Finite Horizon Optimal Tracking Control for a Class of Discrete-Time Nonlinear Systems

Qinglai Wei, Ding Wang, and Derong Liu*

Institute of Automation, Chinese Academy of Sciences, 100190, China
{qinglai.wei,ding.wang,derong.liu}@ia.ac.cn

Abstract. In this paper, a new iterative ADP algorithm is proposed to solve the finite horizon optimal tracking control problem for a class of discrete-time nonlinear systems. The idea is that using system transformation, the optimal tracking problem is transformed into optimal regulation problem, and then the iterative ADP algorithm is introduced to deal with the regulation problem with convergence guarantee. Three neural networks are used to approximate the performance index function, compute the optimal control policy and model the unknown system dynamics, respectively, for facilitating the implementation of iterative ADP algorithm. An example is given to demonstrate the validity of the proposed optimal tracking control scheme.

Keywords: Adaptive dynamic programming, approximate dynamic programming, optimal tracking control, neural networks, finite horizon.

1 Introduction

The optimal tracking problem of nonlinear systems has always been the key focus in the control field in the latest several decades. Traditional optimal tracking control is mostly implemented by feedback linearization [1]. However, the controller designed by feedback linearization technique is only effective in the neighborhood of the equilibrium point. When the required operating range is large, the nonlinearities in the system cannot be properly compensated by using a linear model. Therefore, it is necessary to study the direct optimal tracking control approach for the original nonlinear system. The difficulty for nonlinear optimal feedback control lies in solving the time-varying HJB equation which is usually too hard to solve analytically. In order to overcome the difficulty, in [2], the finite-time optimal tracking control problem was solved via transforming the system model into a sequence of "pseudo-linear" systems. In [3], an infinite horizon approximate optimal tracking controller based on the successive approximation approach was proposed. However, the literature mentioned above is all restricted in the continuous-time domain. There are few results discussing the optimal tracking control problem for discrete-time systems. To the best of our knowledge, only [4] has presented the optimal tracking control scheme in infinite horizon domain. There are no results on the finite horizon optimal tracking control for discrete-time nonlinear systems. This motivates our research.

* This work was supported in part by the NSFC under grants 60904037, 60921061, 61034002, and by Beijing Natural Science Foundation under grant 4102061.

D. Liu et al. (Eds.): ISNN 2011, Part II, LNCS 6676, pp. 620–629, 2011.

As is known, dynamic programming is very useful in solving the optimal control problems. However, due to the "curse of dimensionality", it is often computationally untenable to run dynamic programming to obtain the optimal solution. The approximate dynamic programming (ADP) algorithm was proposed by Werbos [5] as a way to solve optimal control problems forward-in-time. ADP combines adaptive critic design, reinforcement learning technique with dynamic programming. In [5] adaptive dynamic programming approaches were classified into four main schemes: Heuristic Dynamic Programming (HDP), Dual Heuristic Dynamic Programming (DHP), Action Dependent Heuristic Dynamic Programming (ADHDP), also known as Q-learning, and Action Dependent Dual Heuristic Dynamic Programming (ADDHP). Though in recent years, ADP has been further studied by many researchers [6, 7, 8, 9, 10, 12, 11], most results are focus on the optimal regulation problem. In [13], a greedy HDP iteration algorithm to solve the discrete-time Hamilton-Jacobi-Bellman (DT HJB) equation of the optimal regulation control problem for general nonlinear discrete-time systems is proposed, which does not require an initially stable policy. It has been rigorously proved in [13] that the greedy HDP iteration algorithm is convergent. To the best of our knowledge, till now only in [4], ADP was used to solve the infinite-time optimal tracking control problem. There have been no results discussing how to use ADP to solve the finite-time optimal tracking control problem for nonlinear systems.

In this paper, it is the first time to solve finite horizon optimal tracking control problem for a class of discrete-time nonlinear systems using ADP. We firstly transform the tracking problem into an optimal regulation problem, and then a new iterative ADP algorithm can be properly introduced to deal with this regulation problem.

