AIAA 82-4252

Maximum Likelihood Failure Detection of Aircraft Flight Control Sensors

Bernard Friedland*
The Singer Company, Little Falls, New Jersey

A recently developed maximum likelihood failure detection algorithm, based on processing the residuals from an observer designed on the assumption of no failures, is applied to the problem of detecting and correcting gyro or accelerometer failures in the vertical axis of an aircraft flight control system. Design calculations are illustrated with a simplified dynamic model and the effectiveness of the failure detection method is demonstrated by simulation.

Introduction

A NEW approach to failure detection and correction in dynamic systems, frequently known as "analytic redundancy," has been burgeoning over the past few years. In contrast to hardware redundancy—multiple copies of sensors and actuators—analytical redundancy exploits the relationships between different variables in a dynamic system to allow different sensors (or actuators) to serve as backups to each other. Thus, components normally present in a control system can provide a low-cost alternative to redundancy achieved by replicating hardware.

Various algorithms for implementing analytical redundancy have been described in the literature.[1] One class of such algorithms makes use of the "residuals" ("innovations") of an observer F_0 (e.g., a Kalman filter), designed under the hypothesis of no component failure. These residuals are processed by an auxiliary unit, the purpose of which is to detect, identify, and correct for failures. Most of the algorithms[2-4] are based on the sequential likelihood ratio test proposed by Wald[5] and its generalizations.

A different formulation leading to a new algorithm was proposed by Friedland[6-8] for failures in the form of a step change in a bias. Many types of failures, including, for example, ramps (integrals of steps), can be represented as step changes. But, of course, there are other failures (e.g., sudden increase in random noise) which do not fall into this class and for which other methods might be more appropriate.

The algorithm introduced by Friedland also makes use of the residuals of an observer F_0 designed under the hypothesis of no failure, but uses a (simplified) maximum likelihood filter F_1 to process the observer residuals. The second filter, F_1, has the recursive structure of a Kalman filter, but it contains a nonlinear threshold comparison operation. When a particular signal in F_1 known as a "trigger" crosses a corresponding threshold, this is taken as an indication of a failure, i.e., the occurrence of a step change in bias. The gains of F_1 are automatically adjusted from those of a low bandwidth system to those of a high bandwidth system, so that the transition can be rapidly tracked. The gains gradually reduce the bandwidth of F_1 until such time as another step change in bias occurs. Thus it is possible to detect and track several changes in a bias, provided that the occurrences of such changes are relatively infrequent.

The performance of this failure detection strategy in a dynamic system was demonstrated by simulation with a simple example in an earlier paper.[8] The present paper is intended as another demonstration of the capability of this technique, within the context of aircraft flight control. The example has been chosen to faciliatate exposition and to make it possible for the reader to verify the design calculations. The object of this paper is to illustrate the possibility of detecting failures in more than one sensor and identifying the failed sensor, so the use of two sensors, a rate gyro and a normal accelerometer, is presumed, even though the particular performance objective—increased damping—could be achieved with the rate gyro alone.

The investigation reported here is confined to demonstrating the identification of a failed sensor and estimation of the bias change caused by the failure. The issue of how to use the data provided by the algorithm is not addressed here. Since the sensor bias changes can be tracked, it presumably would be possible to continue using a "failed" sensor in a closed-loop system. It probably would be unwise to do so, however, since a bias-shift failure may augur further failures in the sensor shortly thereafter. It would seem prudent to switch the sensor out of the closed-loop system while monitoring its performance. The entire question of closed-loop reconfiguration strategy deserves and is receiving considerable attention.

Process Model

The dynamics of the aircraft in our model are characterized by two state variables; q = pitch rate and α = angle-of-attack, which are assumed to obey the following differential equations:[9]

$$\dot{q} = -\omega_0^2(\alpha - \bar{Q}\delta) \tag{1}$$

$$\dot{\alpha} = -\alpha/\tau + q \tag{2}$$

where ω_0, $\bar{Q}$, and τ are taken to be constant for the purpose of this study. The control input is the deflection of the elevator δ (from the trim condition).

