Neuro-control of Nonlinear Systems with Unknown Input Constraints

Bo Zhao[1], Xinliang Liu[2(✉)], Derong Liu[3], and Yuanchun Li[4]

[1] The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
zhaobo@ia.ac.cn

[2] The Bureau of Informationization Development, Cyberspace Administration of China, Beijing 100044, China
lighterxl@icloud.com

[3] School of Automation, Guangdong University of Technology, Guangzhou 510006, China
derongliu@foxmail.com

[4] Department of Control Science and Engineering, Changchun University of Technology, Changchun 130012, China
liyc@ccut.edu.cn

Abstract. This paper establishes an adaptive dynamic programming algorithm based neuro-control scheme for nonlinear systems with unknown input constraints. The control strategy consists of an online nominal optimal control and a neural network (NN) based saturation compensator. For nominal systems without input constraints, we develop a critic NN to solve the Hamilton-Jacobi-Bellman equation. Hereafter, the online approximate nominal optimal control policy can be derived directly. Then, considering the unknown input constraints as saturation nonlinearity, NN based feed-forward compensator is employed. The ultimate uniform bounded stability of the closed loop system is analyzed via Lyapunov's direct method. Finally, simulation on a torsional pendulum system is provided to verify the effectiveness of the proposed control scheme.

Keywords: Adaptive dynamic programming · Unknown input constraints · Continuous-time nonlinear systems · Stabilizing control · Neural networks

1 Introduction

Optimal control problem has been paid considerable attention to nonlinear systems in the control community. To achieve this objective, adaptive dynamic programming (ADP) algorithm was developed [1] with the aid of NNs forward-in-time. In recent few years, ADP algorithms were developed further to deal with control problems for continuous-time systems [2], discrete-time systems [3],

D. Liu et al. (Eds.): ICONIP 2017, Part I, LNCS 10634, pp. 778–788, 2017.
https://doi.org/10.1007/978-3-319-70087-8_80

trajectory tracking [4], uncertainties [5] and external disturbances [6], time-delay [7], fault-tolerant [8], zero-sum games [9], event-triggered systems [10], etc.

Specially, input constraints are often emerged in many practical systems. They may cause control performance reduction or unstable of the closed-loop system. In order to solve these problems, some ADP methods have been proposed in recent years [11–13]. We can see that most of existing results were concerned with nonlinear systems subject to input constraints with available limited bounds, which are always necessary to design the ADP based control methods directly or indirectly. However, outputs of the actuators may bias or suddenly abrupt in many practical systems. It implies that the bounds of the actuators are uncertain or unknown, which results in application difficulties. Therefore, the optimal control strategy for nonlinear systems in this case is required to be further considered. Thus, this paper establishes an ADP based neuro-control scheme for nonlinear systems in the presence of unknown input constraints.

2 Problem Statement

Consider nominal continuous-time nonlinear systems as

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

where $x \in \mathbb{R}^n$ is the system state vector, $u \in \mathbb{R}^m$ is the control input vector, $f(\cdot)$ and $g(\cdot)$ are locally Lipschitz and differentiable in their arguments with the initial state $x(0) = x_0$ and $f(0) = 0$. System (1) is stable in the sense that there exists a continuous control u to stabilize the system asymptotically.

In the presence of unknown input constraints, (1) can be turned to

$$\dot{x} = f(x) + g(x)\tau, \tag{2}$$

where $\tau = [\tau_1, \tau_2, \ldots, \tau_m]^\mathsf{T} \in \mathbb{R}^m$ is the actual applied control input of (2). It arranges between its lower and upper limits, i.e.,

$$\tau_i = sat(u_i) = \begin{cases} u_{i\max}, & u_i > u_{i\max}, \\ u_i, & u_{i\min} \le u_i \le u_{i\max}, \\ u_{i\min}, & u_i < u_{i\min}, \end{cases} \tag{3}$$

where $i = 1, 2, \ldots, m$, $u_{i\max}$ and $u_{i\min}$ are the unknown upper and lower limit bounds, respectively.

