ON A DETERMINISTIC LEAST SQUARES ESTIMATION THEORY FOR LTI SYSTEMS

Nader Safari-Shad
K. N. Toosi University of Technology
Department of Electrical Engineering
P.O. Box 16765 – 3455
Tehran, Iran 16314
E-Mail: safari@irearn.bitnet

J. Daniel Cobb
University of Wiscosin-Madison
Department of Electrical and
Computer Engineering
1415 Johnson Drive
Madison, WI 53706 – 1691
E-mail: cobb@engr.wisc.edu

ABSTRACT

A deterministic least squares estimation theory of linear time-invariant systems is presented. It is demonstrated that the well-known Kalman filtering results can be obtained for purely deterministic systems with impulsive noises entering into every measurement. Applying known results from singular optimal control theory, the Kalman filtering results are extended to cases where some or all measurements may be deemed noise-free.

KEYWORDS

Kalman filtering, singular stochastic optimal observers, singular optimal control.

1 Introduction

A novel approach to the linear optimal filtering problem was presented by Kalman and Bucy in the early sixties [1]. The approach commonly known as Kalman filtering generates the minimum variance estimate of the state of a dynamic system in the face of random noises in the system and the measurements. However, applications of Kalman filtering in continuous-time have been mostly limited to systems with nonsingular measurement covariance matrices. This has posed limitations in cases where some or all measurements may be deemed noise-free.

To alleviate this difficulty with Kalman filtering in noise-free cases, several approaches have been proposed. For instance, the Luenberger observer theory has been suggested as a natural alternative. Indeed this has been supported by the stochastic optimal observer theory where a minimum variance state estimate is obtained by assuming that some or all initial states are random, [2].

In spite of the interest in stochastic optimal observers, the problem of least squares estimation of systems with some or all noise-free measurements has remained unsolved. Motivated by this observation, we formulate a deterministic optimal estimation problem whereby not only the well-known Kalman filtering results are obtained, but also the singular filtering cases are addressed applying the known results from singular optimal control theory.

2 Problem Formulation and Statement

Consider the completely observable linear time-invariant (LTI) system

$$\begin{aligned} \dot{x} &= Ax \; ; \qquad x(0) = x_0 \\ y &= Cx + \delta D w_0 \end{aligned} \tag{1}$$

where $x \in \mathbf{R}^n$, $y \in \mathbf{R}^p$, $w_0 \in \mathbf{R}^q$ and δ is the Dirac delta function with support at $t = 0$. The triple (A, C, D) is assumed known. Letting $\rho[\cdot]$ denote the rank of a matrix, it is assumed that $\rho[C] = p$ and $0 \leq \rho[D] \leq p$. The initial condition x_0, and the impulsive noise direction, w_0 are assumed unknown.

In our formulation, we consider estimators of the form

$$\begin{aligned} \hat{x} &= \mathcal{H}(y) \\ \mathcal{H} &: \mathbf{D}_p'^{+} \to \mathbf{D}_n'^{+} \end{aligned}$$

where $\mathcal{H}$ belongs to the class of LTI, causal and continuous operators, and $\mathbf{D}_k'^{+}$ denotes the vector space of k-tuples of distributions with support in $[0, \infty)$, [3].

It is shown in [4] that this class of operators is precisely the set of convolution operators with kernels in $\mathbf{D}_{n\times p}'^{+}$. Hence,

$$\hat{x} = \mathcal{H} * y$$

where $*$ denotes the convolution operator as defined in [3] and $\mathcal{H} \in \mathbf{D}_{n\times p}'^{+}$. Under these conditions, we note that our admissible class of estimators includes infinite-dimensional systems; hence, our formulation is more general than that considered in [2].

