System Identification using Kernel Regularization (November 2020)

Dr. Siva Kumar Tadepalli *Superevisor*, Srishti Shukla, *Student*

***Abstract*— Various methods have been proposed for a robust System identification in the recent years.Different classical Parametric predication methods along with some learning approach have come in this scenario.Kernel based Regularisation methods give us some connection between system identification and Machine Learning.In this paper we have summarised some of the resent approaches ,their pro and and cons for the above purpose**

***Index Terms*— linear system identification,non linear sytem identification,kernel ,robust identification,gaussian regression,Bayesian inference**

I. Introduction

Sytem Identification is the process of making mathematical models of systems by monitoring input-output data. For time-invariant linear dynamical systems the impulse reponse is a deconvolution problem or we can say inverse problem.Such problems find their applications in many fields [4] .Some experimental obseravtions have shown unsatisafactory results for prediction error methods (PEM) [15]. A robust identification approach for BIBO-stable linear and time-invariant system has been proposed [14].A Bayesian framework has been used to formulate a probabilistic prior on system's impulse response.The regularisation approach with proper regularisation matrix has shown better accuracy than PEM or maximum likelihood approach [3].For linear systems the earliest regularised approaches were [quote refernce].Atomic and nuclear norms have also been specified in [13]. The introduction of RKHS has been done [10] also conditions for stability has been derived.In this paper a brief description of all these approaches has been divided into parts. .

II. System Identification

System Output is the convolution of input and impulse response.So to find impulse response from previously observed input and output is a deconvolutional problem which is a non trivial task.A minimum error approach is needed so that future predictions should be more accurate . We have a training data(x_i,y_i) where x_i is the input and y_i is the output .The requirement is to design an estimator to predict yi for a given input xi. Mean Square error [9] which is given by

MSE = BIAS^2 + VARIANCE

To handle this trade off between bias and variance so as to minimise the error is the main task of the researches in this paradigm. Bias is the difference between the estimated and mean value of the elaboration of the system. A case in which both of the above coincides is unbiased. The flexibility of the model accounts for its variance .

III. Classical Statistical Methods

A classical PEM (prediction error methods) approach has been used for this purpose [8] A large collection of results on these methods is available quote re.For large number of samples satisfactory

23 November 2020. This work was supported by National Institute Of technology Uttarakhand."

Srishti Shukla is pursuing B.Tech in Electronics and Communication engineering from National Institute Of Technology Uttarakhand,srishti.ecebt18@nituk.ac.in.

Dr.Siva Kumar Tadepalli is Assisstant Professor,Department of Electronics and Communication engineering at National Institute Of technology Uttarakhand ,Srinagar Garhwal sktadepalli@nituk.ac.in .

.

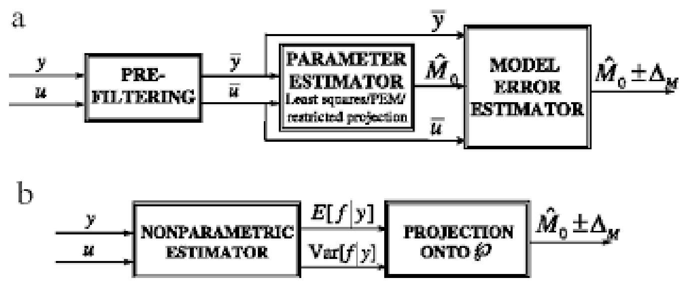

Fig. 1. (a)Shows the earlier approach (b)Depicits the new approach

results have been seen upto some extent .But for short and noisy data they have given in accurate results [2] .Also These models make it very complex for Multiple Input Multiple output system.quote equation if necessary

$$\theta = argmin_{\theta \in D_M} [\sum_{t=1}^{N} ||y(t) - y(t|\theta)||^2] \tag{1}$$

IV. Kernel Based Approaches

Recent researches have shown that a different approach to system identification can show better results [14] .In a way it can be said that some prior information is fed to the process of system identification which is assigning a covariance or Kernel in the language of machine learning. Kernel Based methods have been used for robust identification of BIBO systems .Robust identification can be done in three directions namely stochastic embedding, model-error modelling, set membership identification. All these models start with a low order model and then consider the bias and variance factors. The stochastic embedding approach models the bias error as the realization of a stochastic process [7]. The model-error modelling approach exploits residual analysis in order to characterize undermodeling, whereas set-membership identification determines the worst-case error associated with the nominal model [17] , [6] .

E[f|y] and Var[f|y]are posterior mean and autocovariance of the impulse response, respectively, which are fed to projection model which give the nominal model and its uncertainity.In this approach firstly all the available information is used and then the best possible estimate is given .Here the effect of experimental design and other factors is minimum.

