

Identifiabilité et identification : du Calcul Formel au Calcul Numérique?

François Ollivier (CNRS)

GAGE, Centre de Mathématiques

École polytechnique, 91128 Palaiseau CEDEX, France

francois.ollivier@gage.polytechnique.fr

<http://medicis.polytechnique.fr/gage/ollivier.html>

Résumé

Cet article se propose de poser quelques jalons sur un parcours hypothétique allant du calcul formel au calcul numérique, dans le cadre des problèmes de modélisation. Pour ce faire, on rappelle quelques méthodes classiques ou moins classiques pour tester l'identifiabilité théorique d'un modèle paramétré, en essayant d'envisager leur accélération par l'emploi partiel du calcul numérique et dans quelques cas d'en déduire une méthode d'identification.

En passant, on donne quelques recettes de calcul, qui n'ont rien d'originales, mais qui peuvent néanmoins être utile dans un contexte plus large, pour évaluer les dérivées des solutions d'un système par rapports aux paramètres et aux données initiales.

1 Introduction

Si le titre de cette courte note peut paraître ambitieux, son propos est particulièrement modeste. En effet, si le lien entre calcul formel et calcul numérique tarde à s'établir, en dépit de quelques percées isolées (comme les méthodes de dérivation automatique de programmes FORTRAN développées au sein du projet SAFIR), il faut y chercher des difficultés de fond, qui ne se limitent pas à une différence de culture scientifique, par ailleurs bien réelle, entre « algébristes » et « numériciens ».

Une question centrale est celle de la nature du résultat attendu : exact, approché, probabiliste... Ainsi, les méthodes algébriques permettent de conclure de manière certaine à l'identifiabilité d'un système, éventuellement au prix de calculs lourds, tandis que des calculs numériques de rangs de matrices pourront fournir plus vite une réponse, qui méritera discussion. Si un déterminant calculé en flottant est nettement supérieur à l'incertitude du calcul, on pourra le tenir pour non nul. Mais l'obtention de valeurs, même nombreuses, toutes très proches de zéro, ne fournit en général qu'une indication peu fiable.

On peut en retour questionner l'intérêt d'un résultat d'identifiabilité qui n'est que théorique, comme c'est le cas pour un système où les dérivées des sorties par rapport à certains paramètres sont très petites.

Nous allons donc passer en revue brièvement quelques pistes pour accélérer les calculs d'identifiabilité par des méthodes numériques ou tenter de déduire de méthodes formelles d'identifiabilité des procédés d'identification numérique, non par ferme conviction que ce soit là une voie prometteuse, mais par simple soucis de préciser la topographie du sujet. On considérera ainsi le cas linéaire, et quelques systèmes non linéaires particuliers. Même dans les cas où les méthodes proposées sont en principe applicables, elles se heurtent en pratiques à bien des difficultés de calcul, qui commencent seulement à être surmontées.

Ce travail nous a naturellement conduit au problème du calcul des dérivées de la solution d'un système, tant par rapport aux paramètres que par rapport aux conditions initiales. Des formules classiques apportent une réponse immédiate, mais elles sont difficiles à mettre en œuvre. En effet, dans le cas d'un système d'ordre 1 en n fonctions inconnues, il faut pour chaque dérivée intégrer un système d'ordre 1 en $2n$ fonctions inconnues, ce qui devient lourd si le système de départ était déjà non trivial.

En guise de conclusion, nous avons remarqué que le calcul brutal d'un développement limité pouvait fournir, en particulier à proximité d'un cycle limite une solution très acceptable, sur un intervalle de temps prolongé. On obtient ainsi rapidement une approximation de la solution, polynomiale par morceaux, et fournissant une expression plus maniable de la dépendance par rapport aux paramètres.