Just how well we estimate the state $x(t)$ may be quantified by means of a cost functional that we define next. Consider the $\mathbf{L}^2$-norm squared of the errors incurred in the estimation of the states normalized relative to the unknown parameter vector $\xi_0 = \begin{bmatrix} x_0 \\ w_0 \end{bmatrix} \in \mathbf{R}^{n+q}$, with the weighting function $f(\xi_0)$; i.e.,

$$J = \int_{\mathbf{R}^{n+q}} \int_0^\infty e^T e \; dt \; f(\xi_0) \; dV_{n+q}(\xi_0)$$

$$e = \exp(At)\, x_0 \;\; - \;\; \mathcal{H} * (C \exp(At)\, x_0 + \delta\, D w_0)\,.$$

Here, f is *any* nonnegative measurable function on $\mathbf{R}^{n+q}$ and $dV_{n+q}(\cdot)$ is the differential volume in $\mathbf{R}^{n+q}$.

The choice of the weighting function $f(\xi_0)$ is usually dictated by the nature of the uncertainty ξ_0. Here we restrict ourself to uncertainties confined to a weighted euclidean ball of radius one; i.e., let

$$f(\xi_0) = \begin{cases} 1 & \forall\, \xi_0 \in \Omega_0 \\ 0 & \text{otherwise} \end{cases}$$

where

$$\Omega_0 = \{\xi_0 \in \mathbf{R}^{n+q} \;\; : \;\; \xi_0^T\, \Sigma^{-1} \xi_0 \leq 1\}$$

with Σ a positive definite symmetric matrix. Under these conditions, the cost J reduces to

$$J \;\; = \;\; \alpha_{n,q}\; Tr \int_0^\infty \mathcal{E}\, \Sigma_{11}\, \mathcal{E}^T - 2\mathcal{E}\, \Sigma_{12} D^T\, \mathcal{H}^T + \mathcal{H}\, D \Sigma_{22} D^T\, \mathcal{H}^T \; dt$$

where

$$\mathcal{E} = \exp(At) - \mathcal{H} * C \exp(At)\,,$$

Σ_{ij}'s are the partitioned blocks in Σ, and Tr denotes the trace operator. Indeed, this is verified by observing that

$$\int_{\Omega_0} \xi_0\, \xi_0^T \;\, dV_{n+q}(\xi_0) \;\; = \alpha_{n,q}\; \Sigma$$

where $\alpha_{n,q}$ is a real positive constant given by

$$\alpha_{n,q} = \frac{(\sqrt{2\pi})^{n+q} (\det \Sigma)^{\frac{1}{2}}}{(n+q)(n+q+2)\Gamma(\frac{n+q}{2})}$$

with $\Gamma(\cdot)$ being the gamma function (see [5], Section 5.5 for details).

Noting that $\alpha_{n,q}$ is a positive scalar, the cost J can be normalized relative to it without altering the optimization problem; i.e., write

$$J = Tr \int_0^\infty \mathcal{E}\, \Sigma_{11}\, \mathcal{E}^T - 2\, \mathcal{E}\, \Sigma_{12} D^T\, \mathcal{H}^T + \mathcal{H}\, D \Sigma_{22} D^T\, \mathcal{H}^T \; dt\;.$$

We now define the deterministic optimal estimation problem as the task of minimizing the above cost J relative to $\mathcal{H} \in \mathbf{D}'^{+}_{n \times p}$, subject to the constraint

$$\mathcal{E} = \exp(At) - \mathcal{H} * C \exp(At)\,.$$

3 Duality with Linear Quadratic Optimal Control Problem

Our intention here is to expose a dual relation between the deterministic optimal estimation problem and the well-known linear quadratic (LQ) optimal control problem. To this end, we observe the following relablelling

$$\begin{array}{lll} \tilde{A} \doteq A^T & \tilde{B} \doteq C^T & \tilde{U} \doteq -\mathcal{H}^T \\ \tilde{X} \doteq \mathcal{E}^T & Q \doteq \Sigma_{11} & N \doteq \Sigma_{12} D^T \\ R \doteq D \Sigma_{22} D^T \;\; . & & \end{array}$$

Thus, we have

$$J = Tr \int_0^\infty \tilde{X}^T Q \tilde{X} + 2\tilde{X}^T N \tilde{U} + \tilde{U}^T R \tilde{U} \; dt$$

where

$$\tilde{X} \doteq \exp(\tilde{A}t) + \exp(\tilde{A}t)\tilde{B} * \tilde{U} \; .$$

Evidently, solving the optimal control problem

$$\min_{\tilde{U} \in \mathbf{D}'^{+}_{p \times n}} \; J$$

subject to the constraint imposed on $\tilde{X}$, is dual to solving the primal optimal estimation problem. Henceforth, we refer to the above relabellings simply as the duality relations.

