Journal of Fundamental and Applied Sciences

ISSN 1112-9867

Available online at http://www.jfas.info

Research Article

Special Issue

PERFORMANCE OF PARAMETER-MAGNITUDE BASED INFORMATION CRITERION IN IDENTIFICATION OF LINEAR DISCRETE-TIME MODEL

M. F. Abd Samad*, A. R. M. Nasir

Faculty of Mechanical Engineering, UniversitiTeknikal Malaysia Melaka, 76100 Durian Tunggal, Malacca, Malaysia

Published online: 24 February 2018

ABSTRACT

Information criterion is an important factor for model structure selection in system identification. It is used to determine the optimality of a particular model structure with the aim of selecting an adequate model. There had not been, or scarcely have been, any loss function that evaluates parsimony of model structures (bias contribution) based on the magnitude of parameter or coefficient. The magnitude of parameter could have a big role in choosing whether a term is significant enough to be included in a model and justifies ones' judgement in choosing or discarding a term/variable. This study intends to develop a new information criterion such that the bias contribution is related not only to the number of parameters, but mainly to the magnitude of the parameters. The parameter-magnitude based information criterion (PMIC2) is demonstrated in identification of linear discrete time model. The demonstration is tested using computational software on a number of simulated systems in the form of discrete-time linear regressive models of various lag orders and number of term/variables. It is shown that PMIC2 is able to select the correct the model based on all of the tested datasets.

Keywords: parameter magnitude, information criterion, system identification, discrete-time model, linear regressive model

Author Correspondence, e-mail: mdfahmi@utem.edu.my

doi: http://dx.doi.org/10.4314/jfas.v10i3s.30

1. INTRODUCTION

System identification can be considered a regression problem, where the relationship between input and output variables of a dynamical system has to be estimated. This task is typically accomplished by minimizing a certain information criterion, which measures how well the estimated relationship approximates the one which truly links the available input-output data pairs [1]. Its basic idea is to compare the time dependent responses of the actual system and identified model based on a performance function, hereby referred to as information criterion, giving a measure of how well the model response fits the system response [2].

An identification procedure typically consists in estimating the parameters of different models, and next selecting the optimal model complexity within that set. Increasing the model complexity will decrease the systematic errors, however, at the same time the model variability increases [3].A model accuracy and model parsimony known as variance and bias: $f(J)$=Var(J)+Bias(J) is an important consideration in selecting a model structure [1]. Hence, selecting a model with smallest variance is not a good idea because when the number of parameters increase, the variance will continue to decrease but will present a complex model. At a certain complexity, the additional parameters no longer reduce the systematic errors but are used to follow the actual noise realization on the data [3]. Often, in order to deal with the bias-variance trade-off, the information criterion is augmented with a penalty term intended to guide the search for the "optimal" relationship penalizing undesired regressors, where regressors refer to possible terms and variables. Regularized estimation has been widely applied in the context of system identification [4].

In this paper, the effectiveness of parameter magnitude-based information criterion (PMIC2) will be studied by testing on five simulated dynamic models in the form of discrete-time difference equations model. These models are linear autoregressive models with exogenous input (ARX).

The next sections are as follows: Section 2 introduces system identification; Section 3 explains about information criterions; Section 4 explains the simulated models; Section 5 provides results and discussion and lastly Section 6 concludes the paper along with recommendation of future works.

2. SYSTEM IDENTIFICATION

System identification is the field of building models from measured data. This may be thought of as the inverse of physical modelling in which models are built to simulate data. While physical modelling can enhance the general understanding of physiologic systems, system identification can provide tools for diagnosis and monitoring on an individualized basis [5]. The more accurate the mathematical model identified for a system, the more effective will be the controller designed for it [6].

Several strategies have been proposed to avoid over-parameterization while utilizing all the data for training the model [7]. The most popular strategy is to minimize a theoretically derived formula or criterion, which includes a goodness-of-fit index and a penalty factor for model complexity [5]. System identification can be framed as an optimization problem:

$$\hat{\theta} = \arg \underset{\theta}{min} J_F\ (\theta, D_N)$$

where J_F $(\theta,\ D_N)$ measures how well the model described by parameter θ describes the measured data. A widely used variation of the estimation criterion includes a so-called 'regularization term' in the loss function to be minimized, that is:

$$\hat{\theta} = \arg \underset{\theta}{min} J_F\ (\theta, D_N) + J_R\ (\theta, n)$$

In this case, θ is estimated by trading-off the data fitting term J_F $(\theta,\ D_N)$ and the regularization term J_R (θ, n) which act as a penalty to penalize certain parameter vectors θ which describe 'unlikely' systems [1].

In today's literature, various types of models are proposed for system modelling such as linear autoregressive with exogenous input (ARX) model and nonlinear autoregressive with exogenous input (NARX) model [1].