2 Identifiabilité et algèbre différentielle

2.1 Traduction algébrique du problème

On considère un système algébrique différentiel, de la forme $x'_i = F_i(x, u, t, \theta)$, $i = 1, \dots, n$, où les fonctions F_i sont des fractions rationnelles, dépendant du vecteur d'état x , des commandes u et d'un vecteur de paramètres θ . On suppose mesurées m sorties du système $y_i = G_i(x, u, t, \theta)$. Partant d'un corps de base, $k := \mathbf{R}$ ou $k := \mathbf{R}(t)$, si le système est non stationnaire, un tel système d'équations différentielles algébriques, (implicitelement complété des équations $\theta'_i = 0$) définit un idéal différentiel premier, et donc une extension de corps différentiel $K := k\langle x, u, \theta \rangle$. Sous cette forme, bien des questions d'automatique s'expriment par des relations entre extensions de corps différentielles. Par exemple, l'observabilité locale est équivalente au fait que l'extension $K/k\langle y, u, \theta \rangle$ est algébrique, et l'identifiabilité locale que l'extension $k\langle y, u, \theta \rangle/k\langle y, u \rangle$ est algébrique. Cette traduction permet de ramener les tests d'identifiabilité et d'observabilité à des calculs de dépendance algébrique, susceptible en principe d'être menés par le calcul d'un ensemble caractéristique pour un ordre approprié. Pour ne pas encombrer ce texte de détails techniques, ici trop éloignés de notre propos, nous renvoyons le lecteur curieux à Fliess [1], Diop et Fliess [1], Glad et Ljung [1], ou encore Ollivier [1].

Il suffit de dire qu'un ensemble caractéristique pour un ordre tel que $x \gg \theta \gg y, u$ contient une relation algébrique de la forme $P_i(\theta_1, \dots, \theta_r, u, y)$ pour tout paramètre θ_i si le système est identifiable. Cette propriété, due à Glad et Ljung est assez paradoxale. En effet, si le calcul explicite des polynômes P_i devient vite coûteux, ceux-ci sont en général de peu d'utilité pour l'identification des θ_i . Car, si les P_i ne dépendent pas des dérivées des θ_i , supposés constants, ils sont en général d'un ordre élevé en y , minoré par $\lfloor r/m \rfloor$, si r est le nombre de paramètres et m le nombre de sorties. Or les sorties étant bruitées, il est inconcevable d'espérer évaluer mieux que leur dérivée première. D'autre part, un théorème dû à Joseph Ritt, le fondateur de l'algèbre différentielle permet d'affirmer que tout système est génériquement identifiable, en ce sens qu'il

peut devenir identifiable en modifiant les coefficients de la fraction rationnelle exprimant l'une des sorties de manière arbitrairement petite.

Il est assez troublant de devoir mener des calculs d'autant plus lourds que le système est plus compliqué et que l'identifiabilité semble plus probable. Mais n'est-il pas hasardeux de raisonner en termes probabilistes sur des systèmes décrivant des structures physiques, chimiques ou biologiques et donc fort peu arbitraires?

2.2 Passage à l'algèbre linéaire

On peut conserver l'approche différentielle algébrique de l'identifiabilité, tout en essayant d'alléger les calculs, de diverses manières. On peut par exemple s'inspirer de la méthode de Glad et Ljung, sans toutefois calculer explicitement les P_i . Pour ce faire, on sait calculer l'expression formelle de la dérivée $y_i^{(j)}(0)$, en fonction des paramètres littéraux θ_i , des conditions initiales littérales $x_i(0) = c_i$, et des dérivées formelles $u_i^{(j)}$.

Considérons pour simplifier le cas d'une sortie unique. Génériquement, la matrice jacobienne $(\partial y^{(s)} / \partial x_j | \partial y^{(s)} / \partial \theta_\ell)$ avec $s = 0, \dots, n+r-1$, $j = 1, \dots, n$ et $\ell = 1, \dots, r$ a un déterminant non identiquement nul. Cette propriété suffisante d'identifiabilité n'est pas nécessaire. Le résultat le plus général est donné par l'énoncé suivant, qui est une simple variante du théorème des fonctions implicites.

THÉORÈME 1. — *Un système est globalement identifiablessi les matrices jacobiennes $N := (\partial y_i^{(s)} / \partial x_j)$ et $M := (\partial y_i^{(s)} / \partial x_j | \partial y_i^{(s)} / \partial \theta_\ell)$ (avec $s = 0, \dots, n+r-1$, $j = 1, \dots, n$, $i = 1, \dots, m$ et $\ell = 1, \dots, r$) sont telles que $\text{rg}M = \text{rg}N + r$, où r est le nombre de paramètres.*

Une approche consistant à calculer brutalement des déterminants de matrices formelles est cependant rapidement vouée à l'échec. Le simple calcul des dérivées formelles jusqu'à un ordre suffisant est même un excellent moyen de saturer la mémoire. Il faudra donc le plus souvent se limiter à des évaluations numériques, en limitant au maximum la mémoire utilisée par les calculs formels lors des dérivations. Le travail réalisé dans le cadre du projet SAFIR pour la dérivation de programme FORTRAN fournit une source immédiate d'inspiration, mais il ne semble pas exister d'implantation de ces techniques au sein des systèmes de calcul formel disponibles (cf. Schmidt [1]).

