Information Content Subsetting of Highly Correlated Error Sources

I. M. WEISS*
Rockwell International, Anaheim, Calif.

Credible estimates of a large number of error model terms in missile system postflight guidance analysis are often not possible because of the quality of the data and the high functional correlation between various error sources. For this reason, an information content subsetting technique was devised which grouped the error terms into highly related sets and subsequently used a representative error term from each set in the estimation model. The subsets were chosen by maximizing an information content comparison criterion. The technique uses a criterion based on a measure of the theoretical information contained in the system model and the a priori statistics vs an earlier correlation coefficient technique which only used indications of the correlation in the system model. The relationship between the system models correlation, observability, and information content is discussed. An example illustrates the subsetting technique, and indicates that care must be taken in using theoretical estimation error covariances obtained from the filter.

I. Introduction

CREDIBLE estimates of a large number of error model terms in missile system postflight guidance analysis are often not possible because of the quality of the data and the high functional correlation between various error sources. For this reason, highly correlated error sources are often subsetted in missile system postflight guidance analysis.[1,2] The subsetting is necessary in order to simplify computations and to obtain more accurate results when unmodeled terms are present. Subsetting techniques group the errors into highly related sets and pick one representative error term from each set as a model for all terms of the group. Since the error terms are somewhat different, this procedure does result in modeling errors; however, the devised subsetting procedure will tend to minimize these effects. There are cases in missile system analysis where the grouping of terms into sets may result in significant errors because of the resulting incorrect apportionment of the set errors to the various set members. It is assumed that this difficulty is greatly reduced with the use of the available information by a competent data analyst.

Subsetting in Ref. 1 consisted of calculating the correlation coefficients between the error sensitivity functions of the various error sources. Those error sources which were found to be highly functionally correlated, e.g., 0.95 correlation, were put in the same subset and representative subset uncertainties and subset apportionment coefficients are calculated in the Appendix. Subsequent work calculated the correlation coefficients between the error sensitivity functions divided by the appropriate observation noise variance. This noise normalization procedure is applicable when the statistics of the noise are time varying and/or when indications of the error sources are obtained from different measurement information with their associated noise components, e.g., missile errors measured by different radar measurement channels. The noise normalization procedure will be theoretically justified in the ensuing work. The correlation coefficient methods of subsetting ignores the a priori statistics of the error sources. This paper develops a new subsetting technique based on an information content criterion, which takes into consideration the a priori statistics of the error source as well as their correlation. The information content criterion will give indications of the theoretical change in information or, equivalently, the change in estimation error uncertainty caused by subsetting, which is a more meaningful indicator than the correlation coefficients. The concept of observability will be shown to be related to the correlation of error sources. A scheme is developed which obtains candidate subsets using the correlation coefficient method and then checks these subsets by using an information content criterion. The information content function is made up of terms which are calculated in the correlation coefficient technique, thus making the subsetting extension computationally simple.

This paper consists of sections on problem description and correlation coefficient subsetting, information content and observability, information content subsetting, and an example illustrating the technique.

II. Problem Description and Correlation Coefficient Subsetting

The discrete linear system of interest is described as follows:

$$\mathbf{x}(k+1) = \phi(k+1, k)\mathbf{x}(k) \qquad k = 0, 1, \ldots, l \tag{1}$$

and the discrete linear observations are given as

$$\mathbf{y}(k) = M(k)\mathbf{x}(k) + \mathbf{v}(k) \qquad k = 1, \ldots, l \tag{2}$$

Where $\mathbf{x}(k)$ is the n-vector state at time $t(k)$, ϕ is a $n \times n$ nonsingular transition matrix, $\mathbf{y}(k)$ is the r-vector observation, M is an $r \times n$ nonrandom error propagation matrix, $\mathbf{v}(k)$ is an r-vector white Gaussian noise $\mathbf{v}(k) \sim G(0, R_k)$, $k = 1, \ldots, l$; and $\mathbf{x}_o \sim G(x_o, P_o)$, where $\mathbf{w} \sim G(a, B)$ indicates that the random vector $\mathbf{w}$ has a Gaussian distribution with mean vector $\mathbf{a}$ and covariance matrix $\mathbf{B}$. For the missile postflight guidance analysis problem, ϕ is the identity matrix and, therefore, we are estimating constant terms in the error model. The $\mathbf{y}(k)$ are comparisons between guidance telemetered data and metric tracking data. Regression techniques are used to estimate the error term's contributions to the comparisons. Let us rewrite Eq. (2), breaking the equation down into its component parts, in order to best describe the correlation coefficient subsetting technique for the guidance analysis problem

$$\begin{aligned} y_1(k) &= m_{11}(k)x_1(k) + m_{12}(k)x_2(k) + \cdots + m_{1n}(k)x_n(k) + v_1(k) \\ &\vdots \\ y_r(k) &= m_{r1}(k)x_1(k) + m_{r2}(k)x_2(k) + \cdots + m_{rn}(k)x_n(k) + v_r(k) \end{aligned} \tag{3}$$

