Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-13 (2016) 123–128

Parametric Identification by Means of Consequent Formation of a Conformed Estimations Set

Vladimir A. Fursov*. Andrey V. Gavrilov. Yegor V. Goshin*****

**Samara State Aerospace University, Samara, Russia (e-mail: fursov@ssau.ru)*
***Samara State Aerospace University, Samara, Russia (e-mail: Andrey.W.Gavrilov@gmail.com)*
****Image Processing Systems Institute of the RAS, Samara, Russia (e-mail: goshine@yandex.ru)*

Abstract: The article proposes an algorithm of parametric identification based on the principle of estimation conformity. The main idea is to form all possible estimations on lesser dimension subsystems, then the densest cloud of the estimations is determined. This cloud is used to find the required solution. We introduce a criterion that takes into account both mutual proximity of estimations and their quantity.

Keywords: Linear system identification; Detection, estimation; Non-Gaussian signals and noise.

1. INTRODUCTION

The most common approach to solving parametric identification problems is a statistical one. Within this approach the problem of model parameters estimation is considered the problem of determining the distribution parameters. Moreover, it is usually supposed that all data are samples from a normal collection and the task is to estimate some unknown, but *existing,* probability distribution.

This general scheme, called in this paper the "standard statistical a priori hypothesis", has been criticised by many authors. In particular, an alternative approach is proposed in Alimov (1969; 1966). Significant results have been obtained in robust statistics (Huber, 1964; Tsypkin et al., 1994). The most common robust algorithm, RANSAC (Fischler & Bolles, 1981), is widely and successfully used in problems of image matching (Torr & Murray, 1997). A number of new results is contained in a study by Kopysov (2012).

In cases when the number of observations is small, an approach based on the standard statistical a priori hypothesis is not fully applicable due to the fact that the a priori assumptions about statistical characteristics of probability distributions for a small number of observations are often wrong because of the statistical instability. In this paper, we consider an alternative solution to the problem based only on the assumptions which will be given in section 2. In previous papers (Fursov, 2002; Soifer, 2009) a method of conformed identification, based on these assumptions, is proposed. The idea of the method is to search the solution on some subsystem of lesser dimension, rather than on the initial system. It is reasonable to expect increase of identification accuracy if this subsystem consists of the most noise-free observations. Thus, the task is to find such a subsystem.

To find the most noise-free subsystem we earlier proposed an algorithm which considers two levels of subsystems. The upper-level subsystems are of smaller dimension than the original one, among which the most noise-free subsystem is searched. On each upper-level subsystem a set of all possible subsystems of small dimension (so called lower-level subsystems) is formed by the combination of rows. Then, the mutual proximity function is calculated from estimations on these lower-level subsystems. The upper-level subsystem, for which the value of this function is the lowest, is considered the most suitable for calculation of the required estimation.

Implementation of the described algorithm ensures higher accuracy of the identification if there is a number of observations from which the identification problem can be solved with the required accuracy.

Implementation of the described algorithm faces a number of difficulties associated with a rapid increase in computational complexity because of the increasing dimension of the original system and of the upper-level subsystems. Unfortunately, even with the use of parallel implementation, this algorithm is still computationally difficult.

In this paper we propose a new implementation scheme of the conformity principle. In particular, it is proposed to skip the construction of upper-level subsystems and to form a set of estimates directly on lower-level subsystems. Then, the densest cloud of estimates is searched on this set. This cloud is used to determine the required estimation.

2. PROBLEM DEFINITION

We consider the problem of estimation of the parameter vector $\mathbf{c}$ of the following linear model

$$\mathbf{y} = \mathbf{X}\mathbf{c} + \xi, \tag{1}$$

where $\mathbf{y}$ is an $N\times1$ sample vector, $\mathbf{X}$ is given (or formed from observations) $N\times M$ matrix, and $\xi = [\xi_1, \xi_2, ..., \xi_N]^T$ is an $N\times1$ vector of unknown errors.

The problem is to determine the estimate $\hat{\mathbf{c}}$ using matrix $\mathbf{X}$ and vector $\mathbf{y}$. In cases when there is no information about

10.1016/j.ifacol.2016.07.938

the errors, the least-squares method (LSM) is generally used (Alimov, 1969):

$$\hat{\mathbf{c}} = \left[\mathbf{X}^T\mathbf{X}\right]^{-1}\mathbf{X}^T\mathbf{y}. \quad (2)$$

It is known that the estimate of the LSM is unbiased and efficient if the standard statistical assumptions (Demidenko, 1981) are satisfied. However, in cases when the number of observations is small, these assumptions are often wrong because of the insufficient statistical stability of the probability characteristics. Papers by Fursov (2002) and Soifer (2009) consider a conformed identification method without a priori assumptions about measurement errors distribution.

(I) The matrix **X** and the vector **y** are fixed, i.e. $x_{i,j}, y_i$, $i = \overline{1,N}$, $j = \overline{1,M}$ are known from measurements on the single realisation.

