Brief Papers

Data-Based Controllability and Observability Analysis of Linear Discrete-Time Systems

Zhuo Wang and Derong Liu, *Fellow, IEEE*

***Abstract*— In this brief, we develop data-based methods for analyzing the controllability and observability of linear discrete-time systems which have unknown system parameters. These data-based methods will only use measured data to construct the controllability matrix as well as the observability matrix, in order to verify the corresponding properties. The advantages of our methods are threefold. First, they can directly verify system properties based on measured data without knowing system parameters. Second, our calculation precision is higher than traditional approaches, which need to identify the unknown parameters. Third, our methods have lower computational complexities when constructing the controllability and observability matrices.**

***Index Terms*— Controllability, data-based analysis, linear discrete-time systems, measured data, observability, unknown parameters.**

I. Introduction

In recent years, modern industrial and engineering systems have become more and more complex. Often, one of the direct consequences is the difficulty of decision making [1] or realizing the system's control and optimization with model-based approaches [2]–[7], which need mathematically building the dynamic models of the system. On the other hand, computer technologies, digital sensor technologies, and networking techniques are widely used, which generate a great quantity of historical and real-time data related to the industrial and engineering processes. In this case, technologies for data management such as data mining, data collection, and data fusion have emerged [8]–[11]. All of these have led to the development of data-based methods which have great interest from the systems and controls communities, especially in the research field of control techniques [12]–[14].

Manuscript received March 11, 2011; revised August 26, 2011; accepted September 4, 2011. Date of publication November 18, 2011; date of current version December 13, 2011. This work was supported in part by the National Natural Science Foundation of China, under Grant 60921061 and Grant 61034002, and the Beijing Natural Science Foundation, under Grant 4102061. The acting Editor-in-Chief who handled the review of this brief was F. L. Lewis.

Z. Wang is with the State Key Laboratory of Intelligent Control and Management of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. He is also with the Department of Electrical and Computer Engineering, University of Illinois, Chicago, IL 60607-7053 USA (e-mail: wangzhuo12300@gmail.com).

D. Liu is with the State Key Laboratory of Intelligent Control and Management of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: derong.liu@ia.ac.cn).

Digital Object Identifier 10.1109/TNN.2011.2170219

In control engineering, the study of system characteristics is an important research topic [15]. It is often desirable to determine an input that causes the state vector to transfer from one specified value to another. Naturally, this type of property leads to the concepts of state reachability and controllability [16], [17]. Another interesting property of systems is the ability that determines whether a state vector can be obtained from the output measurements. A method to determine such states by observing the inputs and outputs of the system over some finite interval is quite desirable. This leads to the concept of state observability [16], [17]. Up to now, there have been many works on the state controllability and observability of both linear and nonlinear systems [16]–[19].

In classical methods, to analyze the system properties, one has to do system identification first [20] if the dynamic model of the system is unknown. There are many methods often used for identifying system dynamic models, such as least square methods [21]–[24], neural network methods [25]–[32], and support vector machine methods [33]–[35]. However, these methods spend much time on system identification and cannot avoid identification errors, which affects the overall calculation precision. Specifically, arbitrarily small perturbations in a model may cause a change in the dimension of the associated subspaces. The technique of principal component analysis has been proposed for analyzing signals [36] in order to solve such problems.

In this brief, we will develop data-based methods for analyzing the controllability and observability of linear discrete-time systems, which do not need identification of the system dynamic models. They rely only on the measured state and output data, and provide higher calculation precision as well as lower computational complexity than traditional approaches.

II. System Description and Basic Assumptions

In this brief, we study linear discrete-time systems described by

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

where $x(k) \in R^n$, $u(k) \in R^m$, and $y(k) \in R^p$ are the state, input, and output of (1), respectively.

The matrices A, B, C, and D are unknown and do not have random variables as their elements. In the following sections, we will show that one can directly analyze the system properties by using measured data, without identifying the above matrices. We will also illustrate the advantages of our schemes, including high calculation precision and low computational complexity.

