Solutions to Wiener Filtering and Stationary LQG Problem via H_2 Control Theory - Part II : Discrete - Time System

Tohru KATAYAMA*
(Received June 20, 1990)

Abstract

This paper derives the solutions to the Wiener filtering and stationary LQG problem for a discrete-time system by applying the state-space techniques developed for H_2/H_∞ optimal controls. As mathematical preliminaries, we collect useful operations for the transfer function matrices. We also provide a new proof for the inner-outer factorization algorithm that appears in the discrete-time H_2 optimization.

1. Introduction

This paper, a continuation of Part I [7], considers state-space solutions to the discrete-time Wiener filtering and stationary LQG problem via the H_2 optimization technique. It has been shown [2]—[5] that the general discrete-time H_α ($\alpha=2$ or ∞) control problem is reduced to a model matching problem. The objective is to find a stable transfer function $Q(z)$ such that

$$J=\| T_1(z)-T_2(z)Q(z)T_3(z)\|_\alpha=\text{minimum} \tag{1.1}$$

where $T_1(z)$, $T_2(z)$, $T_3(z)$ are stable matrix functions.

For $\alpha=2$, the H_2 norm of a transfer function matrix $G(z)$ that is analytic on the unit circle is defined by

$$\|G(z)\|_2^2=\frac{1}{2\pi j}\int_{|z|=1}\text{tr}[G^*(z)G(z)]\frac{dz}{z} \tag{1.2}$$

where the asterisk denotes the conjugate transpose

$$G^*(z)=G^T(z^{-1}) \tag{1.3}$$

Moreover, for $\alpha=2$, the solution is particularly simple. In fact, let $T_2(z)$ and $T_3(z)$ respectively be factored as (see Section 2.2)

$T_2(z)=T_{2i}(z)T_{2o}(z)$, T_{2i} : inner, T_{2o} : outer

* Department of Applied Mathematics and Physics, Faculty of Engineering, Kyoto University, Kyoto 606, Japan.

and

$T_3(z) = T_{3co}(z)\, T_{3ci}(z)$, T_{3ci} : co-inner, T_{3co} : co-outer

Then the optimal solution $Q(z)$ is given by [11]

$$Q(z) = T_{2o}^{-1}(z)\,[\,T_{2i}^{*}(z)\,T_1(z)\,T_{3ci}^{*}(z)\,]_+\,T_{3co}^{-1}(z) \tag{1.4}$$

where $[\cdot]_+$ denotes the stable part of a matix function by partial fraction expansion.

We apply the formula (1.4) and state-space techniques to derive solutions to the Wiener filtering and stationary LQG problem. Unlike the continuous-time case, $[\cdot]_+$ may be defined in two ways depending on whether the constant terms are included in it or not. We therefore present strictly causal and causal solutions to both the Wiener filtering and stationary LQG problems.

The organization of this paper is as follows. In Section 2, we introduce discrete-time transfer function matrices, collect useful operations for the transfer function matrices, and present an inner-outer factorization algorithm for a stable function. Section 3 treats the discrete-time Wiener filtering problem. We first present a classical solution based on the spectal factorization and additive decomposition, and then derive the strictly causal and causal solutions using the state-space technique. Section 4 is concerned with the stationary LQG problem, for which the strictly causal and causal solutions are derived by using inner-outer and co-inner-outer factorizations of the transfer matrices appearing in the model matching problem. Concluding remarks are given in Section 5. Appendices include proofs.

2. Mathematical Preliminaries

We present some preliminary results for the transfer matrices and an algorithm of inner-outer factorization for discrete-time systems.

2. 1 Transfer Functions

The state-space realization of a real-rational transfer function is represented by

$$G(z) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := D + C(zI-A)^{-1}B \tag{2.1}$$

where A, B, C, D are constant matrices of dimensions $n\times n$, $n\times m$, $p\times n$, $p\times m$, respectively. Let RH_+^∞ be the class of stable, strictly causal transfer functions of the form

$$G(z) = \sum_{\ell=0}^{\infty} CA^{\ell}Bz^{-(\ell+1)},\ |z| > \max_{1\le i\le n} |\lambda_i(A)| \tag{2.2}$$

where A is stable, and $\lambda_i(A)$ denotes the eigenvalues of A. The complementary space RH_-^∞ is the class of antistable, proper transfer functions represented by

$$\begin{aligned} G(z) &= D - z^{-1}C(z^{-1}I - A^{-1})^{-1}A^{-1}B \\ &= D - \sum_{\ell=0}^{\infty} CA^{-\ell}Bz^{\ell-1},\ |z| < \min_{1\le i\le n} |\lambda_i(A)| \end{aligned} \tag{2.3}$$

where we assume that the constant term belongs to RH_-^∞. We see that if $G(z) \in RH_-^\infty$, then A must be invertibe.

It should be noted that although the constant terms are included in RH_-^∞, it is possible to define RH_+^∞ as the class of stable, proper (not necessarily strictly causal) transfer functions. Then, the complementary space RH_-^∞ must be the class of antistable transfer functions with strictly positive powers of z.

The following are useful formulae for the operations on transfer function matrices [2], [4].

(a) For a nonsingular T,

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{c|c} T^{-1}AT & T^{-1}B \\ \hline CT & D \end{array}\right] \tag{2.4}$$

(b) Suppose that $G(z)$ is square and D is nonsingular. Then

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]^{-1} = \left[\begin{array}{c|c} A-BD^{-1}C & BD^{-1} \\ \hline -D^{-1}C & D^{-1} \end{array}\right] \tag{2.5}$$

(c) A cascade of two transfer matrices is given by

$$\left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right]\left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right] = \left[\begin{array}{cc|c} A_1 & B_1C_2 & B_1D_2 \\ 0 & A_2 & B_2 \\ \hline C_1 & D_1C_2 & D_1D_2 \end{array}\right]$$

$$= \left[\begin{array}{cc|c} A_2 & 0 & B_2 \\ B_1C_2 & A_1 & B_1D_2 \\ \hline D_1C_2 & C_1 & D_1C_2 \end{array}\right] \tag{2.6}$$

(d) Suppose that A is stable and nonsingular. Then $G^*(z) := G^T(z^{-1})$ is expressed as [5]

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]^* = \left[\begin{array}{c|c} A^{-T} & -A^{-T}C^T \\ \hline B^TA^{-T} & D^T-B^TA^{-T}C^T \end{array}\right] \tag{2.7}$$

It should be noted that the properties (a), (b), (c) are the same as those of the continuous-time counterparts. For (d), since A^{-1} exists, we have

$$\begin{aligned} G^*(z) &= D^T + B^T(z^{-1}I - A^T)^{-1}C^T \\ &= D^T - zB^T(zI - A^{-T})^{-1}A^{-T}C^T \\ &= D^T - B^TA^{-T} - B^TA^{-T}(zI - A^{-T})^{-1}A^{-T}C^T \end{aligned}$$

2. 2 Factorization Result

We assume that A is stable and nonsingular, so that $G^*(z)$ is well defined. Thus we have

$$G^*(z)G(z) = \left[\begin{array}{c|c} A^{-T} & -A^{-T}C^T \\ \hline B^TA^{-T} & D^T-B^TA^{-T}C^T \end{array}\right]\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

$$= \left[\begin{array}{cc|c} A & 0 & B \\ -A^{-T}C^TC & A^{-T} & -A^{-T}C^TD \\ \hline \Gamma^TC & B^TA^{-T} & \Delta \end{array}\right] \tag{2.8}$$

where $\Gamma := D - CA^{-1}B$ and $\Delta := \Gamma^T D$. If $G^*(z)G(z) = I_m$, then $G(z)$ is inner. Also, if $G(z)G^*(z) = I_p$, then $G(z)$ is co-inner.

In the following, we present an algorithm for the inner-outer factorization for discrete-time stable transfer matrices. It should be noted that the co-inner-outer factorization of $G(z)$ is easily derived from inner-outer factorization of $G^T(z)$.