2 Paper Preparation

Consider the following discrete-time nonlinear system

$$x_{k+1} = f(x_k) + g(x_k)u_k \tag{1}$$

where $x_k \in \Re^n$ and the input $u_k \in \Re^m$. Here assume that the system is controllable. In this paper, the reference orbit η_k is generated by the n-dimensional autonomous system as $\eta_{k+1} = S(\eta_k)$, where $\eta_k \in \Re^n$, $S(\eta_k) \in \Re^n$. Therefore we define the tracking error as:

$$z_k = x_k - \eta_k. \tag{2}$$

Let $\underline{v}_k$ be an arbitrary finite-horizon tracking control sequence starting at k and let $\mathcal{U}_{z_k} = \left\{\underline{v}_k \colon z^{(f)}\left(z_k, \underline{v}_k\right) = 0\right\}$ be the set of all finite-horizon tracking control sequences of x_k. Let $\mathcal{U}_{z_k}^{(i)} = \left\{\underline{v}_k^{k+i-1} \colon z^{(f)}\left(z_k, \underline{v}_k^{k+i-1}\right) = 0,\ \left|\underline{v}_k^{k+i-1}\right| = i\right\}$ be the set of all finite-horizon admissible control sequences of z_k with length i, where the final state error can be written as $z^{(f)}\left(z_k, \underline{v}_k^{k+i-1}\right) = z_{k+i}$. Then, $\mathcal{U}_{z_k} = \cup_{1 \le i < \infty}\mathcal{U}_{z_k}^{(i)}$. By this notation, a state error z_k is controllable if and only if $\mathcal{U}_{z_k} \neq \emptyset$.

Noticing that the objective in this paper is to design an optimal feedback control policy v_k, which not only renders the state error z_k asymptotically tracking the reference orbit, $i.e.$, z_k asymptotically approaches zero, but also minimizes the performance index function as follow

$$J(z_k, \underline{v}_k^{N-1}) = \sum\nolimits_{i=k}^{N-1} \left\{ z_i^T Q z_i + v_i^T R v_i \right\}, \tag{3}$$

where Q and R are positive-definite matrices. $U(k) = z_k^T Q z_k + v_k^T R v_k$ is the utility function. In addition, we define

$$v_k = u_k - u_{ek}, \tag{4}$$

where u_{ek} is the steady control input expressed as

$$u_{ek} = g^{-1}(\eta_k)(\eta_{k+1} - f(\eta_k)) \tag{5}$$

Combining (2) with (5) , we can get

$$\begin{aligned} z_{k+1} = F(z_k, v_k) = &-S(\eta_k) + f(z_k + \eta_k) + g(z_k + \eta_k)\, v_k \\ &-g(z_k + \eta_k) g^{-1}(\eta_k)(f(\eta_k) - S(\eta_k)). \end{aligned} \tag{6}$$

For any given system state error z_k, the objective of the present finite-horizon optimal control problem is to find a finite-horizon admissible control sequence $\underline{v}_k^{N-1} \in \mathcal{U}_{z_k}^{(N-k)} \subseteq \mathcal{U}_{x_k}$ to minimize the performance index $J\big(z_k, \underline{v}_k^{N-1}\big)$. The control sequence $\underline{v}_k^{N-1}$ has finite length. However, before it is determined, we do not know its length which means that the length of the control sequence $\left|\underline{v}_k^{N-1}\right| = N - k$ is unspecified. This kind of optimal control problems has been called finite-horizon problems with unspecified terminal time.

3 Properties of the Iterative Adaptive Dynamic Programming Algorithm

In this section, a new iterative ADP algorithm is proposed to obtain the finite horizon optimal tracking control for nonlinear systems. The goal of the proposed iterative ADP algorithm is to construct an optimal control policy $v^*(z_k)$, $k = 0, 1, \ldots$, which makes an arbitrary initial state error z_0 to the singularity 0 within finite time, simultaneously makes the performance index function reach the optimum $V^*(z_k)$. Convergence proofs will also be given.

