Polska Akademia Nauk • Instytut Badań Systemowych

AUTOMATYKA STEROWANIE ZARZĄDZANIE

Książka jubileuszowa
z okazji
70-lecia urodzin

PROFESORA KAZIMIERZA MAŃCZAKA

pod redakcją
Jakuba Gutenbauma

Polska Akademia Nauk • Instytut Badań Systemowych

AUTOMATYKA
STEROWANIE
ZARZĄDZANIE

Książka jubileuszowa
z okazji
70-lecia urodzin

PROFESORA KAZIMIERZA MAŃCZAKA

pod redakcją
Jakuba Gutenbauma

Warszawa 2002

Książka jubileuszowa z okazji
70-lecia urodzin
Profesora Kazimierza MAŃCZAKA

Redaktor
prof. dr hab. inż. Jakub Gutenbaum

Warszawa 2002

ISBN 83-85847-78-2

Wydawca: Instytut Badań Systemowych PAN
ul. Newelska 6 01-447 Warszawa
http://www.ibspan.waw.pl

Opracowanie składopisu: Anna Gostyńska, Jadwiga Hartman

Druk: KOMO-GRAF, Warszawa
nakład 200 egz., 34 ark. wyd., 31 ark. druk.

OBSERVER-BASED APPROACHES TO FAULT DIAGNOSIS

Józef Korbicz and Marcin Witczak

Institute of Control and Computation Engineering
University of Zielona Góra
<J.Korbicz, M.Witczak@issi.uz.zgora.pl>

Abstract: *The design and application of the model-based fault diagnosis has received considerable attention during the last few decades. In such a task, the model of the real system of interest is utilized to provide estimates of certain measured and/or unmeasured signals. Due to the still increasing popularity of state-space models, the most popular approach to residual generation is to use observer. Irrespective of the identification method used, there is always the problem of model uncertainty, i.e. the model-reality mismatch. Thus, the observer-based fault diagnosis scheme should provide robustness to model uncertainty. The objective of this work is to review the well-known observers for both linear and non-linear systems.*

Keywords: *observers, non-linear systems, fault diagnosis, model uncertainty.*

1. Introduction

It is well-known that there is an increasing demand for modern systems to become more effective and reliable. This real world's development pressure transformed automatic control, initially perceived as the art of designing a satisfactory system, into the modern science that it is today. The observed increasing complexity of modern systems necessitates the development of new control techniques. Unlike in the past, modern control techniques should take into account the system's safety. This requirement goes beyond the normally accepted safety-critical systems of nuclear reactors and aircraft, where safety is of paramount importance, to less advanced industrial systems.

In a fault diagnosis task, the model of the real system of interest is utilized to provide estimates of certain measured and/or unmeasured signals. Then, in the most usual case, the estimates of the measured signals are compared with their originals, i.e. a difference between the original signal and its estimate is used to form a residual signal. This residual signal can then be employed for *Fault Detection and Isolation* (FDI).

Irrespective of the identification method used, there is always the problem of model uncertainty, i.e. the model-reality mismatch. Thus, the better the model used to represent a system behaviour, the better the chance of improving the reliability and performance in diagnosing faults. Unfortunately, disturbances as well as model uncertainty are inevitable in industrial systems, and hence there exists a pressure creating the need for robustness in fault diagnosis systems. This robustness requirement is usually achieved in the fault detection stage.

In the context of robust fault detection, many approaches have been proposed (Chen and Patton 1999, Patton and Korbicz 1999, Patton et al. 2000). Undoubtedly, the most common approach is to use robust observers, such as the Unknown Input Observer (UIO) (Alcorta et al. 1997, Chen and Patton 1999, Patton et al. 2000), which can tolerate a degree of model uncertainty and hence increase the reliability of fault diagnosis. Unfortunately, much of the work in this subject is oriented towards linear systems. This is mainly because of the fact that the theory for observers (or filters in the stochastic case) is especially well-developed for linear systems.

In addition to that, the existing non-linear extensions of the UIO (Alcorta et al. 1997, Chen and Patton 1999, Seligeer and Frank 2000) require a relatively complex design procedure, even for simple laboratory systems (Zolghardi et al. 1996). Moreover, they are usually limited to a very restricted class of systems.

Another problem is that, even for linear systems, the research concerning UIOs is strongly oriented towards deterministic systems. Indeed, the question of detecting and isolating faults for systems with both modelling uncertainty and the noise has not attracted enough research attention, although most fault diagnosis systems suffer from both modelling uncertainty and the noise. The existing approaches (Chen and Patton 1999, Keller and Darouach 1999), which can be applied to linear stochastic systems, rely on a similar idea to that of the classical Kalman Filter (KF) (Anderson and Moore 1979, Korbicz and Bidyuk 1993).

Another problem arises from the application of fault diagnosis to non-linear stochastic systems. Unfortunately, the only existing approaches to this class of systems consist in the application of the Extended Kalman Filter (EKF). Indeed, the non-linear extensions of the UIO (Alcorta et al. 1997, Chen and Patton 1999, Seliger and Frank 2000) can only be applied to non-linear deterministic systems.

