Introducción a la Geometría Algebraica Computacional

Abraham Martín del Campo

30 de abril de 2020

Índice general

Capítulo 1

Lo básico de las bases de Gröbner

Consideremos un sistema de ecuaciones

$$g_1 = g_2 = \cdots = g_r = 0 \tag{1.1}$$

donde cada g_i es un polinomio, es decir, un elemento del anillo de polinomios con coeficientes complejos $\mathbb{C}[x_1, \ldots, x_n]$. Decimos que el sistema tiene *solución*, si existe un punto $p \in \mathbb{C}^n$ tal que $g_i(p) = 0, \forall i = 1, \ldots, r$. Resulta entonces natural preguntarse si podemos decir de manera efectiva, si el sistema dado tiene al menos una solución o no. Nótese que ya podemos contestar esta pregunta para algunos casos.

Caso 1: (Polinomios lineales)

Supongamos que los polinomios del sistema (1.1) son lineales, es decir, son de la forma

$$g_i = \sum_j \left(a_{i,j} \cdot x_j\right) + b_i = 0$$

con $a_{i,j}, b_i$ números complejos. En álgebra lineal le asociamos al sistema una matriz $\mathbf{A} = (a_{i,j})$, y un vector de términos independientes $\mathbf{b} = (-b_i)$. La eliminación Gaussiana transforma la matriz extendida $(\mathbf{A} \mid \mathbf{b})$, por medio de operaciones elementales, en una matriz escalonada de la cual se pueden leer las soluciones fácilmente.

Ejemplo 1.1. Si el sistema lineal es

$$\begin{aligned} g_1 &: 2x + 3y + 4z - 5 = 0 \\ g_2 &: 3x + 4y + 5z - 2 = 0, \end{aligned}$$

la eliminación Gaussiana toma como pivote, por ejemplo, la indeterminada x en g_1 para eliminarla en g_2, y después toma y como pivote para eliminarla de g_1, y al final obtener

$$\begin{aligned} x &= z - 14 \\ y &= 11 - 2z, \end{aligned}$$

de donde se calculan las soluciones con mayor facilidad que en el sistema original. △

Caso 2.- (Polinomios en una indeterminada)

Cuando los polinomios del sistema (1.1) involucran solamente una indeterminada, entonces podemos hacer uso del algoritmo de Euclides para encontrar el máximo común divisor $f = \text{MCD}(g_1, \ldots, g_r)$ y escribirlo como combinación polinomial:

$$f = h_1 g_1 + \cdots + h_r g_r.$$

De aquí podemos ver que si $p \in \mathbb{C}^n$ es una solución del sistema, entonces cada g_i se anula en p, y por tanto f también se anula en p. Más aún, como $f \mid g_i$ para $i = 1, \ldots, r$, entonces también se cumple que si f se anula en p, entonces cada g_i se debe anular en p.

Ejemplo 1.2. Consideremos los polinomios

$$\begin{aligned} g_1 &= x^3 - x^2 - 2x \\ g_2 &= x^2 - 3x + 2, \end{aligned}$$

El algoritmo de Euclides trata de cancelar términos para ir reduciendo los polinomios. En este caso, el término de mayor grado en g_2 se usa como pivote para cancelar el termino mayor de g_1; así, el algoritmo escribe

$$g_1 - x g_2 = 2x^2 - 4x.$$

Como el término mayor de este polinomio sigue dividiendo al de g_1, se prosigue de la misma manera, y el algoritmo escribe

$$g_1 - x g_2 - 2 g_2 = 2(x - 2)$$

Si escribimos $g_3 = (x - 2)$, entonces podemos ver que tanto g_1 y g_2 son divisibles por g_3. Por lo tanto, tenemos que $\text{MCD}(g_1, g_2) = g_3$. Más aún, podemos escribir a g_3 como combinación de g_1 y g_2 invirtiendo los pasos anteriores.

$$g_3 = \frac{1}{2} g_1 - \frac{(x+2)}{2} g_2.$$

Como g_3 se anula en $x = 2$, entonces tanto g_1 y g_2 deben anularse también. △

En ambos casos, iniciamos con un sistema $\mathcal{G} = \{g_1, \ldots, g_r\}$, y a partir de manipulaciones algebraicas obtenemos otro sistema $\mathcal{F} = \{f_1, \ldots, f_s\}$ que es "equivalente" a $\mathcal{G}$ y "mejor"de alguna manera. Esta es precisamente la filosofía detrás de las bases de Gröbner. Con fin de precisar la idea de sistemas "equivalentes", fijaremos la siguiente notación.

