Defining and Testing Identifiability, Illustrated by a HIV model

H.V.J. Le Meur, CNRS, LAMFA, Univ. de Picardie, 80000 Amiens, France

F. Ollivier, CNRS, LIX, École polytechnique, 91128 Palaiseau Cedex, France

Basic mathematical definitions

Determinant. *Let us recall that the determinant of a matrix is the volume of the parallelepiped defined by its colums (or rows) vectors. The matrix is invertible iff it is non-zero.*

Topology. *Closed sets are sets that contain their border, open sets contain no point of their border. In $\mathbb{R}^n$, closed and bounded sets are compact: any sequence of points admits a converging subsequence.*

Exhaustive summary. *An exhaustive summary of a differential model with s parameters is a function $\rho : \mathbb{R}^s \mapsto \mathbb{R}^q$ such that θ_i is (locally) identifiable iff $\rho(\theta) = \rho(\vartheta)$ implies $\theta_i = \vartheta_i$ (locally).*

Normal form. *An explicit normal form of a differential system Σ is an equivalent system P_i of the form $x_i^{(\alpha_i)} = f_i(x)$, where the f_i do not depend on the main derivatives $x_j^{(\alpha_j)}$ or on derivatives of the main derivatives, e.g. $x_i'' = \sqrt{x_i}$. Implicit normal forms are equations Q that locally define an explicit normal form using the implicit function theorem, e.g. $x_i'' - x_i^2 = 0$.*

Introduction

We question the notion of identifiability and its relation with the actual ability to achieve identification. We show that a boundedness hypothesis on the set of parameter is essential, that amounts in most pratical situation to bounding the magnitude of parameters.

In the formal definition of identifiability, one should be very carefull with the parentheses and the order of quantifiers, mostly concerning the external controls. However, under some good hypotheses, this is unimportant since a single trajectory may stand for all trajectory.

Without controls, a weaker property plays a similar role and is expressed by the non-vanishing of a Wrońskian determinant. Then, eliminating the state variables is enough to get an exhaustive summary and to test identifiability.

This method is close to criteria encountered in Wu *et al* [7] for HIV models that we will consider as an application. Differential algebra also gives rigorous upper bounds on the minimal number of measurement times to test local and global identifiability.

We conclude with references to some computer tools.

Differential algebra

Founded by Joseph Ritt (1893-1951), differential algebra considers differential fields $\mathcal{F}$, *i.e.* fields equipped with a derivation, such as $\mathbb{Q}(t)$ with d/dt, systems of differential polynomial equations, *i.e.* polynomials $\mathcal{F}\{x_1, \ldots, x_n\}$ in some variables x_i and their derivatives $x_i', x_i''', \ldots$ *Characteristic sets* are particular sets of equations, that generalize the notion of *normal form*.

Any set of explicit equations is a characteristic set, defining a *prime differential ideal*. A differential ideal is the set $[\Sigma]$ of all differential equations that are consequences of the initial system σ and prime means that $PQ \in [\Sigma]$ implies $P \in \Sigma$ or $Q \in \Sigma$.

A differential Zariski open set is defined by an inequation $P(x) \neq 0$, where P is a differential polynomial.

An abstract definition of identifiability

In practice, a structure or parametric model is most of the time given by explicit differential equations:

$$x_i' = f_i(x, u, \theta, t), \quad 1 \le i \le n,$$

where the f_i are rational functions, t is the time, satisfying $t' = 1$, the θ_j, $1 \le j \le s$ are constant parameters, so $\theta_j' = 0$, the u_k, $1 \le k \le m$ are control functions, and the x_i, $1 \le i \le n$ are the state variables.

We need to complete this differential system with initial conditions that are rarely discussed:

$$x_i(0) = c_i.$$

We also need to define outputs:

$$y_\ell = g_\ell(x, \theta, t), \quad 1 \le \ell \le r,$$

where the g_ℓ are also rational functions. Let $X[c, \theta, u](t)$ be the unique solution defined by the control u, the parameters θ and the initial condition c. The differential Zariski open set U is such that the functions f and g are defined on U.