The work is organized as follows. Section 2 introduces the basic terminology and concepts of fault diagnosis. Subsequently, in Sections 3 and 4, various observers for both linear and non-linear systems are reviewed with special attention on their drawbacks and advantages. Finally, the last section is devoted to conclusions.

2. Observer-based residual generation

A fault can generally be defined as an unexpected change in a system of interest, e.g a sensor malfunction. All the unexpected variations that tend to degrade the overall performance of a system can also be interpreted as faults. Contrary to the term *failure*, which suggests a complete breakdown of the system, the term *fault* is used to denote a malfunction rather than a catastrophe.

Fault diagnosis can be viewed as a two-stage process, i.e. residual generation and decision making based on this residual (Fig. 1).

If the residuals are properly generated, then fault detection becomes a relatively easy task. Since without fault detection it is impossible to perform fault isolation and consequently fault identification, all efforts regarding the improvement of residual generation seem to be justified. This is the main reason why the research effort of this work is oriented towards fault detection and especially towards residual generation.

There have been many developments in model-based fault detection since the beginning of the 1970s, regarding both the theoretical

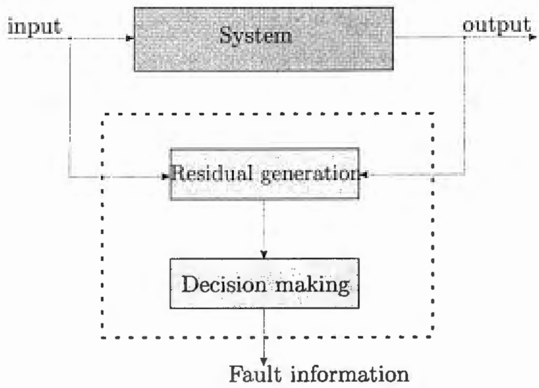

Fig. 1. A two-stage process of fault diagnosis.

context and the applicability to real systems (Chen and Patton 1999, Patton and Korbicz 1999, Patton et al. 2000). In almost all cases, the residual signal is obtained as a difference between system outputs and model outputs, i.e. $\mathbf{r_k} = \mathbf{y_k} - \hat{\mathbf{y}}_\mathbf{k}$. Thus, the better the model used to represent a system behaviour, the better the chance of improving the reliability and performance in diagnosing faults. Unfortunately, disturbances as well as model uncertainty are inevitable in industrial systems, and hence there exists a pressure creating the need for robustness in fault diagnosis systems. This robustness requirement is usually achieved in the fault detection stage, i.e. the problem is to develop residuals generators which should be insensitive (as far as possible) to model uncertainty and real disturbances acting on a system while remaining sensitive to faults.

Although the present work considers the discrete-time systems, some of the techniques are described in a continuous-time form. This is due to the fact that they were originally presented in such a form. However, most of them can relatively easily be applied to discrete-time systems.

The basic idea underlying the observer-based (or filter-based in the stochastic case) approaches to fault detection is to obtain the estimates of certain measured and/or unmeasured signals. Then, in the most usual case, the estimates of the measured signals are compared with their originals, i.e. a difference between the original signal and its estimate is used to form a residual signal $\mathbf{r_k} = \mathbf{y_k} - \mathbf{\hat{y}_k}$ (Fig. 2). To tackle this problem, many different observers (or filters) can be employed, e.g. Luenberger observers, Kalman filters, etc.

In most robust observer-based fault detection schemes, the problem of robustness to both model uncertainty and real disturbances acting on a system has been tackled by the introduction of the concept of an unknown input (Alcorta et al. 1997, Chen and Patton 1999, Selinger and Frank 2000). In spite of this, there is a large spectrum of candidate solutions. This is the main reason why observer-based approaches deserve special attention.

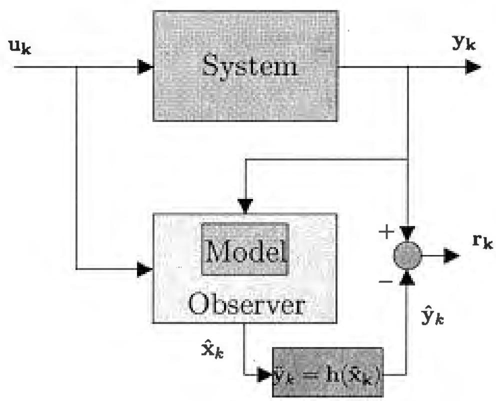

Fig. 2. The principle of observer-based residual generation.

