Global Identifiability of Parametric Differential Models

Chee Yap

Courant Institute, NYU
(Visiting)
Academy of Mathematics & System Sciences
Chinese Academy of Scieces, Beijing

School of Computing, KAIST
Daejeon, S.Korea
May 16, 2018

Joint work with
Hoon Hong, Alexey Ovchinnikov, Gleb Pogudin

Abstract

Global Identifiability of Differential Models

Hoon Hong, Alexey Ovchinnikov, Gleb Pogudin, Chee Yap

(Submitted on 24 Jan 2018)

Many real-world processes and phenomena are modeled using systems of ordinary differential equations with parameters. Given such a system, we say that a parameter is globally identifiable if it can be uniquely recovered from input and output data. The main contribution of this paper is to provide the theory, an algorithm, and software for deciding global identifiability. First, we rigorously derive an algebraic criterion for global identifiability (this is an analytic property), which yields a deterministic algorithm. Second, we improve the efficiency by randomizing the algorithm while guaranteeing probability of correctness.

Subjects:	**Classical Analysis and ODEs (math.CA)**; Commutative Algebra (math.AC); Algebraic Geometry (math.AG); Optimization and Control (math.OC)
MSC classes:	12H05, 14Q20, 68W20, 68W30, 93C15, 93B25, 93B40, 93A30
Cite as:	**arXiv:1801.08112 [math.CA]**
	(or **arXiv:1801.08112v1 [math.CA]** for this version)

Submission history

From: Alexey Ovchinnikov [view email]
[v1] Wed, 24 Jan 2018 18:28:01 GMT (31kb)

Overview

I. Introduction
II. Global Identifiability
III. Our Results
IV. Conclusion

I. Introduction

Introduction

- Modeling of processes in the real world

Introduction

- Modeling of processes in the real world
  - Chemistry: Rate of Reaction

$$NO_3 + CO \implies NO_2 + CO_2.$$

Concentrations of reactants are

$$\langle NO_3\rangle, \langle CO\rangle, \langle NO_2\rangle, \langle CO_2\rangle.$$

Parametric differential equation:

$$\boxed{\frac{d\langle NO_3\rangle}{dt} = -\mu\langle NO_3\rangle \cdot \langle CO\rangle}$$

(the parameter μ is the rate of reaction)

Parameter Estimation Problem: determine μ

There are 3 other equations

Introduction

- Modeling of processes in the real world
  - Biology: Predator-prey model

Introduction

- Modeling of processes in the real world
  - Biology: Predator-prey model

    x_1: prey population (e.g., zebras)

    x_2: predator population (e.g., lions)

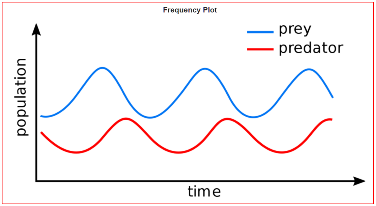
$$\Sigma : \begin{cases} x_1' = \alpha x_1 - \beta x_1 x_2 \\ x_2' = \delta x_1 x_2 - \gamma x_2 \end{cases}$$

(Lottka-Volterra Equations)

where $\alpha, \beta, \delta, \gamma$ are positive parameters.

Introduction

- Biology: Predator-prey model



The solution is periodic (unless there is extinction)

Introduction

- Parametric Differential Algebraic (PDA) Model

$$\Sigma = \Sigma(\mathbf{x}, \mathbf{y}, u; \boldsymbol{\mu}, \mathbf{x}^*) = \begin{cases} \mathbf{x}' & = & \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{y} & = & \mathbf{g}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{x}(0) & = & \mathbf{x}^* \end{cases}$$

Introduction

- Parametric Differential Algebraic (PDA) Model

$$\Sigma = \Sigma(\mathbf{x}, \mathbf{y}, u; \boldsymbol{\mu}, \mathbf{x}^*) = \begin{cases} \mathbf{x}' & = & \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{y} & = & \mathbf{g}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{x}(0) & = & \mathbf{x}^* \end{cases}$$

This is a system of ODE's where

$\mathbf{f} = (f_1, \ldots, f_n)$ and $\mathbf{g} = (g_1, \ldots, g_m)$

are rational functions in $(\mathbf{x}, \boldsymbol{\mu}, u)$

$\mathbf{x}, \mathbf{y}, u$ are differential variables

$\boldsymbol{\mu}, \mathbf{x}^*$ are numerical variables

and $\mathbf{x}', \mathbf{y}', u'$ denotes differenting w.r.t. time t.