3.1 Derivation of the Iterative ADP Algorithm

In the iterative ADP algorithm, the performance index function and control policy are updated by recurrent iteration, with the iteration number i increasing from 0. Let the initial performance index function $V_0(z_k) = 0$ and there exists a control v_k that makes $F(z_k, v_k) = 0$, where z_k is any initial state error. Then, the iterative control $v_0(z_k)$ can be computed as follows:

$$\begin{aligned} v_0(z_k) &= \arg\min_{v_k} \left\{ U(z_k, v_k) + V_0(z_{k+1}) \right\}, \\ &\text{s.t. } z_{k+1} = F(z_k, v_k) = 0 \end{aligned} \tag{7}$$

where $V_0(z_{k+1}) = 0$. The performance index function can be updated as

$$V_1(z_k) = U(z_k, v_0(z_k)) + V_0(F(z_k, v_0(z_k)). \tag{8}$$

For $i = 1, 2, \ldots$, the iterative ADP algorithm will iterate between

$$\begin{aligned} v_i(z_k) &= \arg\min_{v_k} \{U(z_k, v_k) + V_i(z_{k+1})\} \\ &= \arg\min_{v_k} \{U(z_k, v_k) + V_i(F(z_k, v_k))\} \end{aligned} \tag{9}$$

and performance index function

$$\begin{aligned} V_{i+1}(z_k) &= \min_{v_k} \{U(z_k, v_k) + V_i(z_{k+1})\} \\ &= U(z_k, v_i(z_k)) + V_i(F(z_k, v_i(z_k)). \end{aligned} \tag{10}$$

3.2 Properties of the Iterative ADP Algorithm

In the above, we can see that the performance index function $V^*(z_k)$ is replaced by a sequence of iterative performance index functions $V_i(z_k)$ and the optimal control law $v^*(z_k)$ is replaced by a sequence of iterative control law $v_i(z_k)$, where $i \geq 0$ is the iterative index. As (10) is not an HJB equation for $\forall i \geq 0$, generally, the iterative performance index function $V_i(z_k)$ is not optimal. However, we can prove that $V^*(z_k)$ is the limit of $V_i(z_k)$ as $i \to \infty$.

Theorem 1. *Let z_k be an arbitrary state error vector. Suppose that there is a positive integer i such that $\mathcal{U}_{z_k}^{(i)} \neq \emptyset$. Then, for $\mathcal{U}_{z_k}^{(i+1)} \neq \emptyset$, the performance index function $V_i(z_k)$ obtained by (7)–(10) is a nonincreasing convergent sequence for $\forall\, i \geq 1$, i.e., $V_{i+1}(z_k) \leq V_i(z_k)$.*

Proof. We prove this by mathematical induction. First, we let $i = 1$. Then, We have

$$\begin{aligned} V_1(z_k) &= \min_{v_k} \{U(z_k, v_k) + V_0(F(z_k, v_k))\} \\ &= \min_{v_k} \{U(z_k, v_k)\} = U(z_k, v_0(z_k)) \end{aligned} \tag{11}$$

where $V_0(F(z_k, v_0(z_k)) = 0$. The finite horizon admissible control sequence $\underline{v}_k^k = (v_0(z_k))$.

