Solution de systèmes polynomiaux par homotopie

Éric Schost, LIX, École polytechnique

Eric.Schost[at]polytechnique.fr

Vue d'ensemble

Le but de ce projet est de "résoudre" des systèmes de n équations polynomiales de degré 2 à n inconnues (un tel système a généralement 2^n solutions). Il s'agit de mettre en place un cas particulier simplifié de ce qu'on appelle *méthodes par homotopie*.

Résoudre un système $F = F_1, \ldots, F_n$ dans $k[X_1, \ldots, X_n]$ (k est un corps pour l'instant quelconque) reviendra à calculer des polynômes à une variable $P, V_2, \ldots, V_n$ tels que l'ensemble des solutions de $F = 0$ soit décrit par

$$\left\{ \begin{array}{rcl} P(X_1) & = & 0 \\ X_2 & = & V_2(X_1) \\ & \vdots & \\ X_n & = & V_n(X_1); \end{array} \right.$$

on conçoit que sous cette forme, on accède aisément à la plupart des caractéristiques de l'ensemble des solutions (par exemple, trouver les racines de P dans k nous donne les solutions à coordonnées dans k, etc ...). Tous les systèmes n'admettent pas des solutions de cette forme. Cependant, ce sera le cas pour les systèmes qui nous intéresseront : on choisira pour $F_1, \ldots, F_n$ des polynômes "génériques" (tous les coefficients sont aléatoires) de degré 2.

Pour calculer la résolution, on commence par définir un système G de référence. L'objectif est de pouvoir résoudre G "à vue" ; on prend donc

$$G_1 = X_1^2 - X_2, G_2 = X_2^2 - X_3, \quad \ldots, \quad G_{n-1} = X_{n-1}^2 - X_n, G_n = X_n^2 - 1.$$

Il est très facile de résoudre ce système sur le modèle ci-dessus ; on obtient

$$\left\{ \begin{array}{rcl} X_1^{2^n} & = & 1 \\ X_2 & = & X_1^2 \\ & \vdots & \\ X_n & = & X_1^{2^{n-1}}. \end{array} \right.$$

Pour utiliser ce système de référence, on définit une nouvelle variable ε, et on transforme le système F, en introduisant F^ε donné par

$$F_1^\varepsilon = \varepsilon F_1 + (1-\varepsilon) G_1, \quad \ldots, \quad F_n^\varepsilon = \varepsilon F_n + (1-\varepsilon) G_n,$$

En faisant $\varepsilon = 1$, on retrouve F ; en faisant $\varepsilon = 0$, on retrouve G. Ceci mène à l'algorithme suivant :

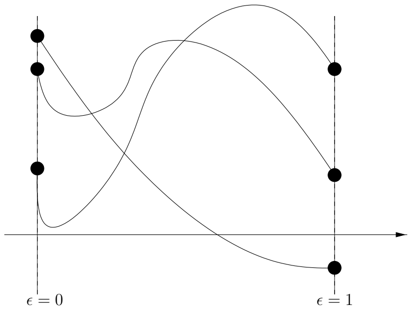
- On résout le système en $\varepsilon = 0$: on a vu que c'est immédiat.

- On en déduit une solution "paramétrée par ε" du système F^ε. Cette solution s'écrit

$$\begin{cases} P^\varepsilon(X_1) &= 0 \\ X_2 &= V_2^\varepsilon(X_1) \\ \vdots & \\ X_n &= V_n^\varepsilon(X_1), \end{cases}$$

où $P^\varepsilon, V_2^\varepsilon, \ldots, V_n^\varepsilon$ sont des polynômes en une variable à coefficients dans $k(\varepsilon)$, c'est-à-dire des fractions rationnelles en ε (voir l'exemple à la fin du sujet).
- On fait $\varepsilon = 1$ dans la solution paramétrée.

Graphiquement, on cherche à déterminer les points de droite, à partir des points de gauche, en "calculant" une description des courbes qui les relient.



C'est bien entendu la seconde étape qui pose toutes les difficultés, et sur laquelle porte l'essentiel du projet.

Calcul de la solution paramétrée

Avant de donner les détails de l'algorithme, on présente rapidement les structures et les outils dont on aura besoin.

Séries formelles sur k. Une série formelle à coefficients dans k est une "somme formelle" de la forme $\sum_{i \geq 0} a_i \varepsilon^i$; on note l'ensemble de ces séries sous la forme $k[[\varepsilon]]$.

Ces séries se comportent comme des polynômes pour les opérations d'addition et de multiplication. Concrètement, on manipule uniquement des séries tronquées, c'est-à-dire avec un nombre fini de termes. On introduit donc une notion de *précision* : la série $a = \sum_{i=0}^{n-1} a_i \varepsilon^i$ sera dite de précision n. Quand on ajoute ou on multiplie deux séries tronquées, on effectue l'opération comme pour des polynômes, puis on tronque le résultat à la précision courante.

