Reprinted by permission from IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. AC-14, No. 1, February 1969
pp. 114-115

Properties of Optimal Linear Smoothing

***Abstract*—The problem is considered of characterizing the improvement in estimating the state of a linear system when filtering is replaced by smoothing. It is found that when the optimal filter is exponentially asymptotically stable, a smoothing lag equal to several time constants associated with this filter yields practically all the possible improvement. The extent of improvement, as measured by an error variance matrix, is also found.**

Introduction

In this correspondence we are concerned with the problem of estimating the state at time t of a lumped linear system of the form

$$\frac{dx(t)}{dt} = F(t)x(t) + G(t)u(t)$$
$$y(t) = H'(t)x(t) \tag{1}$$

where u is Gaussian white noise of mean zero and covariance $Q(t)\delta(t - \tau)$, and the initial state of (1) at time t_0 is a Gaussian random variable of mean zero and covariance a nonnegative definite symmetric matrix P_0. It is assumed that measurements $z(\cdot)$ are available, where

$$z(t) = y(t) + v(t) \tag{2}$$

with v Gaussian white noise of mean zero and covariance $R(t)\delta(t - \tau)$, where R is symmetric positive definite for all t.

We shall also assume that the processes u and v are independent; when this is not the case, a transformation is possible of (1) and (2) which eliminates the dependence, see [1, ch. 16].

It is well known that the Kalman–Bucy filter [1]–[3] will produce a minimum variance linear unbiased estimate $\hat{x}(t/t)$ of $x(t)$, given the measurements $z(t)$ over $[t_0,t]$. But one can envisage situations where as good an estimate as possible of $x(t)$ is desired, and this estimate does not have to be immediately on line, i.e., available at time t. In this case a smoothed estimate of $x(t)$ is desired, produced from knowledge of the measurements over some interval for which t is an interior point.

By way of notation, $\hat{x}(t/t + T)$ will denote the minimum variance linear unbiased estimate of $x(t)$, given $z(t)$ over $[t_0,t + T]$. Intuitively, we note that for larger T a better estimate will result, and in fact it is readily shown that the variance of the estimation error decreases as T increases. The estimate of $x(t)$ which is optimum with respect to T is, of course,

$$\lim_{T\to\infty} \hat{x}(t/t + T)$$

but to obtain such an estimate is obviously impractical. The following questions then arise, motivated by a desire, on the one hand, to produce as good an estimate as possible of $x(t)$ and, on the other hand, to estimate $x(t)$ in finite time.

1) Is there a T (and what is it) such that $\hat{x}(t/t + T)$ is effectively

$$\lim_{T\to\infty} \hat{x}(t/t + T)$$

an

$$E\{[x(t) - \hat{x}(t/t + T)][x(t) - \hat{x}(t/t + T)]'\}$$

is effectively

$$E\{[x(t) - \lim_{T\to\infty} \hat{x}(t/t + T)][x(t) - \lim_{T\to\infty} \hat{x}(t/t + T)]'\} ?$$

2) What is the value of

$$E\{[x(t) - \lim_{T\to\infty} \hat{x}(t/t + T)]\,[x(t) - \lim_{T\to\infty} \hat{x}(t/t + T)]'\} ?$$

Manuscript received July 29, 1968. This work was supported by the Australian Research Grants Committee.

Review of Filtering and Smoothing

A filtered estimate $\hat{x}(t/t)$ of $x(t)$ is provided by the system shown in Fig. 1 [1]–[3]. The matrix $K(t)$ is given by

$$K(t) = P(t)H(t)R^{-1}(t) \tag{3}$$

where $P(t)$ satisfies

$$\dot{P} = FP + PF' - PHR^{-1}H'P + Q, \qquad P(t_0) = P_0. \tag{4}$$

A smoothed estimate $\hat{x}(t/t + T)$ of $x(t)$ is provided by the system shown in Fig. 2 [4]–[6] and the following equations apply:

$$\hat{x}(t/t + T) = \int_t^T A(\sigma)[z(\sigma) - H'(\sigma)\hat{x}(\sigma/\sigma)]d\sigma + \hat{x}(t/t) \tag{5}$$

where

$$A(\sigma) = P(t)\Phi'(\sigma,t)H'(\sigma)R^{-1}(\sigma) \tag{6}$$

and

$$\frac{d\Phi(\sigma,t)}{d\sigma} = [F(\sigma) - K(\sigma)H'(\sigma)]\Phi(\sigma,t), \qquad \Phi(t,t) = I. \tag{7}$$

Notice that $\Phi(\cdot,\cdot)$ is the transition matrix associated with the optimal filter. Notice also that as T increases, the integral in (5) is easy to compute "on line," since no part of the integrand depends explicitly on T.