Next, let us show that there exists a finite horizon admissible control sequence $\hat{\underline{v}}_k^{k+1}$ with length 2 such that $V_1(z_k, \underline{v}_k^k) = \hat{V}_2(z_k, \hat{\underline{v}}_k^{k+1})$. Obviously, $v_0(z_k) \in \mathcal{U}_{z_k}^{(1)}$. The trajectory starting from z_k under the control of $\underline{v}_k^k$ is $z_{k+1} = F(z_k, v_0(z_k)) = 0$. Then, we create a new control sequence $\hat{\underline{v}}_k^{k+1}$ by adding a 0 at the end of sequence $\underline{v}_k^k$ to obtain the control sequence $\hat{\underline{v}}_k^{k+1} = (\underline{v}_k^k, 0)$. Obviously, $|\hat{\underline{v}}_k^{k+1}| = 2$. The state error trajectory under the control of $\hat{\underline{v}}_k^{k+1}$ is $z_{k+1} = F(z_k, v_0(z_k))$, $z_{k+2} = F(z_{k+1}, v_{k+1})$ where $v_{k+1} = 0$. As $z_{k+1} = 0$ and $F(0, 0) = 0$, we have $z_{k+2} = F(z_{k+1}, v_{k+1}) = 0$. So, $\hat{\underline{v}}_k^{k+1}$ is a finite horizon admissible control. Furthermore,

$$\begin{aligned} V_1(z_k, \hat{\underline{v}}_k^k) &= U(z_k, v_k) \\ &= U(z_k, v_k) + U(z_{k+1}, v_{k+1}) = \hat{V}_2(z_k, \hat{\underline{v}}_k^{k+1}). \end{aligned} \tag{12}$$

On the other hand, we have

$$V_2(z_k) = \min_{v_k} \{U(z_k, v_k) + V_1(F(z_k, v_k)\} . \tag{13}$$

According to (11), we have

$$V_2(z_k) = \min_{\underline{v}_k^{k+1}} \{U(z_k, v_k) + U(z_{k+1}, v_{k+1})\} \tag{14}$$

where $z_{k+2} = F(z_{k+1}, v_{k+1}) = 0$. Then we have

$$V_2(z_k) \leq \hat{V}_2(z_k, \hat{\underline{v}}_k^{k+1}). \tag{15}$$

So the theorem holds for $i = 1$. Assume that the theorem holds for any $i = l - 1$, where $l \geq 1$. We have

$$V_l(z_k) = \min_{v_k} \{U(z_k, v_k) + V_{l-1}(F(z_k, v_k))\} \tag{16}$$

where the corresponding finite horizon admissible control sequence is $\underline{v}_k^{k+l-1}$.

Then for $i = l$, we create a control sequence $\hat{\underline{v}}_k^{k+l} = \left\{\underline{v}_k^{k+l-1}, 0\right\}$ with length $l + 1$. Then the state error trajectory under the control of $\hat{\underline{v}}$ is $z_{k+1} = F(z_k, v_l(z_k))$, $z_{k+2} = F(z_{k+1}, v_{l-1}(z_{k+1}))$, . . ., $z_{k+l} = F(z_{k+l}, v_0(z_{k+l})) = 0$, $z_{k+l+1} = 0$. So $\hat{\underline{v}}_k^{k+l}$ is finite horizon admissible control. The performance under the control sequence is

$$\begin{aligned} V_{l+1}(z_k, \hat{\underline{v}}_k^{k+l}) =& U(z_k, v_l(z_k)) + U(z_{k+1}, v_{l-1}(z_{k+1})) \\ &+ \ldots + U(z_{k+l}, v_0(z_{k+l})) + U(z_{k+l+1}, 0) \\ =& \sum_{j=0}^{l+1} U(z_{k+j}, v_{i-j}(z_{k+j})) \end{aligned} \tag{17}$$

where $v_{l-j} = 0$ for all $l < j$.

On the other hand, we have

$$V_{i+1}(z_k) = \min_{v_k} \{U(z_k, v_k) + V_i(F(z_k, v_k)\} = \min_{\underline{v}_k^{k+i}} \left\{ \sum_{j=0}^{i+1} U(z_{k+j}, v_{i-j}(z_{k+j})) \right\}. \tag{18}$$

Then, we have

$$V_{l+1}(z_k) \leq V_{l+1}(z_k, \hat{\underline{v}}_k^{k+l}) = V_l(z_k) \tag{19}$$

The proof is completed.