Denotamos por $S := \mathbb{K}[x_1, \ldots, x_n]$ al anillo de polinomios en n indeterminadas con coeficientes en un campo $\mathbb{K}$. El *ideal generado* por un conjunto de polinomios $\{g_1, \ldots, g_r\}$ es el conjunto

$$\langle g_1, \ldots, g_r \rangle := \{h_1 g_1 + \cdots + h_r g_r \mid h_i \in S\}.$$

De esta manera, decimos que dos sistemas de polinomios $\mathcal{G} = \{g_1, \ldots, g_r\}$ y $\mathcal{F} = \{f_1, \ldots, f_s\}$ son equivalentes si generan el mismo ideal, es decir, si $\langle \mathcal{G} \rangle = \langle \mathcal{F} \rangle$. Un punto $p \in \mathbb{K}^n$ es un *cero de un polinomio* $g \in S$, si g se anula en p, es decir $g(p) = 0$.

Notemos de la definición de ideal, que si $f \in \langle g_1, \ldots, g_r \rangle$ entonces deben existir polinomios $h_1, \ldots, h_r$ tales que

$$f = h_1 g_1 + \ldots + h_r g_r.$$

Por lo tanto, una solución del sistema $\{g_1, \ldots, g_r\}$ debe ser también un cero de f.

En particular, si $1 \in \langle g_1, \ldots, g_r \rangle$ entonces el sistema no tiene solución, ya que el polinomio constante 1 núnca se anula. Esto traduce la pregunta sobre la existencia de soluciones a decidir si un polinomio cae dentro de un ideal dado. Resulta que las bases de Gröbner también son útiles para contestar este tipo de preguntas.

Cabe mencionar que la implicación inversa no siempre es cierta: si un sistema no tiene solución, no necesariamente podemos encontrar al 1 dentro del ideal, a menos que $\mathbb{K}$ sea algebraicamente cerrado. Este es el teorema de los ceros de Hilbert, que veremos más adelante.

En resumen, *las bases de Gröbner generalizan el método de eliminación Gaussiano a sistemas no lineales, y al mismo tiempo generalizan el algoritmo de Euclides a polinomios con más de una indeterminada. Estas bases nos ayudan a decidir si un polinomio cae dentro de un ideal o no.*

1.1. Ideales monomiales y el lema de Dickson

Los cimientos sobre los que se erigen las bases de Gröbner son los ideales monomiales. Trabajar con este tipo de ideales tiene varias ventajas, ya que son objetos esencialmente combinatorios, y algunas preguntas algebraicas se traducen en propiedades sencillas de verificar.

Definición 1.3. Un *monomio* es un polinomio (mónico) con un solo término,

$$x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n} \in S, \quad \text{con } \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n.$$

Definición 1.4. Un *ideal monomial* es un ideal $I \subseteq S$ que satisface cualquiera de las siguientes condiciones (que son equivalentes):

(i) I está generado por monomios.

(ii) Si $f \in I$, entonces cada monomio de f también está en I.

Aunque la definición de ideal monomial dada es intuitiva, determinar si un ideal dado es o no monomial no es una tarea fácil, a menos que el ideal se nos haya presentado con un conjunto de generadores monomiales.

