Parameter identification for systems of ODEs inspired by model theory

Thomas Scanlon

UC Berkeley

11 July 2021

The collaboration

This is joint work with Alexey Ovchinnikov, Anand Pillay, and Gleb Pogudin funded by an NSF FRG grant.

- Computing all identifiable functions of parameters for ODE models, `arXiv:2004.07774`
- Multi-experiment parameter identifiability of ODEs and model theory, `arXiv:2011.10868`

State space models

A state space model is given by a system of ODEs

$$\dot{x} = f(x, \mu, u) \tag{1}$$

$$y = g(x, \mu, u) \tag{2}$$

and possibly a constraint

$$0 = h(x, \mu, u) \tag{3}$$

- x is a vector of state variables
- u is a vector of input variables
- y is a vector of output variables
- μ is a vector of constants called the parameters
- For us, f, g, and h are vectors of rational functions with rational coefficient and will usually omit the constraint Equation 3.

Kermack and McKendrick, *Proceedings of the Royal Society of Edinburgh*, 1927

A Contribution to the Mathematical Theory of Epidemics.

By W. O. Kermack and A. G. McKendrick.

(Communicated by Sir Gilbert Walker, F.R.S.—Received May 13, 1927.)

(From the Laboratory of the Royal College of Physicians, Edinburgh.)

Introduction.

(1) One of the most striking features in the study of epidemics is the difficulty of finding a causal factor which appears to be adequate to account for the magnitude of the frequent epidemics of disease which visit almost every population. It was with a view to obtaining more insight regarding the effects of the various factors which govern the spread of contagious epidemics that the present investigation was undertaken. Reference may here be made to the work of Ross and Hudson (1915–17) in which the same problem is attacked. The problem is here carried to a further stage, and it is considered from a point of view which is in one sense more general. The problem may be summarised as follows: One (or more) infected person is introduced into a community of individuals,

Kermack-McKendrick model with constant parameters

from the consideration of the special case in which ϕ and ψ are constants κ and l respectively.

In this case the equations are

$$\left.\begin{aligned} \frac{dx}{dt} &= -\kappa xy \\ \frac{dy}{dt} &= \kappa xy - ly \\ \frac{dz}{dt} &= ly \end{aligned}\right\} \qquad (29)$$

and as before $x + y + z = \mathrm{N}$.

Kermack-McKendrick model with constant parameters

$$\begin{aligned}\frac{dx}{dt} &= -\kappa xy \\ \frac{dy}{dt} &= \kappa xy - ly \\ \frac{dz}{dt} &= ly \\ x + y + z &= N\end{aligned}$$

- x represents the number of susceptible people in the population
- y represents the number of infected / infectious persons
- z represents the number of recovered (or dead) people
- κ, l, and N are constants representing infectivity, recovery (plus death) rates, and the total population.

Identifiability

We will be interested in the parameter identification problem: can the parameters μ be recovered from the output y? If so, how? Sometimes, this problem goes under the name of system identification.

There are variants of this problem in which the input variables u are assumed to be known or not. There are related problems of determining the state x from the output y or even of inferring the input u from the output y.

We shall interpret recovered from as definable (relative to the theory of differentially closed fields of characteristic zero) from, which by quantifier elimination is equivalent to expressed as a differential rational function of. Moreover, we shall ask (x, y) to be a generic solution of the equations for a sufficiently general (even generic) u. So, given such generic solutions to Equations 1 and 2, we wish to compute $\mathbb{Q}(\mu) \cap \mathbb{Q}\langle u, y\rangle$ and, in particular, wish to determine whether this intersection is $\mathbb{Q}(\mu)$.

Identifiability in the Kermack-McKendrick model

The parameters κ and l are identifiable in the Kermack-McKendrick model if we take both x and y as outputs.

$$\begin{aligned} \frac{dx}{dt} &= -\kappa xy \\ \frac{dy}{dt} &= \kappa xy - ly \\ \frac{dz}{dt} &= ly \\ x + y + z &= N \end{aligned}$$

$$\begin{aligned} \kappa &= \frac{-\dot{x}}{xy} \\ l &= \frac{-(\dot{x} + \dot{y})}{y} \end{aligned}$$

Canonical parameters

There may be obvious reasons why it is impossible to identify the parameters.

- For example, if some transcendental component of μ does not appear in Equation 1 at all, then it would be impossible to compute μ from y. For example, suppose that $\mu = (\mu_1, \mu_2)$ and our equation is $\dot{x} = x^2 + \mu_1$.
- For a less trivial example, it may happen that the system is equivalent to one in which the coefficients are rational functions of μ. For example, our equations might be $\dot{x} = x^2 + \mu_1 + \mu_2$ and $y = x$.

