

Coping with singular transition matrices in estimation and control stability theory†

J. B. MOORE‡ and B. D. O. ANDERSON‡

When stabilizing linear discrete-time finite dimensional systems in control and estimation either optimally or suboptimally, technical difficulties arise in the conventional stability theories for coping with state transition matrices which are permitted to be singular or with eigenvalues arbitrarily small. In overcoming these difficulties, earlier results for feedback stabilization of linear systems and for Kalman filters and regulators are generalized in this paper, with proofs being in fact more direct than those explored earlier.

1. Introduction

The stabilization of a linear system by means of state variable feedback using a simply derived suboptimal law is studied in Kleinman (1970, 1974), Kwon and Pearson (1977, 1978) and Ikeda *et al.* (1975) and by means of a quadratic index minimization procedure in Kalman and Engla (1965), Caines and Mayne (1970) and Hager and Hurowitz (1976). Dual problems arise in state estimator theory (Hager and Hurowitz 1976, Deyst and Price 1968, McGarty 1974, Jazwinski 1970, Anderson 1971). For all these problems, it has become a practice in the discrete-time case to avoid technical difficulties which appear to exist if one permits singular state transition matrices or asymptotically singular ones.

It is intuitively apparent that there should be no difficulties in practice should a transition matrix approach a singular matrix or become singular, for then a linear functional of the state approaches zero or becomes zero. In adaptive control or filtering theory, it is important not to artificially constrain the parameter selections so as to avoid singular or near singular matrices, merely for the sake of an inadequate theory. After all, this would even exclude tapped-delay-line models. Thus there is now added motivation to overcome the technical difficulties of singular or asymptotically singular transition matrices in control and estimation stability theory.

In this paper, not only is the earlier optimal and suboptimal stabilization theory generalized to cope with singular (or approaching singular) state transition matrices in the discrete time case, but more direct derivations are given than in the earlier theory.

More specifically, for the suboptimal feedback stabilization problem, Kleinman (1970) gave a very simple method for stabilizing a continuous-time, time-invariant linear system under a controllability assumption by means of a state variable feedback law involving the controllability Gramian. Results for the discrete time situation were developed by Kleinman (1974), albeit

Received 16 July 1979.

† Supported by the Australian Research Grants Committee.

‡ Department of Electrical Engineering, University of Newcastle, New South Wales 2308, Australia.

with a good deal more complexity in the proof, and an inability to handle the case of singular transition matrices, although it was conjectured that the results apply in this case. The corresponding time-varying problems are considered by Kwon and Pearson (1977, 1978), who point out that a direct generalization to the time-varying case is not at all straightforward. They use a rather lengthy indirect approach involving establishing the instability of an adjoint system; again, the discrete-time results are harder to obtain than the continuous-time ones, and disallow the possibility of singular transition matrices or even non-singular transition matrices with unbounded inverse. Another work in this area is Ikeda *et al.* (1975), which treats continuous-time systems only, which may be time-varying. A minor modification of the law suggested by Kwon and Pearson is used, and the proof that the law is stabilizing is fairly straightforward.

In the next three sections of the paper a direct generalization of Kleinman's results are given for the time-varying case by exploiting a duality between the control problem and an estimation problem of interest in its own right. The results are distinguished from the results of Kleinman (1974) and Kwon and Pearson (1978) in permitting singular or asymptotically singular transition matrices for the discrete-time case and exhibiting in the time-varying situation a Lyapunov function for a closed loop system associated with, but not identical to that appearing in the control and estimation problems. The more general control problem of Kwon and Pearson (1977, 1978) is not studied here (though the methods of this paper would certainly seem applicable). This avoids distraction with comparative inessentials.

For the optimal control and filtering problem in discrete time where the transition matrices can become singular or approach asymptotically a singular matrix, the work of Hager and Hurowitz (1976), valid for infinite dimensional systems also, derives the control results for time-invariant systems and then dualizes to achieve the filtering results. It is claimed that the generalization carried through for time-varying systems, but in our view the generalizations are not possible without the theory of the next three sections of this paper or generalizations of it. In particular the duality between stabilizability and detectability as defined in Hager and Hurowitz (1976) for the time-varying case is not established. Moreover, problems arise since the time interval for filtering, viz. $[0, \infty)$, has a dual for control purposes of $(-\infty, 0]$ leading to difficulties. These difficulties can in part be solved by using intervals $(-\infty, \infty)$, but then a stability assumption on the open-loop system is necessary in the filtering problem.

Simpler derivations of the optimal control and filtering results using the convenient non-singularity assumption on the transition matrices, hidden in the uniform complete controllability and observability assumptions, are given in the earlier work of Caines and Mayne (1970), Deyst and Price (1968), McGarty (1974), Jazwinski (1970) and Anderson (1971). In Anderson (1971), asymptotic stability is discussed when controllability fails, but the details supplied are incomplete.

In § 6 of the paper, we rederive directly the discrete-time Kalman filter stability results for the time-varying case permitting transition matrices with arbitrarily small eigenvalues. The theory can be viewed as building on the work of the earlier sections and generalizes many of the results of Caines and

Mayne (1970), Hager and Hurowitz (1976), Deyst and Price (1968), McGarty (1974), Jazwinski (1970) and Anderson (1971). The derivations are significant in that they are more direct and more complete than earlier versions. Corresponding control results are then achieved by direct derivations carried out using the same technical devices. These generalize the earlier results of Kalman and Engla (1965) and Caines and Mayne (1970) and clarify specializations of the results of Hager and Hurowitz (1976).

2. The control problem

We study the linear system with state equation

$$x_{k+1} = F_k x_k + G_k u_k, \quad k = 0, 1, \dots \quad (2.1)$$

Let $\phi_{k,i}$ for $k \geq i$ denote the transition matrix, i.e.

$$\phi_{i,i} = I \quad \text{and} \quad \phi_{k,i} = F_{k-1} F_{k-2} \dots F_i$$

for $k > i$. To construct a stabilizing control law, we shall assume that (2.1) is uniform with respect to complete reachability (Ikeda *et al.* 1975): with reachability Gramian defined as

$$Y_{k,k-s} = \sum_{i=k-s}^k \phi_{k+1,i+1} G_1 G_1' \phi'_{k+1,i+1} \quad (2.2)$$

then for some $s > 0$, $\alpha > 0$ and all k ,

$$Y_{k,k-s} \geq \alpha I \quad (2.3)$$

The stabilizing feedback control law† we consider, a time-varying generalization of that given in Kleinman (1974), is as follows

$$u_k = -L_k x_k, \quad L_k = G_k' \phi'_{k+s+2,k+1} Y_{k+s+1}^{-1} \phi_{k+s+2,k+1} F_k \quad (2.4)$$

where the existence of the inverse is guaranteed by the reachability assumption (2.3). We seek a proof of the following theorem.