The control objective is to develop an ADP based neuro-control scheme for nonlinear systems subject to unknown input constraints, and ensure all the signals of the closed-loop system to be ultimate uniform bounded (UUB).

3 Online Approximate Optimal Controller Design and Stability Analysis

3.1 Online Nominal Optimal Control

For nominal nonlinear system (1), a feedback control $u_n(x) \in \Psi(\Omega)$ will be designed by finding the stabilizing nominal control $u_n(x)$ which minimizes the infinite-horizon cost function given by

$$V(x_0) = \int_0^\infty U(x(s), u_n(s))ds, \tag{4}$$

where $U(x, u_n) = x^\mathsf{T} Q x + u_n^\mathsf{T} R u_n$ is the utility function, $U(0, 0) = 0$, and $U(x, u_n) \geq 0$ for all x and u_n, in which $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ are positive definite matrices. If the cost function (4) is continuously differentiable, then the infinitesimal versions of (4) is the so-called nonlinear Lyapunov equation

$$0 = U(x, u_n) + \nabla V(x)\dot{x}.$$

Define the Hamiltonian as

$$H\left(x, u_n, \nabla V(x)\right) = U(x, u_n) + \left(\nabla V(x)\right)^\mathsf{T}\left(f(x) + g(x)u_n\right). \tag{5}$$

The optimal cost $V^*(x)$ of (4) can be derived by solving the HJB equation

$$0 = \min_{u_n(x)\in\Psi(\Omega)} H\left(x, u_n, \nabla V^*(x)\right) \tag{6}$$

with $V^*(0) = 0$, and $\nabla V^*(x)$ denotes the partial derivative of the cost function $V^*(x)$ with respect to x, i.e., $\nabla V^*(x) = \frac{\partial V^*(x)}{\partial x}$. If the solution $V^*(x)$ of (6) exists, the optimal control can be obtained by

$$u_n^*(x) = -\frac{1}{2} R^{-1} g^\mathsf{T}(x) \nabla V^*(x). \tag{7}$$

By simple transformation of (7), we have

$$\left(\nabla V^*(x)\right)^\mathsf{T} g(x) = -2\left(u_n^*(x)\right)^\mathsf{T} R. \tag{8}$$

The cost function $V(x)$ will be approximated via NN as

$$V(x) = W_c^\mathsf{T} \sigma_c(x) + \varepsilon_c(x), \tag{9}$$

where $W_c \in \mathbb{R}^{l_1}$ is ideal weight vector, $\sigma_c(x) \in \mathrm{R}^{l_1}$ is the activation function, l_1 is the number of neurons in the hidden-layer, and $\varepsilon_c(x)$ is the approximation error of the NN. Then, the partial gradient of $V(x)$ along with x is

$$\nabla V(x) = \left(\nabla \sigma_c(x)\right)^\mathsf{T} W_c + \nabla \varepsilon_c(x), \tag{10}$$

where $\nabla \sigma_c(x) = \frac{\partial \sigma_c(x)}{\partial x} \in \mathbb{R}^{l_1 \times n}$ and $\nabla \varepsilon_c(x)$ are gradients of the activation function and the approximation error, respectively. Thus, the Hamiltonian can be expressed as

$$H\left(x, u_n, W_c\right) = U(x, u_n) + \left(W_c^\mathsf{T} \nabla \sigma_c(x) + \nabla \varepsilon_c(x)\right)\dot{x}. \tag{11}$$

Combining (6) with (11), we obtain

$$U(x, u_n) + W_c^\mathsf{T} \nabla \sigma_c(x)\dot{x} = e_{cH}, \tag{12}$$

where $e_{cH} = -\nabla\varepsilon_c(x)\dot{x}$ is the residual error caused by NN approximation.