3. INFORMATION CRITERIONS

Model complexity selection is the sub-problem of model selection [8]. Parsimony, working hypotheses, and strength of evidence are three principles that regulate the ability to make inferences [9]. An information criterion can be designed to estimate an expected overall discrepancy, a quantity which reflects the degree of similarity between a fitted approximating model and the generating or 'true model' [10].

This so-called model order selection problem is considerably more difficult than parameter estimation problem. To solve the model order selection problem, one must first establish a set of candidate models (usually models with the same structure but different order) and then determine the best model in the set [5].

In [11], the typical behavior of the test error and training error are presented, as model complexity is varied. The training error tends to decrease whenever the model complexity increases, because the model fits the data harder. If the model is too complex, the test error is high, for the prediction model has large variance. In contrast, the model will underfit, and the prediction model has large bias. Therefore, it becomes an important issue in selecting the optimal complexity of the prediction model.

The PMIC2 is developed from the approach of using parameter magnitude information in information criterion [12]. It includes a bias term or known as penalty function and here will be denoted as PMIC2. It is written as follows:

$$PMIC2 = \sum_{n}(y(t) - \hat{y}(t))^2 + \sum_{j}\frac{1}{\theta_j}$$

where, θ_j is the magnitude of parameter in the model and j is the number of parameter.

4. SIMULATION SETUP

In this simulation, five ARX models are simulated using computer simulation software MATLAB. All models are denoted as Model 1, Model 2, Model 3, Model 4 and Model 5 and each model is further classified as having d.c. level and not having d.c. level. The difference between the two are one having the input and output average subtracted (hence has no d.c. level) and the other not subtracted. The following are the models written as linear regression models, its specifications, number of correct regressors and number of possible regressors:

Model 1:

$y(t) = 0.2y(t-2) + 0.5u(t-1) + 0.8u(t-3) + e(t)$

Specification: l=1, assumed maximum output order, n_y=3, assumed maximum input order, n_u=3

Number of correct regressor = 3 out of 7 (if d.c. level is assumed present) or 6 (if d.c. level is assumed absent)

Number of possible model = 127 (with d.c. level) or 63 (without d.c. level)

Model 2:

$y(t) = 0.1y(t-2) - 0.4y(t-4) + 0.5u(t-1) + 0.7u(t-3) + e(t)$

Specification: l=1, n_y=4, n_u=4

Number of correct regressor = 4 out of 9 (d.c. level present) or 8 (d.c. level absent)

Number of possible model = 511 (with d.c. level) or 255 (without d.c. level)

Model 3:

$y(t) = 0.2y(t-2) - 0.3y(t-4) + 0.6u(t-1) + 0.8u(t-5) + e(t)$

Specification: l=1, n_y=5, n_u=5

Number of correct regressor = 4 out of 11 (d.c. level present) or 10 (d.c. level absent)

Number of possible model = 2047 (with d.c. level) or 1023 (without d.c. level)

Model 4:

$y(t) = 0.1y(t-2) + 0.2y(t-5) - 0.3y(t-6) + 0.2u(t-2) + 0.3u(t-5) + e(t)$

Specification: l=1, n_y=6, n_u=6

Number of correct regressor = 5 out of 13 (d.c. level present) or 12 (d.c. level absent)

Number of possible model = 8191 (with d.c. level) or 4091 (without d.c. level)

Model 5:

$y(t) = 0.1y(t-2) + 0.2y(t-5) - 0.3y(t-7) + 0.2u(t-3) + 0.3u(t-5) + e(t)$

Specification: l=1, n_y=7, n_u=7

Number of correct regressor = 5 out of 15 (d.c. level present) or 14 (d.c. level absent)

Number of possible model = 32767 (with d.c. level) or 16383 (without d.c. level)

The input $u(t)$ is generated from a random uniform distribution in the interval [-1, 1] to represent white signal, while noise $e(t)$ is generated from a random uniform distribution [-0.01, 0.01] to represent white noise. Least squares was used as parameter estimation method. Five hundred data points are generated for each model. All models are penalized by PMIC2 in order to select the correct model.

5. RESULTS AND DISCUSSION

Based on the results, PMIC2 was able to select the correct models for all simulated models and either with or without d.c. level. Only the true regressors were selected and these regressors bear the same parameter values as the simulated models. To further illustrate such ability, Table 1 lists the possible regressors and the selected (true) regressor. Models with d.c. level are denoted as Model 1a, Model 2a, Model 3a, Model 4a and Model 5a while models without d.c. level are denoted as Model 1b, Model 2b, Model 3b, Model 4b and Model 5b.

