[22] B. Friedlander and B. Porat, "The modified Yule-Walker method of spectral estimation," *IEEE Trans. Aerospace Electron. Syst.*, vol. AES-20, pp. 158–173, Mar. 1984.

A Chi-Square Test for Fault-Detection in Kalman Filters

B. D. BRUMBACK AND M. D. SRINATH

***Abstract*—A test for real-time detection of soft failures in navigation systems using Kalman filters has been proposed by Kerr. The test is based on the overlap between the confidence regions associated with two estimates, one obtained from a Kalman filter using on-line measurements, and the other based solely on *a priori* information. An alternate computational technique is presented which is based on constructing a chi-square test statistic from the difference between the two estimates and comparing it to a precomputed threshold. The chi-square test avoids the iterative computations required by the two-ellipsoid method for dimensions of two and higher.**

I. Introduction

Real-time detection of failures in estimation systems is of considerable importance in many problems [1]–[3]. In multisensor systems, multiple locally filtered estimates can be computed, with each estimate being dependent on a subset of the available sensors [4]. These local estimates can then be combined to obtain the global estimate of the system state, and the local estimates can serve as on-line spares. If subtle sensor failures occur that cannot be detected by sensor self-test, local-filter performance will be degraded, and overall system performance affected. An approach is therefore needed to determine the validity of a local-filter estimate computed from sensors that are subject to subtle failures.

We model the problem of failure detection as that of detecting a signal of unknown magnitude that occurs at an unknown time, and assume the following model for the system state $x(k)$ and the observation $z(k)$:

$$x(k+1)=\Phi(k+1,\ k)x(k)+B(k+1)q(k+1)+\upsilon\delta(k,\ \varphi)$$

$$z(k)=H(k)x(k)+r(k) \tag{1}$$

where $q(k)$ and $r(k)$ are independent, zero mean, Gaussian white sequences having covariances of intensity $Q(k)$ and $R(k)$, respectively. The initial state $x(0)$ is a Gaussian random vector independent of $q(k)$ and $r(k)$ and has mean x_0 and covariance P_0. The failure modes are represented as the random vector υ. The failure event is represented by the Kronecker delta $\delta(k, \varphi)$ which is unity for $k = \varphi$, where φ is the time at which the failure occurs, and zero otherwise.

A two-ellipsoid overlap test for failure detection in this system has been presented by Kerr in [5], [6]. Kerr's application was the detection of specific failure modes, such as large gyro bias excursions, in inertial navigation systems. Estimates of the instrument error states were tested to determine if a failure has occurred.

The specific application for which the chi-square test was developed is a multisensor navigation system employing multiple Kalman filters. For this application, it is not necessary to identify specific causes of system failure, but only to determine, in real-time, the validity of a filter output. The instrument-error states can be viewed as inputs driving the inertial-system output error state equations for position error, velocity error, and heading error. Consequently, position or velocity estimates (or both) are tested rather than instrument-error estimates. The alternate test presented here can be shown to be equivalent to the two-ellipsoid test for the one-dimensional case, and is readily extendable to two or more dimensions.

Manuscript received October 10, 1986; revised February 12, 1987.

B. D. Brumback is with General Dynamics, Fort Worth Division, Fort Worth, TX 76101.

M. D. Srinath is with the School of Engineering and Applied Science, Southern Methodist University, Dallas, TX 75275.

IEEE Log Number 8714465.

