Solución de Ecuaciones Polinomiales

M. en C. Andrés A. Galván Navarro [1]

INTRODUCCIÓN

Sea K un campo y $K\,[x_1,...,x_n\,]$ el anillo de polinomios en $x_1,...,x_n$ sobre K. Sí $F_1,...,F_s \in K\,[x_1,...,x_n]$ entonces la variedad definida por $F_1,...,F_s$ en K^n es el conjunto $V\,(F_1,...,F_s\,) = \{(a_1,...,a_n) \,/\, f_I\,(a_1,...,a_n)=0$ para $i = 1, ...,s\,\}$, es decir:

Definición 1. La variedad definida por $F_1,...,F_s$ es el conjunto de soluciones en K^n del sistema de ecuaciones $F_1 = 0, ..., F_s = 0$.

En este trabajo nos proponemos presentar métodos no comunes de solución de sistemas de ecuaciones polinomiales. Para esto, en la Sección 1 definiremos la idea de base de Grobner, que nos permite resolver sistemas de una manera bastante simple y en la Sección 2 usamos técnicas de Álgebra Lineal para obtener dichas soluciones. Como los cálculos que haremos son algo complicados nos auxiliaremos con el paquete computacional Maple; en particular usaremos el comando with (grobner) para acceder a las bases de Grobner, y las demostraciones no presentadas de algúnos teoremas pueden hallarse en los libros de Cox que aparecen en las referencias.

[1] Departamento de Matemáticas y Física. Centro de Ciencias Básicas. U.A.A. Teléfono 910-74-00, Ext. 333. E-mail: agalvan@correo.uaa.mx

SECCIÓN 1. Bases de Grobner y Sistemas de Ecuaciones Polinomiales

Definición 1.1 Por un ideal en $K\,[x_1,...,x_n]$ entendemos un subconjunto A de $K\,[x_1,...,x_n\,]$ tal que:

- Si f y g están en A, $f+g$ está en A.
- Sí f está en $K\,[\,x_1,...,x_n\,]$ y g está en A entonces fg está en A.
- A es no vacío.

Y el ideal generado por $F_1,... ,F_s$ es el conjunto $<F_1,...,F_s> = \{h_1F_1+ ...\, h_sF_s : h_1 ,..., h_s\in K[x_1,...,x_n\,]\,\}$

Ahora, dado el sistema de ecuaciones $F_1 = 0,...,F_s = 0$ a partir de estas ecuaciones si multiplicamos la primera ecuación por h_1, la segunda por h_2 , ..., la s-ésima por h_s y sumamos obtenemos la ecuación $h_1F_1+ ... +h_sF_s = 0$, es decir, la ecuación $h_1F_1 +...+ h_sF_s =0$ es una consecuencia del sistema de ecuaciones $F =,...,F_{s= =} = 0$ y podemos ver el ideal $< F_1,..., F_s>$ como el conjunto de "consecuencias del sistema de ecuaciones dado."

Ahora, dado un ideal I en $K[x_1,...,x_n\,]$ por una base de I, entendemos una colección $\{F_1,...,F_s\,\}$ de polinomios, tales que $I = < F_1,...,F_s >$. Se puede demostrar que todo ideal tiene una base finita.

Teorema 1.2
Si $F_1,...,F_s$ y $G_1,...,G_t$ son elementos de $K\ \ [x_1,...,x_n]$, tales que $<F_1,...,F_s>=<G_1,...,G_t>$ entonces $V\,(F_1,...,F_s) = V(G_1,...,G_t\,)$.

De lo anterior se sigue que a cada ideal I en $K[x_1,...,x_n]$ podemos asociar una variedad, a saber la variedad determinada por una de sus bases.

Definición 1.3 Si $V \subset K^n$ es una variedad, entonces el subconjunto $I(V)=\{g \in K[x_1,...,x_n]$: $g(v)=0$, para todo v que está en $V\}$ es un ideal que se llama el ideal de la variedad V.