It turns out that this problem is related to a problem that we are more familiar with. To clarify, let $\tilde{x}_{i0}^T = [\; 0, \; \ldots, \; 0, \; 1, \; 0, \; \ldots, \; 0 \;]$ where the i-th entry is 1. Also, partition the matrix $\tilde{U}$ as $\tilde{U} = \begin{bmatrix} \tilde{u}_1, & \tilde{u}_2, & \ldots & \tilde{u}_n \end{bmatrix}$ where $\tilde{u}_i \in \mathbf{R}^{p \times 1}$. Then, we can write

$$\tilde{X} = \begin{bmatrix} \tilde{x}_1, & \tilde{x}_2, & \ldots & \tilde{x}_n \end{bmatrix}$$

where

$$\tilde{x}_i = \exp(\tilde{A}t)\tilde{x}_{i0} + \exp(\tilde{A}t)\tilde{B} * \tilde{u}_i$$

for all $i = 1, 2, \ldots, n$. With some matrix algebra, we may write the cost J as

$$J = \sum_{i=1}^{n} J_i$$

where

$$\begin{aligned} J_i &= \int_0^\infty \tilde{x}_i^T Q \, \tilde{x}_i + 2\tilde{x}_i^T N \, \tilde{u}_i + \tilde{u}_i^T R \, \tilde{u}_i \; dt \\ \dot{\tilde{x}}_i &= \tilde{A}\tilde{x}_i \; + \; \tilde{B}\tilde{u}_i \; ; \quad \tilde{x}_i(0) \; = \; \tilde{x}_{i0} \, . \end{aligned}$$

Hence, minimization of J with respect to $\tilde{U}$ can be reduced to minimization of each J_i with respect to $\tilde{u}_i$.

In the following, we shall exploit this duality for obtaining a complete set of solutions to the deterministic optimal estimation problem for the cases where all measurements are noisy (i.e., $\rho[D] = p$), and the cases where some or all measurements are noise-free (i.e., $0 \leq \rho[D] < p$).

4 Regular Case: Completely Noisy Measurements

To present a complete treatment of the deterministic optimal estimation problem, we begin by considering the so-called regular case; i.e., $\rho[D] = p$. This corresponds to a cost functional which is positive definite in the estimation kernel $\mathcal{H}$.

The solution to this case can be readily obtained by invoking the duality with the regular LQ problem (see e.g., [7]).

Theorem 4.1 *For the completely noisy deterministic optimal estimation problem, the optimal $\mathcal{E}$ and $\mathcal{H}$ are given by*

$$\begin{aligned} \mathcal{E}^* &= \exp\left((A - L^* C)\, t\right) \\ \mathcal{H}^* &= \mathcal{E}^* L^* \end{aligned}$$

where

$$L^* = (PC^T + \Sigma_{12}D^T)\,(D\Sigma_{22}D^T)^{-1}$$

and $P \in \mathbf{R}^{n\times n}$ *is the unique positive definite symmetric solution of*

$$AP + PA^T - (PC^T + \Sigma_{12}\,D^T)\,(D\Sigma_{22}\,D^T)^{-1}\,(PC^T + \Sigma_{12}\,D^T)^T + \Sigma_{11} = 0\ .$$

Moreover, the optimal cost is $J = Tr\,P$.

Proof: The proof follows by invoking duality with LQ optimal control problem (see e.g., Chapter 3 in [7]). **Q.E.D.**

In view of the above result, a dynamic state estimator system corresponding to $\mathcal{H}^*$ is realized by

$$\dot{\hat{x}} = A\hat{x} + L^*\,(y - C\hat{x})\ ;\quad \hat{x}(0) = 0\ .$$

Note that this system has the structure of a full-order Luenberger observer where the observer gain matrix L^* is obtained by solving a full-order algebraic Riccati equation for P.

