Consistent estimation of autonomous linear time-invariant systems from multiple experiments

I. Markovsky [1], **R. Pintelon** [1]
[1] Vrije Universiteit Brussel, Department ELEC
Pleinlaan 2, Building K, B-1050 Brussels, Belgium
e-mail: **Ivan.Markovsky@vub.ac.be**

Abstract

Operational modal analysis from impulse response data can alternatively be viewed as an identification of a stable autonomous linear time-invariant system. For example, earthquake response data of civil engineering structures and impulsive excitation of bridges leads to this problem. Identification from a single experiment, however, does not yield a consistent estimator in the output error setting due to the exponential decay of the noise-free signal. Using data from multiple experiments, on the other hand, is not straightforward because of the need to match the initial conditions in the repeated experiments. Consequently, we consider the identification from arbitrary initial conditions and show that consistent estimation is possible in this case. The computational method proposed in the paper is based on analytic elimination of the initial conditions (nuisance parameter) and local optimization over the remaining (model) parameters. It is implemented in a ready to use software package, available from `http://slra.github.io/software.html`

1 Problem formulation

The class of scalar autonomous linear time-invariant systems of order less than or equal to $\mathtt{n}$ is denoted by $\mathscr{L}_{0,\mathtt{n}}$. A system $\mathscr{B} \in \mathscr{L}_{0,\mathtt{n}}$ is a set of trajectories [4]. The statement $y \in \mathscr{B}$ is a short-hand notation for "y is a trajectory of $\mathscr{B}$".

Consider N trajectories

$$\bar{y}^i = \big(\bar{y}^i(1), \dots, \bar{y}^i(T_i)\big), \qquad i = 1, \dots, N,$$

with possibly different lengths $T_1, \dots, T_N$ of a system $\overline{\mathscr{B}} \in \mathscr{L}_{0,\mathtt{n}}$. The data

$$\mathscr{D} := \{\, y^1, \dots, y^N \,\}$$

for the identification problem considered in the paper is generated in the output error setup:

$$y^i = \bar{y}^i + \tilde{y}^i, \qquad \text{where} \quad \bar{y}^i \in \overline{\mathscr{B}} \in \mathscr{L}_{0,\mathtt{n}} \text{ and } \tilde{y}^i \text{ is a zero mean white Gaussian process with covariance matrix } \sigma^2 \text{ and } \tilde{y}^i \text{ is independent of } \tilde{y}^j \text{ for all } i \neq j. \tag{1}$$

Here $\bar{y}^i$ is the "true value" of the trajectory y^i and $\overline{\mathscr{B}}$ is refered to as the "true system". In addition, we assumed that

$$0 < c_1 \leq \|\bar{y}^i\|_2^2 \leq c_2 < \infty. \tag{2}$$

Our aim is to estimate the true system $\overline{\mathscr{B}}$ from the data $\mathscr{D}$ and the prior knowledge that the true system belongs to the model class $\mathscr{L}_{0,\mathtt{n}}$.

Problem 1. (Maximum likelihood identification from multiple trajectories) Given a set of trajectories $\mathscr{D}$ and a model class $\mathscr{L}_{0,\mathtt{n}}$, specified by a natural number $\mathtt{n}$, find a maximum likelihood estimate $\widehat{\mathscr{B}}$ of the true data generating system $\overline{\mathscr{B}}$.

The log likelihood function for the data generating model (1) is

$$L(\widehat{\mathscr{B}},\widehat{\mathscr{D}}) = \begin{cases} \text{const} - \frac{1}{2\sigma^2}\sum_{i=1}^{N} \|y^i - \widehat{y}^i\|_2^2 & \text{if } \widehat{\mathscr{D}} \subset \widehat{\mathscr{B}} \\ -\infty & \text{otherwise.} \end{cases}$$

The maximum likelihood principle leads to the following optimization problem:

$$\begin{aligned} &\text{minimize} \quad \text{over } \widehat{\mathscr{D}} \text{ and } \widehat{\mathscr{B}} \quad \frac{1}{N}\sum_{i=1}^{N} \|y^i - \widehat{y}^i\|_2^2 \\ &\text{subject to} \quad \widehat{\mathscr{D}} \subset \widehat{\mathscr{B}} \in \mathscr{L}_{0,\mathtt{n}}. \end{aligned} \tag{3}$$

