FAULT DETECTION AND IDENTIFICATION USING PARAMETER ESTIMATION TECHNIQUES

Elena M. Cimpoeșu¹, Bogdan D. Ciubotaru², Dan Ștefanoiu³

This paper focuses on the use of parameter estimation techniques for the implementation of real-time Fault Detection and Diagnosis schemes. A detailed analysis of the nonrecursive and recursive Least Squares methods is given in the context of the system diagnosis problem, and a procedure for performing fault detection and identification for multivariable systems is proposed. An application example from the field of aircraft control is considered, which illustrates the suitability of the Recursive Least Squares method with exponential weighting and constant forgetting factor for Fault Detection and Fault Identification / Estimation within complex applications.

Keywords: Fault Detection, Fault Diagnosis, Parameter Estimation, Least Squares, Aircraft Application

1. Introduction

The increasing reliability and availability requirements for safety-critical technical applications have led to the development of a new type of control systems, which are able to ensure the closed-loop stability and the desired performances, both under nominal functioning conditions and in the presence of system faults. They are known as Fault Tolerant Control Systems (FTCS) and can be classified as either Passive (PFTCS) or Active (AFTCS). PFTCS make use of robust off-line controller design in order to ensure that the system is tolerant to certain predefined faults. AFTCS use on-line reconfiguration techniques to compensate for the influence of unpredicted faults; these approaches consist either of choosing a new control law from a precomputed class, or of redesigning the control law based on the impaired system model. Regardless of the reconfiguration strategy, AFTCS rely heavily on Fault Detection and Diagnosis (FDD) schemes to provide accurate, on-time information about the faults that occur within the system [1], [2], [3], [4].

¹PhD Student, Department of Automatic Control and Systems Engineering, University “Politehnica” of Bucharest, Romania
²Lecturer, Department of Automatic Control and Systems Engineering, University “Politehnica” of Bucharest, Romania, e-mail: bogdan.ciubotaru@acse.pub.ro
³Professor, Department of Automatic Control and Systems Engineering, University “Politehnica” of Bucharest, Romania
FDD is a relatively new research field in the context of automatic control engineering, and has known a rapid development during the last decades, mostly due to the increasing industrial safety requirements; this has led to the development of a large number of FDD techniques and associated algorithms. The current FDD approaches can be classified into two categories, namely model-based and data-based schemes.

Model-based FDD approaches use mathematical or schematic representations of the process behavior in order to fulfil their intended purpose. They can be further classified into qualitative and quantitative FDD schemes [2]; other classifications can be found in [5]. Quantitative model-based FDD use mathematical models in order to describe the dynamic behavior of the controlled processes and to identify any deviation from the nominal behavior as a result of a fault occurring within the system. The majority of the techniques that fall within this category are based on parameter estimation, parity equations, state estimation, or a combination of the three. A more detailed description can be found in [6]. This paper focuses solely on model-based FDD schemes that use parameter estimation techniques and continues the work presented in [7]. The motivation of this research consists of the fact that this category of system diagnosis techniques presents strong advantages from the viewpoint of practical implementation, and also with regard to their integration with system reconfiguration / restructuring mechanisms in the context of AFTCS.

The paper is organized as follows: Section 2 presents the existing faults classification criteria and associated models; Section 3 describes the parameter estimation approach to FDD in more detail; Section 4 presents an application example from the field of aircraft control; finally, some concluding remarks are given in Section 5.

2. System Faults

2.1. Faults Classification

System faults are defined as unpermitted deviations of at least one characteristic property of a system variable from its normally accepted behavior. Such events lead to the inability of one or more system components to fulfil their intended purposes and can have catastrophic consequences if they are left unattended [5], [8].

From the point of view of their time dependency, faults are classified as abrupt, incipient or intermittent. Abrupt faults occur instantaneously, often as a result of hardware damage, and can be very severe. Incipient faults represent slow parametric changes often caused by equipment aging, are less severe, but are more difficult to detect. Intermittent faults are faults that appear and disappear repeatedly, for instance due to partially damaged wiring [9].