5 Singular Case: Partially Noise-Free Measurements

This section considers the partially singular deterministic optimal estimation problem; i.e., $0 < \rho[D] < p$. This case is characterized by a cost functional which is partially singular in the estimation kernel $\mathcal{H}$.

Consider the system (1) with the assumption that D has rank $p_1 < p$, $p_1 \neq 0$ and that C and D are already in the form

$$C \;=\; \begin{bmatrix} C_{11} & C_{12} \\ 0 & I_{p_2} \end{bmatrix} \quad D \;=\; \begin{bmatrix} D_1 \\ 0 \end{bmatrix}$$

with $D_1 \in \mathbf{R}^{p_1\times q}$ and $p = p_1 + p_2$. Note that since C has full rank and D is of rank p_1, the required forms for C and D can always be achieved with an appropriate similarity transformation and a suitable coordinate change on y.

Recall that the cost functional is given by

$$J = Tr\int_0^\infty \mathcal{E}\,\Sigma_{11}\,\mathcal{E}^T - 2\mathcal{E}\,\Sigma_{12}D^T\,\mathcal{H}^T + \mathcal{H}D\Sigma_{22}D^T\,\mathcal{H}^T\;dt\,.$$

Denote $D\Sigma_{22}D^T$ by W and note that by our assumption on D,

$$W = \begin{bmatrix} W_{11} & 0 \\ 0 & 0 \end{bmatrix} \geq 0$$

where $W_{11} \doteq D_1\Sigma_{22}D_1^T > 0$. Before we present our main result, we need some preliminaries.

It is convenient to eliminate the cross-term in the cost by introducing a suitable transformation on $\mathcal{H}$. To this end, we partition $\mathcal{H}$ as

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}_1 & \mathcal{H}_2 \end{bmatrix}$$

where $\mathcal{H}_1 \in \mathbf{D}'^{+}_{n\times p_1}$, $\mathcal{H}_2 \in \mathbf{D}'^{+}_{n\times p_2}$. Then we introduce the transformation

$$\begin{aligned} \tilde{\mathcal{H}}_1 &= \mathcal{H}_1 - \mathcal{E}\,\Sigma_{12}\,D_1^T W_{11}^{-1} \\ \tilde{\mathcal{H}}_2 &= \mathcal{H}_2\ . \end{aligned}$$

It is readily verified that

$$J = Tr \int_0^\infty \mathcal{E} V \mathcal{E}^T + \tilde{\mathcal{H}} W \tilde{\mathcal{H}}^T \; dt$$

$$\mathcal{E} = \exp(\hat{A}) - \tilde{\mathcal{H}} * C \exp(\hat{A})$$

where

$$\begin{aligned}
\tilde{\mathcal{H}} &= \begin{bmatrix} \tilde{\mathcal{H}}_1 & \tilde{\mathcal{H}}_2 \end{bmatrix} \\
V &= \Sigma_{11} - \Sigma_{12} \, D_1^T W_{11}^{-1} D_1 \, \Sigma_{12}^T \\
\hat{A} &= \begin{bmatrix} A_{11} - \Sigma_{12,1} \, D_1^T W_{11}^{-1} C_{11} & A_{12} - \Sigma_{12,1} \, D_1^T W_{11}^{-1} C_{12} \\ A_{21} - \Sigma_{12,2} \, D_1^T W_{11}^{-1} C_{11} & A_{22} - \Sigma_{12,2} \, D_1^T W_{11}^{-1} C_{12} \end{bmatrix}
\end{aligned}$$

and $\Sigma_{12} = \begin{bmatrix} \Sigma_{12,1} \\ \Sigma_{12,2} \end{bmatrix}$. Note that in the expression for $\hat{A}$, we have partitioned the matrix A into 4 blocks such that $A_{11} \in \mathbf{R}^{(n-p_2)\times(n-p_2)}$, $A_{22} \in \mathbf{R}^{p_2 \times p_2}$. Partitioning the matrix V into 4 blocks such that V_{11} and V_{22} have identical dimensions to those of A_{11} and A_{22}, respectively, we have

