -\frac{1}{2} R^{-1} g^T(x)V_x^* \quad (17)$$

and

$$v^* = -\frac{1}{2\eta^2} M^{-1} d^T(x)(I - g(x)\sigma^*(x))TV_x^* \quad (18)$$

where $V_x^*$ is $\partial V^*(x)/\partial x$ is solved in the following HJB equation

$$V^*_x f(x) + x^T Q x + g_m^2(x) + \eta^2 W_m^2(x)$$

$$- \frac{1}{4} V^*_x g(x) R^{-1} g^T(x)V_x^* + \left( -\frac{1}{2\eta^2} + \frac{1}{4\eta^4} \right) V_x^T$$

$$\cdot (I - g(x)\sigma^*(x))d(x) M^{-1} d^T(x)(I - g(x)\sigma^*(x))TV_x^* = 0 \quad (19)$$

III. EQUIVALENT ANALYSIS OF THE TRANSFORMATION BETWEEN OPTIMAL AND ROBUST CONTROL PROBLEM

In this section, we design the robust controller for system (1) by solving the optimal controller for system (5) with the performance index (6). The following theorem proves the optimal control policy (17) is the solution of the robust control problem.

**Theorem 1:** Consider the auxiliary system (5) with the performance index (6). Assuming $V^*(x)$ is the solution of the HJB equation (19), if using this solution, the optimal control policy $[u^*, v^*]$ exists and satisfies the following condition

$$v^* M v^* \leq x^T Q x \quad (20)$$

and the positive parameter $\eta$ satisfies

$$0 < \eta < \frac{\sqrt{2}}{2} \quad (21)$$

then, the obtained optimal control policy $u^*$ is a solution of the robust control problem.

**Proof:** Let $[u^*, v^*]$ and $V^*(x)$ be the optimal control policy and optimal performance index obtained in equation (17), (19) and (21) respectively. We will show that $u^*$ is the solution of the robust control problem, i.e., the equilibrium point $x = 0$ of system (1) is globally asymptotically stable for all the admissible uncertainties by using this solution.

From equation (16), we know $V^*(x)$ is a positive definite function, namely, $V^*(x) > 0$ for any $x \neq 0$ and $V^*(x) = 0$ when $x = 0$. Therefore, $V^*(x)$ can be seen as a Lyapunov function.

Using equation (14), we obtain

$$V^*_x f(x) + g(x) u + (I - g(x)\sigma^*(x))d(x(t))v$$

$$= -g_m^2(x) - \eta^2 ||m^T||^2 W_m^2(x) - U(x, u, v) \quad (22)$$

Then, depend on formula (17) and (18), we have

$$2u^* M + V^*_x f(x) = 0 \quad (23)$$

and

$$2\eta^2 v^* M + V^*_x g(x) = 0 \quad (24)$$

Taking the derivative of $V^*(x)$ along the system trajectory, we have

$$\dot{V}^*(x) = \left( \frac{\partial V^*(x)}{\partial x} \right)^T \dot{x}$$

$$= V^*_x f(x) + g(x) u + d(x) W(x)$$

$$= V^*_x f(x) + g(x) u + (I - g(x)\sigma^*(x))d(x) v + g(x) g^*(x) d(x) W(x)$$

$$+ (I - g(x)\sigma^*(x))d(x) (W(x) - v^*)$$

$$= V^*_x f(x) + g(x) u + (I - g(x)\sigma^*(x))d(x) v + g(x) g^*(x) d(x) W(x)$$

$$+ V^*_x (g(x) g^*(x) d(x) W(x))$$

$$+ V^*_x ((I - g(x)\sigma^*(x)) d(x) (W(x) - v^*)) \quad (25)$$

By using (22), (23) and (24), equation (25) becomes

$$\dot{V}^*(x) = -g_m^2(x) - \eta^2 ||m^T||^2 W_m^2(x) - x^T Q x - u^* M u^*$$

$$- \eta^2 v^* M v^* - 2u^* M g^*(x) d(x) W(x) - 2\eta^2 v^* M (W(x) - v^*) \quad (26)$$