III. Data-Based Analysis of Controllability

In this section, we will develop a data-based method for analyzing the state controllability of (1). Let

$$W_c = \big[B, AB, \ldots, A^{n-1}B\big].$$

Then, we have the following controllability criterion.

Lemma 1 ([16, Ch. 3]): The system $x(k+1) = Ax(k) + Bu(k)$ given in (1) is completely state controllable, if

$$\text{rank}\big[W_c\big] = n.$$

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.

In traditional approaches, to verify whether (1) is state controllable, one has to identify the matrices A and B to compute W_c first, and then check its rank. However, if we can find a method that can directly construct W_c without the process of identifying A and B, it will save time in practice and avoid the corresponding identification errors. This idea can be achieved by the application of our data-based method.

Before introducing our main theorem, we do m groups of tests on the system. For all m groups of tests ($1 \le i \le m$), let all the state trajectories start from the same initial state $x^{[i]}(0) = 0 \in R^n$, while the corresponding inputs are chosen as m constant vectors

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

Here, the ith element of $u^{[i]}$ is 1, and all other elements are zeros. Note that in all our m tests, state trajectories start from the origin, and every test uses a constant control input vector according to (2). Then, measure $x^{[i]}(k)$ at time instants $k = 1, 2, \ldots, n$ and store the measured data $x^{[i]}(1), x^{[i]}(2), \ldots, x^{[i]}(n)$. Define

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

where $\bar{X}(k) \in R^{n\times m}$. Note that in the above calculation, we have $\bar{X}(0) = 0$. With these data, we now introduce the following theorem, regarding the state controllability of (1).

Theorem 1: System (1) is completely state controllable if

$$\text{rank}\big[P\big] = n$$

where P is defined in (3).

Obviously, this is a sufficient condition, since the matrix A is not identified in our method.

Proof: With the initial states $x^{[i]}(0) = 0$, when setting the control inputs as $u^{[i]}$ in (2) ($i = 1, 2, \ldots, m$), the state measurements will be

$$\begin{cases} x^{[i]}(1) = Bu^{[i]} \\ x^{[i]}(2) = ABu^{[i]} + Bu^{[i]} \\ \quad\vdots \\ x^{[i]}(k) = A^{k-1}Bu^{[i]} + \cdots + ABu^{[i]} + Bu^{[i]} \\ \quad\vdots \\ x^{[i]}(n) = A^{n-1}Bu^{[i]} + \cdots + ABu^{[i]} + Bu^{[i]} \end{cases} \tag{4}$$

where A and B are assumed to be unknown. From the definition of $u^{[i]}$, we shall have

$$\begin{cases} \big[x^{[1]}(1), x^{[2]}(1), \ldots, x^{[m]}(1)\big] = B \\ \big[x^{[1]}(2), x^{[2]}(2), \ldots, x^{[m]}(2)\big] = AB + B \\ \quad\vdots \\ \big[x^{[1]}(k), x^{[2]}(k), \ldots, x^{[m]}(k)\big] = A^{k-1}B + \cdots + B \\ \quad\vdots \\ \big[x^{[1]}(n), x^{[2]}(n), \ldots, x^{[m]}(n)\big] = A^{n-1}B + \cdots + B. \end{cases} \tag{5}$$

From the definition of $\bar{X}(k)$ and P_j given in (3), we can obtain

$$\begin{cases} P_1 = \bar{X}(1) - \bar{X}(0) = B \\ P_2 = \bar{X}(2) - \bar{X}(1) = AB \\ \quad\vdots \\ P_j = \bar{X}(j) - \bar{X}(j-1) = A^{j-1}B \\ \quad\vdots \\ P_n = \bar{X}(n) - \bar{X}(n-1) = A^{n-1}B. \end{cases} \tag{6}$$

