Applied Mathematical Modelling 37 (2013) 2166–2177

Contents lists available at SciVerse ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

# Conditional parameter identification with different losses of under- and overestimation

Piotr Kulczycki [a,b,*], Malgorzata Charytanowicz [a,c]

[a] Polish Academy of Sciences, Systems Research Institute, Centre of Information Technology for Data Analysis Methods, Poland
[b] Cracow University of Technology, Department of Automatic Control and Information Technology, Poland
[c] Catholic University of Lublin, Institute of Mathematics and Computer Science, Poland

ARTICLE INFO

*Article history:*
Received 29 August 2011
Received in revised form 27 April 2012
Accepted 7 May 2012
Available online 22 May 2012

*Keywords:*
Parameter identification
Asymmetrical losses of estimation errors
Conditional factors
Nonparametric estimation
Statistical kernel estimators
Numerical algorithm

ABSTRACT

In many scientific and practical tasks, the classical concepts for parameter identification are satisfactory and generally applied with success, although many specialized problems necessitate the use of methods created with specifically defined assumptions and conditions. This paper investigates the method of parameter identification for the case where losses resulting from estimation errors can be described in polynomial form with additional asymmetry representing different results of under- and overestimation. Most importantly, the method presented here considers the conditionality of this parameter, which in practice means its significant dependence on other quantities whose values can be obtained metrologically. To solve a problem in this form the Bayes approach was used, allowing a minimum expected value of losses to be achieved. The methodology was based on the nonparametric technique of statistical kernel estimators, which freed the investigated procedure from forms of probability distributions characterizing both the parameter under investigation and conditioning quantities. As a result an algorithm is presented, ready for direct use without further intensive research and calculations.


## 1. Introduction

Parameter identification [1], i.e. assigning a concrete value to a parameter present in a model, despite its very traditional nature, has still great significance in modern scientific and applicational problems. Moreover, its importance continuously increases together with the dominance of model-based methods and the growing, often specific, demands made on models used in science and practice. At the same time, the increasing complexity and novelty of current methods is accompanied by a decrease in the classical understanding of parameter identification as a task of fixing a concrete value of a parameter which exists objectively in reality but is unknown. Here, through investigating, the researcher attempts to get as close as possible to this "true" value. In fact more frequently in contemporary models, their particular parameters describe an entire range of complex phenomena, simplified in a model to one parameter, existing only formally – without concrete physical form. In this situation the quality of parameter identification cannot be evaluated by classical means, obtaining a value as near as possible to an imagined "true" parameter value (since it does not exist), but rather by accounting for the influence of particular parameter values on a considered system, whose part is the investigated model. This moves the mathematical apparatus applied here – present within point estimation – from classical mathematical statistics [2], towards the currently intensively-studied data analysis [3]. Fortunately, the development of modern sophisticated and often specific methods of

* Corresponding author at: Polish Academy of Sciences, Systems Research Institute, Centre of Information Technology for Data Analysis Methods, Poland.
*E-mail addresses:* kulczycki@ibspan.waw.pl, kulczycki@pk.edu.pl (P. Kulczycki).

0307-904X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.apm.2012.05.007

parameter identification is facilitated by the dynamic expansion of contemporary computer technology, supported on the theoretical side by the procedures of advanced information technology [4].

The subject of this paper is an algorithm for parameter identification, i.e. estimation of the value of a parameter occurring in a model, based on four premises:

1. minimization of expected value of losses arising from estimation errors, unavoidable in practice;
2. asymmetry of those losses, i.e. allowing for situations where losses occurring through underestimation are substantially different from losses resulting from overestimation;
3. arbitrariness of probability distributions appearing in the problem;
4. and finally–worth particularly highlighting–conditionality of an identified parameter, that is its significant dependence on a factor (or factors), with values that can be in practice obtained metrologically.

The realization of the first will be through application of the Bayes approach [5].

The second by assuming the loss function resulting from estimation errors, in the asymmetrical form

$$l(\hat{y}, y) = \begin{cases} (-1)^k a_l (\hat{y} - y)^k & \text{for} \quad \hat{y} - y \leqslant 0, \\ a_r (\hat{y} - y)^k & \text{for} \quad \hat{y} - y \geqslant 0, \end{cases} \tag{1}$$

with the given degree $k \in \mathbf{N} \setminus \{0\}$ , where the coefficients $a_l$ and $a_r$ are positive, while $y$ and $\hat{y}$ denote the values of the parameter under consideration and its estimator, respectively. The fact that the coefficients $a_l$ and $a_r$ may differ causes an asymmetry of the above function and enables the inclusion of different losses implied by over- and underestimation of the examined parameter. Limiting the form of function (1) to a polynomial seems not to decrease the generality of considerations in practical applications, offering an effective compromise between precision and complexity of results obtained. Moreover the possibility of change of the polynomial degree $k$ – with respect to that resulting from fundamental research – allows a differing scale of protection against large estimation errors.

The third aspect is realized by applying nonparametric methodology of statistical kernel estimators [6–8] for calculating probability characteristics.

Lastly – and worth highlighting once more – this paper is aimed at the conditional approach, i.e. where the value of the estimated parameter is strongly dependent on a conditional factor, for example in engineering practice it is often a current temperature. If the value of such a factor is metrologically available, then its inclusion can make the used model significantly more precise.

The goal of this paper is the provision of an algorithm for calculating a conditional parameter value, optimal in the sense of minimum expectation value of losses, in particular those different for under- and overestimation. The above value is determined for a fixed (most often current) value of a conditional factor, based on measurements of this parameter obtained earlier for different conditioning values. The algorithm is comprehensive and can be applied directly without detailed knowledge of theoretical aspects, laborious research or analytical calculations. It is sufficient data to take only the measurements of pairs of the model parameter value, and the conditional factor value for which this parameter value was obtained, as well as the quantities introduced in formula (1): the degree $k$ and the ratio of coefficients $a_l/a_r$.

Thus, Section 2 outlines the statistical kernel estimators method. The algorithm worked out is described in Sections 3 and 4, with the asymmetrical linear case in Section 3.1, the asymmetrical quadratic in Section 3.2, and the asymmetrical polynomial (in particular cubic) in Section 3.3. Finally, Section 5 presents the results of experimental verification of the investigated procedure. Section 6 provides a summary of the presented method.

The preliminary version of this paper was presented as [9].

## 2. Preliminaries: statistical kernel estimators

Let the $n$-dimensional random variable $X$ be given, with a distribution characterized by the density $f$. Its kernel estimator $\hat{f} : \mathbf{R}^n \to [0, \infty)$, calculated using experimentally obtained values for the $m$-element random sample

$$x_1, x_2, \ldots, x_m, \tag{2}$$

in its basic form is defined as

$$\hat{f}(x) = \frac{1}{mh^n} \sum_{i=1}^{m} K\left(\frac{x - x_i}{h}\right), \tag{3}$$

where $m \in \mathbf{N} \setminus \{0\}$, the coefficient $h > 0$ is called a smoothing parameter, while the measurable function $K : \mathbf{R}^n \to [0, \infty)$ of unit integral $\int_{\mathbf{R}^n} K(x)\, dx = 1$, symmetrical with respect to zero and having a weak global maximum in this place, takes the name of a kernel. The interpretation of the above definition is illustrated in Fig. 1 for a one-dimensional random variable. In the case of the single realization $x_i$, the function $K$ (transposed along the vector $x_i$ and scaled by the coefficient $h$) represents the approximation of distribution of the random variable $X$ having obtained the value $x_i$. For $m$ independent realiza-