

Learning Globally Stable Dynamics — a Matrix-theoretic Perspective

Peter Benner, Pawan K. Goyal, Siddarth Mamidiseti, Igor Pontes Duff, Süleyman Yildiz

Abstract

1 Motivation

The identification of dynamical systems from data has been considered for several decades, basically ever since *Cybernetics* emerged as a discipline from the wider research area of systems theory [Wiener (1948)]. In systems and control theory, system identification is an important and widely used technique in computer-aided control system design, available in any relevant software package, see, e.g., [Ljung (1999), Sima and Benner (2007), Benner et al. (2010)]. Nevertheless, most available identification tools are for linear systems, and do not necessarily consider constraints like *stability* or other physical properties. On the other hand, due to the advance of machine and deep learning methods, learning dynamical systems from data, in particular nonlinear systems, has recently become a field of renewed and massively growing interest. This is due, on the one hand, to the vast availability of measurement data in areas like the Earth system (weather and climate), biology, sociology, traffic etc., where traditionally no first principle mathematical models are available, but models are needed for prediction, while, on the other hand, the massive advancement in computational power nowadays allows to study large data sets using methods from artificial intelligence.

Nevertheless, the trajectories of time-dependent problems are often constrained by underlying physical principles that are usually not respected by classical black-box methods in machine and deep learning. Thus, physics-informed and physics-enhanced or "scientific" machine learning have emerged as new subdisciplines in the computational sciences and engineering and applied mathematics. Here, we will study in particular the stability of learned dynamical systems. In other words, we answer the question how to *guarantee* that the learned model of a dynamical system has desired stability properties, where we assume that some prior knowledge about the expected stability class is available. We show that stability constraints can be hard-coded into the learning model to guarantee, e.g., (asymptotic/Lyapunov/exponential) stability of linear systems, global asymptotic stability of nonlinear systems, and global stability of Hamiltonian systems. The used ideas stem from partially simple and obvious results in the theory of matrices and tensors. We will show that explicit parameterizations of the learned operators (matrices and tensors) defining the dynamical systems lead to constrained least-squares problems that can be solved using optimization routines that are now widely available in all software stacks for machine learning.

Exemplarily, we will demonstrate this approach for linear, uncontrolled, systems. In the talk, we will focus on the more challenging problems for nonlinear systems.

2 Learning Stable Linear Systems

Consider uncontrolled linear systems of the form

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0, \quad (1)$$

where $A \in \mathbb{R}^{n \times n}$, $x(t) \in \mathbb{R}^n$ is the state of the system at time $t \geq 0$, $\dot{x}$ denotes the derivative of x with respect to time, and $x_0 \in \mathbb{R}^n$ is the initial value.

If time-series data for $x(t)$ and $\dot{x}(t)$ are available, i.e., we have measured or computed vectors

$$x_0 = x(0), \quad x_1 = x(t_1), \quad \dots, \quad x_\ell = x(t_\ell)$$

and

$$x'_0 = \dot{x}(0), \quad x'_1 = \dot{x}(t_1), \quad \dots, \quad x'_\ell = \dot{x}(t_\ell)$$

at a sequence of time points $t_0 = 0 < t_1 < t_2 < \dots < t_\ell$, then a very simple way to infer the matrix A from the available data is to set up a linear least-squares problem using the data matrices

$$\begin{aligned} X &= [x_0, \dots, x_\ell] \in \mathbb{R}^{n \times \ell+1}, \\ X' &= [x'_0, \dots, x'_\ell] \in \mathbb{R}^{n \times \ell+1}. \end{aligned}$$