Presented as Paper 73-867 at the AIAA Guidance and Control Conference, Key Biscayne, Fla., August 20–22, 1973; submitted August 20, 1973; revision received July 18, 1974. The author would like to acknowledge the work of C. M. Shipplett and R. H. Fuessel in this area, and the helpful discussions with S. M. Kawakami and L. H. Pinson, Autonetics personnel of Rockwell International.

Index categories: LV/M Flight Testing; Missile Systems; Tracking Systems.

* Member of Technical Staff—Marine Systems, Autonetics Group. Member AIAA.

and

$$x_1(k+1) = x_1(k) \qquad k = 1,\ldots,l$$
$$\vdots \tag{4}$$
$$x_n(k+1) = x_n(k)$$

where $y_i(k)$ is the ith scalar guidance-radar comparison at the time $t(k)$, $x_j(k)$ is the jth scalar component of $\mathbf{x}(k)$ vector with $x_j(0) \sim G[x_j(0), \sigma_{xj}^2]$, $m_{ij}(k)$ is the error sensitivity term for the jth component of the state vector and the ith radar comparison, and $v_j(k)$ is the noise on $y_j(k)$; uncorrelated white noise components with $v_j(k) \sim G[0, \sigma_j^2(k)]$.

An accuracy problem often occurs in the presence of unmodeled terms when two or more components of $\mathbf{x}(k)$ have error sensitivity terms which are functionally similar and highly correlated† for all guidance-radar comparisons. This is a problem which occurs often in practice and whose optimum solution would require taking additional measurements. Two reasons preventing this approach are the high cost of new instrumentation systems and the unavailability of sites whose geometry is sufficiently unique.[2] Another problem occurs when it is necessary, because of computational requirements, to reduce the number of estimated quantities. Because of these accuracy and/or computational problems, we often find it necessary to subset. High correlation between various error sources and the quality of the data comparisons often made it impossible to obtain credible estimates of error model terms. For the abovementioned reasons, we subset the error terms into highly correlated sets and pick one representative error term from each subset as a model for all terms of the subset.

The noise normalized correlation coefficient subsetting scheme calculates the correlation coefficients, $CC(i,j)$ between the ith and jth term of $\mathbf{x}(k)$ as follows:

$$CC(i,j) = N_{ij}/(D_{ij})^{1/2} \tag{5}$$

where

$$N_{ij} = \sum_{h=1}^{r}\sum_{k=0}^{l} \frac{m_{hi}(k)m_{hj}(k)}{\sigma_h^2(k)} \tag{6}$$

$$D_{ij} = \left(\sum_{h=1}^{r}\sum_{k=0}^{l} \frac{m_{hi}^2(k)}{\sigma_h^2(k)}\right) \times \left(\sum_{h=1}^{r}\sum_{k=0}^{l} \frac{m_{hj}^2(k)}{\sigma_h^2(k)}\right) \tag{7}$$

Those error sources which were found to be highly correlated, e.g., $CC(i,j) \geqq 0.95$, were put in the same subset and representative subset uncertainties and subset apportionment coefficients are calculated (see Appendix).[1] The apportionment coefficients assign portions of the calculated set errors to each of the set members. Section III will introduce the information concepts which are necessary in order to develop the new subsetting scheme.

III. Information Content and Observability

A measure of the information contained in the data and a priori statistics is desired for the problem formulation of Eqs. (1) and (2). Jazwinski,[3] while developing estimates for this problem formulation, obtains an information content function, which is described by the following equations:

$$IF_k \triangleq \phi^T(o,k)P_o^{-1}\phi(o,k) + \xi_{k,1} \tag{8}$$

$$\xi_{k,1} \triangleq \sum_{i=1}^{k} \phi^T(i,k)M^T(i)R_i^{-1}M(i)\phi(i,k) \tag{9}$$

where $\xi_{k,1}$ is the well-known information matrix for the data and IF_k is an information matrix[4] which includes a priori information.

Also

$$[P_k^k]^{-1} = IF_k \tag{10}$$

and

$$E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k^k)(\mathbf{x}_k - \hat{\mathbf{x}}_k^k)'\} = P_k^k \tag{11}$$

† For the purpose of subsetting "high correlation" refers to both high positive and negative correlation.

where $\hat{\mathbf{x}}_k^k$ is the Kalman-Bucy estimate of the state vector and P_k^k is the corresponding error covariance. Therefore, the inverse of Eq. (8) is the closed-form solution of the matrix Ricatti equation for the problem of Eqs. (1) and (2).

