Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems – MTNS 2010 • 5–9 July, 2010 • Budapest, Hungary

Reduction of State Variables Based on Regulation and Filtering Performances

Hidenori Shingin and Yoshito Ohta

Abstract— **This paper provides a component analysis for the state variables of stable linear discrete-time systems based on the control and estimation performance criteria. In the optimal regulation and filtering problems, the trade-offs between the dimension of the control or estimation law and degree of performance degenerations are invariants given as the eigenvalues of the matrices depending on the solutions of both Lyapunov and Riccati equations. This analysis shows the dominant components of the state variables which have major contribution to enhance the performances.**

I. INTRODUCTION

The limitation of data processing capability is a major factor affecting the performance of large scale control systems. The model reduction for the plant and controller is an effective way to overcome this difficulty by decreasing the load of data processing [1]. The method of reduction using the state space representation is based on the component analysis for the state variables [2], [3].

The balance truncation [2] is a method of model reduction for stable systems based on the component analysis using controllability and observability Grammians, which are the solutions of Lyapunov equations. The degrees of contribution of the principal components of the state variable to the input-to-output relation are invariants and given as the eigenvalues of the product of these Grammians. They accord with the nonzero singular values of the Hankel operator, which are the map from the past input to the future output. This analysis enables us to reduce the order of the system while preserving the system property by removing the components which have only minor contribution to the input-to-output relation. Furthermore, some other methods are proposed based on the realizations such as stochastic balance realization [4] and frequency weighted realization [5].

On the other hand, the reduction based on the component analysis for the closed loop property is provided by [6]. The degrees of contribution of the principal components of the state variable to the closed loop property of the LQG control system are invariants and given as the eigenvalues of the product of the solutions of Riccati equations in optimal regulation and filtering problems.

In this paper, we provide a component analysis for the state variables of stable linear discrete-time systems using the both solutions of Lyapunov and Riccati equations based on the control and estimation performance criteria. In optimal control and estimation problems, we show that the trade-offs between the dimension of the control or estimation law and degree of performance degeneration are given by the eigenvalues of the matrices depending on the solutions of both Lyapunov and Riccati equations. This analysis show the dominant components of the state variable which have major contribution to enhance the performances.

The rest of this paper is organized as follows. In Section II, we explain the Lyapunov and Riccati equations used in our analysis. In Sections III and IV, we provide the component analyses of the state variables based on the control and estimation performance criteria using the solutions of these equations. In Section V, we provide an analysis based on both these two criteria. In Section VI, we present concluding remarks.

The notations in this paper are as follows. For a positive definite matrix $P \in \mathbb{R}^{n\times n}$, the eigenvalues of P are ordered as $\lambda_1(P) \geq \cdots \geq \lambda_n(P)$, and the norm of a vector $x \in \mathbb{R}^n$ for P is defined as $\|x\|_P := \sqrt{x^{\mathrm{T}} P x}$ and 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 map f is denoted by $\mathrm{Im}\, f$. The dimension of a linear space $\mathcal{X}$ is denoted by $\dim \mathcal{X}$.

II. LYAPUNOV AND RICCATI EQUATIONS

In this section, we present a brief review of the Lyapunov and Riccati equations used in our analysis. We consider the two kinds of Lyapunov equations

$$X = A^{\mathrm{T}} X A + Q, \tag{1}$$
$$X = A X A^{\mathrm{T}} + W \tag{2}$$

and Riccati equations

$$Y = A^{\mathrm{T}} Y A + Q - A^{\mathrm{T}} Y B (B^{\mathrm{T}} Y B + R)^{-1} B^{\mathrm{T}} Y A, \tag{3}$$
$$Y = A Y A^{\mathrm{T}} + W - A Y C^{\mathrm{T}} (C Y C^{\mathrm{T}} + V)^{-1} C Y A^{\mathrm{T}} \tag{4}$$

with the coefficient matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{\ell\times n}$ and positive definite constant matrices $Q \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times m}$, $W \in \mathbb{R}^{n\times n}$, $V \in \mathbb{R}^{\ell\times \ell}$. To derive our results, we also consider the two kinds of Lyapunov difference equations

$$X_{t+1} = A^{\mathrm{T}} X_t A + Q, \tag{5}$$
$$X_{t+1} = A X_t A^{\mathrm{T}} + W \tag{6}$$

H. Shingin is with Department of Mechanical Engineering, Yamaguchi University, 2-16-1 Tokiwadai, Ube, Yamaguchi 755-8611, Japan shingin@yamaguchi-u.ac.jp

Y. Ohta is with Department of Applied Mathematics and Physics, Kyoto University, 36-1 Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan yoshito_ohta@i.kyoto-u.ac.jp