2 Solution method

A scalar autonomous linear time-invariant system with simple poles can be represented by the "sum-of-exponentials model"

$$\mathscr{B} = \{\, y = \sum_{j=1}^{\mathtt{n}} c_j \exp_{z_j} \mid c \in \mathbb{C}^{\mathtt{n}} \,\}. \tag{4}$$

Here $\exp_z$ is the exponential function $\exp_z(t) := z^t$. The complex numbers $z_1,\dots,z_{\mathtt{n}}$ are the *poles* of the system. In the representation (4), they are assumed to be distinct, *i.e.*, $z_i \neq z_j$, for all $i \neq j$.

A finite trajectory $y = \big(y(1),\dots,y(T)\big)$ of a sum-of-exponentials model (4) can be expressed as

$$y = P_T(\theta)c,$$

where $P_T(\theta)$ is the (extended) Vandermonde matrix

$$P_T(\theta) := \begin{bmatrix} 1 & \cdots & 1 \\ z_1 & \cdots & z_{\mathtt{n}} \\ \vdots & & \vdots \\ z_1^{T-1} & \cdots & z_{\mathtt{n}}^{T-1} \end{bmatrix}$$

and $z_1,\dots,z_{\mathtt{n}}$ are the roots of the polynomial

$$\theta(z) := \theta_1 + \theta_2 z + \cdots + \theta_{\mathtt{n}} z^{\mathtt{n}-1} + z^{\mathtt{n}}.$$

The maximum-likelihood identification problem (3) specialized for the sum-of-exponentials model becomes

$$\begin{aligned} &\text{minimize} \quad \text{over } \widehat{\mathscr{D}},\, \theta,\, c^1,\dots,c^N \quad \frac{1}{N}\sum_{i=1}^{N} \|y^i - \widehat{y}^i\|_2^2 \\ &\text{subject to} \quad \widehat{y}^i = P_{T_i}(\theta)c^i, \quad \text{for } i = 1,\dots,N \end{aligned}$$

or

$$\text{minimize} \quad \text{over } \theta,\, c^1,\dots,c^N \quad \frac{1}{N}\sum_{i=1}^{N} \|y^i - P_{T_i}(\theta)c^i\|_2^2 \tag{5}$$

Applying the variable projections method to (5), leads to N decoupled problems

$$\text{minimize} \quad \text{over } c^i \quad \frac{1}{N}\|y^i - P_{T_i}(\theta)c^i\|_2^2, \qquad \text{for } i = 1,\ldots,N. \tag{6}$$

These are ordinary least squares problems with solutions

$$\hat{y}^i = \underbrace{P_{T_i}(\theta)\big(P_{T_i}^\top(\theta)P_{T_i}(\theta)\big)^{-1}P_{T_i}}_{\Pi_{T_i}(\theta)} y^i,$$

where $\Pi_{T_i}(\theta)$ is an idempotent matrix ($\Pi_{T_i}^2(\theta) = \Pi_{T_i}(\theta)$) Therefore, the cost function of the sum-of-exponentials model is

$$f(\theta) = \frac{1}{N}\sum_{i=1}^{N}(y^i)^\top\big(I - \Pi_{T_i}(\theta)\big)y^i.$$

Proposition 2 (Consistency)**.** *Assuming that the data $\mathscr{D}$ is generated in the output error setup (1), the estimator defined by (5) is strongly consistent,* i.e.,

$$\widehat{\theta} \to \overline{\theta} \textit{ with probability 1 as } N \to \infty.$$

3 Numerical examples

In this section, we illustrate the consistency property of the estimator on a simulation example. The data generating system is a continuous-time linear time-invariant system of order $\mathtt{n} = 6$ with with resonance angular frequencies

$$\bar{\omega}_1 = 2\pi\,80\,\text{rad/s},\ \bar{\omega}_2 = 2\pi\,130\,\text{rad/s},\ \bar{\omega}_3 = 2\pi\,200\,\text{rad/s}$$

and poles' damping ratios

$$\zeta_1 = 0.15, \quad \zeta_2 = 0.1, \quad \zeta_3 = 0.2.$$

The system is sampled with a period $t_s = 10^{-3}$. A specific response $\bar{y}$ is shown in Figure 1.