The matrix $P(t)$ in (4) is not only of use in defining the gain $K(t)$, but it is also a measure of the error associated with the estimate $\hat{x}(t/t)$ through the formula [1]–[3],

$$P(t) = E\{[x(t) - \hat{x}(t/t)][x(t) - \hat{x}(t/t)]'\}. \tag{8}$$

For the smoothed estimate, it can be shown that [4]–[6]

$$E\{[x(t) - \hat{x}(t/t + T)][x(t) - \hat{x}(t/t + T)]'\} = P(t) - P(t)\int_t^T \Phi'(\sigma,t)H(\sigma) \cdot R^{-1}(\sigma)H'(\sigma)\Phi(\sigma,t)d\sigma P(t). \tag{9}$$

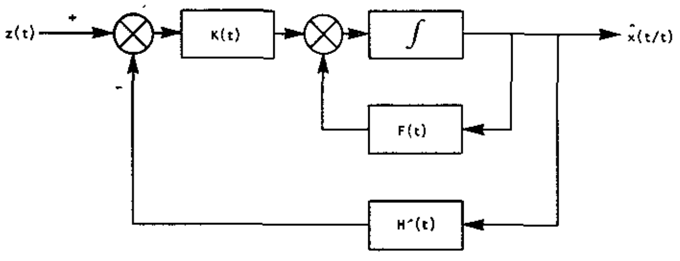

Fig. 1. Optimal filtering.

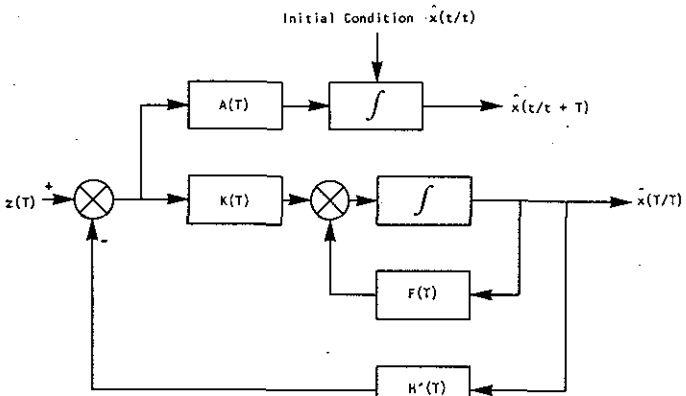

Fig. 2. Optimal smoothing.

This formula indicates that the smoothed error variance decreases monotonically with T.

Practical Implications of the Smoothing Formulas

We look first at the question of choosing T so that $\hat{x}(t/t+T)$ is a good approximation of

$$\lim_{T\to\infty} \hat{x}(t/t+T).$$

Inspection of Fig. 2 shows immediately that the behavior of the gain $A(T)$ for large T is critical, and an obvious requirement is that $A(T)$ approach zero as $T \to \infty$.

Let us impose the reasonable constraints that $H(T)$ and $R^{-1}(T)$ do not approach zero as $T \to \infty$. Then reference to (6) shows immediately that a sufficient condition for $A(T)$ to approach zero as $T \to \infty$ is that the system (7) be asymptotically stable. One way to guarantee this is to have the well-known assumptions [1], [2] that the pair $[F,GQ^{1/2}]$ is uniformly completely controllable, and that the pair $[F,H]$ is uniformly completely observable, together with F, G, H, Q, R, R^{-1} all bounded. These conditions are often satisfied in practice; thus for F, G, H, Q, and R constant, they merely require $[F,GQ^{1/2}]$ to be completely controllable, $[F,H]$ to be completely observable, and R to be nonsingular. Then it is known that the optimal filter is actually exponentially asymptotically stable, i.e., there exist positive constants c_1 and c_2 such that

$$\|\Phi(\sigma,t)\| < c_1 \exp[-c_2(\sigma - t)], \qquad \text{for all } \sigma \geq t \geq t_0. \quad (10)$$