At the very least, if we wish for the parameters to be indentifiable, then they need to be canonical parameters: any other choice of parameters would give an inequivalent system of equations.

Model theory meets modeling at CUNY, Spring 2019

Canonical parameters, model theoretically

The canonical parameter is a standard notion of model theory.

We say that a formula $\phi(x, y)$ has canonical parameters if for any two choices of parameters c and d, we have that $(\forall x)(\phi(x, c) \leftrightarrow \phi(x, d))$ if and only if $c = d$. In this case, we would say that c is the canonical parameter for $\phi(x, c)$.

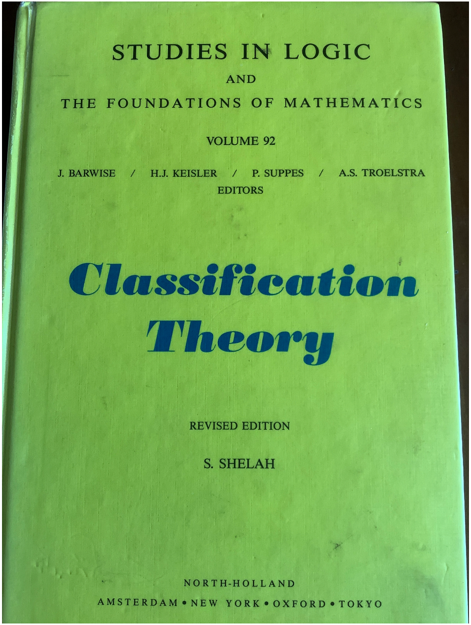
We say that our theory eliminates imaginaries if for each formula $\phi(x, y)$ there is some formula $\psi(x, z)$ so that

- every instance of ϕ is equivalent to an instance of ψ: $(\forall y)(\exists z)(\forall x)(\phi(x, y) \leftrightarrow \psi(x, z))$
- every instance of ψ is equivalent to an instance of ϕ, and
- $\psi(x, z)$ has canonical parameters

For us, the key point is that the theory of differentially closed fields of characteristic zero, DCF_0, eliminates imaginaries. So, every finite system of ODEs (and inequalities) is equivalent to one with canonical parameters.

We will need the related, but more refined, notion of a canonical base of a type in a stable theory.

Canonical bases: Classification Theory



Canonical bases: Theorem III.6.10

THEOREM 6.10 (The Canonicity Theorem for Types): *Let T be stable. In $\mathfrak{C}^{eq}$, for every stationary m-type p there is a set $A = \mathrm{Cb}(p)$ (the canonical base of p) and a stationary type $q \in S^m(A)$, $q = \mathrm{Ctp}(p)$ (= the canonical type of p) such that:*

(1) *p and q are parallel.*

(2) *Any automorphism F of $\mathfrak{C}^{eq}$ takes p to a parallel type iff $F \restriction A =$ identity.*

(3) *For any $a \in \mathfrak{C}^{eq}$, $a \in A$ iff every automorphism F of $\mathfrak{C}^{eq}$ which takes p to a parallel type necessarily satisfies $F(a) = a$.*

(4) *There is $B \subseteq A$, $|B| < \kappa(T)$ over which no type parallel to p forks.*

(5) *If a type in $S^m(C)$ is parallel to p and does not fork over $B \subseteq C$ then $A \subseteq \mathrm{acl}\, B$.*

(6) *$A \subseteq \mathrm{dcl}(\mathrm{Dom}\, p)$.*

Proof. Let $r \in S^m(\mathfrak{C}^{eq})$ be parallel to p. By II, 2.2 for every $\varphi(\bar{x}; \bar{y})$ there is a formula $\psi_\varphi(\bar{y}; \bar{c}_\varphi)$ which defines $r \restriction \varphi$, i.e.,

$$\models \psi_\varphi[\bar{a}; \bar{c}_\varphi] \Leftrightarrow \varphi(\bar{x}; \bar{a}) \in r.$$

Let

$$A = \mathrm{dcl}\{\bar{c}_\varphi / E_{\psi_\varphi} : \varphi \in \mathrm{L}^{eq}\},$$
$$q = r \restriction A.$$

Clearly r is definable over A, hence does not fork over A by 4.11. So q is stationary and q, r are parallel, hence p, q are parallel (this is (1)).

