Automatica 60 (2015) 227–238

Contents lists available at ScienceDirect

Automatica

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

A randomized algorithm for nonlinear model structure selection☆

Alessandro Falsone [1], Luigi Piroddi, Maria Prandini

Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy

ARTICLE INFO

Article history:
Received 10 November 2014
Received in revised form
24 June 2015
Accepted 1 July 2015
Available online 7 August 2015

Keywords:
Model identification
Nonlinear models
Polynomial NARX models
Randomized algorithms
Regressor selection
Prediction error minimization methods

ABSTRACT

The identification of polynomial Nonlinear Autoregressive [Moving Average] models with eXogenous variables (NAR[MA]X) is typically carried out with incremental model building techniques that progressively select the terms to include in the model. The Model Structure Selection (MSS) turns out to be the hardest task of the identification process due to the difficulty of correctly evaluating the importance of a generic term. As a result, classical MSS methods sometimes yield unsatisfactory models, that are unreliable over long-range prediction horizons. The MSS problem is here recast into a probabilistic framework based on which a randomized algorithm for MSS is derived, denoted RaMSS. The method introduces a tentative probability distribution over models and progressively updates it by extracting useful information on the importance of each term from sampled model structures. The proposed method is validated over models with different characteristics by means of Monte Carlo simulations, which show its advantages over classical and competitor probabilistic MSS methods in terms of both reliability and computational efficiency.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

System identification is the process of building a mathematical model of a dynamical system from input–output data (Söderström & Stoica, 1989). In particular, the black-box identification of nonlinear systems is a singularly difficult and challenging problem, since it amounts to solving an optimization problem with a mixed combinatorial (model structure selection) and continuous (parameter estimation) nature, (Billings, 2013; Hong et al., 2008; Sjöberg et al., 1995).

We are here mainly concerned with recursive input/output (I/O) models of the Nonlinear Autoregressive (Moving Average) with eXogenous variables (NAR[MA]X) class, (Leontaritis & Billings, 1985), where the current value of the system output is obtained as a nonlinear functional expansion of lagged input and output (and possibly noise) terms. Polynomial NARX/NARMAX models have earned widespread interest in view of their flexibility and representation capabilities, and several applications are documented (see, *e.g.*, Aguirre, Donoso-Garcia, & Santos-Filho, 2000, Boaghe, Billings, Li, Fleming, & Liu, 2002, Chiras, Evans, & Rees, 2001, Leva & Piroddi, 2002, Loh & Duh, 1996, Palumbo & Piroddi, 2000, Piroddi, Farina, & Lovera, 2012, Pottmann, Unbehauen, & Seborg, 1993, Spinelli, Piroddi, & Li, 2004). Various identification algorithms have been proposed in the literature for NARX/NARMAX models, mainly based on the Prediction Error Minimization (PEM) framework for parameter estimation, which is in fact computationally convenient in view of the linear-in-the-parameters structure of the model. The main difficulty addressed by such algorithms is the selection of an appropriate model structure, considering that functional approximation using families of basis functions often leads to an exponential increase in the number of candidate model structures (curse of dimensionality), a critical issue of polynomial expansions in particular.

Classical model selection techniques based on information criteria, such as the AIC (Akaike Information Criterion), the BIC (Bayesian Information Criterion), and similar indices, appear to be hardly applicable in the nonlinear framework. Essentially, these indices weigh the model accuracy against the model size (number of parameters) and are used in the linear framework to estimate the correct model size. In the nonlinear context, no simple relation between model size and accuracy can be established, because one can construct many models of the same size with very different regressors and these can have quite different performances. As a result, these criteria cannot be used to derive indications on whether to accept or discard a specific term (Chen, Hong, & Harris,

☆ The material in this paper was partially presented at the 53rd IEEE Conference on Decision and Control, December 15–17, 2014, Los Angeles, CA, USA. This paper was recommended for publication in revised form by Associate Editor Juan I. Yuz under the direction of Editor Torsten Söderström.

E-mail addresses: alessandro.falsone@polimi.it (A. Falsone), luigi.piroddi@polimi.it (L. Piroddi), prandini@elet.polimi.it (M. Prandini).

[1] Tel.: +39 02 23994028; fax: +39 02 23993412.

http://dx.doi.org/10.1016/j.automatica.2015.07.023
0005-1098/© 2015 Elsevier Ltd. All rights reserved.

2003; Palumbo & Piroddi, 2000). Regularization criteria like the Least Absolute Shrinkage Selection Operator (LASSO) operate in a similar direction by penalizing the model size in the model identification process. As such they are effective in reducing the model size, but not necessarily in the more difficult task of selecting the appropriate model structure (Bonin, Seghezza, & Piroddi, 2010).

