Résolution de systèmes polynomiaux

Éric Schost
Eric.Schost@polytechnique.fr

April 8, 2004

Le but de ce projet est de présenter quelques techniques de résolution de systèmes d'équations polynomiales. En ce qui nous concerne, sur un exemple, nous serons intéressés par la transformation d'un ensemble d'équations de la forme

$$x^2 + y^2 + z^2 = 1, \quad x^2 + z^2 = y, \quad x = z$$

en un ensemble d'équations équivalentes de la forme

$$z^4 + \frac{1}{2}z^2 - \frac{1}{4} = 0, \quad x = z, \quad y = -2z^2.$$

La deuxième forme est manifestement plus sympathique : on y lit qu'il y a 4 solutions complexes (dont 2 réelles), et il n'est pas difficile d'en trouver des approximations numériques. Par comparaison, il est bien moins évident de répondre à ces questions sur la première formulation. Le deuxième système est une *base de Gröbner* associée aux équations initiales.

On peut considérer comme second exemple le système

$$54 - 5y + 99z - 61xy - 50x^2 - 12yz,$$

$$1 - 47x - 91xy - 47z^2 - 61yz + 41y^2,$$

$$-86 + 23xy - 84x^2 + 19x^2y - 50xyz + 88xz^2 \ ;$$

un des ses bases de Gröbner est de la forme

$$x - P_x(z), \quad y - P_y(z), \quad Q(z),$$

où P_x, P_y sont des polynômes en z de degré 11, et Q un polyôme en z de degré 12. On voit donc que le système initial a 12 solutions ; pour en trouver des approximations, il suffit de trouver des approximations des racines de Q, et les injecter dans P_x et P_y. Ce projet demande d'écrire de quoi calculer de telles bases de Gröbner.

1 Description

1.1 Rappel : division euclidienne

Une idée sous-jacente à la notion de base de Gröbner est une généralisation de la division euclidienne (que l'on connaît pour les polynômes en une variable) à des situations en plusieurs variables. Si k est un corps, pour tous F et G de $k[X]$, il existe d'uniques quotient et reste Q, R tels que $G = QF + R$, avec $\deg R < \deg F$. L'algorithme de division euclidienne est un processus itératif permettant de calculer le reste R.

Notons $\deg F = D$ et $\deg G = D'$ et supposons que D' est au moins égal à D (sinon, il n'y a rien à faire). La boucle élémentaire de la division euclidienne de G par F consiste à *réécrire* G modulo F de manière à faire disparaître son terme de plus haut degré. Pour cela, on pose la soustraction $G - \mathrm{coeff}(G, D')/\mathrm{coeff}(F, D) X^{D'-D} F$, ce qui fait descendre le degré d'au moins 1. On itère ensuite le procédé jusqu'à ce que le degré descende en dessous de celui de F.

Avec plus de recul, l'idée principale derrière ce procédé consiste à associer aux polynômes F et G des *termes de tête*, à savoir les termes de plus haut degré. La division avec reste de G par F passe alors par une suite de réécritures : tant que le terme de tête de F divise le terme de tête de G, on peut annuler ce dernier par une combinaison linéaire adéquate. La technologie des bases de Gröbner est une extension de ce processus en plusieurs variables.

1.2 Ordres monomiaux et division avec reste

Pour étendre les techniques de division avec reste, il faut pouvoir associer à tout polynôme un *terme de tête*. Pour cela, il faut définir une relation d'ordre sur les monômes. On appelle *ordre monomial* une relation d'ordre $<$ sur les monômes de $k[X_1, \ldots, X_n]$ telle que pour tous monômes n, m, m', $1 < n$ et $m < m'$ implique $nm < nm'$. Plus explicitement, soient X^α et X^β deux monômes, avec $\alpha = (\alpha_1, \ldots, \alpha_n)$ et $\beta = (\beta_1, \ldots, \beta_n)$. Des ordres parmi les plus fréquemment utilisés sont les suivants.

Ordre lexicographique (lex). On dit que $X^\beta < X^\alpha$ pour l'ordre lexicographique si $\alpha_i > \beta_i$ pour le premier indice i tel que $\alpha_i \neq \beta_i$. C'est l'ordre du dictionnaire sur les "mots" α et β.

Ordre du degré, raffiné par l'ordre lexicographique inverse (grevlex). On dit que $X^\beta < X^\alpha$ pour cet ordre si $\sum_i \beta_i < \sum_i \alpha_i$, ou, en cas d'égalité, si $\alpha_i < \beta_i$ pour le dernier indice i tel que $\alpha_i \neq \beta_i$.

La notion d'ordre monomial permet d'associer à tout polynôme F son *terme de tête*, c'est le terme correspondant au monôme le plus grand (attention : dans le terme de tête, on prend en compte le coefficient). On peut dès lors écrire un algorithme de division avec reste pour les polynômes à plusieurs variables, qui généralise celui donné dans le cas d'une variable ; étant donnés des polynômes $F_1, \ldots, F_s$ et G, cet algorithme calcule des quotients $Q_1, \ldots, Q_s$ et un reste R tels que $G = Q_1 F_1 + \cdots + Q_s F_s + R$. Le terme de tête du polynôme F_i est noté L_i, pour $i \leq s$.