Clearly an automorphism F of $\mathfrak{C}^{eq}$ takes p to a parallel type iff $F(r) = r$ iff for every φ, $\models \psi_\varphi(\bar{y}; \bar{c}_\varphi) \equiv \psi_\varphi(\bar{y}; F(\bar{c}_\varphi))$ iff for every φ, $\bar{c}_\varphi / E_{\psi_\varphi} = F(\bar{c}_\varphi)/E_{\psi_\varphi}$ iff $F \restriction A =$ identity (this is (2)). If $a \notin A$ then as $A = \mathrm{dcl}\, A$ there is $a' \neq a$, $\mathrm{tp}(a, A) = \mathrm{tp}(a', A)$ hence there is an automorphism F of $\mathfrak{C}^{eq}$, $F \restriction A =$ identity and $F(a) = a'$; so $F(r) = r$, $F(a) \neq a'$ (this is the missing part in (3)). From (3) it is clear that A does not depend on the particular choice of the ψ_φ: Also A depends only on r, i.e., on the equivalence class of p under parallelism. The same is true for q, hence the notation $A = \mathrm{Cb}(p)$, $q = \mathrm{Ctp}(p)$ is justified.

As we could have chosen $\bar{c}_\varphi \in \mathrm{Dom}\, p$ (by II, 2.12) and as $\bar{c}_\varphi / E_{\psi_\varphi} \in$

Some stability theory

Fix a saturated model $M \models T$ of some theory T in a language $\mathscr{L}$, a natural number m, sets $A \subseteq B \subseteq M$ with $|B| < ||M||$, and an m-type $p \in S_m(B)$. We write $\mathrm{x} = (x_1, \ldots, x_m)$.

- p is A-definable if for each formula $\phi(\mathrm{x}, \mathrm{y})$, there is a formula $\psi(\mathrm{y}) \in \mathscr{L}_A$ so that

$$\{\mathrm{b} \in B^{|\mathrm{y}|} \ : \ \phi(\mathrm{x}, \mathrm{b}) \in p\} = \{\mathrm{b} \in B^{|\mathrm{y}|} \ : \ M \models \psi(\mathrm{b})\}$$

- p is stationary if there is a unique B-definable extension of p to a type over M.
- A is a canonical base of p if p is stationary, A-definable, and for any automorphism $\sigma : M \to M$ of M, the types p and $\sigma(p)$ have the same B-definable (respectively, $\sigma(B)$-definable) extensions to M if and only if σ fixes A pointwise. We write $A = \mathrm{Cb}(p)$.

Canonical bases in differentially closed fields

We specialize the definitions to the case of DCF_0.

Let (K, ∂) be a differential field, $k \subseteq K$ a differential subfield, and a a tuple from K. We write $I(\mathrm{a}/k) := \{f \in k\{\mathrm{x}\} \ : \ f(\mathrm{a}) = 0\}$ for the ideal of a over k.

- The type of a of k is stationary if and only if $I(\mathrm{a}/k)$ is absolutely prime. That is, the ideal generated by $I(\mathrm{a}/k)$ in $k^{\mathrm{alg}}\{x\}$ is prime.
- Provided that $\mathrm{tp}(\mathrm{a}/k)$ is stationary, the canonical base of a over k is the differential field of definition of $I(\mathrm{a}/k)$.

The canonical base $\mathrm{Cb}(\mathrm{a}/k)$ may be realized as the differential field generated by the canonical parameters of a formula isolating the type of a over k up to dependence. Algebraically, it may be realized as the differential field generated by the coefficients of the monic differential polynomials in a characteristic generating set for $I(\mathrm{a}/k)$.

Canonical base and parameter identifiability

Let us restrict to a simple case where our equation takes the form

$$\dot{x} = f(x, \mu)$$

$$y = x$$

so that there are no input variables and the state and output variables are identical.

If we set $k = \mathbb{Q}(\mu)$ and let a be a generic solution, then the type of a over k is stationary.

If μ is a canonical parameter for the formula $\dot{x} = f(x, \mu)$, then $k = \mathrm{Cb}(a/k)$.

So, the parameter identifiability problem reduces to asking whether the canonical base $\mathrm{Cb}(a/k)$ is contained in the differential field generated by a, or in more model theoretic terms, in the definable closure of a realization of the generic type of this system.

