Information Sciences 288 (2014) 314–329

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Data-based analysis of discrete-time linear systems in noisy environment: Controllability and observability

Derong Liu *, Pengfei Yan, Qinglai Wei

The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 1 December 2013
Received in revised form 24 June 2014
Accepted 29 July 2014
Available online 12 August 2014

Keywords:
Controllability
Observability
Discrete-time linear system
Data-based method
Noisy environment

ABSTRACT

In this paper, a data-based method is developed for analyzing the controllability and observability of discrete-time linear systems in noisy environment. This method uses measured data to estimate the controllability matrix and the observability matrix without identifying system models. The unbiasedness and consistency of this estimate with measurement noise and system noise are proven, respectively. As the estimated error of system parameters will not accumulate in calculating the controllability matrix and observability matrix, this method has a higher precision than traditional methods, especially in high-dimensional state space. In the simulation, the advantages of the data-based method in accuracy and convergence are illustrated.

1. Introduction

In the past twenty years, with the rapid development of science technology especially the information technology, production equipment and production process have become more and more complex. Traditional control methods, which establish precise mathematical models based on the physical–chemical mechanism to control and forecast the complex processes, become increasingly more difficult. There are large amounts of data which contain the information of production process and equipment operation generated and stored in many companies every day. How to use online and off-line data effectively to control, forecast and evaluate the production process without establishing precise mathematical models, has become an urgent problem to solve.

Data-based control appears in control fields in recent years. Since its birth, it has attracted much attention by many researchers. There are many researches that have been done on data-based control, although the calls are not the same. In 1993, Spall proposed a model-free control based on SPSA (simultaneous perturbation stochastic approximation) [24,25]. In 1994, Hjalmarsson developed iterative feedback tuning (IFT), which optimizes the controller based on gradient iteration with the measured data of closed loop control system [10–12]. In 1995, Safonov proposed unfalsified control (UC), which is a model-free adaptive control method without requiring mathematical models [21,22]. In 2000, Suardabassi and Savaresi studied virtual reference feedback tuning (VRFT) to identify parameters of controllers directly with input and output data [5,7]. Iterative learning control (ILC) proposed by Uchiyama has been one of the hot areas in control fields [6,18]. Most of iterative learning methods are based on compression mapping and fixed point theorem. Lazy learning (LL) is a kind of supervised machine learning [1,4], and was first used for control by Schaal and Atkeson in 1994 [23]. Approximate dynamic programming (ADP) is a hot topic of control theory and practice [3,19,20,29,30]. Q-learning, one of the ADP

* Corresponding author.
E-mail addresses: derong.liu@ia.ac.cn (D. Liu), pengfei.yan@ia.ac.cn (P. Yan), qinglai.wei@ia.ac.cn (Q. Wei).

http://dx.doi.org/10.1016/j.ins.2014.07.048
0020-0255/

schemes, is also a model-free control method that evaluates Q-function with off-line and online data and generates optimal control by Q-function [14,27,28].

Although the methods of data-based control are different, most of them need the assumption that the systems are controllable and observable. Controllability and observability are two basic properties in control theory [2,8,17]. Up to now, there have been many works on the state controllability and observability of linear and nonlinear systems. In 1977, Robert and Arthur extended controllability and observability from linear systems to nonlinear systems [9]. In 1993 and 1996, Liven and Narendra studied the properties of nonlinear dynamical systems using neural networks, including controllability, stabilization, observability, identification and control [15,16]. However, all of these traditional criteria of controllability and observability need to identify system equations, which are in contrary to data-based methods. How to verify these properties under the premise of not establishing models is a problem that needs to be solved, which motivates establishing data-based criteria of controllability and observability. [26] proposed data-based methods for analyzing the controllability and observability of discrete-time linear systems. It offered a novel idea to analyze the system properties with data-based method. However, the method proposed by [26] cannot deal with the problem of measurement noise and system noise. In the real world, the noise is unavoidable. Our work in this paper is to establish data-based criteria of controllability and observability in noisy environment.

In this paper, a data-based method is developed for analyzing the controllability and observability of discrete-time linear systems with noises. First, we use measured input and output data to construct the estimates of the controllability matrix and observability matrix, respectively. Second, the unbiasedness and consistency of the estimates are proven under the assumption of measurement noise and system noise, respectively. At last, the precision of our method is verified to be much smaller than traditional methods by simulations, especially in high-dimensional state space.

This paper is organized as follows. In Section 2, the criteria of controllability and observability are recalled for discrete-time linear systems. In Section 3, the validity of the controllability analysis method is proven in the measurement noisy and system noisy environment, respectively. In Section 4, the validity of observability analysis is proven in noisy environment. In Section 5, the errors of estimate with traditional methods and data-based method are discussed. Section 6 shows the performance of traditional methods and data-based method in two kinds of noisy environment.

2. Preliminaries

In this paper, we consider the following linear discrete-time control system:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k), \\ y(k) = Cx(k) + Du(k), \end{cases} \tag{1}$$

where $x(k) \in R^n, u(k) \in R^m$, and $y(k) \in R^p$ are the state, the input and the output of system (1), respectively. The matrices $A \in R^{n\times n}, B \in R^{n\times m}, C \in R^{p\times n}$ and $D \in R^{p\times m}$ are unknown, time-invariant and have no random variables as their elements.

In control theory, controllability and observability are two basic characteristics of linear systems. Controllability describes the ability of an input to move the state of a system from any initial state to any other final state in a finite time interval. Observability is a measure of how well internal states of a system can be inferred by knowledge of its external outputs [2]. The criteria of controllability and observability of linear systems in classical control theory are as follows.

Lemma 1 [2]. *The system (1) is completely state controllable, if*

$$\mathrm{rank}[W_c] = n,$$

where n is the dimension of the state $x(k)$, and $W_c = [B, AB, \ldots, A^{n-1}B]$ is named controllability matrix of the systems.

This criterion is also a necessary and sufficient condition for the state reachability. Furthermore, if the matrix A is nonsingular, then the above criterion will become a necessary and sufficient condition for the state controllability.

Lemma 2 [2]. *The system (1) is completely state observable, if*

$$\mathrm{rank}[W_o] = n,$$

where $W_o = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$, *$n$ is the dimension of the state $x(k)$, and W_o is called observability matrix of the systems.*

In traditional methods, all of the matrices in (1) need to be identified in order to analyze the system properties. Wang and Liu [26] proposed a novel data-based method, which can directly analyze the controllability and observability of the system by using measured data. However, this method cannot deal with the systems with noises. In this paper, a new data-based method is established to deal with the systems with different kinds of noises, which can also ensure higher calculation precision than traditional methods.

3. Data-Based Controllability Analysis

In this section, we present the data-based criterion of controllability of discrete-time linear systems. The estimates of controllability matrices of the systems with no noise, measurement noise and system noise are provided, respectively, and the unbiasedness and consistency of the estimates in each case are proven.

Compared with traditional approaches which have to identify the matrices A and B to compute W_c for verifying the controllability of the system, Wang and Liu [26] developed a data-based method as follows to obtain the matrix W_c directly.

Lemma 3 [26]. *Do m groups of tests on the system. For all the m groups, let all the initial states* $x^{[i]}(0) \equiv 0 \in R^n$ $(1 \leqslant i \leqslant m)$, *and the corresponding inputs*

$$u^{[i]}(k) \equiv u^{[i]} = [0, \ldots, 1, \ldots, 0] \in R^m (0 \leqslant k \leqslant n-1), \tag{2}$$

where the ith element of $u^{[i]}$ *is 1 and other elements are zeros. Then, measure* $x^{[i]}(k)$ *at time instants* $k = 1, 2, \ldots, n$. *Define*

$$\begin{aligned} &X(k) = [x^{[1]}(k), x^{[2]}(k), \ldots, x^{[m]}(k)] \ (0 \leqslant k \leqslant n), \\ &P_j = X(j) - X(j-1) \ (1 \leqslant j \leqslant n), \\ &P = [P_1, P_2, \ldots, P_n]. \end{aligned} \tag{3}$$

The system (1) is completely state controllable, if

$$\mathrm{rank}[P] = n,$$

where P is defined in (3).

