Presented as late breaking results contribution
21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Kernel based Parametric Modeling with Accurate Step Responses ⋆

Toshiharu Sugie * **Yuki Hori** **

* *Osaka University, Suita, Japan, (e-mail: sugie@jrl.eng.osaka-u.ac.jp)*
** *Kyoto University, Kyoto Japan, (email: hori.yuki.angramain@gmail.com)*

Abstract: This paper is concerned with kernel-based system identification. An approach to estimate a parametric model with given structure is shown, which approximates the step response of the target system. It is composed of two parts. First, taking account of the step response, an IIR model is estimated via kernel regularization with an appropriate input. Second, a parametric model with given structure is estimated from the obtained impulse response. A numerical example is given to demonstrate the effectiveness of the proposed approach.

Keywords: system identification, linear systems, discrete-time systems, kernel regularization, parametric models

1. INTRODUCTION

Since the pioneering work by Pillonetto and De Nicolao (2010), tremendous attention has been paid to the kernel based approach in system identification (see e.g., Pillonetto et al. (2018) and the references there in). The approach enables us to obtain fairly good nonparametric models (i.e., IIR models) from small size noisy I/O data. On the other hand, in many control system design/analysis, compact parametric models would be necessary. Though there are so many identification methods to estimate parametric models, one major drawback there is that the true system is assumed to belong to the given model class. This assumption may not be valid in many practical cases. Consequently, the obtained parametric model could behave quite differently from the true system. In fact, this is one important motivation why the kernel based approach focuses on nonparametric models.

One approach has been proposed to obtain a parametric model based on the estimate of impulse response by the Kernel regularization by J. Wågberg et al. (2018). Their method is robust against both small noisy data and over parametrization. This robustness is important. However, their numerical examples did not demonstrate enough this point (especially for the lower order models). Furthermore, though they focused on impulse response only, the goal of modeling depends on the designer. Sometimes, accuracy of the step response of the model is more important than that of the impulse response. Therefore, it is necessary to cope with this type of requirement.

Based on the above observations, this paper proposes a method to identify a parametric model with given structure whose step response approximates that of the target system well. Furthermore, a numerical example is given to demonstrate its effectiveness.

⋆ This paper is partly supported by JSPS KAKENHI No. JP7H03281.

2. PROBLEM DESCRIPTION

2.1 System Description

Consider the linear shift-invariant discrete-time SISO system described by

$$y(t) = G_\star(q)u(t) + e(t), \quad t = 1, \dots, N \tag{1}$$

where $y(t) \in \mathbb{R}$ is the observed output, $u(t) \in \mathbb{R}$ is the input, $e(t) \in \mathbb{R}$ is a zero mean white noise with variance λ, and $G_\star(q)$ is the transfer function with the shift operator q which can be described as

$$G_\star(q) = \sum_{k=0}^{\infty} g_\star(k) q^{-k} \tag{2}$$

where $g_\star(k)$ is the impulse response at time k. Suppose $u(t) = 0$ holds for any $t \le 0$, then I/O relation of the system is given by

$$y = Ug_\star + e \tag{3}$$

$$y = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}, \quad U = \begin{bmatrix} u(1) & 0 & \cdots & 0 \\ u(2) & u(1) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ u(N) & u(N-1) & \cdots & u(1) \end{bmatrix} \tag{4}$$

$$g_\star = \begin{bmatrix} g_\star(0) \\ g_\star(1) \\ \vdots \\ g_\star(N-1) \end{bmatrix} \tag{5}$$

where $e = [e(1)\ e(2)\ \cdots\ e(N)]^\top$ denotes the noise vector with $\mathrm{E}[e] = 0_N$ and $\mathrm{E}[ee^\top] = \lambda I_N$.

2.2 Model class

Suppose the degrees of the numerator and the denominator, n_b and n_f, are determined by the designer in advance, and we seek the parametric model of the form

Presented as late breaking results contribution
21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

$$G_\theta(q) = \frac{b_0 + b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}}{1 + f_1 q^{-1} + \cdots + f_{n_f} q^{-n_f}} \tag{6}$$

where $\theta = \begin{bmatrix} b_0 \ b_1 \ \cdots \ b_{n_b} \ f_1 \ f_2 \ \cdots \ f_{n_f} \end{bmatrix}^\top \in \mathbb{R}^{n_b+n_f}$ is the parameter vector to be found from the I/O data

$$\mathcal{D} = \{(u(1), y(1)), (u(2), y(2)), \cdots, (u(N), y(N))\}$$

of the target system.