Theorem 2.1 (control system stability)

With $[F_k, G_k]$ uniform with respect to complete reachability, F_k, G_k bounded above, and the feedback gain L_k defined as in (2.4), the closed loop system

$$x_{k+1} = (F_k - G_k L_k) x_k \quad (2.5)$$

is exponentially asymptotically stable.

Method of proof

As a first step, in § 3, we define a dual estimator stability problem and establish the duality. In § 4, an equivalent stability problem is given with a readily constructed Lyapunov function confirmed in § 5. The results are tied together with discussion in § 6.

† An open loop control law which brings an arbitrary state x_k to the zero state at time $k+s+1$ under the reachability assumption is of course more readily derived.

3. A dual estimator problem

A dual problem somewhat more convenient to solve concerns estimator stability. Suppose that the pair $[A_k, C_k]$ is uniform with respect to complete observability (Ikeda *et al.* 1975), i.e. with $\psi_{i,k} = A_{i-1}A_{i-2} \dots A_k$ for $i > k$, $\psi_{i,i} = T$, and the observability Gramian defined as

$$M_{k+t,k} = \sum_{i=k}^{k+t} \psi'_{i,k} C'_i C_i \psi_{i,k} \tag{3.1}$$

then for some $t > 0$, $\beta > 0$ and all k

$$M_{k+t,k} \geq \beta I \tag{3.2}$$

A standard estimator for $x_{j+1} = A_j x_j + u_j$, $y_j = C_j x_j$ implemented with an estimator gain law K_j gives rise to an error equation

$$\xi_{j+1} = (A_j - K_j C_j) \xi_j \tag{3.3}$$

We shall consider this equation where we choose

$$K_j = A_j \psi_{j,j-t-1} M_{j,j-t-1}^{-1} \psi'_{j,j-t-1} C'_j \tag{3.4}$$

Of course, for satisfactory estimation (3.3) should represent an asymptotically stable system, and this is the content of the following result :

Theorem 3.1 (estimator stability)

With $[A_k, C_k]$ uniform with respect to complete observability and A_k, C_k bounded above, then the error system (3.3) with K_j as in (3.4) is exponentially asymptotically stable.

The *duality* between the control problem of Theorem 2.1 and the estimator problem of Theorem 3.1 is *not* defined using the adjoint notion of, e.g. Kwon and Pearson (1977), which requires transition matrix invertibility. Rather we require the relationships

$$A_k \equiv F'_{-k}, \quad C_k \equiv G'_{-k}, \quad s = t \tag{3.5}$$

which we explore in the following lemma.

Lemma 3.1

Under the duality definitions (3.5), the transition matrix definitions, and the controllability and observability Gramian definitions (2.2) and (3.1) :

(i) $\psi_{l,k+1} = \phi'_{-k,-l+1}, \quad M_{l,k+1} = Y_{-k-1,-l} \quad \text{for } l \geq k+1$

and

$$K_k = L'_{-k}, \quad (A_k - K_k C_k) = (F_{-k} - G_{-k} L_{-k})'$$

(ii) Exponential asymptotic stability of (2.5) is equivalent to exponential asymptotic stability of (3.3).

(iii) $[A_k, C_k]$ uniform with respect to complete observability is equivalent to $[F_k, G_k]$ uniform with respect to complete reachability.

(iv) Theorem 2.1 and Theorem 3.1 are dual results.

Proof

(In the main, a discrete time version of results in Ikeda *et al.* (1975).)

$$(i) \quad \begin{aligned} \psi_{i,k+1} &= A_{i-1}A_{i-2} \dots A_{k+1} = F'_{-(i-1)} F'_{-(i-2)} \dots F'_{-(k+1)} \\ &= [F'_{-(k+1)} \dots F'_{-(i-2)} F'_{-(i-1)}]' = \phi'_{-k,-i+1} \end{aligned}$$

The other results of (1) follow directly from this result, the definitions, and (3.5).

(ii) Let $\bar{\psi}_{i,j}$ and $\bar{\phi}_{i,j}$ denote the transition matrices associated with $A_k - K_k C_k$ and $F_k - G_k L_k$. By (i), we have

$$\|\bar{\psi}_{i,k+1}\| = \|\bar{\phi}_{-k,-i+1}\|$$

Therefore, if for all $k < l$ and some $\alpha_1 > 0$, $\lambda \in [0, 1)$, there holds

$$\|\bar{\psi}_{i,k+1}\| \leq \alpha_1 \lambda^{l-(k+1)}$$

it follows that

$$\|\bar{\phi}_{-k,-i+1}\| \leq \alpha_1 \lambda^{l-(k+1)}$$

whence for all $i > j$, $\|\bar{\phi}_{i,j}\| \leq \alpha_1 \lambda^{i-j}$. The converse is trivial.

(iii) Follows from the definitions and the relationship

$$M_{i,k+1} = Y_{-k-1,-i}$$

(iv) Follows from (ii) and (iii) immediately. ▽▽▽

4. An equivalent stability problem

We keep the same notation as above, save that we shall write m for the frequently occurring subscript $j + t + 2$. Consider the system

$$\zeta_{j+1} = \bar{A}_j \zeta_j, \quad \bar{A}_j = A_j - M_{m,j+1}^{-1} \psi'_{m,j+1} C'_m C_m \psi_{m,j+1} A_j \tag{4.1}$$

The following lemma relates (4.1) and the error equation (3.3).

Lemma 4.1

Suppose that $\xi_m = \psi_{m,j} \zeta_j$ for some $j = j_0$. Then equality holds for all $j \geq j_0$. Suppose (3.3) is initialized at time $j_0 + t + 2$. Then there exists ζ_{j_0+t+2} such that $\xi_m = \psi_{m,j} \zeta_j$ for all $j \geq j_0 + t + 2$.