ISBN 978-963-311-370-7

with the initial value $X_0 \geq 0$ and Riccati difference equations

$$Y_{t+1} = A^{\mathrm{T}} Y_t A + Q - A^{\mathrm{T}} Y_t B (B^{\mathrm{T}} Y_t B + R)^{-1} B^{\mathrm{T}} Y_t A, \tag{7}$$

$$Y_{t+1} = A Y_t A^{\mathrm{T}} + W - A Y_t C^{\mathrm{T}} (C Y_t C^{\mathrm{T}} + V)^{-1} C Y_t A^{\mathrm{T}} \tag{8}$$

with the initial value $Y_0 \geq 0$. If A is stable, (1) has a unique positive definite solution X_c, and $X_t \to X_c$ as $t \to \infty$ in (5). Thus if A is stable, (2) has a unique positive definite solution X_o, and $X_t \to X_o$ as $t \to \infty$ in (6). If (A, B) is controllable, (14) has a unique positive definite solution Y_c, and $X_t \to Y_c$ as $t \to \infty$ in (7). Thus if (C, A) is observable, (4) has a unique positive definite solution Y_o, and $Y_t \to Y_o$ as $t \to \infty$ in (8).

III. COMPONENT ANALYSIS BASED ON THE CONTROL PERFORMANCE CRITERION

In this section, we provide a component analysis of the state variables based on the control performance criterion using the solutions of both Lyapunov and Riccati equations. The result shows the dominant components of the state which have major contribution to enhance the control performance.

A. The Setup and the Result

We consider the system

$$x_{t+1} = A x_t + B u_t, \ t \geq 0,$$

where x is the state, and u is the control input. In the optimal regulation problem, the control input is determined as

$$u = f(x_0)$$

by the control law $f : \mathbb{R}^n \to \ell^2$, and the control performance is evaluated by the criterion

$$J(x_0, f) := \sum_{t=0}^{\infty} \left(x_t^{\mathrm{T}} Q \, x_t + u_t^{\mathrm{T}} R \, u_t \right).$$

Then the optimal performance is given as

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

Here we define the degree of performance degeneration due to the dimensional constraint on the control law as

$$\mu_r^* := \min_{\dim \operatorname{Im} f \leq r} \mu(f),$$

where

$$\mu(f) := \max_{x_0 \neq 0} \frac{J(x_0, f)}{J^*(x_0)}.$$

Note that

$$\mu_0^* \geq \mu_1^* \geq \cdots \geq \mu_n^* = 1.$$

Under the above setup, the result of the component analysis is given as the next theorem.

Theorem 1: Suppose that A is stable and (A, B) is controllable. Then

$$\mu_r^* = \lambda_{r+1}(X_c Y_c^{-1}).$$

Furthermore, an optimal control law, that is, a function f_r^* which achieves $\mu(f_r^*) = \mu_r^*$ and satisfies $\dim \operatorname{Im} f_r^* \leq r$ is given by

$$f_r^*(\xi_0) = f_n^*(\hat{\xi}_0),$$
$$T\hat{\xi}_0 = \operatorname{diag}\{I_r, 0_{n-r}\} T \xi_0, \ T = V Y_c^{1/2},$$

where V is the orthogonal matrix such that

$$V Y_c^{-1/2} X_c Y_c^{-1/2} V^{\mathrm{T}} = \operatorname{diag}\left\{ \lambda_1(X_c Y_c^{-1}), \ldots, \lambda_n(X_c Y_c^{-1}) \right\}.$$

We can see that μ_r^* is invariant under linear transformation of the state variable. The dominant components of the state are the first to the r-th elements of the transformed variable $\hat{x}_t = T x_t$. The optimal input under the dimensional constraint can be generated by applying the optimal law under no constrain to the approximate initial state obtained by reducing the non-dominant components.

B. Proof of the Theorem

To prove the above theorem, we consider the finite time optimal regulation problem. In the control in time k, the control input is determined as

$$\tilde{u}_k = f_k(x_0), \ \tilde{u}_k := (u_0, \ldots, u_k)$$

by the control law $f_k : \mathbb{R}^n \to \mathbb{R}^{(k+1)m}$, and the control performance is evaluated by the criterion

$$J_k(x_0, f_k) := \sum_{t=0}^{k} \left(x_t^{\mathrm{T}} Q \, x_t + u_t^{\mathrm{T}} R \, u_t \right). \tag{9}$$

Then the optimal performance is given as

$$J_k^*(x_0) := \min_{f_k} J(x_0, f_k) = x_0^{\mathrm{T}} Y_k x_0$$

for the solution of (7) for $Y_0 = Q$. Here define the degree of performance degeneration due to the dimensional constraint as

$$\mu_{kr}^* := \min_{\dim \operatorname{Im} f_k \leq r} \mu_k(f_k),$$

where

$$\mu_k(f_k) := \max_{x_0 \neq 0} \frac{J_k(x_0, f_k)}{J_k^*(x_0)}.$$

In the following lemmas, let $\tilde{J}_k$ be the function of x_0 and $\tilde{u}_k$ given as the right side of (9) and define

$$\Delta \tilde{J}_k(x_0, \tilde{u}_k) := \tilde{J}_k(x_0, \tilde{u}_k) - J_k^*(x_0).$$

Note that

$$J_k(x_0, f_k) = \tilde{J}_k(x_0, f_k(x_0)), \ J_k^*(x_0) = \min_{\tilde{u}_k} \tilde{J}_k(x_0, \tilde{u}_k).$$

The following result is a counterpart of Lemma 7 which is known in the Kalman filtering theory.