Introduction

- Parametric Differential Algebraic (PDA) Model

$$\Sigma = \Sigma(\mathbf{x}, \mathbf{y}, u; \boldsymbol{\mu}, \mathbf{x}^*) = \begin{cases} \mathbf{x}' & = & \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{y} & = & \mathbf{g}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{x}(0) & = & \mathbf{x}^* \end{cases}$$

(Interpretation of variables)

state variables $\mathbf{x} = (x_1, \ldots, x_n)$

output variables $\mathbf{y} = (y_1, \ldots, y_m)$

input variable u (a scalar)

system parameters $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_t)$

initial values $\mathbf{x}^* = (x_n^*, \ldots, x_n^*)$

Introduction

- Parametric Differential Algebraic (PDA) Model

$$\Sigma = \Sigma(\mathbf{x}, \mathbf{y}, u; \boldsymbol{\mu}, \mathbf{x}^*) = \begin{cases} \mathbf{x}' & = & \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{y} & = & \mathbf{g}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{x}(0) & = & \mathbf{x}^* \end{cases}$$

(Intuitively)

states $\mathbf{x}$ are hidden

output $\mathbf{y}$ are observable

parameters $\boldsymbol{\mu}, \mathbf{x}^*$ are unknown

Introduction

- Recap:

  Σ is system of ODE's

  with parameters and initial values

  (collectively, $\boldsymbol{\theta} := (\boldsymbol{\mu}, \mathbf{x}^*)$)

Σ is a mathematical model of some dynamical physical or bio-chemical phenomenon Ω which the scientist wants to study.

Say the scientists is interested in estimating some parameters $\boldsymbol{\theta}^\ell \subseteq \boldsymbol{\theta}$ in Σ.

It is by no means clear that Σ is "correct" or "suitable".

(There is another topic of model selection which we will not address)

Introduction

- The Problem of Parameter Estimation

The scientist takes 4 steps:

(1) Sets up a PDA Model:
$$\Sigma = \begin{cases} \mathbf{x}' & = & \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{y} & = & \mathbf{g}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{x}(0) & = & \mathbf{x}^* \end{cases}$$

(2) Chooses a subset of parameters

$\boldsymbol{\theta}^\ell \subseteq \boldsymbol{\theta} := (\boldsymbol{\mu}, \mathbf{x}^*)$ of interest

(3) Measures the output $\mathbf{y}_\Omega$ of the physical system Ω.

(4) Uses $\mathbf{y}_\Omega$ and Σ to estimate the values of $\boldsymbol{\theta}^\ell$

Reality: Steps (3) and (4) are messy!

Case Study in Parameter Estimation

- [Miao et al]:

> “Modeling and Estimation of Kinematic Parameters and Replicative Fitness of HIV-1 from Flow-Cytometry-Based Growth Competition”

,

Bulletin of Mathematical Biology (2008) 70: 1749–1771.

Desktop flow cytometer for cell sorting

Case Study in Parameter Estimation

- introduced a nonlinear ODE for the growth of HIV virus,

- investigated identifiability of the ODE system,

- estimated the parameters with wet lab experiments and simulation.

Case Study in Parameter Estimation

- Two types of viruses:

  wild (W) and mutant (M)

- Four cell populations:

  uninfected (T),

  infected by wild virus (T_w)

  infected by mutant virus (T_m)

  infected by both (T_{mw}).

Case Study in Parameter Estimation

- The ODE's

$$\left.\begin{aligned} T' &= (\rho - k_m T_m - k_w T_w - k_R T_{mw})T \\ (T_m)' &= (\rho_m + k_m T - q_m T_w)T_m + 0.25 k_R T_{mw} T \\ (T_w)' &= (\rho_w + k_w T - q_w T_m)T_w + 0.25 k_R T_{mw} T \\ (T_{mw})' &= (\rho_{mw} + 0.5 k_R)T_{mw} + (q_m + q_w)T_w T_m. \end{aligned}\right\} \quad (1)$$

with 11 parameters

$$\left.\begin{array}{l} \rho, \rho_m, \rho_w, \rho_{mw}, \\ k_R, k_m, k_w, k_{mw}, \\ q_m, q_w, q_{mw} \end{array}\right\}$$

Case Study in Parameter Estimation

- Measurements

  In each experiment, the variables T, T_m, T_w, T_{mw}

  are measured at 5 time instances

  ($t = 70, 94, 115, 139, 163 \quad hours$).