3. Observers for linear systems

3.1. Luenberger observers and Kalman filters

Let us consider a linear system described by the following state-space equations:

$$\mathbf{x_{k+1}} = \mathbf{A_k x_k} + \mathbf{B_k u_k} + \mathbf{L_{1,k} f_k}, \quad (1)$$

$$\mathbf{y_{k+1}} = \mathbf{C_{k+1} x_{k+1}} + \mathbf{L_{2,k+1} f_{k+1}}. \quad (2)$$

According to the observer-based residual generation scheme (Fig. 2), the residual signal can be given as:

$$\begin{aligned}\mathbf{r_{k+1}} &= \mathbf{y_{k+1}} - \mathbf{\hat{y}_{k+1}} = \mathbf{C_{k+1}}[\mathbf{x_{k+1}} - \mathbf{\hat{x}_{k+1}}] + \mathbf{L_{2,k+1}f_{k+1}} \\ &= \mathbf{C_{k+1}}[\mathbf{A_k} - \mathbf{K_{k+1}C_{k+1}}][\mathbf{x_k} - \mathbf{\hat{x}_k}] + \mathbf{C_{k+1}L_{1,k}f_k} \\ &\quad -\mathbf{C_{k+}K_{k+1}L_{2,k}f_k} + \mathbf{L_{2,k+1}f_{k+1}}. \end{aligned} \tag{3}$$

To tackle the state estimation problem, the Luenberger observer can be used, i.e.

$$\mathbf{\hat{x}}_{k+1} = \mathbf{A_k\hat{x}_k} + \mathbf{B_ku_k} + \mathbf{K_{k+1}}(\mathbf{y_k} - \mathbf{\hat{y}_k}), \tag{4}$$

where $\mathbf{K_k}$ stands for the so-called gain matrix and it should be obtained in such a way as to ensure an asymptotic convergence of the observer, i.e. $\lim_{k\to\infty}(\mathbf{x_k} - \mathbf{\hat{x}_k}) = \mathbf{0}$ (Paraskevopolous 1996). If this is the case, i.e. $\mathbf{\hat{x}}_k \to \mathbf{x_k}$, the state estimation error $\mathbf{x_k} - \mathbf{\hat{x}_k}$ approaches zero and hence the residual signal (3) is only affected by the fault vector $\mathbf{f_k}$.

A similar approach can be realized in a stochastic setting, i.e. for the systems which can be modelled by:

$$\mathbf{x_{k+1}} = \mathbf{A_kx_k} + \mathbf{B_ku_k} + \mathbf{L_{1,k}f_k} + \mathbf{w_k}, \tag{5}$$

$$\mathbf{y_{k+1}} = \mathbf{C_{k+1}x_{k+1}} + \mathbf{L_{2,k+1}f_k} + \mathbf{v_k}, \tag{6}$$

where $\mathbf{w_k}$ and $\mathbf{v_k}$ are zero-mean white noise sequences with covariance matrices $\mathbf{Q_k}$ and $\mathbf{R_k}$, respectively. In this case, the observer structure can be similar to that of the Luenberger observer (4). To tackle the state estimation problem, the celebrated Kalman filter can be employed (Anderson and Moore 1979, Korbicz and Bidyuk 1993). Finally, the residual signal can be given as:

$$\begin{aligned}\mathbf{r_k} &= [\mathbf{I} - \mathbf{K_{k+1}C_{k+1}}]\mathbf{A_k}[\mathbf{x_k} - \mathbf{\hat{x}_k}] + [\mathbf{I} - \mathbf{K_{k+1}C_{k+1}}]\mathbf{L_{1,k}f_k} \\ &- \mathbf{K_{k+1}L_{2,k}f_{k+1}} + [\mathbf{I} - \mathbf{K_{k+1}C_{k+1}}]\mathbf{w_k} - \mathbf{K_{k+1}v_{k+1}}. \end{aligned} \tag{7}$$

Since the state estimate $\mathbf{\hat{x}}_k$ approaches the real state $\mathbf{x_k}$ (in the mean sense) asymptotically, i.e. $\mathcal{E}(\mathbf{\hat{x}}_k) \to \mathbf{x_k}$, the residual signal is only affected by the faults and the noise.

In both the deterministic (the Luenberger observer) and stochastic (the Kalman filter) cases fault detection can be performed by checking that the residual norm $\|\mathbf{r_k}\|$ exceeds a prespecified threshold, i.e.

$\|\mathbf{r_k}\| > \varepsilon_\mathbf{H}$. In the stochastic case, it is also possible to use more sophisticated, hypothesis-testing approaches such as Generalized Likelihood Ration Testing (GLRT) or Sequential Probability Ratio Testing (SPRT) (Willsky and Jones 1976, Basseville and Nikiforov 1993).

The presented approaches, in spite of their considerable usefulness, suffer from the lack of robustness to model uncertainty. Indeed, in both cases a perfect model of the system is assumed. This problem will be considered in the subsequent sections, where model uncertainty and real disturbances acting on a system are represented by the so-called unknown input.

