

STOCHASTIC OBSERVABILITY AND FAULT DIAGNOSIS OF ADDITIVE CHANGES IN STATE SPACE MODELS

Fredrik Gustafsson

Department of Electrical Engineering
Linköping University, SE-581 83 Linköping, Sweden
Email: fredrik@isy.liu.se

ABSTRACT

We derive a Kalman filter based on data from a sliding window. This is used for a new approach to fault detection and diagnosis, where the state estimate from past data is compared to the state estimate of some of the future data. We suggest a method to judge the quality of diagnosis in a simple way. For fault estimation in the diagnosis, the general concept of *stochastic observability* in linear systems is introduced. Its role on the design step is illustrated on a problem of estimating the true velocity of a car.

1. INTRODUCTION

A standard approach to change detection in parametric models is based on model validation. Assume the measurement is produced by a linear system $y_t = G(q, \theta)u_t + H(q, \theta)e_t$, where u_t is a measured input, e_t is noise, q is the shift operator and θ a parameter vector. Many algorithms are based on the assumption that we have estimated nominal model parameters θ_0 based on a large amount of measurements, and would like to know if a current estimated value $\hat{\theta}$ based on data for a sliding window is consistent with the nominal model. See [1] or [2] for details.

However, many signal processing problems, in particular in navigation related areas, are expressed in state space models. Typical faults include sensor and actuator offsets and drifts, which may all be modeled as additive changes. Change detection, or fault detection as we call it from now on, on state space models are either based on an innovation whiteness test or filter banks, see [2] (Chapters 8,9). The former approach is numerically simple, but not very powerful and not useful for diagnosis. The latter approach is numerically complex.

Here we derive the Kalman filter over a sliding window (Section 3) and suggest a detection and diagnosis algorithm (Section 4) based on this. Section 5 describes a framework for analysing fault diagnosis problems, whose practical implementation is illustrated in Section 6.

2. MODELS AND PROBLEM FORMULATION

The state space model under consideration is

$$x_{t+1} = A_t x_t + v_t, \quad Q_t = \text{Cov}(v_t) \quad (1a)$$

$$y_t = C_t x_t + e_t, \quad R_t = \text{Cov}(e_t). \quad (1b)$$

The covariance matrix R_t is positive definite and Q_t positive semi-definite. No known input is included in the model, since it would not affect the results or complicate the derivations.

Additive state changes can be used for modelling actuator faults and sensor faults, as offsets and drifts, as well as disturbances. The first fault model is

$$x_{t+1} = A_t x_t + v_t + \sigma_{t-k} B_t f, \quad Q_t = \text{Cov}(v_t) \quad (2a)$$

$$y_t = C_t x_t + e_t + \sigma_{t-k} D_t f, \quad R_t = \text{Cov}(e_t). \quad (2b)$$

Here f denotes the fault magnitude and σ_t is the step function, which is one after time k and zero before. An alternative state space model for fault estimation is based on the augmented state vector $\bar{x}_t = (x_t, f)^T$ which becomes

$$x_{t+1} = \begin{pmatrix} A_t & B_t \\ 0 & I \end{pmatrix} x_t + \begin{pmatrix} v_t \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \delta_{t-k-1} f \end{pmatrix} \quad (3a)$$

$$y_t = (C_t \quad D_t) x_t + e_t \quad (3b)$$

Both (2) and (3) are identical, but the latter will be analysed with respect to observability (*is the fault observable from the measurements?*) and the former is used in the proposed algorithm. For a discussion and more information about the models (2) and (3), see [2] (Appendix A and Chapters 8,9).

3. KALMAN FILTER OVER SLIDING WINDOW

Standard references in the field describe how the Kalman filter estimates x_t based on past measurements y_t, y_{t-1}, \dots , see [3, 4, 2]. Assuming Gaussian noise distributions, we get the conditional distribution

$$\hat{x}_{t|t-1} = N(x_t, P_{t|t-1}), \quad (4)$$

from the Kalman filter as one-step predictor.