For illustrative purposes we assume that there are two sensors: a rate gyro which measures the pitch rate and an accelerometer which measures the specific force normal to the longitudinal axis of the aircraft (see Fig. 1):

$$a_N = \bar{V}\dot{\gamma} = V(q - \dot{\alpha})$$

(It is noted that it is not necessary to have both a rate gyro and a normal accelerometer for stability augmentation alone; a rate gyro would be perfectly adequate for this purpose. But the normal accelerometer might well be present for other purposes. And, of course, other combinations of sensors (air

Received Jan. 1982; revision received May 28, 1982.

*Manager, Systems Research, Kearfott Division. Member AIAA.

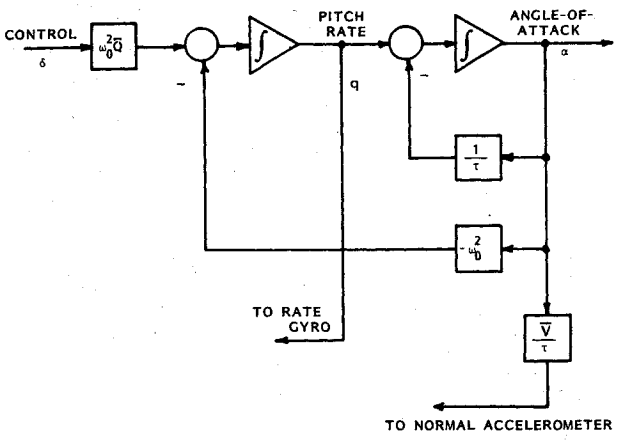

Fig. 1 Aircraft dynamic model.

data, attitude gyros, etc.) used in aircraft flight control systems can be treated by this method.)

Thus, from Eq. (2),

$$a_N = \frac{\bar{V}}{\tau}\alpha$$

and the normal acceleration is proportional to the angle-of-attack α. The scale factor $\bar{V}/\tau$ can be absorbed in the sensor output conversion. Hence the observation equations can be taken as

$$q_m = y_1 = q + b_1 + w_1 \tag{3}$$

$$\alpha_m = y_2 = \alpha + b_2 + w_2 \tag{4}$$

The quantities b_1 and b_2 are the sensor biases (the variables subject to sudden change), and w_1 and w_2 represent white measurement noise, which is assumed to be present at all times.

The dynamics and observations can be expressed in the standard vector-matrix form

$$\dot{x} = Ax + B\delta \tag{5}$$

$$y = Cx + Db + w \tag{6}$$

where

$$x = \begin{bmatrix} q \\ \alpha \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

and

$$A = \begin{bmatrix} 0 & -\omega_0^2 \\ 1 & -1/\tau \end{bmatrix} \qquad B = \begin{bmatrix} \omega_0^2 \bar{Q} \\ 0 \end{bmatrix} \tag{7}$$

$$C = D = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{8}$$

Using this dynamic model, our objective will be to illustrate the calculations needed to design the failure detector and corrector and to demonstrate the performance that can be achieved. This is accomplished in the next two sections.

State Estimation and Failure Detection

Bias-Free State Estimator, F_0

The general theory presented in Refs. 5-7 calls for the design of an observer under the hypothesis that the bias vector b is identically zero and the use of the residuals (innovations) of this filter to drive the failure detector.

The general form of the observer for the process defined by Eqs. (5) and (6) is

$$\dot{\hat{x}} = A\hat{x} + B\delta + Kr \tag{9}$$

where $r = y - C\hat{x}$ is the residual that drives the failure detector.