As a result

$$P = \big[P_1, P_2, \ldots, P_n\big] = \big[B, AB, \ldots, A^{n-1}B\big] = W_c$$

which completes the proof according to Lemma 1. ■

Remark 1: In our method, to begin the tests, we set m zero initial states and m constant input vectors. Actually, the initial states do not have to be zero, but can take any values. The key point is to make all of them have the same value while still using the constant inputs defined in (2). The difference is that we need to do one more group of tests whose input is always zero. In this case, for each of the other m groups, we shall have

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

where all tests choose the same nonzero initial state $x^{[i]}(0) = x^{[0]}(0) \neq 0$ ($1 \le i \le m$). In the mean time, with $u^{[0]}(k) \equiv 0$

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

By defining

$$\bar{X}(k) = \big[x^{[1]}(k) - x^{[0]}(k), \ldots, x^{[m]}(k) - x^{[0]}(k)\big] \; (0 \le k \le n)$$

we can still obtain $B, AB, \ldots, A^{n-1}B$ the same way as in Theorem 1, with the above $(m+1)$ groups of data.

IV. Data-Based Analysis of Observability

After the above theorem on state controllability using the measured data, for the same reason of direct system analysis, we will develop another data-based method for analyzing the observability of (1) in this section. The difference is that, in this method, we will use the measured output data, but not the state data. Let

$$W_o = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

which is the observability matrix. Then, we have the following criterion.

Lemma 2 ([16, Ch. 3]): The system given in (1) is completely state observable, if and only if

$$\text{rank}\big[W_o\big] = n.$$

To verify whether (1) is state observable, the traditional approaches need to identify A and C to construct W_o, and then check its rank. Similar to the case of controllability, we want to find a method that can directly obtain W_o by using the measured data, without the process of identifying the above matrices. This idea can be achieved by the following method.

We set n nonzero initial states of the system as follows:

$$x^{[i]}(0) = [0, \ldots, 0, 1, 0, \ldots, 0]^T \in R^n \ \ (1 \le i \le n) \tag{7}$$

where the ith element of $x^{[i]}(0)$ is 1, while all other elements are zeros. Then, measure the corresponding outputs $y^{[i]}(k)$ at time instants $k = 0, 1, 2, \ldots, n-1$, while setting the inputs as $u^{[i]}(k) \equiv 0 \in R^m$. Then, store the measured output data $y^{[i]}(0), y^{[i]}(1), \ldots, y^{[i]}(n-1)$. Define

$$\bar{Y}(k) = \big[y^{[1]}(k), y^{[2]}(k), \ldots, y^{[n]}(k)\big] \ (0 \le k \le n-1)$$

$$Q = \begin{bmatrix} \bar{Y}(0) \\ \bar{Y}(1) \\ \vdots \\ \bar{Y}(n-1) \end{bmatrix} \tag{8}$$

where $\bar{Y}(k) \in R^{p \times n}$. With these data, we now introduce the following theorem, which is about the state observability.

Theorem 2: Assume that the initial states of (1) can be set as in (7). Then, the system is completely state observable, if and only if

$$\text{rank}\big[Q\big] = n$$

where Q is defined in (8).

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

$$\begin{cases} y^{[i]}(0) = Cx^{[i]}(0) \\ y^{[i]}(1) = CAx^{[i]}(0) \\ \quad\vdots \\ y^{[i]}(k) = CA^k x^{[i]}(0) \\ \quad\vdots \\ y^{[i]}(n-1) = CA^{n-1}x^{[i]}(0) \end{cases} \tag{9}$$

where A and C are unknown. From (8) and (9), we shall have

$$\begin{cases} \bar{Y}(0) = C \\ \bar{Y}(1) = CA \\ \quad\vdots \\ \bar{Y}(k) = CA^k \\ \quad\vdots \\ \bar{Y}(n-1) = CA^{n-1}. \end{cases}$$

Therefore

$$Q = \begin{bmatrix} \bar{Y}(0) \\ \bar{Y}(1) \\ \vdots \\ \bar{Y}(n-1) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = W_o$$

which completes the proof according to Lemma 2. ■

Remark 2: Theorems 1 and 2 illustrate a new perspective of analyzing the controllability as well as the observability of linear discrete-time systems, respectively. They are direct approaches. With our schemes, it is not necessary to identify the parameter matrices, which will save time and will not cause identification errors that affect the calculation precision.