Now we derive an algebraic expression for the state estimate over a sliding window. That is, estimate x_t from $y_t, y_{t+1}, \dots, y_{t+L}$ assuming that the window looks forward in time. The standard model (1) gives

$$\underbrace{\begin{pmatrix} y_t \\ y_{t+1} \\ y_{t+2} \\ \vdots \\ y_{t+L} \end{pmatrix}}_{Y_t^L} = \underbrace{\begin{pmatrix} C_t \\ C_{t+1}A_t \\ C_{t+2}A_{t+1}A_t \\ \vdots \\ C_{t+2}A_{t+L-1} \dots A_t \end{pmatrix}}_{\mathcal{O}_t^L} x_t + \underbrace{\begin{pmatrix} e_t \\ e_{t+1} \\ e_{t+2} \\ \vdots \\ e_{t+L} \end{pmatrix}}_{E_t^L} + \underbrace{\begin{pmatrix} 0 \\ C_{t+1}v_t \\ C_{t+2}(v_{t+1} + A_{t+1}v_t) \\ \vdots \\ C_{t+L}(v_{t+L-1} + A_{t+L-1}v_{t-L-2} + \dots) \end{pmatrix}}_{V_t^L}, \quad (5)$$

or shorter

$$Y_t^L = \mathcal{O}_t^L x_t + E_t^L + V_t^L. \quad (6)$$

The noise terms have covariance matrices (for $L = 2$ here)

$$\text{Cov } E_t \triangleq \bar{R}_t = \begin{pmatrix} R_t & 0 & 0 \\ 0 & R_{t+1} & 0 \\ 0 & 0 & R_{t+2} \end{pmatrix}$$

$$\text{Cov } V_t \triangleq \bar{Q}_t = \begin{pmatrix} 0 & 0 & 0 \\ 0 & C_{t+1}Q_t C_{t+1}^T & C_{t+1}Q_t A_{t+1}^T C_{t+2} \\ 0 & C_{t+2}A_{t+1}Q_t C_{t+1} & \bar{Q}_t^{33} \end{pmatrix},$$

$$\bar{Q}_t^{33} = C_{t+2}Q_{t+1}C_{t+2}^T + C_{t+2}A_{t+1}Q_t A_{t+1}^T C_{t+2}.$$

It follows immediately from the property of Gaussian vectors and least squares theory that the minimum variance state estimate (or Kalman filter over sliding window) is

$$Y_t \in \mathcal{N}(\mathcal{O}_t^L x_t, \bar{R}_t^L + \bar{Q}_t^L) \Rightarrow \quad (7a)$$

$$\hat{x}_{t|t+L} = ((\mathcal{O}_t^L)^T (\bar{R}_t^L + \bar{Q}_t^L)^{-1} \cdot \mathcal{O}_t^L)^{-1} \quad (7b)$$

$$\cdot (\mathcal{O}_t^L)^T (\bar{R}_t^L + \bar{Q}_t^L)^{-1} Y_t^L \quad (7c)$$

$$\in \mathcal{N} \left(x_t, \underbrace{((\mathcal{O}_t^L)^T (\bar{R}_t^L + \bar{Q}_t^L)^{-1} \mathcal{O}_t^L)^{-1}}_{P_{t|t+L}} \right) \quad (7d)$$

$$\hat{x}_{t|t+L}^{obs} = (\mathcal{O}_t^L)^\dagger Y_t^L \quad (7e)$$

$$\in \mathcal{N}(x_t, (\mathcal{O}_t^L)^\dagger (\bar{R}_t^L + \bar{Q}_t^L) ((\mathcal{O}_t^L)^\dagger)^T). \quad (7f)$$

Here we have for comparison also included the observer estimate (not involving any stochastic considerations) $\hat{x}_{t|t+L}^{obs}$. However, it is $\hat{x}_{t|t+L}$ which is the minimum variance estimate and thus the Kalman filter estimate over the sliding window.