Lemma 3 realizes data-based controllability criterion with m groups of special constant inputs defined by (2) and no noise. However, in practice, it is difficult to obtain such m groups of data without any noise, so a new data-based method is presented to deal with the systems with noises and more general inputs.

Theorem 1. *Do M groups of tests on the system (1). For all M groups of tests, let all the initial states* $x^{[i]}(0) \equiv 0 \in R^n$ $(1 \leqslant i \leqslant M)$, *and select any M vectors* $u^{[i]} = [u_{1i}, u_{2i}, \ldots, u_{mi}]^T \in R^m (1 \leqslant i \leqslant M)$. *Let* $U = [u^{[1]}, u^{[2]}, \ldots, u^{[M]}]$, *satisfying that U is row nonsingular* $(\mathrm{rank}(U) = m)$. *For the ith group of tests, the corresponding inputs are*

$$u^{[i]}(k) \equiv u^{[i]} = [u_{1i}, u_{2i}, \ldots, u_{mi}]^T \quad (0 \leqslant k \leqslant n-1). \tag{4}$$

Then, measure $x^{[i]}(k)$ *at time instants* $k = 1, 2, \ldots, n$. *Define*

$$\begin{aligned} &X(k) = [x^{[1]}(k), x^{[2]}(k), \ldots, x^{[M]}(k)] \quad (0 \leqslant k \leqslant n), \\ &\widehat{W}_{cj} = (X(j) - X(j-1))U^T (UU^T)^{-1} \quad (1 \leqslant j \leqslant n), \\ &\widehat{W}_c = [\widehat{W}_{c1}, \widehat{W}_{c2}, \ldots, \widehat{W}_{cn}]. \end{aligned} \tag{5}$$

System (1) is completely state controllable if

$$\mathrm{rank}[\widehat{W}_c] = n,$$

where $\widehat{W}_c$ *is defined in (5).*

Proof. For the ith test, with the initial state $x^{[i]}(0) \equiv 0$ and the control inputs as $u^{[i]}(k)$ in (4), the state measurements will be

$$\begin{cases} x^{[i]}(1) = Bu^{[i]} \\ x^{[i]}(2) = ABu^{[i]} + Bu^{[i]} \\ \vdots \\ x^{[i]}(k) = A^{k-1}Bu^{[i]} + \cdots + ABu^{[i]} + Bu^{[i]} \\ \vdots \\ x^{[i]}(n) = A^{n-1}Bu^{[i]} + \cdots + ABu^{[i]} + Bu^{[i]}. \end{cases} \tag{6}$$

According to (6), the equations can be written in matrix form as

$$\begin{cases} X(1) = BU \\ X(2) = (AB + B)U \\ \vdots \\ X(k) = (A^{k-1}B + A^{k-2}B + \cdots + B)U \\ \vdots \\ X(n) = (A^{n-1}B + A^{n-2}B + \cdots + B)U, \end{cases} \tag{7}$$

where $X(k)$ and U are defined in Theorem 1. As $X(0) = [x^{[1]}(0), x^{[2]}(0), \ldots, x^{[M]}(0)] = 0$, according to (7), we have

$$\begin{cases} X(1) - X(0) = BU \\ X(2) - X(1) = ABU \\ \vdots \\ X(j) - X(j-1) = A^{j-1}BU \\ \vdots \\ X(n) - X(n-1) = A^{n-1}BU. \end{cases}$$

Since U is row nonsingular, the estimate can be obtained

$$\begin{cases} \widehat{W}_{c1} = (X(1) - X(0))U^T(UU^T)^{-1} = B \\ \widehat{W}_{c2} = (X(2) - X(1))U^T(UU^T)^{-1} = AB \\ \vdots \\ \widehat{W}_{cj} = (X(j) - X(j-1))U^T(UU^T)^{-1} = A^{j-1}B \\ \vdots \\ \widehat{W}_{cn} = (X(n) - X(n-1))U^T(UU^T)^{-1} = A^{n-1}B. \end{cases}$$

As a result, the estimate of W_c is

$$\begin{aligned} \widehat{W}_c &= [\widehat{W}_{c1}, \widehat{W}_{c2}, \ldots, \widehat{W}_{cn}] \\ &= [B, AB, \ldots, A^{n-1}B] \\ &= W_c, \end{aligned}$$

where W_c is the controllability matrix of system (1), which is defined in Lemma 1. The proof is completed. □

Theorem 1 realizes the data-based controllability criterion with any M groups of nonzero constant inputs. As $M \gg m$ in practice, the assumption of $\mathrm{rank}(U) = m$ is usually satisfied.

Remark 1. In Theorem 1, we set the initial states of the m groups of tests to be zero. Actually, the initial states can take any values. If the initial states are not zero ($x^{[i]}(0) \neq 0 \in R^n$), just do one more group of test with the input always zero ($u^{[0]}(k) \equiv 0$ for $\forall k \geqslant 0$). Then we have

$$\begin{cases} x^{[0]}(1) = Ax^{[0]}(0) \\ x^{[0]}(2) = A^2x^{[0]}(0) \\ \vdots \\ x^{[0]}(n) = A^nx^{[0]}(0). \end{cases}$$

For each of the other m groups of tests with the initial states nonzero and the input defined by (4), we have

$$\begin{cases} x^{[i]}(1) = Ax^{[0]}(0) + Bu^{[i]} \\ x^{[i]}(2) = A^2x^{[0]}(0) + ABu^{[i]} + Bu^{[i]} \\ \vdots \\ x^{[i]}(n) = A^nx^{[0]}(0) + A^{n-1}Bu^{[i]} + \cdots + ABu^{[i]} + Bu^{[i]}. \end{cases}$$

Redefine $X(k)$ as

$$\begin{aligned} X(k) = [&x^{[1]}(k) - x^{[0]}(k), x^{[2]}(k) - x^{[0]}(k), \ldots, \\ &x^{[n]}(k) - x^{[0]}(k)] \quad (0 \leqslant k \leqslant n). \end{aligned}$$

Then, we can obtain $\widehat{W}_c$ by the same way as in Theorem 1.

Corollary 1. *Assume that the measured data contains measurement noise.*

$$\hat{x}(k) = x(k) + \omega_m(k),$$

where $\omega_m(k) \in R^n$ is measurement noise, and satisfies a normal distribution with expected value 0 and standard deviation σ, i.e., $\omega_m(k) \sim \mathcal{N}(0, \sigma^2 I_n)$. According to Theorem 1, the estimate of W_c is given by,

$$\widehat{W}_c = [\widehat{W}_{c1}, \widehat{W}_{c2}, \ldots, \widehat{W}_{cn}],$$
$$\widehat{W}_{cj} = (\widehat{X}(j) - \widehat{X}(j-1))U^T(UU^T)^{-1}, \ j = 1, 2, \ldots, n,$$

where $\widehat{X}(j) \in R^{n\times M}$ is the measured value of $X(j), W_c = [W_{c1}, W_{c2}, \ldots, W_{cn}] = [B, AB, \ldots, A^{n-1}B]$ is the controllability matrix of (1). Then $\widehat{W}_c$ is an unbiased and consistent estimate of W_c.

Proof. The proof is divided into two parts.