A different classification criterion is the location of the fault within the system. Generally speaking, faults can affect all the system components, namely the sensors, the actuators and the physical plant.
Sensor faults represent incorrect readings from the sensors that the system is equipped with; it is usually possible to perform rapid fault detection and diagnosis, as well as to ensure fault tolerance at the sensors system level, through hardware redundancy. Actuator faults represent partial or total loss of control action; given the increased cost associated to actuation equipment, it is not always possible to rely solely on hardware redundancy in order to ensure fault tolerance at this level. Any other faults that cannot fall within one of these two categories are referred to as component faults; they represent changes in the physical parameters of the system and are often due to structural damage [1].

Finally, faults can be classified either as additive or multiplicative, based on their influence on the dynamic behavior of the system; this will be discussed in further detail in the sequel.

2.2. Fault Models

Using the laws of physics, most engineering systems can be modeled by the following system of equations with differences

\[
\begin{align*}
    x[k + 1] &= g(x[k], u[k], d[k], f[k]), \\
    y[k] &= h(x[k], u[k], d[k], f[k]),
\end{align*}
\]

(1)

where \( g \) and \( h \) are nonlinear functions, \( x[k] \) denotes the vector of system states at the sampling instant \( k \in 0, 1, \ldots \), \( u[k] \) represents the vector of control inputs, \( y[k] \) is the vector of measured system outputs, \( d[k] \) is the vector of unknown inputs which usually contains unmeasurable disturbances, while \( f[k] \) is the vector of faults [8].

However, from the viewpoint of the automatic control problem, the system in its general nonlinear form (1) is difficult to manipulate. For this reason, when the system operates under nominal conditions that are characterized by a state of equilibrium, its corresponding linear form can be obtained through the Taylor series decomposition of functions \( g \) and \( h \) around the equilibrium point, followed by an approximation that consists of keeping only the first order terms of the series.

In this manner, in the vicinity of the equilibrium point, one can obtain the Linear Time Invariant (LTI) model that approximates the behavior of the general system (1) as

\[
\begin{align*}
    x[k + 1] &= A_n x[k] + B_n u[k] + E_d d[k] + E_f f[k], \\
    y[k] &= C_n x[k] + G_d d[k] + G_f f[k].
\end{align*}
\]

(2)

In this representation, the fault \( f[k] \) appears as an additional input into the linear model, thus being classified as additive. Also, in the fault free case, \( f[k] = 0, \forall k \geq 0 \), while in the faulty case, \( f[k] \neq 0, \forall k \geq k_f \), where \( k_f \) is the moment when the matrices \( E_f \) and \( G_f \) become available.
The representation in (2) can be used in order to model a wide range of faults, especially at the sensors and actuators level. Also, it has the advantage of offering the possibility to apply rather straightforward techniques in order to perform fault detection and localization [5].

On the other hand, the assumption that all influences within the system can be treated as additive does not always hold. Since the linear approximation of the general system behavior given by (2) holds in the vicinity of the equilibrium point only, special attention should be paid to any fault that may impair the utility of this linear representation.

In contrast to additive faults, multiplicative faults reflect in changes in the elements of the matrices from the state-space representation. Precisely, after the occurrence of the fault, the matrix triplet \((A_f, B_f, C_f)\) will describe the system behavior

\[
\begin{align*}
\dot{x}[k + 1] &= A_f x[k] + B_f u[k] + E_d d[k], \\
y[k] &= C_f x[k] + G_d d[k].
\end{align*}
\]

The representation in (3) is generally used to model component faults; it is more difficult to manipulate but it presents greater advantages from the viewpoint of fault identification / estimation. Also, it represents a more practical modeling approach for integrated real-time system diagnosis and system reconfiguration, which is currently a major concern for the international research community [2].