1. Poser $Q_1 = \cdots = Q_S = R = 0$
2. **Tant que** $G \neq 0$, **faire**
   (a) Soit L_G le terme de tête de G
   (b) **Si** il existe i tel que L_G est divisible par L_i **faire**
      - $Q_i = Q_i + L_G/L_i$
      - $G = G - (L_G/L_i)F_i$

   (c) **Sinon**
      - $G = G - L_G$
      - $R = R + L_G$

L'exemple suivant montre cependant que sans prendre de précaution supplémentaire, le reste n'est pas unique (il dépend de l'ordre dans lequel on choisit les diviseurs). Dans $\mathbb{Q}[X_1, X_2]$ muni de l'ordre lexicographique, considérons les polynômes $F_1 = \underline{X_1 X_2} + 1, F_2 = \underline{X_2^2} - 1$ et $G = \underline{X_1 X_2^2} - X_1$ (le trait indique le terme de tête). Si on effectue d'abord la réduction par F_2, on obtient 0. Par contre, si on réduit d'abord par F_1, on obtient comme reste $R = -X_2 - X_1$, qui ne se réduit pas à zéro.

1.3 Bases de Gröbner et algorithme de Buchberger

Dans l'exemple précédent, l'unicité du reste a été mise en défaut par l'existence de $R = -X_1 - X_2$ qui peut s'écrire sous la forme $Q_1F_1 + Q_2F_2$, mais dont le terme de tête n'est divisible ni par le terme de tête de F_1 ni par celui de F_2. La définition d'une base de Gröbner vise précisément à éviter ce genre de phénomène.

Définition. *Soit $<$ un ordre monomial sur $k[X_1, \ldots, X_n]$ et $F_1, \ldots, F_r$ des polynômes de $k[X_1, \ldots, X_n]$. Une famille de polynômes $G_1, \ldots, G_s$ forme une* base de Gröbner *pour $F_1, \ldots, F_s$ et pour l'ordre $<$ si*

- *tout G_i s'écrit sous la forme $\sum_j Q_{i,j}F_j$;*
- *le terme de tête de tout élément de la forme $\sum_i Q_iF_i$ est divisible par le terme de tête de l'un des G_i.*

Cette base est réduite *si pour tout i, aucun des monômes de G_i n'est divisible par le terme de tête d'un des G_j, $j \neq i$.*

On montre que pour tous $F_1, \ldots, F_r$, il existe une unique base de Gröbner réduite ; en outre, l'algorithme de division par les polynômes $\{G_1, \ldots, G_s\}$ produit un reste unique, quels que soient les choix effectués : cela permet de généraliser l'algorithme d'Euclide. Enfin, un système d'équations et sa base de Gröbner ont les mêmes ensembles de solutions.

Dans les exemples de l'introduction, la deuxième forme du système est une base de Gröbner pour les équations initiales, pour l'ordre lexicographique. C'est un phénomène général : dans une base lexicographique, on trouve des polynômes en la dernière variable uniquement, en les deux dernières variables, etc ..., ce qui facilite par exemple la résolution numérique, par propagation. Les bases pour l'ordre grevlex n'ont pas cette propriété, mais sont généralement bien plus faciles à calculer.

Soit F un système et G sa base de Gröbner réduite (pour un ordre quelconque). Voici d'autres résultats que l'on peut lire sur G.

- F est inconsistant ssi $G = \{1\}$.
- F a un nombre fini de solutions ssi pour tout i, il existe un polynôme de G dont le terme de tête est $X_i^{\alpha_i}$.
- Si c'est le cas, le nombre de solutions est égal au nombre de monômes qui ne sont divisibles par aucun des termes de tête des éléments de G (attention, on compte ici les solutions avec leurs éventuelles multiplicités).

Il reste à dire comment calculer une base de Gröbner. On passe d'un système $F_1, \ldots, F_r$ à une base de Gröbner par un processus de *complétion* par l'adjonction de *paires critiques*, que l'on appelle S-polynômes. Soient F et G deux polynômes de $k[X_1, \ldots, X_n]$, $aX^\alpha = aX_1^{\alpha_1} \cdots X_n^{\alpha_n}$ et $bX^\beta = bX_1^{\beta_1} \cdots X_n^{\beta_n}$ les termes de tête de F et G. Pour $i = 1, \ldots, n$, soit $\gamma_i = \max(\alpha_1, \beta_i)$, et notons $X^\gamma = X_1^{\gamma_1} \cdots X_n^{\gamma_n}$. On appelle *S-polynôme* de F et G le polynôme

$$\frac{X^\gamma}{aX^\alpha}F - \frac{X^\gamma}{bX^\beta}G.$$

L'idée derrière cette définition est de forcer l'annulation des termes de tête des deux membres par collision.