  Experiment repeated 3 times

  Total of $60 = 3 \times 20$ measurements (over 21 days)

  Let $\mathbf{y}_{wet}$ denote these wetlab measurements

Case Study in Parameter Estimation

The measurements:

Table 2 Measured number of uninfected and infected cells at five time points with three replicates at each time point

Replication	Time (hour)	T (cell)	T_m (cell)	T_w (cell)	T_{mw} (cell)	T (control group)
1	70	32,554,830	134,173	26,180	9,818	28,088,137
	94	46,645,200	481,950	103,950	18,900	46,578,042
	115	64,240,540	1,230,460	309,260	26,320	74,280,528
	139	65,563,680	9,863,280	3,000,480	1,364,580	151,063,920
	163	36,366,400	36,545,600	10,281,600	28,806,400	351,958,208
2	70	35,855,330	158,620	25,235	10,815	–
	94	48,652,100	269,500	73,500	9,800	–
	115	62,989,640	1,081,920	302,680	25,760	–
	139	79,088,100	7,907,900	3,103,100	900,900	–
	163	47,349,120	24,613,680	15,167,880	22,069,320	–
3	70	32,597,373	98,175	22,908	6,545	–
	94	52,059,000	315,000	110,250	15,750	–
	115	62,362,300	847,210	445,900	38,220	–
	139	77,218,680	5,576,620	2,117,920	478,240	–
	163	49,714,560	17,922,240	17,025,120	16,138,080	–

Case Study in Parameter Estimation

- Simulation

  For each choice of parameter value $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$:

  (1) Simulate the (non-parametric) ODE model with $\widehat{\boldsymbol{\theta}}$

  (2) This produces 20 "measurements" $\mathbf{y}_{sim} = \mathbf{y}_{sim}(\widehat{\boldsymbol{\theta}})$

  (20= 5 measurement times $\times$ 4 variables)

  (3) Compute the fitness of $\widehat{\boldsymbol{\theta}}$

  (i.e., least squares difference between $\mathbf{y}_{wet}$ and $\mathbf{y}_{sim}$)

Case Study in Parameter Estimation

Model fitting for each variable: $T = (solid)$, $T_m = (dotted)$,

$T_w = (dashed)$, $T_{mw} = (dashdot)$

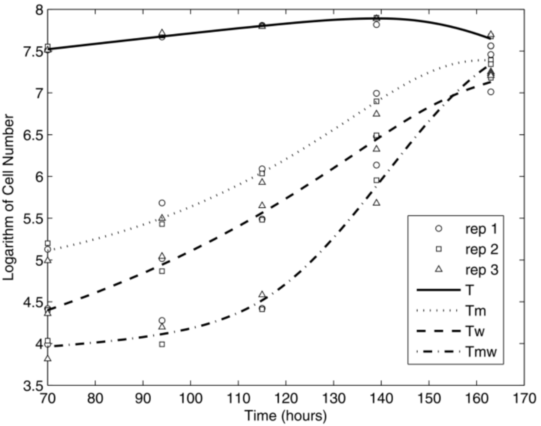


Fig. 1 Experimental data and model fitting results (○ replicate 1, □ replicate 2, and △ replicate 3) and fitted curves for T (solid), T_m (dotted), T_w (dashed) and T_{mw} (dash-dotted).

Case Study in Parameter Estimation

- Search in Parameter Space

  Each parameter $\boldsymbol{\theta}_i \in \boldsymbol{\theta}^\ell$ is searched in some range

  (unclear how range is determined)

  The values $\widehat{\boldsymbol{\theta}}_i$ in the range is picked by

  the Differential Evolution (DE) method

  The value $\widehat{\boldsymbol{\theta}}_i$ with best fitness is picked.