Theorem 5.1 *For the partially singular deterministic optimal estimation problem, the optimal $\mathcal{E}$ and $\mathcal{H}$ are given by*

$$\begin{aligned}
\mathcal{E}^* &= \begin{bmatrix} \exp(A_s t) & \exp(A_s t)\Theta_s \\ 0 & 0 \end{bmatrix} \\
\mathcal{H}^* &= \mathcal{E}^* L_s + \delta L_f
\end{aligned}$$

where

$$\begin{aligned}
A_s &= A_{11} - (P_1 C_s^T + M_s) W_s^{-1} C_s \\
M_s &= \begin{bmatrix} V_{12} & \Sigma_{12,1} D_1^T \end{bmatrix} \\
C_s &= \begin{bmatrix} A_{21} - \Sigma_{12,2} D_1^T W_{11}^{-1} C_{11} \\ C_{11} \end{bmatrix} \\
W_s &= \begin{bmatrix} V_{22} & 0 \\ 0 & W_{11} \end{bmatrix}
\end{aligned}$$

$P_1 \in \mathbf{R}^{(n-p_2)\times(n-p_2)}$ *is the unique positive definite symmetric solution of*

$$A_{11} P_1 + P_1 A_{11}^T - (P_1 C_s^T + M_s) W_s^{-1} (P_1 C_s^T + M_s)^T + \Pi_s = 0$$

$$\Pi_s = V_{11} - V_{12} V_{22}^{-1} V_{12}^T \; .$$

and

$$\begin{aligned}
L_s &= \begin{bmatrix} L_{11s} & L_{12s} \\ L_{21s} & 0 \end{bmatrix} \\
L_{11s} &= (P_1 C_{11}^T + \Sigma_{12,1} D_1^T) W_{11}^{-1} \\
L_{12s} &= -P_1 C_{11}^T W_{11}^{-1} C_{12} + A_{12} + \Theta_s A_{22} \\
&\quad - A_s \Theta_s - (\Sigma_{12,1} + \Theta_s \Sigma_{12,2}) D_1^T W_{11}^{-1} C_{12} \\
L_{21s} &= \Sigma_{12,2} D_1^T W_{11}^{-1} \\
L_f &= \begin{bmatrix} 0 & -\Theta_s \\ 0 & I_{p_2} \end{bmatrix} \\
\Theta_s &= -(P_1 (A_{21} - \Sigma_{12,2} D_1^T W_{11}^{-1} C_{11})^T + V_{12}) V_{22}^{-1} \; .
\end{aligned}$$

Moreover, the optimal cost is $J^* = Tr\, P_1$.

Proof: The proof follows by invoking duality with the singular LQ optimal control problem and the results presented in [6]. **Q.E.D.**

In view of this theorem, a dynamic state estimator system corresponding to $\mathcal{H}^*$ is realized by

$$\begin{aligned} \dot{z} &= A_s z + E_s y\ ; \qquad z(0) = 0 \\ \hat{x} &= G_s z + L_f y \end{aligned}$$

where $z \in \mathbf{R}^{n-p_2}$ and

$$\begin{aligned} E_s &= \begin{bmatrix} E_{1s} & E_{2s} \end{bmatrix} \\ E_{1s} &= L_{11s} + \Theta_s L_{21s} \\ E_{2s} &= L_{12s} \\ G_s &= \begin{bmatrix} I_{n-p_2} \\ 0 \end{bmatrix}. \end{aligned}$$

It is easily recognize that this system has the structure of a reduced-order observer where the triple (A_s, E_s, L_f) are computed by solving a reduced-order algebraic Riccati equation for P_1.

6 Singular Case: Totally Noise-Free Measurements

In this section, we consider the totally singular deterministic optimal estimation problem; i.e., $D = 0$. This case is identified with a cost functional which is totally singular in the estimation kernel $\mathcal{H}$.

Consider the system (1) with the assumption that C has already been transformed into the form $C = \begin{bmatrix} 0 & I_p \end{bmatrix}$. Note that any C which has full rank can always be transformed to this form with an appropriate similarity transformation.

