Adaptive optimal control approach to robust tracking of uncertain linear systems based on policy iteration

Dengguo Xu¹², Qinglin Wang¹ and Yuan Li¹

Abstract
In this study, an optimal adaptive control approach is established to solve the robust output tracking problem of a class of continuous time uncertain linear systems based on the policy iteration (PI) in actor-critic algorithm. First, by augmenting the integral variables of the tracking error into state variables, the robust tracking problem is transformed into a robust control problem of an augmented uncertain linear system. It is proven that the robust control law of the augmented system enables the output of the considered system to track a polynomial time signal asymptotically. Second, an optimal control method in the corresponding auxiliary nominal system is established, and based on the Bellman optimality principle, PI algorithms are proposed to solve online tracking controllers for the matched and the mismatched uncertain systems. Finally, for testing the availability of the proposed approach and theoretical results, two numerical experiments are provided.

Keywords
Policy iteration, uncertain linear system, robust tracking, optimal control

Date received: 23 February 2021; accepted: 11 March 2021

Introduction
Control systems in many practical problems have uncertainties, owing to data errors or disturbances. Therefore, study on the robust control problem of uncertain systems has garnered significant attention. An adaptive optimal control method to design a robust tracking controller based on the policy iteration (PI) of the actor-critic algorithm is introduced in the research of algorithm design. Adaptive optimal control is a control method that automatically adjusts controller’s parameters to optimize the performance of the control system according to the change of parameters or environment. Moreover, it is widely used in control practice.

Since the 1980s, the robust control problem of practical uncertain systems has been a research hotspot. Algebraic Riccati equations (ARE) were utilized to propose a robust control design for some uncertain systems. The robust control of time-varying uncertain linear systems was studied, and the controller parameters were obtained using a one-dimensional search method. The proposed method could be effective if the system were to have satisfied the matching condition. Using a noniterative method, a robust controller was obtained for a matched system. The need for precompensators of the unstable nominal system was a disadvantage of this design method. In addition, other effective robust control-design methods were proposed for general matched uncertain systems. However, only a few study results are available for mismatched systems having an uncertainty in the input matrix.

The robust tracking problem has been investigated in many studies. Schmitendorf et al. proposed a robust tracking control approach for time-invariant uncertain system with unknown constant disturbances. Benson and Schmitendorf considered a robust tracking problem by using observer-based augmented system. Utilizing an improved linear quadratic optimal control method, Shieh et al. designed a robust tracking law for an uncertain linear systems. Solving a robust tracking problem for an uncertain system was turned into an optimal control problem. For linear systems, it is a linear regulator problem, which is solved using an ARE.

¹School of Automation, Beijing Institute of Technology, Beijing, China
²School of Physics and Electrical Engineering, Liupanshui Normal University, Liupanshui, China

Corresponding author:
Yuan Li, School of Automation, Beijing Institute of Technology, No.5, South Street, Zhongguancun, Haidian District, Beijing, P.R. China.
Email: liyuan@bit.edu.cn

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Furthermore, the PI in reinforcement learning (RL) has been utilized extensively in solving control and tracking problems for deterministic systems. In terms of feedback control, we can refer to the literature.\textsuperscript{13–15} There are also some literatures on using PI algorithm to solve tracking problem. An online RL was established to obtain a linear tracking controller in partially unknown linear systems with unstable tracking signals.\textsuperscript{16} By constructing an augmented system comprising the state variable and tracking signal, the tracking problem was turned into an optimal control problem. Therefore, the optimal tracking cost is quadratic as well. An adaptive dynamic programming (ADP) algorithm was proposed to obtain the tracking controller for completely unknown linear systems,\textsuperscript{17} where tracking signals were generated for stabilizing the systems. Moreover, RL and ADP techniques were employed to achieve an optimal output regulation for linear systems.\textsuperscript{18}

On the other hand, the application of RL in the robust control of uncertain systems has been proposed by many researchers.\textsuperscript{19–26} Robust tracking problems exist widely in practical applications. However, to our knowledge, only a few studies were conducted on the PI-based robust tracking of uncertain linear systems. Most existing design methods for solving robust tracking problems are based on impractical nominal systems. Therefore, it is necessary to use the PI algorithm to solve the robust tracking problems of uncertain linear systems with an unknown nominal system matrix. In this study, a PI method was developed to solve the robust tracking control of an uncertain linear system with a polynomial tracking signal. The robust tracking problem was turned into a robust control problem of an augmented system. Additionally, online PI algorithms were developed to solve the considered tracking design problem for matched and mismatched systems.

Our main contributions are as follows. First, we considered the tracking problem of general uncertain linear systems in which both the state matrix and the input matrix were uncertain. In the existing literature,\textsuperscript{12} the design of a robust tracking-control method was proposed for uncertain linear systems only when an uncertainty entered the system matrix. The existing results were extended to a case where the input matrix was uncertain. Second, online PI algorithms were presented to solve tracking problems for the matched uncertain system and the mismatched uncertain systems. Because it is practically difficult to obtain the nominal system matrix information accurately, using the PI algorithm is advantageous. As a result, we extended the PI algorithm to calculate robust tracking control law for the general uncertain linear systems.

The rest of this paper is arranged as follows. We formulate the robust tracking problems and propose some basic results for the issues under consideration in Section 2. Solving robust tracking problem is converted to calculate a robust control law for augmented uncertain systems. In Sections 3 and 4, the robust tracking problem for a matched and mismatched linear system is solved by transforming it into a robust control problem of an augmented system. Online PI algorithms are developed for an augmented uncertain system based on the optimal control of an auxiliary linear system. To support the proposed theoretical framework, we provide numerical experiments with two examples in Section 5. In Section 6, the study is concluded, and the scope for future research is discussed.

Robust tracking control framework

Consider an continuous-time linear system with uncertainty as follows:

\[ \dot{x} = A(s)x + B(l)u \]
\[ y = Cx, \]

where \( x \in \mathbb{R}^n \) is the state vector, \( u \in \mathbb{R}^m \) is the input variable, \( y \in \mathbb{R}^r \) is the system output, \( s \in \mathbb{S}, l \in \mathbb{L} \) are the uncertain parameter vectors, and \( S \) and \( L \) are sets of uncertain parameters. \( A(s) \) is an \( n \times n \) uncertain state parameter matrix, \( B(l) \) is an \( n \times m \) uncertain input parameter matrix, and \( C \) is an \( 1 \times n \) constant output matrix.

The objective of control design is to establish a control input, \( u = Kx \), such that the system output \( y \) asymptotically tracks the desired signal, \( y_r \), for all \( s \in \mathbb{S} \) and \( l \in \mathbb{L} \). The referenced signal is assumed to be a polynomial time signal, \( y_r = a_0 + a_1 t + \cdots + a_d t^{d-1} \), where \( a_0, a_1, \cdots, a_d \) are constants, and \( d \) is a nonnegative integer. In particular, the control design goal is to establish a control input, \( u = Kx \), such that the closed-loop uncertain system (1) is asymptotically stable and the output, \( y = Cx \), can asymptotically track the signal, \( y_r \), for all \( s \in \mathbb{S} \) and \( l \in \mathbb{L} \).