3.2. Unknown input observers

Let us consider a linear system described by the following state-space equations:

$$\mathbf{x_{k+1}} = \mathbf{A_k x_k} + \mathbf{B_k u_k} + \mathbf{E_k d_k} + \mathbf{L_{1,k} f_k}, \quad (8)$$

$$\mathbf{y_{k+1}} = \mathbf{C_k x_{k+1}} + \mathbf{L_{2,k+1} f_{k+1}}, \quad (9)$$

where the term $\mathbf{E_k d_k}$ stands for model uncertainty as well as real disturbances acting on the system. The general structure of an UIO can be given as (Chen and Patton 1999):

$$\mathbf{z_{k+1}} = \mathbf{F_{k+1} z_k} + \mathbf{T_{k+1} B_k u_k} + \mathbf{K_{k+1} y_k}, \quad (10)$$

$$\hat{\mathbf{x}}_{k+1} = \mathbf{z_{k+1}} + \mathbf{H_{k+1}}, \quad (11)$$

then (assuming the fault-free mode, i.e. $\mathbf{f_k} = \mathbf{0}$) the state estimation error is:

$$\mathbf{e_{k+1}} = \mathbf{F_{k+1} e_k} + [\mathbf{I} - \mathbf{H_{k+1} C_{k+1}}]\mathbf{E_k d_k}. \quad (12)$$

From the above equation, it is clear that to decouple the effect of an unknown input from the state estimation error (and consequently from the residual), the following relation should be satisfied:

$$[\mathbf{I} - \mathbf{H_{k+1} C_{k+1}}]\mathbf{E_k} = \mathbf{0}. \quad (13)$$

The necessary condition for the existence of a solution to (13) and is given in (Chen and Patton 1999, p. 72, Lemma 3.1). The remaining task is to design the matrix $\mathbf{K_{1,k+1}}$ so as to ensure the convergence of the observer. This can be realized in a similar way as it is done in the case of the Luenberger observer. Finally, the state estimation error and the residual are given by:

$$\begin{aligned} \mathbf{e}_{k+1} &= \mathbf{F}_{k+1}\mathbf{e}_k + \mathbf{T}_{k+1}\mathbf{L}_{1,k}\mathbf{f}_k \\ &- \mathbf{H}_{k+1}\mathbf{L}_{2,k+1}\mathbf{f}_{k+1} - \mathbf{K}_{1,k+1}\mathbf{L}_{2,k}\mathbf{f}_k, \qquad (14) \\ \mathbf{r}_{k+1} &= \mathbf{C}_{k+1}\mathbf{e}_{k+1} + \mathbf{L}_{2,k+1}\mathbf{f}_{k+1}. \qquad (15) \end{aligned}$$

Since the Kalman filter constitutes a stochastic counterpart of the Luenberger observer, a stochastic counterpart of the UIO can also be developed. A detailed description of such an observer can be found in (Chen and Patton 1999).

3.3. An eigenstructure assignment approach

This section focuses on the problem of designing robust observers using the eigenstructure (eigenvectors and eigenvalues) assignment (Chen and Patton 1999, Chapter 4). The description of the system being considered has the following (continuous-time) state-space form:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{L}_1\mathbf{f}(t) + \mathbf{E}\mathbf{d}(t), \qquad (16) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{L}_2\mathbf{f}(t). \qquad (17) \end{aligned}$$

The observer-based residual generator can be given as:

$$\begin{aligned} \dot{\hat{\mathbf{x}}}(t) &= (\mathbf{A} - \mathbf{K}\mathbf{C})\hat{\mathbf{x}}(t) + (\mathbf{B} - \mathbf{K}\mathbf{D})\mathbf{u}(t) + \mathbf{K}\mathbf{y}(t), \qquad (18) \\ \hat{\mathbf{y}}(t) &= \mathbf{C}\hat{\mathbf{x}}(t) + \mathbf{D}\mathbf{u}(t), \qquad (19) \\ \mathbf{r}(t) &= \mathbf{Q}(\mathbf{y}(t) - \hat{\mathbf{y}}(t)), \qquad (20) \end{aligned}$$

where $\mathbf{Q} \in \mathbb{R}^{\mathbf{p}\times\mathbf{m}}$ $(\mathbf{p} \leq \mathbf{m})$ stands for the residual weighting matrix, which constitutes additional design freedom. When the above residual generator (18)–(20) is applied to the system (16)–(17), then the residual is given by:

$$\begin{aligned} \mathbf{r}(\mathbf{s}) &= \mathbf{Q}\mathbf{L}_2\mathbf{f}(\mathbf{s}) + \mathbf{Q}\mathbf{C}(\mathbf{s}\mathbf{I} - \mathbf{A}\mathbf{K}\mathbf{C})^{-1}(\mathbf{L}_1 - \mathbf{K}\mathbf{L}_2)\mathbf{f}(\mathbf{s}) \\ &+ \mathbf{Q}\mathbf{C}(\mathbf{s}\mathbf{I} - \mathbf{A}\mathbf{K}\mathbf{C})^{-1}\mathbf{E}\mathbf{d}(\mathbf{t}). \qquad (21) \end{aligned}$$

As can be seen from (21), in order to decouple the unknown input from the residual the following relation should be satisfied:

$$\mathbf{Q}\mathbf{C}(\mathbf{s}\mathbf{I} - \mathbf{A}\mathbf{K}\mathbf{C})^{-1}\mathbf{E}\mathbf{d}(\mathbf{t}) = \mathbf{0}. \qquad (22)$$

There are, of course, several different approaches which can be applied to solve (22), e.g. (Chen and Patton 1999, Theorem 4.4, p. 127).