3. Fault Detection and Diagnosis Using Parameter Estimation Techniques

3.1. Nominal and Impaired System Behavior

First, assume that the dynamic behavior of the process under nominal functioning conditions is described by the LTI model

\[
\begin{align*}
\dot{x}_n[k + 1] &= A_n x_n[k] + B_n u_n[k], \\
y_n[k] &= C_n x_n[k],
\end{align*}
\]

having again \(x_n \in \mathbb{R}^n, u_n \in \mathbb{R}^m, y_n \in \mathbb{R}^p, A_n \in \mathbb{R}^{n \times n}, B_n \in \mathbb{R}^{n \times m}, \text{ and } C_n \in \mathbb{R}^{p \times n}\).

Next, consider a multiplicative fault which reflects in changes of the elements, but not in the structure of the matrix triplet from the nominal case. Precisely, the dynamic behavior of the impaired system is described by the LTI model

\[
\begin{align*}
\dot{x}_f[k + 1] &= A_f x_f[k] + B_f u_f[k], \\
y_f[k] &= C_f x_f[k],
\end{align*}
\]

having again \(x_f \in \mathbb{R}^n, u_f \in \mathbb{R}^m, y_f \in \mathbb{R}^p, A_f \in \mathbb{R}^{n \times n}, B_f \in \mathbb{R}^{n \times m}, \text{ and } C_f \in \mathbb{R}^{p \times n}\).

For the purpose of this paper, no unknown inputs were considered for the models describing the nominal and the impaired system behaviors, respectively.
3.2. Basic Design of Fault Detection and Diagnosis Schemes

The main functions performed by any model-based technical diagnosis system are Fault Detection, Fault Isolation, and Fault Identification / Estimation.

Fault Detection consists of determining whether a fault has occurred within the system and of providing the accurate time-stamp of such an event. This is a two-step procedure: first, particular features are generated, describing the behavior of the physical process; secondly, these features are evaluated against the ideal process behavior in order to detect the occurrence of a fault.

In the case of Fault Detection and Diagnosis (FDD) schemes that use parameter estimation techniques, the features are the estimates of the process parameters. Several techniques for features evaluation have been proposed over the years, but one of the most commonly used is the so-called threshold limit test. This consists of comparing the value of the features evaluation function, which can be the norm difference between the reference model parameters and the estimates of the real process parameters, with an appropriately chosen constant $T$ or time-varying $T[k]$ threshold. If the value of the evaluation function exceeds the limit imposed by the threshold, then it is concluded that a fault has occurred; otherwise, the system is considered fault-free.

Fault Isolation consists of determining which system component (or components) is (are) affected by the fault. This is usually performed either by building each feature so that it is sensitive to only one subset of faults and insensitive to the other ones or by designing the features as vectors that have a certain orientation within the features subspace and determining which particular fault signature is the closest to the feature.

No attempt was made in this paper to further analyze the problem of Fault Isolation, since parameter estimation techniques alone provide only a limited amount of information that is useful for this purpose. Recent studies suggest the use of simultaneous state and parameter estimation techniques in order to fulfil this function, see [2] for instance.

Fault Identification / Estimation consists of determining the structure and the parameters of the new model that describes the impaired process behavior.

Parameter estimation techniques offer a strong advantage from this point of view, since they allow for the on-line identification of the fault models; specific methods are presented in the sequel.
3.3. **State-Space Nonrecursive Least Squares Method**

Consider the evolution of the system output over \( N \) samples:

\[
\begin{align*}
y[k-N+1] &= CAx[k-N] + CBu[k-N], \\
y[k-N+2] &= CAx[k-N+1] + CBu[k-N+1], \\
&\vdots \\
y[k] &= CAx[k-1] + CBu[k-1].
\end{align*}
\]

(6)

In (6), the indices \( n \) or \( f \) do not appear explicitly but they apply to the matrix and vector elements for the nominal and the impaired case, respectively.

Further, an assumption is made that the system states are available at any given sample. Since matrix \( C \) describes the relation between the known states and the measured outputs of the system, it can be computed directly as in (5); the parameter estimation problem is then reduced to identifying the elements of matrices \( A \) and \( B \).