(i) The proof of unbiasedness

Since $\widehat{X}(j)$ is the measured value of $X(j)$, according to the definition of $X(j)$ and $\hat{x}(k)$

$$\begin{aligned}\widehat{X}(j) &= [\hat{x}^{[1]}(j), \hat{x}^{[2]}(j), \ldots, \hat{x}^{[M]}(j)] \\ &= [x^{[1]}(j) + \omega_m^{[1]}(j), x^{[2]}(j) + \omega_m^{[1]}(j), \ldots, x^{[M]}(j) + \omega_m^{[1]}(j)] \\ &= X(j) + \Omega(j),\end{aligned}$$

where $\Omega(j) = [\omega_m^{[1]}(j), \omega_m^{[2]}(j), \ldots, \omega_m^{[M]}(j)]$. Since $\omega_m^{[i]}(j), i = 1, 2, \ldots, M$, comes from different groups of tests, it is reasonable to assume that each one of them is independent from the another. So the expected value of $\widehat{W}_{cj}$ is

$$\begin{aligned}E\{\widehat{W}_{cj}\} &= E\{(\widehat{X}(j) - \widehat{X}(j-1))U^T(UU^T)^{-1}\} \\ &= E\{\widehat{X}(j) - \widehat{X}(j-1)\}U^T(UU^T)^{-1} \\ &= E\{X(j) + \Omega(j) - X(j-1) - \Omega(j-1)\}U^T(UU^T)^{-1} \\ &= (X(j) - X(j-1))U^T(UU^T)^{-1} + E\{\Omega(j) - \Omega(j-1)\}U^T(UU^T)^{-1}.\end{aligned} \tag{8}$$

As $E\{\Omega(j)\} = 0, j = 1, 2, \ldots, n$, we get

$$\begin{aligned}E\{\widehat{W}_{cj}\} &= (X(j) - X(j-1))U^T(UU^T)^{-1} \\ &= W_{cj} \\ &= A^{j-1}B.\end{aligned}$$

Therefore, we have

$$E\{\widehat{W}_c\} = W_c,$$

which proves the unbiasedness.

(ii) The proof of consistency

According to the translational invariance of variance and (8), the variance of $\widehat{W}_{cj}$ is

$$\mathrm{Var}\{\widehat{W}_{cj}\} = \mathrm{Var}\{(\Omega(j) - \Omega(j-1))U^T(UU^T)^{-1}\}.$$

Define

$$D(j) = \Delta\Omega(j)U^T(UU^T)^{-1},$$

where $\Delta\Omega(j) = \Omega(j) - \Omega(j-1)$. Then the variance of $\widehat{W}_{cj}$ can be rewritten as

$$\mathrm{Var}\{\widehat{W}_{cj}\} = \mathrm{Var}(D(j)).$$

Let $d_i(j)$ be the ith row of $D(j), i = 1, 2, \ldots, n$, and $\Delta\Omega_i(j)$ is the ith row of $\Delta\Omega(j)$, $i = 1, 2, \ldots, n$. Then we have

$$d_i(j) = \Delta\Omega_i(j)U^T(UU^T)^{-1}.$$

According to the definition of $\Delta\Omega(j)$,

$$E\{d_i(j)\} = E\{\Delta\Omega_i(j)\}U^T(UU^T)^{-1} = 0.$$

So the variance of $d_i(j)$ is

$$\begin{aligned}\mathrm{Var}\{d_i(j)\} &= E\{d_i^T(j)d_i(j)\}\\ &= E\{(UU^T)^{-1}U\Delta\Omega_i^T(j)\Delta\Omega_i(j)U^T(UU^T)^{-1}\}\\ &= (UU^T)^{-1}UE\{\Delta\Omega_i^T(j)\Delta\Omega_i(j)\}U^T(UU^T)^{-1}.\end{aligned} \tag{9}$$

According to the definition of $\Delta\Omega_i(j)$, we can obtain

$$\begin{aligned}E\{\Delta\Omega_i^T(j)\Delta\Omega_i(j)\} &= E\{(\Omega_i(j)-\Omega_i(j-1))^T(\Omega_i(j)-\Omega_i(j-1))\}\\ &= E\{\Omega_i^T(j)\Omega_i(j)\}+E\{\Omega_i^T(j-1)\Omega_i(j-1)\}\\ &= 2\sigma^2 I_M,\end{aligned} \tag{10}$$

where I_M is an identity matrix of order M. As $\Omega(j)$ and $\Omega(j-1)$ are independent, we have $E\{\Omega_i^T(j)\Omega_i(j-1)\} = E\{\Omega_i^T(j-1)\Omega_i(j)\} = 0$. By (9) and (10), we have

$$\mathrm{Var}\{d_i(j)\} = 2\sigma^2(UU^T)^{-1}.$$

According to [13], as $M \to \infty$, we have

$$(UU^T)^{-1} \to 0.$$

So the variance of $d_i(j)$ is

$$\mathrm{Var}\{d_i(j)\} \to 0,\ \text{as } M \to \infty.$$

Furthermore, $\Omega_{i_1}(j)$ and $\Omega_{i_2}(j)$ are independent, when $i_1 \neq i_2$. So we can get

$$\mathrm{Cov}(d_{i_1}(j), d_{i_2}(j)) = 0,\ \text{for } i_1 \neq i_2,$$

where $\mathrm{Cov}(a,b)$ is the covariance of a and b. Then, the following conclusion can be obtained

$$\mathrm{Var}\{\widehat{W}_{cj}\} = \mathrm{Var}\{D(j)\} \to 0,\ \text{as } M \to \infty.$$

This means that, as an estimate of W_{cj}, $\widehat{W}_{cj}$ is consistent. Furthermore, $\widehat{W}_c$ is also the consistent estimate of W_c. Combining (i) and (ii), the proof is completed. □

In practice, the expectation of measurement noise is usually zero. If not, we can estimate the expectation by identification methods. Let the estimate of expectation of measurement noise $\widehat{E}(\omega_m(k)) = \hat{\mu}_k$. Then we have $\omega_m'(k) = \omega_m(k) - \hat{\mu}_k$ with the expected value 0. Then we can analyze the controllability of the system by Corollary 1. In the rest of this article, the expectation of noise is also assumed to be zero.

Corollary 2. *Assume that the system contains noise*

$$x(k+1) = Ax(k) + Bu(k) + \omega_s(k),$$

where $\omega_s(k) \in R^n$ is system noise, and satisfies a normal distribution with expected value 0 and standard deviation σ, i.e., $\omega_s(k) \sim \mathcal{N}(0, \sigma^2 I_n)$. According to Theorem 1, the estimate of W_c is given by,

$$\begin{aligned}&\widehat{W}_c = [\widehat{W}_{c1}, \widehat{W}_{c2}, \ldots, \widehat{W}_{cn}],\\ &\widehat{W}_{cj} = (X(j) - X(j-1))U^T(UU^T)^{-1},\ j = 1,2,\ldots,n,\end{aligned}$$

where $X(j) \in R^{n\times M}$ is the measured state of the system with noise, $W_c = [W_{c1}, W_{c2}, \ldots, W_{cn}] = [B, AB, \ldots, A^{n-1}B]$ is the controllability matrix of (1). Then $\widehat{W}_c$ is an unbiased and consistent estimate of W_c.

Proof. The proof is divided into two parts.

(i) *The proof of unbiasedness*

For the ith test, according to system equation with noise, the state measurements and control inputs have the relationship as follow

$$\begin{cases} x^{[i]}(1) = Bu^{[i]} + \omega_s^{[i]}(0) \\ x^{[i]}(2) = ABu^{[i]} + Bu^{[i]} + A\omega_s^{[i]}(0) + \omega_s^{[i]}(1) \\ \vdots \\ x^{[i]}(k) = A^{k-1}Bu^{[i]} + \cdots + ABu^{[i]} + Bu^{[i]} + A^{k-1}\omega_s^{[i]}(0) \\ \qquad\qquad + \cdots + A\omega_s^{[i]}(k-2) + \omega_s^{[i]}(k-1) \\ \vdots \\ x^{[i]}(n) = A^{n-1}Bu^{[i]} + \cdots + ABu^{[i]} + Bu^{[i]} + A^{n-1}\omega_s^{[i]}(0) \\ \qquad\qquad + \cdots + A\omega_s^{[i]}(n-2) + \omega_s^{[i]}(n-1). \end{cases}$$