Lemma 1: The function $\Delta \tilde{J}_k$ can be represented as

$$\Delta \tilde{J}_k(x_0, \tilde{u}_k) = \| F_k x_0 + \tilde{u}_k \|_{S_k}^2 \tag{10}$$

by some matrix $F_k \in \mathbb{R}^{(k+1)m \times n}$ and positive definite matrix $S_k \in \mathbb{R}^{(k+1)m \times (k+1)m}$ satisfying

$$F_k^{\mathrm{T}} S_k F_k = Z_k, \tag{11}$$

where $Z_k = X_k - Y_k$ and $X_0 = Y_0 = Q$ for the solutions of (5) and (7).

Proof: Since $\tilde{J}_k(x_0, \tilde{u}_k)$ is quadratic with respect to $\tilde{u}_k$, it can be expressed as (9) by some F_k and S_k. Thus $\Delta\tilde{J}_k(x_0, 0) = x_0^{\mathrm{T}} F_k^{\mathrm{T}} S_k F_k x_0$. On the other hand, $J_k^*(x_0) = x_0^{\mathrm{T}} X_k x_0$ and $\tilde{J}_k(x_0, 0) = x_0^{\mathrm{T}} Y_k x_0$. Therefore $x_0^{\mathrm{T}} \left(F_k^{\mathrm{T}} S_k F_k - Z_k\right) x_0 = 0$, which implies (11) since $x_0 \in \mathbb{R}^n$ is arbitrary. ■

The following fact is easily obtained by using the singular value decomposition and orthogonal projection.

Lemma 2: For a matrix $F \in \mathbb{R}^{p\times n}$, positive definite matrices $S \in \mathbb{R}^{p\times p}$, $Y \in \mathbb{R}^{n\times n}$, and class of a map $f : \mathbb{R}^n \to \mathbb{R}^p$ with $p \geq n$,

$$\min_{\dim \operatorname{Im} f \leq r} \max_{x\neq 0} \frac{\|Fx - f(x)\|_S^2}{\|x\|_Y^2} = \lambda_{r+1}(ZY^{-1}),$$

where $Z = F^{\mathrm{T}}SF$. The minimum can be achieved by the map f_r^* such that

$$f_r^*(x) = S^{-1/2} U \Sigma_r V^{\mathrm{T}} Y^{1/2} x,$$

where $\Sigma_r \in \mathbb{R}^{p\times n}$ is given as

$$\Sigma_r = \begin{bmatrix} \hat{\Sigma}_r \\ 0_{(p-n)\times n} \end{bmatrix},$$

$$\hat{\Sigma}_r = \operatorname{diag}\{\lambda_1(ZY^{-1}), \ldots, \lambda_r(ZY^{-1}), 0 \ldots, 0\},$$

and $U, V \in \mathbb{R}^{n\times n}$ are the orthogonal matrices such that

$$S^{1/2} F X^{-1/2} = U \Sigma_n V^{\mathrm{T}}.$$

Furthermore, f_r^* satisfies

$$f_r^*(x) = f_n^*(\hat{x}),$$
$$T\hat{x} = \operatorname{diag}\{I_r, 0_{n-r}\} Tx, \; T = VY^{1/2}.$$

Lemma 3: Suppose that A is stable and (A, B) is controllable. Then

$$\Delta\tilde{J}(x_0, f_n^*(\hat{x}_0)) = \|x_0 - \hat{x}_0\|_{Z_c}^2,$$

where $Z_c = X_c - Y_c$.

Proof: It follows from Lemma 1 that J_k is minimized by the control law

$$f_{kn}^*(x_0) = -F_k x_0,$$

and

$$\begin{aligned}\Delta\tilde{J}_k(x_0, f_{kn}^*(\hat{x}_0)) &= \|F_k(x_0 - \hat{x}_0)\|_{S_k}^2 \\ &= \|x_0 - \hat{x}_0\|_{Z_k}^2.\end{aligned}$$

Define $\phi_k(\tilde{u}_k) := (\tilde{u}_k, 0, 0, \ldots)$. Then

$$\begin{aligned}\tilde{J}(x_0, f_n^*(\hat{x}_0)) &= \lim_{k\to\infty} \tilde{J}(x_0, \phi_k \circ f_{kn}^*(\hat{x}_0)) \\ &= \lim_{k\to\infty} \tilde{J}_k(x_0, f_{kn}^*(\hat{x}_0))\end{aligned}$$

and

$$J^*(x_0) = \lim_{k\to\infty} J_k^*(x_0),$$

thus

$$\Delta\tilde{J}(x_0, f_n^*(\hat{x}_0)) = \lim_{k\to\infty} \Delta\tilde{J}_k(x_0, f_{kn}^*(\hat{x}_0)).$$

