Automatica 92 (2018) 143–154

Contents lists available at ScienceDirect

Automatica

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

Reduction of state variables based on regulation and filtering performances[☆]

Hidenori Shingin[a,*], Yoshito Ohta[b]

[a] Department of Mechanical Engineering, Yamaguchi University, 2-16-1 Tokiwadai, Ube, Yamaguchi 755-8611, Japan
[b] Department of Applied Mathematics and Physics, Kyoto University, 36-1 Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan

ARTICLE INFO

Article history:
Received 14 January 2015
Received in revised form 31 January 2017
Accepted 25 December 2017

Keywords:
Principal component analysis
Model reduction
Optimal regulator
Kalman filter

ABSTRACT

This paper provides a principal component analysis of linear discrete-time systems on the basis of optimal control and estimation. The analysis is to reveal the important state components which remain necessary for reducing performance degradation under dimensional constraints on control and estimation laws. The trade-off relations between the dimension and performance degradation are expressed as system invariants representing the importance of each principal component, which are characterized as the eigenvalues of matrices depending on the solutions of both Lyapunov and Riccati equations. Based on the analysis, the paper also provides model reduction techniques for the systems generating the optimal input and estimate with the desirable properties of stability, reachability, and observability being preserved in the reduced systems.

1. Introduction

As a way of obtaining low-order controllers, reduction of state variables is commonly applied to the plant models or designed controllers. The necessity of low-order controllers arises from lack of computing power available for control. This problem occurs in the situations that the plant is a large-scale system requiring high computing power and that the controller is implemented in embedded processor with low computing power.

The reduction provided in this paper is based on the established balanced truncation (Moore, 1981) summarized as follows. The state elements of the balanced realization are arranged in orders of strengths in terms of the connections to the input and output (Roberts & Mullis, 1987). The order of the system is reduced by eliminating the trailing elements which do not have major contribution to the input-to-output relation. The theoretical basis of this method is a principal component analysis using reachability and observability Gramians. Specifically, the importance of each state element is represented as the square root of the eigenvalue of the product of the Gramians. This index is exactly the singular value of the Hankel operator mapping the past input to the future output, so it is invariant under similarity transformations. A notable advantage of this method is that stability, reachability, and observability are preserved in the reduced systems (Pernebo & Silverman, 1982).

Several other methods of balancing for reduction have been developed along this line. They are based on component analyses using the solutions of some Lyapunov and Riccati equations as alternatives to the reachability and observability Gramians. The positive real balancing (Desai & Pal, 1984) and bounded real balancing (Opdenacker & Jonckheere, 1988) give the reduction methods preserving positive realness and bounded realness, respectively. The LQG balancing (Jonckheere & Silverman, 1983) and H^∞ balancing (Mustafa & Glover, 1991) reveal the elements which contribute strongly to the closed-loop properties. The frequency weighted balancing (Enns, 1984) enables to evaluate the reduction error in frequency domain. The balancing technique is also applicable to the reduction of unstable systems via fractional representation (Meyer, 1990). The details of these component analyses can be found in Antoulas (2005) and Gugercinb and Antoulas (2004), and other techniques such as norm approximation and moment matching also can be seen in Antoulas (2005).

Our principal component analysis reveals the state components strongly contributing to control and estimation performances relative to the cases without these efforts. The contributions are measured by newly introduced indices. Such components are expressed using the solutions of both Lyapunov and Riccati equations, and their contributions are system invariants given as the eigenvalues of matrices depending on these solutions. As in balanced truncation, we propose truncation methods for the systems generating optimal inputs and estimates to remove the elements

[☆] The material in this paper was presented at the 19th International Symposium on Mathematical Theory of Networks and Systems, July 5–9, 2010, Budapest, Hungary. This paper was recommended for publication in revised form by Associate Editor Mario Sznaier under the direction of Editor Richard Middleto.

* Corresponding author.
E-mail addresses: shingin@yamaguchi-u.ac.jp (H. Shingin), yoshito_ohta@i.kyoto-u.ac.jp (Y. Ohta).

https://doi.org/10.1016/j.automatica.2018.03.002
0005-1098/

with minor contributions and show that the desirable properties of stability, reachability, and observability are preserved. Moreover, motivated by the importance of the Hankel operator relating past and future signals, we propose a truncation method based on a component analysis that reveals elements important for determining the future input from the past output.