If the excitation and observation noise spectral density matrices, Q and W, respectively, are known, the observer can be optimized for these noise statistics, and, in this case, the observer becomes a Kalman filter with the gain matrix given by

$$K = PC'W^{-1} \tag{10}$$

where

$$P = AP + PA' - PC'W^{-1}CP + Q \tag{11}$$

As formulated in Eq. (5) no white noise input to the process is assumed. This leads to a covariance matrix P that tends to zero and takes K along with it. The modeling assumption can be changed to include some "process" noise which would produce nonzero steady-state covariance and gain matrices.

The present method, however, does not require that the observer be a Kalman filter; it works for any gain matrix K that produces a stable observer, i.e., for which the eigenvalues of A-KC have negative real parts. In particular, the observer may be a "detection filter" introduced by Beard,[10] the gains of which are designed to facilitate identification of failures. The approach used here embraces the latter viewpoint. Specifically, the gain matrix is selected to meet the following conditions:

1) The gain matrix K must be consistent with a positive-definite covariance matrix, i.e., there must exist a positive-definite spectral density matrix W such that the covariance matrix P is positive-definite. This is easy to accomplish, because from Eq. (10), with $C=I$, $P=KW$. A diagonal gain matrix, K will satisfy this condition for a diagonal spectral density matrix.

2) The resulting observer must be asymptotically stable. This condition is achieved with a diagonal gain matrix

$$K = \begin{bmatrix} K_1 & 0 \\ 0 & K_2 \end{bmatrix} \tag{12}$$

The observer characteristic equation is

$$\left| sI - A_K \right| = 0 \tag{13}$$

with

$$A_K = A - KC = A - K = \begin{bmatrix} -k_1 - \omega_0^2 \\ 1 \quad -1/\tau - K_2 \end{bmatrix} \tag{14}$$

Thus, Eq. (13) becomes

$$(s+K_1)\left(s + \frac{1}{\tau} + K_2\right) + \omega_0^2 = s^2 + \left(K_1 + K_2 + \frac{1}{\tau}\right)s + \omega_0^2 + K_1\left(K_2 + \frac{1}{\tau}\right) = 0 \tag{15}$$

and clearly the observer is asymptotically stable for any nonnegative K_1 and K_2.

3) The residuals generated by the observer should be such as to permit identification of the instrument that fails. This requirement needs elucidation. In general, a sudden transition in the bias of one sensor will create a step change in more than one of the residuals. It is desirable, however, to get a different

failure identification for each sensor failure. In the present (multidimensional) failure detection algorithm, the occurrence of a failure is signalled by one or more components of a "trigger" vector z exceeding a threshold. In general, more than one component of z will exceed the threshold when a failure in a single sensor occurs. The trigger pattern for each sensor failure is determined by a matrix S, which is related to the equivalent observation matrix H as follows

$$S=H'(HFH'+W)^{-1} \tag{16}$$

where F is a matrix, resembling a covariance matrix that appears in the failure detection algorithm. It is shown in Ref. 7 that a different component of the trigger vector is excited for each different bias transition if the matrix S satisfies the condition

$$SH=\Lambda=\text{diag}[\lambda_1,\ldots,\lambda_k] \tag{17}$$

The matrix F tends to zero if the time between failures is long relative to the dynamics of the bias estimator. In that case Eq. (17) becomes

$$H'W^{-1}H=\Lambda \tag{18}$$

Thus, selecting the gain matrix K so that the equivalent observation matrix H satisfies Eq. (18), establishes a one-to-one correspondence between components of the trigger vector z and instrument failures. The "detection filter"[10] performs the same function.

