N72.27620

Smooth Empirical Bayes Estimation of Observation Error Variances in Linear Systems

H.F. Martz, Jr.*

M.W. Lian**

Texas Tech University, Lubbock, Texas 79409

CASE FILE COPY

Introduction

In Reference 1 an empirical Bayes estimator was developed for estimating the unknown random observation error variances in a discrete time linear system. There it was assumed that each unknown variance could be represented as the product of a known nominal value and an unknown random scale factor which is to be estimated.

A continuous empirical Bayes smoothing technique is developed in Reference 2. This technique provides estimators possessing smaller average squared error losses than the type of empirical Bayes estimator employed in Reference 1. A similar smooth empirical Bayes estimator was developed in Reference 3, where a continuous prior density function approximation was "smoothed" through a suitable function of the observation data.

In this note a smooth empirical Bayes estimator is developed for estimating the unknown random scale component of each observation error variance. This estimator will be shown to possess a smaller average squared error loss than the estimator presented in Reference 1.

Scale Factor Estimation

Consider the linear discrete dynamic system given by

$$x_i = \Phi_i x_{i-1} + u_{i-1} \tag{1}$$

augmented by the linear observation - state equation

$$y_i = H_i x_i + v_i, \quad i = 1, 2, \ldots \tag{2}$$

with the same assumptions as in Reference 1. As in Reference 1 the observation error covariance matrix R_i is represented by the diagonal matrix

$$R_i = \text{diag}(r_{i1}^2/\theta_{i1}, \ r_{i2}^2/\theta_{i2}, \ \ldots, \ r_{iq}^2/\theta_{iq}), \tag{3}$$

where q is the number of observation types available, r_{ij}^2 is a known nominal value of the observation error variance associated with the jth observation type at time epoch i. Also, the scale factors { θ_{ij}: i = 1, 2, $\ldots$ } are independent realizations of a random variable Θ_j having a completely unknown and unspecified prior density function $g_j(\theta)$ which is zero on the negative real numbers and which may be different for each observation type.

From the usual Gaussian assumption on v_i in Eq. (2), it follows that

$$z_{ij} = (y_{ij} - h_{ij}x_i)/r_{ij}, \quad i = 1, 2, \ldots, n_j \tag{4}$$

conditional on θ_{ij}, is distributed with probability density function given by

$$f(z_{ij}|\theta_{ij}) = \theta_{ij}^{\frac{1}{2}} (2\pi)^{-\frac{1}{2}} \exp[-\theta_{ij}z_{ij}^2/2] \tag{5}$$

Here h_{ij} is the jth row of H_i and n_j is the total number of observations of type j available up to the present time. For simplicity, we let $n_j = n$ and shall drop the subscript j for the remainder of this section.

The Bayes estimator for θ_n is given by

$$E(\theta_n|z_n) = \frac{\int_0^\infty \theta_n f(z_n|\theta_n) \, g(\theta_n) d\theta_n}{\int_0^\infty f(z_n|\theta_n) \, g(\theta_n) d\theta_n} \tag{6}$$

According to a technique developed in Reference 4 and used in References 2 and 3, the prior density function $g(\theta_n)$ may be estimated by means of the approximation given by

$$g_n(\theta_n) = \frac{K}{nh(2\pi)^{\frac{1}{2}}} \sum_{i=1}^n \exp[-\frac{1}{2}(\frac{\theta_n - \hat{\theta}_i}{h})^2], \quad 0 < \theta_n < \infty \tag{7}$$

where $h = h(n) = n^{-1/5}$, $\hat{\theta}_i$ is a suitably chosen estimate of θ_i to be discussed later, and where

$$K = [1 - \frac{1}{n} \sum_{i=1}^n \Phi(-\hat{\theta}_i/h)]^{-1} \tag{8}$$

In Equation (8), $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

Inserting Equations (5) and (7) into Equation (6), collecting terms, performing the indicated integrations, and simplifying, yields the smooth empirical Bayes estimator for θ_n given by