The first novelty of our analysis is that the dominant components are determined based on optimization indices for evaluating contributions to estimation and control. The analysis (Moore, 1981) justifying the model reduction through balancing ensures the optimality of the component extraction by introducing quantitative indices, but does not investigate influences on estimation and control. The other subsequent works focus mainly on the extensions of balancing for model reduction and thus do not fully justify component extraction. For example, the controller reduction in Jonckheere and Silverman (1983) is based on balancing for optimally estimated and controlled systems, but does not have an optimality index for component extraction. Our objective is to quantify the influences on estimation and control due to component extraction by evaluating the resulting performances relatively to the cases without estimation and control actions.

The second novelty of our analysis is that the dominant components are found for the situation intermediate between the two extremes where control and estimation are not conducted and completely conducted for every component. The dominant components revealed by Jonckheere and Silverman (1983) and Moore (1981) are for these two extreme situations, so they do not always strongly contribute to the performances in the intermediate situation. Our analysis investigates how much each component contributes to the performance relative to the cases of complete control and estimation.

The rest of this paper is organized as follows. Section 2 provides a brief review of Lyapunov and Riccati equations playing an important role in our analyses. Using the solutions of these equations, Sections 3 and 4 respectively conduct state component analyses for optimal control and estimation, and Section 5 conducts an analysis for estimation-based control. Based on the results of these analyses, Section 6 provides methods of model reduction and shows their desirable properties. Section 7 substantiates the results in the preceding sections by numerical examples. Finally, Section 8 presents concluding remarks.

The notations in this paper are as follows. The eigenvalues of a positive definite matrix $P \in \mathbb{R}^{n\times n}$ are denoted by $\lambda_1(P) \geq \cdots \geq \lambda_n(P)$, and the norm of a vector $x \in \mathbb{R}^n$ with respect to P is defined as $\|x\|_P := \sqrt{x^{\mathrm{T}}Px}$, which is simply denoted by $\|x\|$ when $P = I$. The Frobenius norm of a matrix $A \in \mathbb{R}^{m\times n}$ is defined as $\|A\| := \sqrt{\mathrm{tr}A^{\mathrm{T}}A}$. The image of a function f is denoted by $\mathrm{Im}\, f$, and the dimension of a linear space $\mathcal{X}$ is denoted by $\dim \mathcal{X}$.

2. Lyapunov and Riccati equations

In this section, we present a brief review of the Lyapunov and Riccati equations, which will be used in the subsequent analyses. We consider the two Lyapunov equations

$$X = A^{\mathrm{T}}XA + Q, \tag{1}$$

$$\hat{X} = A\hat{X}A^{\mathrm{T}} + W \tag{2}$$

and the Riccati equations

$$Y = A^{\mathrm{T}}YA + Q - A^{\mathrm{T}}YB(B^{\mathrm{T}}YB + R)^{-1}B^{\mathrm{T}}YA, \tag{3}$$

$$\hat{Y} = A\hat{Y}A^{\mathrm{T}} + W - A\hat{Y}C^{\mathrm{T}}(C\hat{Y}C^{\mathrm{T}} + V)^{-1}C\hat{Y}A^{\mathrm{T}} \tag{4}$$

with variables $X, \hat{X}, Y, \hat{Y} \in \mathbb{R}^{n\times n}$ and coefficients $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{\ell\times n}$, $Q \in \mathbb{R}^{n\times n}$, $W \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{\ell\times \ell}$. Here Q and W are positive semidefinite, and R and V are positive definite. For these equations, we will make auxiliary use of the corresponding Lyapunov difference equations

$$X_{k-1} = A^{\mathrm{T}}X_kA + Q, \tag{5}$$

$$\hat{X}_{k+1} = A\hat{X}_kA^{\mathrm{T}} + W \tag{6}$$

and Riccati difference equations

$$\begin{aligned} Y_{k-1} &= A^{\mathrm{T}}Y_kA + Q \\ &\quad - A^{\mathrm{T}}Y_kB(B^{\mathrm{T}}Y_kB + R)^{-1}B^{\mathrm{T}}Y_kA, \end{aligned} \tag{7}$$

$$\begin{aligned} \hat{Y}_{k+1} &= A\hat{Y}_kA^{\mathrm{T}} + W \\ &\quad - A\hat{Y}_kC^{\mathrm{T}}(C\hat{Y}_kC^{\mathrm{T}} + V)^{-1}C\hat{Y}_kA^{\mathrm{T}}. \end{aligned} \tag{8}$$