Case Study in Parameter Estimation

For each θ_i, the value with best fitness is "parameter estimate":

Table 3 Parameter estimation results for the 4D dual infection model (18)

Parameter	Search region lower bound	Search region upper bound	Parameter estimate	Bootstrap 95% confidence interval
ρ (per hour)	−6.0e-02	6.0e-02	1.50e-02	1.29e-02, 1.71e-02
ρ_m (per hour)	−2.0e-01	6.0e-02	−2.29e-02	−4.78e-02, 7.90e-03
ρ_w (per hour)	−6.0e-02	6.0e-02	7.13e-03	−2.96e-02, 4.41e-02
ρ_{mw} (per hour)	−2.0e-01	6.0e-02	5.68e-04	−3.94e-02, 1.83e-02
k_m (per cell per hour)	0	1.0e-08	1.51e-09	9.88e-10, 1.89e-09
k_w (per cell per hour)	0	1.0e-08	1.11e-09	4.01e-10, 1.78e-09
k_R (per cell per hour)	0	1.0e-08	4.36e-10	2.94e-23, 2.00e-09
q_m (per cell per hour)	0	1.0e-08	4.15e-09	2.22e-09, 5.98e-09
q_w (per cell per hour)	0	1.0e-08	1.10e-09	2.87e-11, 2.68e-09

Case Study in Parameter Estimation

- Conclusion

  Wetlab measurements are expensive, sparse and noisy

  "multimillion dollar wetlabs supported by NIH"

  Computational measurements are relatively cheaper

- Motivation of this work:

  How do we know if the parameters $\boldsymbol{\theta}^{\ell}$ are unique determined by "perfect" measurements $\mathbf{y}$?

  We want an algorithm to decide this for a given $(\Sigma, \boldsymbol{\theta}^{\ell})$

  The algorithm should give a positive answer *before* we proceed to wetlab experiments to get $\mathbf{y}_{wet}$ or simulate to get $\mathbf{y}_{sim}$...

II. Global Identifiability

"Eventually, the topic [...of proving non-zeroness...] takes over the whole subject [...of Transcendental Number Theory...]"

— David Masser (2000)

The Problem of Identifiability

- Recall our PDA Model:

$$\Sigma = \Sigma(\mathbf{x}, \mathbf{y}, u; \boldsymbol{\mu}, \mathbf{x}^*) = \begin{cases} \mathbf{x}' & = \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{y} & = \mathbf{g}(\mathbf{x}, \boldsymbol{\mu}, u) \\ \mathbf{x}(0) & = \mathbf{x}^* \end{cases}$$

  - Let $\boldsymbol{\theta} := (\boldsymbol{\mu}, \mathbf{x}^*)$

    be called (numerical) parameters of Σ.

We wants to estimate the values of some $\boldsymbol{\theta}^{\ell} \subseteq \boldsymbol{\theta}$

The Problem of Identifiability

Informally, $\boldsymbol{\theta}^\ell$ is "identifiable" if the (perfect) output $\mathbf{y}$ uniquely determines the value of $\boldsymbol{\theta}^\ell$.

- 2 versions: locally or globally identifiable

The Problem of Identifiability

- Example (predator-prey model again):

$$\Sigma = \begin{cases} x_1' & = \theta_1 x_1 - \theta_2 x_1 x_2 \quad \text{(prey)} \\ x_2' & = -\theta_3 x_2 + \theta_4 x_1 x_2 \quad \text{(predator)} \\ y_1 & = x_1 \\ x_1(0) & = \theta_5 \\ x_2(0) & = \theta_6 \end{cases}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_6)$

Notice: no input (u) and only observe prey population (x_1) 

Henceforth, omit the input u from our models, for simplicity

The Problem of Identifiability

- Example (predator-prey model again):

$$\Sigma = \begin{cases} x_1' = \theta_1 x_1 - \theta_2 x_1 x_2 & \text{(prey)} \\ x_2' = -\theta_3 x_2 + \theta_4 x_1 x_2 & \text{(predator)} \\ y_1 = x_1 \\ x_1(0) = \theta_5 \\ x_2(0) = \theta_6 \end{cases}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_6)$

Suppose the scientist wants to estimate $\boldsymbol{\theta}^{\ell} = (\theta_2, \theta_4)$

("how the predator-prey interaction affect the 2 populations")

Is he on a fool's errant?