Remark 3: In Section II, the system described by (1) is in the form of a deterministic system. However, in real applications, there are always various impacts of noises on the system. Here, we will discuss two kinds of noises.

1) When data are measured by sensors, the sensors often generate measurement noises. Because of this, the measured state and output data are actually

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

where $w_m(k) \in R^n$ and $v_m(k) \in R^p$ are the measurement noises. We assume that these noises have known statistical characteristics. This assumption is based on the fact that, in practical industrial processes, the statistical characteristics of the measurement noises can be determined. Then, we can obtain the unbiased estimates of the real values of $x(k)$ and $y(k)$. In this brief, our discussions are based on such an assumption where all the test data are obtained after this estimation process.

2) If the system has the form

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + w_s(k) \\ y(k) = Cx(k) + Du(k) + v_s(k) \end{cases} \tag{10}$$

where $w_s(k) \in R^n$ and $v_s(k) \in R^p$ are the system noises, we have no other choice than to identify A, B, C, and D first. Obviously, this is in contrary to the main theme of our brief, and we will not discuss this kind of noises further.

V. Analysis of the Calculation Precision

As mentioned in Remark 2, since our methods do not identify A, B, C, and D, the corresponding identification errors will not occur naturally. In this section, we will illustrate this advantage with a numerical example, which is about the state controllability of (1). Recall the state equation $x(k+1) = Ax(k) + Bu(k)$, and set

$$A = \begin{bmatrix} 0.48 & 0.51 & -0.49 \\ 0.26 & -0.32 & 0.29 \\ 0.22 & -0.23 & 0.37 \end{bmatrix}, \quad B = \begin{bmatrix} -0.35 \\ 0.16 \\ -0.21 \end{bmatrix}. \tag{11}$$

Then, we introduce the identification errors ΔA and ΔB, which occurred in the process of identifying A and B

$$\Delta A = \begin{bmatrix} -0.04 & -0.05 & 0.05 \\ 0.02 & 0.08 & -0.02 \\ -0.03 & -0.07 & -0.04 \end{bmatrix}, \quad \Delta B = \begin{bmatrix} 0.02 \\ -0.04 \\ -0.03 \end{bmatrix}.$$

In traditional methods, if we do not round off the calculation results, the controllability matrix will be

$$\begin{aligned} \hat{W}_c &= \left[(B+\Delta B), (A+\Delta A)(B+\Delta B), (A+\Delta A)^2(B+\Delta B)\right] \\ &= \begin{bmatrix} -0.33 & 0.0156 & -0.0004 \\ 0.12 & -0.1860 & 0.0010 \\ -0.24 & -0.1779 & 0.0001 \end{bmatrix}. \end{aligned} \tag{12}$$

As shown in (12), if we round off the elements of $\hat{W}_c$ to two decimal places as those in A and B, then its third column will be all zeros. As a result, $\text{rank}[\hat{W}_c] = 2 < 3$, and the system will not be considered state controllable, which is actually not true.

In our method, since we do not identify A and B, then ΔA and ΔB will not appear in the final result. Recalling (3) and (4), with zero initial states and $u(k) \equiv 1$, we shall have

$$\left[P_1, P_2, P_3\right] = \begin{bmatrix} -0.35 & 0.0165 & -0.0018 \\ 0.16 & -0.2031 & 0.0137 \\ -0.21 & -0.1915 & -0.0205 \end{bmatrix} \tag{13}$$

which has full rank. If we still correct the elements in (13) to two decimal places, the above matrix will become

$$\left[P_1, P_2, P_3\right]_{2d} = \begin{bmatrix} -0.35 & 0.02 & -0.00 \\ 0.16 & -0.20 & 0.01 \\ -0.21 & -0.19 & -0.02 \end{bmatrix}$$

which also has full rank. Therefore, we can conclude that the system defined in (11) is controllable, which is in agreement with the real case.

