8-2004

Kalman Filtering with Uncertain Noise Covariances

Srikiran Kosanam
Cleveland State University, s.kosanam@csuohio.edu

Daniel J. Simon
Cleveland State University, d.j.simon@csuohio.edu

Original Citation

S. Kosanam and D. Simon. (2004). Kalman Filtering with Uncertain Noise Covariances, Intelligent Systems and Control, Honolulu, HI, 375-379.

Kalman Filtering with Uncertain Noise Covariances

Srikiran Kosanam
Department of Electrical Engineering
Cleveland State University
1960 East 24th Street, Cleveland, OH-44115
s.kosanam@csuohio.edu

Dan Simon
Department of Electrical Engineering
Cleveland State University
1960 East 24th Street, Cleveland, OH-44115
d.j.simon@csuohio.edu

***Abstract* - In this paper the robustness of Kalman filtering against uncertainties in process and measurement noise covariances is discussed. It is shown that a standard Kalman filter may not be robust enough if the process and measurement noise covariances are changed. A new filter is proposed which addresses the uncertainties in process and measurement noise covariances and gives better results than the standard Kalman filter. This new filter is used in simulation to estimate the health parameters of an aircraft gas turbine engine.**

Keywords: Kalman filtering, robust filtering, parameter estimation and Ricati equation

I. INTRODUCTION

In a standard Kalman filter, all the system characteristics (i.e., the system model, initial conditions, and noise characteristics) have to be specified a priori. However, if there is uncertainty in any of these characteristics, the filter may not be robust enough. In this paper, an alternate filter is proposed which performs better than the standard Kalman filter for uncertainties in both process and measurement noise covariances.

Most of the recent research in the robust filtering field has dealt with bounded parameter uncertainty or Kalman filtering with an H-infinity norm constraint. Ian R. Petersen [5] is about designing robust state feedback controllers and steady state robust state estimators for uncertain linear systems with norm bounded uncertainties. In this method a guaranteed cost quadratic controller is proposed and a quadratic guaranteed estimator is developed based on the duality. The uncertainties in this work had known upper bounds. Lihua Xie [6] proposes a state estimator which guarantees a bound on estimation error covariance for all admissible uncertainties in the state and output model. Wassim M. Haddad [9] considers parametric uncertainties in plant model. An estimation error bound suggested by multiplicative white noise modeling is utilized for guaranteeing robust estimation over a specified range of parameter uncertainties. In Dah-Jing Jwo [10] the measurement noise covariance matrix and estimation error covariance matrix is identified with a fuzzy method combined with neural networks. Mehra [11] classified the estimation of covariance matrices into four categories: Bayesian, maximum likelihood, correlation and covariance matching. Zhiqian Weng [12] proposes an evolutionary programming technique for uncertain systems with unknown but bounded uncertain parameters by interval systems. This paper deals with minimizing the average estimation error in the presence of uncertainties in process noise and measurement noise covariances. This paper does not assume any bounds on the uncertainties in covariance matrices but is based on the knowledge of the statistics of the uncertainties.

The remaining sections of the paper proceed as follows. Section 2 is concerned with the general state estimation problem. Section 3 introduces uncertainties into process and measurement noise covariances and deals with robust estimation analysis. Some preliminary results are shown in Section 4 for aircraft gas turbine engine. Section 5 presents conclusions and future research issues.

II. THE STATE ESTIMATION PROBLEM

This analysis is based on [1], which applies to continuous time systems, and is extended in this paper to discrete time systems and applied to aircraft gas turbine engine health estimation.

Consider a linear stochastic system represented by

$$\begin{aligned} x_{k+1} &= Ax_k + B_u u_k + B_w w_k \\ y_k &= Cx_k + v_k \end{aligned} \tag{1}$$

Here x is the system state vector, y is the measurement vector, u is the input vector, w is the process noise vector and v is the measurement noise vector. A, B_u, B_w and C are matrices of appropriate dimensions. w and v in this case are assumed to be mutually independent and zero mean white noise. The covariances of w and v are given as

$$\begin{aligned} E[w_k w_k^T] &= Q \\ E[v_k v_k^T] &= R \end{aligned} \tag{2}$$

The state estimate equations before and after the measurements are processed are given as [2]

$$\hat{x}_{k+1}^{-} = A\hat{x}_k^{+} + B_u u_k$$
$$\hat{x}_{k+1}^{+} = \hat{x}_{k+1}^{-} + K_k (y_{k+1} - C\hat{x}_{k+1}^{-}) \tag{3}$$

Where K_k is the Kalman filter gain.