Therefore

$$\begin{aligned}\Delta\tilde{J}(x_0, f_n^*(\hat{x}_0)) &= \lim_{k\to\infty} \|x_0 - \hat{x}_0\|_{Z_k}^2 \\ &= \|x_0 - \hat{x}_0\|_{Z_c}^2.\end{aligned}$$

■

Lemma 4: For the solutions of (5) and (7) for $X_0 = Y_0 = Q$,

$$\mu_{kr}^* = \lambda_{r+1}(X_k Y_k^{-1}).$$

Proof: From Lemma 1,

$$\begin{aligned}\frac{\Delta J_k(x_0, f_k)}{J_k^*(x_0)} &= \frac{\Delta\tilde{J}_k(x_0, f_k(x_0))}{J_k^*(x_0)} \\ &= \frac{\|F_k x_0 + f_k(x_0)\|_{S_k}^2}{\|x_0\|_{X_k}^2}.\end{aligned}$$

By applying Lemma 2, we have

$$\begin{aligned}\mu_{k,r}^* &= \min_{\dim \operatorname{Im} f_k \leq r} \max_{x_0\neq 0} \frac{\Delta J_k(x_0, f_k)}{J_k^*(x_0)} + 1 \\ &= \lambda_{r+1}((X_k - Y_k)\, Y_k^{-1}) + 1 \\ &= \lambda_{r+1}(X_k Y_k^{-1}).\end{aligned}$$

■

Proof of Theorem 1: Define $\psi_k(u) = \tilde{u}_k$. Then

$$\begin{aligned}\mu_r^* &\geq \min_{\dim \operatorname{Im} f_k \leq r} \max_{x_0\neq 0} \frac{J_k(x_0, f_k)}{J^*(x_0)} \\ &= \min_{\dim \operatorname{Im} f \leq r} \max_{x_0\neq 0} \frac{\tilde{J}_k(x_0, \psi_k \circ f(x_0))}{J^*(x_0)} \\ &= \lambda_{r+1}((Y_k - X_k) X^{-1}) + 1 \\ &\to \lambda_{r+1}(YX^{-1}),\end{aligned}$$

where the inequality follows from

$$\tilde{J}(x_0, u) \geq \tilde{J}_k(x_0, \psi_k(u)).$$

Therefore

$$\mu_r^* \geq \lambda_{r+1}(Y_c X_c^{-1}).$$

The equality can be proved by showing that the the right side is achieved as the degree of performance degeneration for the control law given in Theorem 1. Since

$$\begin{aligned}\Delta J(x_0, f_r^*) &= \left\|x_0 - T^{-1} \operatorname{diag}\{I_r, 0_{n-r}\} T x_0\right\|_{Z_c}^2 \\ &= \left\|T^{-1} \operatorname{diag}\{0_r, I_{n-r}\} T x_0\right\|_{Z_c}^2,\end{aligned}$$

it follows that

$$\begin{aligned}\mu_r(f_r^*) &= \max_{x_0\neq 0} \frac{\left\|T^{-1} \operatorname{diag}\{0_r, I_{n-r}\} T x_0\right\|_{Z_c}^2}{\|x_0\|_{X_c}^2} + 1 \\ &= \max_{x_0\neq 0} \frac{\|\Sigma_r x_0\|^2}{\|x_0\|^2} + 1 \\ &= \sigma_{\max}(\Sigma_r) + 1 \\ &= \lambda_{r+1}(Y_c X_c^{-1}),\end{aligned}$$

thus the equality holds. ■

IV. COMPONENT ANALYSIS BASED ON THE ESTIMATION PERFORMANCE CRITERION

In this section, we provide a component analysis of the state variables based on the estimation performance criterion. The result shows the dominant components of the state which have major contribution to enhance the estimation performance.

A. The Setup and the Result

We consider the system

$$x_{t+1} = Ax_t + w_t,\ y_t = Cx_t + v_t,\ t \geq 0,$$

where x is the state, w is the disturbance, v is the noise, and y is the output. Assume that

$$x_0 \sim \mathcal{N}(0, P),\ w_t \sim \mathcal{N}(0, W),\ v_t \sim \mathcal{N}(0, V)$$

and that

$$\begin{gathered} \mathrm{E}\, x_0 w_t^{\mathrm{T}} = 0,\ \mathrm{E}\, x_0 v_t^{\mathrm{T}} = 0, \\ \mathrm{E}\, w_i w_j^{\mathrm{T}} = 0,\ \mathrm{E}\, v_i v_j^{\mathrm{T}} = 0 \text{ for } i \neq j, \\ \mathrm{E}\, w_i v_j^{\mathrm{T}} = 0 \text{ for all } i, j. \end{gathered}$$