In regard to observability, one is asking the basic question of whether estimates of all the state vector components can be obtained from the data. This is equivalent to setting $P_o^{-1} = 0$ in Eq. (8), which means no weight is attached to the a priori statistics of x_o, and then looking at $\xi_{k,1}$. Then from,[3,5] the system is observable when the information matrix

$$\xi_{k,1} = \sum_{i=1}^{k} \phi^T(i,k)M^T(i)R_i^{-1}M(i)\phi(i,k) \tag{9a}$$

is positive definite. If $\xi_{k,1}$ is singular, then certain linear combinations of the elements of $\mathbf{x}(k)$ cannot be determined; there is no information about them in the data. For the guidance analysis problem, with ϕ as the identity matrix, perfect correlation between two of the error sensitivity functions, $CC(i,j) = 1$, will result in the $\xi_{k,1}$ matrix, Eq. (9), not being positive definite, i.e., the nonobservability condition. This is because perfect correlation results in two of the columns of the $\xi_{k,1}$ matrix being equal.

If none of the error sources are perfectly correlated, then theoretically the system will be observable. However, highly correlated error sources do create data analysis problems in the real world, e.g., divergence of estimates. The correlation coefficient subsetting method is an approach to handle the highly correlated error source problem. This technique does not take into consideration all of the information available for subsetting, i.e., a priori statistics; therefore some terms are subsetted which should not be. For this reason, a revised subsetting scheme which used a priori statistics was developed (Sec. IV) with a subsetting criterion based upon a change of information or equivalently a change in estimation error uncertainty.

A simple example, which illustrates the relationship of correlation to estimation error uncertainty, follows. The estimation error uncertainty P_k^k, the inverse of information, will be our comparison function for this report, since this quantity has physical significance to the analyst and results in a meaningful comparison criterion. For the example, two error sources are being estimated from l measurements of each of two guidance-radar comparisons with independent white observation noise.

Example

The equations describing the measurement system are as follows:

$$\begin{aligned} x_1(k+1) &= x_1(k) \\ x_2(k+1) &= x_2(k) \end{aligned} \qquad 0 \leqq k \leqq l \tag{12}$$

$$\begin{aligned} y_1(k) &= m_{11}(k)x_1(k) + m_{12}(k)x_2(k) + v_1(k) \\ y_2(k) &= m_{21}(k)x_1(k) + m_{22}(k)x_2(k) + v_2(k) \end{aligned} \qquad 1 \leqq k \leqq l \tag{13}$$

$$\begin{aligned} v_1(k) &\sim G(0, \sigma_1^2(k)) \\ v_2(k) &\sim G(0, \sigma_2^2(k)) \\ x_1(o) &\sim G(x_1, \sigma_{x_1}^2) \\ x_2(o) &\sim G(x_2, \sigma_{x_2}^2) \end{aligned} \tag{14}$$

where $v_1(k)$, $v_2(k)$, $x_1(o)$, $x_2(o)$ are all uncorrelated.

The information matrix IF_k is obtained from Eqs. (8, 9, and 12–14) as

$$IF_l = \begin{bmatrix} \dfrac{1}{\sigma_{x_1}^2} + \sum\limits_{h=1}^{2}\sum\limits_{k=1}^{l} \dfrac{m_{h1}^2(k)}{\sigma_h^2(k)} & \sum\limits_{h=1}^{2}\sum\limits_{k=1}^{l} \dfrac{m_{h1}(k)m_{h2}(k)}{\sigma_h^2(k)} \\ \sum\limits_{h=1}^{2}\sum\limits_{k=1}^{l} \dfrac{m_{h1}(k)m_{h2}(k)}{\sigma_h^2(k)} & \dfrac{1}{\sigma_{x_2}^2} + \sum\limits_{h=1}^{2}\sum\limits_{k=1}^{l} \dfrac{m_{h2}^2(k)}{\sigma_h^2(k)} \end{bmatrix} \tag{15}$$

and the corresponding x_1 and x_2 estimation error uncertainty terms of the P_k^k matrix are

$$P_{11}(l) = \frac{\left(1 + \sigma_{x_2}^2 \sum\limits_{h=1}^{2}\sum\limits_{k=1}^{l} \dfrac{m_{h2}^2(k)}{\sigma_h^2(k)}\right)}{DN} \sigma_{x_1}^2 \tag{16}$$

$$P_{22}(l) = \frac{\left(1+\sigma_{x_1}{}^2 \sum_{h=1}^{2}\sum_{k=1}^{l} \frac{m_{h1}{}^2(k)}{\sigma_h{}^2(k)}\right)}{DN}\sigma_{x_2}{}^2 \tag{17}$$