4. Observers for non-linear systems

Model linearization is a straightforward way of extending the applicability of linear techniques to non-linear systems. On the other hand, it is well known that such approaches work well when there is no large mismatch between the linearized model and the non-linear system. Two types of linearization can be distinguished, i.e. linearization around the constant state and linearization around the current state estimate. It is obvious that the second type of linearization usually yields better results. Unfortunately, during such linearization the influence of terms higher than linear is usually neglected (as in the case of the extended Luenberger observer and the extended Kalman filter). This disqualifies such approaches for most practical applications. Such conditions have led to the development of linearization-free observers for non-linear systems.

This section briefly reviews the most popular observer-based residual generation techniques for non-linear systems. Their advantages, drawbacks as well as robustness to model uncertainty are discussed.

4.1. Extended Luenberger observers and Kalman filters

Let us consider a non-linear discrete-time system modelled by the following state-space equations:

$$\mathbf{x}_{k+1} = \mathbf{g}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{L}_{1,k}\mathbf{f}_k, \tag{23}$$

$$\mathbf{y}_{k+1} = \mathbf{h}(\mathbf{x}_{k+1}) + \mathbf{L}_{2,k+1}\mathbf{f}_{k+1}. \tag{24}$$

In order to apply the Luenberger observer presented in Section 3.1, it is necessary to linearize equations (23) and (24) around either a constant value (e.g. $\mathbf{x} = \mathbf{0}$) or the current state estimate $\hat{\mathbf{x}}_k$. This second approach seems to be more appropriate as it improves its approximation accuracy as $\hat{\mathbf{x}}_k$ tends to $\mathbf{x}_k$. In this case the approximation can be realized as follows:

$$\mathbf{A}_k = \left.\frac{\partial \mathbf{g}(\mathbf{x}_k, \mathbf{u}_k)}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k=\hat{\mathbf{x}}_k}, \quad \mathbf{C}_k = \left.\frac{\partial \mathbf{h}(\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k=\hat{\mathbf{x}}_k}. \tag{25}$$

As a result of using the Luenberger observer (4), the state estimation error takes the form:

$$\mathbf{e}_{k+1} = [\mathbf{A}_{k+1} - \mathbf{K}_{k+1}\mathbf{C}_k]\mathbf{e}_k + \mathbf{L}_{1,k}\mathbf{f}_k - \mathbf{K}_{k+1}\mathbf{L}_{2,k}\mathbf{f}_k + \mathbf{o}(\mathbf{x}_k, \hat{\mathbf{x}}_k), \tag{26}$$

where $\mathbf{o}(\mathbf{x_k}, \hat{\mathbf{x}}_\mathbf{k})$ stands for the linearization error caused by the approximation (25).

Because of a highly time-varying nature of $\mathbf{A_{k+1}}$ and $\mathbf{C_k}$ as well as the linearization error $\mathbf{o}(\mathbf{x_k}, \hat{\mathbf{x}}_\mathbf{k})$, it is usually very difficult to obtain an appropriate form of the gain matrix $\mathbf{K_{k+1}}$. This is the main reason why this approach is rarely used in practice.

As the Kalman filter constitutes a stochastic counterpart of the Luenberger observer, the extended Kalman filter can also be designed for the following class of non-linear systems:

$$\mathbf{x_{k+1}} = \mathbf{g}(\mathbf{x_k}, \mathbf{u_k}) + \mathbf{L_{1,k}f_k} + \mathbf{w_k}, \tag{27}$$

$$\mathbf{y_{k+1}} = \mathbf{h}(\mathbf{x_{k+1}}) + \mathbf{L_{2,k+1}f_{k+1}} + \mathbf{v_{k+1}}, \tag{28}$$

where, similarly to the linear case, $\mathbf{w_k}$ and $\mathbf{v_k}$ are zero-mean white noise sequences. Using the linearization (25) and neglecting the influence of the linearization error, it is straightforward to use the Kalman filter algorithm described in Section 3.1. The main drawback to such an approach is that it works well only when there is no large mismatch between the model linearized around the current state estimate and the non-linear behaviour of the system.

The EKF can also be used for deterministic systems, i.e. as an observer for the system (23)–(24) (Boutayeb and Aubry 1999). In this case, the noise covariance matrices can be set almost arbitrarily. As was proposed in (Boutayeb and Aubry 1999), this possibility can be used to increase the convergence of an observer.

Apart from difficulties regarding linearization errors, similarly to the case of linear systems, the presented approaches do not take model uncertainty into account. This drawback disqualifies those techniques for most practical applications. Although there are applications for which such techniques work with an acceptable efficiency, e.g. (Kowalczuk and Gunawickrama 2000).

