Kalman–Bucy and Minimax Filtering

ARTHUR J. KRENER

***Abstract*—It is shown that the Kalman–Bucy filter is also a minimax filter.**

I. Introduction

Consider the linear system

$$\dot{x}(t)=A(t)x(t)+B(t)u(t)$$
$$x(0)=x_0$$
$$z(t)=C(t)x(t)+v(t) \quad (1.1)$$

where the state $x(t)$ is $n\times 1$, the driving noise $u(t)$ is $l\times 1$, and observation $z(t)$ and observation noise $v(t)$ are $m\times 1$. The time-varying matrices $A(t)$, $B(t)$, and $C(t)$ are $n\times n$, $n\times l$, and $m\times n$, respectively.

The filtering problem is to "optimally estimate" $x(t)$, $t>0$, given the particular observation history $z(s)$, $0<s<t$, and some general information about x_0, $u(t)$, and $v(t)$, but not their specific values. Henceforth, we shall refer to $(x_0, u(\cdot), v(\cdot))$ as the disturbance triple.

In the *Kalman–Bucy* model, x_0 is assumed to be a Guassian random variable mean 0 and convergence P_0 and $u(t)$ and $v(t)$ are independent Guassian white noise processes with covariances $Q(t)\ \delta(t-s)$ and $R(t)\ \delta(t-s)$. The estimate $\hat{x}(t)$ for $t\in[0,T]$ is that which minimizes the conditional covariance of the error $\tilde{x}(t)=x(t)-\hat{x}(t)$, i.e., $\hat{x}(t)$ minimizes

$$E(b\tilde{x}(t)\tilde{x}'(t)b'|z(s),0<s<t) \quad (1.2)$$

for any $1\times n$ vector b. The form of the Kalman–Bucy filter is well known [1]. It is the solution of stochastic differential equation

$$\frac{d}{dt}\hat{x}(t)=A(t)\hat{x}(t)+F(t)(z(t)-C(t)\hat{x}(t))$$
$$\hat{x}(0)=0 \quad (1.3)$$

where the feedback is given by

$$F(t)=P(t)C'(t)R^{-1}(t) \quad (1.4)$$

for $P(t)$, the solution of the matrix Riccati differential equation

$$\dot{P}=AP+PA'+BQB'-PC'R^{-1}CP$$
$$P(0)=P_0. \quad (1.5)$$

The error covariance (1.2) is nonstochastic, i.e., independent of the particular observations and equal to $P(t)$,

$$E(b\tilde{x}(t)\tilde{x}'(t)b'|z(s),0<s<t)=E(b\tilde{x}(t)\tilde{x}'(t)b')=bP(t)b'. \quad (1.6)$$

In the *minimax* model, we do not make stochastic assumptions regarding the uncertainties, but instead assume that $(x_0,u(\cdot),v(\cdot))$ is an element of a Hilbert space $\mathcal{H}$ and is bounded in norm. For convenience, we assume the bound to be 1, but any bound results in the same estimate, although not the same error. The norm is given by

$$\|x_0,u(\cdot),v(\cdot)\|^2=x_0'P_0^{-1}x_0+\int_0^T u'(s)Q^{-1}(s)u(s)+v'(s)R^{-1}(s)v(s)\,ds \quad (1.7)$$

where P_0 and $Q(s)$ are nonnegative definite and $R(s)$ is positive definite. So that (1.7) is well defined, we adopt the following convention; if x_0 is in the range P_0, $x_0=P_0y$, then

$$x_0'P_0^{-1}x_0=y'x_0,$$

and if not, then

$$x_0'P_0^{-1}x_0=\infty.$$

This definition is independent of the choice of y since the null space of P_0 is orthogonal to the range of P_0. We adopt a similar convention for $u'(s)Q^{-1}(s)u(s)$ and define $\mathcal{H}$ as the space of triples of finite norm,

$$\|x_0,u(\cdot),v(\cdot)\|<\infty.$$

The interpretation of (1.7) is as the energy represented by the disturbance triple.

The minimax filtering problem is to find for each $t\in[0,T]$, the estimate $\hat{x}(t)$ based on the past observations $z(s)$, $0<s<t$ which minimizes the maximum of any linear functional of the error. The maximum is taken as $(x_0,u(\cdot),v(\cdot))$ ranges over those triples of norm less than or equal to one, and which give rise to the particular observations $z(s)$, $0<s<t$. In other words, the minimax estimator $\hat{x}(t)$ minimizes

$$\max\{b\tilde{x}(t):\|x_0,u(\cdot),v(\cdot)\|<1 \text{ and gives rise to } z(s),0<s<t\}.$$

The minimax estimator employs a worst case design philosophy and has a game-theoretic flavor. We assume that our opponent, Nature, chooses the uncertainties in order to hide the true state. Nature is restricted in the total amount energy that she can use. We seek the estimate which minimizes our maximum loss as measured by any linear functional of the error.

From the design point of view, the weighting matrices P_0^{-1}, $Q^{-1}(s)$, and $R^{-1}(s)$ are parameters to be chosen from information about the system being model. They play the same role as the covariances of the Kalman–Bucy model. We shall demonstrate in the next section that if weights are the inverses of the covariances, then the Kalman–Bucy and minimax estimates of the state of (1.1) are identical.