where

$$DN = \left\{\left[1+\sigma_{x_1}{}^2\sum_{h=1}^{2}\sum_{k=1}^{l}\frac{m_{h1}{}^2(k)}{\sigma_h{}^2(k)}+\sigma_{x_2}{}^2\sum_{h=1}^{2}\sum_{k=1}^{l}\frac{m_{h2}{}^2(k)}{\sigma_h{}^2(k)}\right]+ \sigma_{x_1}{}^2\sigma_{x_2}{}^2\left[\left(\sum_{h=1}^{2}\sum_{k=1}^{l}\frac{m_{h1}{}^2(k)}{\sigma_h{}^2(k)}\right)\times\left(\sum_{h=1}^{2}\sum_{k=1}^{l}\frac{m_{h2}{}^2(k)}{\sigma_h{}^2(k)}\right) - \left(\sum_{h=1}^{2}\sum_{k=1}^{l}\frac{m_{h1}{}^2(k)m_{h2}{}^2(k)}{\sigma_h{}^2(k)}\right)^2\right]\right\}$$

Note that the second bracketed term in the denominator (DN) is equal to

$$D_{12}-(N_{12})^2 \tag{18}$$

from Eqs. (6) and (7), and is greater than, or equal to zero. Equality is satisfied if the correlation coefficient $CC(i,j)=1$; therefore, with perfect correlation, no additional information is obtained due to the Eq. (18) term. When the correlation coefficient decreases, this correlation term in Eqs. (16) and (17) increases and the estimation uncertainties go down. Similar behavior can be observed in more complicated problems for the correlation effects on the estimation uncertainties.

The solution to the estimation problem exists even when the system is not observable as the a priori information enables the term in Eq. (8) to be inverted, a necessary manipulation in the estimation scheme. Therefore, the resulting estimation error equations, e.g., Eqs. (16) and (17) with $m_{11}=m_{12}$ and $m_{21}=m_{22}$ (perfect correlation), exist and will apportion the errors relative to a priori statistics. The difficulty in using this approach for apportioning is that the state dimension is not reduced, a characteristic which is desirable in guidance analysis, in order to reduce the cost of processing the comparison data, and in real-time onboard applications to reduce the computational load. Also, unmodeled error terms may result in a deterioration of estimation accuracy.

IV. Information Content Subsetting

The information content technique uses for an initial subsetting of error terms the correlation coefficient results and then checks to determine what the increase in information (decrease in error uncertainty) would be if these terms had not been subsetted. The information matrix calculation discussed in this section is computationally simple, since the matrix elements are the same as those necessary to calculate the correlation coefficients. In order to keep the scheme computationally simple, the information matrix comparisons will be made for one subset at a time, while the other sets are still in subsets obtained by the correlation coefficient technique. Otherwise, in the guidance analysis problem, too many matrix inverses of large order, e.g., 80 × 80, would be required. The criterion chosen for comparing the different error uncertainties from the various subsetting possibilities is an observation normalized trace, which is discussed in this section.

Procedure

1) Subset error terms using the correlation coefficient technique described in Sec. II, saving for the information matrix calculations the individually calculated summation terms in Eqs. (6) and (7).

2) Pick a representative error term for each subset, excluding the 1st subset, and calculate an equivalent error uncertainty for each subset (see Appendix), leaving all the 1st subset members in the state vector.‡

3) Calculate the information matrix for the formulation in Step 2 and take the inverse, obtaining the error uncertainty matrix.

4) Set the error sensitivity terms of all the 1st subset members in the information matrix of Step 3 equal to the value of the most significant subset member§ (see Appendix), and take the inverse of the resulting matrix to obtain the error uncertainty matrix.

5) Compare¶ the error uncertainty matrices of Steps 3 and 4 using a comparison criterion which will be developed, Eq. (20). If the information comparison is insignificant, form the 1st subset and proceed to test one of the remaining untested subsets.

6) If the comparison in 5 is significant, then calculate the error uncertainty matrix which occurs by placing back the actual sensitivity function of the 1st member of the 1st subset in the formulation of Step 4.

7) Repeat Step 6 for each member of the 1st subset, excluding the representative error term; note that the state vector for each iteration of Step 6 only contains the actual sensitivity function of one element of the 1st subset.

8) Compare the error uncertainty matrix of Step 3 with each of those calculated in Steps 6 and 7; those error terms from the subset which result in significant differences in the comparisons should be removed** from the subset.

9) Repeat Steps 2–8 for each of the remaining untested subsets.