4.2. The Tau observer

The observer proposed by Tau (1973) can be applied to a special class of non-linear systems which can be modelled by the following state-space equations:

$$\dot{\mathbf{x}}(t) = \mathbf{Ax(t)} + \mathbf{Bu(t)} + \mathbf{L_1 f(t)} + \mathbf{g(x(t), u(t))}, \qquad (29)$$

$$\mathbf{y(t)} = \mathbf{Cx(t)} + \mathbf{L_2 f(t)}. \qquad (30)$$

This special model class can represent systems with both linear and non-linear parts. The non-linear part is continuously differentiable and locally Lipschitz, i.e.

$$\|\mathbf{g(x(t), u(t))} - \mathbf{g(\hat{x}(t), u(t))}\| \leq \gamma \|\mathbf{x(t)} - \mathbf{\hat{x}(t)}\|. \qquad (31)$$

The structure of the Tau observer can be given as:

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{Ax(t)} + \mathbf{Bu(t)} + \mathbf{g(\hat{x}(t), u(t))} + \mathbf{K(y(t)} - \mathbf{\hat{y}(t))}, \qquad (32)$$

$$\hat{\mathbf{y}}(t) = \mathbf{C\hat{x}(t)}, \qquad (33)$$

where $\mathbf{K} = \mathbf{P}_\theta^{-1}\mathbf{C}^\mathbf{T}$, and $\mathbf{P}_\theta$ is the solution to the Lyapunov equation:

$$\mathbf{A^T P_\theta} + \mathbf{P_\theta A} - \mathbf{C^T C} + \theta \mathbf{P_\theta} = \mathbf{0}, \qquad (34)$$

where θ is a positive parameter, chosen in such a way as to ensure a positive definite solution of (34). Moreover, the Lipschitz constant γ should satisfy the following condition (Schreier et al. 1997):

$$\gamma < \frac{1}{2}\frac{\underline{\sigma}\left(\mathbf{C^T C} + \theta \mathbf{P_\theta}\right)}{\bar{\sigma}\left(\mathbf{P_\theta}\right)}, \qquad (35)$$

where $\bar{\sigma}(\cdot)$ and $\underline{\sigma}(\cdot)$ stand for the maximum and minimum singular values, respectively.

In spite of the fact that the design procedure does not require any linearization, the conditions regarding the Lipschitz constant γ are rather restrictive. This may limit any practical application of such an approach. Another difficulty arises from the lack of robustness to model uncertainty.

4.3. Observers for bilinear and low-order polynomial systems

A polynomial (and, as a special case, bilinear) system description is a natural extension to linear models. Design of observers for bilinear and low order polynomial (up to degree three) systems (Hac 1992, Hou and Pugh 1997, Kinneart 1999, Shields and Ashton 2000) involve only solutions of non-linear algebraic or Ricatti equations. This allows on-line residual generation.

Let us consider a bilinear continuous-time system modelled by the following state-space equations:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(\mathbf{t}) + \sum_{\mathbf{i}=1}^{\mathbf{r}} \mathbf{B}_{\mathbf{i}}\mathbf{u}_{\mathbf{i}}(\mathbf{t})\mathbf{x}\mathbf{t} + \mathbf{E}_1\mathbf{d}(\mathbf{t}), \tag{36}$$

$$\mathbf{y}(\mathbf{t}) = \mathbf{C}\mathbf{x}(\mathbf{t}) + \mathbf{E}_1\mathbf{d}(\mathbf{t}). \tag{37}$$

With a slight abuse of notation, the influence of faults is neglected. However, faults can very easily be introduced without changing the design procedure.

An observer for the system (36)–(37) can be given as (Hou and Pugh 1997):

$$\dot{\zeta}(t) = \mathbf{F}\zeta(\mathbf{t}) + \mathbf{G}\mathbf{y}(\mathbf{t}) + \sum_{\mathbf{i}=1}^{\mathbf{r}} \mathbf{L}_{\mathbf{i}}\mathbf{u}_{\mathbf{i}}(\mathbf{t})\mathbf{y}(\mathbf{t}), \tag{38}$$

$$\hat{\mathbf{x}}(t) = \mathbf{H}\zeta(\mathbf{t}) + \mathbf{N}\mathbf{y}(\mathbf{t}). \tag{39}$$

Hou and Pugh (1997) established the necessary conditions for the existence of the observer (38)–(39). Moreover, they proposed a design procedure involving a transformation of the original system (36)–(37) into an equivalent, quasi-linear one.

An observer for systems which can be described by the state-space equations consisting of both linear and polynomial terms was proposed in (Shields and Ashton 2000). Similarly to the case of the observer (36)–(37), here robustness to model uncertainty is tackled by means of an unknown input.