Une différence importante entre séries et polynômes est la possibilité d'inverser une série formelle *dont le terme constant n'est pas nul.* Par exemple, l'inverse de la série $1 - \varepsilon$ est la série $\sum_{i \geq 0} \varepsilon^i$. La formule suivante donne l'inverse de la série $a = \sum_{i \geq 0} a_i \varepsilon^i$: on pose $b_0 = 1/a_0$ et

$$b_{i+1} = 2b_i - ab_i^2 \bmod \varepsilon^{2^{i+1}}$$

pour $i \geq 0$. Alors b_i est l'inverse de a modulo ε^{2^i}.

Approximation de Padé. Le second outil dont on a besoin est le calcul d'*approximants de Padé.* Si $f = p/q$ est une fraction rationnelle dans $k(\varepsilon)$, le calcul d'inverse ci-dessus permet de calculer le développement de Taylor de f en 0. L'approximation de Padé permet de faire l'inverse : si on connaît le développement de Taylor de f avec suffisamment de termes ($\deg p + \deg q + 1$, en l'occurence), on peut retrouver p

et q (à une constante près). L'algorithme est simple à implanter, c'est une variante de l'algorithme d'Euclide, il est donné au chapitre 5 de *Modern Computer Algebra*, J. von zur Gathen et J. Gerhard, Cambridge University Press, 1999.

Solutions en séries. On fait désormais l'hypothèse que le polynôme $X_1^{2^n} - 1$ a toutes ses racines dans k (on verra plus bas des situations où c'est le cas). Dans ce cas, il est possible d'expliciter toutes les solutions du système G. À partir de ces solutions, on va calculer des solutions sous forme de *séries formelles en* ε pour le système F^ε.

Comme pour l'inverse, on travaille itérativement. On se donne une solution initiale $x^0 = (x_1^0, \ldots, x_n^0)$ pour $F^\varepsilon \mod \varepsilon$, c'est à dire G. On peut alors calculer la suite $x^i = (x_1^i, \ldots, x_n^i)$, où chaque x_j^i est une série à précision 2^i, et telle que $F^\varepsilon(x^i) = 0$ modulo ε^{2^i}. On vérifie facilement que la suite définie par

$$x^{i+1} = x^i - J(x^i)^{-1} F^\varepsilon(x^i) \mod \varepsilon^{2^{i+1}}$$

convient. Ici, J est la *matrice jacobienne* du sysème F^ε par rapport à $X_1, \ldots, X_n$, $J(x^i)$ est son évaluation en x^i, et $J(x^i)^{-1}$ l'inverse de celle-ci. Remarquer qu'il s'agit ici d'une version formelle de l'opérateur de Newton, qui généralise ce qu'on a écrit pour l'inverse d'une série.

Solutions paramétrées. On a maintenant tous les outils nécessaires. La première chose à faire est de calculer toutes les solutions de G ; c'est immédiat, en utilisant l'hypothèse qu'on connaît toutes les racines de $X_1^{2^n} - 1$. Il y a 2^n solutions.

En utilisant l'opérateur de Newton, on calcule 2^n solutions en séries formelles de F^ε, en allant jusqu'à la précision $2^{n+1}+1$. On va noter $x^{(1)}, \ldots, x^{(2^n)}$ ces solutions, où $x^{(i)} = (x_1^{(i)}, \ldots, x_n^{(i)})$.

On peut d'ores et déjà en déduire le polynôme P^ε. On calcule pour cela le polynôme dont les racines sont les $x_1^{(i)}$. Il est facile de voir que ce polynôme coïncide "presque" avec P^ε : ses coefficients sont les développements de Taylor à précision $2^{n+1}+1$ de ceux de P^ε. Par ailleurs, on peut montrer que les coefficients de P^ε ont des numérateurs et dénominateurs de degrés au plus 2^n. Donc pour retrouver P^ε, on calcule un approximant de Padé de chacun des coefficients, en prenant comme degré maximal pour les numérateurs et dénominateurs 2^n.