Lemma 1. *Let $\mu_i(z_k), i = 0, 1 \ldots$ be any sequence of tracking control, and $v_i(z_k)$ is expressed as (9). Define $V_{i+1}(z_k)$ as (10) and $\Lambda_{i+1}(z_k)$ as*

$$\Lambda_{i+1}(z_k) = U(z_k, \mu_i(z_k)) + \Lambda_i(z_{k+1}). \tag{20}$$

Then if $V_0(z_k) = \Lambda_0(z_k) = 0$, we have $V_i(z_k) \leq \Lambda_i(z_k)$, $\forall i$.

According to Theorem 1, we know that the performance index function $V_i(z_k) \geq 0$ is a nonincreasing bounded sequence for iteration index $i = 1, 2, \ldots$. Then we can derive the following theorem.

Theorem 2. *Let z_k be an arbitrary state error vector. Define the performance index function $V_\infty(z_k)$ as the limit of the iterative function $V_i(z_k)$, i.e.,*

$$V_\infty(z_k) = \lim_{i \to \infty} V_i(z_k). \tag{21}$$

Then, we have the following HJB equation

$$V_\infty(z_k) = \min_{v_k}\{U(z_k, v_k) + V_\infty(z_{k+1})\} \tag{22}$$

holds.

Proof. Let $\eta_k = \eta(z_k)$ be any admissible control. According to Theorem 1, for $\forall i$, we have

$$V_\infty(z_k) \leq V_{i+1}(z_k) \leq U(z_k, \eta_k) + V_i(z_{k+1}). \tag{23}$$

Let $i \to \infty$, we have

$$V_\infty(z_k) \leq U(z_k, \eta_k) + V_\infty(z_{k+1}). \tag{24}$$

So

$$V_\infty(z_k) \leq \min_{v_k}\{U(z_k, \eta_k) + V_\infty(z_{k+1})\}. \tag{25}$$

Let $\epsilon > 0$ be an arbitrary positive number. Since $V_i(z_k)$ is nonincreasing for $\forall\, i$ and $\lim_{i\to\infty} V_i(z_k) = V_\infty(z_k)$, there exists a positive integer p such that

$$V_p(z_k) - \epsilon \leq V_\infty(z_k) \leq V_p(z_k). \tag{26}$$

Then, we let

$$\begin{aligned} V_p(z_k) &= \min_{v_k}\{U(z_k, v_k) + V_p(z_{k+1})\} \\ &= U(z_k, v_{p-1}(z_k)) + V_{p-1}(z_{k+1}). \end{aligned} \tag{27}$$

Hence

$$\begin{aligned} V_\infty(z_k) &\geq U(z_k, v_{p-1}(z_k)) + V_{p-1}(z_{k+1}) - \epsilon \\ &\geq U(z_k, v_{p-1}(z_k)) + V_\infty(z_{k+1}) - \epsilon \\ &\geq \min_{v_k}\{U(z_k, v_k) + V_\infty(z_{k+1})\} - \epsilon. \end{aligned} \tag{28}$$

Since ϵ is arbitrary, we have

$$V_\infty(z_k) \geq \min_{v_k}\{U(z_k, v_k) + V_\infty(z_{k+1})\}. \tag{29}$$

Combining (25) and (29) we have

$$V_\infty(z_k) = \min_{v_k}\{U(z_k, v_k) + V_\infty(z_{k+1})\} \tag{30}$$

which proves the theorem.

Next, we will prove that the iterative performance index function $V_i(z_k)$ converges to the optimal performance index function $V^*(z_k)$ as $i \to \infty$.