For later use, denote the set of all parametric models by

$$\mathcal{G}_\theta := \left\{ G_\theta(q) \mid \theta \in \mathbb{R}^{n_b+n_f} \right\} \tag{7}$$

2.3 Problem Statement

Now we consider the criterion

$$J_\star = \|g_\star - g_\theta\|_W^2 := (g_\star - g_\theta)^\top W (g_\star - g_\theta) \tag{8}$$

where $g_\theta = [g_\theta(0) \cdots g_\theta(N-1)]^\top$ denotes the impulse response of the system whose transfer function is $G_\theta(q) \in \mathcal{G}_\theta$ and $W \in \mathbb{R}^{N\times N}$ is the given symmetric matrix. One example of W is $C^\top C$ with

$$C = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 1 & \cdots & \cdots & 1 \end{bmatrix}$$

In this case $J_\star$ means the squared Euclidean norm of the difference of step responses between the true plant $G_\star$ and its model G_θ. For simplicity, here we consider this case only.

As already pointed out by Fujimoto et al. (2017), the choice of input sequence plays a crucial role to minimize $J_\star$. Hence, we should find an appropriate input sequence $u := [u(1), u(2), \cdots, u(N)]^\top$ as well as the parametric model for this purpose. We assume that u should belong to the available input set given by

$$\mathcal{U} = \{u \mid u_\ell \le u(t) \le u_u \ \text{ for } t = 1, 2, \cdots, N\} \tag{9}$$

where $u_\ell \in \mathbb{R}$ and $u_u \in \mathbb{R}$ are the given lower and upper bounds of the available input, respectively, at each time step. Now the problem is stated as follows:

Problem 1 Given $\lambda, W, \mathcal{U}$ and $\mathcal{G}_\theta$, find $u \in \mathcal{U}$ and $G_\theta \in \mathcal{G}_\theta$ such that $J_\star$ is minimized.

Namely, we are going to find both the input sequence and the parametric model so that the resultant step response is as close as possible to the true one.

Note that true system transfer function may not belong to the given system class $\mathcal{G}_\theta$. More importantly, the response of the obtained model is expected to mimic that of the true system even if the system order is different.

3. PROPOSED METHOD

3.1 Estimate of nonparmetric $\bar{g}$ with optimal input

According to Fujimoto and Sugie(2017) we first find $u^* \in \mathcal{U}$ and the nonparametric model $\bar{g} := [\bar{g}(0) \cdots \bar{g}(N-1)]^\top \in \mathbb{R}^N$.

In the kernel regularized approach, we regard that $p(g_\star) \sim \mathcal{N}(0, K)$ holds, where $p(\cdot)$ is the probability density function and $\mathcal{N}(\cdot, \cdot\cdot)$ denotes the normal distribution with mean $(\cdot)$ and covariance matrix $(\cdot\cdot) \in \mathbb{R}^{N\times N}$. K is called the kernel matrix and it may be determined from a preliminary experiment. Then, the posterior distribution of $g_\star$ becomes

$$p(g_\star \mid y) \sim \mathcal{N}(g, \hat{K})$$

where

$$\hat{K} = \left(\lambda^{-1}(U)^\top U + K^{-1}\right)^{-1}, \ g = \lambda^{-1}\hat{K}U^\top y.$$

The above g is the estimate of $g_\star$ based on $\mathcal{D}$.

According to Fujimoto and Sugie (2017), the input u which minimizes

$$\|g_\star - g\|_W^2 p(g_\star|y)p(y),$$

is equal to u^* which minimizes

$$J(u) = \mathrm{Tr}\left(W\left(K^{-1} + \frac{1}{\lambda}U^\top U\right)^{-1}\right). \tag{10}$$

By using a projected gradient mehod, we can get a local minimal point $u^* = [u^*(1) \cdots u^*(N)]^\top \in \mathbb{R}^N$. Then, the corresponding estimate $\bar{g}$ is given by

$$\bar{g} = \lambda^{-1}\left(\lambda^{-1}(U^*)^\top U^* + K^{-1}\right)^{-1}(U^*)^\top y^* \tag{11}$$

$$y^* = U^* g_\star + e \tag{12}$$

$$U^* = \begin{bmatrix} u^*(1) & 0 & \cdots & 0 \\ u^*(2) & u^*(1) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ u^*(N) & u^*(N-1) & \cdots & u^*(1) \end{bmatrix} \tag{13}$$