According to the definition of X and U in Theorem 1, we have

$$\begin{cases} X(1) = BU + \Omega_s(0) \\ X(2) = (AB + B)U + A\Omega_s(0) + \Omega_s(1) \\ \vdots \\ X(k) = (A^{k-1}B + \cdots + AB + B)U + A^{k-1}\Omega_s(0) \\ \qquad\qquad + \cdots + A\Omega_s(k-2) + \Omega_s(k-1) \\ \vdots \\ X(n) = (A^{n-1}B + \cdots + AB + B)U + A^{n-1}\Omega_s(0) \\ \qquad\qquad + \cdots + A\Omega_s(n-2) + \Omega_s(n-1), \end{cases}$$

where $\Omega_s(k) = [\omega_s^{[1]}(k), \omega_s^{[2]}(k), \ldots, \omega_s^{[M]}(k)]$. As $\omega_s(k), k = 0, 1, \ldots, n$, are independent and identically distributed, $\Omega_s(k), k = 0, 1, \ldots, n$, are also independent and identically distributed. Let $\Omega_s(k) = \Omega_s, k = 0, 1, \ldots, n$. Then we can obtain

$$\begin{cases} X(1) = BU + \Omega_s \\ X(2) = (AB + B)U + (A + I_n)\Omega_s \\ \vdots \\ X(k) = (A^{k-1}B + \cdots + AB + B)U \\ \qquad\quad +(A^{k-1} + \cdots + A + I_n)\Omega_s \\ \vdots \\ X(n) = (A^{n-1}B + \cdots + AB + B)U \\ \qquad\quad +(A^{n-1} + \cdots + A + I_n)\Omega_s. \end{cases} \tag{11}$$

As $X(0) = [x^{[1]}(0), x^{[2]}(0), \ldots, x^{[M]}(0)] = 0$, according to (11), we have

$$\begin{cases} X(1) - X(0) = BU + \Omega_s \\ X(2) - X(1) = ABU + A\Omega_s \\ \vdots \\ X(k) - X(k-1) = A^{k-1}BU + A^{k-1}\Omega_s \\ \vdots \\ X(n) - X(n-1) = A^{n-1}BU + A^{n-1}\Omega_s. \end{cases}$$

So the expected value of $\widehat{W}_{cj}$ is

$$\begin{aligned} E\{\widehat{W}_{cj}\} &= E\{(X(j) - X(j-1))U^T(UU^T)^{-1}\} \\ &= E\{(A^{j-1}BU + A^{j-1}\Omega_s)U^T(UU^T)^{-1}\} \\ &= A^{j-1}B + E\{A^{j-1}\Omega_s U^T(UU^T)^{-1}\} \\ &= A^{j-1}B + A^{j-1}E\{\Omega_s\}U^T(UU^T)^{-1} \\ &= A^{j-1}B. \end{aligned}$$

Furthermore, the expected value of $\widehat{W}_c$ can be obtained as

$$\begin{aligned} E\{\widehat{W_c}\} &= E\{[\widehat{W_{c1}}, \widehat{W_{c2}}, \ldots, \widehat{W_{cn}}]\} \\ &= [E\{\widehat{W_{c1}}\}, E\{\widehat{W_{c2}}\}, \ldots, E\{\widehat{W_{cn}}\}] \\ &= [B, AB, \ldots, A^{n-1}B] \\ &= W_c, \end{aligned}$$

which proves the unbiasedness.

(ii) The proof of consistency

Similar to Corollary 1, we can get

$$\mathrm{Var}\{\widehat{W}_{cj}\} = \mathrm{Var}\{A^{j-1}\Omega_s U^T (UU^T)^{-1}\}.$$

Let $D(j) = A^{j-1}\Omega_s U^T(UU^T)^{-1}, H(j) = A^{j-1}\Omega_s$, $d_i(j)$ and $h_i(j)$ are the ith row of $D(j)$ and $H(j)$, respectively. According to the property of multivariate normal distribution,

$$h_i(j) \sim \mathcal{N}(0, \Sigma).$$

As M groups of tests are independent, the elements of Σ are all zeros except the main diagonal. Let $\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_M \end{bmatrix}$ and $\sigma_{\max} = \max\{\sigma_1, \sigma_2, \ldots, \sigma_M\}$. σ_i is determined by A and σ, so $\sigma_{\max}$ is bounded when $M \to \infty$. Then the variance of $d_i(j)$ can be obtained as follows.

$$\begin{aligned} \mathrm{Var}\{d_i(j)\} &= \mathrm{Var}\{h_i(j)U^T(UU^T)^{-1}\} \\ &= E\{(UU^T)^{-1}Uh_i^T(j)h_i(j)U^T(UU^T)^{-1}\} \\ &= (UU^T)^{-1}UE\{h_i^T(j)h_i(j)\}U^T(UU^T)^{-1} \\ &= (UU^T)^{-1}U\Sigma U^T(UU^T)^{-1}. \end{aligned}$$

Taking the norm of $\mathrm{Var}\{d_i(j)\}$ leads to

$$\|\mathrm{Var}\{d_i(j)\}\| \leqslant \|(UU^T)^{-1}U\sigma_{\max}U^T(UU^T)^{-1}\| = \sigma_{\max}\|(UU^T)^{-1}\|.$$

According to [13], as $M \to \infty$, we have

$$(UU^T)^{-1} \to 0.$$

So we have

$$\mathrm{Var}\{d_i(j)\} \to 0, \quad \text{as } M \to \infty.$$

Then, the variance of the $\widehat{W}_{cj}$ is

$$\mathrm{Var}\{\widehat{W}_{cj}\} = \mathrm{Var}\{D(j)\} \to 0, \quad \text{as } M \to \infty.$$

This means that, as the estimate of W_{cj}, $\widehat{W}_{cj}$ is consistent. Furthermore, $\widehat{W_c}$ is a consistent estimate of W_c. Combining (i) and (ii), the proof is completed. □

Remark 2. For any initial state $x(0) \neq 0 \in R^n$, As in Remark 1, do one more group of test with the initial state $x(0)$ and the input $u^{[0]} \equiv 0 \in R^m$ for $\forall k \geqslant 0$. Then we have

$$x^0(n) = A^n x(0). \tag{12}$$

According to the system Eq. (1), we have

$$x(n) = A^n x(0) + A^{n-1}Bu(0) + \cdots + ABu(n-2) + Bu(n-1).$$

Let $x(n) = 0$. Then

$$[B, AB, \ldots, A^{n-1}B]\begin{bmatrix} u(n-1) \\ \vdots \\ u(1) \\ u(0) \end{bmatrix} = -A^n x(0).$$

Combining the definition of W_c and (12), we have

$$W_c \begin{bmatrix} u(n-1) \\ \vdots \\ u(1) \\ u(0) \end{bmatrix} = -x^0(0).$$

Since $\mathrm{rank}(W_c) = n$, the equations have solutions. Especially, when the input space is one dimensional space ($u(k) \in R$), the equations above have a unique solution.

Remark 2 shows us that if a linear discrete-time system is controllable, it can move any initial state to zero within at most n steps. It also provides a new idea to design the controller based on the controllability matrix, which contains more information about the system than the system parameter matrices A and B.

4. Data-Based Observability Analysis

Similar to the controllability, the method for analyzing the observability of the system (1) can be obtained. [26] established data-based criterion of observability by the following lemma.

Lemma 4. *Let n initial states of the system be given as follows,*

$$x^{[i]}(0) = [0, \ldots, 0, 1, 0, \ldots, 0]^T \in R^n \quad (1 \leqslant i \leqslant n), \tag{13}$$

where the ith element of $x^{[i]}(0)$ is 1, and all other elements are zeros. Then, measure the corresponding output $y^{[i]}(k)$ at time instants $k = 0, 1, \ldots, n-1$, while setting the inputs as $u^{[i]}(k) \equiv 0 \in R^m$. Define

$$Y(k) = [y^{[1]}(k), y^{[2]}(k), \ldots, y^{[n]}(k)] \quad (0 \leqslant k \leqslant n).$$

Assume that the initial states of (1) can be set as in (13). Then the system is completely state observable, if and only if

$$\mathrm{rank}[Q] = n.$$

Similar to the previous discussion of controllability, we improved this algorithm to make it applicable to more general data. First, do N groups of tests on the system (1). Select any N linearly independent vectors $x^{[i]} = [x_{1i}, x_{2i}, \ldots, x_{ni}]^T$ $(1 \leqslant i \leqslant N)$ as the initial states of system (1), and let all the inputs $u^{[i]}(k) \equiv 0 \in R^m$ $(1 \leqslant i \leqslant N, 0 \leqslant k \leqslant n-1)$. Let $X = [x^{[1]}, x^{[2]}, \ldots, x^{[N]}]$, and then X is nonsingular. Store the measured output data $y^{[i]}(0), y^{[i]}(1), \ldots, y^{[i]}(n-1)$, for $i = 1, 2, \ldots, N$. Define

$$Y(k) = [y^{[1]}(k), y^{[2]}(k), \ldots, y^{[N]}(k)] \quad (0 \leqslant k \leqslant n-1).$$

Then, according to the definitions of X and $Y(k)$, we can obtain the following matrix

$$\widehat{W}_o = \begin{bmatrix} Y(0)X^T(XX^T)^{-1} \\ Y(1)X^T(XX^T)^{-1} \\ \vdots \\ Y(n-1)X^T(XX^T)^{-1} \end{bmatrix}. \tag{14}$$

Theorem 2. *System (1) is completely state observable if*

$$\mathrm{rank}[\widehat{W}_o] = n,$$

where $\widehat{W}_o$ is defined in (14).