Table 1. Selected Regressor and Parameter Values of PMIC2

	Model 1a	Model 1b	Model 2a	Model 2b	Model 3a	Model 3b	Model 4a	Model 4b	Model 5a	Model 5b
d.c. level										
$y(t-1)$										
$y(t-2)$	0.2	0.2	0.1	0.1	0.2	0.2				
$y(t-3)$										
$y(t-4)$			-0.4	-0.4	-0.3	-0.3				
$y(t-5)$							0.2	0.2	0.2	0.2
$y(t-6)$							-0.3	-0.3		
$y(t-7)$									-0.3	-0.3
$u(t-1)$	0.5	0.5	0.5	0.5	0.6	0.6				

$u(t-2)$							0.2	0.2		
$u(t-3)$	0.8	0.8	0.7	0.7					0.2	0.2
$u(t-4)$										
$u(t-5)$					0.8	0.8	0.3	0.3	0.32	0.32
$u(t-6)$										
$u(t-7)$										

Table 2 shows the value of PMIC2 for variance term, bias term and total of both terms for the selected models. Note that the minimum values of PMIC2 among the possible models become the selected model. It is shown that the bias value is bigger than variance value for all models. Variance and bias is inversely proportional; when variance value is small, the bias value is big and vice versa. Despite such difference, a fine balance between the two enables PMIC2 to select the model which is the same as the given model.

Table 2. The value of PMIC2

Models	Variance	Bias	Total
Model 1a	0.21830	8.20378	8.42209
Model 1b	0.2614	8.2530	8.5144
Model 2a	0.19912	15.95503	16.15416
Model 2b	0.2310	15.9671	16.1981
Model 3a	0.18292	11.22241	11.40533
Model 3b	0.3693	11.2401	11.6093
Model 4a	7.3267	15.7241	23.0508
Model 4b	6.8731	16.0015	22.8746
Model 5a	2.6445	16.0503	18.6947
Model 5b	2.4508	16.2546	18.7054

6. CONCLUSION

From the observation, the PMIC2 proved that it can perform well when selecting the correct model in linear discrete time model. The use of parameter magnitude as basis for bias measure places the criterion at a highly potential position among other information criterions. Further study will be concentrated on comparing it with other information criterions and on non-linear autoregressive models with exogenous input (NARX).

ACKNOWLEDGEMENT

The authors would like to acknowledge all the support from UniversitiTeknikal Malaysia Melaka and Ministry of Higher Education Malaysia for research grant FRGS/1/2015/TK03/FKM/02/F000271.

REFERENCES

[1] L. Ljung, System Identification: Theory for the User, 2nd ed., Prentice Hall, Upper Saddle River(1999).

[2] A. Alfi, and M. M. Fateh, Parameter identification based on a modified PSO applied to suspension system, Journal of Software Engineering & Applications, 3, (2010) 221-229.

[3] F. D. Riddef', R. Pintelon, J.Schoukens, and D. P. Gillikinb, Modified AIC and MDL model selection criteria for short data records, IEEE Transactions on Instrumentation and Measurement, 54 (2004)144-150.

[4] G. Prando, G. Pillonetto and A. Chiuso, "The role of rank penalties in linear system identification," IFAC-Papers online, 48(28) (2015) 1293-1300.

[5] X. Xinshu, L. Ying, and R. Mukkamala, A model order selection criterion with applications to cardio-respiratory-renal systems, IEEE Transactions on Biomedical Engineering, 52(3) (2005) 445-453.

[6] Z. Zibo and F. Naghdy, Application of genetic algorithms to system identification, IEEE International Conference, 2, (1995) 777-782.

[7] H. Akaike, A new look at the statistical model identification, IEEE Transactions on Automatic Control, 19(6)(1974).

[8] K.Kristinsson and G. A. Dumont, System identification and control using genetic algorithms, IEEE Transactions on Systems, Man, and Cybernetics, 22(5) (1992)1033-1046.

[9] L. Xiaoyong, F. Huajing and C. Zhaoxu, A novel cost function based on decomposing least-square support vector machine for Takagi-Sugeno fuzzy system identification, IET Control Theory & Applications, 8(5) (2013)338-347.

[10] A. K. Seghouane and S. I. Amari, The AIC Criterion and Symmetrizing the Kullback–Leibler Divergence, IEEE Transactions on Neural Networks, 18(1)(2007)97-106.

[11] Y. X. Geng and W. Wu, A Bayesian Information Criterion Based Approach for Model Complexity Selection in Speaker Identification, International Conference on Advanced Language Processing and Web Information Technology, 1 (2008)264-268.

[12] M.F. Abd Samad, H. Jamaluddin, R. Ahmad, M.S. Yaacob and A.K.M. Azad, Effect of penalty function parameter in objective function of system identification, International Journal of Automotive and Mechanical Engineering (IJAME), 7(2013)940-954.

How to cite this article:
Abd Samad M F, Nasir A R M. Performance of parameter-magnitude based information criterion in identification of linear discrete-time model. J. Fundam. Appl. Sci., 2018, *10(3S), 345-354.*