Assume that D^TD and Δ are nonsingular. It follows from (2.8) that

$$[G^*(z)G(z)]^{-1} = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{D} \end{array}\right] \tag{2.9}$$

where

$$\tilde{A} = \begin{bmatrix} A - B\Delta^{-1}\Gamma^TC & -B\Delta^{-1}B^TA^{-T} \\ -A^{-T}C^TC + A^{-T}C^TD\Delta^{-1}\Gamma^TC & A^{-T} + A^{-T}C^TD\Delta^{-1}B^TA^{-T} \end{bmatrix} \tag{2.10a}$$

$$\tilde{B} = \begin{bmatrix} B \\ -A^{-T}C^TD \end{bmatrix} \Delta^{-1} \tag{2.10b}$$

$$\tilde{C} = -\Delta^{-1}[\Gamma^TC \quad B^TA^{-T}] \tag{2.10c}$$

$$\tilde{D} = \Delta^{-1} \tag{2.10d}$$

For convenience, we define

$$\begin{aligned} E &:= A - B(D^TD)^{-1}D^TC \\ \Sigma &:= B(D^TD)^{-1}B^T \\ \Pi &:= C^T[I - D(D^TD)^{-1}D^T]C \end{aligned} \tag{2.11}$$

where it may be noted that Σ, Π are nonnegative definite.

Lemma 2.1

Suppose that E in nonsingular. Then $\tilde{A}$ is expressed as

$$\tilde{A} = \begin{bmatrix} E + \Sigma E^{-T}\Pi & -\Sigma E^{-T} \\ -E^{-T}\Pi & E^{-T} \end{bmatrix} := \mathscr{H} \tag{2.12}$$

Proof

A proof is deferred in Appendix A. □

Since $J^{-1}\mathscr{H}^TJ = \mathscr{H}^{-1}$ for $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$, $\mathscr{H}$ of (2.12) is symplectic. Suppose that $\mathscr{H}$ has no eigenvalues on the unit circle. Then $\mathscr{H}$ has n eigenvalues inside the unit disk and n outside the unit disk.

Let $U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}$ be an orthogonal matrix that transforms $\mathscr{H}$ of (2.12) into a real Schur from:

$$\begin{bmatrix} E+\Sigma E^{-T}\Pi & -\Sigma E^{-T} \\ -E^{-T}\Pi & E^{-T} \end{bmatrix}\begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}=\begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}\begin{bmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{bmatrix}$$

where S_{11} is stable and S_{22} is antistable.

Lemma 2.2

Suppose that (E, B) is stabilizable and $\mathscr{H}$ has no eigenvalues on the unit circle. Then, U_{11} is invertible, and $X := U_{21}U_{11}^{-1}$ is nonnegative definite and satisfies the discrete-time ARE

$$X=E^TXE-E^TXB(D^TD+B^TXB)^{-1}B^TXE+\Pi \tag{2.13}$$

Moreover,

$$E_c := E-B(D^TD+B^TXB)^{-1}B^TXE \tag{2.14a}$$

$$=E-\Sigma E^{-T}(X-\Pi) \tag{2.14b}$$

$$=(I+\Sigma X)^{-1}E \tag{2.14c}$$

is stable.

Proof

A proof is found in [8]–[10]. □

The solution X given by Lemma 2.2 is referred to as the stabilizing solution, which is expressed as $X=\mathrm{Ric}\mathscr{H}$. Define

$$K := -(D^TD+B^TXB)^{-1}(B^TXA+D^TC) \tag{2.15}$$

Then, it is easy to see from (2.11) and (2.14) that E_c is expressed as

$$E_c=A+BK \tag{2.16}$$

Lemma 2.3

By the basis change $T=\begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$, (2.9) is reduced to

$$[G^*(z)G(z)]^{-1}=\left[\begin{array}{c|c} T^{-1}\tilde{A}T & T^{-1}\tilde{B} \\ \hline \tilde{C}T & \tilde{D} \end{array}\right]$$

$$=\left[\begin{array}{cc|c} E_c & -\Sigma E^{-T} & B\Delta^{-1} \\ 0 & E_c^{-T} & -A^{-T}(A^TXB+C^TD)\Delta^{-1} \\ \hline K & -\Delta^{-1}B^TA & \Delta^{-1} \end{array}\right] \tag{2.17}$$

Proof

A proof is given in Appendix B. □

By taking the inverse of (2.17), an alternative representation of $G^*(z)G(z)$ is obtained as

$$G^*(z)G(z)=\left[\begin{array}{cc|c} A & 0 & B \\ A^{-T}(A^TXB+C^TD)K & A^{-T} & -A^{-T}(A^TXB+C^TD) \\ \hline -\Delta K & B^TA^{-T} & \Delta \end{array}\right] \tag{2.18}$$

We observe from (2.8) that (2.18) has a factorization

$$G^*(z)G(z)=\left[\begin{array}{c|c} A^{-T} & -A^{-T}(A^TXB+C^TD) \\ \hline B^TA^{-T} & \Delta \end{array}\right]\left[\begin{array}{c|c} A & B \\ \hline -K & I_m \end{array}\right]$$

$$=\left[\begin{array}{c|c} A & B \\ \hline -K & I_m \end{array}\right]^*(R+B^TXB)\left[\begin{array}{c|c} A & B \\ \hline -K & I_m \end{array}\right] \quad (2.19)$$

The following theorem gives a state-space realization for an inner-outer factorization of a stable transfer function.

Theorem 2.1 [5]

Let $G(z):=\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ be a minimal realization with A stable. Let A and $\Delta=(D^T-B^TA^{-T}C^T)D$ be invertible and $D^TD>0$. Let $X=\text{Ric}\ \mathscr{H}$, and K be given by (2.15). Then an inner-outer factorization of $G(z)$ is

$$G(z)=G_i(z)G_o(z) \quad (2.20)$$

where the inner and outer functions are respectively given by

$$G_i(z)=\left[\begin{array}{c|c} A+BK & B \\ \hline C+DK & D \end{array}\right]V^{-1} \quad (2.21)$$

$$G_o(z)=V\left[\begin{array}{c|c} A & B \\ \hline -K & I_m \end{array}\right] \quad (2.22)$$

and where

$$V=(D^TD+B^TXB)^{1/2} \quad (2.23)$$

Proof

A proof is given in Appendix C. □

It may be noted that the derivation of the inner-outer factorization of (2.20) is different from that of [5]. In fact, the above factorization algorithm is derived by noting that $T^{-1}\tilde{A}T$ has a real Schur form as in (2.17) and that the block diagonal elements of the A-matrix of $G^*(z)G(z)$ are not affected under the basis change by $T=\begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$. Thus, even if A is not stable, we can obtain a canonical spectral factor $G_o(z)$ satisfying $G_o^*(z)G_o(z)=G^*(z)G(z)$ as noted in [7].

3. Discrete-Time Wiener Filtering

A discrete-time multivariable Wiener filtering problem is treated in this section. First, we formulate the problem and present the solution based on the spectral factorization and additive decomposition [6]. Then, we derive the state-space solution by converting the Wiener filtering problem into a model matching problem.

3. 1 Wiener Filtering

Suppose that we are given the observed signal $y(t)$ which is the sum of the desired signal $\theta(t)$ and the noise $\nu(t)$, namely

$$y(t)=\theta(t)+\nu(t) \tag{3.1}$$

where $y(t)$, $\theta(t)$, $\nu(t)$, $t=0, \pm1, \cdots$, are p-dimensional zero mean second-order jointly sationary processes. We assume that $\theta(t)$ and $\nu(t)$ are uncorrelated. Let the covariance functions of $y(t)$, $\theta(t)$, $\nu(t)$ be given by $R_{yy}(\ell)$, $R_{\theta\theta}(\ell)$, $R_{\nu\nu}(\ell)$, respectively. Then it follows from (3.1) that

$$R_{yy}(\ell)=R_{\theta\theta}(\ell)+R_{\nu\nu}(\ell),\ \ell=0, \pm1, \cdots$$

where $R_{\alpha\beta}(\ell) := E\{\alpha(t+\ell)\beta^T(t)\}$, and $E\{\cdot\}$ denotes the mathematical expectation. Thus the spectral density matrices satisfy

$$S_{yy}(z)=S_{\theta\theta}(z)+S_{\nu\nu}(z) \tag{3.2}$$

The Wiener filtering problem is to find the least-squares (LS) estimate of the desired signal $\theta(t)$ based on the past observations $Y^t := \{y(\ell), \ell=t-1, t-2, \cdots\}$. This is called the one-step prediction problem, because Y^t includes the data up to time $t-1$ to estimate the desired signal $\theta(t)$. If we denote the LS estimate by $\hat{\theta}(t)$, then the problem is to design the causal filter $W(z)=\Sigma_{\ell=0}^{\infty}W_\ell z^{-\ell}$ such that

$$J=E\{\|\theta(t)-\hat{\theta}(t)\|^2\}=\text{minimum}$$

where the LS estimate is given by

$$\hat{\theta}(t)=\sum_{\ell=0}^{\infty}W_\ell y(t-\ell) \tag{3.3}$$

It may be noted that $W_0\equiv0$ since $W(z)\in RH_+^\infty$. But, if we include $y(t)$ in the data set Y^t, then we have the optimal filtering problem [1]. We also derive the solution for the case of $W_0\neq0$.