The critic NN can be approximated by

$$\hat{V}(x) = \hat{W}_c^{\mathsf{T}}\sigma_c(x), \tag{13}$$

where $\hat{W}_c$ is the estimation of W_c. Then, we have the gradient of $\hat{V}(x)$ as

$$\nabla\hat{V}(x) = (\nabla\sigma_c(x))^{\mathsf{T}}\hat{W}_c. \tag{14}$$

Thus, the approximate Hamiltonian can be derived by

$$H\left(x, u_n, \hat{W}_c\right) = U(x, u_n) + \hat{W}_c^{\mathsf{T}}\nabla\sigma_c(x)\dot{x} = e_c. \tag{15}$$

Define $\tilde{W}_c = W_c - \hat{W}_c$. From (11) and (15), one has

$$e_c = e_{cH} - \tilde{W}_c^{\mathsf{T}}\nabla\sigma_c(x)\dot{x}. \tag{16}$$

To minimize the objective function $E_c = \dfrac{1}{2}e_c^{\mathsf{T}}e_c$, the weight approximation error should be updated by

$$\dot{\tilde{W}}_c = -\dot{\hat{W}}_c = l_c\left(e_{cH} - \tilde{W}_c^{\mathsf{T}}\theta\right)\theta, \tag{17}$$

where $\theta = \nabla\sigma_c(x)\dot{x}$. Thus, $\hat{W}_c$ can be updated by

$$\dot{\hat{W}}_c = -l_c e_c\theta, \tag{18}$$

where $l_c > 0$ is the learning rate of the critic NN.

Thus, according to (7) and (9), the ideal nominal control policy is

$$u_n(x) = -\frac{1}{2}R^{-1}g^{\mathsf{T}}(x)\left(\nabla\sigma_c^{\mathsf{T}}(x)W_c + \nabla\varepsilon_c(x)\right). \tag{19}$$

It can be approximated as

$$\hat{u}_n(x) = -\frac{1}{2}R^{-1}g^{\mathsf{T}}(x)\nabla\sigma_c^{\mathsf{T}}(x)\hat{W}_c. \tag{20}$$

Theorem 1. *For nonlinear system (1), if the weight vector of the critic NN is updated by (18), then the approximation error of the weight vector can be guaranteed to be UUB.*

Proof. Select the Lyapunov function candidate as

$$L_1 = \frac{1}{2l_c}\tilde{W}_c^{\mathsf{T}}\tilde{W}_c. \tag{21}$$

Its time derivative is

$$\begin{aligned}
\dot{L}_1 &= \frac{1}{l_c} \tilde{W}_c^{\mathsf{T}} \dot{\tilde{W}}_c \\
&= \tilde{W}_c^{\mathsf{T}} \left(e_{cH} - \tilde{W}_c^{\mathsf{T}} \theta \right) \theta \\
&= \tilde{W}_c^{\mathsf{T}} e_{cH} \theta - \left\| \tilde{W}_c^{\mathsf{T}} \theta \right\|^2 \\
&\leq \frac{1}{2} e_{cH}^2 - \frac{1}{2} \left\| \tilde{W}_c^{\mathsf{T}} \theta \right\|^2 .
\end{aligned} \tag{22}$$

Assume $\|\theta\| \leq \theta_M$, where $\theta_M > 0$. Hence, $\dot{L}_1 < 0$ as long as $\tilde{W}_c$ lies outside of the compact set

$$\Omega_{\tilde{W}_c} = \left\{ \tilde{W}_c : \left\| \tilde{W}_c \right\| \leq \left\| \frac{e_{cH}}{\theta_M} \right\| \right\}.$$

Therefore, according to Lyapunov's direct method, the approximation error of the weight vector is UUB. This completes the proof.