Definition 1. — *A function H of the parameters θ is algebraically identifiable if there exists a Zariski open set $V \subset U$ such that $\forall (c, \theta, u) \in V$, $\forall (\hat{c}, \hat{\theta}, u) \in U$, $g(X[c, \theta, u]) = g(X[\hat{c}, \hat{\theta}, u])$ implies $H(\hat{\theta}) = H(\theta)$, or equivalently if there exists a differential rational function G such that $H(\theta) = G(u, y, t)$. It is locally identifiable if there exists a differential polynomial P of order 0 in $H(\theta)$ such that $P(H(\theta), u, y, t) = 0$.*

Algebraic identifiability means that θ_i is a differential rational function depending on the derivatives of the controls u and the outputs y. Local identifiability means that it is an algebraic function, so that the value is only locally unique.

One may notice that our definition explicitly involves the initial condition c which is too often omitted in the literature.

Identifiability and identification

Assuming the mathematical model describes perfectly the actual behavior and the noise is zero, is identifiability the guarantee to achieve identification ? And if so, what are the relations between the abstract mathematical identifiability and the potential succes of practical identification processes?

Consider the system corresponding to the following *universal equation*. Assuming that the state x is measured, it is identifiable (we will see why below).

$$x' = \frac{bd}{1 + d^2 - \cos(b(t-a))} \cos\left(e^t\right) \qquad (1)$$

$$x(-a) = c, \qquad (2)$$

where a, b, c and $d > 0$ are real numbers.

Boshernitzan [1] has shown the following result.

Theorem 2. — *For any continuous function $F(t)$ defined on any compact interval E, for any $\epsilon > 0$ there exist a, b, c, d such that $\forall t \in E$ $|X((a, b, c, d), t) - F| < \epsilon$.*

This property implies that any (continuous) function may be approximated by four parameters in a very specific model (1). In presence of noise, any perturbation would give very different parameters, which is possible because the set of parameters is unbounded.

The technical details in the next theorem may be skept. Its meaning is that the value θ_i is unique in any bounded subset, when the precision ϵ is small enough. Other values may exist but for which the size of parameters go to infinity.

Theorem 3. — *If a model $M(\theta)$ is such that θ_i is globally identifiable, for any compact set of parameters $E \subset \mathbb{R}^s$ let the set $I(\theta, \epsilon, T, E) :=$*

$$\{\theta^* \in E \mid \forall 0 \le t \le T \ \|g(X(\theta^*, t), \theta^*) - g(X(\theta, t), \theta)\| < \epsilon\},$$

be the set of parameters θ^* for which the observable field is close to a given observed field up to ϵ. Let $\pi_i(\theta) = \theta_i$, then $\forall T > 0$, $\bigcap_{\epsilon > 0} \pi_i I(\theta, \epsilon, T, E) = \theta_i$.

The proof of the theorem relies on the topological property that characterises compact sets: any sequence admits a converging subsequence.

Irreducible systems

Hong *et al.*[3, prop. 1] have noticed that we have an equivalent definition of identifiability, where the unique value of θ is deduced from the knowledge of the input output behavior, *i.e.* the function that associate an output y to all possible input of control functions u. This important property seems paradoxical as, in practice, one knows during the experiment only one vector of inputs and one vector of outputs.

In fact, in most cases, we can do as if we measured the output functions for all inputs u *and for all initial conditions* c. We can give here a precise meaning to this genericity.

Definition 4. — *A system Σ generating a prime differential ideal $\mathcal{P}$ is irreducible if there exists a Zariski open set U such that for any $(c, \theta, u) \in U$ $P(X[c, \theta, u]) = 0$ implies $P \in \mathcal{P}$. Such solutions are called generic solutions of the ideal $\mathcal{P}$.*

One may prove that a system is irreducible if it admits no rational first integral.