II. Two-Ellipsoid Test

The two-ellipsoid test proposed by Kerr [5], [6] uses two estimates: $\hat{x}_1(k)$, which is the estimate obtained using the measurements $z(k)$ via a Kalman filter, and the estimate $\hat{x}_2(k)$ which is computed from the *a priori* information only. The two estimates are obtained from the following sets of equations:

$$\hat{x}_1(k+1)=[I-G(k+1)H(k+1)]\Phi(k+1,\ k)\hat{x}(k/k)+G(k+1)z(k+1)$$

$$\hat{x}_1(0)=x_0$$

$$P_1(k+1/k)=\Phi(k+1,\ k)P_1(k/k)\Phi^T(k+1,\ k)+B(k+1)Q(k+1)B(k+1) \tag{2}$$

$$P_1(k+1)=[I-G(k+1)H(k+1)]P_1(k+1/k)$$

$$P_1(0)=P_0$$

$$G(k+1)=P_1(k+1/k)H^T(k+1)\cdot[H(k+1)P_1(k+1/k)H^T(k+1)+R(k+1)]^{-1}$$

and

$$\hat{x}_2(k+1)=\Phi(k+1,\ k)\hat{x}_2(k)$$

$$\hat{x}_2(0)=x_0 \tag{3}$$

$$P_2(k+1)=\Phi(k+1,\ k)P_2(k)\Phi^T(k+1,\ k)+B(k+1)Q(k+1)B(k+1)$$

$$P_2(0)=P_0.$$

The two state estimates and associated covariances define two Gaussian probability density functions (p.d.f.'s)

$$p_1(k)=p(x(k)/Z(k)) \tag{4}$$

where $Z(k)$ is

$$Z(k)=z(1),\ \cdots,\ z(k)$$

and

$$p_2(k)=p(x(k)/H_0) \tag{5}$$

which is the p.d.f. of the state conditioned on the hypothesis of no failure (H_0).

The first moments, $\hat{x}_1(k)$ and $\hat{x}_2(k)$, of these two p.d.f.'s may be considered as point estimates of $x(k)$ with uncertainty $P_1(k)$ and $P_2(k)$, respectively. Confidence boundaries can be placed about each estimate and the probability that the true state lies within the confidence regions is given by

$$\text{Prob }[x(k)\in R_1(k)/Z(k)]=\alpha_1$$

$$\text{Prob }[x(k)\in R_2(k)/Z(k)]=\alpha_2 \tag{6}$$

where $R_1(k)$ and $R_2(k)$ are the confidence regions of $\hat{x}_1(k)$ and $\hat{x}_2(k)$, respectively. In one dimension, the regions $R_1(k)$ and $R_2(k)$ are intervals on the real line. For higher dimensions, the confidence regions are ellipsoids. The boundaries of the ellipsoids C_1 and C_2 are the values of $x(k)$ for which

$$y_1(k)=K_1$$

$$y_2(k)=K_2$$

0018-9286/87/0600-0552$01.00

where

$$y_1(k)=(x(k)-\hat{x}_1(k))^T P_1^{-1}(k)(x(k)-\hat{x}_1(k))$$

$$y_2(k)=(x(k)-\hat{x}_2(k))^T P_2^{-1}(k)(x(k)-\hat{x}_2(k)).$$

The interiors of C_1 and C_2 are the values of $x(k)$ such that

$$y_1(k)\le K_1$$

$$y_2(k)\le K_2.$$

The probability that the true state is interior to the ellipsoids is described as

$$\text{Prob }[y_1(k)\le K_1/Z(k)]=\alpha_1$$

$$\text{Prob }[y_2(k)\le K_2/H_0]=\alpha_2. \tag{7}$$

As noted earlier, the estimate $\hat{x}_1(k)$ reflects the on line measurements $Z(k)$ which indicate the actual situation of H_0 (no failures) or H_1 (failure) processed by a Kalman filter that assumes H_0. The estimate $\hat{x}_2(k)$ reflects only the *a priori* information (no on line data) and assumes H_0. As long as the two confidence regions $R_1(k)$ and $R_2(k)$ overlap, the true state may be in both confidence regions, and it is reasonable to conclude that no failures have occurred. If the two confidence regions do not overlap, the true state cannot be in both regions simultaneously and a failure is declared.