Proof

It is trivial to check that

$$\begin{aligned} \psi_{m+1,j+1} (A_j - M_{m,j+1}^{-1} \psi'_{m,j+1} C'_m C_m \psi_{m,j+1} A_j) \\ = (A_m - A_m \psi_{m,j+1} M_{m,j+1}^{-1} \psi'_{m,j+1} C'_m C_m) \psi_{m,j} \end{aligned}$$

or with

$$\begin{aligned} \tilde{A}_m &= A_m - K_m C_m \\ \psi_{m+1,j+1} \tilde{A}_j &= (\tilde{A}_m) \psi_{m,j} \end{aligned} \tag{4.2}$$

Setting $j = j_0$ and multiplying by ζ_{j_0} shows that if $\xi_m = \psi_{m,j} \zeta_j$ for $j = j_0$, the same is true for $j = j_0 + 1$.

Second, set $u_j = M_{m,j}^{-1} \psi'_{m,j} C'_m C_m \xi_m$. Then (3.3) becomes

$$\xi_{m+1} = A_m \xi_m + \psi_{m+1,j} u_m \tag{4.3}$$

It follows that

$$\begin{aligned} \xi_{m+t+2} &= \psi_{m+t+2,m} \xi_m + \sum_{i=m}^{m+t+1} \psi_{m+t+2,i+1} \psi_{i+1,i-t-2} u_i \\ &= \psi_{m+t+2,m} \left(\xi_m + \sum_{i=m}^{m+t+1} \psi_{m,i-t-2} u_i \right) \\ &\in \mathcal{R}[\psi_{m+t+2,m}] \end{aligned}$$

The second claim of the lemma is now immediate on using the first part.

△△△

Lemma 4.2

With A_j, C_j bounded above and $[A_j, C_j]$ uniform with respect to complete observability, the system (4.1) is exponentially asymptotically stable if and only if (3.3) is exponentially asymptotically stable.

Proof

Suppose (4.1) is exponentially asymptotically stable. Let (3.3) commence at time $j_0 + t + 2$. For all $j \geq j_0 + t + 2$, we can write the solution ξ_m of (3.3) as $\xi_m = \psi_{m,j} \zeta_j$ where ζ_j is the solution of (4.1) with suitable initial condition ζ_{j_0+t+2} . Since $\psi_{m,j}$ is bounded, exponential convergence of ζ_j implies exponential convergence of ξ_j .

Conversely, suppose (3.3) is exponentially stable. By Lemma 4.1, with $U_j = M_{m,j+1}^{-1} \psi'_{m,j+1} C'_m C_m$, the evolution of ζ_{j+1} according to (4.1) can be written as

$$\zeta_{j+1} = A_j \zeta_j - U_j \xi_m \tag{4.4}$$

whence

$$\begin{aligned} \zeta_m &= \psi_{m,j} \zeta_j - \sum_{i=j}^{m-1} \psi_{m,i+1} U_i \xi_{i+t+2} \\ &= \xi_m - \sum_{i=j}^{j+t+1} \psi_{m,i+1} U_i \xi_{i+t+2} \end{aligned}$$

This expresses ζ_m as a weighted average (without bounded weights) of $\xi_m, \xi_{m+1}, \dots, \xi_{m+t+1}$. Accordingly, if ξ_j converges exponentially fast, so does ζ_j .

△△△

We now seek a Lyapunov function for (4.1).

5. Lyapunov function

Adopt as a tentative Lyapunov function for (4.1) $V_j = \zeta'_j M_{m-1,j} \zeta_j$ where m is shorthand for $j + t + 2$. Using the easily checked fact that

$$M_{m-1,j} = C'_j C_j + A'_j [M_{m,j+1} - \psi'_{m,j+1} C'_m C_m \psi_{m,j+1}] A_j$$

we obtain

$$\begin{aligned} V_j - V_{j+1} &= \zeta'_j [C'_j C_j + \psi'_{m,j} C'_m C_m \psi_{m,j} - \psi'_{m,j} C'_m C_m \psi_{m,j+1} \\ &\quad \times M_{m,j+1}^{-1} \psi'_{m,j+1} C'_m C_m \psi_{m,j}] \zeta_j \\ &= \zeta'_j \{C'_j C_j + \psi'_{m,j} C'_m [I - C_m \psi_{m,j+1} (M_{m-1,j+1} \\ &\quad + \psi'_{m,j+1} C'_j C_j \psi_{m,j+1})^{-1} \psi'_{m,j+1} C'_m] C_m \psi_{m,j}\} \zeta_j \\ &= \zeta'_j \{C'_j C_j + \psi'_{m,j} C'_m [I + C_m \psi_{m,j+1} M_{m-1,j+1}^{-1} \psi'_{m,j+1} C'_m]^{-1} \\ &\quad \times C_m \psi_{m,j}\} \zeta_j \end{aligned}$$

The last line follows from the matrix inversion lemma. The existence of $M_{m-1,j+1}^{-1}$ is guaranteed by the observability condition (3.2). Now with the assumed bounds on A_j, C_j it follows that for some $\epsilon > 0$ and all j ,

$$[I + C_m \psi_{m,j+1} M_{m-1,j+1}^{-1} \psi'_{m,j+1} C'_m]^{-1} \geq \epsilon^2 I$$

and so

$$V_j - V_{j+1} \geq \zeta_j \bar{C}'_j \bar{C}_j \zeta_j$$

where

$$\bar{C}'_j = [C'_j \quad \epsilon \psi'_{m,j} C'_m]$$

Consequently,

$$V_j - V_m \geq \zeta'_j \left(\sum_{i=j}^{m-1} \bar{\psi}'_{i,j} \bar{C}'_i \bar{C}_i \bar{\psi}_{i,j} \right) \zeta_j$$

where $\bar{\psi}_{i,j}$ is the transition matrix associated with \bar{A}_j .

Now use Lemma A.1, identifying F_j with \bar{A}_j, H_j with \bar{C}_j and K_j with $[0 \quad \epsilon^{-1} M_{m,j+1}^{-1} \psi'_{m,j+1} C'_m]$. For some $\gamma > 0$,

$$\begin{aligned} V_j - V_m &\geq \gamma \zeta'_j \left(\sum_{i=j}^{m-1} \psi'_{i,j} \bar{C}_i \bar{C}'_i \psi_{i,j} \right) \zeta_j \\ &\geq \gamma \zeta'_j \left(\sum_{i=j}^{m-1} \psi'_{i,j} C_i C'_i \psi_{i,j} \right) \zeta_j = \gamma V_j \end{aligned}$$

Exponential decay of V_j , and thus ζ_j , then ξ_j by Lemma 4.2 follow. Accordingly, this establishes Theorem 3.1 and its dual Theorem 2.1.