Because $R = r \cdot r^T$ and $M = m \cdot m^T$, we obtain

$$- u^* M u^* - 2u^* M g^*(x) d(x) W(x)$$

$$= -\|r^T u^* + r^T g^*(x) d(x) W(x)\|^2 \quad (27)$$

and

$$-2\eta^2 v^* M W(x) \leq \eta^2 (\|m^T v^*\|^2 + \|m^T W(x)\|^2) \quad (28)$$

Then, equation (26) becomes

$$\dot{V}^*(x) \leq -g_m^2(x) - \eta^2 ||m^T||^2 W_m^2(x) - x^T Q x - \eta^2 v^* M v^*$$

$$- \|r^T u^* + r^T g^*(x) d(x) W(x)\|^2$$

$$+ \|r^T g^*(x) d(x) W(x)\|^2 + \eta^2 \|m^T V^*(x)\|^2$$

$$+ \eta^2 \|m^T W(x)\|^2 + 2\eta^2 v^* M v^* \quad (29)$$
According to (2) and (11), equation (29) can be modified as following

\[
\dot{V}^*(x) \leq -\left(g^2(x) - \|r^Tq(x)d(x)W(x)\|^2 - x^TQx - \eta^2\|m^T\|^2\left(W^2_m(x) - \|W(x)\|^2\right) + 2\eta^2v^TMv^* - \|r^Tu^* + r^Tg(x)\)d(x)W(x)\|^2\right.
\]
\[
\quad - \eta^2\|m^T\|^2\left(W^2_m(x) - \|W(x)\|^2\right) + 2\eta^2v^TMv^* - x^TQx
\]
\[
\leq 2\eta^2v^TMv^* - x^TQx
\]
\[
= 2\eta^2(v^TMv^* - x^TQx) - (1 - 2\eta^2)x^TQx
\]
(30)

Therefore,

\[
\dot{V}^*(x) \leq -(1 - 2\eta^2)x^TQx < 0
\]
(31)

for any \(x \neq 0\). Thus, \([u^*; v^*]\) can globally asymptotically stabilize the robust control system (1). Namely, \(u^*\) is a solution of system (1). The conclusion is proved.

Note that, if we set \(d(x) = g(x)\), system (5) becomes

\[
\dot{x} = f(x) + g(x)u
\]
(32)

which means the unmatched uncertainty component turns to zero. Therefore, in this special case, problem can be solved without introducing any additional control policies.

IV. CONTROLLER DESIGN BY NEURAL NETWORK

In this section, an iterative approach to solve the HJB equation (19) and obtain the optimal control policy is introduced by using neural network. Two subsections are included. The first one gives the iterative ADP approach for continuous-time nonlinear system with uncertainties to obtain the performance index and the control policy. The neural network implementation is presented in the second subsection.

A. ADP Approach to Approximate the HJB Equation and Control Policy

In the ADP approach, we set the initial value of performance index as \(V^{(0)}(x) = 0\), and the initial admissible control policy as \(U^{(0)} = [u^{(0)}; v^{(0)}]\). Based on this control policy, solve the nonlinear Lyapunov function (14) for \(V^{(1)}(x)\) as

\[
V^{(1)}_x(f(x) + g(x)u^{(0)} + (I - g(x)g^T(x))d(x)v^{(0)}) + g^2_m(x) + \eta^2\|m^T\|^2W^2_m(x) + U(x, u^{(0)}, v^{(0)}) = 0
\]
(33)

According to equation (17) and (18), we can update the control policy by computing

\[
U^{(1)} = \begin{bmatrix} u^{(1)} \\ v^{(1)} \end{bmatrix}
\]

\[
= \begin{bmatrix} -\frac{1}{2}R^{-1}g^T(x)V^{(1)}_x \\ -\frac{1}{2\eta^2}M^{-1}d^T(x)(I - g(x)g^T(x))V^{(1)}_x \end{bmatrix}
\]
(34)

The ADP algorithm, therefore, is obtained by iterating between

\[
V^{(i+1)}_x(f(x) + g(x)u^{(i)} + (I - g(x)g^T(x))d(x)v^{(i)}) + g^2_m(x) + \eta^2\|m^T\|^2W^2_m(x) + U(x, u^{(i)}, v^{(i)}) = 0
\]
(35)

and

\[
U^{(i+1)} = \begin{bmatrix} u^{(i+1)} \\ v^{(i+1)} \end{bmatrix}
\]

\[
= \begin{bmatrix} -\frac{1}{2}R^{-1}g^T(x)V^{(i+1)}_x \\ -\frac{1}{2\eta^2}M^{-1}d^T(x)(I - g(x)g^T(x))V^{(i+1)}_x \end{bmatrix}
\]
(36)

where \(i\) is the iterative index.