The Problem of Identifiability

- Example (predator-prey model again):

$$\Sigma = \begin{cases} x_1' &= \theta_1 x_1 - \theta_2 x_1 x_2 \quad \text{(prey)} \\ x_2' &= -\theta_3 x_2 + \theta_4 x_1 x_2 \quad \text{(predator)} \\ y_1 &= x_1 \\ x_1(0) &= \theta_5 \\ x_2(0) &= \theta_6 \end{cases}$$

Suppose the scientist wants to estimate $\boldsymbol{\theta}^\ell = (\theta_2, \theta_4)$

Yes, he is on a fool's errant!
We can prove that θ_2 is not (even locally) identifiable, but θ_4 is (globally) identifiable

✔

Definition of Identifiability

- Henceforth, fix a PDA model

$$\Sigma_0 = \Sigma_0(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta})$$

- A set of parameters $\widehat{\boldsymbol{\theta}} \in \mathbb{C}^s$ is admissible if $\Sigma_0(\widehat{\boldsymbol{\theta}})$ is well-defined

(i.e., denominators of $\mathbf{f}, \mathbf{g}$ do not vanish)

- Let Ω_0 denote the set of admissible parameters.
- Basic theorem in ODE says:

for all admissible $\widehat{\boldsymbol{\theta}}$,

the system $\Sigma_0(\widehat{\boldsymbol{\theta}})$ has a unique solution.

- Let $(\mathbf{x}, \mathbf{y}) = (X_0(\widehat{\boldsymbol{\theta}}), Y_0(\widehat{\boldsymbol{\theta}}))$ be the unique solution of $\Sigma_0(\widehat{\boldsymbol{\theta}})$

Definition of Identifiability

- Where do the solutions come from?
  - Let $\mathbb{C}^\infty(0)$ denote the set of functions

    $$f : \mathbb{C} \to \mathbb{C}$$

    that is complex analytic at $t = 0$.
  - Solutions comes from $\mathbb{C}^\infty(0)$:

    i.e., the differential variables x_i, y_j take values in $\mathbb{C}^\infty(0)$.
  - In particular,

    $$\boxed{(X_0(\widehat{\theta}), Y_0(\widehat{\theta})) \in (\mathbb{C}^\infty(0))^{n+m}}$$

Definition of Identifiability

- Partition the set Ω_0 into output equivalence classes, where

$$\boxed{\widehat{\boldsymbol{\theta}} \equiv \widetilde{\boldsymbol{\theta}} \quad \text{iff} \quad Y_0(\widehat{\boldsymbol{\theta}}) = Y_0(\widetilde{\boldsymbol{\theta}})}$$

- Let $[\widehat{\boldsymbol{\theta}}] \subseteq \Omega_0$ denote the output equivalence class of $\widehat{\boldsymbol{\theta}}$.
- Let $\theta \in \boldsymbol{\theta}$ be any parameter.

  If $W \subseteq \mathbb{C}^s$, let $\Pi_\theta(W)$ denote

  the projection of W onto the θ-coordinate.

  So $\Pi_\theta(W) \subseteq \mathbb{C}$.

Definition of Identifiability

- Definition of Identifiability
  - A parameter $\theta \in \boldsymbol{\theta}$ is globally identifiable if,

    there is a Zariski-open set U such

    for all output equivalence classes $E \subseteq \Omega_0$,

    the size of the set $\Pi_\theta(E \cap U)$ is one.
  - In above definition, if the size is finite for all E, then

    we say θ is locally identifiable.

  - Clearly, global identifiability implies local identifiability.
  - A set $\boldsymbol{\theta}^\ell$ of parameters is X identifiable if each element of the set is X identifiable.

Necessity of U in definition of Identifiability

- Let $U = V \times W \subseteq \mathbb{C}^s \times \mathbb{C}^\infty(0)$

  We say U is Zariski open if

  $V := \{P(\boldsymbol{\mu}) \neq 0\}$ and $W := \{Q(\boldsymbol{\mu}, u) \neq 0\}$

  where P (Q) is a polynomial (differential polynomial).

- Necessity for V: Let $\Sigma_1 : \{x' = \mu x, y = x, x(0) = x^*\}$.

  Then μ is globally identifiable unless $x^* \neq 0$. Thus we may choose $V = \{x^* \neq 0\}$.

- Necessity for W: Let $\Sigma_2 : \{y = \mu u\}$.

  Then μ is globally identifiable provided $u \neq 0$. So choose $W = \{u \neq 0\}$.

Historical Notes

- The identifiability problem as we formulated is this:

  Given $\Sigma(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta})$ and $\boldsymbol{\theta}^{\ell} \subseteq \boldsymbol{\theta}$,

  is $\boldsymbol{\theta}^{\ell} \subseteq \boldsymbol{\theta}$ globally/locally identifiable?