4. A TWO-FILTER APPROACH TO FAULT DETECTION

The following algorithm is suggested to estimate the fault f in (1). In the sequel, we make the simplifying assumption that $B_t = I$ and $D_t = 0$ so the fault affects all of the state vector but not the sensors signals. The generalization includes more projections that might blur the picture.

4.1. Filter computations

Compute the following estimates:

1. Run a standard Kalman filter for (1) up to time t , not including the measurement y_t . It provides

$$\hat{x}_{t|t-1} \in \mathcal{N}(x_t, P_{t|t-1}).$$

2. Run the sliding window Kalman filter on the measurements from time t to time $t+L$. This provides, according to (7),

$$\hat{x}_{t|t+L} \in \mathcal{N}(x_t + f, P_{t|t+L}).$$

If a fault occurred at time t , then $f \neq 0$. That is, this is a generalization of the well-known approach of two parallel filters, used for parametric models (see Chapter 6 in [2]), to state space models.

3. Compute an estimate of the fault as

$$\hat{f}_t \triangleq \hat{x}_{t|t+L} - \hat{x}_{t|t-1} \in \mathcal{N}(f, P_{t|t-1} + P_{t|t+L}). \quad (8)$$

Here we have used the fact that the two state estimates are independent, since they are based on independent data sets.

4.2. Fault detection

Given (8), the obvious test is to monitor

$$g_t \triangleq \hat{f}_t^T (P_{t|t-1} + P_{t|t+L})^{-1} \hat{f}_t \in \chi^2(\dim(f)) \quad (9)$$

which is χ^2 distributed under the no-fault hypothesis. If g_t exceeds a threshold $g_t > h$, then the next diagnosis step is applied. The expected value of the test statistic after a change is $E(g_t) = f^T (P_{t|t-1} + P_{t|t+L})^{-1} f + \dim(f)$. We can define the *fault to noise ratio* (FNR) as

$$\text{FNR} = \frac{E(g_t)^2}{\text{Var}(g_t)},$$

which yields here

$$\text{FNR} = \frac{(f^T (P_{t|t-1} + P_{t|t+L})^{-1} f + \dim(f))^2}{2 \dim(f)}.$$

It should be clear that the weighting matrix in this expression contains much information about which changes are detectable.

In many cases, detection is an easier task than diagnosis, so it might be worthwhile to collect more data, by increasing L , before applying diagnosis.

4.3. Fault diagnosis

Suppose there are m different fault directions f_1, f_2, \dots, f_m . After detection, we want to isolate the fault, that is, decide which f_i occurred. Normalize the faults as

$$\bar{f}_i = (P_{t|t-1} + P_{t|t+L})^{-1/2} f_i. \quad (10)$$

The scaling of the nominal fault vectors f_i is immaterial, but preferably scaled appropriately. The size of the fault (the norms of the vectors) is unknown, so the actual fault is $f = \sigma f_i$ for some σ and i . Define the fault correlation matrix T with

$$T_t^{(i,j)} = \bar{f}_i^T \bar{f}_j. \quad (11)$$

The off-diagonal elements show the correlation between different faults after the spatial decorrelation (10). Since T_t changes in time, it might be a diagonal matrix at one time instant and a fault entering the system at that time is easy to diagnose, but in another situation this might be impossible.

For diagnosis, the following test statistics can be computed:

$$g_t^i \triangleq \bar{f}_i^T (P_{t|t-1} + P_{t|t+L})^{-1/2} \hat{f}_t \quad (12)$$

Now if the true fault is $f = \sigma f_j$, then we can express $\hat{f}_t = \sigma f_j + w$, where $w \in \mathcal{N}(0, P_{t|t-1} + P_{t|t+L})$ from (8). The distributions of the test statistics are

$$g_t^i \in \mathcal{N}(\bar{f}_i^T \sigma \bar{f}_j, \bar{f}_i^T \bar{f}_i) \quad (13)$$