Notons que si l'on effectue un nombre suffisant de calculs numériques exacts, un résultat de Heintz et Schnorr [1] permet d'affirmer qu'un déterminant est identiquement nul, et qu'on peut obtenir une réponse probabiliste très fiable avec un nombre d'essai plus réduit.

Cette modification de l'approche permet de gagner en efficacité en se ramenant à l'algèbre linéaire. Il a fallu pour ce faire renoncer à exprimer des relations de dépendances, au demeurant inutilisables. Cette transposition nous dispense en partie de supposer que les fonctions F_i sont des fractions rationnelles, si l'on est près à se satisfaire de réponses incertaines données par des calculs approchés en flottants de fonctions transcyclantes quelconques.

2.3 Intégrales premières et symétries

Il existe quelques cas particuliers où l'on peut conclure à l'identifiabilité d'un système avec très peu ou même aucun calcul. La situation la plus simple est celle où toutes les variables d'état sont observées. Dans un tel cas, les coefficients des fractions rationnelles F_i exprimées sous forme réduite sont automatiquement structurellement identifiables, sauf si le système admet une intégrale première polynomiale. Cette situation est impossible, si le système est commandable.

Considérons un système sans intégrale première polynomiale, mais où toutes les variables d'état ne sont pas observées. Supposons que le système est observable. Tester l'observabilité est en général plus aisé, car dans la plupart des modèles, il y a nettement moins de variables d'état que de paramètres. Admettons enfin qu'on ait une expression rationnelle $H(\theta, y, u)$ explicite de l'état en fonction des paramètres et des dérivées des commandes et des sorties. Dans une telle situation, si le système n'est pas identifiable, il existe une loi de groupe, s'exprimant par des formules rationnelles, agissant sur les solutions du système et laissant les sorties invariantes. En effet, soient θ_1 et θ_2 deux valeurs possibles du vecteur de paramètres, pour un même comportement entrée-sortie : en prenant $X := H(\theta_2, G(x, u, t, \theta_1), u)$, on obtient une solution du système $X'_i = F_i(X, u, t, \theta_2)$.

Ceci conduit à une méthode de test consistant à résoudre le système

$$\forall u, x \quad G^{(s)}(x, u, t, \theta_1) = G^{(s)}(H(\theta_2, G(x, u, t, \theta_1), u), u, t, \theta_2), \quad s = 0, \dots, n - 1,$$

de manière à vérifier s'il admet pour θ_1 fixé une infinité de solutions θ_2 . On reconnaît là une variante d'une méthode due à Kalmann [1] dans le cas linéaire et à Vajda *et al.* [1] dans un cadre plus général. Bien que l'expression ci-dessus ne semble guère engageante, on obtient souvent des systèmes particulièrement simples à résoudre, surtout si l'on intervient « manuellement » pour guider les calculs !

On peut bien sûr, comme on l'a fait pour la méthode de Glad et Ljung, « linéariser » cette approche. Ceci peut être fait brutalement, sans hypothèses particulières sur le système, en transformant notre système par son « linéarisé tangent », obtenu à partir de l'opérateur de Kähler :

$$\begin{aligned} dx' &= \frac{\partial F}{\partial x_i} dx_i + \frac{\partial F}{\partial u_j} du_j + \frac{\partial F}{\partial \theta_\ell} d\theta_\ell \\ dy &= \frac{\partial G}{\partial x_i} dx_i + \frac{\partial G}{\partial u_j} du_j + \frac{\partial G}{\partial \theta_\ell} d\theta_\ell \end{aligned}$$

Pour tester l'identifiabilité locale, il suffit de montrer que ce système, complété par les équations $dy_j = 0$ implique $d\theta_\ell = 0$ pour tout ℓ . On peut mener les calculs en utilisant le package Maple Diffalg de François Boulier, permettant de construire l'extension de corps différentiel K/k . Toutefois, il sera bien plus rapide de se contenter d'un calcul numérique obtenu à partir d'un développement en série, à partir de conditions initiales prises au hasard, avec bien sûr tous les inconvénients évoqués ci-dessus.