  - Goes back to the 1950s
    - Goodwin Oscillator (1965)
    - Close connection to Control Theory

    (is closely related to observability)

    (can be reduced to controllability)

Historical Notes

- Evolving of identifiability

• Early History mixed parameter estimation with the structural issue problem.

• Bellman and Åström (1970): structural identifiability

(mixture of decision problem with parameter estimation

via loss function $V(x)$, cf. the least squares above)

• The differential algebraic formulation is formulated by Diop and Fliess (1991).

• Divergence or equivalence of analytic concept versus algebraic concept

Historical Notes

- Trends:
  - (1) Analytic to algebraic formulations

    Practitioners prefer analytic, algorithms people prefer algebraic
  - (2) From to empirical to algorithmic solutions

    Old papers gave techniques (e.g., Laplace Transforms),

    not "algorithms" in the modern sense

Historical Notes

- These trends are related:
  - We start with analytic formulations (intuitive).
  - We need algebraic formulations for algorithms
  - Algebraic characterizations usually lead to decidability, but not necessarily practical solution.
  - To get a truly practical method, we need randomization.
- State of Art Algorithms:

  Several general techniques are available and implemented, but for global identifiablity, they lack completeness for PDA models.

III. Our Results

Overview of Our Work

- State of Art in Local Identifiability

[] Hermann, R., Krener, A. J., 1977. Nonlinear controllability and observability. IEEE Trans. on automatic control 22 (5), 728–740.

[] Karlsson, J., Anguelova, M., Jirstrand, M., 2012. An efficient method for structural identifiability analysis of large dynamic systems. IFAC Proc. Vol. 45 (16), 941–946.

[] Sedoglavic, A., 2002. A probabilistic algorithm to test local algebraic observability in polynomial time. JSC 33 (5), 735–755.

[] Chis, O.-T., Banga, J. R., Balsa-Canto, E., 11 2011. PLOS ONE 6 (11), 1–16.

Overview of Our Work

- Issues we will address:

  - We consider the local identifiability problem solved satisfactorily by Sedoglavic (2002). The global problem remains unsolved.

  - Completely (rigorous) algorithms need the algebraic theory

  - Algebraic formulations of the problem often not connected to the analytic one

    ("parallel theory")

Overview of Our Work

- Results in Our Paper
  - THEORY: We begin with the analytic formulation of global identifiability, then give an effective algebraic characterization

    This implies a deterministic algorithm
  - EFFICIENT ALGORITHM: Further reduce the algebraic characterization into a rigorous and efficient randomized algorithm for global identifiability
  - EXPERIMENTAL WORK: Implemented and show its effectiveness on non-trivial examples in the literature.

    Found errors in some published claims of global identifiability.

Algebraic Preparation

- Differential Ring R:
  - R is a ring with a differentiation map, $(\cdot)' : R \to R$:

    $$(a+b)' = a' + b', \qquad (ab)' = a'b + ab'$$

  - An ideal $I \subseteq R$ is a differential ideal if

    $a \in I$ implies $a' \in I$.

  - Extend the ring $\mathbb{C}[z_1, \dots, z_m]$ into the differential ring $\mathbb{C}\{z_1, \dots, z_m\}$ with infinitely many variables

    $$z_i, z_i', z_i'', \dots, z_i^{(q)}, \dots \qquad (\text{for } i \in [m]).$$

    such that $(z_i^{(q)})' = z_i^{(q+1)}$

  - For $A, B \subseteq R$, the saturation of A at B

    $A : B^\infty := \{r \in R : (\exists b \in B, n \in \mathbb{N})[b^n r \in A]\}$.

Algebraic Preparation

- Method of Prolongation

  Given: $x' = f = f(x, \theta)$,

  $y = g = g(x, \theta)$

  - Keep differentiating y, and plug in x' to get new equations:

$$y' = x' \frac{\partial g(x,\theta)}{\partial x} = x' g_x = f g_x.$$

$$y'' = x'[f g_x]_x = f[f_x g_x + f g_{xx}] = f f_x g_x + f^2 g_{xx}.$$

$$y''' = x'[f f_x g_x + f^2 g_{xx}] = \ldots$$

Main issue: how to halt and draw conclusions.

Reducing Differential Equations into Algebraic Ones

- 4 elements in this technique:

(I) View the differential variables as power series.