3.2 Neural Network Based Unknown Input Constraint Compensation

To tackle the affection of unknown input constraints, NN based compensator is designed in detail as feed-forward control loop in this subsection. The constraint nonlinearity $\delta(x) = u - \tau = [\delta_1, \delta_2, \ldots, \delta_m]^{\mathsf{T}} \in \mathbb{R}^m$ is introduced as

$$\begin{aligned}
u_{\delta_i(x)} &= u_i - \tau_i \\
&= \begin{cases} u_i - u_{i\max}, & u_i > u_{i\max}, \\ 0, & u_{i\min} \leq u_i \leq u_{i\max}, \\ u_i - u_{i\min}, & u_i < u_{i\min}, \end{cases}
\end{aligned} \tag{23}$$

where $i = 1, 2, \ldots, m$.

Thus, the constrained nonlinear system (2) can be transformed into

$$\dot{x} = f(x) + g(x)(u - \delta), \tag{24}$$

which is approximated by a backpropagation NN since $\delta(x)$ is unknown, and it can be expressed by

$$\delta(x) = W_\delta^{\mathsf{T}} \sigma_\delta(x) + \varepsilon_\delta(x), \tag{25}$$

where $W_\delta \in \mathbb{R}^{l_2}$ is the ideal weight vector, $\sigma_\delta(x) \in \mathbb{R}^{l_2}$ is the activation function, l_2 is the number of hidden-layer neurons, and $\varepsilon_\delta(x)$ is the approximation error.

Assumption 1. *The NN approximation error is bounded, i.e.,* $\|\varepsilon_\delta(x)\| \leq \varepsilon_{\delta M}$, *where* $\varepsilon_{\delta M}$ *is a positive constant.*

To determine the unknown weight vector W_δ, (25) is approximated by

$$\hat{\delta}(x) = \hat{W}_\delta^{\mathsf{T}} \sigma_\delta(x), \tag{26}$$

where $\hat{W}_\delta$ is the approximation of W_δ. It can be updated by

$$\dot{\hat{W}}_\delta = -\dot{\tilde{W}}_\delta = \mathit{\Gamma}_\delta \sigma_\delta(x) \left(2u_n^{\mathsf{T}} R - x^{\mathsf{T}} g(x)\right) + k\mathit{\Gamma}_\delta \|x\| \hat{W}_\delta, \tag{27}$$

where $\mathit{\Gamma}_\delta > 0$ and $k > 0$ are both NN learning rates.

Define $\tilde{\delta} = \delta - \hat{\delta}$ as the overall NN approximation error. Thus,

$$\tilde{\delta} = \tilde{W}_\delta^{\mathsf{T}} \sigma_\delta(x) + \varepsilon_\delta(x). \tag{28}$$

Assumption 2. *There exists positive constants δ_M and δ_m such that $\|W_\delta\| \leq \delta_M$ and $\left\|\tilde{W}_\delta\right\| \leq \delta_m$, respectively.*

Hence, the overall control law for nonlinear system (2) can be designed as

$$u = u_n + \hat{\delta}. \tag{29}$$

3.3 Stability Analysis

Theorem 2. *Consider the nonlinear system with unknown input constraints (2), the transformed dynamics (24), as well as the Assumptions 1 and 2, all the signals of the closed-loop system can be guaranteed to be UUB, if the overall control law is designed as (29), which is composed of the online nominal optimal control (19) and constraint compensation (26) with the update law (27).*

Proof. Select the Lyapunov function candidate as

$$L_2 = \frac{1}{2} x^{\mathsf{T}} x + V(x) + tr\left(\frac{1}{2} \tilde{W}_\delta^{\mathsf{T}} \mathit{\Gamma}_\delta^{-1} \tilde{W}_\delta\right), \tag{30}$$

where $tr(\cdot)$ is the trace of the matrix. The time derivative of (30) is

$$\begin{aligned} \dot{L}_2 &= x^{\mathsf{T}} \dot{x} + \dot{V}(x) + tr\left(\tilde{W}_\delta^{\mathsf{T}} \mathit{\Gamma}_\delta^{-1} \dot{\tilde{W}}_\delta\right) \\ &= x^{\mathsf{T}} \left(f(x) + g(x)(u - \delta)\right) + \dot{V}(x) + tr\left(\tilde{W}_\delta^{\mathsf{T}} \mathit{\Gamma}_\delta^{-1} \dot{\tilde{W}}_\delta\right). \end{aligned} \tag{31}$$