Suppose that the spectral density function $S_{yy}(z)$ is analytic on the unit circle, and has a canonical factorization

$$S_{yy}(z)=\Phi_y(z)\Phi_y^T(z^{-1})$$

where $\Phi_y(z)$ is a $p\times p$ outer function, namely, $\Phi_y(z)$, $\Phi_y^{-1}(z)$ are analytic in $|z|\geq1$. It is well known [1], [6] that the transfer function of the optimal filter is given by

$$W(z)=[S_{\theta\theta}(z)\Phi_y^{-T}(z^{-1})]_+\Phi_y^{-1}(z) \tag{3.4}$$

where $[S_{\theta\theta}(z)\Phi_y^{-T}(z^{-1})]_+$ is the matrix function belonging to RH_+^∞ by partial fraction expansion.

We now assume that the observation noise $\nu(t)$ is a white noise with $S_{\nu\nu}(z)=R$ and that $S_{\theta\theta}(z)\to0$ as $z\to\infty$. We see from (3.2) that

$$\begin{aligned} S_{\theta\theta}(z)\Phi_y^{-T}(z^{-1}) &= [S_{yy}(z)-R]\Phi_y^{-T}(z^{-1}) \\ &= \Phi_y(z)-R\Phi_y^{-T}(z^{-1}) \end{aligned} \tag{3.5}$$

where $\Phi_y(z)$ is stable and $R\Phi_y^{-T}(z^{-1})$ is antistable, and $S_{yy}(z)\to R$ as $z\to\infty$. Unlike the continuous-time case, the spectral factor $\Phi_y(z)$ does not tend to $R^{1/2}$ as $z\to\infty$, but

$$\Phi_y(z)\to\Lambda,\ \Phi_y^T(z^{-1})\to\Lambda^{-1}R,\ z\to\infty$$

where $\Lambda := \Phi_y(\infty)$, and Λ is assumed to be nonsingular. Subtracting the constant term

Λ from $\Phi_y(z)$, we get an additive decomposition

$$S_{\theta\theta}(z)\Phi_y^{-T}(z^{-1}) = [\Phi_y(z) - \Lambda] + [\Lambda - R\Phi_y^{-T}(z^{-1})]$$

where $\Phi_y(z) - \Lambda \in RH_+^\infty$ and $\Lambda - R\Phi_y^{-T}(z^{-1}) \in RH_-^\infty$. Thus we have

$$[S_{\theta\theta}(z)\Phi_y^{-T}(z^{-1})]_+ = \Phi_y(z) - \Lambda \tag{3.6}$$

Substituting (3.6) into (3.4) yields the optimal transfer function

$$W(z) = I_p - \Lambda\Phi_y^{-1}(z) \tag{3.7}$$

We see that $W(z)$ is strictly proper, so that $W_0 \equiv 0$. This expression is the same as that of the continuous-time Wiener filter [7].

It should be noted that if we include the current observed signal $y(t)$ in the data set Y^t, then we should modify the additive decomposition so that $[S_{\theta\theta}(z)\Phi_y^{-T}(z^{-1})]_+$ contains all the constant terms in the r.h.s. of (3.5). In fact, from (2.3), the contant term in the antistable function $\Phi_y^{-T}(z^{-1})$ is obtained by

$$\lim_{z\to 0}\Phi_y^{-T}(z^{-1}) = \Phi_y^{-T}(\infty) = \Lambda^{-T}$$

Thus

$$[\Phi_{\theta\theta}(z)\Phi_y^{-T}(z^{-1})]_+ = \Phi_y(z) - R\Lambda^{-T} \tag{3.8}$$

It follows from (3.4) and (3.8) that the optimal transfer function of the filtering problem is given by

$$\begin{aligned} W(z) &= [\Phi_y(z) - R\Lambda^{-T}]\Phi_y^{-1}(z) \\ &= I_p - R\Lambda^{-T}\Phi_y^{-1}(z) \end{aligned} \tag{3.9}$$

It may be noted that $W_0 = W(\infty) = I_p - R(\Lambda\Lambda^T)^{-1} \neq 0$.

In the next section, we derive the state-space solution to the discrete-time Wiener filtering problem for the case where the spectral density function $S_{\theta\theta}(z)$ is rational.

3.2 Model Matching Problem

Consider the case where the desired signal $\theta(t)$ has a rational spectral density, so that we can assume that $\theta(t)$ is generated by a minimal state-space model

$$x(t+1) = Ax(t) + G\xi(t) \tag{3.10}$$

$$\theta(t) = Cx(t) \tag{3.11}$$

where $x(t)$ is the $n\times 1$ state vector, $\xi(t)$ is the $q\times 1$ white noise vector with mean zero and covariance matrix I_q, and A, G, C are $n\times n$, $n\times q$, $p\times n$ constant matrices, respectively. We also assume that A is stable and nonsingular.

Define $\Phi(z) := (zI - A)^{-1}$. Then we see that in z-domain,

$$\begin{aligned} e(z) &= \theta(z) - \hat{\theta}(z) = C\Phi(z)G\xi(z) - \hat{\theta}(z) \\ y(z) &= \theta(z) + \nu(z) = C\Phi(z)G\xi(z) + L\eta(z) \\ \hat{\theta}(z) &= W(z)y(z) \end{aligned} \tag{3.12}$$

where $\nu(z) = L\eta(z)$, and η is a white noise vector with $N(0, I_p)$, and L is nonsingular.

According to the general framework of the model matching problem, (3.12) reduces to (see Fig. 1)

$$\begin{bmatrix} e \\ y \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} v \\ \hat{\theta} \end{bmatrix}$$

$$= \begin{bmatrix} C\Phi(z)G & 0 & -I_p \\ C\Phi(z)G & L & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \\ \hat{\theta} \end{bmatrix} \tag{3.13}$$

and

$$\hat{\theta} = W(z)y \tag{3.14}$$

where $v := \begin{bmatrix} \xi \\ \eta \end{bmatrix}$. Also, the error signal is given by

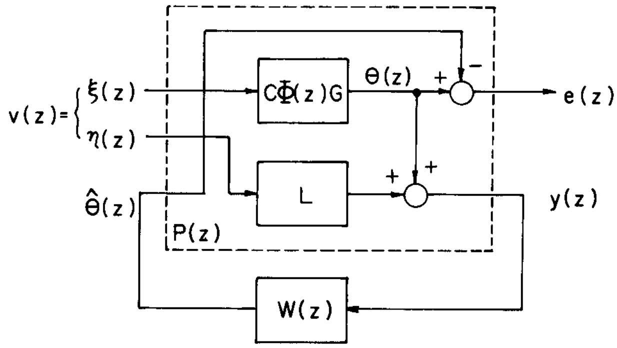

Fig. 1 Standard block diagram for Wiener filtering

$$e(z) = [P_{11}(z) - W(z)P_{21}(z)]v(z) \tag{3.15}$$

The discrete-time Wiener filtering problem is therefore transformed into a standard model matching problem minimizing

$$J = \|P_{11}(z) - W(z)P_{21}(z)\|_2$$

For simplicity, we define $T_1(z) := P_{11}(z)$ and $T_2(z) := P_{21}(z)$. Let a co-inner-outer factorization of $T_2(z)$ be given by $T_2(z) = T_{2co}(z)T_{2ci}(z)$ where $T_{2ci}(z)$ is co-inner and $T_{2co}(z)$ is co-outer. It therefore follows from (1.4) that the optimal filter $W(z)$ is given by

$$W(z) = [T_1(z)T_{2ci}^*(z)]_+ T_{2co}^{-1}(z) \tag{3.16}$$

3. 3 State-Space Solution

It follows from (3.13) that realizations of $T_1(z)$ and $T_2(z)$ are respectively given by

$$T_1(z) = \left[\begin{array}{c|cc} A & G & 0 \\ \hline C & 0 & 0 \end{array}\right] \tag{3.17}$$

and

$$T_2(z)=\left[\begin{array}{c|cc} A & G & 0 \\ \hline C & 0 & L \end{array}\right] \tag{3.18}$$