(II) Prolongate to get new differential equations.

(III) Evaluate differential equations at $t = 0$ to get algebraic equations.

(IV) STOPPING: Naively, when the algebraic equations becomes square, solve for parameters. But there is a trick (which tripped up previous papers).

Reducing Differential Equations into Algebraic Ones

- Example: Identify the parameters μ, x^* in $\boxed{\Sigma}$

  where $m = n = t = 1$

We have 3 differential variables, x, y, u with power series:

$$
\begin{aligned}
x &= x_0 + x_1 t + \frac{x_2}{2!} t^2 + \cdots \\
y &= y_0 + y_1 t + \frac{y_2}{2!} t^2 + \cdots \\
u &= u_0 + u_1 t + \frac{u_2}{2!} t^2 + \cdots
\end{aligned}
$$

View the x_i, y_i, u_i $(i \geq 0)$ as algebraic variables.

Treat y, u as known, i.e., y_i, u_i's are known.

Reducing Differential Equations into Algebraic Ones

- Initially, evaluate $\boxed{\Sigma}$ at $t = 0$:

$$\left.\begin{array}{lcll} x'(0) & : & x_1 & = f(x_0, \mu, u_0) \\ y(0) & : & y_0 & = g(x_0, \mu, u_0) \\ x(0) & : & x_0 & = x^*. \end{array}\right\} \qquad \text{(S1)}$$

This (S1) is an algebraic system with 3 equations and 4 unknowns x_0, x_1, μ, x^*.

Reducing Differential Equations into Algebraic Ones

- Since there are fewer equations than unknowns, we prolongate:

$$\begin{aligned} y &= g(x, \mu, u) \\ y' &= x' g_x + u' g_u \\ &= f g_x + u' g_u \end{aligned}$$

where $g_x = \frac{\partial g}{\partial x}$ and $g_u = \frac{\partial g}{\partial u}$ with g_x, g_u, f are functions in x, μ, u.

Then evaluate y' at $t = 0$:

$$y'(0) : y_1 = x_1 g_x(x_0, \mu, u_0) + u_1 g_u(x_0, \mu, u_0). \qquad \text{(S2)}$$

Now (S1) and (S2) is square: it has 4 equations in 4 unknowns.

Reducing Differential Equations into Algebraic Ones

- Solving gives one of 4 possible outcomes:

  (0) no solutions

  (1) unique solution

  (2) finitely many solutions

  (3) infinitely many solutions

We are tempted to conclude that Σ is (respectively)

(0) inconsistent,

(1) globally identifiable,

(2) locally identifiable,

(3) non-locally identifiable.

Reducing Differential Equations into Algebraic Ones

- We shall exclude (0) by general position assumption.

But are the possibilities (1), (2), (3) correct?

I.e., could further prolongations change the answer?

E.g., Consider the system $S_1 = \{x^2 - 2\}$ and its extension $S_2 = \{x^2 - 2, y^2x + 1\}$.

S_1 has 2 solutions $\{\sqrt{2}, -\sqrt{2}\}$ but S_2 has one $\left\{(-\sqrt{2}, \frac{1}{\sqrt[4]{2}})\right\}$.

E.g., Consider $S_3 = \{x + 1\}$ and $S_4 = \{x + 1, y^2 - 2\}$,

S_3 has 1 solution but its extension S_4 has two.

Reducing Differential Equations into Algebraic Ones

- We show: it suffices to prolongate once more:

$$
\begin{aligned}
y'' &= (x'g_x)' + (u'g_u)' \\
&= (x''g_x + x'(x'g_{xx} + u'g_{xu})) + (u''g_u + u'(x'g_{ux} + u'g_{uu})) \\
&= x''g_x + x'^2 g_{xx} + 2x'u'g_{xu} + u''g_u + u'^2 g_{uu}.
\end{aligned}
$$

Evaluating at $t = 0$ again:

$$
y''(0) : y_2 = x_2 g_x + x_1^2 g_{xx} + 2x_1 u_1 g_{xu} + u_2 g_u + u_1^2 g_{uu}. \qquad (1)
$$

We get a new algebraic unknown x_2, and a system of 5 equations in 5 unknowns.
Its solution (in 4 possibilities) is guaranteed to be stable.