If P_0 and $Q(s)$ are not invertible, then as covariances of Guassian variables, this restricts the disturbances x_0 and $u(s)$ to lie in their range spaces with probability 1. Similarly, if the energy interpretation (1.7) is used, this restricts x_0 and $u(s)$ in exactly the same way.

Schweppe [4] considered a minimax estimator, but used a norm different from (1.7); hence he obtained an estimate different from Kalman–Bucy.

Other characterizations of the Kalman–Bucy filter as a minimax estimator have appeared in the literature. Mintz [2] assumed hybrid uncertainties, the initial condition and observation noise stochastic, and the driving noise deterministic. He weighted the loss function with the norm of the driving noise and obtained a Kalman–Bucy estimate.

Morris [3] also treated hybrid uncertainties and allowed the covariances to vary over compact intervals. He showed that the Kalman–Bucy filter for largest covariances is a minimax solution to these problems.

We treat only continuous time filtering, but the basic result, that the minimum variance estimate is also a minimax estimate, also holds for discrete time or prediction or smoothing problems.

II. The Equivalence of the Kalman–Bucy and Minimax Filters

It is well known (see, for example, Jazwinski [1, ch. 5]) that the minimum variance filter of Kalman–Bucy is also a maximum likelihood estimate for Gaussian disturbances. Moreover, the maximum likelihood estimate of $x(t)$ has a least squares (minimum energy) characterization. The estimate $\hat{x}(t)$ is the endpoint of the trajectory generated by the triple $(x_0,u(\cdot),v(\cdot))$ of minimum energy (1.7) necessary to produce the observations $z(s)$, $0<s<t$.

Let us indicate why this is also the minimax estimate; a formal proof will follow.

The set of disturbance triples giving rise to $z(s)$, $0<s<t$ is an affine subspace of $\mathcal{H}$, and if we also require the norm to be less than or equal to one, we obtain a ball $\mathcal{B}$, in that subspace. The map sending $(x_0,u(\cdot),v(\cdot))$ to $x(t)$ is linear; hence, it carries $\mathcal{B}$ to an ellipsoid $\mathcal{E}$ in

Manuscript received March 12, 1979.
The author is with the Department of Mathematics, University of California, Davis, CA 95616.

0018-9286/80/0400-0291$00.75 ©1980 IEEE

$\mathbf{R}^n$. The minimax estimate is the center of that ellipsoid. The center of $\mathfrak{B}$ is mapped to the center of $\mathfrak{E}$ so that the minimax estimate can be computed as the endpoint of the trajectory generated by the disturbance triple in the center of $\mathfrak{B}$. But from the Hilbert space geometry, the center of $\mathfrak{B}$ is characterized as the element of $\mathfrak{B}$ lying closest to the origin, i.e., the triple of minimum energy necessary to produce the observations.

Now for a formal proof. Let $(\bar{x}_0, \bar{u}(\cdot), \bar{v}(\cdot))$ be the triple of minimum energy (norm) giving rise to the observations $z(s)$, $0 \leq s \leq t$, and let $(x_0, u(\cdot), v(\cdot))$ be any other triple giving rise to the same observation and of norm less than 1. Let $\bar{x}(t)$ and $x(t)$ be the corresponding state trajectories. Then the Kalman–Bucy estimate $\hat{x}(t) = \bar{x}(t)$ [1].

Since $(\bar{x}_0, \bar{u}(\cdot), \bar{v}(\cdot))$ is of minimum energy, it follows that

$$\bar{x}_0 P_0^{-1}(x_0 - \bar{x}_0) + \int_0^t \bar{u}'(s) Q^{-1}(s)(u(s) - \bar{u}(s)) + \bar{v}'(s) R^{-1}(s)(v(s) - \bar{v}(s))\, ds = 0$$

or else the triples

$$(\bar{x}_0, \bar{u}(\cdot), \bar{v}(\cdot)) + \epsilon(x_0 - \bar{x}_0, u(\cdot) - \bar{u}(\cdot), v(\cdot) - \bar{v}(\cdot)),$$

which also gives rise to the observations, would be of less energy for some small positive or negative ϵ. From this, it follows that the norms of $(\bar{x}_0, \bar{u}(\cdot), \bar{v}(\cdot)) \pm (x_0 - \bar{x}_0, u(\cdot) - \bar{u}(\cdot), v(\cdot) - \bar{v}(\cdot))$ are the same. The plus sign is just the triple $(x_0, u(\cdot), v(\cdot))$ and leads to the error

$$\tilde{x}(t) = x(t) - \hat{x}(t) = x(t) - \bar{x}(t).$$

Since $x(t)$ depends linearly on x_0 and $u(\cdot)$, the negative sign leads to the negative of the above error. This shows that $\bar{x}(t)$ is the centroid of the set of all $x(t)$ generated by triples of norm ≤ 1 and giving rise to the observation; and hence, $\bar{x}(t)$ is the minimax estimate.

References

[1] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.
[2] M. Mintz, "A Kalman filter as a minimax estimator," *J. Optimiz. Theory Appl.*, vol. 9, pp. 99–111, 1972.
[3] J. M. Morris, "The Kalman filter: A robust estimator for some classes of linear quadratic problems," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 526–534, 1976.
[4] F. C. Schweppe, "Recursive state estimation: Unknown but bounded errors and systems inputs," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 22–28, 1968.