In this case, the equalities in (6) can be rewritten as

\[
Y[k] = \Psi[k] \hat{\Theta}[k] + E[k],
\]

(7)

where

\[
Y[k] = \begin{bmatrix}
y^T[k-N+1] \\
y^T[k-N+2] \\
\vdots \\
y^T[k]
\end{bmatrix}
\]

contains the measured values of the output variables over the last \( N \) readings,

\[
\Psi^T[k] = \begin{bmatrix}
\psi[k-N+1] & \psi[k-N+2] & \cdots & \psi[k]
\end{bmatrix}
\]

\[
= \begin{bmatrix}
x[k-N] & x[k-N+1] & \cdots & x[k-1] \\
u[k-N] & u[k-N+1] & \cdots & u[k-1]
\end{bmatrix}
\]

(9)

is the identification data matrix containing the known control inputs and system states, while

\[
\hat{\Theta}[k] = \begin{bmatrix}
\hat{A}^T C^T \\
\hat{B}^T C^T
\end{bmatrix}
\]

(10)

contains the values of the process parameters estimated up to the current sample; also, \( E[k] \in \mathbb{R}^{N \times m} \) represents the estimation error matrix associated to the current sample.

The nonrecursive parameter estimation of the Least Squares (LS) method

\[
\hat{\Theta}[k] = \left[ \Psi^T[k] \Psi[k] \right]^{-1} \Psi^T[k] Y[k]
\]

(11)

minimizes the sum of errors

\[
V_{LS} = \sum_{i=k-N+1}^{k} e^T[i] e[i],
\]

(12)
where $e[i] = y[i] - \hat{\Theta}^T[i] \psi[i]$ is the output error at sample $i$.

### 3.4. State-Space Recursive Least Squares Method

The Recursive Least Squares (RLS) parameter estimation algorithm results from the nonrecursive estimation equations for $\hat{\Theta}[k+1]$ and $\hat{\Theta}[k]$, precisely

$$
\hat{\Theta}[k+1] = \hat{\Theta}[k] + \gamma[k] \left( y[k+1] - \psi^T[k+1] \hat{\Theta}[k] \right),
$$

(13)

where $\gamma[k]$ is the correction vector given by

$$
\gamma[k] = P[k+1] \psi[k+1]
= \frac{1}{\psi^T[k+1] P[k] \psi[k+1] + 1} P[k] \psi[k+1],
$$

(14)

with

$$
P[k] = \left( \sum_{i=1}^{k} \psi[i] \psi^T[i] \right)^{-1}
$$

(15)

and

$$
P[k+1] = [I - \gamma[k] \psi^T[k+1]] P[k].
$$

(16)

A typical initialization of this algorithm consists of setting

$$
\hat{\Theta}[0] = O,
$$

$$
P[0] = \alpha I,
$$

with $\alpha$ a large scalar value ($10^2 \div 10^3$). However, in the context of on-line fault detection and diagnosis, a more practical approach consists of setting

$$
\hat{\Theta}_f[0] = \Theta_n,
$$

where $\Theta_n$ is the parameter matrix from the nominal case.

The algorithm converges to the optimal solution $\Theta^*$, which approximates best the real impaired process parameters with regard to the chosen criteria.

### 3.5. Exponential Weighting with Constant Forgetting Factor

This method is suitable for processes with parameters that vary over time as a result of internal or external influences. It can be incorporated with the RLS method by time-depending weighting of the squared errors in the minimization criterion

$$
V_{w,RLS}[k] = \sum_{i=k-N+1}^{k} w[i] e^T[i] e[i];
$$

(17)

this results in a fading memory of the estimation algorithm, which will then associate a higher level of impact to the latest estimations than to the older ones.

By the choice of

$$
w[i] = \lambda^{k-i}, \ 0 < \lambda < 1,
$$

(18)
the errors are weighted with smaller factors for $i = 1$, depending on the choice of $\lambda$, but these factors increase exponentially to 1 for $i = k$.