It is shown in Ref. 11 that the equivalent observation matrix H is given by

$$H=CV+D=V+I \quad \text{[by virtue of Eq. (8)]} \tag{19}$$

where, in the steady state,

$$(A-KC)V=KC$$

or, in this application,

$$(A-K)V=K \tag{20}$$

From Eqs. (19) and (20) we obtain

$$H=(A-K)^{-1}A \tag{21}$$

Thus, to satisfy Eq. (18), select the gain matrix K such that

$$A'(A'-K')^{-1}W^{-1}(A-K)^{-1}A=\Lambda \tag{22}$$

This is the third condition that must be satisfied by the gain matrix K. It turns out, to our delight, that this condition can also be satisfied by a diagonal matrix. In particular, using Eq. (21), we find that

$$H=\frac{1}{\omega_0^2+K_1\left(\frac{1}{\tau}+K_2\right)}\begin{bmatrix}\omega_0^2 & \omega_0^2K_2\\ -K_1 & \omega_0^2+K_1/\tau\end{bmatrix} \tag{23}$$

And, for

$$W=\begin{bmatrix}w_1 & 0\\ 0 & w_2\end{bmatrix}$$

Eq. (22) becomes

$$\frac{1}{\omega_0^2+K_1^2\left(\frac{1}{\tau}+K_2\right)^2}\left[\begin{array}{c|c}\omega_0^4/w_1+K_1^2/w_2 & \omega_0^4K_2/w_1-K_1(\omega_0^2+K_1/\tau)/w_2\\ \hline \omega_0^4K_2/w_1-K_1(\omega_0^2+K_1/\tau)/w_2 & \omega_0^4K_2^2/w_1+(\omega_0^2+K_1/\tau)^2/w_2\end{array}\right]=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$$

To satisfy this condition, it is necessary that the off-diagonal elements of the matrix be zero, i.e.,

$$\frac{w_1}{w_2}=\frac{\omega_0^2K_2}{K_1(\omega_0^2+K_1/\tau)} \tag{24}$$

The matrix H of Eq. (23) and the spectral density matrix W are input data to the failure detection algorithm. The gains, K_1 and K_2, can be selected as desired. Then, by selecting w_1 and w_2 to satisfy Eq. (24), the failure identification requirement, Eq. (18), is achieved. Alternatively, if the spectral density matrix W is used as the starting point, then Eq. (18) is met by selecting the gains in accordance with Eq. (24).

As a numerical example, we assume the open-loop aircraft dynamics are characterized by $\omega_0=10$ rad/s and $\tau=0.5$ s. This corresponds to a "damping ratio" ζ of 0.1. We would like the bias-free filter F_0 to have a higher natural frequency and greater damping, such as would result for a filter characteristic equation

$$(s+\omega_1)^2+\omega_1^2=s^2+2\omega_1 s+2\omega_1^2=0$$

which gives a damping ratio of $1\sqrt{2}$. From Eq. (15) we find that

$$2\omega_1=K_1+K_2+1/\tau$$

$$2\omega_1^2=\omega_0^2+K_1(K_2+1/\tau)$$

Selecting $\omega_1=20$, we find that $K_1=10$ and $K_2=8$, and, for these gains, we obtain, using Eqs. (23) and (24),

$$H=\begin{bmatrix}0.5 & 4.0\\ -0.05 & 0.6\end{bmatrix}$$

and

$$\frac{w_1}{w_2}=66.67$$

The bias-free filter with the diagonal gain matrix satisfies the following differential equations:

$$\dot{\tilde{q}}=-\omega_0^2(\tilde{\alpha}-\tilde{Q}\delta)+K_1(y_1-\tilde{q})$$

$$\dot{\tilde{\alpha}}=-\tilde{\alpha}/\tau+K_2(y_2-\tilde{\alpha})$$

These are shown in the form of an analog computer diagram in Fig. 2, which also shows the residuals as the input to the failure detector.