The estimation error is defined as follows:

$$e_{k+1} \equiv x_{k+1} - \hat{x}^{-}{}_{k+1}$$

From equations (1) and (3) the estimation error satisfies the equation

$$e_{k+1} = (A - AK_k C)e_k + B_w w_k - AK_k v_k \tag{4}$$

Using the noise characteristics in equation (2) the steady state error covariance P becomes solution to the following equation [4]:

$$P = (A - AKC)P(A - AKC)^T + B_w Q B_w{}^T + (AK)R(AK)^T \tag{5}$$

Where P is defined as

$$P = E[ee^T] \tag{6}$$

When R = 0 (no measurement noise), equation (5) becomes

$$X_1 = (A - AKC)X_1(A - AKC)^T + B_w Q_k B_w{}^T \tag{7}$$

Where X_1 is the estimation error covariance due to process noise only.

When Q = 0 (no process noise), equation (5) becomes

$$X_2 = (A - AKC)X_2(A - AKC)^T + (AK)R(AK)^T \tag{8}$$

Where X_2 is the estimation error covariance due to observation noise only.

Adding equations (7) and (8) gives the following:

$$(X_1 + X_2) = (A - AKC)(X_1 + X_2)(A - AKC)^T + B_w Q B_w{}^T + (AK)R(AK)^T \tag{9}$$

This shows that when Q, R are not zero at the same time, the solution P of equation (5) becomes:

$$P = X_1 + X_2 \tag{10}$$

This is the estimation error covariance in the presence of both the process and measurement noise. Thus, it is shown to a linear combination of the estimation error covariance when only one of the noises is present.

This linear combination helps in realizing the performance index of the Kalman Filter, which would be a linear combination of functions of X_1 and X_2.

Therefore, in the standard Kalman filter, the filter gain K minimizes the following performance index [3]:

$$J = tr[E(e_k e_k^T)] = tr(P) = tr(X_1) + tr(X_2) \tag{11}$$

where tr( ) denotes the trace of a matrix. If there are no uncertainties in the process and measurement noise covariances the performance index J attains a global minimum using the standard Kalman filter. But if there were uncertainties in Q and R, J would not attain a minimum. Let us now consider the case where there are uncertainties in Q and R.

III ROBUSTNESS ANALYSIS OF THE KALMAN FILTER

This section considers variations in the process and measurement noise covariances. The covariance matrices of the process noise and measurement noise are assumed to change from nominal covariances Q, R to $\widetilde{Q}, \widetilde{R}$ using scalars α, β as follows:

$$\widetilde{Q} = (1 + \alpha)Q \tag{12}$$
$$\widetilde{R} = (1 + \beta)R \tag{13}$$

where α, β are random variables. α, β are assumed to be zero mean and uncorrelated. The estimation error P changes to $\widetilde{P}$ when the noise covariances change from Q to $\widetilde{Q}$ and R to $\widetilde{R}$. If $\widetilde{P} = P + \Delta P$ is substituted in equation (5),

$$(P + \Delta P) = (A - AKC)(P + \Delta P)(A - AKC)^T + (1 + \alpha)B_w Q_k B_w{}^T + (1 + \beta)(AK)R_k(AK)^T \tag{14}$$

$\widetilde{P}$ is the sum of the solution of the following two equations corresponding to R = 0 and Q = 0 respectively.

$$\widetilde{X}_1 = (A - AKC)\widetilde{X}_1(A - AKC)^T + \alpha B_w Q_k B_w{}^T \tag{15}$$

$$\widetilde{X}_2 = (A - AKC)\widetilde{X}_2(A - AKC)^T + \beta(AK)R(AK)^T \tag{16}$$

Comparing with (7) and (8) we can see that.

$$\widetilde{X}_1 = \alpha X_1, \widetilde{X}_2 = \beta X_2 \tag{17}$$

where X_1 and X_2 are the solutions of equations (7) and (8). Then the change in the estimation error covariance ΔP is described using X_1 and X_2 as follows:

$$\Delta P = \alpha X_1 + \beta X_2 \tag{18}$$

Then the variation of the performance index is as follows:

$$\Delta tr(P) = tr(\Delta P) = \alpha\, tr(X_1) + \beta\, tr(X_2)$$

(19)

Let us now consider the sensitivity of the performance index to α, β. Here α, β are random variables expressing uncertainties of noise covariances as follows:

$$E\{\alpha\} = E\{\beta\} = 0, E\{\alpha\beta\} = 0 \quad (20)$$

$$E\{\alpha^2\} = \sigma_1^2, E\{\beta^2\} = \sigma_2^2 \quad (21)$$

From the above characteristics, the mean of the change of the performance index is given as

$$E\{\Delta tr(P)\} = E\{\alpha\} tr\{X_1\} + E\{\beta\} tr(X_2) = 0 \quad (22)$$