It is now clear that when T is set to equal approximately five times the dominant time constant of the optimal filter, i.e., $5/c_2$, where c_2 is the largest constant which will suffice in (10), then $\hat{x}(t/t+T)$ will approximate

$$\lim_{T\to\infty} \hat{x}(t/t+T).$$

Now consider (9), which is a measure of the goodness of the estimate. With the restrictions as listed previously still applying and thus with (10) holding, it follows that

$$\int_t^T \Phi'(\sigma,t)H(\sigma)R^{-1}(\sigma)H'(\sigma)\Phi(\sigma,t)d\sigma \doteq \int_t^\infty \Phi'(\sigma,t)H(\sigma)R^{-1}(\sigma)H'(\sigma)\Phi(\sigma,t)d\sigma \quad (11)$$

when $T > 5/c_2$. The conclusion is then that $\hat{x}(t/t+(5/c_2))$ provides a good approximation to

$$\lim_{T\to\infty} \hat{x}(t/t+T)$$

in the sense that the associated error variances are approximately the same.

The quantity in (11) is of course a measurement of the improvement to be gained in smoothing. In connection with its computation, it is helpful to note that the quantity on the left-hand side of (11) is $M(t,T)$, where

$$\dot{M} = -M(F - KH') - (F - KH')'M - HR^{-1}H', \quad M(T,T) = 0. \quad (12)$$

When F, G, H, Q, and R are time invariant and $t_0 = -\infty$, it is easy to show that

$$\bar{M} = \lim_{T\to\infty} M(t,T)$$

is independent of t and satisfies the algebraic equation

$$M(F - KH') + (F - KH')'M = -HR^{-1}H' \quad (13)$$

for which solution procedures have been well studied. Notice that in the constant case, $P(t)$ is constant and nonsingular and

$$P^{-1}(F - KH') + (F - KH')'P^{-1} = -HR^{-1}H' - P^{-1}QP^{-1} \quad (14)$$

so that

$$(P^{-1} - M)(F - KH') + (F - KH')'(P^{-1} - M) = -P^{-1}QP^{-1}. \quad (15)$$

The matrix $P^{-1} - M$ is some sort of a measure of the smoothed error variance, which is actually $P[P^{-1} - M]P$. Equation (15) also shows that with Q nonzero, it is never possible to reduce the error variance to zero by smoothing.

Conclusions

It has been shown that a smoothed estimate can normally lead to improvement over a filtered estimate in the sense of having lower error variance. A value for the interval between the time the state is estimated and the time measurements are terminated has been found which allows essentially all the possible improvement to be obtained; the existence of this value depends on the optimal filter being exponentially asymptotically stable—a situation which often occurs.

An expression for the improvement in the error covariance matrix is also deduced.

The application to fixed time lag smoothing is obvious. Sacrifice of several time constants of the optimal filter in obtaining a trajectory estimate leads to an improved estimate (and essentially no further improvement can be achieved).

B. D. O. Anderson
Dept. of Elec. Engrg.
University of Newcastle
Newcastle, N.S.W., Australia

References

[1] R. E. Kalman, "New methods and results in linear filtering and prediction theory," in *Proceedings of the Symposium on Engineering Applications of Probability and Random Functions.* New York: Wiley, 1961.
[2] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME, J. Basic Engrg.*, ser. D, vol. 83, pp. 95–108, March 1961.
[3] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, J. Basic Engrg.*, ser. D, vol. 82, pp. 35–45, 1960.
[4] J. S. Meditch, "On optimal linear smoothing theory," *Information and Control*, vol. 10, pp. 598–615, June 1967.
[5] ——, "On optimal fixed point linear smoothing," *Internat'l J. Control*, vol. 6, no. 2, pp. 189–199, 1967.
[6] T. Kailath, "A Wold-Kolmogorov approach to linear least-squares estimation," Stanford Electronics Labs., Stanford, Calif., Rept. SEL-68-007, 1968.