Simulated behavior of the aircraft in response to white noise on δ and the corresponding observations and state estimates of the state estimator under the hypothesis of no failure are shown in Figs. 3 and 4. Failures in the form of bias shifts are assumed to occur in the normal accelerometer at $t=2$ s and in the rate gyro at $t=4$ s. These are indicated as jumps on the observed quantities in the figures. The response of the filter to these inputs is shown on the traces labeled $\tilde{\alpha}$ and $\tilde{q}$. Since the filter is designed under the hypothesis of no failure, it is oblivious to the occurrence of the transitions. Thus, the outputs of the filter $\tilde{\alpha}$ and $\tilde{q}$ are "contaminated" by these transitions. The effect of this is more evident in the pitch

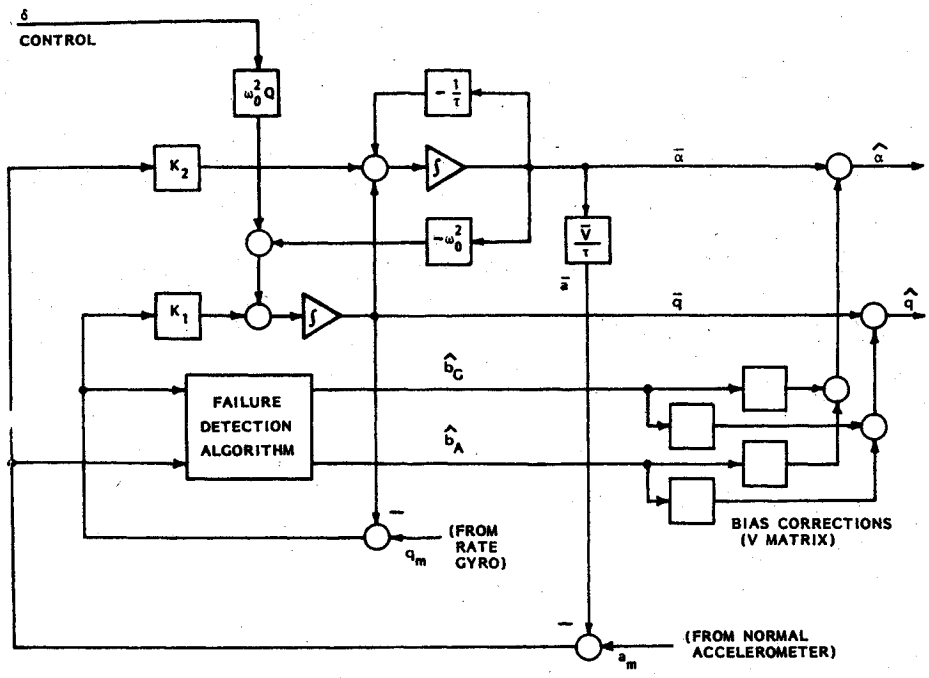

Fig. 2 Filter and failure detector.

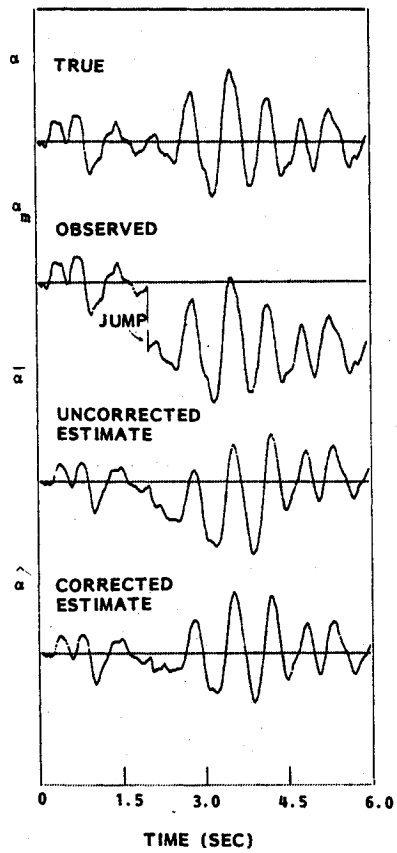

Fig. 3 Angle-of-attack and estimates.

