Adaptive Optimal Control Law for Uncertain Nonlinear Inverted Pendulum System: An Adaptive Dynamic Programing Approach

Dao Phuong Nam[1*], Nguyen Van Huong[1], Ha Duc Minh[1], Nguyen Thanh Long[1]
[1]School of Electrical Engineering, Hanoi University of Science and Technology, Vietnam
[*]Corresponding: nam.daophuong@hust.edu.vn

Abstract **– This work presents the problem of adaptive optimal control law for a class of systems with input disturbance and unknown parameters. The main objective is to find an adaptive optimal control law based on the adaptive dynamic programming (ADP) method. An estimation of attraction region of the closed system is pointed out by using input state stability (ISS) theory. The theoretical analysis and simulation results demonstrate the performance of the proposed algorithm for an inverted pendulum system.**

Keywords - Adaptive dynamic programming (ADP), adaptive optimal control law, inverted pendulum.

INTRODUCTION

In order to design optimal control law for uncertain systems, the approximate/adaptive dynamic programming (ADP) approach is a biologically – inspired, non-model-based, computational method that has been used in numerous researches [1], such that appropriate reinforcement learning systems design by Werbos or neuro-dynamic programming by Bertsekas,… In [5], the control law was obtained after transforming the robust control problem into an optimal control problem. The corresponding optimal control depend on discrete-time HJB equation and it was solved by using a neural network. The proposed control law in [5] ensures closed-loop locally asymptotic stability of uncertain nonlinear system with an inequality condition. In [6], Fan *et al.* proposed the sliding mode controller based on adaptive optimal control theory for partially unknown nonlinear systems with input disturbances. The nearly optimal control design ensures stability of the equivalent sliding-mode dynamics by using policy iteration algorithm [6]. The critic network is utilized to approximate the cost function to overcome the difficulty at the second step of policy iteration algorithm. The proposed controller in [6] ensures closed-loop UUB stability of uncertain nonlinear system depend on the property of bounded signal.

In [7], Jiang *et al.* pointed out the control design based on continuous time systems and the equivalent HJB equation. However, the analytical solution of HJB equation is difficult to be obtained and [7] proposed the PI online technique. The stability analysis of closed-loop system pointed out input state stability (ISS) property and the estimation of attraction region depend on KL functions. In [8], Jiang proposed adaptive optimal control law based on algebraic Ricatti equation for uncertain linear systems without external disturbance. The computational adaptive optimal control algorithm was developed from Kleinman (1968) result. In [9, 2], the discretized model is utilized to propose PI, VI – based output ADP design control techniques. Remarkably, this paper extends [2] and [9] to obtain the robust control law of uncertain systems based on input state stability property. Additionally, we develop a new adaptive optimal control under the framework of the idea of ADP problem and external disturbances.

PROBLEM STATEMENTS

We study a class of discrete-time systems described by:

$$\begin{cases} x_{k+1} = Ax_k + B\left(u_k + \Delta_k\right) + Dv_k \\ v_{k+1} = Ev_k \\ e_k = Cx_k + Fv_k \end{cases} \qquad (1)$$

Where $x_k \in \mathbb{R}^n$ is the measured component of the state available in closed system, $u_k \in \mathbb{R}^m$ is the control input, $y_k = Cx_k \in \mathbb{R}^r$ represents the output of plant, $r_k = -Fv_k \in \mathbb{R}^r$ is the reference signal to be tracked, $e_k \in \mathbb{R}^r$ is tracking error. $v_k \in \mathbb{R}^q$ is the state of the exosystem.

$A \in \mathbb{R}^{n\times n}$; $B \in \mathbb{R}^{n\times m}$; $C \in \mathbb{R}^{r\times n}$; $D \in \mathbb{R}^{n\times q}$; $E \in \mathbb{R}^{q\times q}$; $F \in \mathbb{R}^{r\times q}$; $\Delta_k \in \mathbb{R}^m$ are unknown and $x_k;\Delta_k;v_k$ are unmeasurable.

The control objective is to find an adaptive optimal control law based on an iterative algorithm ensures that tracking errors converge to zero and convergence properties of this iterative algorithm in presence of uncertain and external disturbance in system.