In the existence of constraint nonlinearity, for the second item of (31), we have

$$\dot{V}(x) = \nabla V^{\mathsf{T}}(x) \dot{x} = \nabla V^{\mathsf{T}}(x) \left(f(x) + g(x)u\right) - \nabla V^{\mathsf{T}}(x) g(x) \delta. \tag{32}$$

Define $\tilde{\delta} = \delta - \hat{\delta}$. Then, introduce the overall control law (29), (32) becomes

$$\dot{V}(x) = \nabla V^{\mathsf{T}}(x) \left(f(x) + g(x)u_n\right) - \nabla V^{\mathsf{T}}(x) g(x) \tilde{\delta}. \tag{33}$$

According to (6), (8) and (29), we have

$$\dot{L}_2 = x^\mathsf{T}\left(f(x) + g(x)u_n\right) - x^\mathsf{T} Q x - u_n^\mathsf{T} R u_n + \left(2u_n^\mathsf{T} R - x^\mathsf{T} g(x)\right)\tilde{\delta} + tr\left(\tilde{W}_\delta^\mathsf{T} \Gamma_\delta^{-1} \dot{\tilde{W}}_\delta\right). \tag{34}$$

Since $f(x)$ is locally Lipschitz, there exists positive constants D_f such that $\|f(x)\| \leq D_f \|x\|$. Assume that $\|g(x)\| \leq D_g$. Thus, (34) becomes

$$\dot{L}_2 \leq D_f \|x\|^2 + \frac{1}{2}\|x\|^2 + \frac{1}{2}D_g^2 \|u_n\|^2 - \lambda_{\min}(Q)\|x\|^2 - \lambda_{\min}(R)\|u_n\|^2 + \left(2u_n^\mathsf{T} R - x^\mathsf{T} g(x)\right)\tilde{\delta} + tr\left(\tilde{W}_\delta^\mathsf{T} \Gamma_\delta^{-1} \dot{\tilde{W}}_\delta\right), \tag{35}$$

where $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue of the matrix.

Combining (25), (26) and (27), we have

$$\begin{aligned}\dot{L}_2 \leq{} & D_f \|x\|^2 + \frac{1}{2}\|x\|^2 + \frac{1}{2}D_g^2 \|u_n\|^2 - \lambda_{\min}(Q)\|x\|^2 - \lambda_{\min}(R)\|u_n\|^2 \\ & + \left(2u_n^\mathsf{T} R - x^\mathsf{T} g(x)\right)\tilde{\delta} - tr\left(\tilde{W}_\delta^\mathsf{T}(\sigma_\delta(x)\left(2u_n^\mathsf{T} R - x^\mathsf{T} g(x)\right) + k\|x\|\hat{W}_\delta)\right) \\ ={} & -\left(\lambda_{\min}(Q) - D_f - \frac{1}{2}\right)\|x\|^2 - \left(\lambda_{\min}(R) - \frac{1}{2}D_g^2\right)\|u_n\|^2 \\ & + \left(2u_n^\mathsf{T} R - x^\mathsf{T} g(x)\right)\varepsilon_\delta(x) - k\|x\|\, tr\left(\tilde{W}_\delta^\mathsf{T}\left(W_\delta - \tilde{W}_\delta\right)\right).\end{aligned} \tag{36}$$