Some definitions and basic assumptions\textsuperscript{1–27} are elaborated as follows.

**Assumption 1.** Nominal values \( s_0 \in \mathbb{S} \) and \( l_0 \in \mathbb{L} \), such that \( (A(s_0), B(l_0)) \) is stabilizable and \( (A(s_0), C) \) observable.

**Definition 1.** System (1) satisfies matched condition in system matrix if, for every \( s \in \mathbb{S} \), there is a matrix, \( \phi(s) \), such that

\[ A(s) - A(s_0) = B(l_0)\phi(s), \]

where \( \phi(s) \in \mathbb{R}^{n \times n} \).

**Definition 2.** System (1) satisfies the matched condition in input matrix if, for every \( l \in \mathbb{L} \), there is a matrix \( \phi(l) \), such that

\[ B(l) - B(l_0) = B(l_0)\phi(l), \]

where \( \phi(l) \in \mathbb{R}^{m \times m} \), and \( \phi(l) \geq 0 \).
Definition 3. When system (1) satisfies conditions (2) and (3), for all $s \in S$ and $l \in \mathbb{L}$, it is called a matched uncertain linear system.

Assumption 2. For the linear system (1) with matching uncertainties satisfying conditions (2) and (3), there is a known matrix, $M$, such that $\phi^T(s) \phi(s) \lesssim M \geq 0$ for every $s \in S$.

If system (1) does not satisfy matching conditions (2) and (3), the pseudo inverse, $B(l)_o^+$, of nominal input matrix $B(l)_0$ is introduced to decompose the uncertain system matrix and input the matrix into the following matched and mismatched parts.

$$
A(s) - A(s_0) = B(l)_0B(l)_0^+ [A(s) - A(s_0)] + [I - B(l)_0B(l)_0^+] [A(s) - A(s_0)],
$$

and

$$
B(l) - B(l)_0 = B(l)_0B(l)_0^+ [B(l) - B(l)_0] + [I - B(l)_0B(l)_0^+] [B(l) - B(l)_0],
$$

where $B(l)_0^+ = [B^T(l)_0B(l)_0]^{-1}B^T(l)_0$.

If system (1) is a mismatched uncertain system, let the uncertainty of the system be subject to the following assumptions.

Assumption 3. A positive semidefinite matrix, $F$, exists, such that

$$
[A(s) - A(s_0)]^T [B(l)_0^+]^T B(l)_0^+ [A(s) - A(s_0)] \leq F.
$$

Assumption 4. A positive semidefinite matrix, $H$, exists, such that

$$
[A(s) - A(s_0)]^T [A(s) - A(s_0)] \leq H.
$$

Assumption 5. A positive semidefinite matrix, $G$, exists, such that

$$
[A(s) - A(s_0)]^T C^T C [A(s) - A(s_0)] \leq G.
$$

The above assumptions are common in the study of uncertain systems^1,2,6,27.

To solve the robust tracking problem, we introduce a new variable, $e = y - y_r$, as an error. Based on the integrals of the error variable, the following new state variables are defined:

$$
\dot{q}_1 = e, \dot{q}_2 = q_1, \ldots, \dot{q}_d = q_{d-1}
$$

Now, define the augmented system state as

$$
X(t) = [x^T(t) \quad q_1 \quad q_2 \quad \cdots \quad q_d]^T.
$$

Using the augmented system state, an uncertain augmented linear system is constructed as

$$
\dot{X} = T(s)X + B(l)u + Ny_r,
$$

where

$$
T(s) = \begin{bmatrix} A(s) & O_{n \times d} \\ C & O_{d \times d} \end{bmatrix},
$$

$$
B(l) = \begin{bmatrix} B(l) \\ O_{d \times m} \end{bmatrix}, N = \begin{bmatrix} O_{e \times 1} \\ -1 \\ \vdots \\ 0 \end{bmatrix}.
$$

Here, $O_{e \times j} \in \mathbb{R}^{e \times j}$ represents a zero matrix. Equation (8) can be regarded as a linear nonhomogeneous uncertain system with a nonhomogeneous term, $Ny_r$. The considered robust tracking problem could be converted to the following robust stabilization problem. For every $s \in S$ and $l \in L$, find a control input, $u = KX$, such that the augmented homogeneous uncertain linear system

$$
\dot{X} = T(s)X + B(l)u
$$

is asymptotically stable.

Lemma 1. If system (1) satisfies conditions (2) and (3), then the augmented uncertain system (9) is a matched uncertain linear system.

Proof. Assume that system (1) satisfies the system matrix matched condition (2) and the input matrix matched condition (3). Consequently, $A(s) - A(s_0) = B(l)_0 \phi(s)$. Therefore,

$$
T(s) - T(s_0) = \begin{bmatrix} A(s) - A(s_0) & O_{n \times d} \\ O_{d \times n} & O_{d \times d} \end{bmatrix}
$$

$$
= \begin{bmatrix} B(l)_0 \\ O_{d \times m} \end{bmatrix} \begin{bmatrix} \phi(s) & O_{m \times d} \end{bmatrix}
$$

$$
= B(l)_0 \Phi(s),
$$

where $B(l)_0 = \begin{bmatrix} B(l) \\ O_{d \times m} \end{bmatrix} \in \mathbb{R}^{(n + d) \times m}$, and

$$
\Phi(s) = [\phi(s) \quad O_{m \times d}] \in \mathbb{R}^{m \times (n + d)}.\n$$

This implies that the augmented system (3) satisfies the system matrix-matched condition. Moreover,

$$
\dot{B}(l) - \dot{B}(l)_0 = \begin{bmatrix} B(l) - B(l)_0 \\ O_{d \times m} \end{bmatrix}
$$

$$
= \begin{bmatrix} B(l)_0 \phi(l) \\ O_{d \times m} \end{bmatrix} = \tilde{B}(l)_0 \phi(l),
$$

which implies that the augmented system (9) satisfies the input matrix-matched condition. Hence, the proof is complete.

Lemma 2. Denote the new state as $q = [q_1 \quad q_2 \quad \cdots \quad q_d]^T$. For all $s \in S$, $l \in L$, supposing that the state feedback controller $u = KX = [\dot{K}x \quad K_i q]$
can stabilize the system $\dot{X} = T(s)X + \overline{B}(l)u$, then $u = Kx$ will stabilize system (1), and $y \to y_{r}$ as $t \to \infty$.