From the above example, we can see that our methods are robust to identification errors, and they have higher calculation precision. Therefore, the round-off errors will not affect our schemes, especially in the case when the largest nonvanishing minor of the controllability (observability) matrix has a value close to zero.

VI. Analysis of Computational Complexity

To compare the computational complexities between the traditional methods and ours, we still take the controllability matrix W_c as an example. The observability case has a similar conclusion, and therefore it will not be discussed here. In this section, to simplify the analysis, we regard both a summation (subtraction) and a multiplication of two matrix elements as one operation.

When multiplying A^k $(1 \le k \le n-2)$ with A, each element of the product is obtained by n element multiplications and $(n-1)$ element summations, while there are n^2 elements in the produced matrix. Therefore, to compute A^{k+1}, there are $n^2(2n-1)$ operations. When multiplying A^l $(1 \le l \le n-1)$ with B, each element of the product is obtained by n multiplications and $(n-1)$ summations. Since there are mn elements in $A^l B$, it is necessary to do $mn(2n-1)$ operations, and there are $(n-1)$ such matrix multiplications. In summary, to compute $W_c = [B, AB, \ldots, A^{n-1}B]$, there are totally

$$\begin{aligned} &mn(2n-1)(n-1) + \sum_{i=1}^{n-2} i\left[n^2(2n-1)\right] \\ &= \frac{1}{2}\left[2n^5 - 7n^4 + (7+4m)n^3 - (2+6m)n^2 + 2mn\right] \end{aligned}$$

operations. Then, the traditional methods have computational complexity of $O(n^5)$.

In our methods, recall (6), there are totally mn^2 element subtractions, while other operations like (5) are only arranging the measured data. So, our methods have computational complexity of $O(mn^2)$. It is clear that our data-based methods have less computational complexity, which is obviously another advantage.

VII. Conclusion

In this brief, we developed two data-based methods to analyze the controllability and the observability of linear discrete-time systems. These data-based methods use the measured state and output data to directly construct the controllability matrix as well as the observability matrix, which are applied to verify the corresponding properties. Our methods not only have low computational complexity, but also improve the calculation precision. They present a quite new idea in solving the problems in control theory. Our methods are feasible to the study of characteristics of deterministic systems, and robust to system identification errors. For systems with measurement noises, if the statistical characteristics of the noises are known, our methods are still applicable to their analysis.

However, there is still much work to do. As we can see, (1) is assumed to be time-invariant, since for time-varying systems one can determine neither the controllability nor the observability by checking only the rank of W_c or W_o. For systems having noises in the form of (10), we still have to identify the parameter matrices. We hope that our work will inspire more works in data-based approaches for control and analysis of systems, in particular for complex nonlinear systems in the future.

References

[1] H. W. Wang, C. Qi, Y. C. Wei, B. Li, and S. Zhu, "Review on data-based decision making methodologies," *Acta Autom. Sinica*, vol. 35, no. 6, pp. 820–833, Jun. 2009.

[2] P. Young and A. Chotai, "Data-based mechanistic modeling, forecasting, and control," *IEEE Control Syst.*, vol. 21, no. 5, pp. 14–27, Oct. 2001.

[3] P. C. Young, "The data-based mechanistic approach to the modelling, forecasting and control of environmental systems," *Annu. Rev. Control*, vol. 30, no. 2, pp. 169–182, Nov. 2006.

[4] Z. S. Hou and J. X. Xu, "On data-driven control theory: The state of the art and perspective," *Acta Autom. Sinica*, vol. 35, no. 6, pp. 650–667, Jun. 2009.