For one dimension, the confidence regions are intervals and the test for overlap is direct. The confidence interval associated with $p_1(k)$ is

$$\hat{x}_1(k)\pm n(k)(P_1(k))^{1/2}$$

and the confidence interval associated with $p_2(k)$ is

$$\hat{x}_2(k)\pm n(k)(P_2(k))^{1/2}$$

where $n(k)$ is chosen to meet a prespecified probability of false alarm [7]. The test for overlap of the two intervals is

$$|\beta(k)|<n(k)\{[P_1(k)]^{1/2}+[P_2(k)]^{1/2}\} \tag{8}$$

where

$$\beta(k)=\hat{x}_1(k)-\hat{x}_2(k). \tag{9}$$

For two or more dimensions the test for overlap is not as simple. Iterative computations are required and details are presented in [5], [6].

The two-ellipsoid overlap test depends on the estimation errors $e_1(k)$ and $e_2(k)$ being Gaussian. As long as the filter is linear, even if it is suboptimal, these errors will be Gaussian since the noise processes $q(k)$ and $r(k)$ are Gaussian. The Gaussian noise processes are present under both no failure and failure conditions. For a suboptimal filter, this test requires that the filter covariances $P_1(k)$ and $P_2(k)$ be accurately interpretable as the covariances of the estimation errors $e_1(k)$ and $e_2(k)$, respectively. This is necessary for the actual implementation of Kerr's test to attain the false-alarm and detection probabilities predicted by analysis. This approach to failure detection does not depend on the residuals being white in the unfailed mode.

The two-ellipsoid overlap test is a long-term test for failure detection. The two confidence regions defined by the two estimates and associated error covariances reflect all expected system behavior since filter initialization (which is also system "start time"). The measurements reflect the actual situation with respect to hypothesis H_0 (failure) or H_1 (no failure). If a failure has occurred its effect will eventually be detectable via the measurements $z(k)$. Since the estimate $\hat{x}_1$ includes information contained in the measurements and $\hat{x}_2$ does not, any failure that causes the two confidence regions to become disjoint will eventually be detected. There is no way that the test will "track" the error caused by a failure rather than detect it. This characteristic makes the two estimate approach well-suited to fault-detection in systems subject to "soft" failures.

III. An Alternate (Chi-Square) Test for Failure Detection

The two-ellipsoid overlap test can be viewed as a geometric method to determine when two estimates of the same parameter, $x(k)$, agree within the expected uncertainties of the estimates. An alternate approach to test agreement of two estimates is to consider the distribution of $\beta(k)$ defined by (9). The result is a closed form, even for the multidimensional case.

Define the estimation errors $e_1(k)$ and $e_2(k)$ as

$$e_1(k)=\hat{x}_1(k)-x(k)$$

$$e_2(k)=\hat{x}_2(k)-x(k). \tag{10}$$

Now (9) can be written as

$$\beta(k)=e_1(k)-e_2(k). \tag{11}$$

Since each filter is linear, each estimate is unbiased, so that

$$E\{\beta(k)\}=E\{e_1(k)-e_2(k)\}=0. \tag{12}$$

The covariance of $\beta(k)$

$$\begin{aligned}B(k)&=E\{\beta(k)\beta^T(k)\}\\&=E\{e_1(k)e_1^T(k)-e_1(k)e_2^T(k)\\&\quad-e_2(k)e_1^T(k)+e_2(k)e_2^T(k)\}\\&=P_1(k)+P_2(k)-P_{12}(k)-P_{12}^T(k)\end{aligned} \tag{13}$$

where

$$P_{12}(k)=E\{e_1(k)e_2^T(k)\}=P_{21}^T(k). \tag{14}$$

Finally, $\beta(k)$ is Gaussian since it is the linear combination of two Gaussian random variables $e_1(k)$ and $e_2(k)$. Since $\beta(k)$ is Gaussian with zero mean and covariance $B(k)$, given by (13), its distribution is completely defined.