(II) There is an uncertainty of the *conditioning* properties of the matrix **X** and the *statistical* characteristics of the vector $\boldsymbol{\xi}$.

(III) The norm of the error vector $\boldsymbol{\xi} = \left[\varepsilon_1, \varepsilon_2, ..., \varepsilon_N\right]^T$ is limited by:

$$\left\|\boldsymbol{\xi}\right\|_2 \le R_\xi, \quad (3)$$

i.e. $\boldsymbol{\xi} \in \boldsymbol{\Xi}$, $\quad \boldsymbol{\Xi} = \left\{\boldsymbol{\xi} : \left(\boldsymbol{\xi}^T\boldsymbol{\xi}\right)^{1/2} = \left\|\boldsymbol{\xi}\right\|_2 \le R_\xi\right\}$.

(IV) In system (3), the most noise-free subsystem is contained, for which the estimates

$\hat{\mathbf{c}}$ can be constructed with the desired accuracy**.**

(V) There is an exact model of a system corresponding to (2):

$$\mathbf{y}^* = \mathbf{X}\mathbf{c}, \quad (4)$$

where $\mathbf{y}^* = \mathbf{y} - \boldsymbol{\xi}$.

These assumptions specify the conditions when the proposed approach can be applied. In particular, the assumption (I) is necessary for the implementation of the scheme proposed below. The assumption (IV) allows us to search the most noise-free subsystem.

As for the a priori uncertainty of the conditioning properties in assumption (II) it is really typical for small subsystems. Regarding this fact, a preliminary conditioning estimation is performed. These issues were considered by authors in (Fursov & Gavrilov, 2004). In this paper, it is assumed that the preliminary elimination of ill-conditioned subsystems is performed and almost degenerate subsystems are absent.

The assumption (III) is also immaterial in the present paper. This assumption is important for estimated accuracy predictions. This issue was also covered in the previous papers. Despite the fact that the assumptions (II) and (III) are not essential within the issues this paper considers, they are given in the same way as in previous papers because they are necessary for the implementation of overall conformed identification technology.

Let us emphasize that the assumptions I, IV, V are the most essential in terms of the issues considered in this paper.

In contrast to the two-level scheme of identification method implementation described by Soifer (2009), in the present work we skip the stage of top-level subsystem formation and searching among them for the most noise-free one. In this case we only form a large number of small dimension subsystems, which are called lower-level subsystems in Soifer's paper. Instead of finding the most noise-free upper-level subsystem, we search the most conformed subset among all estimates on lower-level subsystems. This required subset must contain the least number of errors and it is used to form the required estimate.

It is further assumed that these subsystems have the smallest possible dimension, i.e. these subsystems are expressed in $M \times M$ square matrices $\mathbf{X}_k$. The algorithm for generating such subsystems can be described as a multiplication of system (1) on the left by the rectangular-$M \times N$ matrix $\mathbf{G}_k$, each row of which contains only one 1, while the remaining elements are 0s:

$$\mathbf{y}_k = \mathbf{X}_k\mathbf{c}_k + \boldsymbol{\xi}_k, \quad k = 1, 2, ... \quad (5)$$

where $\mathbf{y}_k = \mathbf{G}_k\mathbf{y}$, $\mathbf{X}_k = \mathbf{G}_k\mathbf{X}$, $\boldsymbol{\xi}_k = \mathbf{G}_k\boldsymbol{\xi}$.

In this case, each subsystem (5) will include all rows of the original system (1), the numbers of which coincide with the numbers of matrix $\mathbf{G}_k$ columns containing 1s. For given matrix dimensions, the set of all possible subsystems will consist of C_N^M subsystems. By calculating estimate (2) for each thus formed subsystem (3), we can get the set C_N^M of all possible estimates (Lawson & Hanson, 1974).

Since it is assumed that each subsystem contains only M lines, it is possible that among them either degenerate or ill-conditioned subsystems will randomly appear. One reason for the ill-conditioning is an almost linear dependence of subsystems' rows. The solution of the identification problem is a point of intersection of hyperplanes in the space of parameters, which are expressed by the lines of system (1) and/or subsystem (3).

Thus, we proceed from the fact that as a result of calculation of estimates on the set of all possibilities (except for the pre-eliminated ill-conditioned subsystems), a set Θ, consisting of K estimates $\hat{\mathbf{c}}_k$, is formed:

$$\Theta = \left\{\hat{\mathbf{c}}_k \in \Theta\right\} \quad k = \overline{1,K}, \quad K \le C_N^M. \quad (6)$$

The next task is to select a subset Θ_L of suitable estimates from the set:

$$\Theta_L = \left\{\hat{\mathbf{c}}_l \in \Theta_L\right\} \quad \Theta_L \subset \Theta, \quad l = \overline{1,L}, \quad L < K, \quad (7)$$