According to Assumptions 1 and 2, and suppose that $\left\|2u_n^\mathsf{T} R - x^\mathsf{T} g(x)\right\| \leq \upsilon$, (36) becomes

$$\dot{L}_2 \leq -\left(\lambda_{\min}(Q) - D_f - \frac{1}{2}\right)\|x\|^2 - \left(\lambda_{\min}(R) - \frac{1}{2}D_g^2\right)\|u_n\|^2 + \varepsilon_{\delta M}\upsilon - k\|x\|\left(\delta_M\delta_m - \delta_m^2\right). \tag{37}$$

Let $A = \lambda_{\min}(Q) - D_f - \frac{1}{2}$, $B = k\left(\delta_M\delta_m - \delta_m^2\right)$. From (37), we can conclude that $\dot{L}_2 \leq 0$ when the state x lies outside of the compact set $\Omega_x = \left\{x : \|x\| \leq \frac{-B+\sqrt{B^2+4A\varepsilon_{\delta M}\upsilon}}{2A}\right\}$ with $\lambda_{\min}(Q) > D_f + \frac{1}{2}$ and $\lambda_{\min}(R) \geq \frac{1}{2}D_g^2$ hold. It implies that all the signals of the closed-loop system with unknown actuator saturation can be guaranteed to be UUB. This completes the proof.

4 Simulation Study

Consider a torsional pendulum system [14] with unknown constrained control input $\tau \in \mathbb{R}$, which is described as

$$\begin{cases} \dfrac{\mathrm{d}\theta}{\mathrm{d}t} = \omega, \\ J\dfrac{\mathrm{d}\omega}{\mathrm{d}t} = \tau - Mgl\sin\theta - f_d\dfrac{\mathrm{d}\theta}{\mathrm{d}t}, \end{cases}$$

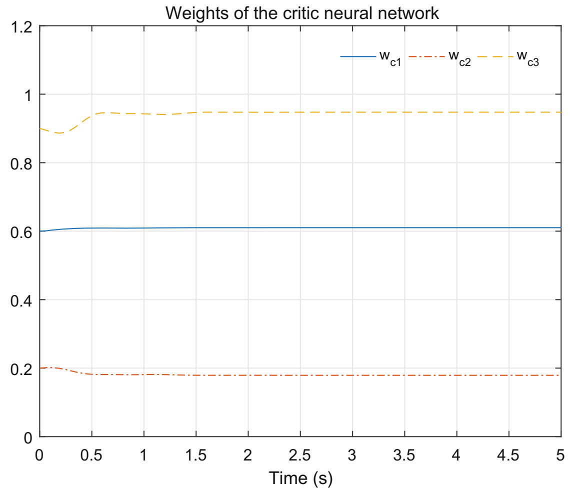

Fig. 1. Weights of critic neural network

where $M = \frac{1}{3}\,\mathrm{kg}$ and $l = \frac{2}{3}\,\mathrm{m}$ are the mass and length of the pendulum bar, respectively. The angle θ and the angular velocity ω are the system states. Let $J = \frac{4}{3}Ml^2$, $f_d = 0.2$ and $g = 9.8\,\mathrm{m/s^2}$ be the rotary inertia, frictional factor and gravitational acceleration, respectively. $\tau \in \mathbb{R}$ is the actual applied control input with unknown constraint. In this simulation, it is chosen as

$$\tau = sat(u) = \begin{cases} 0.1, & u > 0.1, \\ u, & -0.1 \le u \le 0.1, \\ -0.1, & u < -0.1. \end{cases} \tag{38}$$