In order to obtain a co-inner-outer factorization of $T_2(z)$, we apply Theorem 2.1 to

$$T_2(z)^T=\left[\begin{array}{c|c} A^T & C^T \\ \hline G^T & 0 \\ 0 & L^T \end{array}\right]$$

From (2.11), we get $E=A^T$, $\Sigma=C^T(LL^T)^{-1}C$, and $\Pi=GG^T$. The ARE associated with $T_2(z)^T$ is therefore given by

$$Y=AYA^T-AYC^T(LL^T+CYC^T)^{-1}CYA^T+GG^T \tag{3.19}$$

Define

$$H=-AYC^T(LL^T+CYC^T)^{-1} \tag{3.20}$$

$$V=(LL^T+CYC^T)^{1/2} \tag{3.21}$$

It then follows from (2.22) that

$$T_{2o}(z)^T=\left[\begin{array}{c|c} A^T & C^T \\ \hline -V^TH^T & V^T \end{array}\right]$$

so that

$$T_{2co}(z)=\left[\begin{array}{c|c} A & -HV \\ \hline C & V \end{array}\right] \tag{3.22}$$

Hence, a co-inner function is found to be

$$T_{2ci}(z)=\left[\begin{array}{c|cc} A_H & G & HL \\ \hline V^{-1}C & 0 & V^{-1}L \end{array}\right] \tag{3.23}$$

where $A_H:=A+HC$ is stable.

We see from (3.17) and (3.23) that

$$\begin{aligned} T_1(z)T_{2ci}^*(z)&=\left[\begin{array}{c|cc} A & G & 0 \\ \hline C & 0 & 0 \end{array}\right]\left[\begin{array}{c|cc} A_H & G & HL \\ \hline V^{-1}C & 0 & V^{-1}L \end{array}\right]^* \\ &=\left[\begin{array}{c|cc} A & G & 0 \\ \hline C & 0 & 0 \end{array}\right]\left[\begin{array}{c|c} A_H^{-T} & -A_H^{-T}C^T \\ \hline G^TA_H^{-T} & -G^TA_H^{-T}C^T \\ L^TH^TA_H^{-T} & L^T-L^TH^TA_H^{-T}C^T \end{array}\right]V^{-T} \\ &=\left[\begin{array}{cc|c} A_H^{-T} & 0 & -A_H^{-T}C^T \\ GG^TA_H^{-T} & A & -GG^TA_H^{-T}C^T \\ \hline 0 & C & 0 \end{array}\right]V^{-T} \\ &=\left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline \bar{C} & \bar{D} \end{array}\right] \end{aligned} \tag{3.24}$$

where A_H is assumed to be nonsingular.

Define $T=\begin{bmatrix} I & 0 \\ Y & I \end{bmatrix}$ using the stabilizing solution of (3.19). Note also that the ARE of (3.19) is rewritten as

$$(-Y+G^TG)A_H^{-T}+AY=0$$

It then follows that

$$T^{-1}\bar{A}T=\begin{bmatrix} A_H^{-T} & 0 \\ 0 & A \end{bmatrix}$$

$$T^{-1}\bar{B}=\begin{bmatrix} -A_H^{-T}C^TV^{-T} \\ AYC^TV^{-T} \end{bmatrix}$$

$$\bar{C}T=[\,CY \quad C\,]$$

Hence, (3.24) becomes

$$T_1(z)T_{2ci}^*(z)=\left[\begin{array}{cc|c} A_H^{-T} & 0 & -A_H^{-T}C^TV^{-T} \\ 0 & A & AYC^TV^{-T} \\ \hline CY & C & 0 \end{array}\right]$$

$$=C(zI-A)^{-1}AYC^TV^{-T}-CY(zI-A_H^{-T})^{-1}A_H^{-T}C^TV^{-T} \tag{3.25}$$

We see that the first term of the r.h.s. of (3.25) obviously belongs to RH_+^∞, since A is stable and A_H^{-T} is antistable. Thus, if we insist that the filter is strictly causal, then

$$[\,T_1(z)T_{2ci}^*(z)\,]_+=\left[\begin{array}{c|c} A & -HV \\ \hline C & 0 \end{array}\right] \tag{3.26}$$

since $AYC^T(VV^T)^{-1}=-H$.

Theorem 3.1

The optimal filter, or the one-step predictor, is given by

$$W(z)=-C(zI-A-HC)^{-1}H \tag{3.27}$$

Proof

It follows from (3.16), (3.22) and (3.26) that

$$W(z)=\left[\begin{array}{c|c} A & -HV \\ \hline C & 0 \end{array}\right]\left[\begin{array}{c|c} A & -HV \\ \hline C & V \end{array}\right]^{-1}$$

$$=I_p-V\left[\begin{array}{c|c} A & -HV \\ \hline C & V \end{array}\right]^{-1}$$

$$=I_p-\left[\begin{array}{c|c} A+HC & H \\ \hline C & I_p \end{array}\right]$$

This completes the proof. □

On the other hand, if we do not assume that the optimal filter is strictly causal, then the operator $[\,\cdot\,]_+$ should include the constant terms as well as the strictly causal part in (3.25). As in Section 3.1, we see that the constant term equals CYC^TV^{-T}, by taking $z \to 0$ in the second term in the r.h.s. of (3.25). Thus we get

$$[\,T_1(z)T_{2ci}^*(z)\,]_+=\left[\begin{array}{c|c} A & -HV \\ \hline C & CYC^TV^{-T} \end{array}\right] \tag{3.28}$$

Theorem 3.2

The optimal filter is given by

$$W(z)=I_p-(LL^T)(VV^T)^{-1}-(LL^T)(VV^T)^{-1}C(zI-A-HC)^{-1}H \tag{3.29}$$

Proof

It follows from (3.22) and (3.28) that

$$H(z)=\left[\begin{array}{c|c} A & -HV \\ \hline C & CYC^TV^{-T}\end{array}\right]\left[\begin{array}{c|c} A & -HV \\ \hline C & V\end{array}\right]^{-1}$$

$$=I_p+[\,CYC^TV^{-T}-V]\left[\begin{array}{c|c} A & -HV \\ \hline C & V\end{array}\right]^{-1}$$

$$=I_p+[\,CYC^TV^{-T}-V]\left[\begin{array}{c|c} A+HC & H \\ \hline V^{-1}C & V^{-1}\end{array}\right]$$

$$=I_p+[\,CYC^T(VV^T)^{-1}-I_p]\left[\begin{array}{c|c} A_H & H \\ \hline C & I_p\end{array}\right]$$

Since $CYC^T(VV^T)^{-1}-I_p=-(LL^T)(VV^T)^{-1}$, we have (3.29). □

4. Stationary LQG Problem

In this section, we deal with the stationary LQG problem for the discrete-time system via the technique employed in Section 3. The LQG problem is converted into a model matching problem based on a doubly coprime factorization, and then the optimal controller is derived by applying the inner-outer and co-inner-outer factorizations.