Proof. With the initial states $x^{[i]}(0) = x^{[i]}$, and the inputs $u^{[i]}(k) \equiv 0$, the output of (1) will be

$$\begin{cases} y^{[i]}(0) = Cx^{[i]}(0) = Cx^{[i]} \\ y^{[i]}(1) = CAx^{[i]}(0) = CAx^{[i]} \\ \vdots \\ y^{[i]}(n-1) = CA^{n-1}x^{[i]}(0) = CA^{n-1}x^{[i]}, \end{cases}$$

where A and C are unknown. From the definition of $x^{[i]}$, we have

$$\begin{cases} [y^{[1]}(0), y^{[2]}(0), \ldots, y^{[n]}(0)] = C[x^{[1]}, x^{[2]}, \ldots, x^{[n]}] \\ [y^{[1]}(1), y^{[2]}(1), \ldots, y^{[n]}(1)] = CA[x^{[1]}, x^{[2]}, \ldots, x^{[n]}] \\ \vdots \\ [y^{[1]}(n), y^{[2]}(n), \ldots, y^{[n]}(n)] = CA^{n-1}[x^{[1]}, x^{[2]}, \ldots, x^{[n]}]. \end{cases}$$

According to the definitions of X and U, we can get

$$\begin{cases} Y(0) = CX \\ Y(1) = CAX \\ \vdots \\ Y(n-1) = CA^{n-1}X. \end{cases}$$

Since X is nonsingular, we can obtain

$$\widehat{W}_o = \begin{bmatrix} Y(0)X^T(XX^T)^{-1} \\ Y(1)X^T(XX^T)^{-1} \\ \vdots \\ Y(n-1)X^T(XX^T)^{-1} \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = W_o,$$

where W_o is the observability matrix of (1). According to Lemma 2, the proof is completed. □

Corollary 3. *Assume that the measured data of output $y(k)$ contains measurement noise,*

$$\hat{y}(k) = y(k) + v_m(k),$$

where $v_m(k) \in R^p$ is the measurement noise, and satisfies a normal distribution with expected value 0 and standard deviation σ, i.e., $v_m(k) \sim \mathcal{N}(0, \sigma^2 I_p)$. According to Theorem 2, the estimate of W_o is given by,

$$\widehat{W_o} = [\widehat{W}_{o1}, \widehat{W}_{o2}, \ldots, \widehat{W}_{on}],$$
$$\widehat{W}_{oj} = \widehat{Y}(j-1)X^T(XX^T)^{-1}, \; j = 1, 2, \ldots, n,$$

where $\widehat{Y}(j) \in R^{p\times M}$ is the measured value of $Y(j)$. Then $\widehat{W}_o$ is an unbiased and consistent estimate of W_o.

Proof. The proof is divided into two parts.

(i) *The proof of unbiasedness*

Since $\widehat{Y}(j)$ is the measured value of $Y(j)$, according to the definition of $Y(j)$ and $\hat{y}(k)$

$$\begin{aligned} \widehat{Y}(j) &= [\hat{y}^{[1]}(j), \hat{y}^{[2]}(j), \ldots, \hat{y}^{[M]}(j)] \\ &= [y^{[1]}(j) + v_m^{[1]}(j), y^{[2]}(j) + v_m^{[2]}(j), \ldots, y^{[M]}(j) + v_m^{[M]}(j)] \\ &= Y(j) + V(j), \end{aligned}$$

where $V(j) = [v_m^{[1]}(j), v_m^{[2]}(j), \ldots, v_m^{[M]}(j)]$. Since $v_m^{[i]}(j), i = 1, 2, \ldots, M$, comes from different groups of tests, it is reasonable to assume that each one of them is independent from others,

$$\begin{aligned} E\{\widehat{W}_{oj}\} &= E\left\{\widehat{Y}(j-1)X^T(XX^T)^{-1}\right\} \\ &= E\left\{\widehat{Y}(j-1)\right\}X^T(XX^T)^{-1} \\ &= E\{Y(j-1) + V(j-1)\}X^T(XX^T)^{-1} \\ &= Y(j-1)X^T(XX^T)^{-1} + E\{V(j-1)\}X^T(XX^T)^{-1}. \end{aligned}$$

As $E\{V(j)\} = 0$ $(j = 0, 1, \ldots, n)$, we have

$$\begin{aligned} E\{\widehat{W}_{oj}\} &= Y(j-1)X^T(XX^T)^{-1} \\ &= W_{oj} = CA^{j-1}. \end{aligned}$$

Therefore, we can conclude

$$E\{\widehat{W}_o\} = W_o.$$

The unbiasedness is proven.

(ii) *The proof of consistency*

According to the translational invariance of variance, we can obtain

$$\mathrm{Var}\{\widehat{W}_{oj}\} = \mathrm{Var}\{V(j-1)X^T(XX^T)^{-1}\}.$$

As $V(j), j = 0, 1, \ldots, n$, are independent and identically distributed, we can let $V(j) = V$. Define

$$D = VX^T(XX^T)^{-1}.$$

Let d_i be the ith row of D and v_i be the ith row of V. So we have

$$d_i = v_i X^T(XX^T)^{-1}.$$

According to the assumption of V, we have

$$E\{d_i\} = 0.$$

So the variance of d_i is

$$\begin{aligned}\mathrm{Var}\{d_i\} &= E\left\{d_i^T d_i\right\} \\ &= E\left\{(XX^T)^{-1}X \cdot v_i^T v_i \cdot X^T(XX^T)^{-1}\right\} \\ &= (XX^T)^{-1}X \cdot E\{v_i^T v_i\} \cdot X^T(XX^T)^{-1}.\end{aligned} \tag{15}$$

According to the definition of v_i and V, we can obtain

$$E\{v_i^T v_i\} = \sigma^2 I_M. \tag{16}$$

Substituting (16) into (15), $\mathrm{Var}\{d_i\}$ can be written as

$$\mathrm{Var}\{d_i\} = \sigma^2(XX^T)^{-1}.$$

According to [13], as $M \to \infty$, we have

$$(XX^T)^{-1} \to 0.$$

So we can obtain

$$\mathrm{Var}\{d_i\} \to 0, \quad \text{as } M \to \infty.$$

Then we have

$$\mathrm{Var}\{\widehat{W}_{oj}\} = \mathrm{Var}\{D\} \to 0, \quad \text{as } M \to \infty.$$

This means that, $\widehat{W}_{oj}$ is a consistent estimate of W_{oj}. Furthermore, $\widehat{W}_o$ is also the consistent estimate of W_o. Combining (i) and (ii), the proof is completed. □

Corollary 4. *Assume that the system contains noise*

$$y(k) = Cx(k) + v_s(k),$$

where $v_s(k) \in R^p$ is system noise, and satisfies a normal distribution with expected value 0 and standard deviation σ, i.e., $v_s(k) \sim \mathcal{N}(0, \sigma^2)$. According to Theorem 2, the estimate of W_o is given by,

$$\begin{aligned}&\widehat{W}_o = [\widehat{W}_{o1}, \widehat{W}_{o2}, \ldots, \widehat{W}_{on}], \\ &\widehat{W}_{oj} = Y(j-1)X^T(XX^T)^{-1}, \; j = 1, 2, \ldots, n,\end{aligned}$$

where $Y(j) \in R^{p\times M}$ is the measured output of the system with noise. Then, $\widehat{W}_o$ is an unbiased and consistent estimate of W_o.