The estimation algorithm remains the same as for the simple RLS method, precisely (13), but the correction vector becomes

$$
\gamma[k] = \frac{1}{\psi^T[k + 1]P[k]\psi[k + 1] + \lambda}P[k]\psi[k + 1],
$$

with $P[k]$ given by (15) and

$$
P[k + 1] = \left[I - \gamma[k]\psi^T[k + 1]\right]P[k]\frac{1}{\lambda}.
$$

The choice of the forgetting factor $\lambda$ is very important and depends on the nature of the process. In the case of rapidly varying parameters, where only small noise is tolerated, this value should be small, whereas a larger value can be considered in the case of processes with rapidly varying parameters.

4. Application Example

4.1. Nominal System

Consider the open-loop longitudinal model of a Class III large transport aircraft Boeing 747 in Class I cruise flight. The aircraft flight condition is assumed to be a straight and constant level flight at a fixed altitude of 40000 ft and 0.8 Mach. The corresponding steady-state speed is 774 fps and the aircraft mass is 19792 slug. The CoG (center of gravity) position is considered to be (0.25, 0, 0) of m.a.c. (mean aerodynamic chord).

The aircraft longitudinal dynamics is described by the state vector

$$
x = \begin{bmatrix} u & w & q & \theta \end{bmatrix}^T,
$$

where $u$ (fps) and $w$ (fps) represent the inertial velocities in the $x -$ and $z -$ directions of the body-axis reference-frame $F_B$, while $q$ (rad/s) and $\theta$ (rad) represent respectively the pitch rate and the pitch angle. The control input $\delta_E$ (rad) is the elevator deflection.

The linearized discrete-time open-loop model of the system using a sample period $T_s = 0.5$ sec and a zero-order hold on the inputs method is given by the following set of matrices, see [10],

$$
A_n = \begin{bmatrix}
0.9964 & 0.0069 & -2.5172 & -16.0714 \\
-0.0305 & 0.7724 & 310.9238 & 0.2842 \\
0.0001 & -0.0004 & 0.7268 & -0.0005 \\
0 & -0.0001 & 0.4359 & 0.9999
\end{bmatrix},
$$

$$
B_n = \begin{bmatrix}
0.4635 & -105.5749 & -0.5027 & -0.1324
\end{bmatrix}^T,
$$

$$
C_n = I_4.
$$
4.2. Impaired System

Now consider an abrupt fault that does not compromise the closed-loop system stability. Let the impaired system be described by the following set of matrices

\[
A_f = \begin{bmatrix}
0.9466 & 0.0065 & -2.3914 & -15.2678 \\
-0.0290 & 0.7338 & 295.3776 & 0.2700 \\
0.0001 & -0.0004 & 0.6905 & -0.0005 \\
0 & -0.0001 & 0.4141 & 0.9499
\end{bmatrix},
\]

\[
B_f = \begin{bmatrix}
0.3476 \\
-79.1812 \\
-0.3770 \\
-0.0993
\end{bmatrix}^T,
\]

\[
C_f = C_n = I_4,
\]

that were obtained by considering a 25% loss-of-effectiveness in the elevator deflection, which can be modeled in the post-fault system by multiplying the control distribution matrix with \((1 - \tau_f) = 0.75\) - this leads to \(B_f = (1 - \tau_f)B_n\); also, the fault was considered to affect the system matrix so that \(A_f = (1 - \gamma_f)A_n\), where \(\gamma_f = 5\%\). The fault was introduced into the system at time \(t_f = t_0 + 125\sec\), where \(t_0 = 0\sec\) represents the initial simulation time.

4.3. System Diagnosis Using the Recursive Least Squares Method

Given the nature of the system diagnosis problem, the algorithm was implemented so that it runs continuously, regardless of the existence of faults within the system, thus providing estimates of the system parameters at each sampling step. The RLS algorithm initialization was performed by setting \(\hat{\Theta}[0] = O_{5,4} , P[0] = 10^2 \times I_5\).