It is of interest to attempt to formulate a Lyapunov function for the estimation error equations (3.3). For the case when $\psi_{k,l}$ is non-singular and bounded below in norm, we have that

$$V_j = \zeta'_j M_{m-1,j} \zeta_j = \xi'_m (\psi'_{m,j} M_{m-1,j} \psi_{m,j}^{-1}) \xi_m$$

is also a Lyapunov function for (3.3). However, when $\psi_{m,j}$ is permitted to approach a singular matrix or be a singular matrix, then it is clear from the relationship $\xi_m = \psi_{m,j} \zeta_j$ that ξ_m will approach zero faster than exponentially. For this situation, the standard Lyapunov theory for exponential stability is not applicable, and even the tentative Lyapunov function

$$V_j = \xi'_m (\psi_{m,j} M_{m-1,j}^{-1} \psi'_{m,j})^\# \xi_m$$

where $\#$ denotes a pseudo inverse does not yield the desired stability results. Likewise there is difficulty formulating a Lyapunov function for the control problem.

Theorems 2.1 and 3.1 place no requirement on the degree of stability of the closed-loop. Borrowing an idea used for continuous-time systems in Ikeda *et al.* (1975), we can secure a certain degree of stability as follows. In lieu of the quantity M in (3.1), we use

$$M_{k+t,k}^\rho = \sum_{i=k}^{k+t} \rho^{i-k} \psi'_{i,k} C'_i C_i \psi_{i,k}$$

Here, ρ is an arbitrary real constant in $(1, \infty)$. Equation (3.2) implies that $M_{k+t,k}^\rho$ is also strictly bounded away from zero. The estimator gain law K_j of (3.4) is replaced by

$$K_j = \rho^{t+1} A_j \psi_{j,j-t-1} (M_{j,j-t-1}^\rho)^{-1} \psi'_{j,j-t-1} C'_j$$

In (4.1), $M_{m,j+1}^{-1}$ is replaced by $\rho^{t+1} (M_{j,j-t-1}^\rho)^{-1}$ and the conclusions of § 4 remain valid. Finally, the Lyapunov function $\zeta'_j M_{m-1,j}^\rho \zeta_j$ is adopted for the modified (4.1), and one finds that

$$V_j - V_{j+1} \geq \frac{\rho - 1}{\rho} V_j$$

It is easily checked that larger ρ lead to more rapid exponential convergence for ζ_j . With $\xi_m = \psi_{m,j} \zeta_j$, ξ_m converges at least at the same rate, but may converge faster if $\psi_{m,j}$ is singular.

6. Optimal estimation and control

Existing theory for deriving stability results in optimal discrete-time filtering and control problems for quadratic indices and time varying linear systems requires that the state one-step transition matrices ϕ are non-singular with bounded inverse, see e.g. Caines and Mayne (1970), Deyst and Price (1968), McGarty (1974), Jazwinski (1970) and Anderson (1971). Here we relax this assumption obtaining conditions for exponential and non-exponential asymptotic stability.

Consider the stochastic signal model

$$x_{k+1} = A_k x_k + w_k, \quad z_k = C_k x_k + v_k, \quad E[x_0 x'_0] = P_0 \tag{6.1}$$

for w_k, v_k stochastic processes with zero mean independent of x_0 and

$$E \left[\begin{bmatrix} w_k \\ v_k \end{bmatrix} \begin{bmatrix} w'_t & v'_t \end{bmatrix} \right] = \begin{bmatrix} Q_k & 0 \\ 0 & R_k \end{bmatrix} \delta_{kt}, \quad Q_k \geq 0, R_k > 0 \tag{6.2}$$

A convenient filter structure is

$$\hat{x}_{k+1|k+1} = \bar{A}_k \hat{x}_{k|k} + \mathcal{K}_{k+1} z_{k+1}, \quad \bar{A}_k = A_k - \mathcal{K}_{k+1} C_{k+1} A_k \tag{6.3}$$

for some $\{\mathcal{K}_k\}$. The corresponding one-step ahead predictor is

$$\hat{x}_{k+1|k} = \tilde{A}_k \hat{x}_{k|k-1} + A_k \mathcal{K}_k z_k, \quad \tilde{A}_k = A_k - A_k \mathcal{K}_k C_k \tag{6.4}$$

The error covariance of the filter and predictor are respectively

$$\Sigma_{k|k} = \bar{A}_{k-1} \Sigma_{k-1|k-1} \bar{A}'_{k-1} + \mathcal{K}_k R_k \mathcal{K}'_k + (I - \mathcal{K}_k C_k) Q_k (I - \mathcal{K}_k C_k)' \tag{6.5 a}$$

$$\Sigma_{k+1|k} = \tilde{A}_k \Sigma_{k|k-1} \tilde{A}'_k + A_k \mathcal{K}_k R_k \mathcal{K}'_k A'_k + Q_k \tag{6.5 b}$$

The Kalman filter and one-step-ahead predictor minimize the above error covariances over all $\{\mathcal{K}_k\}$ with an optimal gain which to avoid undue repetition of equations is still denoted \mathcal{K}_k , but is given from

$$\mathcal{K}_k = \Sigma_{k|k-1} C'_k (C_k \Sigma_{k|k-1} C'_k + R_k)^{-1} = \Sigma_{k|k} C'_k R_k^{-1} \quad (6.6)$$

with $\Sigma_{k|k-1}$ here denoting the minimum error covariance. We shall use the now standard equations

$$\Sigma_{k+1|k} = A_k \Sigma_{k|k} A'_k + Q_k, \quad \Sigma_{0|-1} = P_0 \quad (6.7 a)$$

$$\Sigma_{k|k} = \Sigma_{k|k-1} - \Sigma_{k|k-1} C'_k (C_k \Sigma_{k|k-1} C'_k + R_k)^{-1} C_k \Sigma_{k|k-1} \quad (6.7 b)$$

Upper bounds on $\Sigma_{k|k}$, \mathcal{K}_k

Lemma 6.1

With A_k , C_k , Q_k , R_k and R_k^{-1} bounded above and $[A_k, C_k]$ uniform with respect to complete observability, then for the Kalman filter and one-step ahead predictor above, the error covariance $\Sigma_{k|k}$ and $\Sigma_{k|k-1}$ and gain \mathcal{K}_k are bounded above.