4.1 Problem Statement

Let a discrete-time linear stochastic system be described by

$$x(t+1)=Ax(t)+Bu(t)+G\xi(t)$$
$$y(t)=Cx(t)+L\eta(t) \tag{4.1}$$

where $x(t)$ is the $n\times 1$ state vector, $u(t)$ is the $m\times 1$ control vector, $y(t)$ is the $p\times 1$ output observation vector, $\xi(t)$ is the $q\times 1$ process noise, and $\eta(t)$ is the $p\times 1$ observation noise. A, B, C, G, L are constant matrices of dimensions $n\times n$, $n\times m$, $p\times n$, $n\times q$, $p\times p$, respectively. We assume that $\xi(t)$ and $\eta(t)$ are stationary white Gaussian noises with means zero and the covariance matrices

$$E\{\xi(k)\xi^T(\ell)\}=I_q\delta_{k\ell}$$
$$E\{\eta(k)\eta^T(\ell)\}=I_p\delta_{k\ell} \tag{4.2}$$

where $\delta_{k\ell}$ is the Kronecker delta. It is also assumed that the initial state $x(0)$ and noises $\xi(t)$, $\eta(t)$ are independent.

Let Q and R be $n\times n$ nonnegative definite and $m\times m$ positive definite matrices, respectively. Then the stationary LQG problem is to minimize the steady-state average cost

$$J=E\{x^T(t)Qx(t)+u^T(t)Ru(t)\}\ \text{as}\ t\to\infty \tag{4.3}$$

with respect to a causal controller. It is required that the closed-loop system is internally asymptotically stable. We assume for simplicity that the admissible controller is strictly causal, although, as in Section 3.3, we can derive the optimal controller that is causal, but not strictly causal. We also assume that $(A, B, Q^{1/2})$, (A, G, C) are minimal, where $Q=Q^{T/2}Q^{1/2}$.

4. 2 Model Matching Problem

We define

$$e(t):=\begin{bmatrix} Q^{1/2}x(t) \\ R^{1/2}u(t) \end{bmatrix},\quad v(t):=\begin{bmatrix} \xi(t) \\ \eta(t) \end{bmatrix} \tag{4.4}$$

Then, from (4.1) and (4.4), in z-domain,

$$\begin{bmatrix} e \\ y \end{bmatrix}=\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}\begin{bmatrix} v \\ u \end{bmatrix}$$

$$=\begin{bmatrix} Q^{1/2}\Phi(z)G & 0 & Q^{1/2}\Phi(z)B \\ 0 & 0 & R^{1/2} \\ C\Phi(z)G & L & C\Phi(z)B \end{bmatrix}\begin{bmatrix} \xi \\ \eta \\ u \end{bmatrix} \tag{4.5}$$

where $\Phi(z):=(zI-A)^{-1}$ (see Fig. 2).

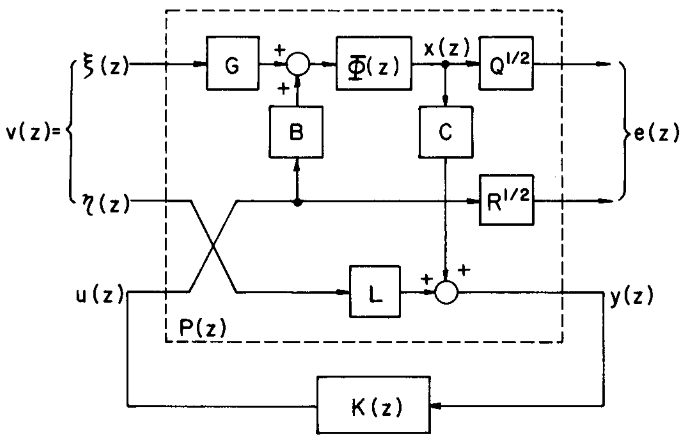

Fig. 2 Block diagram for stationary LQG problem

Let the admissible controller be given by $K(z)$. Then

$$u=K(z)y \tag{4.6}$$

It therefore follows from (4.5) and (4.6) that

$$e=[P_{11}+P_{12}K(z)(I-P_{22}K(z))^{-1}P_{21}]v \tag{4.7}$$

Hence, by the Parseval Theorem, minimizing J of (4.3) is equivalent to minimizing the H_2 norm

$$\|e\|_2^2=\frac{1}{2\pi j}\int_{|z|=1}e^*(z)e(z)\frac{dz}{z}$$

with respect to $K(z)$. Since the covariance matrix of $v(t)$ is identity, the stationary LQG control problem reduces to minimizing

$$J=\|P_{11}+P_{12}K(z)(I-P_{22}K(z))^{-1}P_{21}\|_2 \tag{4.8}$$

Doubly coprime factorizations of discrete-time transfer matrices are the same as those of continuous-time transfer matrices, so that $P_{22}(z):=C\Phi(z)B$ is factored as [2], [4]

$$P_{22}(z)=N_2M_2^{-1}=\tilde{M}_2^{-1}\tilde{N}_2$$

where

$$\begin{bmatrix} \tilde{X}_2 & -\tilde{Y}_2 \\ -\tilde{N}_2 & \tilde{M}_2 \end{bmatrix}\begin{bmatrix} M_2 & Y_2 \\ N_2 & X_2 \end{bmatrix}=\begin{bmatrix} M_2 & Y_2 \\ N_2 & X_2 \end{bmatrix}\begin{bmatrix} \tilde{X}_2 & -\tilde{Y}_2 \\ -\tilde{N}_2 & \tilde{M}_2 \end{bmatrix}=I_{p+m}$$

It is well known [2], [4] that the set of all proper rational stabilizing controllers is parametrized as

$$\begin{aligned} K(z)&=[Y_2(z)-M_2(z)Q(z)][X_2(z)-N_2(z)Q(z)]^{-1} \\ &=[\tilde{X}_2(z)-Q(z)\tilde{N}_2(z)]^{-1}[\tilde{Y}_2(z)-Q(z)\tilde{M}_2(z)] \end{aligned} \tag{4.9}$$

where $Q(z)$ is an arbitrary transfer matrix belonging to RH_+^∞. Substituting (4.9) into (4.8) and rearranging the terms yield

$$J=\|T_1(z)-T_2(z)Q(z)T_3(z)\|_2 \tag{4.10}$$

where

$$T_1(z)=P_{11}+P_{12}M_2\tilde{Y}_2P_{21}$$

$$T_2(z)=P_{12}M_2$$

$$T_3(z)=\tilde{M}_2P_{21}$$

Now we choose F, H so that $A_F:=A+BF$, $A_H:=A+HC$ are stable. Then, the state-space realizations of M_2, N_2, $\tilde{M}_2$, $\tilde{N}_2$, X_2, Y_2, $\tilde{X}_2$, $\tilde{Y}_2$ are given by [4], [11]

$$M_2=\left[\begin{array}{c|c} A_F & B \\ \hline F & I_m \end{array}\right],\quad N_2=\left[\begin{array}{c|c} A_F & B \\ \hline C & 0 \end{array}\right]$$

$$\tilde{M}_2=\left[\begin{array}{c|c} A_H & H \\ \hline C & I_p \end{array}\right],\quad \tilde{N}_2=\left[\begin{array}{c|c} A_H & B \\ \hline C & 0 \end{array}\right]$$

$$X_2=\left[\begin{array}{c|c} A_F & -H \\ \hline C & I_p \end{array}\right],\quad Y_2=\left[\begin{array}{c|c} A_F & -H \\ \hline F & 0 \end{array}\right]$$

$$\tilde{X}_2=\left[\begin{array}{c|c} A_H & -B \\ \hline F & I_m \end{array}\right],\quad \tilde{Y}_2=\left[\begin{array}{c|c} A_H & -H \\ \hline F & 0 \end{array}\right] \tag{4.11}$$

Since, from (4.5),

$$P_{11}(z)=\left[\begin{array}{c|cc} A & G & 0 \\ \hline Q^{1/2} & 0 & 0 \\ 0 & 0 & 0 \end{array}\right],\quad P_{12}(z)=\left[\begin{array}{c|c} A & B \\ \hline Q^{1/2} & R^{1/2} \end{array}\right],\quad P_{21}(z)=\left[\begin{array}{c|cc} A & G & 0 \\ \hline C & 0 & L \end{array}\right]$$

it follows that

$$T_1(z)=\left[\begin{array}{cc|cc} A_F & -BF & G & 0 \\ 0 & A_H & G & HL \\ \hline Q^{1/2} & 0 & 0 & 0 \\ R^{1/2}F & -R^{1/2}F & 0 & 0 \end{array}\right] \tag{4.12a}$$

$$T_2(z)=\left[\begin{array}{c|c} A_F & B \\ \hline Q^{1/2} & 0 \\ R^{1/2}F & R^{1/2} \end{array}\right] \tag{4.12b}$$

$$T_3(z)=\left[\begin{array}{c|cc} A_H & G & HL \\ \hline C & 0 & L \end{array}\right] \tag{4.12c}$$

It may be noted that the derivations of (4.12) are the same as those of the continuous-time counterparts [7].