$$\in \mathcal{N}(\sigma T_t^{(i,j)}, T_t^{(i,i)}) \quad (14)$$

under hypothesis H_j of a fault $f = \sigma f_j$. Here the role of the diagnosability matrix T_t is clear. Any large off-diagonal element can cause an incorrect diagnosis decision by the randomness of the estimated \hat{f}_t . More specifically, we can define the FNR as

$$\text{FNR} = \sigma^2 \frac{(T_t^{(i,j)})^2}{T_t^{(i,i)}}.$$

A good practical procedure would be as follows: After a fault is detected, compute the normalized fault directions \bar{f}_i , and the correlation matrix T_t . Then, compute the test statistics g_t^i and maximize for diagnosis. Check the risk of incorrect decision. This can be done since T_t contains all information about the risk of incorrect decisions. If the risk is too large, try to increase L .

5. THE STOCHASTIC OBSERVABILITY GRAMIAN

This section describes an alternative procedure for diagnosis based on estimation of the faults. The method is primarily aimed at analysis and design purposes.

Assume that some fault detection algorithm, for instance the one in Section 4.2, has shown that a fault has occurred. A Kalman filter is then applied on the augmented state space model (3). The question now is if the faults are observable. For notational simplicity, let A, B, C, Q, R denote the state space matrices of (3).

For time invariant systems, *observability* is equivalent to full rank of the observability matrix (assume $L \geq \dim(x) - 1$)

$$\mathcal{O} \triangleq \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^L \end{pmatrix}$$

The straightforward generalization of observability to time-varying systems follows from (5) as the condition of full rank of

$$\mathcal{O}_t \triangleq \begin{pmatrix} C_t \\ C_{t+1}A_t \\ C_{t+2}A_{t+1}A_t \\ \vdots \\ C_{t+2}A_{t+L-1} \dots A_t \end{pmatrix},$$

from which the square *observability Gramian* is defined in the control literature (see for example [5]):

$$G_t \triangleq (\mathcal{O}_t^L)^T \mathcal{O}_t^L. \quad (15)$$

For stochastic systems, the definition is

$$G_t \triangleq (\mathcal{O}_t^L)^T (\bar{R}_t^L)^{-1} \mathcal{O}_t^L. \quad (16)$$

However, in practice these measures are a bit dull, since they do not say anything about the FNR; they just answer yes or no.

To answer the question which (fault) states are observable and distinguishable from noise, the following *stochastic observability Gramian* is suggested:

$$\bar{G}_t \triangleq (\mathcal{O}_t^L)^T (\bar{R}_t^L + \bar{Q}_t^L)^{-1} \mathcal{O}_t^L = P_{t|t+L}^{-1}. \quad (17)$$

It follows from (7) that it coincides with the inverse covariance matrix of the state estimate from a sliding window which includes L measurements. The eigenvalues of \bar{G}_t gives important information about the FNR, and in particular which fault vectors f are not distinguishable from noise for a given excitation. In the Kalman filter design, we may turn off adaptation of the fault states at certain time intervals, if the condition number of \bar{G}_t is poor!

6. EXAMPLE: VELOCITY ESTIMATION

The following example illustrates a case where the Gramian (15) is well conditioned but the stochastic Gramian (17) is

ill-conditioned. As expected, the Kalman filter will not provide useful estimates unless some conditions on excitation are satisfied. The stochastic Gramian can thus be used as a design tool to examine which combinations of design parameters Q , R and excitation (here C_t) are permissible.

The following model describes a very simplified model for estimating absolute velocity in a car [6]. Two sensors are used, one that measures angular velocity $\omega_t = v_t/r_0$ of a wheel with nominal radius r_n (true value is $r_0 = r_n - \delta_r$) and one accelerometer that measures acceleration \dot{v}_t with an unknown offset δ_a . It is in this study assumed that the accelerometer is mounted on the wheel such that their velocity vectors are identical. That is, the fault state is $f = (\delta_a, \delta_r)^T$, and the state space model (3) becomes:

$$x_t = \begin{pmatrix} \dot{v}_t \\ v_t \\ \delta_a \\ \delta_r \end{pmatrix} \quad (18a)$$

$$x_{t+1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ T_s & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} x_t + \begin{pmatrix} T_s & 0 & 0 \\ T_s^2/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} v_t \quad (18b)$$

$$y_t = \begin{pmatrix} a_x \\ \omega_t r_n \end{pmatrix} = \begin{pmatrix} \dot{v}_t + \delta_a + e_t^1 \\ v_t + \omega \delta_r + e_t^2 \end{pmatrix} \quad (18c)$$

$$= \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \omega_t \end{pmatrix} x_t + e_t \quad (18d)$$