In order to implement this work, some following assumptions must be given out:

Assumption 1: The pair $\left(A;B\right)$ is controllable.

Assumption 2: The transmission zeros condition holds, i.e, $rank\begin{bmatrix} A-\lambda I & B \\ C & 0 \end{bmatrix} = n+r; \forall \lambda \in \delta\left(E\right)$.

Assumption 3: The pairs $\left(\bar{C};\bar{A}\right)$ and $\left(F;E\right)$ are observable, where $\bar{C}=\begin{bmatrix}C & F\end{bmatrix};\bar{A}=\begin{bmatrix}A & D\\ 0 & E\end{bmatrix}$.

Assumption 4: There exists a constant value $\Delta>0$ such that $\left\|\Delta_k\right\|\le\Delta,\forall k\ge0$ and we note that $\Delta_k\in\mathbb{R}^m$ is the additional term in system described in [2]

PROPOSED CONTROL LAW

We propose the adaptive optimal control law based on the next lemmas, theorems described as follows:

Theorem 1 [3]: Let K_0 be any stabilizing feedback gain matrix, and repeat the following steps for $k=0;1;...$

(1) Solve for the real symmetric positive definite solution P_k of the Lyapunov equation:

$$\left(A-BK_k\right)^T P_k\left(A-BK_k\right)-P_k+C^TQC+K_k^TRK_k=0 \quad (2)$$

(2) Update the feedback gain matrix by: $K_{k+1}=\left(R+B^TP_kB\right)^{-1}B^TP_kA \quad (3)$

Then, the following properties are obtained:

(1) $A-BK_k$ is Hurwitz.

(2) $P^*\le P_{k+1}\le P_k$

(3) $\lim\limits_{k\to\infty}K_k=K^*;\lim\limits_{k\to\infty}P_k=P^*$

Remark 1: It is clear that $K^*=\left(R+B^TP^*B\right)^{-1}B^TP^*A$ with the symmetric matrix $P^*>0$ is the unique solution of the well-known discrete-time algebraic Riccati equation:

$$A^TPA-P+C^TQC-A^TPB\left(R+B^TPB\right)^{-1}B^TPA=0 \quad (4)$$

From that, authors in [2] have been proposed the **PI based output ADP design** as follows**:**

Suppose $\Delta_k=\bar{\Delta}_k$ is available during the learning phase. Define $w_k=u_k+\bar{\Delta}_k$.

The reconstruction:

Letting $m_k=\left[x_k^T;v_k^T\right]^T;\bar{B}=\left[B^T;0_{m\times q}^T\right]^T$, the system (1) can be rewritten : $\begin{cases}m_{k+1}=\bar{A}m_k+\bar{B}w_k\\ e_k=\bar{C}m_k\end{cases}$

The state of system can be written in the form of hist

$$m_k=\bar{A}^{n+q}m_{k-(n+q)}+\begin{bmatrix}\bar{B} & \bar{A}\bar{B} & ... & \bar{A}^{n+q-1}\bar{B}\end{bmatrix}\begin{bmatrix}w_{k-1}\\ w_{k-2}\\ ...\\ w_{k-(n+q)}\end{bmatrix}$$

$$\begin{bmatrix}e_{k-1}\\ e_{k-2}\\ \vdots\\ e_{k-(n+q)}\end{bmatrix}=\begin{bmatrix}\bar{C}\bar{A}^{n+q-1}\\ \vdots\\ \bar{C}\bar{A}\\ \bar{C}\end{bmatrix}m_{k-(n+q)}$$
$$+\begin{bmatrix}0 & \bar{C}\bar{B} & \bar{C}\bar{A}\bar{B} & ... & \bar{C}\bar{A}^{n+q-2}\bar{B}\\ 0 & 0 & \bar{C}\bar{B} & \cdots & \bar{C}\bar{A}^{n+q-3}\bar{B}\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ 0 & ... & & 0 & \bar{C}\bar{B}\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}w_{k-1}\\ w_{k-2}\\ \vdots\\ w_{k-(n+q)+1}\\ w_{k-(n+q)}\end{bmatrix}$$

$$\Rightarrow m_k=\bar{A}^{n+q}m_{k-(n+q)}+B_1\bar{w}_k,\bar{e}_k=C_1m_{k-(n+q)}+\Gamma\bar{w}_k \quad (5)$$

where

$\bar{e}_k=\left[e_{k-1}^T,e_{k-2}^T,...,e_{k-(n+q)}^T\right]^T$,

$\bar{w}_k=\left[w_{k-1}^T,w_{k-2}^T,...,w_{k-(n+q)}^T\right]^T$,

$B_1=\begin{bmatrix}\bar{B} & \bar{A}\bar{B} & ... & \bar{A}^{n+q-1}\bar{B}\end{bmatrix}$