Estimation Error Uncertainty

The information matrix at t_l is obtained for the general problem in a manner similar to that of Eq. (15), as follows:

$$\begin{bmatrix} \left(\frac{1}{\sigma_{x_1}{}^2}+\left(\sum_{h=1}^{r}\sum_{k=1}^{l}\frac{m_{h1}{}^2(k)}{\sigma_h{}^2(k)}\right)\right) & \cdots & \left(\sum_{h=1}^{r}\sum_{k=1}^{l}\frac{m_{h1}(k)m_{hn}(k)}{\sigma_r{}^2(k)}\right) \\ \left(\sum_{h=1}^{r}\sum_{k=1}^{l}\frac{m_{h1}(k)m_{h2}(k)}{\sigma_r{}^2(k)}\right) & \ddots & \vdots \\ \vdots & & \vdots \\ \left(\sum_{h=1}^{r}\sum_{k=1}^{l}\frac{m_{hn}(k)m_{h1}(k)}{\sigma_r{}^2(k)}\right) & \cdots & \left(\frac{1}{\sigma_{x_n}{}^2}+\left(\sum_{h=1}^{r}\sum_{k=1}^{l}\frac{m_{hn}{}^2(k)}{\sigma_h{}^2(k)}\right)\right) \end{bmatrix} \tag{19}$$

and, of course, from Eq. (10) the estimation error uncertainty $P_l{}^l$ is just the inverse of this matrix, noticing again that all the terms in Eq. (19) are obtainable from the summations in the correlation calculations of Eqs. (6) and (7).

Comparison Criterion

We want to compare two general estimation error covariance matrices P_a and P_b where the matrix elements are denoted as follows:

$$P_a = \{P_a{}^{ij}\} \qquad P_b = \{P_b{}^{ij}\}$$

The comparison of interest is of the diagonal elements of these matrices, since the diagonal elements are the estimation error uncertainties for the state vector components. A comparison of the traces of P_a and P_b appears to be meaningless since the individual diagonal terms have different dimensions and are,

‡ The state vector will now include representative error terms for all subsets excluding the 1st subset, nonsubsetted error terms remaining from Step 1, and the individual error terms from the 1st subset.

§ The a priori covariances of each of the 1st subset members must be normalized in order to maintain consistency of units with the substituted most significant subset member (see Appendix).

¶ This comparison checks to see if a significant information change would result from resubsetting the 1st subset.

** The procedure in Steps 4–8 may be repeated using another representative error term, and the composite information used to ascertain subset members.

thus, not compatible. A comparison criterion, observation normalized trace, which normalizes the diagonal terms in order to obtain a meaningful comparison is defined as follows:

$$CF = \left(\frac{f_1{}^2 P_a{}^{11} + \cdots + f_n{}^2 P_a{}^{nn}}{f_1{}^2 P_b{}^{11} + \cdots + f_n{}^2 P_b{}^{nn}} \right)^{1/2} \tag{20}$$

where f_i is the absolute value of the error sensitivity function from the error propagation matrix at the sample time t_l (see Appendix), which will convert the 1st error term uncertainty into an observation uncertainty in the units of the guidance-radar comparison.

This criterion, which is a weighted trace, compares the estimation errors as they would propagate into the observation comparisons and is appealing to the theorist as well as the data analyst. Since the estimation scheme minimizes a weighted trace,[5] the choice of this subsetting comparison criterion appears to have a theoretical justification.

The level of significance for the comparison will depend on the particular application involved. Values close to one for the comparison criterion of Eq. (20) CF will indicate that the theoretical information content of the two estimation systems being compared is similar, a desirable result.

V. Example

In order to demonstrate the information content subsetting scheme, a computer program was written which followed the procedure in Sec. IV. The example used was that of Eqs. (3) and (4) with five state vector components $x_i(k)$, $i = 1, \ldots, 5$, three comparison measurements $y_j(k)$, $j = 1, \ldots, 3$, and corresponding error sensitivity functions $m_{ji}(k)$ as indicated in Fig. 1. The a priori statistics of $x_i(k)$ were

$$\sigma_{x_{i=1,\ldots,5}} = 1, 3, 0.5, 1, 1$$

and the observation noise was

$$\sigma_{j=1,2,3} = 0.2, 0.3, 0.4, \text{respectively}$$

The coordinates of Fig. 1 are in units of feet per unit error of x_i and discrete time t_k. First, the correlation coefficients were calculated, $CC(i,j)$, from Eqs. (5–7), Step 1. The correlations between x_1 and x_2, x_3, x_4, x_5 are tabulated as follows:

$$CC(1,2) = -0.986$$
$$CC(1,3) = -0.947$$
$$CC(1,4) = 0.483$$
$$CC(1,5) = 0.423$$

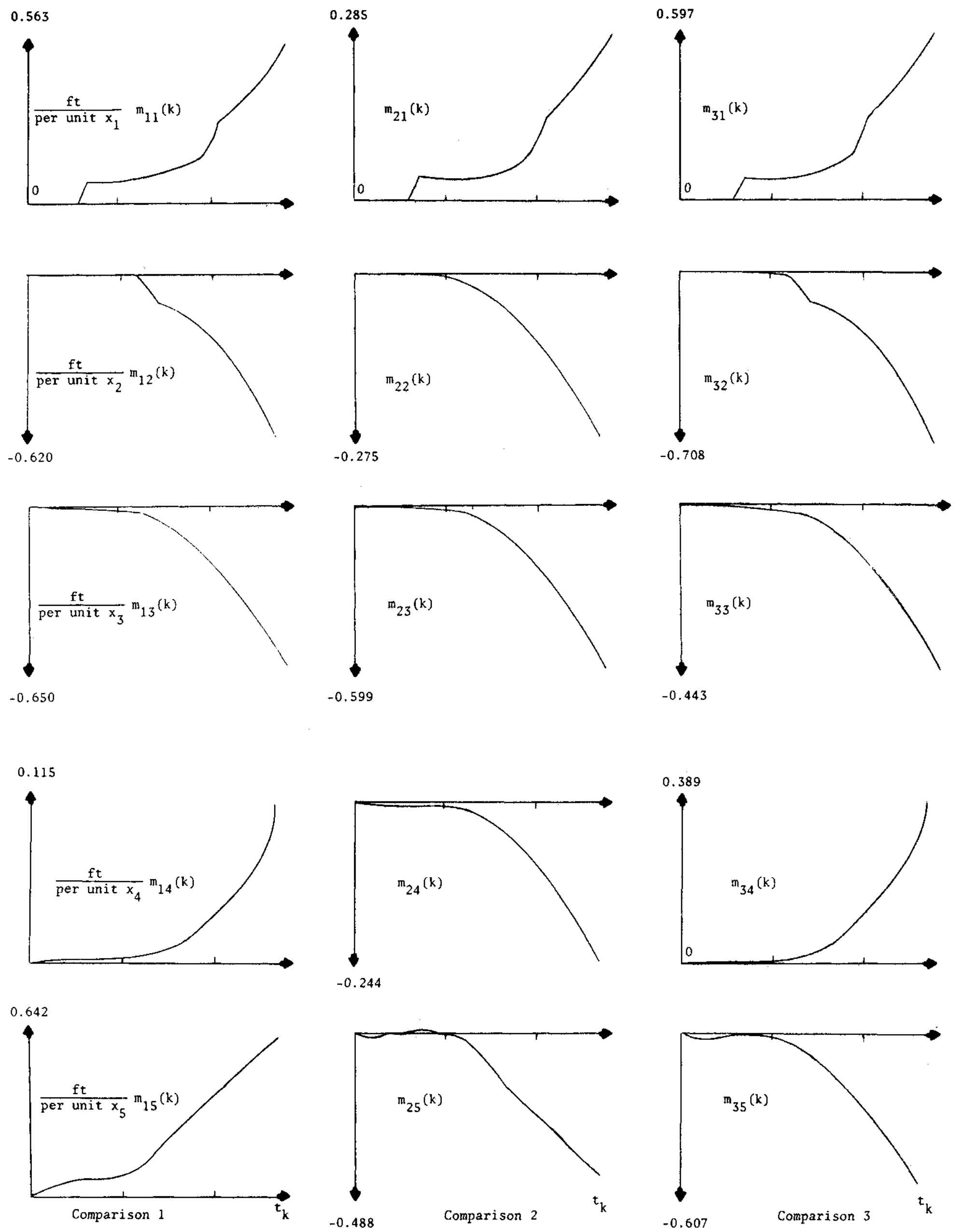

Fig. 1 Error sensitivity functions.

Table 1 Actual performance (rms estimation error)

	Proper model state vector contains x_1, x_2, x_3, x_4, x_5	Correlation coef. subset · state vector contains (x_1, x_2^*, x_3) x_4, x_5	Information content subset · state vector contains $x_1, (x_2^*, x_3)$ x_4, x_5	Nonsuggested subset · state vector contains (x_1, x_2^*) x_3, x_4, x_5
x_1	0.66	1.11	0.65	1.12
x_2	0.64	1.18	0.78	1.00
x_3	0.38	0.56	0.61	0.39
x_4	0.53	0.61	0.61	0.53
x_5	0.12	0.12	0.12	0.12

Table 3 Theoretical estimates vs actual performance

	Information content subset $x_1, (x_2^*, x_3), x_4, x_5$ Appendix method[a]	Information content subset $x_1, (x_2^*, x_3), x_4, x_5$ Step 6 method[a]	Actual performance $x_1, (x_2^*, x_3), x_4, x_5$ Monte Carlo
x_1	0.68	0.68	0.65
x_2	0.52	0.74	0.78
x_3	0.09	0.45	0.61
x_4	0.42	0.42	0.61
x_5	0.12	0.12	0.12

[a] Theoretical performance.

