

## Adaptive Estimation via Sequential Processing

R.M. HAWKES AND J.B. MOORE

**Abstract**—The computational advantages of processing vector measurement data one component at a time (termed sequential processing) in adaptive estimation schemes involving banks of Kalman filters is investigated.

### I. INTRODUCTION

The linear filtering of vector measurement data one component at a time (termed sequential processing) via Kalman filtering techniques has certain advantages over the more standard Kalman filter algorithms in which the components of a vector measurement are processed in one batch (termed simultaneous processing) [1]–[4]. For the case when the additive measurement noises in the components of the data vector are uncorrelated there is a computational saving in sequential processing which could be as much as 50 percent [1].

In view of the results of [1]–[4], it is not surprising that there are advantages to the sequential processing of data in adaptive estimation schemes involving banks of Kalman filters [6], [7]. What is surprising, however, is the extent of these advantages. Not only is there a possible computational advantage associated with the Kalman filtering as such, but there is inherently a significant computational advantage in processing the innovations of these filters to achieve the *a posteriori* densities. This latter aspect is now explored.

### II. ADAPTIVE ESTIMATION

Consider the linear discrete-time state-space signal model for the  $m$ -vector measurements  $z_k$ ,

$$x_{k+1} = F_k x_k + G_k w_k, \quad z_k = H_k' x_k + v_k \quad (1)$$

where  $v_k$  and  $w_k$  are independent zero-mean white-noise processes with covariance matrices  $R_k = R_k' > 0$  and  $Q_k = Q_k' > 0$ , respectively.

In the signal model (1),  $F_k$ ,  $H_k$ ,  $G_k$ ,  $Q_k$ , and  $R_k$  are known functions of a parameter vector  $\theta_s$  where  $\theta_s \in \{\theta_s\}_{s=1}^N$  with *a priori* probabilities  $p(\theta_s)$ ,  $s = 1, \dots, N$ . The adaptive estimator [6] consists of a bank of  $N$  Kalman filters where each filter is matched to each possible member of  $\{\theta_s\}_{s=1}^N$  with outputs  $\hat{x}_k|k, \theta_s$  weighted according to the *a posteriori* probabilities  $p(\theta_s|Z_k)$ ,  $s = 1, \dots, N$  where  $Z_k = \{z_0, z_1, \dots, z_k\}$  as

$$\hat{x}_k|k = \sum_{s=1}^N \hat{x}_k|k, \theta_s p(\theta_s|Z_k). \quad (2)$$

The results of [1]–[4] can be readily applied to obtain  $\hat{x}_k|k, \theta_s$  by sequential processing. If  $R_k$  is not block diagonal then it is necessary to introduce the LDU factorization [4], [5] as

$$R_k = J_k \bar{R}_k J_k', \quad \bar{R}_k = \text{diag} \{ \bar{R}_k^1 \bar{R}_k^2 \dots \bar{R}_k^m \} \quad (3)$$

where  $J_k$  is in lower block triangular form. Also let us introduce the

Manuscript received July 24, 1974. This work was supported in part by the ARGC Committee and in part by the RRB.

The authors are with the Department of Computing and Control, Imperial College, London, England, on leave from the Department of Electrical Engineering, University of Newcastle, NSW, Australia.

Copyright 1975, by the Institute of Electrical and Electronics Engineers, Inc.

PRINTED IN THE U.S.A. Annals No. 502AC025

definitions of  $\bar{H}_k^i$  from

$$\bar{H}_k^i = J_k^{-1} H_k^i, \quad \bar{H}_k = [\bar{H}_k^1 \bar{H}_k^2 \cdots \bar{H}_k^r]. \quad (4)$$

(Of course,  $J_k^{-1}$  exists since  $R_k$  is positive definite.)

Now consider a signal model for  $\bar{z}_k = J_k^{-1} z_k$  where we partition the measurement vector  $\bar{z}_k = J_k^{-1} z_k$  and its additive white-noise term  $\bar{v}_k = J_k^{-1} v_k$  into components  $\bar{z}_k^i$  and  $\bar{v}_k^i$  such that

$$\bar{z}_k^i = (\bar{H}_k^i)' x_k + \bar{v}_k^i. \quad (5)$$

Here  $E[\bar{v}_k^i \bar{v}_k^j] = \bar{R}_k^i \delta_{ki}$ . Let us also denote the measurements  $\bar{z}_0, \bar{z}_1, \dots, \bar{z}_k$  as  $\bar{Z}_k$ . We now show how  $p(\theta_s | Z_k) = p(\theta_s | \bar{Z}_k)$  can be calculated from the innovation process of the Kalman filter matched to  $\theta_s$ .