$,C_1=\left[\left(\bar{C}\bar{A}^{n+q-1}\right)^T,\left(\bar{C}\bar{A}^{n+q-2}\right)^T,...,\left(\bar{C}\bar{A}\right)^T,\bar{C}^T\right]^T$,

$$\Gamma=\begin{bmatrix}0 & \bar{C}\bar{B} & \bar{C}\bar{A}\bar{B} & ... & \bar{C}\bar{A}^{n+q-2}\bar{B}\\ 0 & 0 & \bar{C}\bar{B} & \cdots & \bar{C}\bar{A}^{n+q-3}\bar{B}\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ 0 & ... & & 0 & \bar{C}\bar{B}\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

Because of **assumption 3** and Caley-Hamilton theorem, we obtain: $\bar{A}^{n+q}=MC_1$

The left inverse of C_1 exists:

$$C_1^*C_1=\left(C_1^TC_1\right)^{-1}C_1^TC_1=I_{n+q}$$

We can obtain: $M=\bar{A}^{n+q}C_1^*+G\left(I-C_1C_1^*\right)$

From (5) we have:

$$m_k=\bar{A}^{n+q}\left(C_1^TC_1\right)^{-1}C_1^T\bar{e}_k+\left(B_1-\bar{A}^{n+q}\left(C_1^TC_1\right)^{-1}C_1^T\Gamma\right)\bar{w}_k \quad (6)$$

From (1) we obtain: $\begin{cases}v_{k+1}=Ev_k\\ r_k=-Fv_k\end{cases}$

It is observed that: $r_{k-j}=-Fv_{k-j}=-FE^{q-j}v_{k-q}$

$\Rightarrow\left[r_{k-1}^T;r_{k-2}^T;...;r_{k-q}^T\right]$

$=v_{k-q}^T\left[\left(-FE^{q-1}\right)^T;\left(-FE^{q-2}\right)^T;...;\left(-F\right)^T\right]$

$\Rightarrow\bar{r}_k=F_1v_{k-q}$

where $\bar{r}_k=\left[r_{k-1}^T;r_{k-2}^T;...;r_{k-q}^T\right]^T$, $F_1=\left[\left(-FE^{q-1}\right)^T;\left(-FE^{q-2}\right)^T;...;\left(-F\right)^T\right]^T$.

In the similar way, we have: $v_k=E^q\left(F_1^TF_1\right)^{-1}F_1^T\bar{r}_k$ $\quad(7)$

The regulator equation:

$$\begin{cases}XE=AX+BU+D\\0=CX+F\end{cases}$$

Using **assumption 2** implies the regulator equation is solvable for any matrix $D;F$.

Define $\varepsilon_k=x_k-Xv_k$, from (1) we have:

$\varepsilon_k=x_k-Xv_k$

$\Rightarrow\varepsilon_{k+1}=A\varepsilon_k+B\left(u_k-Uv_k+\Delta_k\right)$ $\quad(8)$

$e_k=y_k-r_k=Cx_k+Fv_k=C\left(\varepsilon_k+Xv_k\right)+Fv_k$

$\Rightarrow e_k=C\varepsilon_k$ $\quad(9)$

From (6) we infer that:

$\varepsilon_k=x_k-Xv_k=\left[I_n,-X\right]m_k$

$$=\left[I_n,-X\right]\begin{bmatrix}\bar{A}^{n+q}\left(C_1^TC_1\right)^{-1}C_1^T\bar{e}_k\\+\left(B_1-\bar{A}^{n+q}\left(C_1^TC_1\right)^{-1}C_1^T\Gamma\right)\bar{w}_k\end{bmatrix}$$

$$\Rightarrow\varepsilon_k=M_e\bar{e}_k+M_w\bar{w}_k=\left[M_e,M_w\right]\begin{bmatrix}\bar{e}_k\\\bar{w}_k\end{bmatrix}=M\bar{z}_k$$

where $M_e=\left[I_n,-X\right]\bar{A}^{n+q}\left(C_1^TC_1\right)^{-1}C_1^T$, $M_w=\left[I_n,-X\right]\left(B_1-\bar{A}^{n+q}\left(C_1^TC_1\right)^{-1}C_1^T\Gamma\right)$, $M=\left[M_e,M_w\right],\bar{z}_k=\begin{bmatrix}\bar{e}_k\\\bar{w}_k\end{bmatrix}$

From we have $v_k=E^q\left(F_1^TF_1\right)^{-1}F_1^T\bar{r}_k=N\bar{r}_k$

Define:

$\bar{P}_j=M^TP_jM;\bar{K}_j=K_jM;\bar{U}=UN;A_j=A-BK_j$.

From (8) we have:

$$\varepsilon_{k+1}=A_j\varepsilon_k+B\left(u_k-\bar{U}\bar{r}_k+\bar{K}_j\bar{z}_k+\Delta_k\right)\quad(10)$$

In the case $\Delta_k=\bar{\Delta}_k$,define $p_k=w_k-\bar{U}\bar{r}_k+\bar{K}_j\bar{z}_k$, we have:

$\varepsilon_{k+1}=A_j\varepsilon_k+B\left(w_k-\bar{U}\bar{r}_k+\bar{K}_j\bar{z}_k\right)=A_j\varepsilon_k+Bp_k$

We consider that:

$\bar{z}_{k+1}^T\bar{P}_j\bar{z}_{k+1}-\bar{z}_k^T\bar{P}_j\bar{z}_k=\varepsilon_{k+1}^TP_j\varepsilon_{k+1}-\varepsilon_k^TP_j\varepsilon_k$

$=\left(A_j\varepsilon_k+Bp_k\right)^TP_j\left(A_j\varepsilon_k+Bp_k\right)-\varepsilon_k^TP_j\varepsilon_k$

Define: $S_{1,j}=B^TP_jB,S_{2,j}=B^TP_jB\bar{U}$, $S_{3,j}=M^TA_j^TP_jB,S_{4,j}=M^TA_j^TP_jB\bar{U}$, ,we $S_{5,j}=\bar{U}^TB^TP_jB\bar{U}$

rewrite:

$\bar{z}_{k+1}^T\bar{P}_j\bar{z}_{k+1}-\bar{z}_k^T\bar{P}_j\bar{z}_k=$

$-\left(e_k^TQe_k+\bar{z}_k^T\bar{K}_j^TR\bar{K}_j\bar{z}_k\right)+\left[w_k^TS_{1,j}w_k+\bar{z}_k^T\bar{K}_j^TS_{1,j}\bar{K}_j\bar{z}_k\right]$

$-\left[w_k^TS_{2,j}\bar{r}_k+\bar{z}_k^T\bar{K}_j^TS_{2,j}\bar{r}_k\right]+2\left[\bar{z}_k^TS_{3,j}w_k+\bar{z}_k^TS_{3,j}\bar{K}_j\bar{z}_k\right]$

$-2\bar{z}_k^TS_{4,j}\bar{r}_k+\bar{r}_k^TS_{5,j}\bar{r}_k$ $\quad(13)$

Applying Kronecker product representation gives:

$\bar{z}_k^T\bar{P}_j\bar{z}_k=\left(\bar{z}_k^T\otimes\bar{z}_k^T\right)vec\left(\bar{P}_j\right)$,

$w_k^TS_{1,j}w_k+\bar{z}_k^T\bar{K}_j^TS_{1,j}\bar{K}_j\bar{z}_k$

$=\left[w_k^T\otimes w_k^T+\left(\bar{K}_j\bar{z}_k\right)^T\otimes\left(\bar{K}_j\bar{z}_k\right)\right]vec\left(S_{1,j}\right)$,

$w_k^TS_{2,j}\bar{r}_k+\bar{z}_k^T\bar{K}_j^TS_{2,j}\bar{r}=\left(\bar{r}_k^T\otimes w_k^T+\bar{r}_k^T\otimes\left(\bar{K}_j\bar{z}_k\right)^T\right)vec\left(S_{2,j}\right)$