[5] C. Wang, T. R. Chen, and T. F. Liu, “Deterministic learning and data-based modeling and control,” *Acta Autom. Sinica*, vol. 35, no. 6, pp. 693–706, Jun. 2009.
[6] R. H. Chi and Z. S. Hou, “A model-free periodic adaptive control for freeway traffic density via ramp metering,” *Acta Autom. Sinica*, vol. 36, no. 7, pp. 1029–1033, Jul. 2010.
[7] J. Zhang, H. B. Xu, and H. B. Zhang, “Observer-based sampled-data control for a class of continuous nonlinear systems,” *Acta Autom. Sinica*, vol. 36, no. 12, pp. 1780–1787, Dec. 2010.
[8] M. Xiang and W. R. Shi, “A cluster data management algorithm based on data correlation of wireless sensor networks,” *Acta Autom. Sinica*, vol. 36, no. 9, pp. 1343–1350, May 2010.
[9] F. Wu, Y. Zhong, and Q. Y. Wu, “Mining frequent patterns over data stream under the time decaying model,” *Acta Autom. Sinica*, vol. 36, no. 5, pp. 674–684, May 2010.
[10] K. F. Wang, Z. J. Li, and S. M. Tang, “Visual traffic data collection approach based on multi-features fusion,” *Acta Autom. Sinica*, vol. 37, no. 3, pp. 322–330, Mar. 2011.
[11] X. K. Dong, Y. C. Fang, and Y. D. Zhang, “An improved AFM dynamic imaging method based on data fusion of neighboring point set,” *Acta Autom. Sinica*, vol. 37, no. 2, pp. 214–221, Feb. 2011.
[12] G. O. Guardabassi and S. M. Savaresi, “Virtual reference direct design method: An off-line approach to data-based control system design,” *IEEE Trans. Autom. Control*, vol. 45, no. 5, pp. 954–959, May 2000.
[13] S. Tan and J. F. Zhang, “Adaptive measured-data based linear quadratic optimal control of stochastic systems,” *Int. J. Control*, vol. 80, no. 10, pp. 1676–1689, 2007.
[14] J. X. Xu and Z. S. Hou, “Notes on data-driven system approaches,” *Acta Autom. Sinica*, vol. 35, no. 6, pp. 668–675, Jun. 2009.
[15] K. Ogata, *Modern Control Engineering*, 5th ed. Upper Saddle River, NJ: Prentice-Hall, 2009.
[16] P. J. Antsaklis and A. N. Michel, *Linear Systems.* Berlin, Germany: Birkhäuser, 2006.
[17] E. Hendricks, O. Jannerup, and P. H. Sørensen, *Linear Systems Control: Deterministic and Stochastic Methods*. Berlin, Germany: Springer-Verlag, 2008.
[18] S. Barnett, *Introduction to Mathematical Control Theory*. New York: Oxford Univ. Press, 1975.
[19] T. Li, *Controllability and Observability for Quasilinear Hyperbolic Systems*. Springfield, MO: AIMS, 2010.
[20] J. N. Juang, *Applied System Identification*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
[21] I. Markovsky, J. C. Willems, S. V. Huffel, B. D. Moor, and R. Pintelon, “Application of structured total least squares for system identification and model reduction,” *IEEE Trans. Autom. Control*, vol. 50, no. 10, pp. 1490–1500, Oct. 2005.
[22] M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[23] F. Ding, P. X. Liu, and G. Liu, “Multiinnovation least-squares identification for system modeling,” *IEEE Trans. Syst., Man, Cybern., Part B: Cybern.*, vol. 40, no. 3, pp. 767–778, Jun. 2010.
[24] X. Liu and J. Lu, “Least squares based iterative identification for a class of multirate systems,” *Automatica*, vol. 46, no. 3, pp. 549–554, Mar. 2010.
[25] O. Nelles, *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. New York: Springer-Verlag, 2001.
[26] W. Yu and X. Li, “Some new results on system identification with dynamic neural networks,” *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 412–417, Mar. 2001.
[27] X. M. Ren, A. B. Rad, P. T. Chan, and L. L. Wai, “Identification and control of continuous-time nonlinear systems via dynamic neural networks,” *IEEE Trans. Ind. Electron.*, vol. 50, no. 3, pp. 478–486, Jun. 2003.
[28] S. Yilmaz and Y. Oysal, “Fuzzy wavelet neural network models for prediction and identification of dynamical systems,” *IEEE Trans. Neural Netw.*, vol. 21, no. 10, pp. 1599–1609, Oct. 2010.
[29] H. L. Wei, S. A. Billings, Y. F. Zhao, and L. Z. Guo, “An adaptive wavelet neural network for spatio-temporal system identification,” *Neural Netw.*, vol. 23, no. 10, pp. 1286–1299, Dec. 2010.
[30] D. Z. Cheng, H. S. Qi, and Z. Q. Li, “Model construction of Boolean network via observed data,” *IEEE Trans. Neural Netw.*, vol. 22, no. 4, pp. 525–536, Apr. 2011.
[31] A. Alessandri, M. Baglietto, G. Battistelli, and M. Gaggero, “Moving-horizon state estimation for nonlinear systems using neural networks,” *IEEE Trans. Neural Netw.*, vol. 22, no. 5, pp. 768–780, May 2011.
[32] B. Subudhi and D. Jena, “Nonlinear system identification using memetic differential evolution trained neural networks,” *Neurocomputing*, vol. 74, no. 10, pp. 1696–1709, May 2011.
[33] L. Wang, *Support Vector Machines: Theory and Applications*. Berlin, Germany: Springer-Verlag, 2005.
[34] I. Steinwart and A. Christmann, *Support Vector Machines*. New York: Springer-Verlag, 2008.
[35] L. Zhang, Y. G. Xi, and W. D. Zhou, “Identification and control of discrete-time nonlinear systems using affine support vector machines,” *Int. J. Artif. Intell. Tools*, vol. 18, no. 6, pp. 929–947, 2009.
[36] B. Moore, “Principal component analysis in linear systems: Controllability, observability, and model reduction,” *IEEE Trans. Autom. Control*, vol. 26, no. 1, pp. 17–32, Feb. 1981.