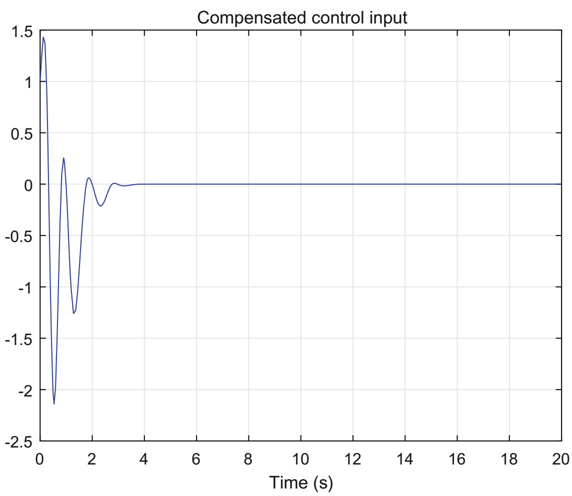

Fig. 2. Compensated control input

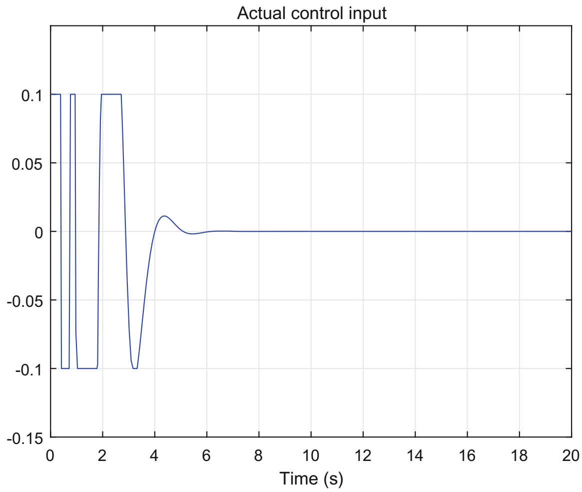

Fig. 3. Actual control input

Define $x = [x_1, x_2]^\mathsf{T} = [\theta, \omega]^\mathsf{T} \in \mathbb{R}^2$ as the state vector of the torsional pendulum system, whose initial state vector be $x_0 = [1, -1]^\mathsf{T}$. In this simulation, the cost function is approximated by a critic NN, whose weight vector is denoted as $\hat{W}_c = [\hat{W}_{c1}, \hat{W}_{c2}, \hat{W}_{c3}]^\mathsf{T}$, and its initial value is chosen as $\hat{W}_{c0} = [0.6, 0.2, 0.9]^\mathsf{T}$. The activation function of the critic NN is chosen as $\sigma_c(x) = \left[x_1^2,\ x_1 x_2,\ x_2^2\right]$. Let $Q = 10I_2$ and $R = 10I$, where I_n denotes identity matrix with n dimensions, the learning rate of the critic NN be $l_c = 0.0002$, the learning rates of the NN for constraint compensator be $\Gamma_\delta = 0.01$ and $k = 1$, respectively.

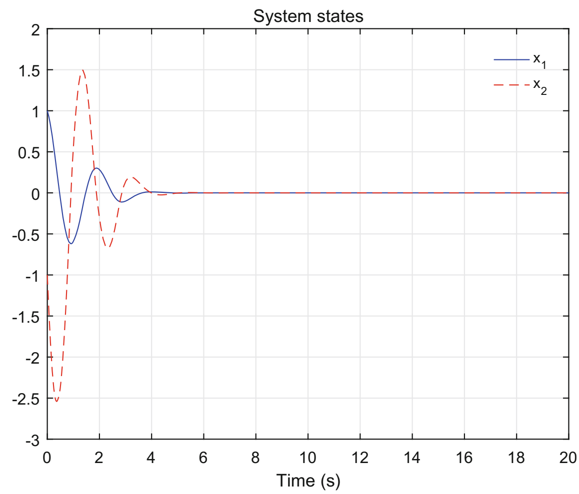

Fig. 4. System states

The simulation results are shown as Figs. 1, 2, 3 and 4. From Fig. 1, we can see that the weights of the critic NN converge to $[0.6108, 0.1764, 0.9576]^{\mathsf{T}}$. As shown in Fig. 2, the NN based feed-forward compensator (26) is employed to overcome the negative effect brings by the unknown input constraints. We can see that the actual control input illustrated in Fig. 3 is limited within the bounded values. Fortunately, with the proposed ADP based stabilizing control scheme, the system states are still convergence as Fig. 4.