Proof†

Consider the suboptimal moving average estimate

$$\hat{x}_k^s = \psi_{k,k-t} M_{k,k-t}^{-1} \sum_{i=k-t}^k \psi'_{i,k-t} C'_i R_i^{-1} z_i$$

Expressing z_i as a linear combination of v_i , x_{k-t} and w_j for $k-t \leq j \leq i-1$ yields

$$\hat{x}_k^s - x_k = \sum_{i=k-t}^k (\alpha_{i,k} w_i + \beta_{i,k} v_i)$$

for some $\alpha_{i,k}$, $\beta_{i,k}$ bounded uniformly above for all $k-t \leq i \leq k$ from the boundedness assumptions.

As a consequence $\Sigma_{k|k}^s = E[(\hat{x}_k^s - x_k)(\hat{x}_k^s - x_k)']$ is bounded above. From optimality $\Sigma_{k|k} \leq \Sigma_{k|k}^s$ and the desired result follows. $\nabla \nabla \nabla$

An upper bound on $\Sigma_{k|k}^{-1}$

Lemma 6.2

With A_k , C_k , Q_k , R_k and R_k^{-1} bounded above, $[A_k, C_k]$ uniform with respect to complete observability, and $[A_k, D_k]$ uniform with respect to complete reachability for any D_k satisfying $D_k D'_k = Q_k$, then $\Sigma_{k|k}^{-1}$ exists for $k > N$ for some finite N and is bounded above.

† The proof in Jazwinski (1970) requires the non-singularity of A_k (and has errors as pointed out by later authors). The above proof avoids the difficulties encountered in Jazwinski (1970). An alternative proof using the suboptimal scheme of § 2 is also straightforward.

Proof†

First we claim that under the boundedness and observability conditions of the lemma,

$$\Sigma_{k|k} \geq \epsilon^2 \Sigma_{k|k-1} \tag{6.8}$$

for ϵ sufficiently small. This result follows from (6.7 b) if

$$S'_k C'_k (C_k S_k S'_k C'_k + R_k)^{-1} C_k S_k \leq (1 - \epsilon^2) I$$

where S_k satisfies $S_k S'_k = \Sigma_{k|k-1}$. Equivalently, the result follows using the Appendix Lemma A.3, if

$$C_k S_k S'_k C'_k (1 - \epsilon^2)^{-1} \leq C_k S_k S'_k C'_k + R_k$$

or more simply if $\epsilon^2 (1 - \epsilon^2)^{-1} C_k \Sigma_{k|k-1} C'_k \leq R_k$. For ϵ sufficiently small, this inequality holds under the assumption of the lemma and the result of Lemma 6.1, so that our claim above is established.

The inequality (6.8) leads by induction to the inequality

$$\Sigma_{k+1|k} \geq \bar{\Sigma}_{k+1}, \quad \bar{\Sigma}_{k+1} \triangleq \epsilon^2 A_k \bar{\Sigma}_k A'_k + Q_k, \quad \bar{\Sigma}_0 \triangleq \Sigma_{0|0} \tag{6.9}$$

from (6.7 a). Now the reachability condition of the lemma is equivalent to the condition that $[\epsilon A_k, D_k]$ be uniform with respect to complete reachability for arbitrary ϵ (see Appendix Lemma A.4). This latter condition ensures that $\bar{\Sigma}_k > \alpha I$ for some α and $k > N$ for some N and thus that $\Sigma_{k|k}$ is likewise bounded away from zero. ▽▽▽

Stability equivalence of predictor and filter

Lemma 6.3

With A_k, C_k, \mathcal{K}_k bounded above, then the filter (6.3) (optimal or otherwise) is asymptotically stable [exponentially asymptotically stable] if and only if the corresponding one-step-ahead predictor (6.4) is asymptotically stable [exponentially asymptotically stable].

Proof‡

This follows the pattern of proofs for Lemma 4.1 and 4.2. Corresponding to the system (3.3) and (4.1) we have

$$\xi_{j+1} = \tilde{A}_j \xi_j, \quad \zeta_{j+1} = \bar{A}_j \zeta_j \tag{6.10}$$

Corresponding to (4.2)-(4.4) we have the relationship

$$A_j \bar{A}_{j-1} = \tilde{A}_j A_{j-1}, \quad \xi_{j+1} = A_j \xi_j + A_j u_j, \quad \zeta_{j+1} = A_j \zeta_j - U_j \xi_{j+1}$$

for some \bar{u}_j, \bar{U}_j . Thus corresponding to Lemma 4.1 we have that if $\xi_{j+1} = A_j \xi_j$ for some j_0 , then the relationship holds for all $j \geq j_0$ and if (6.8 a) is initialized at $j_0 + 1$, there exists ζ_{j_0} such that $\xi_{j+1} = A_j \zeta_j$ for all $j \geq j_0$. Further details are easily deduced. ▽▽▽

† Less direct proofs requiring non-singularity of A_k are given elsewhere, see Jazwinski (1970, Chap. 7).

‡ In the optimal case, it is known that $\hat{x}_{k+1|k} = F_k \hat{x}_{k|k}$ from taking the appropriate conditional expectations on (6.1) and the lemma is trivially established.

We now have the machinery to establish exponential stability of the filter under certain conditions. Before doing this, we shall set up machinery for establishing non-exponential asymptotic stability.

The case of singular $\Sigma_{k|k}$

In the absence of the reachability condition of Lemma 6.2, $\Sigma_{k|k}$ could well be singular. The following lemma provides the key result we shall need.