4. 3 Solution to Discrete-Time LQG Problem

In order to apply the formula (1.4), we need the inner-outer and co-inner-outer factorizations of $T_2(z)$ and $T_3(z)$, respectively.

From (4.12b), E, Σ, Π defined by (2.11) are given by

$E=A,\ \Sigma=BR^{-1}B^T,\ \Pi=Q$

Thus the ARE associated with $T_2(z)$ is given by

$$X=A^TXA-A^TXB(R+B^TXB)^{-1}BXA+Q \tag{4.13}$$

Since $(A,\ B,\ Q^{1/2})$ is minimal, (4.13) has the unique stabilizing solution X. Also, from (2.15), we get

$$K_2:=-(R+B^TXB)^{-1}(B^TXA_F+RF)$$
$$=-F-(R+B^TXB)^{-1}B^TXA$$

It therefore follows from Theorem 2.1 that

$$T_{2o}(z)=V_2\left[\begin{array}{c|c} A_F & B \\ \hline -K_2 & I_m \end{array}\right] \tag{4.14}$$

where, from (2.20), $V_2:=(R+B^TXB)^{1/2}$. Since X is the stabilizing solution, if we take $F=-(R+B^TXB)^{-1}B^TXA$, then $A_F:=A+BF$ is stable. Thus the simplest form of the outer function is given by $T_{2o}(z)=V_2$, so that the inner function is

$$T_{2i}(z)=\left[\begin{array}{c|c} A_F & B \\ \hline Q^{1/2} & 0 \\ R^{1/2}F & R^{1/2} \end{array}\right]V_2^{-1} \tag{4.15}$$

A co-inner-outer factorization of $T_3(z)$ is obtained by applying Theorem 2.1 to

$$T_3(z)^T=\left[\begin{array}{c|c} A_H^T & C^T \\ \hline G^T & 0 \\ L^TH^T & L^T \end{array}\right] \tag{4.16}$$

It follows from (4.16) and (2.11) that

$E=A^T,\ \Sigma=C^T(LL^T)^{-1}C,\ \Pi=GG^T$

so that the ARE associated with $T_3(z)^T$ is

$$Y=AYA^T-AYC^T(LL^T+CYC^T)^{-1}CYA^T+GG^T \tag{4.17}$$

Note that (4.17) is the same as (3.19), which is employed for the state-space solution to the Wiener filtering problem. We see that since (A, G, C) is minimal, (4.17) has the unique stabilizing solution Y. Hence, from (2.15), we get

$$\begin{aligned} K_3^T :&= -(LL^T+CYC^T)^{-1}(CYA_H^T+LL^TH^T) \\ &= -H^T-(LL^T+CYC^T)^{-1}CYA^T \end{aligned}$$

Thus, from Theorem 2.1,

$$T_{3o}(z)^T = V_3^T\left[\begin{array}{c|c} A_H^T & C^T \\ \hline -K_3^T & I_p \end{array}\right]$$

where

$$V_3^T=(LL^T+CYC^T)^{1/2} \tag{4.18}$$

The co-outer function is therefore given by

$$T_{3co}(z)=\left[\begin{array}{c|c} A_H & -K_3 \\ \hline C & I_m \end{array}\right]V_3 \tag{4.19}$$

Here, if we take $H=-AYC^T(LL^T+CYC^T)^{-1}$ then $A_H := A+HC$ is stable, because Y is the stabilizing solution. Thus, the simplest co-outer function is $T_{3co}(z)=V_3$, and the corresponding co-inner function becomes

$$\begin{aligned} T_{3ci}(z) &= T_{3co}^{-1}(z)T_3(z) \\ &= V_3^{-1}\left[\begin{array}{c|cc} A_H & G & HL \\ \hline C & 0 & L \end{array}\right] \end{aligned} \tag{4.20}$$

In the following, we evaluate $T_{2i}^*(z)T_1(z)T_{3ci}^*(z)$, where F and H are fixed as $F=-(R+B^TXB)^{-1}B^TXA$ and $H=-AYC^T(LL^T+CYC)^{-1}$. We see from (4.12a) and (4.15) that

$$\begin{aligned} T_{2i}^*T_1 &= \left[\begin{array}{c|c} A_F & BV_2^{-1} \\ \hline Q^{1/2} & 0 \\ R^{1/2}F & R^{1/2}V_2^{-1} \end{array}\right]^* \left[\begin{array}{cc|cc} A_F & -BF & G & 0 \\ 0 & A_H & G & HL \\ \hline Q^{1/2} & & 0 & 0 \\ R^{1/2}F & -R^{1/2}F & 0 & 0 \end{array}\right] \\ &= V_2^{-T}\left[\begin{array}{c|cc} A_F^{-T} & A_F^{-T}Q^{T/2} & A_F^{-T}F^TR^{T/2} \\ \hline -B^TA_F^{-T} & -B^TA_F^{-T}Q^{T/2} & R^{T/2}-B^TA_F^{-T}F^TR^{T/2} \end{array}\right] \\ &\times \left[\begin{array}{cc|cc} A_F & -BF & G & 0 \\ 0 & A_H & G & HL \\ \hline Q^{1/2} & 0 & 0 & 0 \\ R^{1/2}F & -R^{1/2}F & 0 & 0 \end{array}\right] \\ &= V_2^{-T}\left[\begin{array}{ccc|cc} A_F & -BF & 0 & G & 0 \\ 0 & A_H & 0 & G & HL \\ A_F^{-T}(Q+F^TRF) & -A_F^{-T}F^TRF & A_F^{-T} & 0 & 0 \\ \hline C_1 & C_2 & C_3 & 0 & 0 \end{array}\right] \end{aligned} \tag{4.21}$$

where

$$C_1=-B^TA_F^{-T}(Q+F^TRF)+RF$$
$$C_2=-RF+B^TA_F^{-T}F^TRF$$
$$C_3=-B^TA_F^{-T}$$

Introducting the basis change

$$T=\begin{bmatrix} I & 0 & I \\ 0 & 0 & I \\ -X & I & -X \end{bmatrix}$$

and using the ARE of (4.13), (4.21) becomes

$$T_{2i}^*T_1=V_2^{-T}\left[\begin{array}{ccc|cc} A_F & 0 & -HC & 0 & -HL \\ 0 & A_F^{-T} & 0 & XG & 0 \\ 0 & 0 & A_H & G & HL \\ \hline 0 & -B^TA_F^{-T} & -B^TXA & 0 & 0 \end{array}\right]$$

$$=V_2^{-T}\left[\begin{array}{cc|cc} A_F^{-T} & 0 & XG & 0 \\ 0 & A_H & G & HL \\ \hline -B^TA_F^{-T} & -B^TXA & 0 & 0 \end{array}\right] \qquad (4.22)$$

It therefore follows from (4.20) and (4.22) that

$$T_{2i}^*T_1T_{3ci}^*=V_2^{-T}\left[\begin{array}{cc|cc} A_F^{-T} & 0 & XG & 0 \\ 0 & A_H & G & HL \\ \hline -B^TA_F^{-T} & -B^TXA & 0 & 0 \end{array}\right]\left[\begin{array}{c|cc} A_H & G & HL \\ \hline V_3^{-1}C & 0 & V_3^{-1}L \end{array}\right]^*$$

$$=V_2^{-T}\left[\begin{array}{cc|cc} A_F^{-T} & 0 & XG & 0 \\ 0 & A_H & G & HL \\ \hline -B^TA_F^{-T} & -B^TXA & 0 & 0 \end{array}\right]$$

$$\times\left[\begin{array}{c|c} A_H^{-T} & -A_H^{-T}C^T \\ \hline G^TA_H^{-T} & -G^TA_H^{-T}C^T \\ L^TH^TA_H^{-T} & L^T-L^TH^TA_H^{-T}C^T \end{array}\right]V_3^{-T}$$

$$=V_2^{-T}\left[\begin{array}{ccc|c} A_H^{-T} & 0 & 0 & B_1 \\ XGG^TA_H^{-T} & A_F^{-T} & 0 & B_2 \\ A_{31} & 0 & A_H & B_3 \\ \hline 0 & B^TA_F^{-T} & B^TXA & 0 \end{array}\right]V_3^{-T} \qquad (4.23)$$

where

$$A_{31}=G^TGA_H^{-T}+HLL^TH^TA_H^{-T}$$
$$B_1=A_H^{-T}C^T,\ B_2=XGG^TA_H^{-T}C^T$$
$$B_3=GG^TA_H^{-T}C^T-HLL^T+HLL^TH^TA_H^{-T}C^T$$