$\bar{z}_k^TS_{3,j}w_k+\bar{z}_k^TS_{3,j}\bar{K}_j\bar{z}_k=\left(w_k^T\otimes\bar{z}_k^T+\left(\bar{K}_j\bar{z}_k\right)^T\otimes\bar{z}_k^T\right)vec\left(S_{3,j}\right)$

$\bar{z}_k^TS_{4,j}\bar{r}_k=\left(\bar{r}_k^T\otimes\bar{z}_k^T\right)vec\left(S_{4,j}\right),\bar{r}_k^TS_{5,j}\bar{r}_k=\left(\bar{r}_k^T\otimes\bar{r}_k^T\right)vec\left(S_{5,j}\right)$

Define:

$\Phi_P\left(k\right)=\bar{z}_k^T\otimes\bar{z}_k^T-\bar{z}_{k+1}^T\otimes\bar{z}_{k+1}^T$,

$\Phi_1\left(k\right)=w_k^T\otimes w_k^T+\left(\bar{K}_j\bar{z}_k\right)^T\otimes\left(\bar{K}_j\bar{z}_k\right)$,

$\Phi_2\left(k\right)=\bar{r}_k^T\otimes w_k^T+\bar{r}_k^T\otimes\left(\bar{K}_j\bar{z}_k\right)^T$,

$\Phi_3\left(k\right)=w_k^T\otimes\bar{z}_k^T+\left(\bar{K}_j\bar{z}_k\right)^T\otimes\bar{z}_k^T$, $\Phi_4\left(k\right)=\bar{r}_k^T\otimes\bar{z}_k^T$

, $\Phi_5\left(k\right)=\bar{r}_k^T\otimes\bar{r}_k^T$, $\Phi_{QR}\left(k\right)=e_k^TQe_k+\bar{z}_k^T\bar{K}_j^TR\bar{K}_j\bar{z}_k$

Consequently, we have:

$\Psi_jS_j=\Pi_j$ $\quad(14)$

where

$$\Psi_j = \begin{bmatrix} \Phi_P(k_{j,1}) & \Phi_1(k_{j,1}) & \Phi_2(k_{j,1}) & \Phi_3(k_{j,1}) & \Phi_4(k_{j,1}) & \Phi_5(k_{j,1}) \\ \Phi_P(k_{j,2}) & \Phi_1(k_{j,2}) & \Phi_2(k_{j,2}) & \Phi_3(k_{j,2}) & \Phi_4(k_{j,2}) & \Phi_5(k_{j,2}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Phi_P(k_{j,s}) & \Phi_1(k_{j,s}) & \Phi_2(k_{j,s}) & \Phi_3(k_{j,s}) & \Phi_4(k_{j,s}) & \Phi_5(k_{j,s}) \end{bmatrix}$$

$$, S_j = \begin{bmatrix} vec(\bar{P}_j) \\ vec(S_{1,j}) \\ vec(S_{2,j}) \\ vec(S_{3,j}) \\ vec(S_{4,j}) \\ vec(S_{5,j}) \end{bmatrix}, \Pi_j = \begin{bmatrix} \Phi_{QR}(k_{j,1}) \\ \Phi_{QR}(k_{j,2}) \\ \vdots \\ \Phi_{QR}(k_{j,s}) \end{bmatrix}$$

Assumption 5: There exists $s^* > 0$ such that for all $s > s^*$ we have

$$rank\left(\begin{bmatrix} \varpi_{k_{j,1}}^T \otimes \varpi_{k_{j,1}}^T \\ \varpi_{k_{j,2}}^T \otimes \varpi_{k_{j,2}}^T \\ \vdots \\ \varpi_{k_{j,s}}^T \otimes \varpi_{k_{j,s}}^T \end{bmatrix}\right) = \frac{\mu(\mu+1)}{2}, \text{with} \quad \varpi_k = \begin{bmatrix} \bar{z}_k \\ u_k \\ \bar{r}_k \end{bmatrix}$$

and $\mu = (m+r)(n+q)+m+rq$.