Therefore, the initial subset was chosen to contain x_1, x_2, and x_3.

The representative error term for the subset was chosen to be x_2 and the associated representative member normalization carried out. The comparison check of Step 5 ascertained that significant information change would result from resubsetting, i.e., the cost function (CF), Eq. (20), had a value of 0.83. The level of significance for CF in the example is chosen as 0.95. The cost function with x_1 replaced in the state vector was 0.98 and with x_3 replaced in the state vector, was 0.83, Steps 6–8. Therefore, it was decided to replace x_1 in the state vector. The resulting state vector would consist of x_1, x_4, x_5, and x_2 subsetted with x_3.

A Monte Carlo simulation of 100 trials was run in order to check the actual performance of the various subsetting possibilities and to demonstrate the subsetting scheme's capability. For each trial, Gaussian random numbers were selected for the x_i and the $v_j(k)$ random variables; and the observations, $y_j(k)$, were generated using these numbers. These observations were filtered, obtaining estimates of the state vector, using the various possible subsetted estimation systems. The results are indicated in Table 1. The column entries indicate for each run how the components of the state vector are modeled in the estimation system. The subsets are indicated by parentheses and the subset representative by an asterisk, e.g., (x_1, x_2^*, x_3), x_4, x_5 of column 2 means x_1, x_2, x_3 are subsetted with x_2 as their representative member and then this representative x_2 along with x_4 and x_5 are estimated. The subset estimate is then apportioned to all the subset members, the results listed in the column. The results in the table indicate actual performance for the suggested information content subset comparable to that of the proper model and better than the correlation coefficient method.

The theoretical estimates of subset error covariances (Appendix) were found to be unreasonably small as they are just composite numbers which do not take into consideration the difficulties associated with estimation in the presence of highly correlated error sources. However, the error covariance calculated in Step 6 turned out to be of a reasonable level. Tables 2 and 3 show the comparisons of these numbers for the subsets obtained for the correlation coefficient and information content techniques. The tables show that the calculated estimation error of Step 6 is closer to the actual filter performance obtained from the Monte Carlo simulation than that of the method described in the Appendix. Inherent in the overly optimistic error covariances of the Appendix's method is the possibility of divergent estimates of the state.[6]

Table 2 Theoretical estimates vs actual performance

	Correlation subset (x_1, x_2^*, x_3) x_4, x_5 Appendix method[a]	Correlation subset (x_1, x_2^*, x_3) x_4, x_5 Step 6 method[a]	Actual performance (x_1, x_2^*, x_3) x_4, x_5 Monte Carlo
x_1	0.04	0.81	1.11
x_2	0.11	0.92	1.18
x_3	0.02	0.45	0.56
x_4	0.41	0.41	0.61
x_5	0.12	0.12	0.12

[a] Theoretical performance.

VI. Conclusion

An improved technique for subsetting highly correlated error sources occurring in missile system postflight guidance analysis is obtained. The technique uses a criterion based on a measure of the theoretical information contained in the system model and the a priori statistics vs the correlation coefficient technique which only uses indications of the correlation in the system model. The information content technique is computationally simple in that it uses terms already calculated in the correlation coefficient technique in order to refine the subsets obtained from that method. A comparison criterion is devised in order to compare the theoretical information content of the various subset possibilities. An example illustrates the applicability of the technique and indicates that care must be taken in using theoretical estimation error covariances obtained from the filter.

Appendix: Selection of Representatives and Apportionment of Coefficients††

Once the subsets have been established, the next steps are to obtain a representative error function for the subset and to determine an a priori uncertainty for the representative.[1] Let the error sources in a given subset be numbered from 1 to n and let the representative be numbered zero. The unit error functions will then use $m_1(k)$ to $m_n(k)$ and $m_0(k)$. The a priori uncertainties will then be σ_{x_1} to σ_{x_n} and σ_{x_0}. The estimated coefficients will be denoted $\hat{x}_1$ to $\hat{x}_n$ and $\hat{x}_0$, while the estimated uncertainties are $\sigma_{\hat{x}_1}$ to $\sigma_{\hat{x}_n}$ and $\sigma_{\hat{x}_0}$. The known values are $m_1(k)$ to $m_n(k)$ and σ_{x_1} to σ_{x_n}. Then the regression procedure will provide $\hat{x}_0$ and $\sigma_{\hat{x}_0}$: we must determine a scheme which will yield $\hat{x}_1$ to $\hat{x}_n$ and $\sigma_{\hat{x}_1}$ to $\sigma_{\hat{x}_n}$.