In the optimal filtering problem, the estimate $\hat{x}_t$ of x_t is determined as

$$\hat{x}_t = g_t(\tilde{y}_t),\ \tilde{y}_t := (y_0, \ldots, y_t)$$

by the estimation law $g_t : \mathbb{R}^{(t+1)\ell} \to \mathbb{R}^n$, and the estimation performance is evaluated by the criterion

$$J_t(\xi, g_t) := \mathrm{E}\left[\xi^{\mathrm{T}}(x_t - \hat{x}_t)\right]^2$$

for some $\xi \in \mathbb{R}^n$. Then the optimal performance is given as

$$J_t^*(\xi) := \min_{g_t} J_t(\xi, g_t) = \xi^{\mathrm{T}} P_t\, \xi,$$

where

$$P_t := \mathrm{cov}(e_t) = Y_t - Y_t C^{\mathrm{T}}(C Y_t C^{\mathrm{T}} + V)^{-1} C Y_t$$

for the solution of (8) for $Y_0 = P$. Here we define the degree of performance degeneration due to the dimensional constraint as

$$\nu_r^* := \limsup_{t\to\infty} \nu_{rt}^*,\ \nu_{rt}^* := \min_{\dim \mathrm{Im}\, g_t \leq r} \nu_t(g_t),$$

where

$$\nu_t(g_t) := \max_{\xi \neq 0} \frac{J_t(\xi, g_t)}{J_t^*(\xi)}.$$

Under the above setup, the result of the component analysis is given as the next theorem.

Theorem 2: Suppose that A is stable and (C, A) is observable. Then

$$\nu_r^* = \lambda_{r+1}(X_o Z_o^{-1}).$$

Furthermore, an optimal estimation law, that is, a function g_{rt}^* which achieves $\nu_t(g_{rt}^*) = \nu_{rt}^*$ and satisfies $\dim \mathrm{Im}\, g_{rt}^* \leq r$ is given by

$$\begin{gathered} g_{rt}^*(\tilde{\eta}_t) = \hat{\xi}_t,\ T_t \hat{\xi}_t = \mathrm{diag}\left\{I_r, 0_{n-r}\right\} T_t \xi_t, \\ \xi_t = g_{nt}^*(\tilde{\eta}_t),\ T_t = V_t Y_t^{1/2} V_t^{\mathrm{T}}, \end{gathered}$$

where V_t is the orthogonal matrix such that

$$\begin{aligned} & V_t Y_t^{-1/2} P_t Y_t^{-1/2} V_t^{\mathrm{T}} \\ & = \mathrm{diag}\left\{\lambda_1(P_t Y_t^{-1}), \ldots, \lambda_n(P_t Y_t^{-1})\right\}. \end{aligned}$$

We can see that ν_r^* is invariant under linear transformation of the state variable. The dominant components of the state are the first to the r-th elements of the transformed variable $\hat{x}_t = T x_t$. The optimal estimate under the dimensional constraint can be generated by reducing the non-dominant components of the estimate obtained by applying the optimal law under no constrain.

B. Proof of the Theorem

Since x_0, w, and v are all Gaussian, we can restrict our attention to the case that the estimation law is linear, i.e. given as

$$g_k(\tilde{\eta}_k) = \hat{G}_k \tilde{\eta}_k,\ \tilde{\eta}_k := (\eta_0, \ldots, \eta_k) \tag{12}$$

by some matrix $\hat{G}_k \in \mathbb{R}^{n\times(k+1)\ell}$. This follows from the next Lemma, which is an easy consequence of the optimality of the linear estimation for Gaussian random variables.

Lemma 5: For random vectors x in $\mathbb{R}^m$ and y in $\mathbb{R}^n$ such that

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^{\mathrm{T}} & \Sigma_{yy} \end{bmatrix}\right),\ \Sigma_{xx} > 0,\ \Sigma_{yy} > 0,$$

and class of a map $g : \mathbb{R}^n \to \mathbb{R}^m$. Then

$$\min_{\dim \mathrm{Im}\, g \leq r} \mathrm{E}\, \|x - g(y)\|^2 = \sum_{i=r+1}^{n} \lambda_i\left(\Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{xy}^{\mathrm{T}}\right).$$

The minimum can be achieved by the map

$$g_r(\eta) = G_r \eta,$$

where

$$\begin{gathered} \bar{G}_r = G_r \Sigma_{yy}^{1/2},\ M = \Sigma_{xy}\Sigma_{yy}^{-1},\ \bar{M} = M \Sigma_{yy}^{1/2}, \\ T_i = U \begin{bmatrix} \Sigma_i & 0_{i\times(n-i)} \\ 0_{(m-i)\times i} & 0_{(m-i)\times(n-i)} \end{bmatrix} V^{\mathrm{T}}, \\ \Sigma_i = \mathrm{diag}\left\{\sigma_1(\bar{M}), \ldots, \sigma_i(\bar{M}), 0, \ldots, 0\right\} \in \mathbb{R}^{k\times k} \end{gathered}$$

with $k = \mathrm{rank}\, M$, and $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ are the orthogonal matrices such that $U\bar{M}V^{\mathrm{T}} = \Sigma_k$.