De lo anterior se sigue que hay una relación íntima entre variedades e ideales y que es posible asociar a cada variedad un ideal y a cada ideal una variedad.

Para resolver el sistema de ecuaciones $F_1(x) = ...= F_s(x) = 0$ consideremos el ideal generado por $\{F_1,...,F_s\}$ en $K[x]$. Como en $K[x]$ cada ideal es principal existe h en $K[x]$ tal $< F_1,...,F_s > = < h >$, y las soluciones del sistema dado son las soluciones de la ecuación $h = 0$...

Para resolver el sistema de ecuaciones $F_1 = ...= F_s = 0$ donde $F_1,..., F_s \in K[x_1,..., x_n]$ entonces usaremos el ideal $< F_1,..., F_s >$.

¿Pero cómo decidir si $F \in K[x_1,..., x_n]$ está en $<F,…,F>$?

Para responder esta pregunta necesitaremos generalizar el algoritmo de la división en una variable, es decir, necesitaremos un procedimiento que nos permita dividir F por $F_1,...,F_s$. Para esto daremos la siguiente definición:

Definición 1.4 Un orden monomial en $K[x_1,...,x_n]$ es un orden total en los monomios en $K[x_1,...,x_n]$, que es un buen orden y que es compatible con la multiplicación.

Son órdenes monomiales los siguientes:

1. Orden lexicográfico.
   Si $a = (a_1...a_n)$ y $b = (b_1...b_n)$ son elementos de Z^n. Decimos que $a >_{lex} b$ si en el vector $a-b$ el primer elemento no cero más a la izquierda es positivo.
2. Orden Graduado Lexicográfico
   Si $a =(a_1...a_n)$ y $b = (b_1...b_n)$ son elementos de Z^n y $/a/ = \sum a_i$
   Decimos que $a >_{grlex} b$ si y sólo si $/a/>/b/$ ó {si $/a/=/b/$, $a >_{lex} b$ }
3. Orden graduado lexicográfico inverso
   Si $a=(a_1...a_n)$ y $b= (b_1...b_n)$ son elementos de Z^n y $/a/ = \sum a_i$

Decimos que $a >_{grevlex} b$ si y sólo si $/a/>/b/$ ó { si $/a/=/b/$ en el vector $a-b$, el elemento no cero más a la derecha es negativo}.

Usando estos órdenes podemos como en el caso de una variable escribir un polinomio con los términos ordenados de grado mayor a grado menor.

Ejemplo Sea $f = 4xy^2z+4z^2-5x^3+7x^2z^2$ un polinomio en $C[x,y,z]$

Entonces: f escrito ordenadamente según el orden lexicográfico se ve así:

$f = -5x^3+7x^2z^2+4xy^2z+4z^2$.

Y según el *grlex* , así: $f = 7x^2z^2+4xy^2z-5x^3+4z^2$.

El hecho que un orden monomial nos permita escribir un polinomio ordenadamente, también nos permite desarrollar un algoritmo de división como en el caso de polinomios de una variable.

Teorema 1.5

Sean $F,F_1,...,F_s \in K[x_1,..., x_n]$. Supongamos que en $K[x_1,..., x_n]$ tengamos definido un orden monomial. Entonces existen $a_1,...,a_s$, $r \in K[x_1,...,x_n]$ tales que $F= a_1F_1+...+a_sF_s+r$, donde r es una suma de términos $a_\alpha x^\alpha$ tales que $a_\alpha x^\alpha$ es menor que los términos $LT(F_1), ..., LT(F_s)$.

El problema es que el resultado de esta división depende de la manera en que los divisores $F_1,..., F_s$ sean ordenados y si $I = < F_1,..., F_s >$, este algoritmo tampoco nos permite decidir si un polinomio $F \in K[x_1,...,x_n]$ está o no en el ideal. Afortunadamente, entre las muchas bases que tiene un ideal hay una llamada base de Grobner que nos permite mejorar la eficiencia del algoritmo anterior.