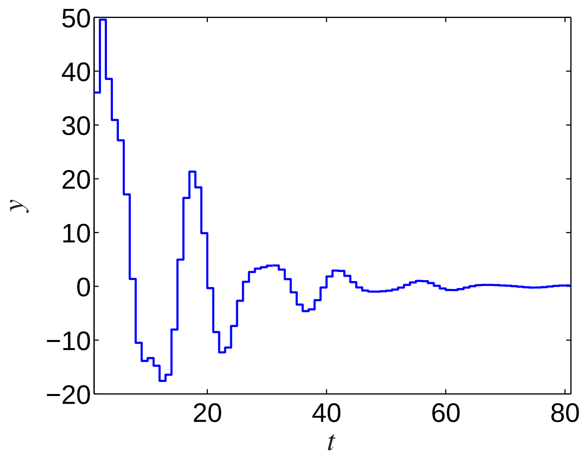

Figure 1: Specific trajectory of $\overline{\mathscr{B}}$.

The identification data $\mathscr{D}$ is generated via the output error model (1) with random true trajectories $\bar{y}^1,\ldots,\bar{y}^N$. The lengths $T_1,\ldots,T_N$ of the responses are determined, so that the trajectories have sufficient decay.

The number of experiments N varies from 1 to 50 and the signal to noise ratio is 100. For each value of N, the identification experiment is repeated $K = 200$ times and the average parameter error

$$e = \sqrt{\frac{1}{K}\sum_{k=1}^{K} \|\overline{\theta} - \widehat{\theta}^k\|_2^2}.$$

is computed. The results obtained (see Figure 2) show the convergence of the average parameter estimation error. The convergence rate is close to $1/\sqrt{N}$.

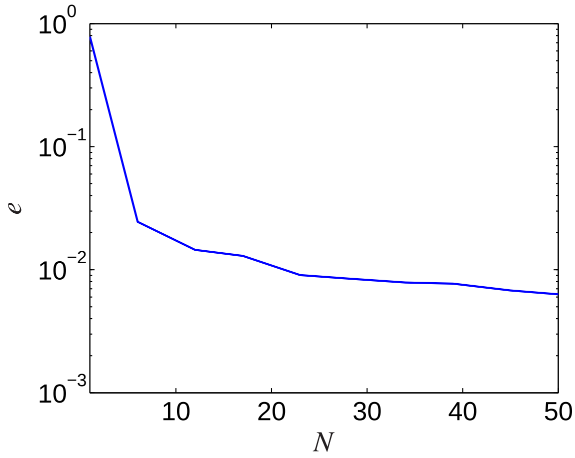

Figure 2: Error convergence.

Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement number 258581 "Structured low-rank approximation: Theory, algorithms, and applications", the Research Foundation Flanders (FWO-Vlaanderen), the Flemish Government (Methusalem Fund, METH1), and the Belgian Federal Government (Interuniversity Attraction Poles programme VII, Dynamical Systems, Control, and Optimization).

References

[1] I. Markovsky, *A software package for system identification in the behavioral setting*, Control Engineering Practice **21** (2013), 1422–1436.

[2] I. Markovsky, J. Goos, K. Usevich, and R. Pintelon, *Realization and identification of autonomous linear periodically time-varying systems*, Automatica (2014).

[3] I. Markovsky and K. Usevich, *Software for weighted structured low-rank approximation*, J. Comput. Appl. Math. **256** (2014), 278–292.

[4] J. Polderman and J. C. Willems, *Introduction to mathematical systems theory*, Springer-Verlag, 1998.

[5] J. Schoukens, G. Vandersteen, Y. Rolain, and R. Pintelon, *Frequency response function measurements using concatenated subrecords with arbitrary length*, IEEE Trans. on Instr. and Measurement **61** (2012), no. 10, 2682–2688.