UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Faculty of Mechanical Engineering

PARAMETER MAGNITUDE-BASED INFORMATION CRITERION FOR OPTIMUM MODEL STRUCTURE SELECTION IN SYSTEM IDENTIFICATION

Abdul Rahman bin Mohd Nasir

Master of Science in Mechanical Engineering

2020

PARAMETER MAGNITUDE-BASED INFORMATION CRITERION FOR OPTIMUM MODEL STRUCTURE SELECTION IN SYSTEM IDENTIFICATION

ABDUL RAHMAN BIN MOHD NASIR

A thesis submitted
in fulfillment of the requirements for the degree of Master of Science
in Mechanical Engineering

Faculty of Mechanical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2020

DECLARATION

I declare that this thesis entitled "Parameter Magnitude-Based Information Criterion for Optimum Model Structure Selection in System Identification" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature	:	....
Name	:	....
Date	:	....

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Master of Science in Mechanical Engineering.

Signature	:……………...
Supervisor Name	:……………………… ....
Date	:…………………….. ....

DEDICATION

To my beloved mother and father

ABSTRACT

Model structure selection is among one of the steps in system identification and in order to carry out this, information criterion is developed. It plays an important role in determining an optimum model structure with the aim of selecting an adequate model to represent a real system. A good information criterion not only evaluates predictive accuracy but also the parsimony of model. There are many information criteria those are widely used such as Akaike information criterion (AIC) and Bayesian information criterion (BIC). On bias evaluation, these criteria only tackle on the number of parameters in a model. There scarcely have been any information criterion 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 one's judgement in choosing or discarding a term/variable. This study presents the comparison between parameter-magnitude based information criterion 2 (PMIC2), PMIC (an earlier version of its kind), AIC and BIC in selecting a correct model on simulated data and real data. For simulated data, PMIC2 was compared to AIC and BIC using enumerative approach and genetic algorithm. The test were made to a number of simulated systems in the form of discrete-time models of various linearity, lag orders and number of terms/variables. Then, PMIC2 was tested in selecting a good model to represent a real system based on gas furnace data and the results is compared to PMIC. The selected model was then tested using correlation test for model validation. Overall conclusion, it is shown that PMIC2 is able to select a more parsimonious model, yet adequately accurate, than AIC, BIC and PMIC.

ABSTRAK

Pemilihan struktur model adalah salah satu langkah dalam pengenalpastian sistem dan untuk melaksanakannya, kriteria maklumat perlu dibangunkan. Ia memainkan peranan penting dalam menentukan struktur model yang optimum dengan matlamat memilih model yang mencukupi untuk mewakili sistem sebenar. Kriteria maklumat yang baik bukan sahaja menilai ketepatan ramalan tetapi juga kekikiran model. Terdapat banyak kriteria maklumat yang digunakan secara meluas seperti kriteria maklumat Akaike (AIC) dan kriteria maklumat Bayesian (BIC). Untuk menilai struktur model, kriteria-kriteria ini hanya mengambil kira bilangan parameter dalam model sahaja dan hampir tiada kriteria maklumat yang menilai struktur model berdasarkan magnitud parameter atau pekali. Magnitud parameter mungkin mempunyai peranan yang besar dalam memilih sama ada sesuatu terma cukup penting untuk dimasukkan ke dalam model dan menyokong penilaian dalam memilih atau membuang sesuatu terma. Kajian ini membentangkan perbandingan antara kriteria maklumat berasaskan magnitud parameter 2 (PMIC2), PMIC (versi awal jenisnya), AIC dan BIC dalam memilih model yang betul pada data simulasi dan data sebenar. Untuk data simulasi, PMIC2 dibandingkan dengan AIC dan BIC menggunakan pendekatan perhitungan dan algoritma genetik. Ujian dibuat pada beberapa sistem simulasi dalam bentuk model masa-diskret dengan pelbagai kelelurusan, tertib susulan dan bilangan terma/pemboleh ubah. Kemudian, PMIC2 diuji dalam memilih model yang baik untuk mewakili sistem sebenar berdasarkan data relau gas dan hasilnya dibandingkan dengan PMIC. Model yang dipilih kemudian diuji menggunakan ujian korelasi untuk pengesahan model. Kesimpulan keseluruhan, telah terbukti bahawa PMIC2 dapat memilih model yang lebih ringkas, tetapi mencukupi dari segi ketepatan, berbanding AIC, BIC dan PMIC.