The iterative results can converge to the optimal performance index and optimal control according to [10], [25]. The following subsection describes the neural network based implementation.

B. Neural-network-based Implementation

Critic neural network is used to implement the ADP approach in this paper with a three-layer network structure [26]. The purpose of the critic network is to approximate the performance index (16) which can be formulated as

\[
\dot{V}(x) = \omega^T_{c2}Ψ(h_c)
\]
(37)

where \(\omega_{c2}\) is the weight matrix between hidden and output layer of critic network. \(h_c = \omega^T_{c1}[x_T, u^T, \dot{v}^T]^T\), to which \(\omega_{c1}\) denotes the weight matrix between the input and hidden layer of critic network. \(\dot{u}\) and \(\dot{v}\) are the estimated control policies. Note that \(\omega_{c1}\) is fixed during the implementation process in this paper.

\(Ψ(\cdot)\) is a sigmoid function that can be described as

\[
Ψ(\cdot) = \frac{1 - e^{-\cdot}}{1 + e^{-\cdot}}
\]
(38)

The optimal performance index can be formulated as

\[
V^*(x) = \omega^T_{c2}Ψ(h_c) + \epsilon(x)
\]
(39)

where \(\omega_{c2}\) is the optimal weight matrix between hidden and output layer and \(\epsilon(x)\) is the approximate error of critic network. The derivative of the optimal performance index with respect to \(x\) is

\[
V_x = \frac{∂V^*(x)}{∂x} = \left(\frac{∂Ψ(h_c)^T}{∂x}\right)^T \omega^*_{c2} + \frac{∂ε(x)}{∂x} \dot{x} + g^2_m(x)
\]
(40)

then the nonlinear Lyapunov equation (14) turns into

\[
\left(\frac{∂Ψ(h_c)^T}{∂x}\right)^T \omega^*_{c2} + \frac{∂ε(x)}{∂x} \dot{x} + g^2_m(x)
\]
\[
+ η^2\|m^T\|^2W^2_m(x) + U(x, u, v) = 0
\]
(41)

However, the ideal weight matrix is unknown, equation (37) is applied to approximate the results. In this situation, equation (14) is reconstructed in terms of the estimated performance index as

\[
\left(\frac{∂Ψ(h_c)^T}{∂x}\right)^T \omega_{c2} \dot{x} + g^2_m(x) + η^2\|m^T\|^2W^2_m(x)
\]
\[
+ U(x, u, v) = ε_c
\]
(42)
Adjusting the weights is to minimize the squared error

\[ E_c = \frac{1}{2} e_c^2 \]  

(43)

Hence, we obtain the critic network weight adjustments for the hidden to the output layer

\[
\dot{\omega}_{c2} = -\beta_c \frac{\partial E_c}{\partial \omega_{c2}} = -\beta_c \frac{\partial E_c}{\partial e_c} \frac{\partial e_c}{\partial \omega_{c2}}
\]

(44)

where \( \beta_c > 0 \) is the learning rate of critic network.

The control policy should also be updated iteratively to obtain the optimal value that minimizes the current performance index. Therefore, we have

\[
\dot{U} = \begin{bmatrix}
\dot{u}
\dot{v}
\end{bmatrix} = \begin{bmatrix}
-\frac{1}{2} R^{-1} g^T(x) \dot{V}_x \\
-\frac{1}{2 \gamma^2} M^{-1} d^T(x) (I - g(x) g^T(x)) \dot{V}_x
\end{bmatrix}
\]