The observability matrix is

$$\mathcal{O}_t = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \omega_t \\ 1 & 0 & 1 & 0 \\ T_s & 1 & 0 & \omega_{t+1} \\ 1 & 0 & 1 & 0 \\ 2T_s & 1 & 0 & \omega_{t+2} \end{pmatrix}$$

It is easily seen that the observability matrix has full rank *if and only if* ω_t is not a linear function of time (including a constant). The optimistic conclusion would be that a Kalman filter successfully estimates the fault states δ_a and δ_r . However, care must be taken such that the process noise on the fault states does not integrate during periods of poor excitation. Figure 1 shows a simulated velocity profile and the condition number of the stochastic observability Gramian G_t^{10} and Kalman covariance $P_{t|t}$, respectively. As expected, the excitation is better the higher acceleration. This has been confirmed experimentally as well in a test car.

7. CONCLUSIONS

The contributions are:

- A method to compute the Kalman filter over a sliding window, here derived in an algebraic way without the need to specify initial conditions.

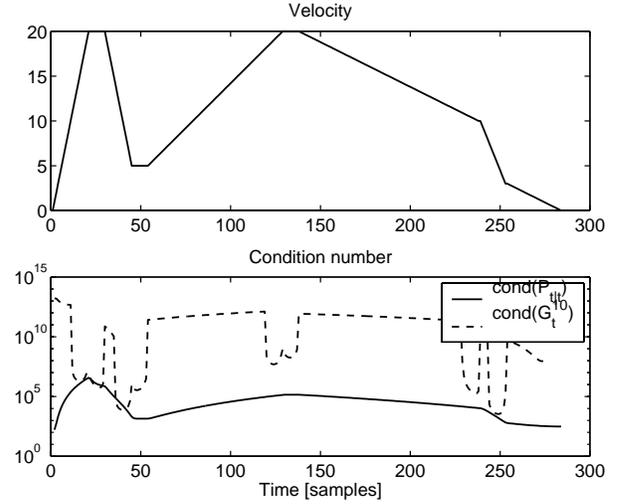


Fig. 1. Velocity profile and condition numbers.

- An algorithm based on a predictive Kalman filter and the sliding window Kalman filter. Additive faults are here detected and diagnosed more or less in a standard way. For diagnosis, we suggested to analyse a certain correlation matrix (T), which contains sufficient information for calculating the risk of false diagnosis. From T , the fault to noise ratio (FNR) can be computed.
- The generalization of observability using the sliding window Kalman filter. This approach also provides a measure of FNR, useful for filter design.
- An example highlighting the need for feedforward excitation measures of fault states to the Kalman gain.

8. REFERENCES

- [1] M. Basseville and I.V. Nikiforov, *Detection of abrupt changes: theory and application*, Information and system science series. Prentice Hall, Englewood Cliffs, NJ., 1993.
- [2] Fredrik Gustafsson, *Adaptive filtering and change detection*, John Wiley & Sons, Ltd, 2000.
- [3] B.D.O. Anderson and J.B. Moore, *Optimal filtering*, Prentice Hall, Englewood Cliffs, NJ., 1979.
- [4] T. Kailath, A.H. Sayed, and B. Hassibi, *Linear estimation*, Information and System Sciences. Prentice-Hall, Upper Saddle River, New Jersey, 2000.
- [5] T. Kailath, *Linear systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
- [6] F. Gustafsson and S. Ahlqvist, "Sensor fusion system," Swedish patent application nr 0002212-9, 2000.