3.2 Estimate of parametric $G_\theta(q)$

Once the input u^* and its corresponding output y^* are obtained, we estimate $G_\theta(q)$ based on these I/O data. In stead of minimizing

$$J_\star(\theta) = \|g_\theta - g_\star\|_W^2, \tag{14}$$

we try to minimize its expectation by regarding $g_\star$ as a random variable same as in the previous subsection. As a result, we determine θ by

$$\hat{\theta} = \arg\min \|g_\theta - \bar{g}\|_W^2 \tag{15}$$

4. SIMULATION

This section shows the effectiveness of the proposed method through a numerical example.

Suppose that true system is given by the transfer function ($n_f = 3$,$n_b = 2$) below.

$$G_\star(q) = \frac{3.871q^2 - 4.968q + 2.549}{q^3 - 1.507q^2 + 0.3258q + 0.2506} \tag{16}$$

The data length N and the input amplitude constraint (u_ℓ, u_u) are given by

$$N = 80 \tag{17}$$

$$u_l = -3, u_u = 3 \tag{18}$$

As a preliminary experiment, we inject a white binary input sequence $u_w \in \mathbb{R}^N$ and observe the corresponding output $y_w \in \mathbb{R}^N$. Choosing a Tuned Correlated Kernel of the form

$$K_{i,j} = \alpha\beta^{\max(i,j)}, \quad \alpha > 0, \ 0 < \beta < 1$$

we determine these hyper parameters (α, β) by empirical Bayes based on the I/O data $\mathcal{D}_w := \{u_w, y_w\}$.

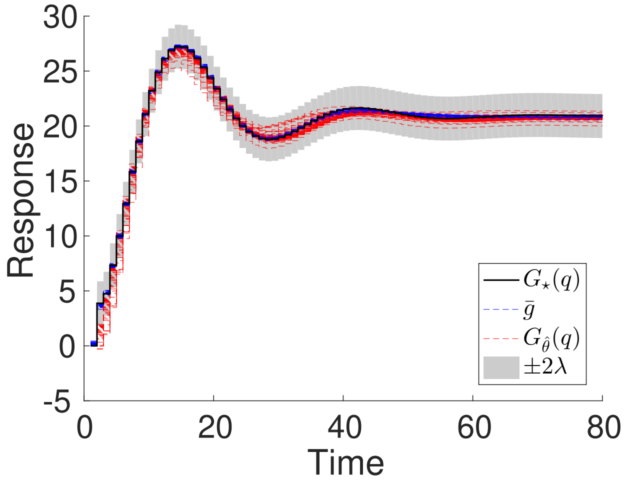

Fig. 1. Step responses in case of $n_f = 2$, (Proposed Method)

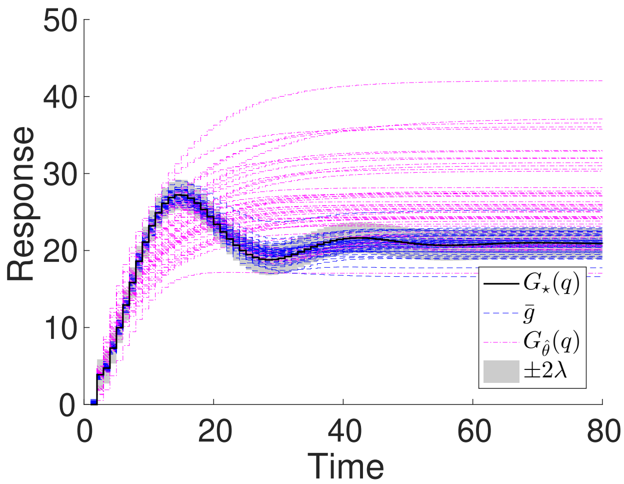

Fig. 2. Step responses in case of $n_f = 2$, (Existing Method)

Then calculate the optimal input u^* and the corresponding parameter θ by exploiting the constrained Particle Swarm Optimization (see Maruta et al. (2013)). We ran the simulation 50 times with different noises of variance $\lambda = 1$.

First, we set $n_f = 2$ (with $n_b = n_f - 1$), which is lower than the true degree. Fig. 1 shows 50 step responses of the estimated model by the proposed method. They are shown by the red broken lines. The true response is shown by the black line, and the step responses estimated by $\bar{g}$ are shown by the blue broken lines. The gray shaded area shows the noise amplitude. For comparison purpose, the results based on $\mathcal{D}_w$ with the method by J. Wågberg et al. (2018) are given by Fig. 2 . These two figures clearly demonstrate the effectiveness of the proposed method. Even in the case of lower order model, the proposed method provides us a fairly good model from the viewpoint of step responses.