Proof. Under the conditions of this corollary, it is equivalent to Corollary 3. The proof is omitted. □

5. Precision Analysis

As mentioned previously, our method does not identify system matrices. So the corresponding identification errors of A, B, C and D will not present in the result. Taking the controllability for example, assume that the estimates of A and B in traditional methods are

$$\widetilde{A} = A + \Delta A,$$
$$\widetilde{B} = B + \Delta B.$$

By the above formula, the estimate of W_c can be obtained as

$$\begin{aligned}\widetilde{W}_c &= [\widetilde{B}, \widetilde{A}\widetilde{B}, \ldots, \widetilde{A}^{n-1}\widetilde{B}] \\ &= [\widetilde{W}_{c1}, \widetilde{W}_{c2}, \ldots, \widetilde{W}_{cn}].\end{aligned}$$

So we have

$$\Delta\widetilde{W}_c = \widetilde{W}_c - W_c.$$

For the ith block of $\widetilde{W}_c$, we have

$$\begin{aligned}\widetilde{W}_{ci} &= \widetilde{A}^{i-1}\widetilde{B} \\ &= (A + \Delta A)^{i-1}(B + \Delta B) \\ &= (A^{i-1} + \mathcal{C}_{i-1}^{1}A^{i-1}\Delta A + \cdots \\ &\quad + (\Delta A)^{i-1})(B + \Delta B).\end{aligned}$$

where $\mathcal{C}_n^j$ is binomial coefficient. Ignore the high order terms of errors, then we have

$$\begin{aligned}\widetilde{W}_{ci} &\approx (A^{i-1} + (i-1)A^{i-1}\Delta A)(B + \Delta B) \\ &= A^{i-1}B + (i-1)A^{i-1}\Delta AB + A^{i-1}\Delta B \\ &\quad + (i-1)A^{i-1}\Delta A\Delta B \\ &\approx A^{i-1}B + (i-1)A^{i-1}\Delta AB + A^{i-1}\Delta B.\end{aligned}$$

Let $\Delta\widetilde{W}_{ci} = \widetilde{W}_{ci} - W_{ci}$, and we can get

$$\begin{aligned}\Delta\widetilde{W}_{ci} &= \widetilde{W}_{ci} - W_{ci} \\ &\approx (i-1)A^{i-1}\Delta AB + A^{i-1}\Delta B.\end{aligned}$$

It can be seen from the above result that $\Delta\widetilde{W}_{ci}$ will grow exponentially with the index i increasing. Therefore, we can conclude that the estimation error $\|\Delta\widetilde{W}_c\|$ in traditional methods grows exponentially with the dimension of state space n. The criteria of controllability in traditional methods may be infeasible in the high dimension space.

To overcome this difficulty, the data-based method is proposed, which can keep the estimation error growing linearly. The estimate of W_c can be obtained with the Theorem 1 and its corollaries,

$$\widehat{W}_c = [\widehat{W}_{c1}, \widehat{W}_{c2}, \ldots, \widehat{W}_{cn}].$$

For each $\widehat{W}_{ci}, i = 1, \ldots, n$, its error matrix is

$$\Delta\widehat{W}_{ci} = \widehat{W}_{ci} - W_{ci}.$$

The estimation error of W_c is

$$\begin{aligned}\|\Delta\widehat{W}_c\| &= \|\widehat{W}_c - W_c\| \\ &\leqslant \sum_{i=1}^{n}\|\Delta\widehat{W}_{ci}\| \\ &\leqslant n\hat{\Delta}.\end{aligned}$$

where $\hat{\Delta} = \max\{\|\Delta\widehat{W}_{c1}\|, \|\Delta\widehat{W}_{c2}\|, \ldots, \|\Delta\widehat{W}_{cn}\|\}$. Therefore, we can conclude that the upper bound of estimation error $\|\Delta\widehat{W}_c\|$ is linear growth with the dimension of state space n.

6. Simulation Study

In this section, two examples of systems containing measurement noise and system noise respectively are provided. The estimation errors with traditional methods and data-based method are illustrated in the simulation results.

6.1. Discrete-Time Linear Systems with Measurement Noise

Assume that the measured data contains measurement noise

$$\begin{cases}\hat{x}(k) = x(k) + \omega_m(k), \\ \hat{y}(k) = y(k) + v_m(k),\end{cases}$$

where $\omega_m(k) \sim N(0, \sigma_{m1} I_n)$ and $v_m(k) \sim N(0, \sigma_{m2} I_p)$. In this example, we set $\sigma_m = \sigma_{m1} = \sigma_{m2}$. The system equation is defined by (1).The elements of A, B, C take random values within the interval $[-1, 1]$ and $D = 0$.

For controllability analysis, let the dimension of state space and input space be $n = 3$ and $m = 2$, respectively. Besides, the standard deviation of measurement noise is $\sigma_m = 0.05$. Let the number of the measured data N vary from 101 to 5000, the estimation errors of W_c with traditional methods and data-based method are shown in Fig. 1(a). Then, let n vary from 3 to 6, while setting $N = 1000$ and keeping other parameters unchanged. The change of estimation errors is shown in Fig. 1(b). Letting m vary from 2 to 5 with $N = 1000$ and other parameters unchanged, the change of estimation errors is shown in Fig. 1(c). Letting σ_m vary in $[0.01, 0.02, 0.035, 0.05, 0.07, 0.1]$ with $N = 1000$ and other parameters unchanged, the change of estimation errors is shown in Fig. 1(d).

For observability analysis, let the dimension of state space, input space and output space be $n = 4, m = 3$ and $p = 3$, respectively. Besides, the standard deviation of measurement noise is $\sigma_m = 0.05$. Letting the number of the measured data N vary from 101 to 5000, the estimation errors of W_o with traditional methods and data-based method are shown in Fig. 2(a). Then, letting n vary from 4 to 7 with $N = 1000$ and other parameters unchanged, the change of estimation errors is shown in Fig. 2(b). Letting p vary from 3 to 6 with $N = 1000$ and other parameters unchanged, the change of estimation errors is shown in Fig. 2(c). Let σ_m vary in $[0.01, 0.02, 0.035, 0.05, 0.07, 0.1]$, while setting $N = 1000$ and keeping other parameters unchanged. The change of estimation errors is shown in Fig. 2(d).

Under the condition that measured data contains measurement noise, error convergence rate of the data-based method is faster than traditional methods as shown in Fig. 1(a) and Fig. 2(a). At the same time, these two figures also confirm the consistency of the data-based method. Estimation error of traditional methods increases exponentially with the dimension of the state space, while the error of data-based method increases linearly as shown in Fig. 1(b) and Fig. 2(b). This advantage can be seen more clearly in high-dimensional space. It is shown in Fig. 1(c)–(d) that the dimension of input space m and the variance of measurement noise σ^2 influence the estimation errors of these two methods linearly. Nevertheless, the error of the data-based method is less and has a lower increasing rate than traditional methods. Similar results are shown in Fig. 2(c)–(d).

6.2. Discrete-Time Linear Systems with System Noise

Consider the following system:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + \omega_s(k), \\ y(k) = Cx(k) + Du(k) + v_s(k), \end{cases}$$

where the elements of A, B, C take random value within the interval $[-1, 1]$ and $D = 0, \omega_s(k) \sim N(0, \sigma_{s1} I_n)$ and $v_s(k) \sim N(0, \sigma_{s2} I_p)$. In this example, we set $\sigma_s = \sigma_{s1} = \sigma_{s2}$.