**Proof.** Because (9) is an augmented system corresponding to system (1), $u = Kx$ will stabilize system (1). By determining the d-order differential on both sides in the augmented system (8), it yields

$$X^{d+1} = T(s)X^{d} + \overline{B}(l)KX^{d}$$

$$= \left( T(s) + \overline{B}(l)K \right)X^{d}.$$ (10)

Considering that $u = KX$ can stabilize system (9), the matrix $T(s) + \overline{B}(l)K$ is Hurwitz stable. It follows from (10) that $X^{d} \to 0$ as $t \to \infty$. This implies that $u^{d} = e \to 0$ as $t \to \infty$. Therefore, we have $y \to y_{r}$ as $t \to \infty$. Hence, the proof is complete.

Lemma 2 shows that the robust tracking problem of system (1) can be transformed to a robust stabilization problem of the augmented homogeneous system (9).

**Remark 1.** Tan et al.\textsuperscript{12} considered the robust tracking of an uncertain linear system without input uncertainty. In this study, we considered a system with uncertainties entering the input matrix, which is an extension of the existing study outcomes.

Here, the system is divided into matched and mismatched cases, and the robust tracking problem of the system is discussed.

**Matched uncertain linear systems**

In this section, PI algorithms are developed to calculate robust tracking control law for linear systems with matched uncertainties. The problem is transferred into stabilizing an augmented uncertain system which contains the original system states and tracking signal. Based on solving an optimal control problem with augmented nominal system and predefined performance index, the PI algorithms are proposed to obtain robust tracking feedback control.

**Robust Stabilization of an augmented linear system with uncertainty**

Here, we discuss system (1) when it satisfies the matched conditions. Furthermore, robust stabilization for the augmented uncertain systems is transformed into calculating an optimal control with nominal system and predefined performance index. The optimal control problem is solved using the PI method, and the required robust tracking control law is obtained.

For the nominal linear system

$$\dot{X} = T(s_{0})X + \overline{B}(l_{0})u,$$ (11)

we construct an optimal-control problem. Acquire a controller, $u = Kx$, that minimizes the following performance index

$$V(X(t_{0}), u(\cdot)) = \int_{t_{0}}^{\infty} [X^{T}QX + u^{T}u]dt,$$ (12)

where $Q = \bar{M} + I \succeq 0$ and $\bar{M}$ is the supremum of uncertainty $\Phi^{T}(s)\Phi(s)$, and $I$ is an identity matrix with proper dimensions. Actually, $\bar{M} = \begin{bmatrix} M & O \\ O & O \end{bmatrix}$, which based on the following derivation

$$\Phi^{T}(s)\Phi(s) = \begin{bmatrix} \varphi(s) \\ \varphi^{T}(s) \end{bmatrix}^{T} \begin{bmatrix} \varphi(s) & O_{m \times d} \end{bmatrix}$$

$$= \begin{bmatrix} \varphi^{T}(s)\varphi(s) \\ O \\ O \end{bmatrix}$$

$$\leq \begin{bmatrix} M & O \\ O & O \end{bmatrix}$$

Here, $O$ denotes a zero matrix of an appropriate dimensions.

According to the optimal control theory, $x_{0} = -B^{T}(l_{0})PX$ is the solution of the optimal control problem (11) and (12), which the positive matrix $P$ satisfies the following ARE:

$$T^{T}(s_{0})P + PT(s_{0}) + Q - \bar{B}(l_{0})\overline{B}^{T}(l_{0})P = 0.$$ (13)

**Theorem 1.** Denote $K = -B^{T}(l_{0})P$. If system (1) is a matched uncertain system with conditions (2) and (3), then the solution $u = KX$, in optimal control problem (11) and (12) can stabilize the augmented uncertain system (9). That is, for every $s \in \mathbb{S}$, $l \in \mathbb{L}$, the closed-loop system,

$$\dot{X} = [T(s) + \overline{B}(l)K]X$$ (14)

is asymptotically stable.

**Proof.** It follows from Lemma 1 that the augmented system (9) satisfies the following matching conditions:

$$T(s) - T(s_{0}) = \bar{B}(l_{0})\varphi(s)$$ (15)

and

$$\bar{B}(l) - \bar{B}(l_{0}) = \delta(l)$$ (16)

Choosing the Lyapunov function as $V(X) = X^{T}PX$ and taking its time derivative, along the closed-loop system (14), we obtain

$$\frac{dV}{dt} = \dot{X}^{T}PX + X^{T}\dot{X}$$

$$= X^{T}[T(s) + \overline{B}(l)K]^{T}PX + X^{T}P[T(s) + \overline{B}(l)K]X.$$
It follows from the optimal control gain, $K = -B^T(l_0)P$, and Riccati equation (13), that
\[
\frac{dV}{dt} = -X^T MX - X^T X - X^T K^T K X
- 2X^T K^T \Phi(s) X - 2X^T K^T \phi(l) K X
= -X^T M X - X^T X - X^T K^T K X
- 2X^T K^T \Phi(s) X - 2X^T K^T \phi(l) K X
- X^T \Phi(s) \Phi(s) X + X^T \Phi(s) \Phi(s) X
- X^T (M - \Phi(s) \Phi(s)) X - 2X^T K^T \phi(l) K X
- X^T (K + \Phi(s))^T (K + \Phi(s)) X - X^T X.
\]

We immediately conclude that $\frac{dV}{dt} < 0$, because $\Phi^T(s) \Phi(s) \leq M$ and $\phi(l) \geq 0$. Therefore, for all $s \in S, l \in L$, closed-loop system (14) is asymptotically stable. Hence, the proof is complete.

**PI algorithm and its convergence**

This subsection has two parts. First, based on the performance function (12), an offline policy iteration algorithm is proposed to solve the robust tracking problem. Then, an integral reinforcement formula is developed based on the Bellman equation. An online real-time PI algorithm is developed to solve the robust tracking problem with unknown nominal system matrix.

For any initial time, $t$, rewrite the performance function (12) as
\[
X^T P X = \int_t^\infty [X^T Q X + u^T u] dt
\tag{17}
\]
Differentiating (17) with the trajectories of system (11) yields
\[
H^T P + PH + Q + K^T K = 0,
\tag{18}
\]
where $H = T(s_0) + B(l_0) K$.

Using (18) and $K = -B^T(l_0) P$, an off-line PI algorithm for robust tracking problem is obtained.

**Algorithm 1.** Offline PI algorithm for robust tracking of matched uncertain linear systems

1. **Initialization:** Select an initial stabilization control gain, $K_0$.
2. **Policy evaluation:** Solve $P_i$ in the equation, for a given control gain $K_i$,
\[
H_i^T P_i + P_i H_i + Q + K_i^T K_i = 0,
\tag{19}
\]
where $H_i = T(s_0) + B(l_0) K_i$.
3. **Policy improvement:** Compute $K_{i+1}$ using
\[
K_{i+1} = -B^T(l_0) P_i.
\tag{20}
\]

By alternative iterating (19) and (20), Algorithm 1 can be used to calculate the robust tracking law of uncertain systems, which is an extension of Kleinman’s algorithm. The convergence proof of Algorithm 1 is identical to that of Kleinman’s algorithm.