Second, we set $n_f = 4$ (with $n_b = n_f - 1$), which is higher than the true degree. The results corresponding to the proposed method is shown in Fig. 3, and the those corresponding to the existing one is shown in Fig. 4. In this case, the existing method yields better results

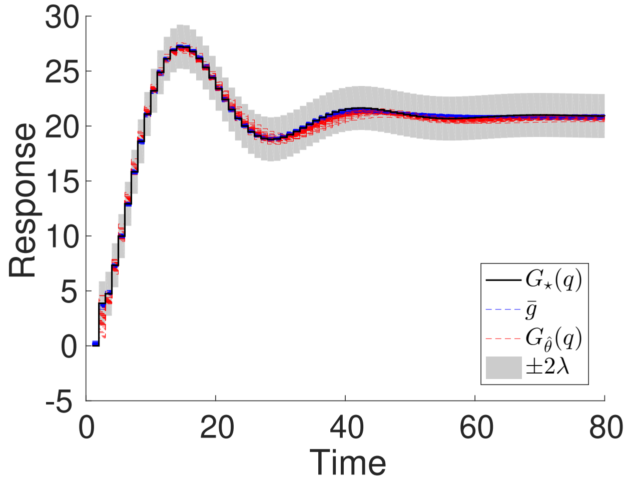

Fig. 3. Step responses in case of $n_f = 4$ (Proposed Method)

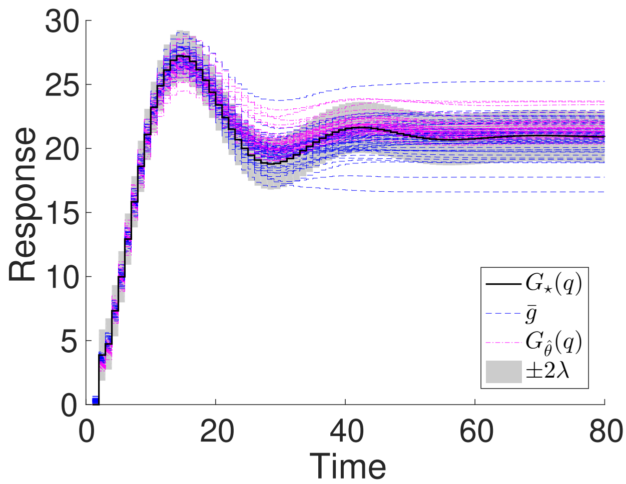

Fig. 4. Step responses in case of $n_f = 4$ (Existing Method)

compared to the lower degree case. However, the variance of step responses is still big. On the contrary, the proposed method exhibits much better results.

These results demonstrate that the proposed method provides fairly accurate step responses even if the system order is different from the true one. This robustness seems to be crucial in practice when we estimate the parametric models.

5. CONCLUSION

This paper gives an approach for identification of parametric model, which has two important properties. One is that the obtained model is robust against the system order mismatch, and the other is that the method reflects the modeling purpose (such as accuracy of the step response) more directly. For the first property, the kernel regularized method is utilized, and for the second property, the input is chosen according to the purpose.

REFERENCES

Pillonetto, G., and De Nicolao, G. (2010). A new kernel-based approach for linear system identification, *Auto-*

matica, **46**-1, 81/93

Pillonetto,G., Dinuzzo, F., Chen,T., De Nicolao,G., and Ljung, L. (2014). Kernel methods in system identification, machine learning and function estimation: A survey, *Automatica*, **50**-3, 657/682

Wågberg, J., Zachariah, D. and Schön, T.B. (2018). Regularized parametric system identification: a decision-theoretic formulation, *Proc. of Annual American Control Conference*, 1895/1900

Fujimoto, Y., and Sugie, T. (2017). Goal-Oriented Input Design for Kernel-Based System Identification, *Proc. of IEEE 56th Annual Conference on Decision and Control*, 3398/3402

Maruta, I., Kim, T.-H., Song, D., and Sugie, T. (2013). Synthesis of fixed-structure robust controllers using a constrained particle swarm optimizer with cyclic neighborhood topology, *Expert Systems with Application*, **40**-9, 3595/3605