By **assumption 5**, then S_j can be uniquely determined by: $S_j = \left(\Psi_j^T \Psi_j\right)^{-1} \Psi_j^T \Pi_j$ $\quad (15)$

Assumption 2 implies that B is in full column rank, so $S_{1,j}$ is the a nonsingular matrix, then:

$$\bar{U} = S_{1,j}^{-1} S_{2,j}; \bar{K}_{j+1} = \left(R + S_{1,j}\right)^{-1} S_{3,j}^T \qquad (16)$$

Algorithm 1:

Select a stabilizing K_0 and a threshold $\nu > 0$.

$j \leftarrow 0$.

Apply a bounded control policy μ_k on $\left[0; k_{0;0}\right)$.

repeat

Apply $u_k = -\bar{K}_j \varepsilon_k + e_k$ on $\left[k_{j;0}; k_{j;s}\right]$ with e_k an exploration noise. Solve $\bar{P}_j; \bar{K}_{j+1}$ from $(14)-(16)$:

$\Psi_j S_j = \Pi_j \qquad (14)$; $S_j = \left(\Psi_j^T \Psi_j\right)^{-1} \Psi_j^T \Pi_j \quad (15)$;

$$\bar{U} = S_{1,j}^{-1} S_{2,j}; \bar{K}_{j+1} = \left(R + S_{1,j}\right)^{-1} S_{3,j}^T \qquad (16)$$

$j \leftarrow j+1$

until $\left|\bar{P}_j - \bar{P}_{j-1}\right| < \nu$

$j^* \leftarrow j$. We obtain the approximated optimal control policy:

$$u_k = -\bar{K}_{j^*} \bar{z}_k + \left(S_{1,j^*}^{-1} S_{2,j^*}\right) \bar{r}_k$$

Theorem 2: Let $\bar{K}_0$ be any stabilizing feedback gain matrix, and let $\left(\bar{P}_j; \bar{K}_{j+1}; \bar{U}\right)$ be obtained from **algorithm 1**. Then, under assumption 5, we have $\lim_{j\to\infty} \bar{K}_j = \bar{K}^*; \lim_{j\to\infty} \bar{P}_j = \bar{P}^*$, with $\bar{K}^* = K^* M, \bar{P}^* = M^T P^* M$.

Proof:

From $(11);(16)$ one see that the $\left(P_k; K_{k+1}\right)$ obtained from $(2);(3)$ must satisfy the condition (14). In addition, by **assumption 4**, it is unique determined. Therefore, from **theorem 1**, we have $\lim_{j\to\infty} \bar{K}_j = \bar{K}^*; \lim_{j\to\infty} \bar{P}_j = \bar{P}^*$, with $\bar{K}^* = K^* M, \bar{P}^* = M^T P^* M$.

Lemma 1: Let $x \in \mathbb{R}^p, y \in \mathbb{R}^q$ and M, N are appropriately dimensioned matrices, then for any positive number θ and every appropriately dimensioned matrix $X(t)$ satisfying $X^T(t)X(t) \le I$ we have:

$$2x^T MXNy \le \theta x^T MM^T x + \theta^{-1} y^T N^T Ny$$

Theorem 3: The approximated optimal control policy $u_k = -\bar{K}_{j^*} \bar{z}_k + \left(S_{1,j^*}^{-1} S_{2,j^*}\right) \bar{r}_k$ can obtain the system $\varepsilon_{k+1} = A\varepsilon_k + B\left(u_k - Uv_k + \Delta_k\right)$ $\quad (8)$ is ISS. Moreover, we obtain that the tracking error of system (1) attracts to region.

Proof:

We have: $u_k = -\bar{K}_{j^*} \bar{z}_k + \left(S_{1,j^*}^{-1} S_{2,j^*}\right) \bar{r}_k = -K_{j^*} \varepsilon_k + Uv_k$.