The parameter matrix \(\hat{\Theta}[k]\) estimated at each sample represents the feature generated for the purpose of performing fault detection. The feature evaluation function is given by

\[
\|\Theta_n - \hat{\Theta}[k]\|_2,
\]

with

\[
\Theta_n = \begin{bmatrix}
A_n^T C_n^T \\
B_n^T C_n^T
\end{bmatrix}
\]

\[
\text{and } \hat{\Theta}[k] = \begin{bmatrix}
A^T[k] \hat{C}^T[k] \\
B^T[k] \hat{C}^T[k]
\end{bmatrix}.
\]

The evolution of the features evaluation function over the simulation time is illustrated in Fig. 1. The value of this function is close to zero in the fault-free case, after the algorithm converges to the optimal solution of the parameter estimation problem, but increases significantly when the fault occurs. By setting the threshold value at \(T = 10\), the fault is detected almost instantaneously and the corresponding alarm signal is available at time \(t_{FD} = 126.5\sec\). The evolution of the 2-norm of the output estimation error over the simulation time is illustrated in Fig. 2. A testing signal that approximates a white gaussian noise was added to the input in order to stimulate the process so that the parameter estimation algorithm would provide correct results.
It is obvious that, while the RLS method is highly efficient from the viewpoint of rapid fault detection, it is not suitable for fault diagnosis. The parameter estimation algorithm fails to converge to a solution that estimates the impaired system parameters with a satisfactory level of accuracy.

4.4. System Diagnosis Using the Recursive Least Squares Method with Exponential Weighting and Constant Forgetting Factor

The basic design and implementation of the system diagnosis scheme based on RLS with exponential weighting and constant forgetting factor remains unchanged compared to the classical approach presented in the subsection 4.3. The only difference is the introduction of the forgetting factor $\lambda$; this enables a higher weight to be associated to the latest identification data, thus minimizing the influence of older, out of date, estimations on the current one.

Fig. 3 illustrates the evolution of the features evaluation function for various values of $\lambda$. The results show that using a smaller value of $\lambda$ leads to a decrease in the time needed to compute the parameter estimates with an acceptable level of accuracy. The case of $\lambda = 1$ corresponds to the classical RLS method and is shown only for comparison.

For a value of $\lambda = 0.90$ and the same threshold value $T = 10$ as in the previous scenario, the fault detection alarm signal is available at time $t_{FD} = 125.5 \text{ sec}$.

Fig. 4 shows the evolution of the estimation error and further demonstrates the efficiency of this method in the context of FDD. Also, Fig. 5 illustrates the real and the estimated dynamic behaviors of the system state variables in response to a unitary step input.

The RLS method with exponential weighting and constant forgetting factor provides a solution that satisfies

$$\|y[k] - \hat{\Theta}^T[k]\psi[k]\|_2 < 10^{-3}$$
at time $t_{FDD} = 210.5\text{sec}$.

The estimated system matrices are

$$\hat{A} = \begin{bmatrix} 0.9468 & 0.0063 & -2.3565 & -15.1717 \\ -0.0291 & 0.7339 & 295.3635 & 0.2311 \\ 0.0001 & -0.0004 & 0.6905 & -0.0005 \\ 0 & -0.0001 & 0.4141 & 0.9498 \end{bmatrix} \quad \text{and} \quad \hat{B} = \begin{bmatrix} 0.3451 \\ -79.1802 \\ -0.3770 \\ -0.0993 \end{bmatrix},$$

while the estimation error is $4.0416 \times 10^{-4}$.
5. Conclusion

This paper focused on the use of parameter estimation techniques for the design of system diagnosis schemes that are suitable for real-time implementation. The main design principles were discussed and several approaches based on the Least Squares (LS) method were presented in detail. A procedure for performing fault detection and identification for multivariable systems using the Recursive Least Squares (RLS) method with exponential weighting and constant forgetting factor was also proposed. An application example from the field of aircraft control was used in order to illustrate the main theoretical principles presented in the paper. The simulation results show that, with a certain degree of tuning specific to the application, the method considered is suitable for solving the Fault Detection and the Fault Identification / Estimation problems within complex applications.

Acknowledgement

The work of Bogdan D. Ciubotaru has been co-funded by the Sectorial Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labor, Family and Social Protection through the Financial Agreement POSDRU/89/1.5/S/62557.

REFERENCES