The following result is known in the Kalman filtering theory.

Lemma 6: Suppose that the estimation law is given as (12). Then the covariance of the error can be represented as

$$\mathrm{cov}(e_k) = (G_k - \hat{G}_k) S_k (G_k - \hat{G}_k)^{\mathrm{T}} + Y_k \tag{13}$$

by some matrix $G_k \in \mathbb{R}^{n\times(k+1)m}$ and positive definite matrix $S_k \in \mathbb{R}^{(k+1)m\times(k+1)m}$ satisfying

$$G_k S_k G_k^{\mathrm{T}} = Z_k, \tag{14}$$

where $Z_k = X_k - Y_k$ for the solutions of (6) and (8) for $X_0 = Y_0 = P$.

The following fact is immediate from the theory of singular value decomposition.

Lemma 7: For a matrix $G \in \mathbb{R}^{n\times p}$ and positive definite matrices $S \in \mathbb{R}^{p\times p}$, $Y \in \mathbb{R}^{n\times n}$ with $p \geq n$,

$$\min_{\mathrm{rank}\,\hat{G} \leq r} \max_{x\neq 0} \frac{\|(G-\hat{G})^{\mathrm{T}}x\|_S^2}{\|x\|_Y^2} = \lambda_{r+1}(ZY^{-1}), \qquad (15)$$

where $Z = F^{\mathrm{T}}SF$. The minimum can be achieved by

$$\hat{G}_r^* = Y^{1/2}U\Sigma_r V^{\mathrm{T}} S^{1/2},$$

where $\Sigma_r \in \mathbb{R}^{n\times p}$ is given as

$$\Sigma_r = \left[\begin{array}{cc} \hat{\Sigma}_r & 0_{n\times(p-n)} \end{array}\right],$$
$$\hat{\Sigma}_r = \mathrm{diag}\,\{\lambda_1(ZY^{-1}),\ldots,\lambda_r(ZY^{-1}),0\ldots,0\},$$

and $U, V \in \mathbb{R}^{n\times n}$ are the orthogonal matrices such that

$$Y^{-1/2}FS^{1/2} = U\Sigma_n V^{\mathrm{T}}.$$

Furthermore, the map $g_r^*(y) = \hat{G}_r^* y$ satisfies

$$Tg_r^*(y) = \mathrm{diag}\,\{I_r, 0_{n-r}\}\,Tg_n^*(y),\ T = U^{\mathrm{T}}Y^{-1/2}.$$

Lemma 8: For the solutions of (6) and (8) for $X_0 = Y_0 = P$,

$$\nu_{kr}^* = \lambda_{r+1}(X_k Y_k^{-1}).$$

Proof: From Lemma 6,

$$\frac{\Delta J_k(\xi, g_k)}{J_k^*(\xi)} = \frac{\|(G_k-\hat{G}_k)^{\mathrm{T}}\xi\|_{S_k}^2}{\|\xi\|_{Y_k}^2}.$$

By applying Lemma 7, we have

$$\begin{aligned}\nu_{kr}^* &= \min_{\dim \mathrm{Im}\, g_k \leq r} \max_{\xi\neq 0} \frac{\Delta J_k(\xi,g_k)}{J_k^*(\xi)} + 1\\ &= \min_{\mathrm{rank}\,\hat{G}_k \leq r} \max_{\xi\neq 0} \frac{\|(G_k-\hat{G}_k)^{\mathrm{T}}\xi\|_{S_k}^2}{\|\xi\|_{Y_k}^2} + 1\\ &= \lambda_{r+1}((X_k - Y_k)Y_k^{-1}) + 1\\ &= \lambda_{r+1}(X_k Y_k^{-1}).\end{aligned}$$

■

Proof of Theorem 2: The proof is immediate from Lemma 8 since $\lambda_{r+1}(X_k Y_k^{-1}) \to \lambda_{r+1}(X_o Y_o^{-1})$. ■

V. COMPONENT ANALYSIS BASED ON THE CONTROL AND ESTIMATION PERFORMANCE CRITERIA

In this section, we provide a component analysis of the state variables based on the both control and estimation performance criteria by investigating the effect of the past observed output to the future control performance under the dimensional constraint on the control law.