Un léger problème se pose pour les polynômes V^ε : leurs coefficients (dans $k(\varepsilon)$) ont des degrés de l'ordre de 4^n, ce qui est beaucoup trop. Pour contourner cette difficulté, on définit $W_j^\varepsilon = (V_j^\varepsilon \times (P^\varepsilon)') \mod P^\varepsilon$, qui vit aussi dans $k(\varepsilon)[X_1]$, et on montre que ce choix est le bon : les coefficients de W_j^ε ont un degré au plus en 2^n. Par ailleurs, on voit facilement que $W_j^\varepsilon(X_1)$ vaut

$$\sum_{i \le 2^n} x_j^{(i)} \prod_{\ell \ne i} (X_1 - x_j^{(\ell)}), \tag{1}$$

si on calcule "à précision infinie". On commence donc par calculer cette expression avec des coefficients dans $k[[\varepsilon]]$ à précision $2^{n+1}+1$, et on en déduit les coefficients de W_j^ε par approximation de Padé en degré 2^n. On fait alors $\varepsilon = 0$ dans W_j^ε pour obtenir un W_j, et on retrouve V_j par la formule $W_j = (V_j \times P') \mod \mathrm{P}$ (ce dernier calcul nécessite un pgcd étendu).

Racines de l'unité dans k. Ce qui précède demande d'avoir certaines racines de l'unité dans k. Un choix naturel est alors de travailler dans un corps fini de la forme $\mathbb{Z}/p\mathbb{Z}$, avec p un premier bien choisi. En effet, si p est un premier de la forme $\ell 2^n + 1$, on peut montrer que $X^{2^n} - 1$ a toutes ses racines dans k.

Par exemple, on pourra prendre $p = 29 \cdot 2^{57} + 1$ et $\omega = 21$. Alors, ω est une racine 2^{57}-ième de 1 dans $\mathbb{Z}/p\mathbb{Z}$. Dans ce cas, pour tout $n \leq 57$, on pose $\tau = \omega^{2^{57-n}}$; les racines de $X^{2^n} - 1$ sont les puissances $1, \tau, \ldots, \tau^{2^n-1}$, et elles sont distinctes.

Exemple

On choisit $n = 2$, on travaille modulo 9001, et le système à résoudre est

$$F_1 = 8996X_2^2 + 66X_2X_1 + 99X_2 + 77X_1^2 + 54X_1 + 8940 = 0,$$

$$F_2 = 31X_2^2 + 8989X_2X_1 + 8975X_2 + 8951X_1^2 + 8983X_1 + 8939 = 0.$$

Modulo 9001, on les a 4 racines quatrièmes de l'unité $1, 9000, 1237, 7764$. On en déduit que les solutions de $G : \{X_1^2 - X_2, X_2^2 - 1\}$ sont

$$(1,1), (9000,1), (1237,9000), (7764,9000).$$

En partant de $(1,1)$, on obtient successivement par l'opérateur de Newton les solutions de F^ε en "séries tronquées", jusqu'à la précision 9 :

$$(1 + 6670\varepsilon, 1 + 4569\varepsilon)$$

$$(1 + 6670\varepsilon + 4674\varepsilon^2 + 8487\varepsilon^3, 1 + 4569\varepsilon + 4028\varepsilon^2 + 1959\varepsilon^3)$$

$$\begin{gathered}(1 + 6670\varepsilon + 4674\varepsilon^2 + 8487\varepsilon^3 + 6058\varepsilon^4 + 4716\varepsilon^5 + 1854\varepsilon^6 + 1415\varepsilon^7,\\ 1 + 4569\varepsilon + 4028\varepsilon^2 + 1959\varepsilon^3 + 5533\varepsilon^4 + 4596\varepsilon^5 + 5107\varepsilon^6 + 1568\varepsilon^7)\end{gathered}$$

$$\begin{gathered}(1 + 6670\varepsilon + 4674\varepsilon^2 + 8487\varepsilon^3 + 6058\varepsilon^4 + 4716\varepsilon^5 + 1854\varepsilon^6 + 1415\varepsilon^7 + 8903\varepsilon^8,\\ 1 + 4569\varepsilon + 4028\varepsilon^2 + 1959\varepsilon^3 + 5533\varepsilon^4 + 4596\varepsilon^5 + 5107\varepsilon^6 + 1568\varepsilon^7 + 840\varepsilon^8)\end{gathered}$$

On fait de même avec les 3 autres solutions. Ceci permet de reconstituer le polynôme ayant pour racines les coordonnées X_1 de ces solutions, à coefficients à précision 9 :

$$\begin{aligned}&X_1^4 + (96\varepsilon + 2400\varepsilon^2 + 5521\varepsilon^3 + 82\varepsilon^4 + 3518\varepsilon^5 + 3262\varepsilon^6 + 8605\varepsilon^7 + +5763\varepsilon^8)X_1^3\\ &+(8793\varepsilon + 7463\varepsilon^2 + 7042\varepsilon^3 + 6137\varepsilon^4 + 789\varepsilon^5 + 3829\varepsilon^6 + 4951\varepsilon^7 + 5255\varepsilon^8)X_1^2\\ &+(114\varepsilon + 5848\varepsilon^2 + 3466\varepsilon^3 + 8386\varepsilon^4 + 2077\varepsilon^5 + 4952\varepsilon^6 + 364\varepsilon^7 + 4994\varepsilon^8)X_1\\ &+9000 + 321\varepsilon + 5223\varepsilon^2 + 1843\varepsilon^3 + 5697\varepsilon^4 + 4214\varepsilon^5 + 7588\varepsilon^6 + 3976\varepsilon^7 + 2096\varepsilon^8.\end{aligned}$$