In Algorithm 1, we can solve ARE (13) by iteratively computing (19) and (20). However, the implementation of the algorithm needs to know the information of the nominal system, and the process of calculating the controller can only be realized offline. Here, an online PI algorithm is developed to solve ARE (13) with an unknown nominal system matrix, $T(s_0)$.

For any initial time, $t$, the optimal cost in (12) can be rewritten as
\[
V(X(t)) = \int_t^\infty [X^T Q X + u^T u] dt
= \int_t^{t + \Delta t} [X^T Q X + u^T u] dt
+ \int_{t + \Delta t}^\infty [X^T Q X + u^T u] dt,
\]
that is,
\[
V(X(t)) = \int_t^{t + \Delta t} [X^T Q X + u^T u] dt + V(X(t + \Delta t))
\]
So we can have
\[
X^T(t) P X(t) = \int_t^{t + \Delta t} [X^T Q X + u^T u] dt
+ X^T(t + \Delta t) P X(t + \Delta t).
\tag{21}
\]
In equation (21), only the matrix $P$ is unknown. we can calculate the elements of the matrix $P$ through the system trajectory data. Consequently, an online PI algorithm sloving robust tracking problem for uncertain linear system is obtained.

**Algorithm 2.** The online PI algorithm for robust tracking of matched uncertain linear systems

1. **Initialization:** Select an initial stabilization control law, $K_0$.
2. **Policy evaluation:** Compute $P_i$ from
\[
X^T(t) P X(t) = \int_t^{t + \Delta t} [X^T Q X + X^T K_i^T K_i X] dt
+ X^T(t + \Delta t) P X(t + \Delta t).
\tag{22}
\]
3. **Policy improvement:** Compute $K_{i+1}$
\[
K_{i+1} = -B^T(l_0) P_i.
\tag{23}
\]

**Remark 2.** Algorithm 2 is an online PI algorithm based on the RL. Through the online trajectory data of system (11), the matrix $P_i$ can be solved using (22) and the least-squares method. Through iterative calculation and using (22) and (23), the robust control gain, $K_i$ for augmented uncertain linear system (9) is obtained. Moreover, according to Lemma 2, the system robust
tracking law is expressed as $u = Kx$ by decomposing the robust control law into $KY = [Kx \quad K_i q]$. As an advantage, Algorithm 2 need not know the nominal system matrix and can effectively avoid dimension disaster. In [13], by the use of online PI, the linear quadratic regulator with unknown system matrix is calculated. We developed this algorithm to solve the robust tracking problem for an uncertain linear system.

The convergence in Algorithm 2 is proven as follows:

**Theorem 2.** We assume that $(T(s_0) + \tilde{B}(l_0)K_i)$ is stable, and we solve $P_i$ using equation (22), which equals to obtaining a solution for the following equation:

$$[T(s_0) + \tilde{B}(l_0)K_i]^TP_i + P_i[T(s_0) + \tilde{B}(l_0)K_i] + Q + K_i^TP_iK_i = 0$$  \hspace{1cm} (24)

**Proof.** Dividing by $\Delta t$ on both sides of (22) and taking a limit, we attain

$$0 = \lim_{\Delta t \to 0} \frac{X^T(t + \Delta t)P_iX(t + \Delta t) - X^T(t)P_iX(t)}{\Delta t} + \lim_{\Delta t \to 0} \frac{\int_{t}^{t + \Delta t} X^T(Q + K_i^TP_iK_i)Xdt}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{\int_{t}^{t + \Delta t} \frac{dX^T(t)P_iX(t)}{dt}dt}{\Delta t} + \lim_{\Delta t \to 0} \frac{\int_{t}^{t + \Delta t} X^T(Q + K_i^TP_iK_i)Xdt}{\Delta t}$$

$$= X^T(T(s_0) + \tilde{B}(l_0)K_i)^TP_iX(t) + X^T[TP_i(T(s_0) + \tilde{B}(l_0)K_i)]X(t) + X^T(T(s_0) + \tilde{B}(l_0)K_i)^TP_iX(t).$$

Thus, (22) implies (24).

Furthermore, take into account the asymptotically stable systems, $X = (T(s_0) - \tilde{B}(l_0)K_i)X$. Selecting the Lyapunov function as $V_i(X) = X^T P_i X$, and taking its time derivative along the yields

$$\frac{d}{dt}(X^TP_iX) = X^T(T(s_0) + \tilde{B}(l_0)K_i)^TP_iX(t) + P_i(T(s_0) + \tilde{B}(l_0)K_i)X(t).$$

Calculating definite integral from $t$ to $t + \Delta t$ leads to

$$\int_{t}^{t + \Delta t} X^T(Q + K_i^TP_iK_i)Xdt = X^T(t + \Delta t)P_iX(t + \Delta t) - X^T(t)P_iX(t),$$

which is (22). Hence, the proof is complete.

Therefore, the integral reinforcement relation (22) is equivalent to equation (19) in Algorithm 1, which is an extension of Kleinman’s algorithm. This indicates the convergence of Algorithm 2.

**Mismatched uncertain linear system**

The robust uncertain linear system (1) is discussed in this section. When the matched conditions (2) and (3) are not satisfied in system, we consider the design of robust tracking control law. The robust tracking problem is transformed into designing a robust control law for an augmented uncertain linear system. By calculating iteratively the optimal control law for an extended nominal system with a properly defined performance index, the online PI method is used to obtain robust tracking control law.

Obviously, the augmented system (9) is a mismatched uncertain linear system when the uncertain system (1) does not match conditions (2) and (3). In addition, the uncertainty of the augmented uncertain linear system can be decomposed into a matched part and a mismatched part according to (4) and (5):

$$T(s) - T(s_0) = [A(s) - A(s_0) \quad O]$$

$$= [B(l_0)B(l_0)^+ [A(s) - A(s_0)] + O$$

$$+ [I - B(l_0)B(l_0)^+][A(s) - A(s_0)]$$

$$= B(l_0)B(l_0)^+ [T(s) - T(s_0)$$

$$+ [I - B(l_0)B(l_0)^+][T(s) - T(s_0)]$$

and

$$\tilde{B}(l) - B(l_0) = B(l_0)B(l_0)^+ [B(l) - B(l_0)]$$

$$+ [I - B(l_0)B(l_0)^+][B(l) - B(l_0)],$$

where $B(l_0) = [B(l_0) \quad O], B(l_0)^+ = [B(l_0)^+ \quad O], I = [I \quad O], O$ is a zero matrix of appropriate dimensions.