Lemma 6.4

With quantities as defined above, the following identities hold :

$$(\Sigma_{k|k}^\# + \Sigma_{k|k}^\# \Sigma_{k|k} C'_k R_k^{-1} C_k \Sigma_{k|k} \Sigma_{k|k}^\#)^\# \geq \Sigma_{k|k} - \Sigma_{k|k} C'_k R_k^{-1} C_k \Sigma_{k|k} \tag{6.11}$$

$$\Sigma_{k-1|k-1} - \Sigma_{k-1|k-1} \bar{A}'_{k-1} (\Sigma_{k|k}^\# + C'_k R_k^{-1} C_k) \bar{A}_{k-1} \Sigma_{k-1|k-1} \geq 0 \tag{6.12}$$

Proof

By orthogonal transformation of the coordinate basis at each k , we can assume without loss of generality that $\Sigma_{k|k} = \Sigma_{k|k}^1 \dot{+} 0$ with $\Sigma_{k|k}^1$ non-singular. Let $C_k = [C_k^1 \ C_k^2]$ with C_k^1 having number of columns equal to the dimension of $\Sigma_{k|k}^1$. Then

$$\begin{aligned} & \Sigma_{k|k} - \Sigma_{k|k} C'_k R_k^{-1} C_k \Sigma_{k|k} \\ & \leq \Sigma_{k|k} - \Sigma_{k|k} C'_k [R_k + C_k^1 \Sigma_{k|k}^1 C_k^1]^{-1} C_k \Sigma_{k|k} \\ & = \begin{bmatrix} \Sigma_{k|k}^1 - \Sigma_{k|k}^1 C_k^1 [R_k + C_k^1 \Sigma_{k|k}^1 C_k^1]^{-1} C_k^1 \Sigma_{k|k}^1 & 0 \\ & 0 & 0 \end{bmatrix} \\ & = \begin{bmatrix} \{(\Sigma_{k|k}^1)^{-1} + C_k^1 R_k^{-1} C_k\}^{-1} & 0 \\ & 0 & 0 \end{bmatrix} \\ & = (\Sigma_{k|k}^\# + \Sigma_{k|k}^\# \Sigma_{k|k} C'_k R_k^{-1} C_k \Sigma_{k|k} \Sigma_{k|k}^\#)^\# \end{aligned}$$

To prove (6.12), we proceed as follows. From (6.5 a) and (6.6), we have

$$\Sigma_{k|k} - \Sigma_{k|k} C'_k R_k^{-1} C_k \Sigma_{k|k} \geq \bar{A}_{k-1} \Sigma_{k-1|k-1} \bar{A}'_{k-1}$$

and so, using (6.11),

$$(\Sigma_{k|k}^\# + \Sigma_{k|k}^\# \Sigma_{k|k} C'_k R_k^{-1} C_k \Sigma_{k|k} \Sigma_{k|k}^\#)^\# - \bar{A}_{k-1} \Sigma_{k-1|k-1} \Sigma_{k-1|k-1}^\# \Sigma_{k-1|k-1} \bar{A}'_{k-1} \geq 0$$

Now

$$\begin{aligned} & \mathcal{N}[\Sigma_{k|k}^\# + \Sigma_{k|k}^\# \Sigma_{k|k} C'_k R_k^{-1} C_k \Sigma_{k|k} \Sigma_{k|k}^\#] \\ & = \mathcal{N}[\Sigma_{k|k}] \subset \mathcal{N}[\Sigma_{k-1|k-1} \bar{A}'_{k-1}] \end{aligned}$$

with the containment following from (6.5 a). Also, we have trivially

$$\mathcal{N}[\Sigma_{k-1|k-1}^\#] \subset \mathcal{N}[\bar{A}_{k-1} \Sigma_{k-1|k-1}]$$

Lemma A.3 of the Appendix then yields

$$\Sigma_{k-1|k-1} - \Sigma_{k-1|k-1} \bar{A}'_{k-1} \times (\Sigma_{k|k} \# + \Sigma_{k|k} \# \Sigma_{k|k} C'_k R_k^{-1} C_k \Sigma_{k|k} \Sigma_{k|k} \#) \bar{A}_{k-1} \Sigma_{k-1|k-1} \geq 0$$

Using again the fact that $\mathcal{N}[\Sigma_{k|k}] \subset \mathcal{N}[\Sigma_{k-1|k-1} \bar{A}'_{k-1}]$, (6.12) follows. $\nabla \nabla \nabla$

Equation (6.12) is almost the final form we are seeking. In fact, we shall use the following reformulation of it.

Lemma 6.5

Suppose $\Sigma_{k|k}$ is non-singular† for some $k = K$. Then for $k > K$,

$$\mathcal{R}[\Sigma_{k|k}] \supset \mathcal{R}[\bar{\Psi}_{k,K}]$$

and if $\zeta_{k+1} = \bar{A}_k \zeta_k$ for $k \geq K$ with ζ_K arbitrary,

$$\zeta'_k \Sigma_{k|k} \# \zeta_k \geq \zeta'_k \bar{A}'_k (\Sigma_{k+1|k+1} \# + C'_{k+1} R_{k+1}^{-1} C_{k+1}) \bar{A}_k \zeta_k \tag{6.13}$$

Proof

Equation (6.5 a) yields $\Sigma_{k|k} = \bar{\Psi}_{k,K} \Sigma_{K|K} \bar{\Psi}'_{k,0} +$ non-negative terms, and the range inclusion is satisfied. Then (6.13) follows from (6.12) with k replaced by $k + 1$, and the fact that $\zeta_k \in \mathcal{R}[\bar{\Psi}_{k,K}]$.

Stability results

Kalman filter stability theorem. With A_k, C_k, R_k and R_k^{-1} bounded above, and with $[A_k, C_k]$ uniform with respect to complete observability, the optimal linear filter and one-step ahead predictor are asymptotically stable. If in addition $\Sigma_{k|k} \#$ is bounded above for all $k > N$ for some N (see Lemma 6.2), then exponential asymptotic stability follows.

Proof

Let K be an integer for which $\Sigma_{k|k}$ is non-singular. If no such integer exists in $[0, \infty)$, extend the definition of the signal model backwards by one unit so that $F'_{-1} = 0, C_{-1} = 0$, the signal model state variance at time -1 is I , and $Q_{-1} = P_0$. Then $\Sigma_{-1|-1} = I$, and the covariance and gain matrix sequences are unaltered on $[0, \infty)$.

Associate with the closed-loop system $\zeta_{k+1} = \bar{A}_k \zeta_k$ the tentative Lyapunov function defined in $[K, \infty)$

$$V_k = \zeta'_k \Sigma_{k|k} \# \zeta_k \tag{6.14}$$

Using Lemma 6.5,

$$V_k - V_{k+1} = \zeta'_k [\Sigma_{k|k} \# - \bar{A}'_k \Sigma_{k+1|k+1} \# \bar{A}_k] \zeta_k \geq \zeta'_{k+1} C'_{k+1} R_{k+1}^{-1} C_{k+1} \zeta_{k+1}$$

whence

$$V_k - V_{k+t} \geq \zeta'_{k+1} \left(\sum_{i=k+1}^{k+t+1} \bar{\Psi}'_{i,k+1} C'_i R_i^{-1} C_i \bar{\Psi}_{i,k+1} \right) \zeta_{k+1}$$

† This could be guaranteed by a reachability assumption, or by having $\Sigma_{0|0}$ non-singular. It can also follow if some states are reachable, and $\Sigma_{0|0}$ is singular but not zero.