3 Systèmes linéaires et polynomiaux

Après avoir évoqué brièvement la manière dont le recours au calcul numérique pouvait accélérer le test d'identifiabilité, on va envisager le problème de l'identification dans quelques cas particuliers. Prenons le cas simple d'un système linéaire stationnaire $X' = A(\theta)X + BU$, $Y = C(\theta)X$. On peut par un calcul simple de modèle différentiel éliminer l'état et calculer un système différentiel linéaire $y_i^{(r_i)} = L_i(y, u, \theta)$, où les L_i sont linéaires en y, u et leurs dérivées ont des coefficients rationnels en θ . Soit $C_p(\theta)$ l'ensemble de ces coefficients. On sait (cf. *e.g.* Ollivier [1], chap. 5) que le système est identifiablessi chaque θ_ℓ est algébrique sur le corps engendré par les fractions $C_p(\theta)$. *A priori*, si ce système fournit un test d'identifiabilité, on peut craindre qu'il ne soit pas d'un grand secours pour identifier, en raison des dérivées d'ordre élevé qu'il contient. Toutefois, dans le cas linéaire stationnaire où nous nous trouvons, il est envisageable, au moins en théorie, d'intégrer les équations un nombre de fois suffisante pour remplacer toute les dérivées par des intégrales

Supposons réalisée une telle opération avec une précision suffisante. Nous obtenons des valeurs approchées des coefficients C_P , à partir desquelles on peut espérer obtenir une ou plusieurs estimations des paramètres θ . Toutefois, il nous faut prendre garde au fait qu'un tel système est en général surdéterminé, de sorte qu'une modification même minime de la valeur des paramètres fait qu'il n'y a plus de solution du tout. Quelques travaux récents apportent toutefois des possibilités de résolution pour ce type de problème, fréquent en pratique. On pourra se reporter à une conférence de Dedieu, disponible en vidéo sur le Web (<http://www.msri.org/>), et intitulée de manière provocatrice « Solving algebraic systems with no solution », ou aux articles de Giusti et Schost [1] ou Bondyfalat *et. al.* [1]. Une telle approche est encore en principe envisageable si l'on considère un système où les F_i sont des polynômes et où tout l'état est observé. Il est en revanche impossible de la généraliser, sauf cas très particuliers, pour des systèmes faisant intervenir des fractions, ce qui est le cas général lorsqu'on élimine l'état.

4 Utiliser la méthode de Newton ?

On a évoqué ci-dessus l'usage du linéarisé tangent pour tester l'identifiabilité locale. Le système linéaire tangent permet également d'évaluer, par intégration numérique, les dérivées des fonctions d'état par rapport aux paramètres et aux conditions initiales. Considérons en effet le système en $2n$ fonctions inconnues

$$\begin{aligned} \frac{dx'_i}{dt} &= F_i(x, u, t, \theta) \\ &= \frac{\partial F_i}{\partial x_i} dx_i + \frac{\partial F_i}{\partial u_j} du_j + \frac{\partial F_i}{\partial \theta_\ell} d\theta_\ell. \end{aligned}$$

Si on l'intègre en prenant $d\theta_j = 0$, pour tout j , $dx_i(0) = 1$ et $dx_j(0) = 0$ pour $j \neq i$, on obtient des solutions $dx_j(t) = \partial x_j(t)/\partial c_i$ où $x_i(0) = c_i$ est la condition initiale pour la i^e fonction inconnue. Si l'on choisit $dx_i = 0$ pour tout i , $d\theta_\ell(0) = 1$ et $d\theta_q(0) = 0$ pour $q \neq \ell$, on obtient des solutions $dx_j(t) = \partial x_j(t)/\partial c_\ell$.

Cette méthode permet donc en principe d'avoir une évaluation numérique des dérivées des fonctions de sortie par rapports aux paramètres et aux conditions initiales. À proximité des valeurs correspondant à la sortie mesurée, il est donc envisageable de converger vers les valeurs réelles par une méthode de type Newton. Toutefois, sans indication précise sur de bonnes valeurs d'initialisation, on n'obtiendra en général qu'un minimum local. Des essais multiples entraînent une grande lourdeur de calcul, puisqu'il faut intégrer à chaque étape $n + r$ systèmes d'ordre 2n. Notons que cette méthode classique s'itère, pour obtenir des dérivées d'ordre supérieur, elle permet par exemple d'obtenir les formules de Poincaré pour le problème des centres. Toutefois, la lourdeur des opérations dissuade d'envisager un développement en série de la solution par rapport aux conditions initiales.

On peut toutefois en déduire des indications utiles, comme la dérivée de la période d'un cycle limite par rapport à un paramètre ou une condition initiale. Pour des systèmes simples possédant des solutions exactes, la précision des valeurs numériques des dérivées est tout à fait satisfaisante.