The Results

- Recall $\Sigma_0(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta})$

  with $x_i' = f_i(\mathbf{x}, \boldsymbol{\theta})$ and $y_j = g_j(\mathbf{x}, \boldsymbol{\theta})$.

- Let Q be the LCM of the denominators of f_i, g_j so that

  $f_i = F_i/Q$ and $g_j = G_j/Q$ for $i \in [n], j \in [m]$

- $T_0 := \{(Qx_i' - F_i)^{(k)}, (Qy_j - G_j)^{(k)} : i \in [n], j \in [m], k \geq 0\}$

  is a triangular set.

- $J_0 := (T_0) : Q^\infty$ is the saturation of T_0 at Q.

  Then J_0 is a prime ideal.

The Results

- Algebraic Characterization
  - Let $R_0 := \mathbb{C}[\mu]\{\mathbf{x}, \mathbf{y}\}$

    $S_0 := R_0/J_0$ is an integral domain.

    F_0 is quotient field of S_0,

    E_0 subfield of F_0 generated by the image of $\mathbb{C}\{\mathbf{y}\}$

PROPOSITION:

A parameter θ in Σ_0 is locally (resp., globally) identifiable

iff

$[E_0(\theta) : E_0]$ *is finite (is equal to one).*

The Results

- Effective Algebraic Characterization

- PROPOSITION is not effective as E_0 comes

  from quotient field of an infinitely generated algebra S_0.

- Solution: if $\mathbf{h} = (h_1, \ldots, h_m) \in \mathbb{N}^m$, we define the set $S_{\mathbf{h}}$ incrementally into two stages:

  (1) First put $(Qy_j - G_j)^{(k)}$ into $S_{\mathbf{h}}$ for $k \in \{0, \ldots, h_j\}$ and $j \in [m]$.

  (2) While $x_i^{(j)}$ appears in $S_{\mathbf{h}}$ but $(Qx'_k - G_k)^{(j-1)} \notin (S_{\mathbf{h}})$, add it to $S_{\mathbf{h}}$.

- Note that step (2) terminates. Moreover, we define

$$R_{\mathbf{h}}, \qquad J_{\mathbf{h}}$$

analogously to R_0, J_0 but with the orders of $\mathbf{x}, \mathbf{y}$ bounded as in $S_{\mathbf{h}}$.

The Results

- We can bound $h_1 + h_2 + \cdots + h_m \leq |\boldsymbol{\theta}|$,

and to test if any θ is globally identifiable, we check if the zero set of $J_{\mathbf{h}}$ projected to the coordinates $(\theta, \mathbf{y}_{\mathbf{h+1}})$ has a generic fibre of cardinality one.

- This gives a decision procedure based on computing the rank of the Jacobian of $S_{\mathbf{h}}$.

- We take the final step to practical efficiency, by considering a randomized algorithm: Instead of generic fibre, we use a random fibre, and reduce to *checking the consistency of a system of polynomial equations, and inequations*.

The Results

- Empirical Work

• Our software is implemented in Maple, and publicly available from https://github.com/pogudingleb/Global_Identifiability

Input is $(\Sigma, \boldsymbol{\theta}^\ell, p \in (0,1))$ where $\boldsymbol{\theta}^\ell$ are known to be locally identifiable.

It outputs a subset of $\boldsymbol{\theta}^\ell$ representing the globally identifiable parameters. This output is correct with probability p.

• Refer to paper for examples, including comparison to publicly available software (e.g., DAISY).

Using our software, we have discovered errors in published claims about global identifiability.

IV. Conclusion

Conclusion

- Current Work

  – We have started with an analytic formulation of Global Identifiability, and given algebraic characterizations.

  – We gave the first complete algorithm for Global Identifiability.

  – Experimentally, we verified that it is practical and efficient, and improves on known software.

Conclusion

- Future Work

– I think another kind of “local identifiablity” is more useful: how to determine if a parameter is identifiable within a given region of parameter space.

– The truly challenging research is in parameter estimation. This is of primary interest to scientists, but the area is largely without proper theory.

Thanks for Listening!

*"Algebra is generous,
she often gives more than is asked of her."*

— Jean Le Rond D'Alembert (1717-83)

"To Generalize is to be an Idiot. To Particularize is the Alone Distinction of Merit – General Knowledges are those Knowledges that Idiots possess."

— William Blake (1757 – 1827)

Annotations to Sir Joshua Reynolds's Discourses, pp. xvii – xcviii