Robust Algebraic Parameter Estimation via Gaussian Process Regression ⋆

Oren Bassik * **Alexander Demin** ** **Alexey Ovchinnikov** ***

* *Graduate Center, City University of New York, 365 Fifth Avenue, New York, NY 10016, USA (e-mail: obassik@gradcenter.cuny.edu).*
** *HSE University, Moscow, Russia (e-mail: asdemin_2@edu.hse.ru).*
*** *Department of Mathematics, CUNY Queens College, and Ph.D. Programs in Mathematics and Computer Science, CUNY Graduate Center, New York, NY, USA (e-mail: aovchinnikov@qc.cuny.edu).*

Abstract: To estimate parameters in ODE systems from noisy data, we propose adding a Gaussian Process Regression (GPR) step to an approach based on differential algebra. This replaces the interpolation step in the original algorithm and allows for more reliable estimation even in the presence of noise. We tested the method on a suite of benchmark problems from multiple disciplines. With GPR, we are able to retain the robust features inherent in the algebraic approach, while extending applicability to realistic data with noise.

Keywords: Parameter and state estimation, software for system identification, continuous time system estimation

1. INTRODUCTION

Parameter estimation for systems of ordinary differential equations is a fundamental problem in systems modeling and dynamics. A typical approach to estimating parameters from experiments uses nonlinear optimization to search for parameters which minimize error against observed data. This inherits various difficulties from nonlinear optimization: the need for good initial guesses, getting stuck in local minima, and only finding a single solution.

An algorithm outlined in (Ovchinnikov et al., 2023) combines differential algebra with baryrational interpolation and multivariate polynomial systems solving. This method does not suffer from the non-robustness inherent in nonlinear optimization. It needs no initial guesses, requires little input from the user, and ideally finds all solutions (in the case of local identifiability.)

This differential algebraic method shows excellent performance on noise-free or synthetic data, but due to the reliance on interpolation degrades severely with even minimal measurement noise. To overcome this limitation, we replace the baryrational interpolation with Gaussian Process Regression (GPR) (see, for example, Rasmussen and Williams (2006)) using a squared exponential kernel, and automatically learn the noise hyperparameter from the data.

This simple replacement significantly improves robustness to measurement noise in observed data. We tested the method on a suite of benchmark problems from multiple disciplines. With GPR, we are able to retain the robust features inherent in the algebraic approach, while extending applicability to realistic data with noise.

⋆ This work was partially supported by the NSF under grants CCF-2212460 and DMS-1853650.

2. STATEMENT OF PROBLEM

We are given an ODE model

$$\begin{gathered}\mathbf{x}' = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p})\\ \mathbf{y} = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p})\\ \mathbf{x}(0) = \mathbf{x}_0\end{gathered}$$

where $\mathbf{x}$ is a vector of state variables, $\mathbf{u}$ is a vector of control variables, $\mathbf{p}$ is a vector of unknown parameters, x_0 is a vector of unknown initial conditions, and $\mathbf{f}$ and $\mathbf{g}$ are rational functions in $\mathbf{x}, \mathbf{p}, \mathbf{u}$.

We are given a set of measurements of the form $(t_i, \mathbf{y}_i)$ for $i = 1, 2, \ldots, N$. We seek to estimate the parameters $\mathbf{p}$ and the unknown initial conditions $\mathbf{x}_0$ from the measurements.

3. METHOD

We give a bird's eye view of the method, and track a toy problem through the algorithm. While the toy problem is simple, the method generalizes fully to higher-dimensional systems. The main constraint on the input system is that all functions must be rational.

3.1 Toy problem

Let $x' = a^2x^2 + b$ and $y = x^2 + x$. We seek to estimate the parameters a, b, and the initial condition $x(0)$ from measurements of y taken at several points.

3.2 Step 1: Differentiate the ODE system

$$\begin{aligned} x' &= a^2x^2 + b \\ x'' &= 2a^2xx' \end{aligned} \qquad \begin{aligned} y &= x^2 + x \\ y' &= 2xx' + x' \\ y'' &= 2(x'x' + xx'') + x'' \end{aligned}$$

Step 2: Approximate the derivatives from data Pick a time point t_i. Approximate the derivatives of the measurements y at t (see Section 4 for the details). For example, let

$$y_0 \approx y(t_i),\ y_1 \approx y'(t_i),\ y_2 \approx y''(t_i)$$

and also denote $x(t_i)$, $x'(t_i)$, and $x''(t_i)$ by x_0, x_1, and x_2, respectively. To be clear, the y_i's are real number constants.