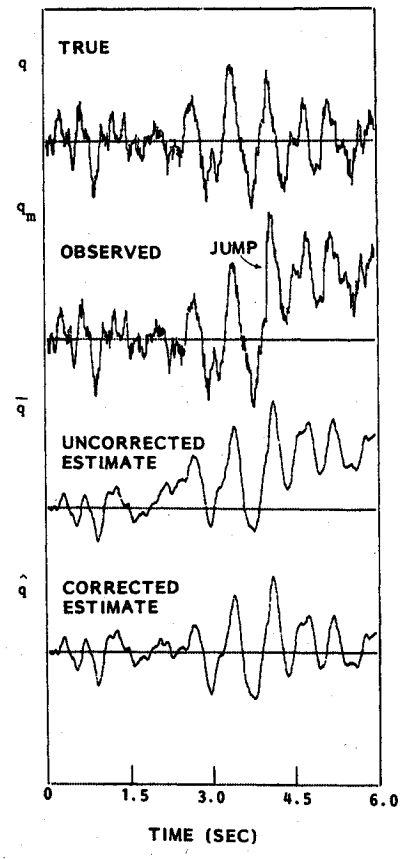

Fig. 4 Pitch rate and estimates.

rate estimate $\hat{q}$, which tends to follow the observed pitch rate q_m. The transfer function from the observed to the estimated angle of attack is high pass, so the transition in the accelerometer bias has no steady state effect.

The last trace on Figs. 3 and 4 show the corrected estimates, obtained as explained below.

Failure Detection and Correction

The residuals of the bias-free filter shown in Fig. 2 are the inputs to the failure detection algorithm shown in Fig. 5, the details of which are given in Ref. 7. This algorithm requires the equivalent observation matrix H and the observation noise matrix designated as R. The former is the matrix defined by Eq. (23), and has the numerical value given by Eq. (25). The latter is proportional to the spectral density matrix W, i.e.,

$$R = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \quad \text{with } r_1 = 66.67\, r_2$$

The algorithm also requires a set of thresholds for each of the triggers and a set of variances σ_{wi}^2 by which the diagonal elements of the bias estimation covariance matrix (P in Ref. 7) are increased as a result of a detected transition, i.e., when the trigger exceeds the corresponding threshold. If the thresholds are all set to zero, the algorithm will always take the right-hand (YES) branch and reverts to the linear (Kalman filtering) algorithm for the estimation of a random walk (b = white noise) observed through noise. The distinctive, nonlinear behavior of the bias estimator results when the thresholds are set at nonzero levels.

The benefits of the nonlinear (maximum-likelihood) algorithm of Fig. 5 are illustrated for this application in Figs. 6 and 7. The upper trace shows the performance of the nonlinear algorithm when the thresholds are set to realistic, nonzero values approximately half the value of the trigger generated by corresponding instrument failures. The following three sets of traces correspond and show the performance of linear filters (i.e., thresholds set to zero) for three levels of process, noise σ_{wi}. These are designated as "low gain," "middle gain," and "high gain." The high gain case corresponds to the same σ_{wi} used in the nonlinear filter, the middle gain case uses σ_{wi} set at 1/10 of the high gain values, and the low gain case uses $\sigma_{wi} = 1/100$ of the high gain values.

Figures 6 and 7 show that the steps are tracked to some extent by the linear filters with each set of gains, as well as by the nonlinear filter. The nonlinear filter, however, shows a significant advantage in that it combines a rapid transient response with an absence of noise in the steady state. Consider the estimates of the pitch rate bias shown in Fig. 7, for example. The transient response speed of the nonlinear filter is comparable to that of the linear filter with the "middle gain" values. The latter filter has a noisier steady-state ouput. A lower steady-state noise level can be achieved with a linear filter by reducing the gain, but this reduces the speed of response. A heuristic explanation of why the nonlinear estimator can exhibit performance unattainable in a linear filter is that the bandwidth of the latter is fixed. In order to produce a fast transient response, it must have a high bandwidth and hence must permit the passage of much of the white noise present in the residual. Reducing the bandwidth to reduce steady-state output noise necessarily compromises transient response. The bandwidth of the nonlinear filter, on the other hand, is not fixed. When the occurence of a transition is detected, the covariance matrix P is increased, (see