Error sources which are in the same subset do not necessarily (or even usually) have identical units. Therefore, it is necessary to multiply the coefficients and uncertainties ($\hat{x}$, $\hat{\sigma}$, σ) by the corresponding unit error curves $[m(k)]$ to get all the data in tracker coordinates. Defining

$$\hat{y} = \hat{x}m(k)$$
$$\hat{\beta} = \hat{\sigma}m(k) \tag{A1}$$
$$\lambda = \sigma m(k)$$

then $\hat{y}$ is the estimated error propagation in tracker coordinates, while $\hat{\beta}$ and λ are the propagations of the estimation uncertainty and a priori uncertainty in tracker coordinates.

†† The apportionment scheme and following explanation was suggested by C. M. Shipplett.

The representative error function will be chosen as that which has the maximum λ_i value, i.e., pick the x_i which corresponds to the

$$\max_{i,k,h} |\sigma_{x_i} m_{h_i}(k)|$$

as the representative error function x_o; where $\hat{i}$, $\hat{k}$, $\hat{h}$ correspond to the maximizing values of these parameters and h corresponds to the index on the radar or coordinate measurement. Since all the error sources of interest at the present moment are in the same subset (highly correlated), the inference is that the ratio's of unit error functions are essentially the same at all times and in all coordinates. Therefore, for the purpose of subsetting the following parameters will be defined:

$$\begin{aligned} f_o &= |m_{\hat{h}\hat{i}}(\hat{k})| \\ f_j &= |m_{\hat{h}j}(\hat{k})| \end{aligned} \tag{A2}$$

The f_j's are scalars and are the absolute value of x_j error sensitivity function at time f_o occurs (t_l).

The a priori uncertainty for the representative error source must be such that it accounts for the uncertainties in all component error sources. These sources are assumed to be independent, so their a priori propagations into tracker coordinates can be root sum-squared and the following equation results:

$$\sigma_{x_o} = \frac{1}{f_o}\left(\sum_{i=1}^{n} \sigma_{x_i}{}^2 f_i^2\right)^{1/2} \tag{A3}$$

The scheme for apportioning coefficients required that the total error in tracker coordinates from the subset must be equal to the sum of the contributions of the members of the subset:

$$\hat{y}_o = \sum_{1}^{n} \hat{y}_i \tag{A4}$$

If $\hat{y}_i = w_i \hat{y}_o$, then the weighing functions, which are calculated below, must satisfy

$$\sum_{i=1}^{n} w_i = 1$$

Therefore

$$\hat{x}_i = w_i(f_o/f_i)\hat{x}_o \tag{A5}$$

and the apportionment of the estimation uncertainty is

$$\hat{\beta}_i = w_i \hat{\beta}_o$$

or

$$\hat{\sigma}_i = w_i(f_o/f_i)\hat{\sigma}_0 \tag{A6}$$

The apportionment of the estimation uncertainty must be interpreted carefully. It does not really represent the estimation uncertainty for the apportioned coefficients. That cannot be determined. It represents merely a subdivision of the total subset estimation uncertainty.

The method which has been chosen for apportionment is to make all component error sources have equally likely deviations; assuming a priori mean values of zero

$$\hat{y}_i/\lambda_i = \hat{y}_j/\lambda_j \tag{A7}$$

From this requirement, it can be determined that

$$w_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} = \frac{f_i \sigma_{x_i}}{\sum_{j=1}^{n} f_j \sigma_{x_j}} \tag{A8}$$

completing the derivation of Eqs. (A5) and (A6).

References

[1] Goodman, A. F., "Estimation of Coefficients in a Linear Model by Extended Iterative Weighted Least Squares," *Naval Research Logistic Quarterly*, Vol. 18, No. 2, June 1971, pp. 243–276.

[2] Rodems, R. C., "The Problem of Using Real Data is Statistical Filtering," *Proceedings of the IEEE Symposium on Nonlinear Estimation Theory and its Applications*, San Diego, Calif., 1970, pp. 70–90.

[3] Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[4] Van Trees, H. L., *Detection Estimation and Modulation Theory, Part I*, Wiley, New York, 1968.

[5] Nahi, N. E., *Estimation Theory and Application*, Wiley, New York, 1969.

[6] Fitzgerald, R. J., "Error Divergence in Optimal Filtering Problems," Second IFAC Symposium on Automatic Control in Space, Vienna, Austria, 1967.