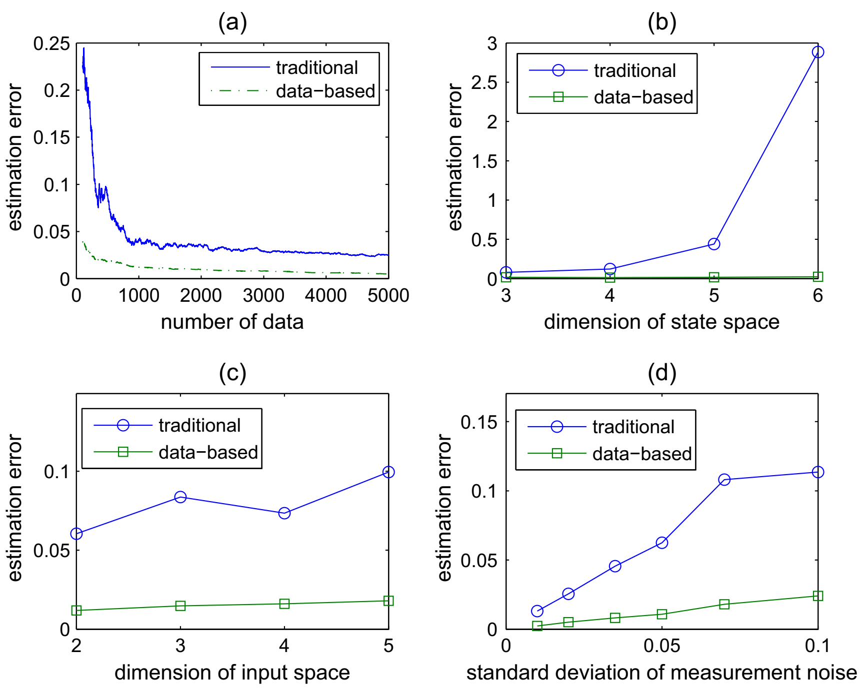

Fig. 1. Estimation errors of controllability with measurement noise. (a) The change of estimation errors with respect to N. (b) The change of estimation errors with respect to n. (c) The change of estimation errors with respect to m. (d) The change of estimation errors with respect to σ_m.

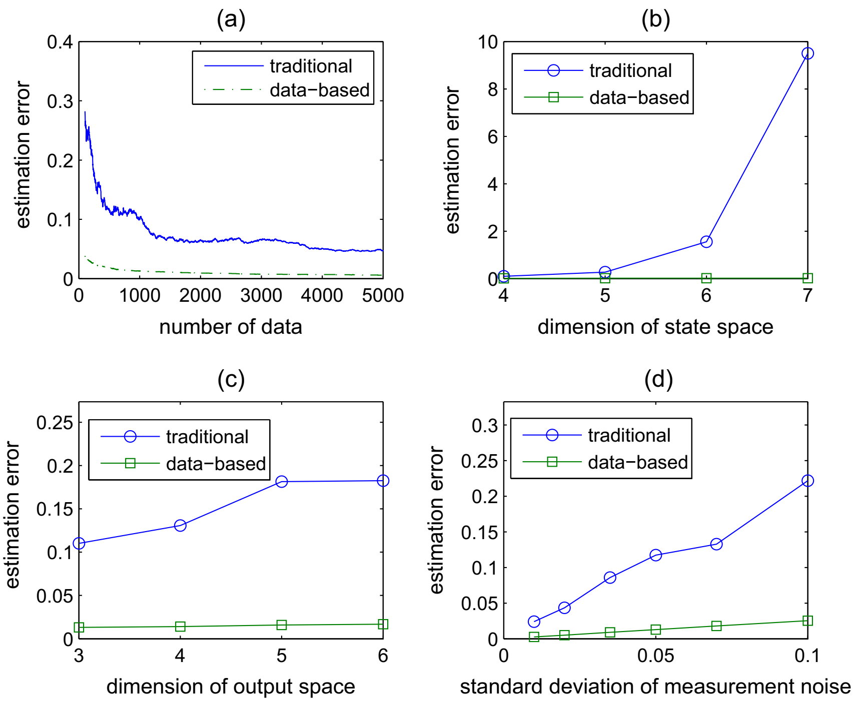

Fig. 2. Estimation errors of observability with measurement noise. (a) The change of estimation errors with respect to N. (b) The change of estimation errors with respect to n. (c) The change of estimation errors with respect to p. (d) The change of estimation errors with respect to σ_m.

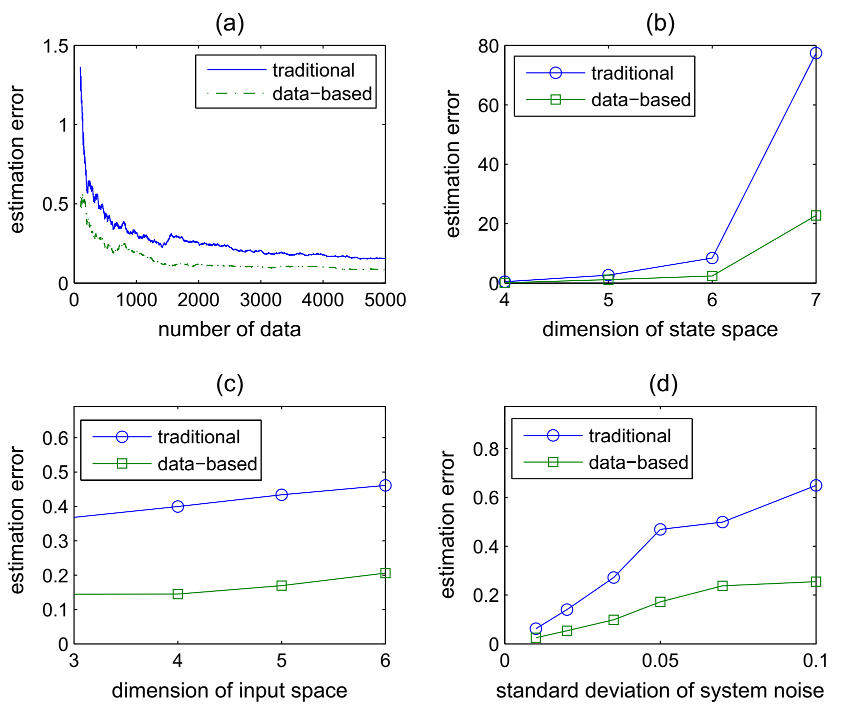

Fig. 3. Estimation errors of controllability with measurement noise. (a) The change of estimation errors with respect to N. (b) The change of estimation errors with respect to n. (c) The change of estimation errors with respect to m. (d) The change of estimation errors with respect to σ_s.

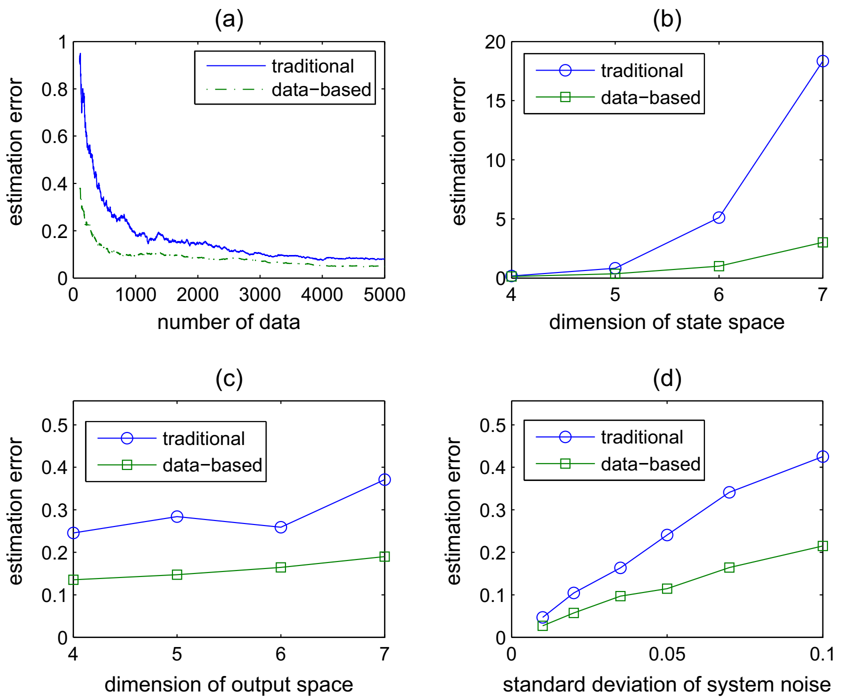

Fig. 4. Estimation errors of observability with measurement noise. (a) The change of estimation errors with respect to N. (b) The change of estimation errors with respect to n. (c) The change of estimation errors with respect to p. (d) The change of estimation errors with respect to σ_s.

