Mini-projects for the SOCN course
"Behavioral approach to systems theory"

Ivan Markovsky

"The challenge is finding a balance between exploring new ideas and acquiring the skills to realize them."

Ken Robinson

As we've seen in the exercises, in case of exact data the data-driven methods solve the problems exactly, *i.e.*, they give equivalent results to model-based methods that use the data-generating system $\mathscr{B}$ instead of the data w_{d}. The equivalence between data-driven and model-based method no longer holds true when w_{d} is inexact. Indeed, the true data-generating system can not be recovered from inexact data. Moreover, different identification methods deliver different models. The fundamental difficulty in dealing with inexact data is that the problems are nonconvex. All currently known methods can be viewed as heuristics for solving the underlying nonconvex optimization problem.

The mini-projects in Section 1 explore the performance of the methods in case of inexact data w_{d}. We use different types of prior knowledge and observe its effect on the accuracy. The theoretically best performance is the one of methods using the true data-generating system (or equivalently exact data). Of particular interest is a comparison of the direct and indirect approaches. For the indirect approach, you can use methods implemented in the System Identification Toolbox of Matlab in order to obtain the model from the data. Section 2 has two projects related to practical applications. One is from the field of metrology and the other is about free fall in a gravitational field. Section 3 lists three current research projects of mine: recursive algorithm for exact identification, data-driven frequency response estimation, and computing distance measures between systems. You can also propose to work on a mini-project of your own choice (for example, a subproblem appearing in your own research topic).

1 Further experiments with the direct data-driven method

Noisy data w_{d}

consider simulated data in the classical errors-in-variables, output error, and ARMAX setups

what are the maximum-liklihood estimators in these cases?

how does the naive direct data-driven method compare with them?

what modifications of the naive method can improve the performance?

High-order linear time-invariant system

consider simulated data from a high-order linear time-invariant system

define an optimal low-order approximation, how to solve the resulting approximation problem? are there connections to the case of noisy data?

how does the naive direct data-driven method compare with the optimal approximation method?

what modifications of the naive method can improve the performance?

Unstable / nonminimum-phase systems

consider a system for which some input/outupt partitionings lead to a BIBO stable / minimum-phase representation while others to unstable / nonminimum-phase ones, how to exploit this fact?

how to simulate a bounded trajectory of a system with unstable input/outupt partitioning?

what is the meaning of the roots of $R(z)$?

how does the unstable / nonminimum-phase dynamics affect the numerical computations in problems such as simulation, Kalman filtering/smoothing, model reduction, and system identification?

Stiff systems

stiff systems exibit behavior at very different time-scales (fast and slow dynamics), they are notoriously difficult to simulate

how do the direct data-driven methods work in case of stiff systems?

if there are problems, where are they and how to deal with them?

Real-data from DAISY

how do direct data-driven methods compare with model-based ones on these real-data sets?

in my experience direct data-driven methods uniformly outperform model-based methods, why is this?

2 Applications

Sensor speed-up

Free fall in a gravitational field

3 Research projects

Denoising methods

structured low-rank approximation methods

nuclear norm relaxations

low-rank approximation (SVD truncation-type) methods

shrinkage methods

Uncertainty quantification of the denoising methods

statistical approach (consistency, confidence bounds, ...)

numerical linear algebra approach (perturbation analysis)

Improving the computational efficiency of data-driven methods

comparison of the computational efficinecy of model-based and direct data-driven methods

is a model needed for doing fast computations?

using fast methods for structured matrix computations in direct data-driven methods

Using Hankel vs page matrix in data-driven methods

Using a trajectory matrix with $\ell+1$ block-rows leads to an unstructured data matrix, so that in this case approximation by truncation of the singular value decomposition yields optimal approximation. This observation allows us to avoid the nonconvex optimization of the structured low-rank approximation problem, however, data from multiple short experiments is less informative than data from one long experiment with the same total number of samples. Empirical results show that overall low-rank approximation of the Hankel matrix gives more accurate model parameter estimates than low-rank approximation of the trajectory matrix. This contradicts empirical results reported in [2, 1] on using the trajectory matrix in data-driven MPC control so that a further research in this direction is needed.

Distance between systems

definition and computation of a distance between LTI systems

connection to the gap metric

applications (model-reduction, distance to uncontrollability, ...)

Data-driven state construction

state = intersection of past and future

Hankel matrix as a non-minimal state sequence

how to obtain a minimal state sequence? how to obtain a balanced state sequence?

Recursive algorithm for exact identification

In [6] we did not delve into numerical computations issues related to the implementation of Algorithm 1. This topic has connections with work on numerical linear algebra methods for Hankel structured matrices, see for example [3, 4, 5]. In particular incorporating approximation in Algorithm 1 is a interesting topic for further work.