So the mean of the variation of the performance index is zero. The variance of the change of the performance index becomes

$$Var\{\Delta tr(P)\} = E\{(\alpha\, tr(X_1) + \beta\, tr(X_2))^2\}$$
$$= \sigma_1^2 (tr(X_1))^2 + \sigma_2^2 (tr(X_2))^2 \quad (23)$$

Considering (22) and (23), if we minimize the variance of the change of the performance index, we make the filter robust to changes in Q and R. Ideally we would like to have the performance index for the filter with uncertain Q and R to be (23). But we want to balance the performance of the estimator under nominal conditions (nominal Q and R) with the performance of the estimator under off-nominal conditions (off-nominal Q and R). In order to achieve this balance, we want to have the performance index be a combination of the standard cost function and the variance of the change in the standard cost function.

From the the performance index for the robust Kalman filter can be given as:

$$J = \rho_1 \{ tr(X_1) + tr(X_2) \}$$
$$+ \rho_2 \{ \sigma_1^2 (tr(X_1))^2 + \sigma_2^2 (tr(X_2))^2 \} \quad (24)$$

where ρ_1 and ρ_2 provide relative weighting to nominal performance and robustness. This results in a new Kalman gain to minimize the new performance index. The robust Kalman filter is developed with the steady state gain of the standard Kalman filter and using the hybrid gradient descent algorithm a new Kalman gain is realized which minimizes the new performance index. Using this new Kalman gain the estimation error will be found to be less than the standard Kalman filter when there is an uncertainty in the noise covariances.

The hybrid gradient descent algorithm to realize the gain can be summarized as follows

To find the minimum of new J , $\frac{\partial J}{\partial K}$ has to be found. But $\frac{\partial J}{\partial K}$ cannot be found analytically as J is not an explicit function of K . J is an analytical function of X_1 and X_2 i.e. J= f(X_1, X2) and X_1 and X_2 are numerical functions of K.

Therefore $$\frac{\partial J}{\partial K} = \frac{\partial J}{\partial X_1}\frac{\partial X_1}{\partial K} + \frac{\partial J}{\partial X_2}\frac{\partial X_2}{\partial K} \quad (25)$$

$\frac{\partial J}{\partial X_1}$ and $\frac{\partial J}{\partial X_2}$ given analytically which are given as

$$\frac{\partial J}{\partial X_1} = \rho_1 I + 2\rho_2 \sigma_1 tr(X_1) I, \quad (26)$$
$$\frac{\partial J}{\partial X_2} = \rho_1 I + 2\rho_2 \sigma_1 tr(X_2) I$$

where I stands for the identity matrix of appropriate dimension. $\frac{\partial X_1}{\partial K}, \frac{\partial X_2}{\partial K}$ are computed numerically. The calculation of these partial derivatives is complex as the numerator and denominator are both matrices. In order to compute these partials each element of K is perturbed from its nominal value and then the new X_1 and X_2 are calculated.

This calculation of partial of X_1 and X_2 with respect to K is not straightforward as both X_1, X2 and K are all matrices. In order to compute these partials each element of K is perturbed and the corresponding change in X_1 and X_2 is calculated numerically. So every time X_1 and X_2 are calculated a discrete time Ricatti equation has to be solved, which is computationally very expensive. The calculation of partial of X_1 with K is shown here for limiting the space. Similar results apply to calculate the partial of X_2.

$$\frac{\partial X_1(i,j)}{\partial K(i,j)} = \frac{X_1 - \Delta X_1}{\varepsilon K(i,j)} \quad (27)$$

where ΔX_1 is the change in X_1 caused by perturbation of ith row jth column element of K. This perturbation is carried out for all the elements of K. Then, these partials are so multiplied that the change in each element of X1 for change in all the elements of K is achieved. After this partial is evaluated the gradient descent steps are given as follows.

Step1. Start with nominal value of K as standard Kalman gain

Step2. Compute $\frac{\partial J}{\partial K} \quad at \quad K = K_i$

Step3. If $\left\| \frac{\partial J}{\partial K} \right\| < Tolerance, Quit$

Step4. $K_{i+1} = K_i - \varepsilon \frac{\partial J}{\partial K}$

Step5. Go to step 2

Special care has to be taken to come up with the gradient descent step size and the perturbation size to find the partial derivative.

Computationally this method is time consuming but this is a straightforward method of realizing a new Kalman gain. Computationally it may be better to use efficient search algorithms than gradient descent.