From (8) we imply that:

$$\begin{aligned} \varepsilon_{k+1} &= A\varepsilon_k + B\left(u_k - Uv_k + \Delta_k\right) \\ &= A\varepsilon_k + B\left(-K_{j^*} \varepsilon_k + Uv_k - Uv_k + \Delta_k\right) \\ &= \left(A - BK_{j^*}\right)\varepsilon_k + B\Delta_k = A_{j^*} \varepsilon_k + B\Delta_k \end{aligned}$$

Consider the Lyapunov candidate function: $V_k = \varepsilon_k^T P_{j^*} \varepsilon_k$ we have:

$$\begin{aligned} V_{k+1} - V_k &= \varepsilon_{k+1}^T P_{j^*} \varepsilon_{k+1} - \varepsilon_k^T P_{j^*} \varepsilon_k \\ &= \left(A_{j^*} \varepsilon_k + B\Delta_k\right)^T P_{j^*} \left(A_{j^*} \varepsilon_k + B\Delta_k\right) - \varepsilon_k^T P_{j^*} \varepsilon_k \\ &= \varepsilon_k^T \left(A_{j^*}^T P_{j^*} A_{j^*} - P_{j^*}\right)\varepsilon_k + 2\Delta_k^T B^T P_{j^*} A_{j^*} \varepsilon_k + \Delta_k^T B^T P_{j^*} B\Delta_k \\ &= -\varepsilon_k^T \left(C^T QC + K_{j^*}^T RK_{j^*}\right)\varepsilon_k \\ &\quad + 2\Delta_k^T B^T P_{j^*} A_{j^*} \varepsilon_k + \Delta_k^T B^T P_{j^*} B\Delta_k \end{aligned}$$

$$\leq -\lambda_{\min}\left(C^TQC + K_{j^*}^T RK_{j^*}\right)\|\varepsilon_k\|^2$$
$$+2\Delta_k^T B^T P_{j^*} A_{j^*}\varepsilon_k + \left\|\left(B^T P_{j^*} B\right)^{1/2}\right\|^2 \|\Delta_k\|^2$$

Applying **lemma 1**, we have:

$$V_{k+1} - V_k \leq -\lambda_{\min}\left(C^TQC + K_{j^*}^T RK_{j^*}\right)\|\varepsilon\|^2$$
$$+\left[\frac{2\left\|B^T P_{j^*} A_{j^*}\right\|^2}{\lambda_{\min}\left(C^TQC + K_{j^*}^T RK_{j^*}\right)}\|\Delta_k\|^2 + \frac{1}{2}\lambda_{\min}\left(C^TQC + K_{j^*}^T RK_{j^*}\right)\|\varepsilon_k\|^2\right]$$
$$+\left\|\left(B^T P_{j^*} B\right)^{1/2}\right\|^2 \|\Delta_k\|^2$$

$$V_{k+1} - V_k \leq -\frac{1}{2}\lambda_{\min}\left(C^TQC + K_{j^*}^T RK_{j^*}\right)\|\varepsilon_k\|^2$$
$$+\left[\frac{2\left\|B^T P_{j^*} A_{j^*}\right\|^2}{\lambda_{\min}\left(C^TQC + K_{j^*}^T RK_{j^*}\right)} + \left\|\left(B^T P_{j^*} B\right)^{1/2}\right\|^2\right]\|\Delta_k\|^2$$

It is clear that $\lambda_{\min}\left(P_{j^*}\right)\|\varepsilon_k\|^2 \leq V_k \leq \lambda_{\max}\left(P_{j^*}\right)\|\varepsilon_k\|^2$, so that the system (8) is ISS.

Remark 1: It is different from [2], we obtain the adaptive optimal control law for discrete-time systems affected by external disturbances.

Remark 2: The proposed optimal control law guarantees the ISS property of the closed-loop system in presence of uncertain parameters and external disturbances.

Remark 3: The proposed optimal control law absolutely developed for continuous time

SIMULATION RESULTS

M	Weight of car	8.378	kg
M	Weight of link	0.051	kg
B	Friction coefficient	12.98	Ns/m
L	½ length of link	0.325	m
I	Inertial moment of link	1.796×10^{-03}	$kg.m^2$
G	Gravity	9.81	m/s^2

Table 1. The Parameters of Inverted Pendulum

In this section, we apply the proposed adaptive optimal control law to an inverted pendulum on a cart described as (17) and table 1. The simulation results in Fig. 1 show the convergence of matrix P and K of proposed algorithm and the tracking errors converge to zero.