For controllability analysis, let the dimension of state space and input space be $n = 4$ and $m = 3$, respectively. Besides, the standard deviation of measurement noise is $\sigma_m = 0.05$. Let the number of the measured data N vary from 101 to 5000, and we can obtain estimation error of W_c with traditional methods and data-based method shown in Fig. 3(a). Then, letting n vary from 4 to 7 with $N = 1000$ and other parameters unchanged, the change of estimation errors is shown in Fig. 3(b). Letting m vary from 3 to 6 with $N = 1000$ and other parameters unchanged, the change of estimation errors is shown in Fig. 3(c). Letting σ_s vary in $[0.01, 0.02, 0.035, 0.05, 0.07, 0.1]$ with $N = 1000$ and other parameters unchanged, the change of estimation errors is shown in Fig. 3(d).

For observability analysis, let the dimension of state space, input space and output space be $n = 4, m = 3$ and $p = 4$, respectively. Besides, the standard deviation of measurement noise is $\sigma_m = 0.05$. Let the number of the measured data N vary from 101 to 5000, and we can obtain the estimation error of W_o with traditional methods and data-based method shown in Fig. 4(a). Then, letting n vary from 4 to 7 with $N = 1000$ and other parameters unchanged, the change of estimation errors is shown in Fig. 4(b). Letting p vary from 4 to 7 with $N = 1000$ and other parameters unchanged, the change of estimation error is shown in Fig. 4(c). Letting σ_s vary in $[0.01, 0.02, 0.035, 0.05, 0.07, 0.1]$ with $N = 1000$ and other parameters unchanged, the change of estimation errors is shown in Fig. 4(d).

Under the condition that the system contains system noise, error convergence rate of the data-based method is also faster than traditional methods as shown in Fig. 3(a) and Fig. 4(a). These two figures also confirm the consistency of the data-based method. Under this condition, estimation error of these two methods increases exponentially with the dimensions of the state space n as shown in Fig. 3(b) and Fig. 4(b). Nevertheless, the data-based method has much less error than traditional methods, especially in high-dimensional state space. It is shown in Fig. 3(c)–(d) that the dimension of input m and the variance of system noise σ^2 influence the estimation errors of these two methods linearly. However, the error of the data-based method is less and has a increasing lower rate than traditional methods. Similar results are shown in Fig. 4(c)–(d).

7. Conclusions

In this paper, we developed a data-based method with general data to analyze the controllability and the observability of discrete-time linear systems in noisy environment. This method uses the measured input and output data to estimate the controllability matrix and observability matrix directly for analyzing the corresponding properties. Compared with traditional methods, we have shown that the developed method possesses has a higher precision. Furthermore, compared with the method of [26], it eases the requirements on input/output data and can analyze the controllability and the observability effectively in noisy environment. We hope that this method can be used to analyze nonlinear systems in the future and inspire new data-based methods for control systems.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grants 61034002, 61233001, 61273140, 61304086, and 61374105, in part by Beijing Natural Science Foundation under Grant 4132078, and in part by the Early Career Development Award of SKLMCCS.

References

[1] D.W. Aha, Editorial: lazy learning, Artif. Intell. Rev. 11 (1997) 7–10.
[2] P.J. Antsaklis, A.N. Michel, Linear Systems, Birkhauser, Berlin, Germany, 2006.
[3] D. Bertsekas, J. Tsitsiklis, Neuro-dynamic Programming, Athena Scientific, Belmont, MA, USA, 1996.
[4] G. Bontempi, M. Birattari, H. Bersini, Lazy learning for local modelling and control design, Int. J. Control 72 (1999) 643–658.
[5] M.C. Campi, S.M. Savaresi, Direct nonlinear control design: the virtual reference feedback tuning (VRFT) approach, IEEE Trans. Autom. Control 51 (2006) 14–27.
[6] Y. Chen, C. Wen, Iterative Learning Control: Convergence, Robustness and Applications, Springer-Verlag, London, UK, 1999.
[7] G.O. Guardabassi, S.M. Savaresi, Virtual reference direct design method: an off-line approach to data-based control system design, IEEE Trans. Autom. Control 45 (2000) 954–959.
[8] E. Hendricks, O.E. Jannerup, P.H. Sorensen, Linear Systems Control: Deterministic and Stochastic Methods, Springer-Verlag, Berlin, Germany, 2008.
[9] R. Hermann, A. Krener, Nonlinear controllability and observability, IEEE Trans. Autom. Control 22 (1977) 728–740.
[10] H. Hjalmarsson, Control of nonlinear systems using iterative feedback tuning, in: Proceedings of the American Control Conference, vol. 4, Philadephia, USA, 1998, pp. 2083–2087.
[11] H. Hjalmarsson, Iterative feedback tuningłan overview, Int. J. Adapt. Control Signal Process. 16 (2002) 373–395.
[12] H. Hjalmarsson, S. Gunnarsson, M. Gevers, A convergent iterative restricted complexity control design scheme, in: Proceedings of the 33rd IEEE Conference on Decision and Control, vol. 2, Lake Buena Vista, USA, 1994, pp. 1735–1740.
[13] T.L. Lai, H. Robbins, C.Z. Wei, Strong consistency of least squares estimates in multiple regression II, J. Multiv. Anal. 9 (1979) 343–361.
[14] J.M. Lee, J.H. Lee, Approximate dynamic programming-based approaches for input–output data-driven control of nonlinear processes, Automatica 41 (2005) 1281–1288.
[15] A.U. Levin, K.S. Narendra, Control of nonlinear dynamical systems using neural networks: controllability and stabilization, IEEE Trans. Neural Netw. 4 (1993) 192–206.
[16] A.U. Levin, K.S. Narendra, Control of nonlinear dynamical systems using neural networks, part II: Observability, identification, and control, IEEE Trans. Neural Netw. 7 (1996) 30–42.
[17] D. Li, Controllability and Observability for Quasilinear Hyperbolic Systems, American Institute of Mathematical Sciences, Springfield, MO, USA, 2010.
[18] K.L. Moore, Iterative Learning Control for Deterministic Systems, Springer-Verlag, New York, USA, 2006.
[19] J.J. Murray, C.J. Cox, G.G. Lendaris, R. Saeks, Adaptive dynamic programming, IEEE Trans. Syst. Man Cybernet. 32 (2002) 140–153.
[20] W.B. Powell, Approximate Dynamic Programming: Solving the Curses of Dimensionality, Wiley, New York, USA, 2007.
[21] M. Safonov, T.C. Tsao, The unfalsified control concept and learning, IEEE Trans. Autom. Control 42 (1997) 843–847.
[22] M.G. Safonov, T.C. Tsao, The unfalsified control concept: a direct path from experiment to controller, in: B.A. Francis, A.R. Tannenbaum (Eds.), Feedback Control, Nonlinear Systems, and Complexity, Springer-Verlag, New York, USA, 1995, pp. 196–214.
[23] S. Schaal, C.G. Atkeson, Robot juggling: implementation of memory-based learning, Control Syst. 14 (1994) 57–71.
[24] J.C. Spall, Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, IEEE Trans. Autom. Control 37 (1992) 332–341.
[25] J.C. Spall, J.A. Cristion, Model-free control of general discrete-time systems, in: Proceedings of the 32nd IEEE Conference on Decision and Control, San Antonio, USA, 1993, pp. 2792–2797.
[26] Z. Wang, D. Liu, Data-based controllability and observability analysis of linear discrete-time systems, IEEE Trans. Neural Netw. 22 (2011) 2388–2392.
[27] C.J.C.H. Watkins, Learning from delayed rewards, Ph.D. thesis, University of Cambridge, England, 1989.
[28] C.J.C.H. Watkins, P. Dayan, Q-learning, Machine Learn. 8 (1992) 279–292.
[29] P.J. Werbos, Advanced forecasting methods for global crisis warning and models of intelligence, General Syst. Yearbook 22 (1977) 25–38.
[30] P.J. Werbos, A menu of designs for reinforcement learning over time, in: W.T. Miller, R.S. Sutton, P.J. Werbos (Eds.), Neural Networks for Control, MIT Press, Cambridge, MA, 1991, pp. 67–95.