4.4. Non-linear unknown input observers

This section presents an extension of the unknown input observer for linear systems described in Section 3.2. Such an extension can be applied to systems which can be modelled by the following state-space equations:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{a}(\mathbf{x}(\mathbf{t})) + \mathbf{B}(\mathbf{x}(\mathbf{t}))\mathbf{u}(\mathbf{t}) + \mathbf{E}_1(\mathbf{x}(\mathbf{t}), \mathbf{u}(\mathbf{t}))\mathbf{d}(\mathbf{t}) \\ &+ \mathbf{K}_1(\mathbf{x}(\mathbf{t}), \mathbf{u}(\mathbf{t}))\mathbf{f}(t),\end{aligned} \tag{40}$$

$$\mathbf{y}(\mathbf{t}) = \mathbf{c}(\mathbf{x}(\mathbf{t})) + \mathbf{E}_2(\mathbf{u}(\mathbf{t}))\mathbf{d}(\mathbf{t}) + \mathbf{K}_2(\mathbf{x}(\mathbf{t}))\mathbf{f}(\mathbf{t}). \tag{41}$$

For notational convenience, the dependence of time t is neglected (e.g. $\mathbf{u} = \mathbf{u}(\mathbf{t})$).

The underlying idea is to design an unknown input observer for the system (40)-(41) without model linearization. For that purpose the following observer structure is proposed (Alcorta and Frank 1997, Selinger and Frank 2000):

$$\dot{\hat{\mathbf{z}}} = \mathbf{l}(\hat{\mathbf{z}}, \mathbf{y}, \mathbf{u}, \dot{\mathbf{u}}), \tag{42}$$

$$\mathbf{r} = \mathbf{m}(\hat{\mathbf{z}}, \mathbf{y}, \mathbf{u}), \tag{43}$$

where

$$\mathbf{z} = \mathbf{T}(\mathbf{x}, \mathbf{u}). \tag{44}$$

From (40)-(41) and (42)-(43), it can be seen that the residual $\mathbf{r}$ is governed by:

$$\mathbf{r} = \mathbf{m}(\mathbf{T}(\mathbf{x}, \mathbf{u}), \mathbf{c}(\mathbf{x}) + \mathbf{E}_2(\mathbf{u})\mathbf{d} + \mathbf{K}_2(\mathbf{x})\mathbf{f}, \mathbf{u}). \tag{45}$$

Taking the time derivative of (44) yields:

$$\dot{\mathbf{z}} = \frac{\partial \mathbf{T}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}}\dot{\mathbf{x}} + \frac{\partial \mathbf{T}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}}\dot{\mathbf{u}}. \tag{46}$$

Substituting (40) into (46) leads to:

$$\begin{aligned} \dot{\mathbf{z}} &= \frac{\partial \mathbf{T}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}}(\mathbf{a}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\mathbf{u} + \mathbf{E}_1(\mathbf{x}, \mathbf{u})\mathbf{d} + \mathbf{K}_1(\mathbf{x}, \mathbf{u})\mathbf{f}) \\ &+ \frac{\partial \mathbf{T}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}}\dot{\mathbf{u}}. \end{aligned} \tag{47}$$

From the above equation, it is clear that the unknown input decoupling condition can be stated as:

$$\forall \mathbf{x}, \mathbf{u} \quad \frac{\partial \mathbf{T}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}}\mathbf{E}_1(\mathbf{x}, \mathbf{u}) = \mathbf{0}. \tag{48}$$

The unknown input decoupling problem can now be resolved by analytically solving a set of linear first order partial differential equations (48).

Moreover, if any fault $\mathbf{f}$ is to be reflected by the transformed model, it must be required that:

$$\forall \mathbf{x}, \mathbf{u} \quad \mathrm{rank}\left(\frac{\partial \mathbf{T}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}}\mathbf{K}_1(\mathbf{x}, \mathbf{u})\right) = \mathrm{rank}\,(\mathbf{K}_1(\mathbf{x}, \mathbf{u})). \tag{49}$$

The effect of an unknown input can be decoupled from the output signal (41) in a similar way (Selinger and Frank 2000).

The main drawback to the proposed approach is that it requires a relatively complex design procedure, even for simple laboratory systems (Zolghardi et al. 1996). This may limit most practical applications of non-linear input observers. Other problems may arise from the application of the presented observer to non-linear discrete-time systems.

5. Conclusions

Observer-based techniques constitute one of the most popular ways of residual generation. As for almost all techniques, the robustness to model uncertainty and other factors which may lead to an unreliable fault detection is tackled by means of an unknown input. The popularity of observer comes also from the fact that they are widely used in modern control systems. This means that an observer can be employed for both fault diagnosis and control purposes. Such consideration leads directly to fault tolerant control.

There are efficient approaches to robust observer-based residual generation for linear systems (e.g. unknown input observers) but the existing solutions for non-linear systems are not mature yet. There are of course many approaches which can be applied to certain classes of non-linear systems, e.g. for bilinear or polynomial systems. However, this requirement limits the applicability of such approaches. On the other hand, the existing non-linear extensions of the UIO (Selinger and Frank 2000), which can be applied to a wider class of systems, require a relatively complex design procedure, even for simple laboratory systems (Zolghardi et al. 1996).