A test for the occurrence of a failure consists of comparing the scalar test statistic

$$\lambda(k)=\beta^T(k)B^{-1}(k)\beta(k) \tag{15}$$

to a constant threshold. The test statistic $\lambda(k)$ is chi-square distributed with n degrees of freedom. The test for failure detection is

$$\lambda(k)>K_2 \qquad \text{failure } (H_1)$$

$$\lambda(k)\le K_2 \qquad \text{no failure } (H_0) \tag{16}$$

where the threshold K_2 is determined from tables of the chi-square distribution and

$$\text{Prob }[\lambda>K_2/H_0]=P_{fa} \tag{17}$$

so that constant probability of false alarm P_{fa} is maintained.

In order to compute the test statistic $\lambda(k)$, it is necessary to determine $B(k)$. Equation (13) requires the cross-covariance $P_{12}(k)$ between the errors in the two estimates $\hat{x}_1(k)$ and $\hat{x}_2(k)$. The errors in the two estimates are correlated because of the initial condition and the system process noise, which is a common input to both error processes. The time evolution of the two errors is described as

$$e_1(k+1/k)=\Phi(k+1,\ k)e_1(k/k)-B(k+1)q(k+1)$$

$$e_2(k+1/k)=\Phi(k+1,\ k)e_2(k/k)-B(k+1)q(k+1). \tag{18}$$

When the filter processes a measurement, the error in $\hat{x}_1$ changes according to

$$e_1(k+1/k+1)=[I-G(k+1)H(k+1)]e_1(k+1/k)+G(k+1)z(k+1). \tag{19}$$

The error in $\hat{x}_2$ does not depend on the measurements, so

$$e_2(k+1/k+1)=e_2(k+1/k). \tag{20}$$

The covariance of $e(k) = [e_1(k)e_2(k)]^T$ is

$$E\{e(k)e^T(k)\} = P(k) = \begin{pmatrix} P_1(k) & P_{12}(k) \\ P_{12}^T(k) & P_2(k) \end{pmatrix}. \tag{21}$$

The covariance, $P(k)$, evolves as

$$P(k+1/k+1) = \Psi P(k/k)\Psi^T + \Gamma Y(k+1)\Gamma^T \tag{22}$$

where

$$\Psi = \begin{pmatrix} [I-G(k+1)H(k+1)]\Phi(k+1, k) & 0 \\ 0 & \Phi(k+1, k) \end{pmatrix}$$

$$\Gamma = \begin{pmatrix} -[I-G(k+1)H(k+1)]B(k+1) & G(k+1) \\ -B(k+1) & 0 \end{pmatrix}$$

$$Y(k+1) = \begin{pmatrix} Q(k+1) & 0 \\ 0 & R(k+1) \end{pmatrix}.$$

The initial conditions for each estimate are

$$\hat{x}_1(0) = \hat{x}_2(0) = E\{x(0)\} = x_0$$

and the initial covariance $P(0)$ is

$$P(0) = \begin{pmatrix} P_0 & P_0 \\ P_0 & P_0 \end{pmatrix}. \tag{23}$$

Equation (22) can be expanded to yield the transition equations for the submatrices of $P(k)$ shown in (21). For the optimal filter gain, identical state models for the two estimates $\hat{x}_1$ and $\hat{x}_2$, and the initial covariance of (23) it can then be shown that

$$P_{12}(k) = P_1(k), \qquad k = 0, 1, 2, \cdots. \tag{24}$$

Substituting (24) into (13) yields

$$B(k) = P_2(k) - P_1(k). \tag{25}$$

Only $P_1(k)$ and $P_2(k)$ must be computed in order for the chi-square test statistic $\lambda(k)$ to account for the correlation between the errors $e_1(k)$ and $e_2(k)$. Two n-state covariance models can be computed instead of the single $2n$-state covariance model of (22).