$$\begin{bmatrix}\dot{x}\\ \ddot{x}\\ \dot{\phi}\\ \ddot{\phi}\end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0\\ 0 & \dfrac{-\left(I+ml^2\right)b}{I\left(M+m\right)+Mml^2} & \dfrac{m^2gl^2}{I\left(M+m\right)+Mml^2} & 0\\ 0 & 0 & 0 & 1\\ 0 & \dfrac{-mlb}{I\left(M+m\right)+Mml^2} & \dfrac{mgl\left(M+m\right)}{I\left(M+m\right)+Mml^2} & 0\end{bmatrix}\begin{bmatrix}x\\ \dot{x}\\ \phi\\ \dot{\phi}\end{bmatrix}$$
$$+\begin{bmatrix}0\\ \dfrac{I+ml^2}{I\left(M+m\right)+Mml^2}\\ 0\\ \dfrac{ml}{I\left(M+m\right)+Mml^2}\end{bmatrix}u \qquad (17)$$
$$y = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}x\\ \dot{x}\\ \phi\\ \dot{\phi}\end{bmatrix} + \begin{bmatrix}0\\ 0\end{bmatrix}u$$

CONCULUSION

This paper presents an adaptive optimal control algorithm for practical online implementation of discrete-time systems with unknown system dynamics and external disturbance. The proposed algorithm pointed out the ISS and convergence properties. The theory analysis and simulation results illustrate the effectiveness of proposed algorithm.

REFERENCES

[1] Z. P. Jiang and Y. Jiang, “Robust adaptive dynamic programming for linear and nonlinear systems: An overview,” *European Journal of Control*, vol. 19, no. 5, pp. 417–425, 2013.

[2] Weinan Gao, Jiang, “Adaptive optimal output regulation via output-feedback: An adaptive dynamic programing approach,” *IEEE 55th Conference on Decision and Control (CDC)*, Las Vegas, USA, pp.5845 – 5850, 2016.

[3] G. Hewer., “An iterative technique for the computation of the steady state gains for the discrete optimal regulator,” *IEEE Transactions on Automatic Control*, vol. 16, no. 4, pp.382–384, 1971

[4] P. Lancaster and L. Rodman.: Algebraic Riccati Equations. New York, NY: Oxford University Press Inc., 1995

[5] D. Wang, D. Liu, H. Li, B. Luo, and H. Ma., “An approximate optimal control approach for robust stabilization of a class of discrete-time nonlinear systems with uncertainties,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 5, pp. 713–717, 2016

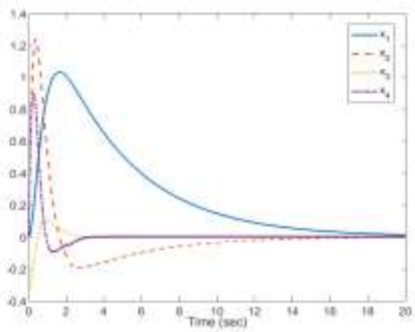
[6] Q. Y. Fan and G. H. Yang., “Adaptive actor-critic design-based integral sliding-mode control for partially unknown nonlinear systems with input

disturbances," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 27, no. 1, pp. 165–177, 2016

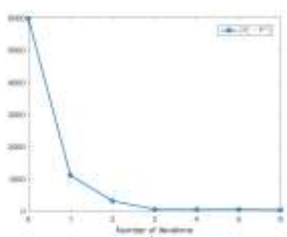
[7] Y. Jiang and Z. P. Jiang., "Robust adaptive dynamic programming and feedback stabilization of nonlinear systems," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 25, no. 5, 882–893, 2014

[8] Y. Jiang and Z. P. Jiang., "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica,* vol. 48, no. 10, 2699–2704, 2012

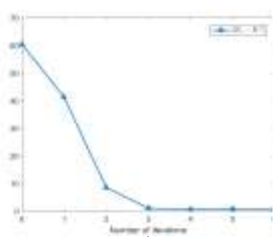
[9] W. Gao, Y. Jiang, Z. P. Jiang, and T. Chai., "Output-feedback adaptive optimal control of interconnected systems based on robust adaptive dynamic programming," *Automatica*, vol. 72, pp. 37–45, 2016



(a)



(b)



(c)

Fig. 1 Convergence of matrix P, K and tracking errors