Again, by the basis change

$$T=\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ Y & 0 & I \end{bmatrix}$$

and by the ARE of (4.17), we get

$$T_{2i}^{*}T_{1}T_{3i}^{*}=V_{2}^{-T}\left[\begin{array}{ccc|c} A_{H}^{-T} & 0 & 0 & A_{H}^{-T}C^{T} \\ XGG^{T}A_{H}^{-T} & A_{F}^{-T} & 0 & XGG^{T}A_{H}^{-T}C^{T} \\ 0 & 0 & A_{H} & 0 \\ \hline B^{T}XAY & B^{T}A_{F}^{-T} & B^{T}XA & 0 \end{array}\right]V_{3}^{-T}$$

$$=V_{2}^{-T}\left[\begin{array}{cc|c} A_{H}^{-T} & 0 & A_{H}^{-T}C^{T} \\ XGG^{T}A_{H}^{-T} & A_{F}^{-T} & XGG^{T}A_{H}^{-T}C^{T} \\ \hline B^{T}XAY & B^{T}A_{F}^{-T} & 0 \end{array}\right]V_{3}^{-T} \tag{4.24}$$

Since A_H, A_F are stable, the r.h.s. of this equation is antistable. Thus we get

$$[\,T_{2i}^{*}(z)\,T_{1}(z)\,T_{3ci}^{*}(z)\,]_{+}=0$$

so that, from (1.4), $Q_{opt}(z)=0$.

Theorem 4.1

The transfer function of the optimal controller is given by

$$K(z)=\left[\begin{array}{c|c} A_F & -H \\ \hline F & 0 \end{array}\right]\left[\begin{array}{c|c} A_F & -H \\ \hline C & I_p \end{array}\right]^{-1} \tag{4.25}$$

$$=-F(zI-A-BF-HC)^{-1}H$$

Proof

A proof is immediate from $K(z)=Y_2(z)X_2^{-1}(z)$. □

It may be noted that (4.25) is the transfer matrix of the well-known optimal stationary LQG regulator for the discrete-time system.

4. 4 Causal Optimal Controller

The optimal controller of (4.25) is strictly causal, so that it consists of the noise-free regulator and the Kalman filter that produces the one-step predicted estimate. If we include constant terms in the causal part $[\,\cdot\,]_{+}$, then, by taking $z\rightarrow 0$ in (4.24), we get

$$[\,T_{2i}^{*}T_{1}T_{3ci}^{*}]_{+}=V_{2}^{-T}B^{T}XAYC^{T}V_{3}^{-T}$$

Since $T_{2o}(z)=V_2$, $T_{3co}(z)=V_3$, we get from (1.4)

$$\begin{aligned} Q_{opt}(z) &= (V_2^{T}V_2)^{-1}B^{T}XAYC^{T}(V_3V_3^{T})^{-1} \\ &= (R+B^{T}XB)^{-1}B^{T}XAYC^{T}(LL^{T}+CYC^{T})^{-1} \\ &= FA^{-1}H \end{aligned} \tag{4.26}$$

Theorem 4.2

The transfer function of the optimal controller, not strictly causal, is given by

$$K(z)=-FA^{-1}H-FA^{-1}A_{H}(zI-A-BF-HC-BFA^{-1}HC)^{-1}A_{F}A^{-1}H \tag{4.27}$$

Proof

It follows from (4.9), (4.11) and (4.26) that

$$K(z)=[Y_2(z)-M_2Q_{opt}(z)][X_2(z)-N_2Q_{opt}(z)]^{-1}$$

$$=\left[\begin{array}{c|c} A_F & H+BFA^{-1}H \\ \hline -F & -FA^{-1}H \end{array}\right]\left[\begin{array}{c|c} A_F & H+BFA^{-1}H \\ \hline -C & I_p \end{array}\right]^{-1}$$

$$=\left[\begin{array}{c|c} A_F & H+BFA^{-1}H \\ \hline -F & -FA^{-1}H \end{array}\right]\left[\begin{array}{c|c} A_F+HC+BFA^{-1}HC & H+BFA^{-1}H \\ \hline C & I_p \end{array}\right]$$

$$=-\left[\begin{array}{cc|c} A_F & HC+BFA^{-1}HC & H+BFA^{-1}H \\ 0 & A_F+HC+BFA^{-1}HC & H+BFA^{-1}H \\ \hline -F & FA^{-1}HC & FA^{-1}H \end{array}\right]$$

By the basis change $T=\begin{bmatrix} I & I \\ 0 & I \end{bmatrix}$, we get

$$K(z)=-\left[\begin{array}{cc|c} A_F & 0 & 0 \\ 0 & A_F+HC+BFA^{-1}HC & H+BFA^{-1}H \\ \hline F & F+FA^{-1}HC & FA^{-1}H \end{array}\right]$$

$$=-\left[\begin{array}{c|c} A_F+HC+BFA^{-1}HC & H+BFA^{-1}H \\ \hline F+FA^{-1}HC & FA^{-1}H \end{array}\right]$$

This completes the proof. □

It is not difficult to show that the above controller is formed by the noise-free optimal regulator and the Kalman filter that produces the filtered estimate [1].

5. Conclusions

We have developed solutions to the discrete-time Wiener filtering and stationary LQG problem by converting them to model matching problems and applying the H_2 control theory. Both strictly causal and causal solutions are derived based on the state-space technique. It may be noted that the state-space technique developed for the model matching problem [2], [4] is very powerful in manipulating various transfer matrices.

The derivation of the inner-outer factorization in Section 2 is applicable to the spectral factorization of $\gamma^2 I-G^*(z)G(z)$ that appears in H_∞ optimizations. This will be presented elsewhere.

Appendix A

The proof of Lemma 2.1 is based on the use of the matrix inversion lemma. Define $\tilde{A}=\begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{21} \\ \tilde{A}_{12} & \tilde{A}_{22} \end{bmatrix}$ Then from (2.10a) and (2.11), we get [5]

$$\begin{aligned}
\tilde{A}_{11} &= A - B(D^TD - B^TA^{-T}C^TD)^{-1}(D^T - B^TA^{-T}C^T)C \\
&= E + B(D^TD)^{-1}D^TC \\
&\quad - B(D^TD)^{-1}[I - B^TA^{-T}C^TD(D^TD)^{-1}]^{-1}(D^T - B^TA^{-T}C^T)C \\
&= E + B(D^TD)^{-1}[I - B^TA^{-T}C^TD(D^TD)^{-1}]^{-1} \\
&\quad \times \{[I - B^TA^{-T}C^TD(D^TD)^{-1}]D^T - D^T + B^TA^{-T}C^T\}C \\
&= E + B(D^TD)^{-1}[I - B^TA^{-T}C^TD(D^TD)^{-1}]^{-1} \\
&\quad \times B^TA^{-T}C^T[I - D(D^TD)^{-1}D^T]C \\
&= E + B(D^TD)^{-1}B^T[A^T - C^TD(D^TD)^{-1}B^T]^{-1}\Pi \\
&= E + \Sigma E^{-T}\Pi \qquad \text{(A1)}
\end{aligned}$$

$$\begin{aligned}
\tilde{A}_{12} &= -B(D^TD - B^TA^{-T}C^TD)^{-1}B^TA^{-T} \\
&= -B(D^TD)^{-1}B^T[A^T - C^TD(D^TD)^{-1}B^T]^{-1} \\
&= -\Sigma E^{-T} \qquad \text{(A2)}
\end{aligned}$$