V. FOR LINEAR AND NON LINEAR SYSTEMS

For the linear systems it has been shown that the selection of regularisation matrix plays a crucial role [3].It has also been shown that a lower order model is estimated by classical techniques PEM and for higher order model the regularisation approach works better.Transfer function estimation has also been carried out by using FIR model. The use of stable spline kernel in regularisation can be clearly seen [14], [2] as it models the impulse response as a zero mean gaussian process.

The use of Bayesian frame work has been done in some researches [14] they have implanted this issue in a completely Bayesian system. Specifically, another probabilistic earlier has been figured

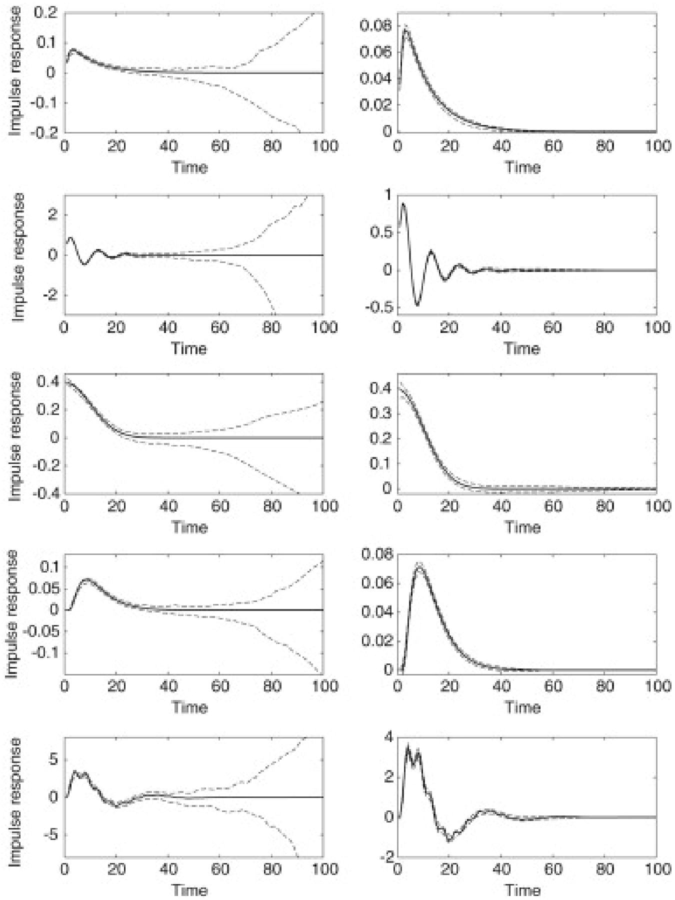

Fig. 2. True impulse response (solid line) and estimated impulse response(dashed line) using stable spline kernel ,here white noise has been given as input [2]

legitimately on the obscure drive reaction f , as opposed to on the predisposition mistake. This earlier, in some sense, is the most un-submitting one that consolidates data on both coherence of f and framework BIBO-security. Also diagonal correlated kernel with positive correlation and stable spline kernel belong to the class of exponentially convex and locally stable kernels [2].These outcomes have motivated to build a more generalised kernel structure to predict the properties of impulse response. Weiner systems are linear systems having a static non linearity,their identification has been into research [12].

For non linear system the use of gaussian regression has been done ,where the system is realised as a gaussian field [11].In non linear cases neural network approach has also been proposed by researchers [18]

VI. NUCLEAR AND ATOMIC NORMS

Recent researches have been in nuclear form for system identification .The nuclear norm (or trace norm) of a matrix is the sum of its singular values Atomic norms have been considered as regularizers for system identification in the past years . The function to be reconstructed is described as the sum of a (possibly infinite) number of basis functions which are called atoms.An advantage offered by this approach is that it gives the best convex penalty when the function is sum of few atoms.This approach has found applications in signal processing and machine vision for regard estimation of sparse vectors and low-rank matrices [1] Some researchers have shown performance of ReLS (regularides least squares)equipped with atomic, nuclear or kernel-based norms via numerical studies [13].The literarture also highlights that the use of inappropriate regularisers can result into large variance estimators obtained. The recent research [19]

VII. SOME OTHER KERNEL BASED METHODS

Kernel based methods basically include the gaussian regression [16], regularisation networks ,support vector machines and smoothing splines [20]

The regularisation network makes use of Regularised Kernel Hilbert Spaces(RKHS).

$$g = argmin_{f \in H}[\sum_{i=1}^{N}(y_i - f(x_1))^2/N + \gamma||f||^2] \qquad (2)$$

The appropriate use of RKHS can handle the balance between bias and variance [10].Recent research has introduced the RKHS of dynamic system also considering BIBO stability.This point of view builds up a strong connection

between framework recognizable proof and AI, with center

on the issue of gaining from models. Such structure has

prompted straightforward inductions of RKHSs BIBO dependability conditions. In

the nonlinear situation, remembering for the regularized assessor too

other soundness ideas, for example in the Sontag's sense (Sontag, 2008),

seems a significant future exploration bearing.