Theorem 3. *Let the performance index function $V_i(z_k)$ be defined by (10). If the system state error z_k is controllable, then the performance index function $V_i(z_k)$ converges to the optimal performance index function $V^*(z_k)$ as $i \to \infty$, i.e.,*

$$V_i(z_k) \to V^*(z_k). \tag{31}$$

Proof. As

$$V^*(z_k) = \min\left\{V(z_k, \underline{v}_k) \colon \underline{v}_k \in \mathcal{U}_{z_k}^{(i)}\right\}, \;\; i = 1, 2, \ldots . \tag{32}$$

we have

$$V^*(z_k) \leq V_i(z_k). \tag{33}$$

Then, let $i \to \infty$, we have

$$V^*(z_k) \leq V_\infty(z_k). \tag{34}$$

Let $\epsilon > 0$ be an arbitrary positive number. Then there exists a finite horizon admissible control sequence η_q such that

$$V_q(z_k) \leq V^*(z_k) + \epsilon. \tag{35}$$

On the other side, according to Lemma 1, for any finite horizon admissible control η_q, we have

$$V_\infty(z_k) \leq V_q(z_k) \tag{36}$$

holds.

Combining (35) and (36), we have

$$V_\infty(z_k) \leq V^*(z_k) + \epsilon. \tag{37}$$

As ϵ is arbitrary positive number, we have

$$V_\infty(z_k) \leq V^*(z_k). \tag{38}$$

According to (34) and (38), we have

$$V_\infty(z_k) = V^*(z_k). \tag{39}$$

The proof is completed.

Then we can derive the following corollary.

Corollary 1. *Let the performance index function $V_i(z_k)$ be defined by (10). If the system state error z_k is controllable and Theorem 3 holds, then the iterative control law $v_i(z_k)$ converges to the optimal control law $v^*(z_k)$.*

3.3 The Procedure of the Algorithm

Now we summarize the iterative ADP algorithm for the time-variant optimal tracking control problem as:

Step 1. Give $x(0)$, $i_{\max}$, ε, desired trajectory η_k.
Step 2. Set $i = 0$, $V_0(z_k) = 0$.
Step 3. Compute $v_0(z_k)$ by (7) and $V_1(z_k)$ by (8).
Step 4. Set $i = i + 1$.
Step 5. Compute $v_i(z_k)$ by (9) and $V_{i+1}(z_k)$ by (10).
Step 6. If $|V_{i+1}(z_k) - V_i(z_k)| < \varepsilon$ then go to step 8, else go to step 7.
Step 7. If $i > i_{\max}$ then go to step 8, otherwise go to step 6.
Step 8. Stop.

4 Simulation Study

Consider the following affine nonlinear system

$$x_{k+1} = f(x_k) + g(x_k)u_k \tag{40}$$

where $x_k = \begin{bmatrix} x_{1k} & x_{2k} \end{bmatrix}^T$, $u_k = \begin{bmatrix} u_1(k) & u_2(k) \end{bmatrix}^T$,

$$f(x_k) = \begin{bmatrix} 0.2x_{1k}\exp(x_{2k}^2) \\ 0.3x_{2k}^3 \end{bmatrix}, \; g(x_k) = \begin{bmatrix} -0.2 & 0 \\ 0 & -0.2 \end{bmatrix}.$$