Approximate Dynamic Programming for Optimal Stationary Control with Control-Dependent Noise

Yu Jiang, *Student Member*, *IEEE*, and
Zhong-Ping Jiang, *Fellow*, *IEEE*

***Abstract*— This brief studies the stochastic optimal control problem via reinforcement learning and approximate/adaptive dynamic programming (ADP). A policy iteration algorithm is derived in the presence of both additive and multiplicative noise using Itô calculus. The expectation of the approximated cost matrix is guaranteed to converge to the solution of some algebraic Riccati equation that gives rise to the optimal cost value. Moreover, the covariance of the approximated cost matrix can be reduced by increasing the length of time interval between two consecutive iterations. Finally, a numerical example is given to illustrate the efficiency of the proposed ADP methodology.**

***Index Terms*— Approximate dynamic programming, control-dependent noise, optimal stationary control, stochastic systems.**

I. Introduction

Reinforcement learning [1] is one of the most important branches in learning theory. Roughly speaking, it is concerned with how an agent improves decisions at each step to achieve some long-term goal, based on interactions with and rewards received from the environment. One technique to solve reinforcement learning problems is called adaptive/approximate dynamic programming (ADP), pioneered by Werbos [2]–[5]. The strategy consists of estimating the cost function online, and making further decisions through successive iterations.

Reinforcement learning and ADP are extensively studied for Markov decision processes for both the discrete and continuous state and action spaces, see [6]–[9], for example.

Manuscript received January 13, 2011; revised June 2, 2011; accepted August 14, 2011. Date of publication September 26, 2011; date of current version December 13, 2011. This work was supported in part by the U.S. National Science Foundation, under Grant DMS-0906659 and Grant ECCS-1101401.

Y. Jiang is with the Department of Electrical and Computer Engineering, Polytechnic Institute of New York University, Brooklyn, NY 11201 USA (e-mail: yu.jiang@nyu.edu).

Z.-P. Jiang is with the Department of Electrical and Computer Engineering, Polytechnic Institute of New York University, Brooklyn, NY 11201 USA. He is also with the College of Engineering, Beijing University, Beijing 100214, China (e-mail: zjiang@poly.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNN.2011.2165729