Step 3: Form the system of equations Use the approximated derivatives y_0, y_1, y_2 to form a system of polynomial equations:

$$\begin{array}{ll} x_1 = a^2 x_0^2 + b & y_0 = x_0^2 + x_0 \\ & y_1 = 2x_0 x_1 + x_1 \\ x_2 = 2a^2 x_0 x_1 & y_2 = 2(x_1 x_1 + x_0 x_2) + x_2 \end{array}$$

Care must be taken in this and previous steps to get a square system. Here we have 5 equations in 5 unknowns: (x_0, x_1, x_2, a, b).

Step 4: Solve the system of equations Use any method to solve the system of equations to find the indeterminates a, b, x_0, x_1, x_2. Ideally, we find all solutions. In this case, we expect at least two solutions, because of the symmetry between a and $-a$. Our software supports many solvers, including Gröbner-based and homotopy continuation solvers.

Step 5: Backsolve and filter the solutions If t_i was not the initial time, then for each solution we found, we can backsolve the ODE to find the initial conditions $x(0)$. We can also forward solve and compare to the measurements to calculate error.

4. THE DATA PROCESSING STEP

It is a theorem that for locally identifiable parameters, using (a) perfect approximations of the derivatives in step 2, and (b) a certified polynomial root finder, we can recover all parameters (Hong et al., 2020). Both of these issues are significant numerical analysis problems in their own right. In (Ovchinnikov et al., 2023) we tested this algorithm on synthetic data sampled from precise ODE solutions and found that for estimating derivatives, AAA baryrational interpolation followed by algorithmic differentiation of the interpolant outperformed many alternatives (including polynomial interpolants, splines, Fourier interpolation, and finite differences).

However, AAA is an interpolation scheme, and as such, even small amounts of measurement noise (10^{-8}) cause overfitting, oscillation, and particularly poor estimates for derivatives of observables.

In this work, we modify our algorithm by replacing the interpolation step with a Gaussian Process Regression. This is a standard regression used in machine learning which estimates (by maximizing likelihood) the mean function and noise most likely to have generated the observed data. Specifically, we use a squared-exponential kernel and learn the regression hyperparameters (noise variance and lengthscale) by tuning to the data. The mean function returned from the regression is smooth (infinitely differentiable), and we apply algorithmic differentiation to this mean function.

5. RESULTS

We tested the extended algorithm on a suite of benchmark problems, including some population dynamics models (Lotka-Volterra, SEIR) and some systems from biology (Crauste NELM model, HIV dynamics, Fitzhugh-Nagumo). For each system, we generated synthetic data and added varying levels of noise (from 1e-8 to 1e-2.) We ran parameter estimation using both AAA interpolation and the GPR-enhanced method, as well as a baseline comparison vs. a traditional loss-function minimizer estimation package.

For our benchmarking set, we observe that the approach using GPR is significantly better than our original interpolation-based method, shows similar accuracy to IQM and SciML, and that AMIGO2 typically showed better performance. While estimation loses precision with higher noise (as it must), the algebraic approach is no longer handicapped vs optimization based approaches.

6. DISCUSSION AND FUTURE DIRECTIONS

The algorithm above is currently bottlenecked by solving the polynomial system, and as such, the integration of GPR does not significantly affect the computation run time of our method. This modest and low-cost addition significantly increases the applicability of algebraic parameter estimation, as most real-world data is sampled imprecisely. Because GPR is self-tuning, the algorithm retains its “hands-off” nature, demanding little from the user. Given these benefits, we have added GPR to our default suite of interpolators.

Gaussian process regression gives not only an estimate of noise variance, but actually furnishes confidence bounds around the mean function and is used for uncertainty quantification. While most GPR packages do not provide such uncertainty quantification for higher derivatives, an interesting avenue for future research could be to use this information, for picking optimal timepoints or providing confidence intervals on parameter estimates.

SOFTWARE AVAILABILITY

The code for this work is available at `https://github.com/orebas/ODEParameterEstimation`.

REFERENCES

Hong, H., Ovchinnikov, A., Pogudin, G., and Yap, C. (2020). Global identifiability of differential models. *Communications on Pure and Applied Mathematics*, 73(9), 1831–1879.

Ovchinnikov, A., Bassik, O., Berman, Y., Go, S., Hong, H., Ilmer, I., Rackauckas, C., Soto, P., and Yap, C. (2023). Robust parameter estimation for rational ordinary differential equations. URL `https://arxiv.org/abs/2303.02159`.

Rasmussen, C.E. and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning.* The MIT Press.