Definición 1.6 Sea I en $K[x_1,...,x_n]$ un ideal, por una base de Grobner de I, entendemos un conjunto $G = \{g_1,...,g_s\}$ de elementos de $K[x_1,...,x_n]$ tales que $<LT(g_1)...,LT(g_s)>=<LT(I)>$, para un orden monomial en $K[x_1,...,x_n]$, donde $LT(g_1), ..., LT(g_s)$ son los términos líderes de los elementos $g_1, ...,g_s$ y $LT(I)$ representa el conjunto de los términos líderes de los elementos de I.

Teorema 1.7

(Algoritmo de la División). Dados $F,F_1,...,F_s$ polinomios en $K[x_1,...,x_n]$ con $K[x_1,...,x_n]$ con un orden monomial. Entonces existen $a_1,...,a_s$, y r en $K[x_1...x_n]$, tales que $F = a_1F_1+ ... +a_sF_s+ r$, donde r es el único elemento tal que ninguno de sus términos es divisible por uno, de los términos lideres de los polinomios $F_1, ... ,F_s$

Ahora, usando este resultado somos capaces de desarrollar un algoritmo que nos permite calcular las bases de Grobner de un ideal y que nos lleva a las siguientes:

APLICACIONES

1. Resolver el sistema de ecuaciones lineales:
   $3x – 6y –2z = 0$
   $2x – 4y + 4w = 0$
   $X – 2y – z – w = 0$
   Solución. Consideremos el ideal $I=<3x–6y–2z,2x–4y+4w, x–2y–z–w>$ en $C[x,y,z]$ con el lexorder y hallemos una base de Grobner para él. $G=\{x-2y+2w,z+3w\}$ es tal base y las soluciones del sistema son $x=-2y-2w$, $y= y$, $z= -3w$, $w = w$.
2. Sea $I=<xz-y^2, x^3-z^2>$ en $C[x,y,z]$, sea $F=-4x^2y^2z^2+y^6+3z^5$. Y supongamos que en $C[x,y,z]$ está definido el orden grlex. Decidir si F está en I o no.
   Solución: Hallemos una base de Grobner de I, con respecto al orden lexicográfico. Esta es: $G=\{xz-y^2,x^3-z^2,x^2y^2-z^3, xy^4-z^4, y^6–z^5\}$ y apliquemos el algoritmo de la división. Si dividimos F por G obtenemos que el residuo es cero, por lo tanto F está en el ideal.
3. Resolver el sistema
   $x^2+ y^2+z^2=1, \quad x^2+ z^2 = y , \quad x = y$
   Solución: La base de Grobner del ideal $I=<x^2+y^2+z^2-1,x^2+z^2-y, x-y >$ con respecto al orden lexicográfico es:
   $g_1 =x-z, \quad g_2=-y+2z^2, \quad g_3=z^4-½ z^2 – ¼$ y es fácil hallar la solución del sistema de estas ecuaciones.

SECCIÓN 2. Solución de Ecuaciones Polinomiales y Algebra Lineal

Sea $I=<f_1,...f_s>$ un ideal en $K[x_1,...,x_n]$. Construyamos el anillo cociente $A=K[x_1,...,x_n]/I$. Para esto, sea $G = \{g_1,...,g_t\}$ una base de Grobner de I. en $K[x_1,...,x_n]$ definamos la siguiente relación de equivalencia $f \approx g$ si y sólo si $f-g$ está en I.

Sea $K[x_1,...,x_n]/I$ el conjunto de clases de equivalencia definidas por ≈, y si f es un elemento de $K[x_1,...,x_n]$, la clase de f la denotamos por $[f]$.

Ahora, hagamos $K[x_1,...,x_n]/I$ un anillo, definiendo:

$[f]+[g]=[f+g]$ y $[f]*[g]=[f*g]$. En particular; si $\lambda \in K$, $\lambda*[f]= [\lambda*f]$, entonces $K[x_1,...,x_n]/I$ es también un espacio vectorial sobre K.