Throughout this paper, we assume that A is asymptotically stable, that is, all the eigenvalues of A have modulus less than one. In this case, (1) has a unique positive definite solution X, and $X_k \to X$ as $k \to -\infty$ for any $X_0 \geq 0$. In the same way, (2) has a unique positive definite solution $\hat{X}$, and $\hat{X}_k \to \hat{X}$ as $k \to \infty$ for any $\hat{X}_0 \geq 0$. Furthermore, if $Q = C^{\mathrm{T}}C$ and (C, A) is observable, (3) has a unique positive definite solution Y, and $Y_k \to Y$ as $k \to -\infty$ for any $Y_0 \geq 0$. Similarly, if $W = BB^{\mathrm{T}}$ and (A, B) is reachable, (4) has a unique positive definite solution $\hat{Y}$, and $\hat{Y}_k \to \hat{Y}$ as $k \to \infty$ for any $\hat{Y}_0 \geq 0$. These positive definite solutions are stabilizing solutions, that is, the corresponding matrices $F = A + BK$ for $K = -\left(B^{\mathrm{T}}YB + R\right)^{-1}B^{\mathrm{T}}YA$ and $G = A + LC$ for $L = -A\hat{Y}C^{\mathrm{T}}\left(C\hat{Y}C^{\mathrm{T}} + V\right)^{-1}$ are asymptotically stable. We shall also use the finite-time versions of these matrices and their convergence properties given by $F_k = A + BK_k \to F$ for $K_k = -\left(B^{\mathrm{T}}Y_kB + R\right)^{-1}B^{\mathrm{T}}Y_kA \to K$ and $G_k = A + L_kC \to G$ for $L_k = -A\hat{Y}_kC^{\mathrm{T}}\left(C\hat{Y}_kC^{\mathrm{T}} + V\right)^{-1} \to L$. The similarity transformation by a nonsingular matrix $T \in \mathbb{R}^{n\times n}$ defined as $(A, B, C) \mapsto (TAT^{-1}, TB, CT^{-1})$ induces the change of these solutions and the related matrices; for example X and Y are changed to $T^{-\mathrm{T}}XT^{-1}$ and $T^{-\mathrm{T}}YT^{-1}$, respectively, if $Q = C^TC$, and $\hat{X}$ and $\hat{Y}$ are changed to $T\hat{X}T^{\mathrm{T}}$ and $T\hat{Y}T^{\mathrm{T}}$, respectively, if $W = BB^{\mathrm{T}}$.

3. Component analysis for optimal control

In this section, we provide a principal component analysis for optimal control to reveal the state components which are important for performance improvement. The principal components and their importance are represented by the solutions of both Lyapunov and Riccati equations.

3.1. The setup and the result

We consider the linear discrete-time system

$$x_{k+1} = Ax_k + Bu_k, \tag{9}$$

$$y_k = Cx_k, \tag{10}$$

where $x_k \in \mathbb{R}^n$ is the state, $u_k \in \mathbb{R}^m$ is the input, and $y_k \in \mathbb{R}^\ell$ is the output. Here it is assumed that A is asymptotically stable, and (C, A) is observable. The control law for this system is the function $f : \mathbb{R}^n \to \ell^2$ which determines the control input as

$$u = f(x_0).$$

The optimal regulation problem is to find the control law which minimizes the performance function

$$J(x_0, f) := \sum_{k=0}^{\infty} \left(x_k^{\mathrm{T}}Qx_k + u_k^{\mathrm{T}}Ru_k\right),$$

where $Q = C^{\mathrm{T}}C$. As is well known, the optimal performance is

$$J^*(x_0) = \min_f J(x_0, f) = x_0^{\mathrm{T}}Yx_0,$$