Spurious annihilators

Algorithm 1 in [6] detects the annihilators of degree up to $\ell_{\max} := L_{\max} - 1$. Assuming that the data is generated by a linear time-invariant system $\bar{\mathscr{B}} = \ker R(\sigma) \in \partial\mathscr{L}^{q,n}_{m,\ell}$, the identified model $\widehat{\mathscr{B}} := \mathscr{B}_{\text{MPUM}}(w_{\text{d}})$ contains the annihilators $R^1, \dots, R^g$ of $\bar{\mathscr{B}}$ of degree up to $\ell_{\max}$. Depending on the data w_{d}, however, $\widehat{\mathscr{B}}$ may contain additional annihilators of degree up to $\ell_{\max}$ that are not annihilators of $\bar{\mathscr{B}}$. We call these annihilators *spurious*.

Definition 1 (Spurious annihilators)**.** Let the data w_{d} be generated by a system $\bar{\mathscr{B}} \in \partial\mathscr{L}^{q}_{(m,\ell,n)}$, $\bar{\mathscr{R}}$ be the set of annihilators of $\bar{\mathscr{B}}$, and $\widehat{\mathscr{R}}$ be the set of annihilators of $\mathscr{B}_{\text{MPUM}}(w_{\text{d}})$. The elements of the set difference $\widehat{\mathscr{R}} \setminus \bar{\mathscr{R}}$ are spurious annihilators of $\mathscr{B}_{\text{MPUM}}(w_{\text{d}})$ (with respect to $\bar{\mathscr{B}}$).

Classical identifiability results such, as the fundamental lemma, distinguish spurious annihilators from the true system's annihilators by degree separation. This mini-project investigates identifiability conditions that filter spurious annihilators based on other types of prior knowledge about the true system $\bar{\mathscr{B}}$.

Existence of spurious annihilators prevents identifiability. Indeed, by definition, $\mathscr{B}_{\text{MPUM}}(w_{\text{d}}) = \bar{\mathscr{B}}$ if and only if $\mathscr{B}_{\text{MPUM}}(w_{\text{d}})$ has no spurious annihilators. Consequently, all identifiability results give conditions ensuring that the model $\widehat{\mathscr{B}}$ does not include spurious annihilators. A prototypical example of a spurious annihilator is an annihilator for a set of input variables. In the fundamental lemma such annihilators are avoided by the persistency of excitation assumption for the inputs. Indeed, persistency of excitation of u_{d} of order L implies that there are no annihilators of order $L-1$. The fundamental lemma uses a separation of the annihilators into spurious and non-spurious based on degree: any annihilator of degree larger than an a priori known threshold degree $\ell_{\max}$ is spurious.

Robustifying the fundamental lemma

For the persistency of excitation condition to hold true 1) $a_i \neq 0$, for all i and 2) $\lambda_{u,i} \neq \lambda_{u,j}$, for all $i \neq j$. A way of robustifying these conditions is 1) $a_i > \varepsilon$, for some user defined tolerance ε, and 2) choose the $\lambda_{u,i}$'s "well spread".

Input design using the fundamental lemma

Using the input model representation of a persistently exciting signal the freedom of choosing an input that satisfies the conditions of the fundamental lemma is equivalent to choosing the input model $\mathscr{B}_u$ and the initial conditions $x_{u,\mathrm{ini}}$. The input model representation of the class of sufficiently exciting inputs can be used then for input design under user defined specifications, such as frequency band, maximum/minimum value bounds, *etc*.

Data-driven representation of continuous-time systems

what is the equivalent of the generalized persistency of excitation in continuous-time

Using Chebfun in continuous-time system theory

an algorithm for computing $\int_0^{T-L} w(t,\tau)g(\tau)d\tau$

Comparison of methods for input estimation

comparison of the direct data-driven method with the method of Gakis and Smith

References

[1] A. Agarwal et al. "Model agnostic time series analysis via matrix estimation". In: *Proc. ACM on Measurement and Analysis of Computing Systems* 2.3 (2018), pp. 1–39.

[2] J. Coulson, J. Lygeros, and F. Dörfler. "Distributionally robust chance constrained data-enabled predictive control". In: *IEEE Trans. Automat. Contr.* 67 (2022), pp. 3289–3304.

[3] G. Heinig and P. Jankowski. "Kernel structure of block Hankel and Toeplitz matrices and partial realization". In: *Linear Algebra Appl.* 175 (1992), pp. 1–30.

[4] G. Heinig and K. Rost. *Algebraic methods for Toeplitz-like matricies and operators*. Birkhauser, 1984.

[5] T. Laudadio, N. Mastronardi, and P. Van Dooren. "The Generalized Schur Algorithm and Some Applications". In: *Axioms* 7 (2018), p. 81.

[6] I. Markovsky and F. Dörfler. "Identifiability in the Behavioral Setting". In: *IEEE Trans. Automat. Contr.* (2023). DOI: 10.1109/TAC.2022.3209954.