Robust Parameter Estimation for Biological Systems

Justin Krueger[1], Mihai Pop[2], and Matthias Chung[3]

Short Abstract — **The use of modeling to interpret data and predict dynamics has significantly contributed to advances in the understanding of cellular regulatory systems, but quantitative methods must continue to adapt to handle the increasing complexity of this understanding. To this end, we have developed a method to address many of the issues associated with traditional parameter estimation techniques for differential equations models of complex dynamical systems. By defining a "nearby" problem, our method improves significantly on the efficiency and robustness of traditional approaches while providing accurate parameter estimates.**

Keywords — **Parameter estimation, Differential equations, Optimization.**

I. PROBLEM STATEMENT

QUANTITATIVE experimentation and predictive modeling have become significant tools for enhancing biological research. Particularly in the study of cellular regulatory systems, the use of modeling has helped to analyze collected data and focus future research [3]. With experimental methods continuing to improve, the related quantitative techniques must evolve to remain relevant. This means, for example, developing methods for broader ranges of problems and bigger data sets.

In general, the problem statement for parameter estimation for differential equation models involves finding the model parameters that minimize some measurement of error between the solution of the initial value problem and the experimental data set, i.e.,

$$\min_{p} \left\| m(y) - d \right\|_2^2 \text{ subject to } y' = f(t, y, p), \ y(0) = y_0,$$

where d is the given data, t the time, and y the state solution of the initial value problem with parameters p. The function m projects the state solution onto the data d.

Traditional methods such as single shooting methods are used for solving parameter estimation problems for biological systems. These methods require a sequential scheme for solving the initial value problem and have been shown to be inefficient and to lack robustness [7].

[1]Department of Mathematics, Virginia Tech. E-mail: kruegej2@vt.edu
[2]Department of Computer Science, University of Maryland, College Park. E-mail: mpop@umiacs.umd.edu
[3]Department of Mathematics, Virginia Tech. E-mail: mcchung@vt.edu

II. ROBUST PARAMETER ESTIMATES

To improve on the efficiency and robustness of single shooting methods, we build on previous alternatives [2,4,5]. We replace the state solution y with a spline approximation s uniquely defined by spline parameters q and relax the hard constraints to achieve the new problem statement

$$\min_{p,q} \left\| m(s(q)) - d \right\|_2^2 + \lambda \left\| s'(q) - f(t, s(q), p) \right\|_2^2 + \alpha \left\| c(p, s(q)) \right\|_2^2.$$

This equation defines a "nearby" problem that can be solved using more robust numerical methods.

We test our method on *in silico* data generated by generalized Lotka-Volterra equations. We then test our method on data from an intestinal microbiota experiment, and we compare our results to a published parameterized model [1,6]. In the first case, we recover both the parameters and data, and in the second, our method provides improved data recovery relative to the published results.

III. CONCLUSION

Our approach of defining a "nearby" problem eliminates the computational inefficiency and robustness-limiting step of numerically solving an initial value problem at every optimization step. In doing so, we have not sacrificed parameter or data recovery. This suggests our method improves on the capabilities of traditional parameter estimation methods and is a valuable tool, particularly for the increasingly complex parameter estimation problems like those found in cellular regulatory systems.

REFERENCES

[1] Buffie CG, et al. (2012) Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to Clostridium difficile-induced colitis. *Infect Immun* **80**(1), 62-73.

[2] Chung M, Haber E (2012) Experimental design for biological systems. *SIAM J Control Optim* **50**(1), 471-489.

[3] Diaz P, et al. (2009) A mathematical model of the immune system's role in obesity-related chronic inflammation. *SIURO* **2**(2).

[4] Poyton AA, et al. (2006) Parameter estimation in continuous-time dynamic models using principal differential analysis. Computers & Chemical Engineering **30**(4), 698-708.

[5] Ramsay JO (1996) Principal differential analysis: data reduction by differential operators. *J R Stat Soc Series B Stat Methodol* **58**(3), 495-508.

[6] Stein RR, et al. (2013) Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota. *PLoS Comput Biol* **9**(12).

[7] Stoer J (2002) *Introduction to Numerical Analysis.* Springer-Verlang, New York.