$$\begin{aligned}
\tilde{A}_{21} &= -A^{-T}C^TC + A^{-T}C^TC[D^TD - B^TA^{-T}C^TD]^{-1}(D^T - B^TA^{-T}C^T)C \\
&= -[A^{-T} + A^{-T}C^TD(D^TD - B^TA^{-T}C^T)^{-1}B^TA^{-T}]C^TC \\
&\quad + A^{-T}C^TD[D^TD - B^TA^{-T}C^TD]^{-1}D^TC \\
&= -[A^T - C^TD(D^TD)^{-1}B^T]^{-1}C^TC \\
&\quad + [A^T - C^TD(D^TD)^{-1}B^T]^{-1}C^TD(D^TD)^{-1}D^TC \\
&= -E^{-T}C^T[I - D(D^TD)^{-1}D^T]C \\
&= -E^{-T}\Pi \qquad \text{(A3)}
\end{aligned}$$

$$\begin{aligned}
\tilde{A}_{22} &= A^{-T} + A^{-T}C^TD(D^TD - B^TA^{-T}C^TD)^{-1}B^TA^{-T} \\
&= [A^T - C^TD(D^TD)^{-1}B^T]^{-1} \\
&= E^{-T} \qquad \text{(A4)}
\end{aligned}$$

This completes the proof of Lemma 2.1. □

Appendix B

We see that, from (2.12),

$$\begin{aligned}
T^{-1}\tilde{A}T &= \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}\begin{bmatrix} E+\Sigma E^{-T}\Pi & -\Sigma E^{-T} \\ -E^{-T}\Pi & E^{-T} \end{bmatrix}\begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \\
&= \begin{bmatrix} E+\Sigma E^{-T}(\Pi - X) & -\Sigma E^{-T} \\ 0 & -X\Sigma E^{-T} + E^{-T} \end{bmatrix} \qquad \text{(B1)}
\end{aligned}$$

where the (2,1)-block is zero from the discrete-time ARE of (2.13). It follows from (2.14b), (2.14c) and (2.16) that the (1,1)-block of (B1) is given by $E_c = A + BK$, and the (2,2)-block is E_c^{-T}. Also, from (2.10b),

$$\begin{aligned}
T^{-1}\tilde{B} &= \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}\begin{bmatrix} B \\ -A^{-T}C^TD \end{bmatrix}\Delta^{-1} \\
&= \begin{bmatrix} B\Delta^{-1} \\ -A^{-T}(A^TXB + C^TD)\Delta^{-1} \end{bmatrix} \qquad \text{(B2)}
\end{aligned}$$

and, from (2.10c),

$$\tilde{C}T=-\Delta^{-1}[\Gamma^T C \quad B^T A^{-T}]\begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$$

$$=-\Delta^{-1}[\Gamma^T C+B^T A^{-T}X \quad B^T A^{-T}] \tag{B3}$$

We evaluate the (1,1)-block of (B3). Similarly to the derivation of (A1),

$$\begin{aligned}(D^TD)\Delta^{-1}\Gamma^T C &= (D^TD)(D^TD-B^TA^{-T}C^TD)^{-1}(D^T-B^TA^{-T}C^T)C \\ &= [I-B^TA^{-T}C^TD(D^TD)^{-1}]^{-1}(D^T-B^TA^{-T}C^T)C \\ &= D^TC+[I-B^TA^{-T}C^TD(D^TD)^{-1}]^{-1} \\ &\quad \times \{D^T-B^TA^{-T}C^T-[I-B^TA^{-T}C^TD(D^TD)^{-1}]D^T\}C \\ &= D^TC-[I-B^TA^{-T}C^TD(D^TD)^{-1}]^{-1} \\ &\quad \times B^TA^{-T}C^T[I-D(D^TD)^{-1}D^T]C \\ &= D^TC-B^TE^{-T}\Pi \end{aligned} \tag{B4}$$

Also,

$$\begin{aligned}(D^TD)\Delta^{-1}B^TA^{-T}X &= (D^TD)[D^TD-B^TA^{-T}C^TD]^{-1}B^TA^{-T}X \\ &= B^TE^{-T}X \end{aligned} \tag{B5}$$

Thus combining (B4) and (B5) yields

$$\begin{aligned}\alpha := (D^TD)\Delta^{-1}(\Gamma^TC+B^TA^{-T}X) \\ = D^TC+B^TE^{-T}(X-\Pi) \end{aligned} \tag{B6}$$

But, from (2.13), (2.14a) and (2.15)

$$\begin{aligned}E^{-T}(X-\Pi) &= XE-XB(D^TD+B^TXB)^{-1}B^TXE \\ &= X(A+BK) \end{aligned} \tag{B7}$$

Substituting (B7) into (B6) gives

$$\begin{aligned}\alpha &= D^TC+B^TXA+B^TXBK \\ &= -(D^TD+B^TXB)K+B^TXBK \\ &= -(D^TD)K \end{aligned}$$

Thus (B3) becomes

$$\tilde{C}T=[K \quad -\Delta^{-1}B^TA^{-T}]$$

This completes the proof of Lemma 2.3. □

Appendix C

It follows from the hypotheses of the theorem that $\mathscr{H}$ has no eigenvalues on the unit circle, so that Lemma 2.2 applies. It is easy to see that $G_o(z)$ of (2.22) is outer, since $A+BK$ is stable. Thus $G_i(z)$ is obtained as

$$\begin{aligned}G_i(z) &= G(z)G_o^{-1}(z) \\ &= \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]\left[\begin{array}{c|c} A+BK & BV^{-1} \\ \hline K & V^{-1} \end{array}\right]\end{aligned}$$

$$=\left[\begin{array}{cc|c} A+BK & 0 & BV^{-1} \\ BK & A & BV^{-1} \\ \hline DK & C & DV^{-1} \end{array}\right] \tag{C1}$$

Introducing the basis change $T=\begin{bmatrix} I & 0 \\ I & I \end{bmatrix}$, (C1) reduces to

$$G_i(z)=\left[\begin{array}{cc|c} A+BK & 0 & BV^{-1} \\ 0 & A & 0 \\ \hline C+DK & C & DV^{-1} \end{array}\right]=\left[\begin{array}{c|c} A+BK & B \\ \hline C+DK & D \end{array}\right]V^{-1}$$

This completes the proof of Theorem 2.1. □

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, N. J.: Prentice-Hall, 1979.

[2] J. C. Doyle, "Lecture notes in advances in multivariable control," ONR/Honeywell Workshop, Minnenapolis, MN, 1984.

[3] J. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-space solutions to standard H_2 and H_∞ control problem," *IEEE Trans. Automat. Control*, vol. AC-34, no. 8, pp. 831-847, Aug. 1989.

[4] B. A. Francis, *A Course in H_∞ Control Theory*. Berlin: Springer-Verlag, 1987.

[5] D.-W. Gu, M. C. Tsai, S. D. O'Young and I. Postlethwaite, "State-space formulae for discrete-time H_∞ optimization," *Int. J. Control*, vol. 19, no. 5, pp. 1683-1723, May 1989.

[6] T. Kailath, *Lectures on Wiener and Kalman Filtering*. New York: Springer, 1981.

[7] T. Katayama and Y. Washikita, "Solutions to Wiener filtering and stationary LQG problem via H_2 optimal control theory-Part I: Continuous-time system," *Memoris of the Faculty of Engineering*, Kyoto University, vol. 52, pt. 3, pp. 219-244, July 1990.

[8] V. Kucera, "The discrete Riccati equation of optimal control," *Kybernetika*, vol. 8, no. 5, pp. 430-447, 1972.

[9] A. J. Laub, "A Schur method for solving algebraic Riccati equation," *IEEE Trans.* vol. AC-24, no. 6, pp. 913-921, Dec. 1979.

[10] T. Papas, A. J. Laub and N. R. Sandel, Jr., "On the numerical solution of the discrete-time algebraic Riccati equation," *IEEE Trans. Automat. Control*, vol. AC-25, no. 4, pp. 631-641, Oct. 1980.

[11] M. Vidyasagar, *Control System Synthesis: A Factorization Approach*. Cambridge, Mass.: The MIT Press, 1985.