**Lemma 3.** If Assumptions 3 and 4 hold, then the uncertainty of the augmented uncertain linear system (9) satisfies the bounded conditions as follows:

$$[T(s) - T(s_0)]^T [B(l_0)^+]^T B(l_0)^+ [T(s) - T(s_0)] \leq F \quad (25)$$

$$[T(s) - T(s_0)]^T [T(s) - T(s_0)] \leq \tilde{H} \quad (26)$$

where $F = [F \quad O]$ and $\tilde{H} = [H \quad O]$.  

**Proof.** Denote $A = A(s) - A(s_0)$. From Assumption 3, we can obtain
\[
[T(s) - T(s_0)]^T [B(l_0)^+]^T [B(l_0)^+ [T(s) - T(s_0)] = \\
\begin{bmatrix}
\dot{A}^T & O \\
O & O
\end{bmatrix}
\begin{bmatrix}
B(l_0)^+ \\
O
\end{bmatrix}^T \\
\times [B(l_0)^+ O] \\
\begin{bmatrix}
\dot{A} & O \\
O & O
\end{bmatrix} \\
= \begin{bmatrix}
\dot{A}^T [B(l_0)^+]^T B(l_0)^+ \dot{A} & O \\
O & O
\end{bmatrix} \\
= \begin{bmatrix}
F & O \\
O & O
\end{bmatrix}
\]

Therefore, (25) holds. In addition, we can achieve

\[
[T(s) - T(s_0)]^T [T(s) - T(s_0)] = \\
\begin{bmatrix}
\dot{A}^T & O \\
O & O
\end{bmatrix}
\begin{bmatrix}
\dot{A} & O \\
O & O
\end{bmatrix} = \\
\begin{bmatrix}
H & O \\
O & O
\end{bmatrix} = \hat{H}.
\]

Therefore, (26) holds, thus completing the proof.

We construct the following optimal control problem. For an extended nominal linear system,

\[
\dot{X} = T(s_0)X + \bar{B}(l_0)u + (I - \bar{B}(l_0)\bar{B}(l_0)^+)v, \quad (27)
\]

find an augmented state feedback law, \(u = KX\), \(v = LX\), in order to minimize the performance index

\[
\int_0^\infty [X^T (F + H + \beta^2 I)X + u^T u + v^T v] dt, \quad (28)
\]

where \(\beta \geq 0\) is a design parameter.

We denote \(\bar{B} = [\bar{B}(l_0) \ (I - \bar{B}(l_0)\bar{B}(l_0)^+)]\) and \(\bar{u} = [u \ v]\). Then rewriting (27) yields

\[
\dot{X} = T(s_0)X + \bar{B}\bar{u},
\]

and (28) can be rewritten as

\[
\int_0^\infty [X^T (F + H + \beta^2 I)X + \bar{u}^T \bar{u}] dt.
\]

By computing the optimal control law, it follows that

\[
\begin{bmatrix}
u \\
v
\end{bmatrix} = -\bar{B}^T PX = \begin{bmatrix}
-\bar{B}^T (l_0)P \\
-(I - \bar{B}(l_0)\bar{B}(l_0)^+) P
\end{bmatrix} X, \quad (29)
\]

where \(P\) is a positive definite matrix solving the following ARE

\[
PT(s_0) + T^2(s_0)P + F + H + \beta^2 I = PB\bar{B}^T P = 0 \quad (30)
\]

The following theorem illustrates that optimal control law (29) can stabilize the mismatched uncertain linear system (9).

**Theorem 3.** We assume that \(u = KX\) and \(v = LX\) are the solutions to the optimal control problems (27) and (28). If parameter \(\beta\) can be chosen to hold the following conditions:

\[
\beta^2 I - 2L^T L \succ 0, \quad [\bar{B}(l) - \bar{B}(l_0)]\bar{B}^T (l_0) \succ 0, \quad (31)
\]

then the optimal control law, \(u = KX\), with \(K = -\bar{B}^T (l_0)P\) can stabilize the mismatched augmented uncertain system (9). That is, for every \(s \in S, l \in L\), the mismatched closed-loop system

\[
\dot{X} = [T(s) + \bar{B}(l)K]X \quad (32)
\]

is asymptotically stable.

**Proof.** Choosing the Lyapunov function as \(V(X) = X^T PX\) and taking its time derivative, along the closed-loop system (32), we obtain

\[
\frac{dV}{dt} = X^T [(T^T(s)P + PT(s)) + 2PB(l)K]X
\]

By decomposing the matched and mismatched parts in the augmented uncertain system (9), we will have

\[
\frac{dV}{dt} = X^T [T^2(s_0)P + PT(s_0)]X
\]

According to the ARE (30)

\[
\frac{dV}{dt} = -X^T (F + H + \beta^2 I)X + X^T [P[B(l_0)\bar{B}^T (l_0)
\]

\[
\quad + (I - \bar{B}(l_0)\bar{B}(l_0)^+) (I - \bar{B}(l_0)\bar{B}(l_0)^+) ]PX
\]

\[
\quad + X^T (T(s) - T(s_0)) [T^T(l_0)\bar{B}^T (l_0)] + \bar{B}^T (l_0)PX
\]

\[
\quad + X^T (T(s) - T(s_0)) [T^T(l_0)\bar{B}^T (l_0) + T^2(s_0)P]X
\]

\[
\quad + 2X^T [P[B(l_0)\bar{B}(l_0)^+] + 2X^T [P[B(l) - \bar{B}(l_0)]K]X.
\]

By \(K = -\bar{B}^T (l_0)P\), we will have
\[
\frac{dV}{dt} = -X^T(\tilde{F} + \tilde{H} + \beta^2 I)X + X^TK^TKX \\
+ X^TP(I - \tilde{B}(l_0)\tilde{B}(l_0)^+)(I - \tilde{B}(l_0)\tilde{B}(l_0)^+)^TYPX \\
- X^T(T(s) - T(s_0))^T(\tilde{B}(l_0)^+)^T KKX \\
+ X^T(T(s) - T(s_0))^T(I - \tilde{B}(l_0)\tilde{B}(l_0)^+)^TYPX \\
- X^TK^TB(l_0)^+(T(s) - T(s_0))X \\
+ X^T(T(s) - T(s_0))^T(I - \tilde{B}(l_0)\tilde{B}(l_0)^+)^TYPX \\
- 2X^TK^TKX - 2X^TPB(l) - B(l_0)\tilde{B}(l_0)PX.
\]

By \(L = -(I - \tilde{B}(l_0)\tilde{B}(l_0)^+)^TP\), it follows that
\[
\frac{dV}{dt} = -X^T(\tilde{F} + H + \beta^2 I)X - X^TK^TKX + X^TL^TX \\
- 2X^TB(l_0)^+(T(s) - T(s_0))X \\
- 2X^T(T(s) - T(s_0))^TX \\
- 2X^TPB(l) - B(l_0)\tilde{B}(l_0)PX.
\]