IV SIMULATION RESULTS

The performance of an aircraft gas turbine engine deteriorates over time. This deterioration reduces the fuel economy of the engine[13]. To determine the health of the engine, data is periodically collected and is used to decide maintenance schedules. The data is then used to come up with the linearized model of the engine using the DIGTEM software, a public domain turbofan software simulation developed by the NASA Glenn Research Center [7],[14]. Three seconds of data are collected at 10 Hz every flight. In order to check the application of this paper, 50 flights are simulated.. As the system is highly complex evaluate, the results presented here are not optimal, as the algorithm had not converged in a given time frame (see Figure 1). It took 96 hours of computation on a Pentium-IV, 1.8 GHz, 256 MB RAM system for the system to complete the iterations shown in Figure 1.This is because for every iteration in gradient descent algorithm there are 1152(2x2x12x24) Ricati equations to be solved in MATLAB© (the Kalman gain matrix is 12 x 24, there are 2 Riccati equations to solve for [X_1 and X_2], and 2 evaluations of each Riccati equation is required to approximate the partial derivative). Each Ricati takes about 5.5 seconds to be solved. Figure 3 illustrates the nonlinearity of performance index with respect to two different elements of K for a turbofan health parameter estimation problem. Although the gradient descent algorithm did not converge, the suboptimal results verify that the robust Kalman filter may provide an attractive filtering option when there are uncertainties in the noise covariances. The results here are shown only for the presence of measurement noise uncertainty but no process noise uncertainty as the system model obtained by DIGTEM is assumed to be accurate. In other words, we used $\sigma_1^2 = 0$ and $\sigma_2^2 = 1$ in equation (21). The nonlinearity of the performance index with two different elements of K is given in Figure3.

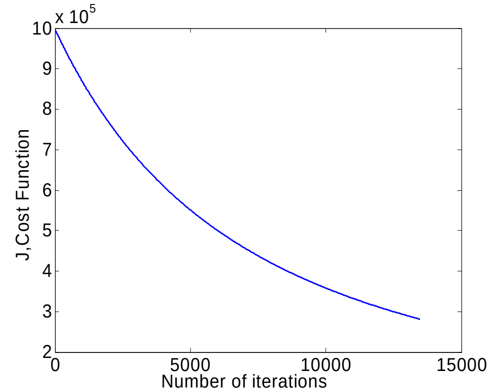

Figure 1 Cost function vs. number of iterations. This represents four days of CPU time

The new Kalman gain realized after these iterations resulted in a filter that was unstable. So we had to realize a new gain which is a linear combination of the standard Kalman gain (K_s)and the robust Kalman gain(K_r). So the new gain is given as follows

$$K_{new} = \eta K_s + (1-\eta) K_r \tag{28}$$

Where η is determined heuristically to give a stable but robust filter. η which is used for the simulation result shown in Figure 2 is 0.7. The health parameters that are to be estimated are as follows.

- Fan airflow
- Fan efficiency
- Compressor airflow
- Compressor efficiency
- High pressure turbine airflow
- High pressure turbine enthalpy change
- Low pressure turbine airflow
- Low pressure turbine enthalpy change

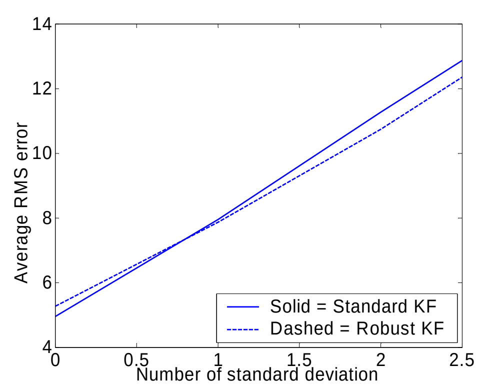

Figure 2 Performance of filters for various measurement noise perturbations

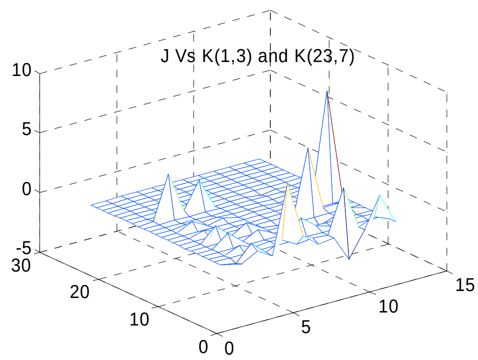

Figure 3 Performance index as a function of two elements of K

In the condition that there are no uncertainties in the noise covariances, the standard Kalman filter is expected to perform better than the robust Kalman filter. Simulation results are in accordance with this theory. These results are shown in Table 2. Although the robust filter does not perform as well as the standard filter, the drop off in performance is slight, and it may be worth the extra robustness that is obtained as seen in Table 1.