Lemma A.2 ensures the observability of $[\bar{A}_k, C_k]$ provided that $[I - \mathcal{K}_k C_k]^{-1}$ is bounded. This bound is available since

$$\begin{aligned} \det [I - \mathcal{K}_k C_k] &= \det [I - \Sigma_{k|k-1} C'_k (C_k \Sigma_{k|k-1} C'_k + R_k)^{-1} C_k] \\ &= \det [I - (C_k \Sigma_{k|k-1} C'_k + R_k)^{-1} C_k \Sigma_{k|k-1} C'_k] \\ &= \frac{\det R_k}{\det (C_k \Sigma_{k|k-1} C'_k + R_k)} \\ &> \alpha \end{aligned}$$

for some constant α . Observability of $[\bar{A}_k, C_k]$ ensures that

$$V_k - V_{k+t} \geq \beta \zeta'_{k+1} \zeta_{k+1}$$

for some $\beta > 0$. It is trivial to conclude that $\zeta_k \rightarrow 0$.

In case $\Sigma_{k|k}^\#$ is bounded above, we see that V_k decays exponentially in the following way.

$$V_k - V_{k+t} \geq \beta \zeta'_{k+1} \zeta_{k+1} \geq \gamma V_{k+1} \geq \gamma V_{k+t}$$

where the second inequality comes from the bound on $\Sigma_{k|k}^\#$, and the third from the monotone character of V_k . Of course, γ is a positive constant. This inequality gives $V_{k+t} \leq (\gamma + 1)^{-1} V_k$, which yields the exponential convergence of V_k . Then the first inequality yields exponential convergence of ζ_k . ▽▽▽

Remarks

(1) This theorem generalizes an earlier result of Anderson (1971), in that it relaxes the requirement that A_k^{-1} be bounded above.

(2) The theorem can be further generalized by relaxing the observability requirement to a detectability one. This will appear in a companion paper which exploits a number of the lemmas in this paper. See also Hager and Hurowitz (1976) where a dualization of control results is exploited, but full details are given only for the time-invariant case.

Optimal regulator stability

It is known (Kalman and Engla 1965, Caines and Mayne 1970), that with F_k, G_k bounded and $[F_k, G_k]$ uniform with respect to complete reachability†, the solution of the Riccati equation for $k \leq T$

$$P_{k+1, T} = F'_{k|} [P_{k, T} - P_{k, T} G_k (R_k + G'_k P_{k, T} G_k)^{-1} G'_k P_{k, T}] F_k + Q_k$$

for $P_{T, T}$ given is bounded uniformly in T and

$$\bar{P}_k = \lim_{T \rightarrow \infty} P_{k, T}$$

exists. Moreover, for the system

$$x_{k+1} = F_k x_k + G_k u_k$$

† Actually the stronger condition $[F_k, G_k]$ uniformly completely reachable is used, but the relaxation noted here is trivially established using the method of proof in Caines and Mayne (1970).

the control law which minimizes the index

$$V_k = \sum_{i=0}^{\infty} (x'_i Q_i x_i + u'_i R_i u_i)$$

is

$$u_i^* = -K_k x_k, \quad K_k = -[R_k + G'_k \bar{P}_{k+1} G_k]^{-1} G'_k \bar{P}_{k+1} F_k$$

where now \bar{P}_{k+1} satisfies the Riccati equation above. Here, we show that if in addition, $[F_k, D_k]$ is uniform with respect to complete observability for any $D'_k D_k = Q_k$, and Q_k is bounded, then the closed-loop system

$$x_{k+1} = \bar{F}_k x_k, \quad \bar{F}_k = F_k - G_k K_k$$

is exponentially asymptotically stable.

Consider the tentative Lyapunov function $V_k = x'_k \bar{P}_k x_k$. Then

$$V_k - V_{k+1} = x'_k Q_k x_k + x'_k [(\bar{P}_k - Q_k) - \bar{F}'_k \bar{P}_{k+1} \bar{F}_k] x_k$$

The second term is now studied. From the Riccati equation for \bar{P}_k we have

$$\begin{aligned} (\bar{P}_k - Q_k) - \bar{F}'_k \bar{P}_{k+1} \bar{F}_k &= \bar{F}'_k \bar{P}_{k+1} G_k (G'_k \bar{P}_{k+1} G_k + R_k)^{-1} G'_k \bar{P}_{k+1} F_k \\ &= F'_k [I - \bar{P}_{k+1} G_k \bar{R}_k^{-1} G'_k] [\bar{P}_{k+1} G_k \bar{P}_k^{-1} G'_k \bar{P}_{k+1}] F_k \\ &= F'_k \bar{P}_{k+1} G_k \bar{R}_k^{-1} [\bar{R}_k - G'_k \bar{P}_{k+1} G_k] \bar{R}_k^{-1} G'_k \bar{P}_{k+1} F_k \\ &= (F'_k \bar{P}_{k+1} G_k \bar{R}_k^{-1}) R_k (\bar{R}_k^{-1} G'_k \bar{P}_{k+1} F_k) \end{aligned}$$

where \bar{R}_k denotes $G'_k \bar{P}_{k+1} G_k + R_k$. Thus, with $Q_k = D'_k D_k$, and

$$\bar{D}'_k = [D'_k \quad F'_k \bar{P}_{k+1} G_k \bar{P}_k^{-1} R_k^{1/2}]$$

then $\bar{F}_k = F_k - G_k R_k^{-1/2} \bar{D}_k$, $V_k - V_{k+1} = x'_k \bar{D}'_k \bar{D}_k x_k$ and with $\bar{F}_{i,k}$ the transition matrix of the closed-loop system, then

$$V_k - V_{k+t} = x'_k \left(\sum_{i=k}^{k+t} \bar{F}'_{i,k} \bar{D}'_k \bar{D}_k \bar{F}_{i,k} \right) x_k$$

Now $[F_k, D_k]$ uniform with respect to complete observability implies that $[F_k, \bar{D}_k]$ has the same property, and by the Appendix Lemma A.2, that $[\bar{F}_k, \bar{D}_k]$ also has the same property. Thus for t sufficiently large,

$$V_k - V_{k+t} \geq \beta x'_k x_k$$

for some $\beta > 0$. Straightforward arguments under the boundedness assumptions give that $V_k \rightarrow 0$ exponentially as $k \rightarrow \infty$ and then so does x_k as claimed.