The test statistic $\lambda(k)$ is defined by (15). However, it can be computed using the Cholesky decomposition [7],

$$B(k) = LL^T$$

by solving a triangular system of equations, and computing $\lambda(k)$ as an inner product of a vector with itself.

IV. Conclusions

The two-estimate technique developed by Kerr is suitable for use as a local-filter self-test that determines the validity of the filter estimate and covariance model by computing two estimates of the system state, one based on on line data and another based solely on *a priori* information. This technique is a long-term test, since at each check-time all integrated effects since system start-time are considered. The long-term aspect suggests that the two-estimate technique is well-suited to failure detection with sensors that are subject to soft failures, such as instrument bias shifts that require time to integrate into the measured variables. The test can also be used with suboptimal filters since it does not depend on the residuals being white in the unfailed case.

For the particular case where the two estimates are computed based on the same model of the system state, and assuming the filter is optimal, the two covariances completely define the joint covariance of the errors in the two estimates. The threshold for the chi-square test can be obtained from tables of the chi-square distribution, chosen according to the number of degrees of freedom in the chi-square variable. The chi-square test yields a closed form for two and higher dimensions, and is computationally straightforward.

Finally, it is important to note that the chi-square test requires knowledge of the cross-covariance of the errors in the two estimates. For the case where the two estimates are the Kalman filter estimate and the time-extrapolated *a priori* estimate, both computed assuming the same state model, the cross-covariance is the same as the filter covariance (and, hence, does not require additional computation). An interesting application is the comparison of two Kalman filter estimates, each based on different measurements, to determine if they agree within the confidence limits for both. It is possible to compute the cross-covariance of the errors in these two estimates, however, Kerr's two-ellipsoid overlap test may require less computational effort and provide a framework for approximation methods.

References

[1] A. S. Wilsky, "A survey of design methods for failure detection in dynamic systems," *Automatica*, vol. 12, pp. 601-611, 1976.

[2] C. B. Chang and K. P. Dunn, "On GLR detection and estimation of unexpected inputs in linear discrete systems," *IEEE Trans. Automat. Contr.*, vol. AC-24, no. 3, pp. 499-501, 1979.

[3] E. Y. Chow, "A failure detection system design methodology," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA.

[4] B. D. Brumback, "A decentralized fault-tolerant approach to estimation in multisensor navigation systems," Ph.D. dissertation, Southern Methodist Univ., Dallas, TX.

[5] T. H. Kerr, "Real-time failure detection: A nonlinear optimization problem that yields a two-ellipsoid overlap test," *J. Optimiz. Theory Appl.*, vol. 22, no. 4, pp. 509-536, 1977.

[6] ——, "Statistical analysis of a two-ellipsoid overlap test for real-time failure detection," *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 4, pp. 762-773, 1980.

[7] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*. New York: Academic, 1977.

Stochastic Teams with Nonclassical Information Revisited: When is an Affine Law Optimal?

RAJESH BANSAL AND TAMER BAŞAR

Abstract—**In this note we consider a parameterized family of two-stage stochastic control problems with nonclassical information patterns, which includes the well-known 1968 counterexample of Witsenhausen. We show that whenever the performance index does not contain a product term between the decision variables, the optimal solution is linear in the observation variables. The parameter space can be partitioned into two regions in one of which the optimal solution is linear, whereas in the other it is inherently nonlinear. Extensive computations using two-point piecewise constant policies and linear plus piecewise constant policies provide numerical evidence that nonlinear policies may indeed outperform linear policies when the product term is present.**

I. Introduction

In the context of team decision theory it is now customary to distinguish problems on the basis of their information structure, which may be of the classical, quasi-classical, or nonclassical type.

Manuscript received February 3, 1986; revised October 29, 1986. Paper recommended by Past Associate Editor, R. R. Tenney. This work was supported in part by the U.S. Air Force Office of Scientific Research under Grant AFOSR-84-0056.

The authors are with the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801.

IEEE Log Number 8613536.