Figure 2 shows the average RMS estimation error for the two filters for the various changes in the measurement noise covariance. As expected, when the covariance change is zero, the standard Kalman filter outperforms the robust Kalman filter. However, as the covariance changes more and more, the robust filter gains more and more performance relative to the standard filter.

Health Parameter #	Standard KF RMS error	Robust KF RMS error
1	14.0	12.8
2	16.6	17.3
3	8.15	7.3
4	12.5	12.3
5	6.7	6.7
6	12.1	11.0
7	9.3	8.5
8	10.5	9.7
Average	11.3	10.7

Table 1 – Health parameter estimation errors (percent) when the variation in the measurement noise covariance is two standard deviations; η=0.7

Health Parameter #	Standard KF RMS error	Robust KF RMS error
1	5.5	4.6
2	6.7	7.1
3	2.9	3.0
4	5.6	5.7
5	2.6	5.5
6	4.7	5.0
7	4.4	4.1
8	6.8	6.7
Average	4.9	5.2

Table 2 – Health parameter estimation errors (percent) when there is no change in the measurement noise covariance; η=0.7

V CONCLUSION

For the turbofan problem the health parameter estimates with the robust gain are better than the estimates with the standard Kalman gain if the measurement noise covariance increases by more than one standard deviation. The performance of the robust filter may improve if the solution to the gradient descent algorithm converges. Also, better performance may be obtained for other variations of η_1 and η_2.

Other issues to be pursued in this area are the use of genetic algorithms instead of gradient descent for better convergence, and the feasibility of the application of the current algorithm for a time varying filter.

The next immediate step is to look at the possibility of weighting data coming from different sensors depending on the confidence levels on the sensors. Right now β in (14) is same for all the elements in R. This would not be the case when we have different confidence levels on different sensors. This issue is currently being pursued. Another step is to extend this work to other turbofan simulation software [8].

VI REFERENCES

[1] Shuichi Sasa, Robustness of a Kalman filter against uncertainties of Noise Covariances, Proceedings of the American Control Conference, pp.2344-2348, Philadelphia, Pennsylvania, June 1998

[2] B. Anderson and J. Moore, Optimal Filtering (Prentice Hall, Englewood Cliffs, New Jersey, 1979)

[3] R. Stengel, Optimal Control and Estimation (Dover Publications, New York, 1994)

[4] A. Gelb, Applied Optimal Estimation (MIT Press, 1974)

[5] Ian R. Petersen and Duncan C.McFarlane, Optimal Guaranteed Cost Control and Filtering for Uncertain Linear Systems, IEEE Transactions on Automatic Control, pp. 1971-1977, September 1994

[6] Lihua Xie and Yeng Chai Soh, Robust Kalman filtering for uncertain systems, Systems and Control Letters 22 (1994), pp.123-129

[7] D. Simon and D.L. Simon, ASME Turbo Expo 2003, Aircraft Turbofan Engine Health Estimation Using Constrained Kalman Filtering, Atlanta, GA, paper GT2003-38584, June 2003

[8] Khary I. Parker and Ten-Heui Guo, Development of a Turbofan Engine Simulation in a Graphical Simulation Environment.

[9] Wassim M. Haddad and Dennis S. Bernstein, Robust, Reduced-Order, Nonstrictly Proper State Estimation Via the Optimal Projection Equations with guaranteed Cost Bounds, IEEE Transactions on Automatic Control, Vol. 33,No. 6, June 1998

[10] Dah-Jing Jwo and Hung-Chin Huang, Fuzzy neural network aided adaptive extended Kalman filtering for GPS navigation, Aerospace and Science Technology, submitted for publication.

[11] R.K. Mehra Approaches to adaptive filtering, IEEE transactions on Automatic Control. AC-17 (1972), pp 693-698

[12] Zhiqian Weng, Guanrong Chen, Leang S.Shieh, Johan Larsson, Evolutionary Programming Kalman filter, Information Sciences 129 (2000), pp 197-210

[13] Allan J. Volponi, Hans DePold, Ranjan Ganguli, Chen Daguang, The use of Kalman Filter and Neural Network Methodologies in Gas Turbine Performance Diagnostics : A Comparative Study, proceedings of ASME TURBOEXPO 2000, 2000 Munich Germany.

[14] D. Simon and D.L. Simon, Aircraft Turbofan Engine Health Estimation Using Constrained Kalman Filtering, ASME Journal of Engineering for Gas Turbines and Power, in print.