5 En guise de conclusion

Au cours de mes tentatives de calcul, j'ai été conduit à manipuler des solutions en séries formelles pour certains systèmes biologiques possédant des cycles limites, par exemple un modèle

de la production de la protéine PER proposé par Goldbeter. J'ai remarqué que si, à distance du cycle, ces séries, médiocrement convergentes, ne donnaient des résultats fiables que pour des temps de l'ordre de la demi-heure, on avait en revanche sur le cycle lui-même des solutions précises jusqu'à 6 ou 8 heures de sorte que trois ou quatre polynômes suffisaient à représenter une variable d'état sur la période circadienne. (Ceci implique en contrepartie que l'identification numérique est particulièrement difficile sur le cycle.) Faute d'avoir eu le temps de réaliser une étude systématique, je me garderai d'affirmer qu'il s'agit d'un fait général, toutefois, lorsqu'une telle situation se présente, on peut en tirer parti de bien des manières.

D'une part, certaines opérations que l'on est naturellement amené à faire de manière répétée, telles que le tracé de courbe, s'en trouvent notablement accélérées. D'autre part, même dans les cas où il faut se contenter d'un pas d'intégration d'une taille moins spectaculaire, l'intégration numérique à partir d'un développement formel offre peut-être dans certains cas une méthode compétitive, qui a l'avantage de permettre un calcul direct des dérivées par rapport aux paramètres, par simple dérivation des coefficients de la série.

6 Références bibliographiques

BOULIER (François),

[1] *Étude et implantation de quelques algorithmes en algèbre différentielle*, thèse de l'Université des Sciences et Techniques de Lille, juin 1994.

BONDYFALAT (Didier), MOURRAIN (Bernard) et PAN (Victor Y.),

« Controlled iterative methods for solving polynomial systems », actes de ISSAC'98, ACM Press, 1998.

BOULIER (François), LAZARD (Daniel), OLLIVIER (François) et PETITOT (Michel),

[1] « Representation for the radical of a finitely generated differential ideal », dans les actes de ISSAC'95, Montréal, Québec, 1995, A.H.M. Levelt éd., ACM Press, New York, ISBN : 0-89791-699-9, p. 158-166.

DIOP (Sette) et FLIESS (Michel),

« On nonlinear observability », actes de la 1^e Conférence Européenne d'Automatique, Hermès, 1991.

GLAD (S.T.) et LJUNG (L.),

[1] « Parametrization of Non-linear Model Structures as Linear Regressions », dans les actes de IFAC World Congress, Tallin, août 1990.

FLIESS (Michel),

[1] « Automatique et corps différentiels », Forum Math., **1** 227–238, 1989.

GIUSTI (Marc) et SCHOST (Éric),

« Solving some overdetermined polynomial systems », preprint soumis à ISSAC'99, 1999.

HEINTZ (Joos) et SCHNORR (C.P.),

[1] « Testing polynomials which are easy to compute », dans les actes de *12th Ann. ACM Symp. on Computing*, 1980, p. 262-268 ; reproduit dans *Logic and Algorithmic. An International Symposium held in Honour of Ernst Specker*, Monographie no 30 de l'Enseignement de Mathématique, Genève, 1982, p. 237-254.

KALMAN (R.E.), FALB (P.L.) et ARBIB (M.A.),

- [1] *Linear optimal control systems*, McGraw-Hill, New-York, 1969.
- OLLIVIER (François),
- [1] *Le problème de l'identifiabilité structurelle globale : étude théorique, méthodes effectives et bornes de complexité*, thèse de l'École polytechnique, juin 1990.
- RITT (Joseph Fels),
- [1] *Differential Algebra*, Amer. Math. Soc. Colloq. Publ., vol. 33, A.M.S., New-York, 1950.
- SCHMIDT (Nicole),
- [1] *Définition automatique : application à un problème d'optimisation en météorologie*, thèse de doctorat en science de l'université de Nice, 1993.
- VAJDA (Sandor), GODFREY (Keith R.) et RABITZ (Herschel),
- [1] « Similarity transformation approach to identifiability analysis of nonlinear compartmental models », *Mathematical Biosciences*, **93**, 217–248, Elsevier, 1989.
- WALTER (éric),
- [1] *Identifiability of State Space Models*, Lecture Notes in Biomathematics no 46, Springer-Verlag, Berlin et New-York, 1982.