Robust Akaike's Criterion for Model Order Selection

Vladimir Stojanovic[1*], Vojislav Filipovic[1]
[1]Faculty of Mechanical and Civil Engineering in Kraljevo, Department of Energetics and Automatic Control, University of Kragujevac (Serbia)

The paper considers the model order selection (Output Error model) of the system with constant parameters. Ad hoc selection of model order leads to overparametrization or parsimony problem. To avoid these problems, different selection criterions of the model are used: AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and FPE (Final Prediction Error Criterion). In this paper, Akaike's criterion is used, which is obtained by minimization of the Kullback-Leibler information distance. The criterion is basically a generalization of the maximum likelihood method. It is assumed that the stochastic disturbance in the model belongs to the class of ε-contaminated distributions. In such conditions the originally proposed AIC criterion cannot be applied. By determining the least favourable probability density for a given class of probability distribution represents a base for design of the robust version of AIC criterion. Simulations illustrate the behavior of the proposed criterion.

Keywords: model order selection, output error model, ε-contaminated distributions, robust Akaike's criterion

1. INTRODUCTION

Obtaining system models based on the fundamental laws of physics is a difficult problem. In order to facilitate the controller design, for the obtained model, different simplifications of the model are performed. Most often it is a procedure of linearization around the equilibrium point. However, there are the systems that cannot be linearized around an equilibrium point, because there is no equilibrium point. If a linear approximation is found, the resulting model will be valid only for a small region around the linearization point. As an alternative approach the design of controllers is largely based on the use of mathematical models that are obtained during the process of system identification [1,2].

Most identification algorithm assume that the model structure is a priori known. As is well known, a fundamental difficulty in statistical analysis is the choice of an appropriate model and determining the order of a model. In recent years, the necessity of introducing the concept of model has been recognized and the problem is posed how to choose the "best approximating" model among a class of competing models with different numbers of parameters by a suitable model selection criterion given a data set. Also, there is presently a great deal of interest in simple criteria represented by parsimony of parameters for choosing one of a set of competing models to describe a given data set. Therefore, the best model is the one with least complexity, or equivalently the highest information. For example, parameter parsimony requires that the smallest number of factors is chosen, such that the corresponding model fits the data. The selection of a parsimonious model, in general, is a nontrivial problem without the aid of model selection criteria.

Several information theoretic criterion have been proposed for structure selection in linear dynamic input output models. The model which minimizes the criterion is then chosen as the best model from the available set. Examples of the classical criterions are the Final Prediction Error (FPE), Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC). These techniques find a tradeoff between goodness of fit and model complexity. The performance of an order-selection criterion is optimal if the model of the selected order is the most accurate model in the considered set of estimated models. Note that this is not necessarily the true model order. If the true process is, e.g. tenth-order, where the last six parameters are insignificant, the estimated fourth-order model will be the most accurate.

Used way for deriving model selection criteria is based on the quantification of "how close are" the probability density of the generating model and the probability density of the fitted approximating model. Several coefficients or "measures" have been introduced in the literature for this quantification. The Kullback-Leibler information distance is the most frequently used information theoretic coefficient for measuring divergence or separation between two probability densities [3]. The Akaike's information criterion (AIC) is a commonly used tool for choosing between alternative models [4].

Here, those results are extended on the case when the measurement noise is a non-Gaussian. Justification of this approach was confirmed in practice [5,6]. Namely, in measurements there are rare, inconsistent observations with the largest part of population of observations (outliers). The presence of outliers can considerably degrade the performance of linearly recursive algorithms based on the assumptions that measurements have a Gaussian distribution.

The synthesis of robust algorithms is of primary interest. The synthesis is based on Huber's theory of robust statistics [6]. As a generator of a recursive algorithm, according Huber's theory, it is defined the functional based on the least favourable probability distribution for a given class of probability distribution. Robust recursive algorithms in the identification of

*Corresponding author: stojanovic.v@mfkv.kg.ac.rs