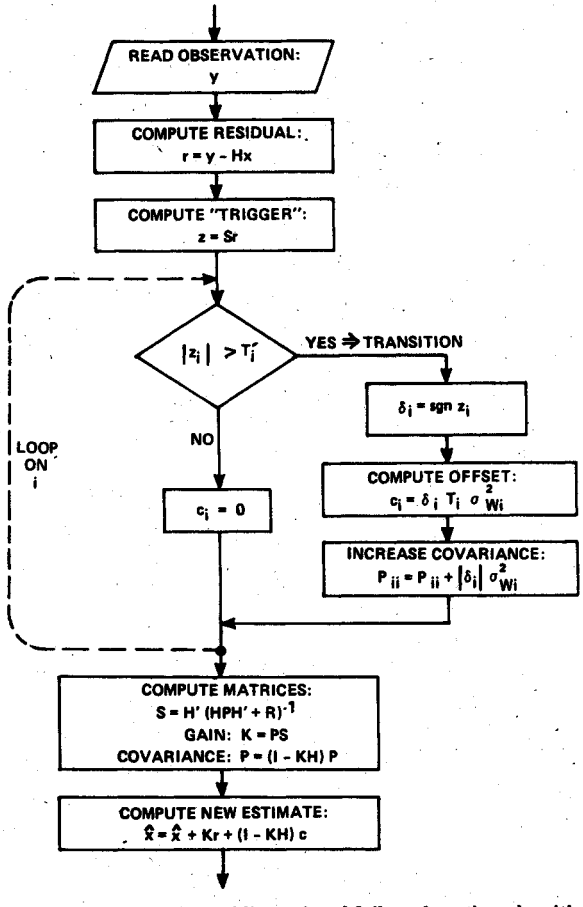

Fig. 5 Flow chart of multidimensional failure detection algorithm.

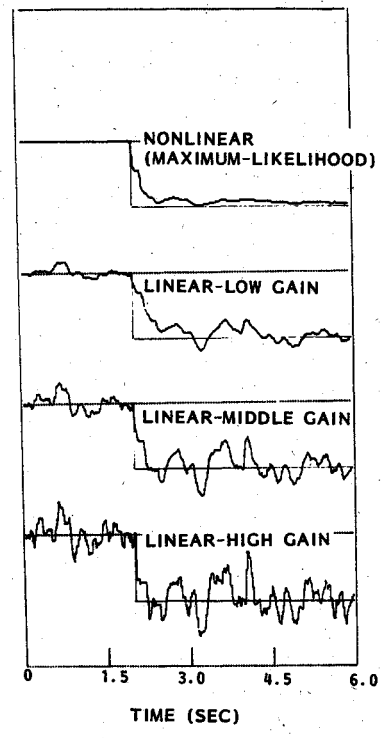

Fig. 6 Accelerometer bias estimates.

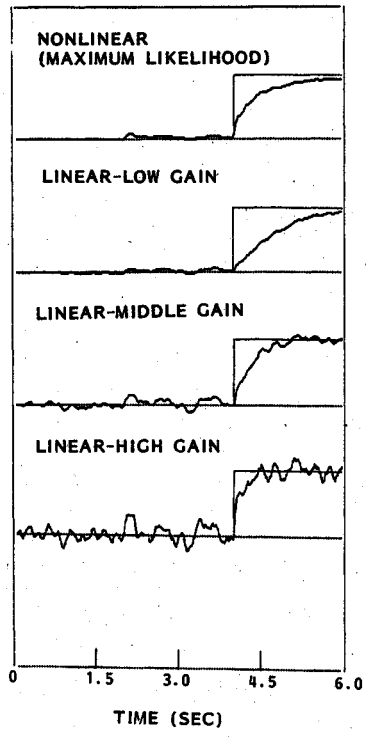

Fig. 7 Pitch rate bias estimates.