Because
\[
X^TK^TKX - 2X^TK^TB(l_0)^+(T(s) - T(s_0))X \\
= -X^T[K - B(l_0)^+(T(s) - T(s_0))]^T \\
\times [K - B(l_0)^+(T(s) - T(s_0))]X \\
+ X^T[B(l_0)^+(T(s) - T(s_0))]^T \\
\times [B(l_0)^+(T(s) - T(s_0))]X \\
\leq X^T[B(l_0)^+(T(s) - T(s_0))]^T \\
\times [B(l_0)^+(T(s) - T(s_0))],
\]
and
\[
- 2X^TL^T(T(s) - T(s_0))X \\
\leq X^T(T(s) - T(s_0))^T(T(s) - T(s_0))X \\
+ X^TL^TLX \\
\leq X^TL^TX + X^THX,
\]

hence
\[
\frac{dV}{dt} = -X^T(\tilde{F} + \tilde{H} + \beta^2 I)X + X^TL^TX \\
- X^TK^TKX - 2X^TK^TB(l_0)^+(T(s) - T(s_0))X \\
- 2X^T(T(s) - T(s_0))^TL^TX \\
- 2X^TPB(l) - B(l_0)\tilde{B}(l_0)PX \\
\leq -X^T(F + H + \beta^2 I)X + 2X^TL^TX + X^THX \\
+ X^T[B(l_0)^+(T(s) - T(s_0))]^T \\
\times [B(l_0)^+(T(s) - T(s_0))]X \\
- 2X^TPB(l) - B(l_0)\tilde{B}(l_0)PX \\
\leq -2X^TPB(l) - B(l_0)\tilde{B}(l_0)PX \\
- X^T(\beta^2 I - 2L^TL)X.
\]

According to the conditions in (31)
\[
\frac{dV}{dt} < 0.
\]

Therefore, the mismatched uncertain linear system (9) is asymptotically stable for every \(s \in S\) and \(l \in L^3\). That is, \(u = -\tilde{B}(l_0)^PX\) is a robust controller for the augmented mismatched uncertain system (9). Hence, the proof is complete.

**Remark 3.** Generally, the pseudo-inverse of matrix \(B(l_0)^+\), will exist if its column vectors are not linearly dependent.\(^{29}\) In practical application, the input matrix \(B(l_0)\), is usually a matrix with column full-rank. Therefore, the pseudo-inverse of the input matrix \(B(l_0)\) is generally satisfied. Furthermore, the pseudo-inverse \(B(l_0)^+\) satisfies \(B(l_0)^+ B(l_0) = I\). However, it does not satisfy \(B(l_0)B(l_0)^+ = I\).

Here, we propose a robust tracking PI algorithm based on the RL for mismatched linear systems.

For any initial time, \(t\), the optimal performance function (28) can be written as
\[
V(X(t)) = \int_t^\infty [X^T \tilde{Q} X + u^T \tilde{u}]dt = \int_t^{\infty} [X^T \tilde{Q} X + \bar{u}^T \bar{u}]dt \\
+ \int_{t + \Delta t}^{\infty} [X^T \tilde{Q} X + u^T \tilde{u}]dt,
\]
where \(\tilde{Q} = F + H + \beta^2 I\).

Similar to Algorithm 2, the following online PI algorithm can be utilized to calculate the robust tracking law of a mismatched uncertain linear system.

**Algorithm 3.** The online PI algorithm for robust tracking of mismatched uncertain linear systems

1. Initialization: Select an initial stabilization control gain, \(\bar{K}_0\).

2. Policy evaluation: Compute \(P_i\) from
\[
X^T(t)P_iX(t) = \int_t^{t + \Delta t} [X^T \tilde{Q} X + X^T K_i^TK_iX]dt \\
+ X^T(t + \Delta t)P_iX(t + \Delta t). \tag{35}
\]

3. Policy improvement: Compute \(K_{i+1}\)
\[
K_{i+1} = -\tilde{B}^TP_i. \tag{36}
\]

The convergence of the algorithm is proven similar to that of the matched system, which is omitted here.

**Remark 4.** Robust tracking design is a challenging problem,\(^{30,33}\) particularly when the signals to track is unstable. We consider the tracking problem of a class of polynomial signals that may be unstable. Algorithm 3 is an online robust control method, which can be available for mismatched uncertain systems with unknowing nominal state matrix. Using the least squares method, the matrix \(P_i\) can be solved using (35)
through the online trajectory data of system (27). We used the PI algorithm in the RL to solve this problem, which is a novel approach. In Theorem 3, it is clear that if parameter $\beta$ is selected to be large relatively, then the conditions in (31) easily matches in many practical applications.

**Numerical experiments**

The results of two numerical experiments are demonstrated to examine the viability of the theoretical frameworks. Moreover, the proposed PI algorithms are applied to solve the robust tracking of matched and mismatched uncertain systems effectively.

**Example 1.** A matched uncertain linear system is considered as follows:

$$
\dot{x} = \begin{bmatrix}
0 & 2 \\
2 + s & s
\end{bmatrix} x + \begin{bmatrix}
0 \\
l
\end{bmatrix} u,
$$

(37)

where $s \in [0, 3]$ and $l \in [1, 3]$ are uncertain parameters. The reference signal is assumed to be $y_r = t + 5$. The control objectives of system is to design a state feedback law, $u = Kx$, which can make the system output $y = Cx$ asymptotically track the reference signal for every $s \in [0, 3]$ and $l \in [1, 3]$.

We denote $A(s) = \begin{bmatrix} 0 & 2 \\ 2 + s & s \end{bmatrix}$, $B(l) = \begin{bmatrix} 0 \\ l \end{bmatrix}$, and $C = [1, 0]$, respectively. Let $s_0 = 0$ and $l_0 = 1$. According to (8), we will have

$$
T(s) = \begin{bmatrix} 0 & 2 & 0 \\ 2 + s & s & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},
B(l) = \begin{bmatrix} 0 \\ l \\ 0 \\ 0 \end{bmatrix},
$$

$$
N = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}.
$$

Then

$$
T(s) - T(s_0) = \begin{bmatrix} 0 & 0 & 0 \\ s & s & 0 \\ 0 & 0 & 0 \end{bmatrix},
$$

(38)

and

$$
B(l) - B(l_0) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},
$$

(39)

which implies that the parameters perturbation in system (37) are matched. Obviously, $\phi(l) = (l - 1)\geq 0$ for any $l \in [1, 3]$.