Abstract failure of single experiment identifiability

- In general stable theories it is “rare” for the canonical base of a type to be definable (or even algebraic) from a single realization.
- Theories where this always happens are (provably) degenerate or closely related to linear algebra.
- Computing the canonical base from one solution even for ordinary polynomial equations can be impossible. If $f(\mathrm{x}, \mathrm{y})$ is a polynomial over $\mathbb{Q}$ for which $f(\mathrm{x}, \mathrm{b})$ is always absolutely irreducible, then a generic solution to $f(\mathrm{a}, \mathrm{b}) = 0$ will be stationary over $\mathbb{Q}(\mathrm{b}) = \mathrm{Cb}(\mathrm{a}/\mathbb{Q}(\mathrm{b}))$. If $\mathrm{x} = (x_1, \ldots, x_n)$ has $n > 2$, then it is not possible to compute b from a.

A failure of single-experiment identifiability

A standard example relative to the theory of algebraically closed fields showing that the canonical base need not be algebraic over a single realization of a type comes families of lines in the plane.

Let $b, c, d \in C$ be three algebraically independent elements. Set $e := bd + c$ and let $k = \mathbb{Q}(d, e)$.

The type $\mathrm{tp}((b, c)/k)$ is generated by the formulae $x_2 + dx_1 = e$ and $p(x_1) \neq 0$ for $p \in k[x] \smallsetminus \{0\}$. So, $\mathrm{Cb}((b, c)/k) = k$.

In particular, $\mathrm{Cb}((b, c)/k) \not\subseteq \mathbb{Q}(b, c)$.

Consider a satisfying $\partial(a) = ba + c$. A simple computation shows that (a, b) is the generic solution to the following system.

$$\begin{aligned} \dot{x_1} &= x_1 x_2 - dx_2 + e \\ \dot{x_2} &= 0 \end{aligned}$$

This system violates single experiment identifiability.

Multi-experiment identifiability

Given an input-output system

$$\begin{aligned}\dot{x} &= f(x, \mu, u) \\ y &= g(x, \mu, u)\end{aligned}$$

one might ask whether the parameters μ are identifiable from multiple independent experiments.

- Of course, as before we must assume that the parameters μ are canonical.
- If the answer is yes, then we would like to compute a bound on the number of experiments needed.
- The bounds may depend on whether we vary the input variable or not between the various experiments.

Model theory meets identifiability problems in systems biology at AIM, Summer 2019

Computing canonical bases from Morley sequences

There is a very general model theoretic result which says that multi-experiment parameter identification is always possible, even if we allow our parameters to be nonconstant.

Theorem

In any totally transcendental theory if p is a stationary type over the set B, then there is a number N so that if $\mathrm{a}_1, \ldots, \mathrm{a}_N$ is a sequence of independent realizations of p, then $\mathrm{Cb}(p)$ is definable from $\langle \mathrm{a}_1, \ldots, \mathrm{a}_N \rangle$. Moreover, if b is a tuple from which $\mathrm{Cb}(p)$ is definable and the Lascar rank of b is $s < \omega$, then it suffices to take $N = s + 1$.

Interpreting the general theorem for systems of ODEs

- The theory of differentially closed fields of characteristic zero is the quintessential example of a totally transcendental theory.
- Independence may be defined differential algebraically: if $\mathrm{tp}(\mathrm{a}/L)$ is stationary over the differential field L and M is an differential extension field, then a is independent from L over M if $I(\mathrm{a}/M)$ generates $I(\mathrm{a}/L)$. A sequence $\langle \mathrm{a}_1, \ldots, \mathrm{a}_n \rangle$ is independent over the differential field M if for each $i < n$, a_{i+1} is stationary over M and independent from $M\langle \mathrm{a}_1, \ldots, \mathrm{a}_i \rangle$ over M.
- In a differentially closed field, an element c is definable from some tuple b just in case $c \in \mathbb{Q}\langle \mathrm{b} \rangle$.
- The Lascar rank is a dimension defined using (in)dependence. For us, the main point is that the Lascar rank of a tuple b is bounded above by $\mathrm{tr.deg}\, \mathbb{Q}\langle \mathrm{b} \rangle$ and when b is a tuple of constants, the Lascar rank is equal to this transcendence degree.

Algorithms

- Interpreting the Shelah reflection principle as Lagrange interpolation, we produced an algorithm to compute an upper bound on the number of independent experiments that would be necessary to recover the parameters.
- Using general results on canonical bases, we produced an algorithm to compute the identifiable parameters. That is, even if the parameters in our original model are not identifiable from a singe experiment, there may be nontrivial rational combinations of those parameters which are identifiable and we work out what those will be.
- These algorithms have been implemented in the Julia language and are available at `https://github.com/pogudingleb`.

Extensions

- The general results from model theory we have described work equally well for PDEs, but what specific consequences they have and the computational approach remain to be investigated.
- Extensions to difference equations and difference-differential equations should be possible, but here the model theory is somewhat more complicated.