Fig. 5). This increases the estimator gain K and hence its bandwidth. In the time interval between transitions, however, the covariance matrix becomes smaller; this reduces the estimator gain and its bandwidth. It is the ability of the nonlinear estimator to change its covariance matrix in response to a detected transition that gives the nonlinear estimator its desirable properties.

The estimate of the bias can be used to correct the state estimate, in accordance with the relationship[10]

$$\hat{x}=\bar{x}+V\hat{b}$$

where V is the matrix given by Eq. (20). For this example,

$$V=\begin{bmatrix} -0.5 & 4.0 \\ -0.05 & -0.4 \end{bmatrix}$$

The results of using this correction with $\hat{b}$ from the nonlinear filter shown by the last traces in Figs. 3 and 4. It is observed that the offsets due to the biases in the uncorrected estimates $\bar{\alpha}$ and $\bar{q}$ are removed when the correction $V\hat{b}$ is applied.

Conclusion

This paper has presented a digital simulation of the effectiveness of a new failure detection and correction technique to detection of failures of sensors in an aircraft flight control system.

The aircraft model was intentionally simplified to permit the reader to follow and verify the calculations. The design calculations for a more complete and realistic aircraft model follow the same pattern as those illustrated in this paper, but would require the aid of a computer to perform. There is no indication that the results for a more realistic model would be any different from those presented here.

References

[1] Willsky, A.S., "A Survey of Design Methods for Failure Detection in Dynamic Systems," *Automatica,* Vol. 12, No. 6, Nov. 1976, pp. 601-611.

[2] Deckert, J.C., Desai, M.N., Deyst, J.J., and Willsky, A.S., "F-8 DFBW Sensor Failure Identification Using Analytic Redundancy," *IEEE Transactions on Automatic Control,* Vol. AC-22, No. 5, Oct. 1977, pp. 795-803.

[3] Chien, T. and Adams, M.B., "A Sequential Failure Detection Technique and Its Application," *IEEE Transactions on Automatic Control,* Vol. AC-21, No. 5, Oct. 1976, pp. 750-757.

[4] Caglayan, A.K., "Simultaneous Failure Detection and Estimation in Linear Systems," *Proceedings of the 19th Conference on Decision and Control,* Vol. 2, Dec. 1980, pp. 1038-1041.

[5] Wald, A., *Sequential Analysis,* Wiley, New York, 1947.

[6] Friedland, B., "Maximum-Likelihood Estimation of a Process with Random Transitions (Failures)," *IEEE Transactions of Automatic Control,* Vol. AC-24, No. 6, Dec. 1979, pp. 932-937.

[7] Friedland, B., "Multidimensional Maximum-Likelihood Failure Detection and Estimation," *IEEE Transactions on Automatic Control,* Vol. AC-26, No. 2, April 1981, pp. 567-570.

[8] Friedland, B. and Grabousky, S., "Estimating Sudden Changes of Biases in Linear Dynamic Systems," *IEEE Transactions on Automatic Control,* Vol. AC-27, No. 1, Feb. 1982, pp. 237-240.

[9] Bryson, A.E. and Ho, Y.-C., *Applied Optimal Control,* Blaisdell Publishing Co., Waltham, Mass., 1969, p. 171.

[10] Beard, R.V., "Failure Accommodation in Linear Systems Through Self-Reorganization," Ph.D., Thesis, MTV-71-1, Massachusetss Institute of Technology, Feb. 1971.

[11] Friedland, B., "Notes on Separate-Bias Estimation," *IEEE Transactions on Automatic Control,* Vol. AC-23, No. 4, Aug. 1978, pp. 735-738.