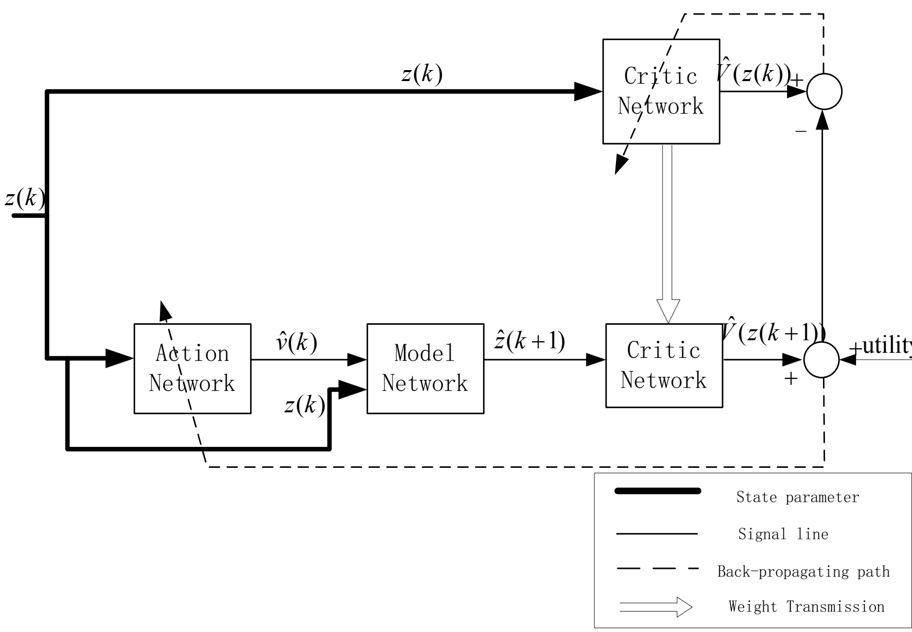

Fig. 1. The structure diagram of the algorithm

The desired trajectory is set to $\eta(k) = [\sin(k + \frac{\pi}{2}) \quad 0.5\cos(k)]^T$. We use neural network to implement the iterative ADP algorithm. We choose three-layer neural networks as the critic network, the action network and the model network with the structure 2-8-1, 2-8-2 and 6-8-2 respectively. The initial weights of action network, critic network and model network are all set to be random in $[-1, 1]$. It should be mentioned that the model network should be trained first. For the given initial state $x(0) = [1.5 \quad 1]^T$, we train the model network for 10000 steps under the learning rate $\alpha_m = 0.05$. After the training of the model network completed, the weights keep unchanged. Then the critic network and the action network are trained for 5000 steps so that the given accuracy $\varepsilon = 10^{-6}$ is reached. In the training process, the learning rate $\beta_a = \alpha_c = 0.05$. The structure diagram of the algorithm is shown in Fig. 1.

The convergence curve of the performance index function is shown in Fig.2(a). The state trajectories are given as Fig. 2(b) and Fig. 2(c). The corresponding control curves are given as Fig. 2(d).