Ahora vamos a tratar de resolver un sistema de m ecuaciones en n incógnitas sobre C, donde C son los números complejos, en el caso en que el sistema tenga sólo un número finito de soluciones. En este caso, el espacio vectorial $A=C[x_1,...,x_n]/I$ tiene dimensión finita

Teorema 2.1

1. Si f está en $C[x_1,...,x_n]$, definimos $m_f:A \rightarrow A$, así: $[g] \rightarrow [fg]$, para cada $g \in A$. Entonces m_f es una transformacion lineal de A en A.
2. Si f,g están en $C[x_1,...,x_n]$, $m_f= m_g$, si y sólo sí $f-g$ está en I. En particular m_f es la función 0, sólo si f está en I.

Teorema 2.2

Si f,g están en $C[x_1,...,x_n]$, entonces $m_{f+g}= m_f+ m_g$ y 2. $m_{f+g} = m_f * m_g$, donde * es la composición de las funciones.

Ahora es fácil ver que si $h(t) = \sum c_i t^i$ está en $C[t]$, entonces la expresión $h(f) = \sum_i c_i f^i$ donde f está en $C[x_1, ..., x_n]$ tiene sentido. Y similarmente $h(m_f) = \sum_i c_i(m_f)^i$, donde m_f es una matriz de m_f.

También tenemos el siguiente colorario: Si h está en $C[t]$ y f está en $C[x_1,....,x_n]$, entonces $m_{h[t]} = h(m_f)$.

Recordando que si f está en $C[x_1,...,x_n]$, entonces f determina su clase en A, como A es de dimensión finita sobre C, existen $b_1,...b_l$, en C no todos ceros, tales que $\sum b_i [f]^i= 0$.

Luego $\sum b_i f^i$ está en I,y $\sum b_i f^i$ se anula en $V(I)$. Y si $h(t)$ está en $C[t]$, $h(m_f) = 0$, si y sólo si $h([f]) = 0$ en A.

Ahora, dada una matriz M, dxd, sobre K. El conjunto I_M de todos los polinomios $h(t)$ tales que $h(M) = 0$, la matriz 0, es un ideal; el generador de ese ideal se llama el polinomio mínimo de M, y como ese polinomio divide a todos los elementos de $I_{M,}$ divide al polinomio característico de M.

Sea h_f el polinomio mínimo de la multiplicación m_f: $A \rightarrow A$. Entonces tenemos 3 conjuntos interesantes de números:

1. Las raíces de la ecuación $h_f(t)=0$.
2. Los valores característicos de la matriz m.
3. Los valores de la función f de $V(I)$ en C que queremos hallar.

Teorema 2.3

Sea $V = V(I)$ una variedad afín en C^n y supongamos que en $C[x_1,...,x_n]$, tenemos definido un orden monomial. Entonces son equivalentes:

1) V es finita.
2) Para cada i, $1 \leq i \leq n$, existe $mi \geq 0$ tal que $xim \in <LT(I)>$
3) Sea G una base de Gröbner de I. Para cada i, $1 \leq i \leq n$ existe $mi \geq 0$ tal que $xim = LT(g)$, para un $g \in G$.
4) El C - espacio vectorial generado por $\{x\alpha : x\alpha\} \notin <LT(I)> \}$ es de dimensión finita.
5) El C- espacio vectorial $C[x1,...,xn\ ] / I$ es de dimensión finita.

Un ideal I que satisface una de las cinco condiciones anteriores se llama un ideal cero dimensional.

Teorema 2.4

Sea $I \in C[x_{1,}...,x_n\ ]$ un ideal cero dimensional, $f \in C[x_{1,}...,x_n]$, y h_f el polinomio mínimo de $m_f : A \rightarrow A$.

Entonces para un $\lambda \in C$, son equivalentes:

a. λ es una raíz de la ecuación $h_f = 0$.
b. λ es un valor característico de m_f.
c. λ es un valor de la función $f : V(I) \rightarrow C$.