According to (38), we will have

$$
\Phi^T(s)\Phi(s) = \begin{bmatrix} s^2 & s^2 & 0 & 0 \\ s^2 & s^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \leq \begin{bmatrix} 9 & 9 & 0 & 0 \\ 9 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \bar{M}.
$$

(40)

Therefore, the robust tracking problem is finally transformed into solving an optimal control of the nominal system. For the augmented nominal system,

$$
\dot{X} = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} X + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} u,
$$

design a state feedback law, $u = KX$, in order to minimize the performance index

$$
V(X(t_0), u(.)) = \int_{t_0}^{\infty} [X^T Q X + u^T u] dt
$$

(41)

where

$$
Q = \bar{M} + I = \begin{bmatrix} 10 & 9 & 0 & 0 \\ 9 & 10 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
$$

(42)

By applying Algorithm 2, a robust tracking-control gain is obtained for system (37). We chose the initial control gain as $[-14.5 \quad -8 \quad -20 \quad -6]$. The initial condition of the augmented nominal system was selected as $X_0 = [0.2 \quad 0.15 \quad 0]^T$. By using MATLAB, after four iterations, the positive definite matrix $P$ and the control gain converge to the optimal solutions as follows:

$$
P = \begin{bmatrix} 13.4446 & 8.0880 & 11.4490 & 3.2698 \\ 8.0880 & 6.5092 & 3.6855 & 1.0071 \\ 11.4490 & 3.6855 & 18.9043 & 6.0825 \\ 3.2698 & 1.0071 & 6.0825 & 3.6346 \end{bmatrix},
$$

and

$$
K = [-8.0880 \quad -6.5092 \quad -3.6855 \quad -1.0071].
$$

We used MATLAB to solve the ARE directly, which is the implementation of Algorithm 1. Consequently, the following $P$ matrix is obtained.

$$
P = \begin{bmatrix} 13.3014 & 8.0402 & 11.2417 & 3.2466 \\ 8.0402 & 6.4931 & 3.6167 & 1.0000 \\ 11.2417 & 3.6167 & 18.5987 & 6.0402 \\ 3.2466 & 1.0000 & 6.0402 & 3.6167 \end{bmatrix}.
$$
By comparison, the online PI algorithm has the same effect as the existing offline methods. But the online PI method does not need to know the nominal system matrix, which is the advantage of algorithm 2. It clearly shows that PI algorithm is effective in solving robust tracking problem. Figure 1 shows the evolution of the tracking-control signal. The convergence process of the elements in the $P$ matrix is shown in Figure 2. It shows that optimal cost is obtained at time $t = 4s$ after four updates of the controller parameters. The evolution of system output tracking reference signal with different parameters, $s$ and $l$, are presented in Figure 3, which show that robust tracking is achieved for uncertain system (37).

**Example 2.** A mismatched uncertain system is considered as follows:

$$\begin{align*}
\dot{x} &= \begin{bmatrix} s & s - 1 \\ 2 & -3 \end{bmatrix} x + \begin{bmatrix} 0 \\ l \end{bmatrix} u, \\
y &= \begin{bmatrix} x \\ 0 \end{bmatrix},
\end{align*}
$$

where $s \in [-2, 2]$ and $l \in [1, 3]$ are uncertain parameters. The reference signal is assumed to be $y_r = 3$. The control objectives of system is to design a state feedback law, $u = Kx$, which can make the system output $y = Cx$ asymptotically track the reference signal for every $s \in [-2, 2]$ and $l \in [1, 3]$. We denote $A(s) = \begin{bmatrix} s & s - 1 \\ 2 & -3 \end{bmatrix}$, $B(l) = \begin{bmatrix} 0 \\ l \end{bmatrix}$, and $C = [1, 0]$, respectively. Let $x_0 = 0$ and $l_0 = 1$. According to (8), we will have

$$\begin{align*}
T(s) &= \begin{bmatrix} s & 0 \\ 2 & 0 \\ -3 & 1 \\ 0 & 0 \end{bmatrix}, B(l) = \begin{bmatrix} 0 \\ l \end{bmatrix}, N = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \\
T(s) - T(s_0) &= \begin{bmatrix} s & 0 \\ 2 & 1 \\ -3 & 0 \\ 0 & 0 \end{bmatrix}, B(l) - B(l_0) = \begin{bmatrix} 0 \\ l - 1 \end{bmatrix}.
\end{align*}$$

It is clear that the system is a mismatched system.

$$\begin{align*}
\bar{B}(l_0)^+ &= (\bar{B}^T(l_0)\bar{B}(l_0))^{-1}\bar{B}^T(l_0) = [0 \ 1 \ 0] \\
I - \bar{B}(l_0)\bar{B}(l_0)^+ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}
\end{align*}$$

We chose $\beta = 1$. Denote $\Delta T(s) = T(s) - T(s_0)$. Therefore, we had

$$\begin{align*}
\Delta T(s)^T(\bar{B}(l_0)^+)^T\bar{B}(l_0)^+\Delta T(s) &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \bar{F}, \\
\Delta T(s)^T\Delta T(s) &= \begin{bmatrix} s^2 & s^2 & 0 \\ s^2 & s^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \leq \begin{bmatrix} 4 & 4 & 0 \\ 4 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix} = H.
\end{align*}$$

Therefore, the positive definite weight matrix $Q$ is calculated as follows:

$$Q = \bar{F} + H + \beta^2 I = \begin{bmatrix} 5 & 4 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

By applying Algorithm 3, a robust tracking-control gain is calculated for the mismatched uncertain system (43). The initial augmented control law is selected as

$$K = \begin{bmatrix} -2 & -2 \\ 2 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}.$$  

The initial condition of the augmented nominal system is selected as $x_0 = [3, -2, -8]^T$. By using MATLAB,
after four iterations, the positive definite matrix $P$ and
the augmented control gain converge to the optimal
solutions as follows:

$$P = \begin{bmatrix}
2.6301 & 0.4074 & 0.2535 \\
0.4074 & 0.6118 & -0.0741 \\
0.2535 & -0.0741 & 1.0090
\end{bmatrix}$$

and

$$K = \begin{bmatrix}
-0.4074 & -0.6118 & 0.0741 \\
-2.6301 & -0.4074 & -0.2535 \\
0 & 0 & 0 \\
-0.2535 & 0.0741 & -1.0090
\end{bmatrix}$$

To solve the ARE directly using MATLAB software, the following $P$ matrix is obtained:
It clearly shows that PI algorithm is effective in solving robust tracking problem for mismatched uncertain system. Figure 4 shows the evolution of the tracking-control signal. The convergence process of the elements in the $P$ matrix is shown in Figure 5. The evolution of the output and the reference trajectory with different parameters $s$ and $l$ are presented in Figure 6. As a result, a robust output tracking is achieved for mismatched uncertain system (43).

**Conclusion**

In this study, RL-based online PI algorithms were proposed to calculate robust tracking control law for uncertain linear systems. It was based on an online policy iteration without using a nominal system matrix. The robust tracking problem was transformed into solving an optimal control with a predefined cost function. Based on the corresponding augmented linear system, offline and online PI algorithms were established to obtain a robust tracking controller. Numerical experiments were presented to demonstrate the effectiveness of the theoretical results. The proposed method may be developed to solve tracking problems of uncertain discrete-time systems, which may be the subject of our future research.