An efficient method for tackling the Model Structure Selection (MSS) task for NARX/NARMAX models was first suggested in Korenberg, Billings, Liu, and McIlroy (1988), based on an incremental model building procedure (*forward regression*). In detail, at each algorithm iteration a new term is included in the model based on an importance index, the Error Reduction Ratio (ERR), which evaluates the local accuracy improvement that can be gained by adding the term to the current model. The method also exploits Orthogonal Least Squares (OLS) to decouple the estimation of the various regressors. Accordingly, it is denoted Forward Regression Orthogonal Estimator (FROE). Several variants of this method have been introduced in the literature using both forward and backward regression schemes (see, *e.g.*, Bonin et al., 2010, Billings, Chen, & Korenberg, 1989, Farina & Piroddi, 2012, Guo, Guo, Billings, & Wei, 2015, Haber & Unbehauen, 1990, Hong et al., 2008, Li, Peng, & Irwin, 2005, Mao & Billings, 1999, Piroddi & Spinelli, 2003 and Wei & Billings, 2008). The combinatorial optimization performed by the FROE in the space of all possible models follows a greedy scheme, so that there is no guarantee of convergence to the global minimum. Several other drawbacks, that may ultimately prevent the convergence towards the correct model, have been pointed out, *e.g.*, in Aguirre and Billings (1995), Billings and Aguirre (1995) and Piroddi and Spinelli (2003). They are essentially related to the inadequacy of the ERR index to express in an absolute way the importance of a regressor. Indeed, such measure depends on the specific model to which the regressor is to be added. Notice also that the PEM paradigm guarantees the unbiasedness of the parameter estimates only in ideal conditions, where the system is persistently excited and the model structure (including the disturbance model) exactly matches that of the target system, a condition which is typically not met in the model building process, precisely because the model is constructed iteratively.

Recently, some novel approaches have been introduced to address the nonlinear identification problem, based on randomized algorithms (Tempo, Calafiore, & Dabbene, 2004). An algorithm based on the Expectation Maximization (EM) approach is presented in Baldacchino, Anderson, and Kadirkamanathan (2012), that employs the particle filter to handle nonlinearities and jointly perform MSS and parameter estimation. In Baldacchino, Anderson, and Kadirkamanathan (2013), both tasks are dealt with in a unified Bayesian framework, that is suitable for describing the uncertainty in both parameters and structure. Structure and parameter variations are performed based on a statistical acceptance/rejection mechanism. Posterior distributions are inferred using the Reversible Jump Markov Chain Monte Carlo (RJMCMC) procedure. The introduction of random sampling favors the convergence to the global minimum. However, MCMC methods are known to require a burn-in period for the Markov chain to converge to the desired stationary distribution, and this calls for many iterations.

In this paper, a novel iterative randomized algorithm is introduced for the identification of nonlinear systems, based on a different probabilistic reformulation of the MSS problem. The method is here described for NARX models only, although the extension to the NARMAX case can be envisaged (and is a matter of current research endeavors). A Bernoulli random variable is associated to each regressor. These random variables are assumed to be independent and, at each iteration, the proposed algorithm generates a set of models, each one being independently extracted from the joint distribution of all regressors. More precisely, an extracted model will contain a specific regressor if the value taken by the Bernoulli random variable associated to that regressor is 1. Then, the parameters of the extracted models are estimated, and the performances of the parameterized models evaluated in terms of a suitable index based on the prediction and simulation errors. Finally, the Bernoulli distribution of each regressor is updated based on the performances of the entire population of extracted models. More precisely, a regressor probability is increased if, on average, the extracted models that contain that specific regressor perform better than those that do not, and decreased in the opposite case. The algorithm converges to a limit distribution corresponding to a specific model structure. Some examples are analyzed by means of Monte Carlo simulations to show the effectiveness of the adopted probabilistic formulation, and to illustrate the improved reliability of the proposed algorithm compared to currently available randomized methods.

The proposed approach has some features in common with evolutionary methods, such as genetic algorithms (GA) (Rodriguez-Vazquez, Fonseca, & Fleming, 2004), in that it exploits randomness in choosing potential regressors and processes populations of models. More in detail, a GA selects the fittest individuals in the current population and manipulates them to generate a new population, using specific pair-wise operators. In our framework, the "fitness" of each regressor is evaluated from an aggregate analysis of the whole population. All individuals of the population contribute to the evaluation of the regressors, either reinforcing or discouraging their selection. Then, the new population is generated from scratch, based on the aggregate information derived from the current population.

A preliminary version of this work is given in Falsone, Piroddi, and Prandini (2014). The present paper significantly extends that contribution both from a theoretical and a methodological viewpoint. More specifically, the iterative algorithm is here better formalized within an appropriately defined probabilistic framework, and its convergence properties are established. Furthermore, a more extensive assessment of the performance of the proposed approach has been carried out via numerical examples taken from the literature. In particular, the behavior of the algorithm is analyzed in critical operating conditions, such as when a slowly varying input signal is used or when the randomized procedure is only allowed to get partial information on the correct model structure. In the case of a slowly varying input signal, the possible advantages related to the use of the simulation error for performance evaluation purposes are also discussed.

The rest of the paper is organized as follows. Section 2 provides the basic framework and notation for nonlinear system identification of NARX models and briefly reviews the main approaches in the literature. Section 3 discusses the crucial issue of how to evaluate the importance of each regressor. The proposed method is illustrated in Section 4 and then tested in Section 5. Finally, some concluding remarks are drawn in Section 6.

2. Preliminaries

2.1. The NARX model class

A NARX model (Leontaritis & Billings, 1985) is described by the following input/output recursive equation:

$$y(k) = f(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u)) + e(k) \quad (1)$$

where $y(k)$, $u(k)$, and $e(k)$ are the output, input, and (white) noise signals, respectively, n_y and n_u being suitable maximum lags, and $f(\cdot)$ is an unknown nonlinear function.

Download English Version:

https://daneshyari.com/en/article/695348

Download Persian Version:

https://daneshyari.com/article/695348

Daneshyari.com