The above stability result derivation is relatively direct compared with earlier treatments (Jalman and Engla 1965, Caines and Mayne 1970).

It might be thought that duality between control and filtering could be explored usefully to avoid duplicating some results—as is done in Hager and Hurowitz (1976). However there are limitations to this in the time-varying case since the time interval $[0, \infty)$ for filtering is dualized as $(-\infty, 0]$ for control. The study of intervals $(-\infty, \infty)$ has its own limitations including the imposition of signal model stability for the filtering problems.

7. Conclusions

In this paper we have demonstrated that when transition matrices approach singular matrices or are singular, the associated stability theory need not be any less elegant. On the contrary, confrontation rather than side-stepping of the apparent difficulties leads to a deeper insight into the underlying structure of the stability questions.

The techniques and results of this paper now form a solid basis to generate more general stability results in time-varying linear estimation and control. Current research is aimed at relaxing the controllability and observability assumptions to stabilizability and detectability ones, thereby giving a time-varying generalization of the most recent time-invariant system results and fleshing out the theoretical possibilities suggested in Hager and Hurowitz (1976).

Appendix

Lemma A.1 (invariance of observability under bounded feedback)

$[F_j, H_j]$ is uniform with respect to complete observability if and only if $[F_j + K_j H_j, H_j]$ has the same property for arbitrary bounded K_j .

Proof

With $\mathcal{H}'_{j+t,j} = [H'_j \phi'_{j+1,j} H'_{j+1} \dots \phi'_{j+t,j} H'_{j+t}]'$, the observability Gramian $M_{j+t,j}$ of $[F_j, H_j]$ is

$$\mathcal{H}'_{j+t,j} \mathcal{H}_{j+t,j}$$

and that of $[F_j + K_j H_j, H_j]$ is

$$\mathcal{H}'_{j+t,j} X'_j X_j \mathcal{H}_{j+t,j}$$

where X_j is a bounded lower triangular matrix with 1's on the diagonal, as shown by a simple inductive argument. It follows easily that one observability Gramian is bounded away from zero if and only if the other is.

The above lemma shows that an (optimum) one-step ahead predictor has the same observability property as the signal model. The following lemma does this for the filter.

Lemma A.2

$[F_j, H_j]$ is uniform with respect to complete observability if and only if for any sequence K_j such that $(I + H_j K_j)^{-1}$ is bounded, $[(I + K_{j+1} H_{j+1}) F_j, H_j]$ is uniform with respect to complete observability.

Proof

With $\mathcal{H}'_{j+t,j}$ as in the proof of Lemma 1, one finds that the observability Gramian of $[(I + K_{j+1} H_{j+1}) F_j, H_j]$ is

$$\mathcal{H}'_{j+t,j} X'_j X_j \mathcal{H}_{j+t,j}$$

where X_j is a bounded lower block triangular matrix with diagonal blocks $I, I + H_{j+1} K_{j+1}, I + H_{j+2} K_{j+2} \dots$. The claim then follows easily.

Lemma A.3

Suppose that $A = A' \geq 0$, $B = B' \geq 0$, $\mathcal{N}[A] \subset \mathcal{N}[C']$ and $\mathcal{N}[B] \subset \mathcal{N}[C]$. With $X^\#$ denoting the pseudo-inverse of X ,

$$A - CBC' \geq 0 \Leftrightarrow B^\# - C'A^\#C \geq 0$$

Proof

Verify that

$$\begin{bmatrix} A - CBC' & 0 \\ 0 & B^\# \end{bmatrix} = \begin{bmatrix} I & -CB \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ C' & B^\# \end{bmatrix} \begin{bmatrix} I & 0 \\ -BC' & I \end{bmatrix}$$

where

$$A - CBC' \geq 0 \Leftrightarrow \begin{bmatrix} A & C \\ C' & B^\# \end{bmatrix} \geq 0 \Leftrightarrow B^\# - C'A^\#C \geq 0$$

Lemma A.4

Let $[F_k, G_k]$ be uniform with respect to complete reachability. The $[\beta F_k, G_k]$ for any $\beta > 0$, has the same property.

Proof

From the definition (2.2) it is immediate that

$$0 < \beta_1 Y_{k,k-s} \leq Y_{k,k-s}^\beta \leq \beta_2 Y_{k,k-s} < \infty$$

for some β_1, β_2 since

$$Y_{k,k-2}^\beta = \sum_{i=k-s}^k \beta^{2(k-i)} \phi_{k+1,i+1} G_i G_i' \phi_{k+1,i+1}'$$

▽▽▽

REFERENCES

- ANDERSON, B. D. O., 1971, *J. Franklin Inst.*, **291**.
 CAINES, P. E., and MAYNE, D. Q., 1970, *Int. J. Control*, **12**, 785.
 DEYST, J. J., JR., and PRICE, C. F., 1968, *I.E.E.E. Trans. autom. Control*, **13**, 702.
 HAGER, W., and HUROWITZ, L., 1976, *SIAM J. Control Optim.*, **14**, 295.
 IKEDA, M., MAEDA, H., and KODAMA, S., 1975, *SIAM J. Control*, **13**, 304.
 JAZWINSKI, A. H., 1970, *Stochastic Processes and Filtering Theory* (New York : Academic Press).
 KALMAN, R., and ENGLA, T., 1965, *A User's Manual for ASP, C*, RIAS Rep. NASA, Baltimore, Md.
 KLEINMAN, D. L., 1970, *I.E.E.E. Trans. autom. Control*, **15**, 692 ; 1974, *Ibid.*, **19**, 252.
 KWON, W. H., and PEARSON, A. E., 1977, *I.E.E.E. Trans. autom. Control*, **22**, 838 ; 1978, *Ibid.*, **23**, 479.
 MCGARTY, T. P., 1974, *Stochastic Systems and State Estimation* (New York : Wiley).