5 Conclusion

In this paper, a neuro-control scheme of nonlinear systems with unknown input constraints was developed by using NN compensation based ADP algorithm. This strategy is utilized to solve the stabilizing problem without any priori knowledge of the bounds of input constraints, as well as the initial stabilizing control and the persisting of excitation condition, which are always required in traditional ADP methods.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grants 61603387, U1501251, 61533017, 61773075, 61374051, 61374105 and 61503379, in part by the Scientific and Technological Development Plan Project in Jilin Province of China under Grants 20150520112JH and 20160414033GH, and in part by Beijing Natural Science Foundation under Grant 4162065.

References

1. Werbos, P.J.: Approximate dynamic programming for real time control and neural modeling. In: White, D.A., Sofge, D.A. (eds.) Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches. Van Nostrand Reinhold, New York (1992)
2. Wang, D., He, H., Zhao, B., Liu, D.: Adaptive near-optimal controllers for nonlinear decentralised feedback stabilisation problems. IET Control Theory Appl. **11**(6), 799–806 (2017)
3. Lin, Q., Wei, Q., Zhao, B.: Optimal control for discrete-time systems with actuator saturation. Optim. Control Appl. Methods (2017). doi:10.1002/oca.2313
4. Zhao, B., Liu, D., Yang, X., Li, Y.: Observer-critic structure-based adaptive dynamic programming for decentralised tracking control of unknown large-scale nonlinear systems. Int. J. Syst. Sci. **48**(9), 1978–1989 (2017)
5. Gao, W., Jiang, Y., Jiang, Z.P., Chai, T.: Output-feedback adaptive optimal control of interconnected systems based on robust adaptive dynamic programming. Automatica **72**, 37–45 (2016)
6. Fan, Q., Yang, G.: Adaptive actor-critic design-based integral sliding-mode control for partially unknown nonlinear systems with input disturbances. IEEE Trans. Neural Netw. Learn. Syst. **27**(1), 165–177 (2016)
7. Zhang, H., Song, R., Wei, Q., Zhang, T.: Optimal tracking control for a class of nonlinear discrete-time systems with time delays based on heuristic dynamic programming. IEEE Trans. Neural Netw. **22**(12), 1851–1862 (2011)

8. Zhao, B., Liu, D., Li, Y.: Online fault compensation control based on policy iteration algorithm for a class of affine non-linear systems with actuator failures. IET Control Theory Appl. **10**(15), 1816–1823 (2016)
9. Fu, Y., Fu, J., Chai, T.: Robust adaptive dynamic programming of two-player zero-sum games for continuous-time linear systems. IEEE Trans. Neural Netw. Learn. Syst. **26**(12), 3314–3319 (2015)
10. Wang, D., Mu, C., He, H., Liu, D.: Event-driven adaptive robust control of nonlinear systems with uncertainties through NDP strategy. IEEE Trans. Syst. Man Cybern. Syst. (2016). doi:10.1109/TSMC.2016.2592682
11. Abu-Khalaf, M., Lewis, F.L., Huang, J.: Neurodynamic programming and zero-sum games for constrained control systems. IEEE Trans. Neural Netw. **19**(7), 1243–1252 (2008)
12. Heydari, A., Balakrishnan, S.N.: Finite-horizon control-constrained nonlinear optimal control using single network adaptive critics. IEEE Trans. Neural Netw. Learn. Syst. **24**(1), 145–157 (2013)
13. Liu, D., Yang, X., Wang, D., Wei, Q.: Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints. IEEE Trans. Cybern. **45**(7), 1372–1385 (2015)
14. Zhao, B., Liu, D., Li, Y.: Observer based adaptive dynamic programming for fault tolerant control of a class of nonlinear systems. Inf. Sci. **384**, 21–33 (2017)