For (1), a possible formulation of an *operator inference problem* (*OpInf*) is then

$$A_* := \arg \min_A \|X' - AX\|_F^2 + \mathcal{R}(A) \quad (2)$$

with a potential regularization term $\mathcal{R}(A)$, e.g., for classical Tikhonov regularization (aka “ridge regression” in the machine learning community) $\mathcal{R}(A) = \beta \|A\|_F^2$, where $\beta > 0$ would be a regularization parameter. Here, one might also use a sparsity-promoting norm, e.g., $\mathcal{R}(A) = \beta \|A\|_q$, $q = 0, 1$. Note that for $\beta = 0$, or in general, $\mathcal{R}(A) \equiv 0$, (1) represents a standard linear least-squares problem, as the Frobenius norm turns into the standard Euclidian vector norm once $X' - AX$ is vectorized. From this, it is also obvious that this becomes a classical overdetermined system once $\ell \geq n$. Then minimization problem (1) can be solved using classical solution methods for linear least-squares problems like pivoted QR factorization or singular value decomposition (SVD).

The OpInf framework described so far has seen numerous applications and extensions in recent years. One of its major drawbacks still is that the inferred models can not be guaranteed to be stable (regarding local or global, asymptotic or Lyapunov stability) even if it is known that the data-generating model has a certain stability property. Some stability promoting weak enforcement strategies have been suggested, see, e.g., [Kaptanoglu et al. (2021)], but so far no certified stable models could be produced using the OpInf problem (2). Our goal is to introduce a parametrization of stability directly in the least-squares problem and in this way to obtain guaranteed stable models, following [Goyal et al. (2023b)]. Note that a very similar idea is also used in model reduction methods for port-Hamiltonian systems [Schwerdtner and Voigt (2023)].

The inference of asymptotically stable linear models with guarantee is based on the following somewhat surprising result, that was stated in [Gillis and Sharma (2017)], whereby related partial results can also be found in earlier literature. First, note that asymptotic stability of linear systems of the form (1) is fully characterized by the spectrum of A , $\Lambda(A)$, and in particular by the property that $\Lambda(A) \subset \mathbb{C}^-$, i.e., that the spectrum of A is fully contained in the open left half of the complex plane. The proof of the theorem is entirely based on this characterization.

Theorem 2.1 ([Gillis and Sharma (2017)]). *A matrix $A \in \mathbb{R}^{n \times n}$ is asymptotically stable (Hurwitz, Lyapunov stable) if and only if it can be represented as*

$$A = (J - R)Q,$$

where $J = -J^T$ and $R = R^T$, $Q = Q^T$ are both positive definite.

Now, it is straightforward to replace the linear least-squares problem (2) by the following inference problem:

$$(J_*, R_*, Q_*) := \arg \min_{\substack{J=-J^T \\ R=R^T \succ 0 \\ Q=Q^T \succ 0}} \|X' - (J - R)QX\|_F^2 + \mathcal{R}(J, Q, R). \quad (3)$$

Unfortunately, this problem is relatively hard to solve due to the positive definiteness constraints. Therefore, in [Goyal et al. (2023b)], we suggest to use the following parametrization:

$$J = S - S^T, \quad R = L^T L, \quad Q = K^T K,$$

where $S \in \mathbb{R}^{n \times n}$ is a general square matrix and $L, K \in \mathbb{R}^{n \times n}$ are upper-triangular matrices (or Cholesky factors). Then the OpInf problem for linear systems becomes

$$(S_*, L_*, K_*) := \arg \min_{\substack{L, K \text{ upper} \\ \text{triangular}}} \left(\|X' - (S - S^T - L^T L)K^T KX\|_F^2 + \mathcal{R}(L, K, S) \right). \quad (4)$$

If this problem can be solved, the obtained matrix

$$A_* = (S_* - S_*^T - L_*^T L_*) K_*^T K_*$$

is guaranteed to be asymptotically stable due to Theorem 2.1 and solves (2) under the asymptotic stability constraint. The price paid for inferring a model with stability certificate is that even for a zero regularizer, the inference problem (4) is nonlinear in the entries of L, K . Fortunately, with the advance of machine learning methods, such problems can be solved using (stochastic) gradient descent methods implemented in tools like PyTorch.¹

In our talk, we will describe how this result and the related stability-guaranteed OpInf problem can be extended to controlled systems $\dot{x} = Ax + Bu$, where u is a control input [Pontes Duff et al. (2024)] and to parameter dependent systems, where $A = A(\mu)$ depends on a parameter vector $\mu \in \mathbb{R}^q$.