Under these conditions, the cost functional is given by

$$J = Tr \int_0^\infty \mathcal{E}\, V\, \mathcal{E}^T\ dt$$

where

$$\mathcal{E} = \exp{(At)} - \mathcal{H} * C \exp{(At)}$$

and $V \doteq \Sigma_{11}$. To present our result, we need to partition the matrices A and V into 4 blocks such that $A_{11}, V_{11} \in \mathbf{R}^{(n-p)\times(n-p)}$ and $A_{22}, V_{22} \in \mathbf{R}^{p\times p}$. Then we have

Theorem 6.1 *For the totally singular deterministic optimal estimation problem, the optimal* $\mathcal{E}$ *and* $\mathcal{H}$ *are given by*

$$\begin{aligned} \mathcal{E}^* &= \begin{bmatrix} \exp{(A_s t)} & \exp{(A_s t)}\Theta_s \\ 0 & 0 \end{bmatrix} \\ \mathcal{H}^* &= \mathcal{E}^* L_s + \delta L_f \end{aligned}$$

where

$$\begin{aligned} A_s &= A_{11} + \Theta_s A_{21} \\ \Theta_s &= -(P_1 A_{21}^T + V_{12}) V_{22}^{-1} \end{aligned}$$

$P_1 \in \mathbf{R}^{(n-p)\times(n-p)}$ is the unique positive definite symmetric solution of

$$A_{11}P_1 + P_1A_{11}^T - (P_1A_{21}^T + V_{12})V_{22}^{-1}(P_1A_{12}^T + V_{12})^T + V_{11} = 0$$

and

$$\begin{aligned} L_s &= \begin{bmatrix} L_{1s} \\ 0 \end{bmatrix} \\ L_{1s} &= A_{12} + \Theta_s A_{22} - A_s\Theta_s \\ L_f &= \begin{bmatrix} -\Theta_s \\ I_p \end{bmatrix}. \end{aligned}$$

Moreover, the optimal cost is given by $J^* = Tr\, P_1$.

Proof: Again the proof follows by invoking the duality with LQ optimal control problem and applying the known results presented in [6].

In view of this theorem, a dynamic state estimator system corresponding to $\mathcal{H}^*$ is realized by

$$\begin{aligned} \dot{z} &= A_s z + L_{1s} y \; ; \qquad z(0) = 0 \\ \hat{x} &= G_s z + L_f y \end{aligned}$$

where $z \in \mathbf{R}^{n-p}$ and

$$G_s = \begin{bmatrix} I_{n-p} \\ 0 \end{bmatrix}.$$

It is easily recognized that this system has the structure of a minimal-order observer where the triple (A_s, L_{1s}, L_f) are computed by solving a reduced-order algebraic Riccati equation for P_1.

7 Conclusion

We have formulated and solved a deterministic optimal state estimation problem of LTI systems with completely noisy, partially noise-free, and totally noise-free measurements, using the method of least squares. In each case, it is shown that the optimal estimator has the structure of a Luenberger observer where the gains are computed by solving a corresponding full-order, reduced-order, and minimal-order algebraic Riccati equation.

References

[1] R. E. Kalman and R. S. Bucy, *"New Results in Linear Filtering and Prediction Theory,"* Journal of Basic Engineering, pp 95-108, 1961.

[2] H. Maeda and H. Hino, *"Design of Optimal Observers for Linear Time-invariant Systems,"* INT. J. Contr. Vol. 19, no.5, pp 993-1004, 1974.

[3] L. Schwartz, **"Mathematics for the Physical Sciences,"** Addison-Wesley, 1966.

[4] L. Schwartz, **"Theorie des Distributions,"** (rev. ed.), Hermann, Paris, 1966.

[5] W. Fleming, **"Functions of Several Variables,"** New York: Springer-Verlag, 1977.

[6] J. C. Willems, A. Kitapci, and M. Silverman, *"Singular Optimal Control,"* SIAM J. Control and Optimization, Vol. 24, no. 2, March 1986.

[7] B. D. O. Anderson and J. B. Moore, **"Optimal Control: Linear Quadratic Methods,"** New Jersey: Prentice-Hall, 1990.