Demostración: $a \Leftrightarrow b$, se sigue de la definición de valor característico de una transformación lineal.

$b \rightarrow c$. Sea λ un valor característico de m_f, entonces existe un vector $z \neq 0$ en A, tal que $[f- \lambda]$ $[z] = 0$.

Para llegar a una contradicción, supongamos que λ no es un valor de f en $V(I)$, es decir, supongamos que $V(I) = \{p_1,...,p_m\}$ y que $f(p_i) \neq \lambda$ para $i = 1,...,m$. Sea $g = f- \lambda$ una función tal que $g(p_i) \neq 0$ para cada i.

Como dados p_1, ..., p_m existe una función g_i tal que $g_i(p_j) = 0$ si $i \neq j$ y $g_i(p_j) = 1$, si $i = j$. Consideremos el polinomio $g' = \Sigma(1/g(p_i))g_i$. Entonces $g'g(p_i) = 1$ para todo i. Y $1\text{-}g'g \in I\ (V(I))$, lo que implica que $g'g)^l \in I$ $(1\text{-}$ para un $l \geq 1$. Desarrollando la expresión anterior y reduciendo términos semejantes tenemos que $1\hat{g}\ g \in I$ para $\hat{g} \in C[x_{1,}...,x_n]$. De donde se sigue que g es invertible lo que es una contradicción.

$c \Leftrightarrow a$. Sea $\lambda = f(p)$ para $p \in V(I)$. Como $h_f(m_f) = 0$, entonces $h_f([f]) = [0]$, $h_f(t) \in I$, y $h_f(f)$ se anula en cada punto de V(I). Por lo tanto, $h_f(\lambda) = h_f(f(p)) = 0$.

Corolario 2.5. Sea I en $C[x_{1,}...,x_n\ ]$ un ideal cero dimensional. Entonces los valores característicos de la multiplicación m_x en A, coinciden con las x_i coordenadas de los puntos en $V(I)$.

Además sustituyendo $t = x_i$ en el polinomio mínimo h_x tenemos el único generador mónico del ideal $I \cap C\ [x_i\ ]$.

Ahora notamos que además de los vectores característicos izquierdos de una matriz, existen los derechos, y ellos están relacionados entre sí, como veremos a continuación. Sea λ un valor característico de M y v un vector característico asociado a él, entonces $Mv = \lambda v$.

Trasponiendo tenemos $M^t v^t = \lambda v^t$ y si hacemos $v^{tt} = w$, tenemos $wM = \lambda w$.

Para un ideal I, hay una fuerte conexión entre los puntos de $V(I)$ y los vectores característicos derechos de la matriz m_f, relativamente a una base monomial B que se obtiene a partir de la base de Grobner de I.

Supongamos que el ideal I es radical. En este caso como la dimensión de A es m, el número de puntos de $V(I)$. Podemos escribir la base B, anteriormente mencionada así:

$B = \{\ [x^{\alpha(1)}], \dots, [x^{\alpha(m)}]\}$

Usando esta base, supongamos que m_f es la matriz de la multiplicación de A por f. Entonces podemos relacionar los vectores característicos derechos de m_f con los puntos de $V(I)$, como sigue:

Teorema 2.6

Sea f en $C[x_1,...,x_n]$ elegido de tal forma que los valores $f(p)$ de f en $V(I)$ sean distintos.

Entonces los subespacios característicos derechos de M_f son de dimensión 1, y generados por los vectores renglones $(p^{\alpha(1)}, ..., p^{\alpha(m)})$ para p en $V(I)$.

Demostración

Sea $m_f = (m_{ij})$. Entonces para cada j entre 1 y m, $[x^{\alpha(j)}f] = m_f\ [x^{\alpha(j)}] = m_{ij}[x^{\alpha(j)}]+...+ m_{mj}[x^{\alpha(m)}]$.