$$\overset{\sim}{\theta}_n = E_n(\theta_n|z_n) = \left(\frac{3}{2}h\right)\frac{\sum_{i=1}^{n} \exp[-a_i]\, U(2,b_i)}{\sum_{i=1}^{n} \exp[-a_i]\, U(1,b_i)} \tag{9}$$

where

$$a_i = (3z_n^4 h^4 + 8z_n^2 \hat{\theta}_i^2 h^2 + 4\hat{\theta}_i^2 - 4h^3 z_n^2 \hat{\theta}_i)/16h^2 \tag{10}$$

$$b_i = (z_n^2 h^2 - 2\hat{\theta}_i)/2h \tag{11}$$

and where $U(c,x)$ is the parabolic cylinder function[5,6] defined by

$$U(c,x) = [e^{-x^2/4}/\Gamma(c + 1/2)]\int_0^\infty e^{-xy-y^2/2}y^{c-1/2}dy \tag{12}$$

Numerous asymptotic expansions exist for evaluating $U(c,x)$. In particular the form given in Reference 5 in Section 19.12.3, in conjunction with the asymptotic expansion in Section 13.1.2, was used here for evaluating Equation (12).

The entire preceding development was undertaken on the assumption that the true state vector x_n is known. Specific-ally, x_n was used in obtaining z_n in Equation (4). This is

clearly not the case, and the estimate for x_n given by

$$\bar{x}_n = \phi_n \hat{x}_{n-1} \tag{13}$$

where $\hat{x}_{n-1}$ is the Kalman, Empirical Bayes filter estimate for x_{n-1} given in Reference 1, is used.

Now consider the estimates $\hat{\theta}_i$ required in Equation (9). The preceding estimates obtained from Equation (9) for $i = 1, 2, \ldots, n-1$, i.e., $\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_{n-1}$, can be substituted for $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{n-1}$ respectively when calculating $\tilde{\theta}_n$. Also, the maximum likelihood estimate for θ_n given by $1/z_n^2$ may be substituted for $\hat{\theta}_n$ in Equation (9) or an estimate of θ_n such as that provided by Reference 1 could be employed. For simplicity the maximum likelihood estimate was used here.

Performance Comparison

Monte Carlo simulation was employed to examine the average squared error loss incured in using the estimator given by Equation (9). For purposes of comparison the same dynamic system exemplified in Reference 1 was again used here. The squared estimation errors $(\tilde{\theta}_n - \theta_n)^2$ were averaged for all five observation types over twenty-five replications for the standard estimator presented in Reference 1 as well as for the estimator given here. The average squared errors are plotted in Figure 1 at every fifth observation stage for both types of estimators. We observe that except for the first five stages, the smooth empirical Bayes estimator has an

averaged squared error which is somewhat smaller than the standard empirical Bayes estimator. The same basic results were observed for numerous simulation runs using different sets of parameters. Thus it appears that the estimator in Equation (12) is somewhat superior in average squared error performance to the estimator in Reference 1 and any gains should, in theory at least, be passed on to the Kalman state estimation procedure presented in Reference 1.

References

1.   Martz, H.F. and Born, G.H., "Empirical Bayes Estimation of Observation Error Variances in Linear Systems", _AIAA Journal_, Vol. 9, No. 6, June 1971, pp. 1183-1187.

2.   Bennett, G.K. and Martz, H.F., "A Continuous Empirical Bayes Smoothing Technique", to appear in _Biometrika_.

3.   Martz, H.F. and Kamat, S.J., "Empirical Bayes State Estimation in Discrete Time Linear Systems", _AIAA Journal_, Vol. 9, No. 8, August 1971, pp. 1654-1656.

4.   Parzen, E., "On Estimation of a Probability Density Function and Mode", _Annals of Mathematical Statistics_, Vol. 33, 1962, pp. 1065-1076.

5.   Abramowitz, M. and Stegun, I.A., _Handbook of Mathematical Functions_, National Burequ of _Standards Applied Mathematics Series_, No. 55, 1964.

6.   Gradshteyn, I.S. and Ryzhik, I.M., _Table of Integrals, Series, and Products_, Academic Press, New York, 1965.