A. The Setup and the Result

We consider the system

$$x_{t+1} = Ax_t + w_t + Bu_t,\ y_t = Cx_t + v_t,$$
$$u_t = 0 \text{ for } t < 0,$$
$$w_t = 0 \text{ and } v_t = 0 \text{ for } t \geq 0.$$

The assumptions on w and v are same as those in Section IV. We also assume that $x_t \sim \mathcal{N}(0, \mathrm{cov}(x_t))$, $\mathrm{cov}(x_t)$ is bounded, $\mathrm{E}\,x_t w_t^{\mathrm{T}} = 0$, and $\mathrm{E}\,x_t v_t^{\mathrm{T}} = 0$. In the optimal control problem, the control input is determined as

$$(u_0, u_1, \ldots) = h(y_0, y_{-1}, \ldots),$$

by the control law $h : \ell^\infty \to \ell^\infty$, and the control performance is evaluated by the criterion

$$J(h) = \mathrm{E}\sum_{t=0}^{\infty}\left(x_t^{\mathrm{T}}Q\,x_t + u_t^{\mathrm{T}}R\,u_t\right).$$

Then the optimal performance is given as

$$J^* := \min_h J(h) = \mathrm{tr}\,(X_cY_o + X_oY_c - Y_cY_o).$$

Here we define the performance degeneration due to the dimensional constraint as

$$\Delta J_r^* := \min_{\dim \mathrm{Im}\, h \leq r} \Delta J(h),$$

where

$$\Delta J(h) := J(h) - J^*.$$

Note that

$$\Delta J_0^* \geq \Delta J_1^* \geq \cdots \geq \Delta J_n^* = 0.$$

Under the above setup, the result of the component analysis is given as the next theorem.

Theorem 3: Suppose that A is stable, (A, B) is controllable, and (C, A) is observable. Then

$$\Delta J_r^* = \sum_{i=r+1}^{n} \lambda_i(Z_cZ_o), \qquad (16)$$

where $Z_c = X_c - Y_c$ and $Z_o = X_o - Y_o$. Furthermore, an optimal control law, that is, a function h_r^* which achieves $\Delta J(h_r^*) = \Delta J_r^*$ and satisfies $\dim \mathrm{Im}\, h_r^* \leq r$ is given by

$$h_r^*(\eta_0, \eta_{-1}, \ldots) = f_n^*(\hat{\xi}_0^*),\ T\,\hat{\xi}_0^* = \mathrm{diag}\,\{I_r, 0_{n-r}\}\,T\,\hat{\xi}_0,$$
$$\hat{\xi}_0 = g_n^*(\eta_0, \eta_{-1}, \ldots),\ T = VZ_c^{1/2},$$

where V is the orthogonal matrix such that

$$VZ_c^{1/2}Z_oZ_c^{1/2}V^{\mathrm{T}} = \mathrm{diag}\,\{\lambda_1(Z_cZ_o),\ldots,\lambda_n(Z_cZ_o)\}.$$

We can see that ΔJ_r^* is invariant under linear transformation of the state variable. The dominant components of the state are the first to the r-th elements of the transformed variable $\hat{x}_t = Tx_t$. The optimal input under the dimensional constraint can be generated by applying the optimal control law under no constraint to the approximate of the estimate of the initial value obtained by reducing the non-dominant components of the estimate generated by the optimal estimation law under no constrain.

B. Proof of the Theorem

To prove the above theorem, we consider the finite time optimal control problem. In the control over finite time steps from 0 based on the observation over finite time steps to 0, the control input is determined as

$$\tilde{u}_k = h_{jk}(\tilde{y}_j),$$
$$\tilde{u}_k := (u_0, \dots, u_k),\ \tilde{y}_j := (y_{-j}, \dots, y_0)$$

by the control law $h_{jk} : \mathbb{R}^{(j+1)\ell} \to \mathbb{R}^{(k+1)m}$, and the control performance is evaluated by the criterion

$$J_{jk}(h_{jk}) = \mathrm{E} \sum_{t=0}^{k} \left(x_t^{\mathrm{T}} Q\, x_t + u_t^{\mathrm{T}} R\, u_t \right).$$

Then the optimal performance is given as

$$J_{jk}^* := \min_{h_{jk}} J_{jk}(h_{jk}) = \mathrm{tr}\,(X_k^c Y_j^o + X_k^o Y_j^c - Y_k^c Y_j^o),$$

where X_t^c and X_t^o are the solutions X_t of (5) and (6) for $X_0^c = X_j^o$ and $X_0^o = 0$, and Y_t^c and Y_t^o are the solutions Y_t of (7) and (8) for $Y_0^c = X_j^o$ and $Y_0^o = 0$, respectively. Here we define the performance degeneration due to the dimensional constraint as

$$\Delta J_{jkr}^* := \min_{\dim \mathrm{Im}\, h_{jk} \le r} \Delta J_{jkr}(h_{jk}),$$

where

$$\Delta J_{jk}(h_{jk}) := J_{jk}(h_{jk}) - J_{jk}^*.$$

Lemma 9: For the matrices $Z_k^c = X_k^c - Y_k^c$ and $Z_j^o = X_j^o - Y_j^o$,

$$\Delta J_{jk,r}^* = \sum_{i=r+1}^{n} \lambda_i(Z_k^c Z_j^o).$$

Proof: From the results in optimal control theory, it follows that

$$J_{jk}(h_{jk}) = \mathrm{tr}(Y_k^c X_j^o) + \mathrm{E}\,\|F_k^c x_0 + \tilde{u}_k\|_{S_k^c}^2$$

and

$$\min_{\dim \mathrm{Im}\, h_{jk} \le r} \mathrm{E}\,\|F_k^c x_0 + \tilde{u}_k\|_{S_k^c}^2 = \mathrm{tr}\, Y_j^o (X_k^c - Y_k^c) + \min_{\dim \mathrm{Im}\, h_{jk} \le r} \mathrm{E}\,\|F_k^c \hat{x}_0 + \tilde{u}_k\|_{S_k^c}^2 .$$