3 Outlook: Nonlinear Systems

While for linear systems, local and global stability as well as asymptotic, Lyapunov, and exponential stability concepts are equivalent, these have to be distinguished for nonlinear systems. Obviously, using the parametrization of A from (4) to an operator inference problem for a nonlinear system with linear part $Ax(t)$, the inferred system will be locally Lyapunov stable with Lyapunov function

$$V(x) := \frac{1}{2} x^T Q x.$$

In [Goyal et al. (2023a)], we additionally study the identification of globally Lyapunov stable quadratic systems as well as systems with bounded domain of attraction. This is again based on explicit parameterizations of matrices and tensors. These results will be presented in the talk as well as ideas on how to guarantee global stability of Hamiltonian systems, where we employ techniques from deep learning and elementary properties of symplectic matrices [Goyal et al. (2023c)]. The theoretical findings will be illustrated by several numerical examples.

¹<https://pytorch.org/>

References

- [Benner et al. (2010)] Benner, P., Kressner, D., Sima, V., and Varga, A. (2010). Die SLICOT-Toolboxen für MATLAB. *at-Automatisierungstechnik*, 58(1), 15–25.
- [Gillis and Sharma (2017)] Gillis, N. and Sharma, P. (2017). On computing the distance to stability for matrices using linear dissipative Hamiltonian systems. *Automatica*, 85, 113–121.
- [Goyal et al. (2023a)] Goyal, P., Pontes Duff, I., and Benner, P. (2023a). Guaranteed stable quadratic models and their applications in SINDy and operator inference. e-print arXiv:2308.13819.
- [Goyal et al. (2023b)] Goyal, P., Pontes Duff, I., and Benner, P. (2023b). Stability-guaranteed learning of linear models. e-print arXiv:2301.10060.
- [Goyal et al. (2023c)] Goyal, P., Benner, P., and Yıldız, S. (2023c). Deep learning for structure-preserving universal stable Koopman-inspired embeddings for nonlinear canonical Hamiltonian dynamics. e-print arXiv:2308.13835.
- [Kaptanoglu et al. (2021)] Kaptanoglu, A., Callahan, J., Aravkin, A., Hansen, C., and Brunton, S. (2021). Promoting global stability in data-driven models of quadratic nonlinear dynamics. *Physical Review Fluids*, 6(9), 094401.
- [Ljung (1999)] Ljung, L. (1999). *System Identification – Theory for the User*. Prentice-Hall, Upper Saddle River, NJ, 2nd edition.
- [Peherstorfer and Willcox (2016)] Peherstorfer, B. and Willcox, K. (2016). Data-driven operator inference for nonintrusive projection-based model reduction. *Computer Methods in Applied Mechanics and Engineering*, 306, 196–215.
- [Pontes Duff et al. (2024)] Pontes Duff, I., Goyal, P., and Benner, P. (2024). Stability-certified learning of control systems with quadratic nonlinearities. *IFAC-PapersOnLine*, 58(17), 151–156..
- [Schwerdtner and Voigt (2023)] Schwerdtner, P. and Voigt, M. (2023). SOBMOR: Structured optimization-based model order reduction *SIAM Journal on Scientific Computing* 45(2), A502–A529.
- [Sima and Benner (2007)] Sima, V. and Benner, P. (2007). Fast system identification and model reduction solvers. In B.R. Andrievsky and A.L. Fradkov (eds.), *Preprints 9th IFAC Workshop “Adaptation and Learning in Control and Signal Processing” (ALCOSP 2007), August, 29-31, 2007, Saint Petersburg, Russia*.
- [Wiener (1948)] Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT Press, Cambridge, Massachusetts.