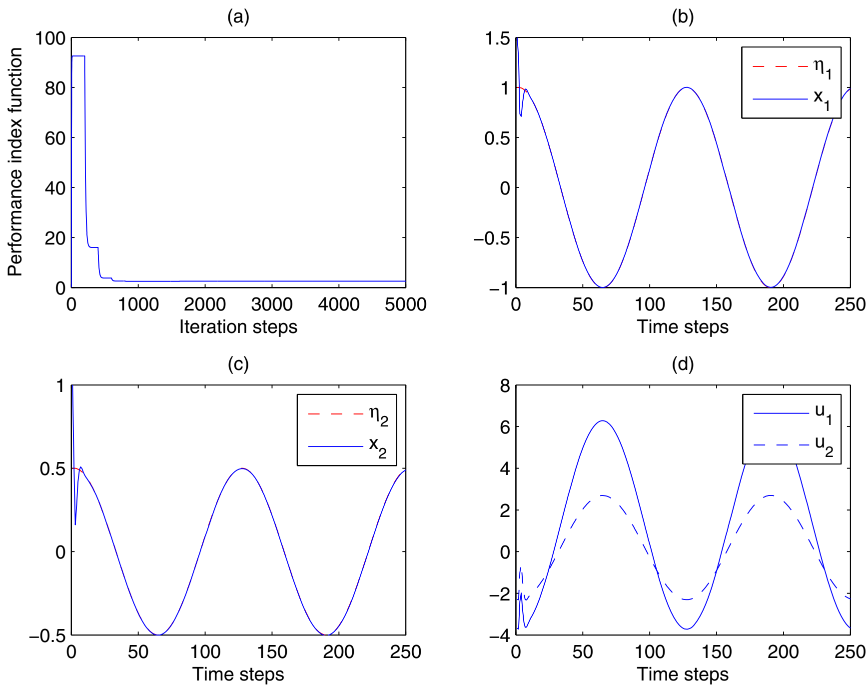

Fig. 2. The results of the algorithm

5 Conclusions

In this paper we propose an effective algorithm to solve the optimal finite horizon tracking control problem for a class of discrete-time systems. First, the tracking problems are transformed as regulation problem. Then the iterative ADP algorithm is introduced to deal with the regulation problem with rigorous convergence analysis. Three neural networks are used as parametric structures to approximate the performance index function, compute the optimal control policy and model the unknown system respectively, i.e. the critic network, the action network and the model network. The construction of model network make the scheme can be use to control the plant with unknown dynamics. The simulation study have successfully demonstrated the upstanding performance of the proposed tracking control scheme for various discrete-time nonlinear systems.

References

1. Ha, I.J., Gilbert, E.G.: Robust tracking in nonlinear systems. IEEE Transactions Automatic Control 32, 763–771 (1987)
2. Cimen, T., Banks, S.P.: Nonlinear optimal tracking control with application to super-tankers for autopilot design. Automatica 40, 1845–1863 (2004)
3. Gao, D., Tang, G., Zhang, B.: Approximate optimal tracking control for a class of nonlinear systems with disturbances. In: Proceedings of 6th World Congress on Intelligent Control and Automation, Dalian, China, vol. 1, pp. 521–525 (2006)
4. Zhang, H.G., Wei, Q.L., Luo, Y.H.: A novel infinite-time optimal tracking control scheme for a class of discrete-time nonlinear systems via the greedy HDP iteration algorithm. IEEE Transactions on System, Man, and cybernetics-Part B: Cybernetics 38, 937–942 (2008)
5. Werbos, P.J.: A menu of designs for reinforcement learning over time. In: Miller, W.T., Sutton, R.S., Werbos, P.J. (eds.) Neural Networks for Control, pp. 67–95. MIT Press, Cambridge (1991)
6. Liu, D.R., Zhang, Y., Zhang, H.: A self-learning call admission control scheme for CDMA cellular networks. IEEE Trans. Neural Networks 16, 1219–1228 (2005)
7. Wei, Q.L., Zhang, H.G., Dai, J.: Model-free multiobjective approximate dynamic programming for discrete-time nonlinear systems with general performance index functions. Neurocomputing 72, 1839–1848 (2009)
8. Zhang, H.G., Wei, Q.L., Liu, D.R.: An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. Automatica 47(1), 207–214 (2011)
9. Zhang, H.G., Wei, Q.L., Liu, D.R.: On-Line Learning Control for Discrete Nonlinear Systems Via an Improved ADDHP Method. In: Liu, D., Fei, S., Hou, Z.-G., Zhang, H., Sun, C. (eds.) ISNN 2007. LNCS, vol. 4491, pp. 387–396. Springer, Heidelberg (2007)
10. Wei, Q.L., Liu, D.R., Zhang, H.G.: Adaptive Dynamic Programming for a Class of Nonlinear Control Systems with General Separable Performance Index. In: Sun, F., Zhang, J., Tan, Y., Cao, J., Yu, W. (eds.) ISNN 2008, Part II. LNCS, vol. 5264, pp. 128–137. Springer, Heidelberg (2008)
11. Zhang, H., Wei, Q., Liu, D.: An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. Automatica 47, 207–214 (2011)
12. Wang, F., Jin, N., Liu, D., Wei, Q.: Adaptive dynamic programming for finite-horizon optimal control of discrete-time nonlinear systems with ϵ-error bound. IEEE Transactions on Neural Networks 22, 24–36 (2011)
13. Al-Tamimi, A., Abu-Khalaf, M., Lewis, F.L.: Adaptive critic designs for discrete-time zero-sum games with application to H_∞ control. IEEE Trans. Systems, Man, and Cybernetics-Part B: Cybernetics 37, 240–247 (2007)