Therefore

$$\begin{aligned}
\Delta J_{jkr}^* &= \min_{\dim \mathrm{Im}\, h_{jk} \le r} \mathrm{E}\,\|F_k^c\, \mathrm{E}\,[x_0|\tilde{y}_j] + h_{jk}(\tilde{y}_j)\|_{S_k^c}^2 \\
&= \min_{\dim \mathrm{Im}\, h_{jk} \le r} \mathrm{E}\,\left\|F_k^c F_j^o \tilde{y}_j + h_{jk}(\tilde{y}_j)\right\|_{S_k^c}^2 \\
&= \min_{\mathrm{rank}\, G_j \le r} \mathrm{E}\,\left\|F_k^c F_j^o \tilde{y}_j + G_j \tilde{y}_j\right\|_{S_k^c}^2 \\
&= \min_{\mathrm{rank}\, H_j \le r} \left\|S_k^{c\,1/2} F_k^c F_j^o S_j^{o\,1/2} + H_j\right\|^2 \\
&= \sum_{i=r+1}^{n} \lambda_i(Z_k^c Z_j^o).
\end{aligned}$$

■

Proof of Theorem 3: From Lemma 7, $\lim_{k\to\infty} \lim_{j\to\infty} \Delta J_{jkr}^*$ is equal to the right side of (16), so the theorem can be proved by showing that it is also equal to ΔJ_r^*. Let $\Delta \hat{J}_{jkr}^*$ be the minimum of the performance degeneration corresponding to the case where $w_t = 0$ and $v_t = 0$ for $t < -j$. It follows that $\lim_{j\to\infty} \Delta J_{jkr}^* = \Delta J_{\infty kr}^*$ since $J_{jkr}^* \le J_{\infty kr}^*$ follows from $\lim_{j\to\infty} \Delta J_{jkr}^* = \lim_{j\to\infty} \Delta \hat{J}_{jkr}^* \ge \Delta J_{\infty kr}^*$, and $J_{jkr}^* \ge J_{\infty kr}^*$ follows from $J_{jkr}^* \le J_{\infty kr}^*$. Therefore it is verified that $\lim_{k\to\infty} \lim_{j\to\infty} \Delta J_{jkr}^* = \Delta J_{\infty\infty r}^*$ since $\lim_{k\to\infty} \lim_{j\to\infty} \Delta J_{jkr}^* \le \Delta J_{\infty\infty r}^*$ follows from $J_{\infty kr}^* \le J_{\infty\infty r}^*$ and $\lim_{k\to\infty} J_{\infty kr}^* = J_{\infty\infty r}^*$, and $\lim_{k\to\infty} \lim_{j\to\infty} \Delta J_{jkr}^* \ge \Delta J_{\infty\infty r}^*$ from $\Delta J_{\infty\infty r}^* \ge \lim_{k\to\infty} \lim_{j\to\infty} \Delta J_{jkr}^*$. ■

VI. CONCLUSIONS

We have provided a component analysis for the state variables of stable linear discrete-time systems based on the regulation and filtering performance criteria using the solutions of both Lyapunov and Riccati equations. The analysis shows the dominant components of the state variables which have major contribution to enhance the performances, and degree of contribution can be represented as invariants the eigenvalues depending on these solutions.

REFERENCES

[1] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, SIAM, 2005.

[2] B. C. Moore, "Principal component analysis in linear systems — controllability, observability, and model reduction, " *IEEE Transactions on Automatic Control*, Vol. 26, No. 1, pp. 17–32, 1981.

[3] R. A. Roberts and C. T. Mullis, *Digital Signal Processing*, Addison-Wesley, 1987.

[4] U. Desai and D. Pal, "A transformation approach to stochastic model reduction, " *IEEE Transactions on Automatic Control*, Vol. 29, No. 12, pp. 1097–1100, 1984.

[5] D. F. Enns, "Model reduction with balanced realizations —An error bound and a frequency weighted generalization, " *Proceedings of the 23rd IEEE Conference on Decision and Control*, pp. 127–132, 1984.

[6] E. A. Jonckheere and L. M. Silverman, "A new set of invariants for linear systems —Application to reduced order compensator design, " *IEEE Transactions on Automatic Control*, Vol. 28, No. 10, pp. 953–964, 1983.