One may eliminate state variables by computing a characteristic set for the differential ideal associated to our initial system, with an ordering such that the state variables x_i and all their derivatives are greater than the y_j. It contains a set of differential equations $Q_j(y, u, \theta, t)$. We may chose to extend the order on derivatives to an order on monomials and assume that the coefficient of the main monomial in Q_j is 1. Then let $\rho_{j,\ell}(\theta)$, $1 \le \ell \le p_j$ denote the coefficients of the p_j monomials in Q_j.

Theorem 5. — *If the system is irreducible, then the $\rho_{j,\ell}$ are algebraically identifiable.*

The main idea of the proof is easy: if some $\rho_{j,\ell}$ is not identifiable, then there exists $\vartheta \neq \theta$ such that $P := Q_j(y, u, \theta, t) - Q_j(y, u, \vartheta, t) \notin \mathcal{P}$ but $P(X[c, \theta, u]) \in \mathcal{P}$. In other words, the independence of monomials implies the equality of their coefficients.

Controlability is a sufficient condition of irreducibility that is easy to check. But what can we do when there is no control?

Using the Wrońskian

We borrow the following HIV model to Wu *et al.* where it is assumed that only V is measured.

$$\begin{cases} T'(t) = \lambda - \rho T(t) - \beta T(t) V(t) & , T(0) = T_0 \\ U'(t) = +\beta T(t) V(t) - \delta U & , U(0) = U_0 \\ V'(t) = N\delta U(t) - c V(t) & , V(0) = V_0. \end{cases} \qquad (3)$$

Eliminating T and U, one gets an equation of minimal order in V alone.

$$P(,V) := VV^{(3)}(t) - (V' - \rho V - \beta V^2)(V'' + (\delta + c)V' + \delta c V) - N\lambda\delta\beta V + \delta c V' + (\delta + c)V'' = 0. \qquad (4)$$

We see that only $N\lambda$, $\delta + c$ and δc appear in this system. So only the vector of new parameters $\theta := (\rho, \beta, \kappa := N\lambda\delta, \mu := \delta + c, \nu := \delta c)$ can be possibly identifiable. Their local identifiability is deduced by Wu *et al.* from the non-vanishing of the Jacobian determinant:

$$\left| \frac{\partial P^k}{\partial \theta_j} | 0 \le k \le 4, 1 \le j \le 5 \right|,$$

using the implicit function theorem.

Besides the main monomial VV''', equation (4) contains the set of 12 monomials $M := \{V'V'', V'^2, VV', V^2V'', V^2V', V^3, VV'', VV', V^2, V'', V', V\}$. By the proof of th 5, if the coefficients $\rho_{i,j}(\theta)$ are not identifiable, there is a non-trivial relation with constant coefficients between these monomials : $P(\theta, V) - P(\vartheta, V) = 0$. The existence of such a relation with V analytic is equivalent to the vanishing of the Wrońskian determinant:

$$W(M) := \left| m^k | m \in M, 0 \le k \le \#M - 1 \right|$$

Theorem 6. — *Let $\mathcal{B}$ is a characteristic set of the prime ideal $\mathcal{P}$ associated to the model Σ for a ranking that eliminates the state and let the Q_i be the elements of $\mathcal{B}$ that do not depend on the state variables x. Let M_i be the set of monomials in Q_i, then for any $1 \le i \le r$, if $W(M_i) \neq 0$, the coefficients $\rho_{i,j}(\theta)$ are identifiable.*

Testing the non-vanishing of the Wrońskian

The expression of the derivatives in the Wrońskian may be huge. To test if the determinant is non-zero modulo the equations of the system, we need to reduce it, producing an increase of the size. An easy way to solve the problem is to replace $x_i(t)$ by power series solution, associated to random integer coefficients and initial conditions. One may also make computations modulo some great prime integer p. This gives a probabilistic answer. If one gets a non-zero evaluation the determinant is non-zero. If the evaluation is 0, we can reduce the probability of failure by repeating the experiment and increasing the size of random integer coefficients.