At last, RN union to the ideal indicator has been

demonstrated under suppositions customized to framework distinguishing proof,

additionally bringing up the connection among consistency and RKHS

soundness.

Some of the approaches [5] include the direct study of non linear case for Support Vector Machines(SVM).In large scale problems memory hinderance limits the usage of higher order optimization problems.Optimization via coordinate descent has came as a discussed approach in machine learning and statistics.

REFERENCES

[1] T. Chen, M. S. Andersen, L. Ljung, A. Chiuso, and G. Pillonetto. System identification via sparse multiple kernel-based regularization using sequential convex optimization techniques. *IEEE Transactions on Automatic Control*, 59(11):2933–2945, 2014.

[2] Tianshi Chen and Lennart Ljung. On kernel structures for regularized system identification (i): a machine learning perspective**this work has been supported by a research grant for junior researchers no. 621-2014-5894 and the linnaeus center cadics, both funded by the swedish research council, and the erc advanced grant learn, no. 267381, funded by the european research council.http://www.hamecmopsys.ens2m.fr. *IFAC-PapersOnLine*, 48(28):1035 – 1040, 2015. 17th IFAC Symposium on System Identification SYSID 2015.

[3] Tianshi Chen, Henrik Ohlsson, and Lennart Ljung. On the estimation of transfer functions, regularizations and gaussian processes—revisited. *Automatica*, 48(8):1525 – 1535, 2012.

[4] Giuseppe De Nicolao, Giovanni Sparacino, and Claudio Cobelli. Nonparametric input estimation in physiological systems: Problems, methods, and case studies. *Automatica*, 33(5):851 – 870, 1997.

[5] Francesco Dinuzzo. Analysis of fixed-point and coordinate descent algorithms for regularized kernel methods. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 22:1576–87, 08 2011.

[6] A. Garulli, A. Vicino, and G. Zappa. Conditional central algorithms for worst case set-membership identification and filtering. *IEEE Transactions on Automatic Control*, 45(1):14–23, Jan 2000.

[7] G. C. Goodwin, M. Gevers, and B. Ninness. Quantifying the error in estimated transfer functions with application to model order selection. *IEEE Transactions on Automatic Control*, 37(7):913–928, 1992.

[8] Lennart Ljung. Prediction error estimation methods. *Circuits, Systems, and Signal Processing*, 21:11–21, 01 2002.

[9] Lennart Ljung, Tianshi Chen, and Biqiang Mu. A shift in paradigm for system identification. *International Journal of Control*, 93(2):173–180, 2020.

[10] G. Pillonetto. System identification using kernel-based regularization: New insights on stability and consistency issues. *Autom.*, 93:321–332, 2018.

[11] G. Pillonetto, M. H. Quang, and A. Chiuso. A new kernel-based approach for nonlinearsystem identification. *IEEE Transactions on Automatic Control*, 56(12):2825–2840, 2011.

[12] Gianluigi Pillonetto. Consistent identification of wiener systems: A machine learning viewpoint. *Automatica*, 49(9):2704 – 2712, 2013.

[13] Gianluigi Pillonetto, Tianshi Chen, Alessandro Chiuso, Giuseppe De Nicolao, and Lennart Ljung. Regularized linear system identification using atomic, nuclear and kernel-based norms: The role of the stability constraint. *Automatica*, 69:137 – 149, 2016.

[14] Gianluigi Pillonetto and Giuseppe De Nicolao. A new kernel-based approach for linear system identification. *Automatica*, 46(1):81 – 93, 2010.

[15] Gianluigi Pillonetto, Francesco Dinuzzo, Tianshi Chen, Giuseppe De Nicolao, and Lennart Ljung. Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, 50(3):657 – 682, 2014.

[16] Carl Edward Rasmussen. Gaussian processes for machine learning. MIT Press, 2006.

[17] Wolfgang Reinelt, Andrea Garulli, and Lennart Ljung. Comparing different approaches to model error modeling in robust identification. *Automatica*, 38:787–803, 01 2001.

[18] Shun-Feng Su and F. . Y. P. Yang. On the dynamical modeling with neural fuzzy networks. *IEEE Transactions on Neural Networks*, 13(6):1548–1553, 2002.

[19] Rajiv Singh, Mario Sznaier, and Lennart Ljung. An atomic norm minimization framework for identification of parameter varying nonlinear arx models. *IFAC-PapersOnLine*, 52(28):1 – 6, 2019. 3rd IFAC Workshop on Linear Parameter Varying Systems LPVS 2019.

[20] G. Wahba. *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1990.