The *a posteriori* probabilities  $p(\theta_s | \bar{Z}_k)$  are given from the relationship

$$p(\theta_s | \bar{Z}_k) = \frac{p(\bar{z}_k | \bar{Z}_{k-1}, \theta_s) p(\theta_s | \bar{Z}_{k-1})}{\sum_{j=1}^N p(z_k | \bar{Z}_{k-1}, \theta_j) p(\theta_j | \bar{Z}_{k-1})}. \quad (6)$$

The conditional probabilities  $p(\bar{z}_k | \bar{Z}_{k-1}, \theta_s)$  are in turn expanded using Bayes rule as

$$p(\bar{z}_k | \bar{Z}_{k-1}, \theta_s) = \prod_{j=1}^r p(\bar{z}_k^j | \bar{Z}_{k-1}, \bar{z}_k^1, \bar{z}_k^2, \dots, \bar{z}_k^{j-1}, \theta_s). \quad (7)$$

These probabilities can be calculated from the pseudoinnovations process of the bank of  $N$  Kalman filters via the formula [7]

$$p(\bar{z}_k | \bar{Z}_{k-1}, \theta_s) = \prod_{j=1}^r |\Omega_k^j(\theta_s)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} [\nu_k^j(\theta_s)]' [\Omega_k^j(\theta_s)]^{-1} \cdot [\nu_k^j(\theta_s)] \right\} \quad (8)$$

$$\Omega_k^j(\theta_s) = [\bar{H}_k^j(\theta_s)]' \Sigma_k^j(\theta_s) [\bar{H}_k^j(\theta_s)] + \bar{R}_k^j(\theta_s) \quad (9)$$

$$\nu_k^j(\theta_s) = \bar{z}_k^j - [\bar{H}_k^j(\theta_s)]' \hat{x}_{k|k-1, \theta_s}. \quad (10)$$

The above algorithm is, of course, a sequential estimation algorithm and should be compared to the standard simultaneous algorithm [6], [7]

$$p(z_k | Z_{k-1}, \theta_s) = |\Omega_k(\theta_s)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \nu_k(\theta_s) \Omega_k^{-1}(\theta_s) \nu_k(\theta_s) \right\} \quad (11)$$

$$\Omega_k(\theta_s) = H_k'(\theta_s) \Sigma_k(\theta_s) H_k(\theta_s) + R_k(\theta_s) \quad (12)$$

$$\nu_k(\theta_s) = z_k - H_k'(\theta_s) \hat{x}_{k|k-1, \theta_s}. \quad (13)$$

The significant point for us here is that the sequential processing algorithm (8) is clearly simpler to work with than the algorithm (11)—at least given the appropriate filter bank in each case. It is certainly less effort to calculate  $\prod_{j=1}^r |\Omega_k^j(\theta_s)|^{-\frac{1}{2}}$  than  $|\Omega_k(\theta_s)|^{-\frac{1}{2}}$  and again it is certainly less effort to calculate  $\sum_{j=1}^r [\nu_k^j(\theta_s)]' [\Omega_k^j(\theta_s)]^{-1} [\nu_k^j(\theta_s)]$  than  $\nu_k(\theta_s) \Omega_k^{-1}(\theta_s) \nu_k(\theta_s)$  given the  $\nu$  and  $\Omega^{-1}$  terms. For the case when the  $\nu_k^j(\theta_s)$  are scalar and  $r = m$  the memory saving is  $m$  versus  $m^2$  and the computation saving is  $2m$  versus  $m(m+1)$ .

### III. DISCUSSION

We conclude that sequential processing for adaptive estimation is more efficient than the corresponding simultaneous processing algorithms—at least when the sequential filtering algorithms are no less efficient than the simultaneous filtering algorithms. Of course when the sequential filtering algorithms are less efficient, there is a tradeoff involved, the details of which will be omitted here.

### REFERENCES

- [1] R.A. Singer and R.G. Sea, "Increasing the computational efficiency of discrete Kalman filters," *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-16, pp. 254-257, June 1971.
- [2] H.W. Sorenson, "Kalman filtering techniques," in *Advances in Control Systems*, C.T. Leondes, Ed. New York: Academic, 1966.
- [3] J. Mendel, "Computational requirements for a discrete Kalman filter," *IEEE Trans. Automat. Contr.*, vol. AC-16, p. 748-758, Dec. 1971.
- [4] P. Kaminsky, A. Bryson, and S. Schmidt, "Discrete square root filtering: A survey of current techniques," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 727-735, Dec. 1971.
- [5] E. Isaacson and H. Keller, *Analysis of Numerical Methods*. New York: Wiley, 1966, pp. 34-37.
- [6] D.G. Lainiotis, "Optimal adaptive estimation: Structure and parameter adaptation," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 160-170, Apr. 1971.
- [7] F.L. Sims and D.G. Lainiotis, "Recursive algorithm for the calculation of adaptive Kalman filter weighting coefficients," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-14, pp. 215-218, Apr. 1969.