À l'aide de cette définition, l'*algorithme de Buchberger* consiste à adjoindre au système initial de générateurs tous les S-polynômes que l'on peut en déduire, jusqu'à stabilisation du processus. On montre que cet algorithme calcule une base de Gröbner.

1. Poser $G = F$
2. **Faire**
   (a) Poser $G' = G$.
   (b) **Pour** p, q dans G', **faire**
      - Soit R le reste du S-polynôme de p, q par G'
      - **Si** $R \neq 0$, **faire** $G = G \cup \{R\}$.

   Jusqu'à ce que $G = G'$.

À la sortie, la base G n'est pas réduite : on réduit G en considérant successivement tous les éléments g de G, et en réduisant g par $G - \{g\}$: si le reste r est différent de g, on remplace g par r dans G (si $r = 0$, on supprime g de G). On montre que le processus s'arrête, et calcule une base réduite.

2 Travail demandé

La première étape à réaliser pour ce projet est l'implantation des algorithmes décrits ci-dessus : calcul d'une base de Gröbner réduite, et calcul du nombre de solutions d'un système, si ce nombre de solutions est fini.

Jusqu'ici, on n'a pas spécifié qui était le corps k : concrètement, on pense à des corps finis $\mathbb{Z}/p\mathbb{Z}$ (p premier) ou à $\mathbb{Q}$. Pour pouvoir traiter ces deux cas simultanément (et potentiellement n'importe quel autre corps), on écrira du code générique, en utilisant les mécanismes d'héritage. De même, on s'attachera à écrire un code le plus générique possible vis-à-vis de l'ordre monomial utilisé.

Lorsqu'on travaille sur $\mathbb{Q}$, la taille des coefficients "explose" rapidement : il ne suffira pas de travailler avec des `int` ni des `long`, et on aura besoin d'entiers en précision arbitraire. Une solution est d'utiliser la bibliothèque `BigInteger`.

À titre indicatif, vous pourrez tester votre code sur les systèmes *cycliques* : pour n fixé, on obtient le système cyclique d'ordre n en posant les équations

$$\begin{array}{rcl}
X_1 + X_2 + \cdots + X_n & = & 0, \\
X_1X_2 + X_2X_3 + X_3X_4 + \cdots + X_nX_1 & = & 0, \\
X_1X_2X_3 + X_2X_3X_4 + \cdots + X_nX_1X_2 & = & 0, \\
 & \vdots & \\
X_1X_2\cdots X_{n-1} + \cdots + X_nX_1\cdots X_{n-2} & = & 0, \\
X_1\cdots X_n & = & 1.
\end{array}$$

Jusqu'à quelle valeur de n pouvez vous effectuer les calculs ?

Enfin, pour aller plus loin, on propose les extensions suivantes :

Changement d'ordre : l'expérience montre que les bases pour l'ordre lexicographique sont plus "intéressantes" que celles pour l'ordre grevlex, mais plus longues à calculer. Les techniques de changement d'ordre consistent à calculer une base grevlex pour commencer, puis à en déduire une base lexicographique. L'ingrédient essentiel est de l'algèbre linéaire.

Critères de Buchberger et techniques modulaires : une large part des calculs consiste à réduire des S-polynômes par la base en cours de construction. S'il y a moyen de prédire qu'une réduction donne 0, on peut directement rejeter le S-polynôme en question, et c'est autant de temps de gagné.

Les critères de Buchberger premettent d'effectuer ce type de prédiction. Une autre possibilité, lorsqu'on travaille sur $\mathbb{Q}$, est de commencer par réduire le système modulo un nombre premier p et d'effectuer le calcul modulo p, en gardant trace des réductions à 0.

Isolation des racines réelles : Enfin, lorsqu'on travaille sur $\mathbb{Q}$, dans le cas où le système a un nombre fini de solutions (complexes), on propose d'écrire un algorithme pour approcher les solutions réelles (il existe également des algorithmes pour trouver les racines complexes, mais ils sont plus délicats à mettre en place). L'ingrédient principal est la règle des signes de Descartes pour les polynômes en une variable.

Références. Une référence très abordable est le livre de D. Cox, J. Little, D. O'Shea, *Ideals, varieties and algorithms*, Springer, 1992.

Pour les extensions décrites ci-dessus, on pourra consulter le livre de D. Cox, J. Little, D. O'Shea, *Using algebraic geometry*, Springer, 1998. Avant d'aborder ces extensions, il est vivement conseillé de prendre contact avec l'auteur du sujet pour obtenir davantage d'explications, et un guide plus précis de la littérature.

Enfin, pour vérifier vos résultats, n'hésitez pas à utiliser un logiciel comme Maple, installé dans les salles informatiques de l'X, qui permet d'effectuer tous ces calculs.