ACKNOWLEDGEMENTS

First and foremost, I would like to take this opportunity to express my sincere acknowledgement to my supervisor Dr. Md. Fahmi bin Abd Samad @ Mahmood from the Faculty of Mechanical Engineering Universiti Teknikal Malaysia Melaka (UTeM) for his essential supervision, support and encouragement towards the completion of this thesis.

I would also like to express my greatest gratitude to Dr. Mohd Khairi bin Mohamed Nor from Faculty of Mechanical Engineering, co-supervisor of this project for his cooperation and supports. Special thanks to Ministry of Higher Education for the financial support throughout this project.

Special thanks to all my peers, my beloved mother, father and siblings for their moral support. Lastly, thank you to everyone who had been to the crucial parts of realisation of this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AIC	-	Akaike information criterion
AR	-	AutoRegressive
ARMAX	-	AutoRegressive Moving Average with eXogenous input
ARX	-	AutoRegresive with eXogenous input
BF	-	Bayes factor
BIC	-	Bayesian information criterion
CA	-	Celular automata
CV	-	Cross validation
EC	-	Evolutionary computation
EI	-	Error index
EP	-	Evolutionary programming
ES	-	Evolution strategies
GA	-	Genetic algorithm
GP	-	Genetic programming
HC	-	Hill-climbing
MDL	-	Minimum description length information criterion
MEM	-	Maximum entropy method
MLE	-	Maximum-likelihood estimation
NAR	-	Nonlinear AutoRegressive
NARMAX	-	Nonlinear AutoRegressive Moving Average with eXogenous input
OF	-	Objective function
PDF	-	probability density function
PMIC	-	Parameter magnitude-based information criterion
PSO	-	Particle swarm optimisation

SGA	-	Simple genetic algorithm
SISO	-	Single-input-single-output

LIST OF SYMBOLS

a_j	-	absolute value of the parameter for term j
Bias(J)	-	Penalty term that penalise model parsimony
c	-	Constant criterion
d	-	Time delay
$e(t)$	-	Noise value at time
$F_*^{\ l}$	-	Nonlinear function
f(J)	-	Loss function
$J_F\ (\theta, D_N)$	-	Measure how well the model described by parameter
$J_R\ (\theta, \text{n})$	-	Penalty term that penalise model parsimony
j, k	-	The number of parameters in the statistical model
l	-	Degree of nonlinearity
L	-	Maximised value of the likelihood function for the estimated model
$lchrom$	-	Length of bit string chromosome
M	-	Model order
$maxgen$	-	Maximum number of generations
n	-	Number of observations
N	-	The number of data
n_u	-	Maximum orders of lag for input
n_y	-	Maximum orders of lag for output
pen	-	The value of penalty applied for having the variable/term
$penalty$	-	Fixed value termed penalty function parameter
$popsize$	-	Population size
r	-	The number of responses measured for each individual

		associated with the parameter
RSS	-	The residual sum of square
t	-	Time
$u(t)$	-	Input value at time
Var(J)	-	The variance for the estimated model
V_N	-	Criterion or cost function considering N data
$y(t)$	-	Output value at time
$\hat{y}(t)$	-	k-step-ahead predicted output
Z^N	-	Function of the parameter vector
ε	-	Residual or prediction error
$\emptyset$	-	Regressor vector
θ	-	Parameter vector
$\hat{\theta}$	-	Estimated parameter vector
$\delta(\tau)$	-	Kronecker delta

LIST OF PUBLICATIONS

Samad, M. F. A., and Nasir, A. R. M., 2018, Performance of Parameter-Magnitude Based Information Criterion in Identification of Linear Discrete-Time Model. *Journal of Fundamental and Applied Sciences*, 10(3S), pp. 345-354.

Samad, M. F. A., and Nasir, A. R. M., 2018, Implementation of Parameter Magnitude-Based Information Criterion in Identification of a Real System. *The Defence S&T Technical Bulletin*, 11(1), pp. 99-106.