(45)

where \( \dot{V}_x \) is the approximated partial derivative of performance index. According to equation (37), \( \dot{V}_x \) can be formulated as

\[
\dot{V}_x = \frac{\partial V(x)}{\partial v} \left[ \begin{array}{c}
\dot{u} \\
\dot{v}
\end{array} \right] = \frac{\partial V(x)}{\partial v} \left[ \begin{array}{c}
\dot{u} \\
\dot{v}
\end{array} \right] + \frac{\partial V(x)}{\partial u} \left[ \begin{array}{c}
u \\
v
\end{array} \right] + \frac{\partial V(x)}{\partial v} \left[ \begin{array}{c}
u \\
v
\end{array} \right]
\]

(46)

to which \( \omega_{c1}(x) \) is the fixed weight matrix of \( x \) component for input to hidden layer of critic network. The choice of fixed (non-adaptive) \( \omega_{c1}(x) \) is often sub-optimal, according to the analysis in [19], and we made it to simplify the overall algorithm.

Note that we use critic network to approximate equation (35) and use equation (45) to obtain the control policy. These two parts are used iteratively to achieve the optimal value. In addition, implementing this approach does not involve the system function \( f(x) \).

V. SIMULATION RESULTS

One example is developed in this section to demonstrate the performance of the proposed method. This example is taken from [22] with some modifications.

Consider the continuous-time nonlinear systems

\[
\begin{align*}
\dot{x}_1 &= x_2 + \lambda_1 x_1 \cos \left( \frac{1}{x_2 + \lambda_2} \right) + \lambda_3 x_2 \sin \left( \lambda_4 x_1 x_2 \right) \\
\dot{x}_2 &= u
\end{align*}
\]

(47)

where \( \lambda_1, \lambda_2, \lambda_3 \) and \( \lambda_4 \) are the unknown parameters. The term \( d(x)W(x) = \lambda_1 x_1 \cos \left( \frac{1}{x_2 + \lambda_2} \right) + \lambda_3 x_2 \sin \left( \lambda_4 x_1 x_2 \right) \) reflects the unknown perturbations in the system. For convenience, we assume \( \lambda_1 \in [-1, 0.2], \lambda_2 \in [-100, 100], \lambda_3 \in [-0.2, 1], \) and \( \lambda_4 \in [-100, 0] \). Therefore, the robust control system (47) can be reconstructed as

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
0 & 0
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
0 & 1 \\
0 & 1
\end{bmatrix} u + \begin{bmatrix}
1 & 0
\end{bmatrix} W(x)
\]

(48)

to which

\[
W(x) = \lambda_1 x_1 \cos \left( \frac{1}{x_2 + \lambda_2} \right) + \lambda_3 x_2 \sin \left( \lambda_4 x_1 x_2 \right)
\]

(49)

Clearly, in this situation, \( g^*(x) = (g^T(x)g(x))^{-1} g^T(x) = g^T(x) = [0, 1] \). Set \( Q, R, r \) and \( m \) are the identity matrices with appropriate dimensions. Therefore, \( \|r^T g^*(x) d(x) W(x)\|^2 = 0 \leq g_m^2(x) \)

(50)

\[
\|W(x)\|^2 \leq x_1^2 + x_2^2 \leq W_m^2(x)
\]

(51)

and

\[
(I - g(x) g^T(x)) d(x) = \begin{bmatrix}
0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0
\end{bmatrix} = \begin{bmatrix}
1 & 0
\end{bmatrix}
\]

(52)

Choose \( \eta = 0.5 \sqrt{2}/2 \). Then, we can obtain the auxiliary optimal control system

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
0 & 0
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
0 & 1 \\
0 & 1
\end{bmatrix} u + \begin{bmatrix}
1 & 0
\end{bmatrix} v
\]

(53)

we need to find out the optimal control policy \([u^*; v^*]\) to minimize the performance index

\[
\int_0^\infty \left( 1.5 x_1^T x + u^T u + 0.5 v^T v \right) dt
\]

(54)

During the implementation process of the proposed method, we construct the critic network with the neuron structure of \( 4 \times 5 \times 1 \) which is an one-hidden-layer network. The initial weights of the critic network are chosen randomly in \([0, 0.5]\), and the initial control policy \([v_0; v_0]\) is stochastically chosen in \([0, 1]\). Furthermore, set the learning rate of critic network be \( \beta_c = 0.01 \), and the initial state of the system be \( x_0 = [1, 0.5]^T \). The sampling time is chosen as \( 0.05s \). Equation (44) and (45) are applied iteratively to achieve the optimal weights of critic network.