The computation of power series can be done with a near linear asymptotic complexity algorithm, due to Sedoglavic[2]. Van der Hoeven's algorithm [6] has greater asymptotic complexity but is often better in practice.

The main idea of Sedoglavic's algorithm is to generalise Newton's method. A key ingredient is to be able to compute the derivative of $X[c, \theta]$ with respect to θ: $\partial X / \partial \theta_i$. It is solution to the linearized system

$$\sum_{j=1}^{n} \frac{\partial P_i}{\partial x_j} dx_j, \ dx_j(0) = 0.$$

With such a setting, one may also test local identifiability by computing the jacobian determinant $|\partial X[c, \theta](t_i) / \partial \theta_j|$ using s random observation points t_i as proposed in Wu *et al.* Interval or ball arithmetic can confirm it is non-zero, but the vanishing will remain dubious.

Minimal number of observation points

For an algebraically identifiable system, the θ_i can be expressed as differential rational functions of the outputs y, of order at most $n + s$, that is the number of initial conditions c and parameters θ. This is indeed the order of the system, completed with the equations $\theta_i' = 0$, and so the maximal order of any characteristic set. This suggests that $n + s + 1$ generic observation points may be enough to compute the parameters. But one need to keep in mind that for some system, no identification is possible with noisy data without bounding the magnitude of parameters.

Some implementation

Sedoglavic's method has been implemented in Maple and is available:
http://www.lifl.fr/~sedoglav/Software/ObservabilityTest/

The algorithms in the Maple DifferentialAlgebra Package are also available in the C library BLAD:
https://pro.univ-lille.fr/francois-boulier/logiciels/blad/?print=484

A Sage interface BMI is also available:
https://pro.univ-lille.fr/francois-boulier/logiciels/bmi/?print=484

The free computer algebra system Mathemagix provides efficient implementations of up to date fast algorithms for exact and approximate computations, including power series solutions of ODEs and ball arithmetics.
http://www.mathemagix.org/www/mmdoc/doc/html/main/index.en.html

An experimental package allows to compute a linearized system in a form that allows numerical integration in Maple:
http://www.lix.polytechnique.fr/~ollivier/AD_Web/D_ODE_tools/D_ODE_tools.mpl
http://www.lix.polytechnique.fr/~ollivier/AD_Web/D_ODE_tools/Test_D_ODE_tools.html

References

[1] Michael Boshernitzan, "Universal formulae and universal differential equations", *Ann. of Math.*, Second Series, Vol. 124, No. 2 (Sep., 1986), pp. 273-291.

[2] Alin Bostan, Frédéric Chyzak, François Ollivier, Éric Schost, Bruno Salvy and Alexandre Sedoglavic, "Fast computation of power series solutions of systems of differential equations", *Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms*, 1012–1021, La Nouvelle-Orléans, 2007.

[3] H. Hong, A. Ovchinnikov, G. Pogudin and C. Yap, *Global Identifiability of Differential Models*, arXiv:1801.08112, 2018.

[4] J.F. Ritt, *Differential Algebra*, Amer. Math. Soc., 1950.

[5] F. Ollivier and A. Sedoglavic, "Algorithmes efficaces pour tester l'identifiabilité locale", Conférence Internationale Francophone d'Automatique, IEEE, 811-816, 2002.

[6] Joris van der Hoeven, "Relax, but Don't be Too Lazy", *J. Symbolic Comput.*, Volume 34, Issue 6, 479–542, 2002.

[7] H. Wu, H. Zhu, H. Miao and A.S. Perelson, "Parameter identifiability and estimation of HIV/AIDS dynamics models", *Bull. Math. Biol.*, 70, 785–799, 2008.