Notation

t	time
k	discrete time
$\mathbf{x_k}, \hat{\mathbf{x}}_\mathbf{k}$ $(\dot{\mathbf{x}}(\mathbf{t}), \dot{\hat{\mathbf{x}}}(\mathbf{t})) \in \mathbb{R}^\mathbf{n}$	state vector and its estimate
$\mathbf{y_k}, \hat{\mathbf{y}}_\mathbf{k} \in \mathbb{R}^\mathbf{m}$	output vector and its estimate
$\mathbf{e_k} \in \mathbb{R}^\mathbf{n}$	state estimation error
$\varepsilon_\mathbf{k} \in \mathbb{R}^\mathbf{m}$	output error (residual)
$\mathbf{u_k} \in \mathbb{R}^\mathbf{r}$	input vector
$\mathbf{d_k} \in \mathbb{R}^\mathbf{q}$	unknown input vector, $q \leq m$
$\mathbf{w_k}$, $\mathbf{v_k}$	process and measurement noise
$\mathbf{Q_k}$, $\mathbf{R_k}$	covariance matrices of $\mathbf{w_k}$ and $\mathbf{v_k}$
$\mathbf{f_k} \in \mathbb{R}^\mathbf{s}$	fault vector
$\mathbf{g}(\cdot)$, $\mathbf{h}(\cdot)$	non-linear functions
$\mathbf{E_k} \in \mathbb{R}^{\mathbf{n}\times\mathbf{q}}$	unknown input distribution matrix
$\mathbf{L_{1,k}}$, $\mathbf{L_{2,k}}$	faults distribution matrices

References

Alcorta Garcia E., Frank P. M. (1997) Deterministic nonlinear observer-based approaches to fault diagnosis. *Control. Eng. Practice.* **5**, 5, 663–670.

Anderson B.D.O., Moore J. B.(1979) *Optimal Filtering.* New Jersey: Prentice-Hall.

Basseville M., Nikiforov I.V.(1993) *Detection of Abrubt Changes: Theory and Applications.* New York: Prentice Hall.

Boutayeb M., Aubry D. (1999) *A strong tracking extended Kalman observer for nonlinear discrete-time systems.* IEEE Trans. Automat. Contr., **44**, 8, pp. 1550-1556.

Chen J., Patton R. J. (1999) *Robust Model-based Fault Diagnosis for Dynamic Systems.* London: Kluwer Academic Publishers.

Hac A. (1992) *Design of disturbance decoupled observers for bilinear systems.* ASME J. Dynamic Sys. Measure. Contr., **114**, 556-562.

Hou M., Pugh A.C. (1997) *Observing state in bilinear systems: an UIO approach.* – Proc. IFAC Symp.: *Fault Detection, Supervision and Safety of Technical Processes: SAFEPROCESS'97*, Hull, UK, **2**, 783-788.

Keller J. Y., Darouach M. (1999) *Two-stage Kalman estimator with unknown exogenous inputs.* Automatica, **35**, 339-342.

Kinneart M. (1999) *Robust fault detection based on observers for bilinear systems.* Automatica, **35**, 1829-1834.

Korbicz J., Bidyuk P. (1993) *State and Parameter Estimation. Digital and Optimal Filtering. Applications.* Zielona Góra: Technical University Press.

Kowalczuk Z., Gunawickrama K.(2000) *Leak detection and isolation for transission pipelines via non-linear state estimation.* Proc. IFAC Symp.: *Fault Detection, Supervision and Safety of Technical Processes: SAFEPROCESS 2000*, Budapest, Hungary, **2**, 943-948.

Paraskevopoulos P.N. (1991a) *Digital Control Systems.* London: Prentice Hall.

Patton R. J., Frank P., Clark R. N. (2000) *Issues of Fault Diagnosis for Dynamic Systems.*Berlin: Springer-Verlag.

Patton R. J., Korbicz J. (1999) *Advances in Computational Intelligence for Fault Diagnosis Systems.* Special issue of Int. J. Appl. Math. Comput. Sci., **9**, 3.

Schreier G., Ragot J., Patton R.J., Frank P.M. (1997) *Observer design for a class of non-linear systems.* Proc. IFAC Symp.: *Fault Detection, Supervision and Safety of Technical Processes: SAFEPROCESS'97*, Hull, UK, **1**, 483-488.

Seliger R., P. Frank (2000) *Robust observer-based fault diagnosis in non-linear uncertain systems.* In: Issues of Fault Diagnosis for Dynamic Systems (Patton, R. J., P. Frank and R. N. Clark, Eds.) Berlin: Springer-Verlag.

Shields D.N., Ashton S. (2000) *A fault detection observer method for non-linear systems.* Proc. IFAC Symp.: *Fault Detection, Supervision and Safety of Technical Processes: SAFEPROCESS 2000*, Budapest, Hungary, **1**, 226-231.

Tau F.E. (1973) *Observing the state of non-linear dynamic systems.* Int. J. Contr., **17**, 3.

Willsky A.S., Jones H.L. (1976) *A generalized likelihood ratio approach to the detection of jumps in linear systems.* IEEE Trans. Automat. Contr., **21**, 108-121.

Zolghardi A., Henry D., Monision M. (1996) *Design of nonlinear observers for fault diagnosis. A case study.* Control. Eng. Practice. **4**, 11, 1535-1544.

ISBN 83-85847-78-2