![Fig. 1. Weights of the critic network.](image-url)
The trajectories of the weights of critic network between hidden and output layer are shown in Fig.1. We can see the weights converge to
\[ \omega_{c2}^* = [0.1352, 0.2608, 0.5923, 0.4374, 0.1816]^T \]  

(55)

Based on the converged weights of critic network, we can use equation (17) and (18) to obtain the optimal control policy as
\[ U^* = \begin{bmatrix} u^* \\ v^* \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} [0 \ 1] V_x^* \\ -[1 \ 0] V_x^* \end{bmatrix} \]  

(56)

where
\[ V_x^* = \frac{1}{2} \omega_{c2}^T \left( 1 - \Psi^2(h_c) \right) \omega_{c1}(x) \]  

(57)

Note that \( v^* \) is not the direct control policy of the robust control system (47), but it plays an important role not only in solving the optimal weights of critic network, but also in finding the optimal control policy \( u^* \) for system (47) based on equation (56).

Now, we simulate the close-loop system with unknown perturbation based on optimal control policy (56). Set the initial state be \( x_0 = [2, -2]^T \) and the sampling time be 0.05s. The simulation results for the following four cases are given in Fig.2-Fig.5, respectively.

Case 1: \( \lambda_1 = -1, \lambda_2 = -100, \lambda_3 = 0, \lambda_4 = -100 \);

Case 2: \( \lambda_1 = 0.2, \lambda_2 = 100, \lambda_3 = 1, \lambda_4 = -1 \);

Case 3: \( \lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 0, \lambda_4 = 0 \);

Case 4: \( \lambda_1 = -0.2, \lambda_2 = -100, \lambda_3 = -0.2, \lambda_4 = -100 \);

From the results (Fig.2-Fig.5), we can observe that under the action of the optimal control policy \( u^* \), the states of above four cases can converge to the equilibrium point. In other words, the optimal control policy \( u^* \) designed by neural-network-based iterative ADP method is the robust control policy of the original system with unknown perturbation and can globally asymptotically stabilize the nonlinear system with admissible uncertainties. Verification of the condition (20) of Theorem 1 in the above four cases can be observed from Fig.6-Fig.9. We can know these cases all satisfy the condition \( v^T M v^* \leq x^T Q x \).

In Fig.2-Fig.9, we use four specific cases to show the convergence of the robust control system in the action of optimal control \( u^* \) which is obtained in the equivalent optimal control system. From the results, we can clearly observe that a set of perturbation parameters (\( \lambda_1, \lambda_2, \lambda_3 \) and \( \lambda_4 \)) comes with a new state trajectory. Without loss of generality, we randomly choose the perturbation parameters within their boundaries, respectively. Then, the root mean square error (RMSE) is measured based on each set of parameters.

We repeat this process for 5,000 times and plot the histogram of RMSE for these 5,000 round of \( x_1 \) and \( x_2 \).
Verification of condition (Case 1) in Fig. 6.

Verification of condition (Case 2) in Fig. 7.

Verification of condition (Case 3) in Fig. 8.

Verification of condition (Case 4) in Fig. 9.

Histogram of RMSE of $X_1$ in Fig. 10.

Histogram of RMSE of $X_2$ in Fig. 11.

showed in Fig. 10 and Fig. 11, respectively. From the results, we can observe that the RMSE for $x_1$ and $x_2$ can focus on a small range of errors, which means the states can converge regardless of the values of the parameters within their boundaries, which indicates our control approach is robust.

VI. CONCLUSION

In this paper, we designed a robust controller for continuous-time nonlinear system with unknown perturbation by using ADP approach. Specifically, we converted the original control system into an auxiliary optimal control system with an appropriate performance index and solved for the corresponding stable optimal control policy. The equivalence of the transformation between these two systems was proved. ADP approach was applied to iteratively improve the performance index and control policy, and an adaptive critic neural network was employed in this method. An example was given to verify the effectiveness of the proposed approach.

REFERENCES