Ahora si p en $V(f_1,...,f_s)$ es fijo y evaluamos esta ecuación en p tenemos:

$$p^{\alpha(j)}f(p) = m_{1j}p^{\alpha(1)}+...+ m_{mj}p^{\alpha(m)}$$

Y si hacemos esto para $j = 1,..., m$ tenemos que: $f(p)\,(p^{\alpha(1)},..., p^{\alpha(m)}) = (p^{\alpha(1)},..., p^{\alpha(m)})\, m_f$, y como uno de los elementos básicos de A es 1, $(p^{\alpha(1)},... ,p^{\alpha(m)})$ es un vector característico de m_f.

Como por hipótesis, todos los valores de $f(p)$ son distintos para p en $V(I)$, la matriz m_f tiene m valores característicos distintos, y los subespacios característicos determinados por m_f son de dimensión 1.

Para terminar usaremos la proposición anterior para hallar los puntos en $V(I)$ para un ideal I de dimensión 0.

Primero supongamos que I es radical, reemplazando I por radical de I, si es necesario.

Entonces calculemos una base de Grobner G de I como es usual, y consideremos la función $f = c_1 x_1+...+ c_n x_n$, donde los elementos $c_1,...,c_n$, son enteros elegidos aleatoriamente. Esto garantiza casi siempre, que los valores $f(p)$ son distintos para p en $V(I)$.

Relativamente a la base B hallemos la matriz m_f, un valor característico λ de ella y un vector característico derecho v asociado a él.

Este vector característico cuando se combina con la base de Grobner G hace trivial hallar una solución p en $V(I)$.

Para ver esto, notemos que por el Teorema 2.5
(*) $v = \lambda(p^{\alpha(1)},..., p^{\alpha(m)})$ para un $\lambda \neq 0$ y p en $V(I)$.

Escribamos $p = (a_1, ..., a_n)$. Queremos hallar las coordenadas a_i de p en términos de las coordenadas de p. La ecuación * implica que cada coordenada de v es de la forma $\lambda p^{\alpha(j)}$..Por el teorema 2.3, para cada i, $i =1,\dots,n$ existe un m_i tal que un polinomio g_i en la base de Grobner G tiene por término líder $a\, x_i^c$, donde $c=m_i$.

Si $m_i > 1$ entonces $[x_i]$ está en B, luego λa_i es una coordenada de v, y como $[1]$ está en B, λ también es coordenada de v, por lo tanto $a_i = \lambda a_i / \lambda$ y hemos obtenido la coordenada a_i.

Si $m_i = 1$, la variable x_i no aparece en B. Entonces usando la base de Grobner como guía, supongamos que $x_1 > ... > x_n$, y que $i_1 > ... > i_1$ sean las variables que aparecen a la potencia 1. Entonces para $j = 1,...,l$, existe un g_j en G, tal que $g_j = x_u$ + suma de términos que contienen x_i con $i > i_j$ y $u= i_j$

Y si evaluamos estas g_j en $p = (a_1, ... ,a_n)$ tenemos que $0 = a_i$ + suma de términos que contienen a_i, con $i > i_j$ y de allí es facil hallar a_i para $i = i_{1,\dots,}\ i_{l.}$

CONCLUSIONES

En la sección anterior, sección 2, hemos usado técnicas de álgebra lineal para resolver sistemas de ecuaciones, ellas han trabajado muy bien y nos han mostrado un camino alternativo para la solución de sistemas de ecuaciones, tan eficientes han sido que ellas se están considerando para la próxima generación de sistemas algebraicos computacionales.

REFERENCIAS

1. D. Cox, J. Little and D. O'Shea. *Using Algebraic Geometry*, Springer, 1997.

2. D. Cox, J. Little and D. O'Shea. *Ideals, Varieties and Algorithms* Springer, 1996.

3. I. Emiris and A. Rege. "Monomial bases and polynomial systems Solving", in *Proceedings of International symposium on Symbolic and Algebraic computation*, ACM Press, 1994, 114-122.

4. H. Statter. *Multivariate polynomials equations and matrix eigen problems*, Preprint, 1993.