**Acknowledgements**

The authors thank the editors and reviewers for their valuable comments. The authors also would like to thank Editage (www.editage.com) for English language editing.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partially supported by the Research Fund for High-level Talents of Liupanshui Normal University under grant LPSSYKYJ202001, and the National Natural Science Found of China under grant 61463002.

**ORCID iD**

Dengguo Xu https://orcid.org/0000-0002-5735-2002

$$P = \begin{bmatrix}
2.6232 & 0.4071 & 0.2373 \\
0.4071 & 0.6076 & -0.0730 \\
0.2373 & -0.0730 & 0.9687
\end{bmatrix}$$
References

1. Zhou K, Doyle JC, Glover K, et al. Optimal and robust control, Volume 40. New Jersey: Prentice Hall, 1996.

2. Lewis FL and Syrmos VL. Optimal control. Hoboken: John Wiley & Sons, 2012.

3. Petersen IR and Hollot CV. A riccati equation approach to the stabilization of uncertain linear systems. Automatica 1986;22(4): 397–411.

4. Schmitendorf W. Designing stabilizing controllers for uncertain systems using the riccati equation approach. IEEE Trans Automatic Control 1988; 33(4): 376–379.

5. Schmitendorf WE. A design methodology for robust stabilizing controllers. AIAA J Guid Control Dyn 1987; 10(2): 250–254.

6. Jabbari F and Schmitendorf W. A noniterative method for the design of linear robust controllers. IEEE Trans Automatic Control 1990; 35(8): 954–957.

7. Tsay SC. Robust control for linear uncertain systems via linear quadratic state feedback. Syst Control Lett 1990; 15(3): 199–205.

8. Dolphus R and Schmitendorf W. A non-iterative riccati approach to robust control design. In: American control conference, San Diego, CA, USA, 1990, pp. 916–918. IEEE.

9. Schmitendorf WE and Barmish BR. Robust asymptotic tracking for linear systems with unknown parameters. Automatica 1986; 22(3): 355–360.

10. Benson RW and Schmitendorf WE. Augmented system approach to measurement-based robust tracking. In: Proceedings of 1994 American control conference, 29 June–July 1994, vol.3, pp. 2960–2964. Baltimore, MD, USA: IEEE.

11. Shieh NC, Liang KZ and Mao CJ. Robust output tracking control of an uncertain linear system via a modified optimal linear-quadratic method. J Optimiz Theory App 2003; 117(3): 649–659.

12. Tan H, Shu S and Lin F. An optimal control approach to robust tracking of linear systems. Int J Control 2009; 82(3): 525–540.

13. Vrabie D, Pastravanu O, Abu-Khalaf M, et al. Adaptive optimal control for continuous-time linear systems based on policy iteration. Automatica 2009; 45(2): 477–484.

14. Bertsekas DP. Approximate policy iteration: a survey and some new methods. J Control Theory Appl 2011; 9(3): 310–335.

15. Kiumarsi B, Vamvoudakis KG, Modares H, et al. Optimal and autonomous control using reinforcement learning: a survey. IEEE Trans Neural Netw Learn Syst 2018; 29(6): 2042–2062.

16. Modares H and Lewis FL. Linear quadratic tracking control of partially-known continuous-time systems using reinforcement learning. IEEE Trans Automatic Contr 2014; 59(11): 3051–3056.

17. Qin C, Zhang H and Luo Y. Online optimal tracking control of continuous-time linear systems with unknown dynamics by using adaptive dynamic programming. Int J Control 2014; 87(5): 1000–1009.

18. Gao W and Jiang ZP. Adaptive dynamic programming and adaptive optimal output regulation of linear systems. IEEE Trans Automat Contr 2016; 61(12): 4164–4169.

19. Liu L, Liu Y, Tong S, et al. Integral barrier Lyapunov function based adaptive control for switched nonlinear systems. Sci China Inform Sci 2020; 63(3): 132203:1–132203:14.

20. Xie X, Zhou Q, Yue D, et al. Relaxed control design of discrete-time takagi – sugeno fuzzy systems: an event-triggered real-time scheduling approach. IEEE Trans Syst Man Cybern Syst 2018; 48(12): 2251–2262.

21. Xu D, Wang Q and Li Y. Optimal guaranteed cost tracking of uncertain nonlinear systems using adaptive dynamic programming with concurrent learning. Int J Control Autom Syst 2020; 18(5):1116–1127.

22. Liu Y, Tang L, Tong S, et al. Reinforcement learning design-based adaptive tracking control with less learning parameters for nonlinear discrete-time mimo systems. IEEE Trans Neural Netw Learn Syst 2014; 26(1): 165–176.

23. Wang D, He H and Liu D. Adaptive critic nonlinear robust control: a survey. IEEE Trans Cybern 2017; 47(10): 3429–3451.

24. Wang D, Liu D, Li H, et al. An approximate optimal control approach for robust stabilization of a class of discrete-time nonlinear systems with uncertainties. IEEE Trans Syst Man Cybern Syst 2016; 46(5): 713–717.

25. Jiang H, Zhang H, Cui Y, et al. Robust control scheme for a class of uncertain nonlinear systems with completely unknown dynamics using data-driven reinforcement learning method. Neurocomputing 2018; 273: 68–77.

26. Yang X, He H and Zhong X. Adaptive dynamic programming for robust regulation and its application to power systems. IEEE Trans Ind Electron 2018; 65(7): 5722–5732.

27. Lin F. Robust control design: an optimal control approach, Volume 18. Chichester: John Wiley & Sons, 2007.

28. Kleinman D. On an iterative technique for riccati equation computations. IEEE Trans Autom Control 1968; 13(1): 114–115.

29. Ben-Israel A and Greville TNE. Generalized inverses: theory and applications. New York: Springer, 2003.

30. Yu P, Wu M, She J, et al. Robust tracking and disturbance rejection for linear uncertain system with unknown state delay and disturbance. IEEE/ASME Trans Mechatron, 2018, 23(3): 1445–1455.

31. Liu L, Liu Y, Li D, et al. Barrier Lyapunov function based adaptive fuzzy FTC for switched systems and its applications to resistance inductance capacitance circuit system. IEEE Trans Cybern 2020; 50(8): 3491–3502.

32. Zhang J, Shi Z, Yang X, et al. Robust adaptive control for continuous wheel slip rate tracking of vehicle with state observer. Measurement Control 2020; 53(7-8):1331–1341.

33. Li R, Wang X, Liu X, et al. Adaptive neural tracking control for uncertain switched nonlinear non-lower triangular system with disturbances and dead-zone input. Int J Control Autom Syst 2020; 18(6):1445–1452.

34. Mu C and Zhang Y. Learning-based robust tracking control of quadrotor with time-varying and coupling uncertainties. IEEE Trans Neural Netw Learn Syst 2020; 31(1): 259–273.