On applique alors une approximation de Padé sur chacun des coefficients, pour obtenir P^ε :

$$\begin{aligned}X_1^4 \quad &+\frac{2346\varepsilon^4 + 2660\varepsilon^3 + 3938\varepsilon^2 + 4726\varepsilon}{\varepsilon^4 + 976\varepsilon^3 + 1411\varepsilon^2 + 5186\varepsilon + 2487}X_1^3\\ &+\frac{2916\varepsilon^4 + 8872\varepsilon^3 + 1851\varepsilon^2 + 4762\varepsilon}{\varepsilon x^4 + 976\varepsilon^3 + 1411\varepsilon^2 + 5186\varepsilon + 2487}X_1^2\\ &+\frac{1207\varepsilon^4 + 8180\varepsilon^3 + 4499\varepsilon^2 + 4487\varepsilon}{\varepsilon^4 + 976\varepsilon^3 + 1411\varepsilon^2 + 5186\varepsilon + 2487}X_1\\ &+\frac{4696\varepsilon^4 + 6406\varepsilon^3 + 8269\varepsilon^2 + 1053\varepsilon + 6514}{\varepsilon^4 + 976\varepsilon^3 + 1411\varepsilon^2 + 5186\varepsilon + 2487}\end{aligned}$$

En faisant $\varepsilon = 1$, on retrouve $P = X_1^4 + 7910X_1^3 + 1877X_1^2 + 5851X_1 + 3439$. Pour obtenir W_2, on commence par calculer le développement de Taylor de W_2^ε, en utilisant l'équation (1) ; on trouve

$$\begin{aligned}&(152\varepsilon + 3151\varepsilon^2 + 7932\varepsilon^3 + 2801\varepsilon^4 + 6656\varepsilon^5 + 8728\varepsilon^6 + 667\varepsilon^7 + 5977\varepsilon^8)X_1^3\\ &+(8923\varepsilon + 1644\varepsilon^2 + 2127\varepsilon^3 + 3934\varepsilon^4 + 8030\varepsilon^5 + 4864\varepsilon^6 + 170\varepsilon^7 + 5966\varepsilon^8)X_1^2\\ &+(4 + 8421\varepsilon + 2873\varepsilon^2 + 1813\varepsilon^3 + 1603\varepsilon^4 + 998\varepsilon^5 + 1190\varepsilon^6 + 1118\varepsilon^7 + 3333\varepsilon^8)X_1\\ &+120\varepsilon + 1936\varepsilon^2 + 2120\varepsilon^3 + 5755\varepsilon^4 + 8628\varepsilon^5 + 1888\varepsilon^6 + 4008\varepsilon^7 + 2380\varepsilon^8.\end{aligned}$$

On applique une approximation de Padé aux coefficients pour retrouver W_2^{ε} :

$$\frac{3998\varepsilon^4+8412\varepsilon^3+1851\varepsilon^2+8983\varepsilon}{\varepsilon^4+976\varepsilon^3+1411\varepsilon^2+5186\varepsilon+2487}X_1^3+$$
$$\frac{6475\varepsilon^4+6053\varepsilon^3+2711\varepsilon^2+4036\varepsilon}{\varepsilon^4+976\varepsilon^3+1411\varepsilon^2+5186\varepsilon+2487}X_1^2+$$
$$\frac{8732\varepsilon^4+6768\varepsilon^3+2455\varepsilon^2+442\varepsilon+947}{\varepsilon^4+976\varepsilon^3+1411\varepsilon^2+5186\varepsilon+2487}X_1+$$
$$\frac{693\varepsilon^4+136\varepsilon^3+548\varepsilon^2+1407\varepsilon}{\varepsilon^4+976\varepsilon^3+1411\varepsilon^2+5186\varepsilon+2487}.$$

En faisant $\varepsilon=1$, on trouve $W_2=6951X_1^3+7703X_1^2+6268X_1+8630$; par un calcul de pgcd étendu, on en déduit $V_2=4359X_1^3+2251X_1^2+2523X_1+5498$. Donc, les solutions du système F_1, F_2 sont décrites par :

$$\begin{cases} X_1^4+7910X_1^3+1877X_1^2+5851X_1+3439=0, \\ X_2=4359X_1^3+2251X_1^2+2523X_1+5498. \end{cases}$$

En factorisant le premier polynôme, on trouve par exemple que $(2353, 5333)$ et $(2334, 4888)$ sont solutions du système (F_1, F_2).