Samad, M. F. A., and Nasir, A. R. M., 2017. Parameter Magnitude-Based Information Criterion in Identification of Discrete-Time Dynamic System. *Journal of Mechanical Engineering Vol. SI*, 4(1), pp. 119-128.

Samad, M. F. A., and Nasir, A. R. M., 2017. Parameter-Magnitude Based Information Criterion for Identification of Linear Discrete-Time Model. *Proceedings of Innovative Research and Industrial Dialogue'16*, pp. 129-130.

Samad, M. F. A., and Nasir, A. R. M., 2017, Discrete-Time System Identification Based on Novel Information Criterion Using Genetic Algorithm. *Journal of Fundamental and Applied Sciences*, 9(7S), pp. 584-599.

Samad, M. F. A., and Nasir, A. R. M., 2017, Comparison of Information Criterion on Identification of Discrete-Time Dynamic System. *Journal of Engineering and Applied Sciences*, 12(1SI), pp. 5660-5665.

CHAPTER 1

INTRODUCTION

1.1 Background

The field of system identification has received a lot of attention over the last two decades. It is now a fairly mature field, and many powerful methods are at the disposal of control engineers (Kristinsson and Dumont, 1992). System identification deals with the problem of building mathematical models of dynamic systems that based on observed data from the system (Ljung, 1999).

In system identification, data acquisition, model structure selection, parameter estimation and model validation are the main steps in approximating a system (Ljung, 1999). Model structure selection is the one of the stage in system identification and it refers to the determination of the variables and terms to be included in a model. Basically, having adequate predictive accuracy to the system response and yet parsimonious in structure is a criteria in describing an optimum model. A model preferred as a parsimonious model structure since it has less number of variables and/or terms, system analysis and control becomes easier (Samad, 2009).

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 loss function, which measures how well the estimated relationship approximates the one which truly links the available input-output data pairs (Prando et al., 2015).

1.2 Problem statement

Till today, there are not many loss functions that can be used in selecting model structure in system identification. Among the well-known ones, namely Akaike information criterion (AIC) and Bayesian information criterion (BIC), accounts for both variance contribution and bias contribution in the selection. However, these loss functions were developed for a general class of problems such that they require information of number of samples and number of parameters. These information are inadequate in making correct selection of model structure.

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 plays a big role in choosing whether a term is significant enough to be included in a model and justifies a judgement in choosing or discarding a term/variable.

When selecting a model structure, two considerations need to be evaluated. One is model accuracy and the other one is model parsimony (Söderström and Stoica, 1989, Ljung, 1999). Two components are common in a loss function - variance and bias: $f(J) = \mathrm{Var}(J) + \mathrm{Bias}(J)$, where $f(J)$ is the loss function, $\mathrm{Var}(J)$ is the variance which is the maximised value of the likelihood functions for the estimated model and $\mathrm{Bias}(J)$ is the penalty term that penalise the parsimony of the model. Although many well-known functions such as AIC and BIC are used widely, these loss functions do not account for individual term significance by the magnitude of parameter, which may cause the bias not to be well-defined towards selecting a parsimonious model, without much sacrifice to accuracy. On bias evaluation, these criteria only tackle on the number of parameters in a model. Having two models of different structures but same number of parameters will constitute the same bias evaluation. Unless the magnitudes of the parameters are considered, a more parsimonious model structure may be selected among the two.

1.3 Research objectives

The objectives that have been decided in this research are as follow:

1. To propose a loss function with a criterion such that it relates a models evaluation directly to the property of the model, especially to the magnitudes of parameters associated with individual terms.
2. To establish a potential for better balance between accuracy and parsimony, particularly in discrete-time difference model.
3. To compare analytically the model structure selected by the developed function to other loss function (information criterion) on simulated data and real data.

1.4 Research scopes

Due to wide development of study in the field of system identification, the scopes of this research are:

1. Discrete-time difference equation models are used in this study.

   Discrete time models (or time series model) become a practical choice because of the assumption that the output of a system is a realisation of the variables at instants of time. The typical data acquisition practice is also inline with the assumption. Difference equation model is the simplest interpretation of a system's process in the group of discrete-time models from a study of difference equation models, it shown that difference equation models are representative of many other types of models