Ejemplo 1.5. El ideal $I = \langle x^3, y^2 \rangle$ es monomial, y contiene al polinomio $y^2 - x^3$. Con esto vemos que los ideales monomiales contienen más que monomios. Observemos también que aunque el ideal $J = \langle y^2 - x^3, y^2 \rangle$ no está definido por monomios, este ideal resulta monomial, ya que se puede verificar rápidamente que $J = \langle x^3, y^2 \rangle = I$. △

Dado que la única información necesaria para definir un monomio x^α es el exponente $\alpha \in \mathbb{N}^n$, los monomios en S están en biyección con los puntos en $\mathbb{N}^n$.

Definición 1.6. Dado un polinomio $f \in S$, definimos su *soporte* como el conjunto

$$\operatorname{spte}(f) := \{\alpha \in \mathbb{N}^n \mid x^\alpha \text{ es un monomio de } f\}.$$

Más aún, si $I \subseteq S$ es un ideal (no necesariamente monomial), definimos su soporte como

$$\operatorname{spte}(I) := \bigcup_{f \in I} \operatorname{spte}(f).$$

Dado que el soporte de un ideal está definido como la unión de los soportes de los polinomios contenidos en él, tenemos que si $f \in I$ entonces $\operatorname{spte}(f) \subseteq \operatorname{spte}(I)$. Para los ideales monomiales la relación es más fuerte, ya que por la segunda parte de la Definición 1.4, los monomios de f también tienen que pertenecer al ideal. En este caso, tenemos una identificación entre ideales monomiales y su conjunto de exponentes:

$$I \text{ monomial} \quad \longleftrightarrow \quad \{\alpha \in \mathbb{N}^n \mid x^\alpha \in I\} = \operatorname{spte}(I). \tag{1.2}$$

Si consideramos un ideal $I = \langle x^\alpha \rangle$ generado por un único monomio, y un polinomio $f \in S$, podemos decidir si f está o no en I muy fácilmente, ya que $f \in I$ si podemos escribir $f = h \cdot x^\alpha$ con $h \in S$, es decir, si x^α divide a f. Nótese que esto sucede si y solo si x^α divide a cada monomio de f. Para ideales monomiales en general, la divisibilidad también caracteriza a los conjuntos generadores, como resumimos en la siguiente observación.

Observación 1.7. *Si I es monomial, un conjunto de monomios $G \subseteq I$ es generador si y sólo si, cada uno de los monomios de I es divisible por alguno de los de G.*

Gracias a la biyección (1.2), podemos traducir la pertenencia en un ideal monomial I a su versión discreta $\operatorname{spte}(I)$. Empecemos observando que $x^\alpha \mid x^\beta$ implica que $\alpha_i \leq \beta_i$ para toda$i = 1, \ldots, n$, y viceversa. Por lo tanto, la divisibilidad de monomios está en correspondencia con un orden parcial en $\mathbb{N}$. Así, tenemos que para ideales monomiales, si $\alpha \in \operatorname{spte}(I)$ y $\alpha \leq \beta$, entonces $\beta \in \operatorname{spte}(I)$.

Ejemplo 1.8. Observemos que $xyz \mid x^3yz^2$ se traduce en el soporte como la desigualdad $(1,1,1) \leq (3,1,2)$. Inversamente, elementos no comparables $(3,1,2) \not\leq (2,3,1)$ están en biyección con monomios no divisibles $x^3yz^2 \nmid x^2y^3z$. △

Ordenes parciales en los naturales

Recordar un poquito de órdenes, como qué es un orden parcial, uno total, una anticadena, el lema de Zorn, un refinamiento de órdenes, que un buen orden implica inducción.

Es posible que la mejor manera de poner esta información es en un apéndice. –Abraham

Si consideramos un conjunto de monomios que generan un ideal monomial, podemos extraer de ahí un conjunto de generadores que sea mínimo bajo divisibilidad, ya que la divisibilidad define un orden parcial en $\mathbb{N}^n$, lo que demuestra el siguiente lema.

Lema 1.9. *Los ideales monomiales tienen un único conjunto mínimo de generadores.*

Una ventaja de trabajar con la biyección entre ideales monomiales I y su soporte, es que podemos visualizar algunas propiedades de I. Por ejemplo, la minimalidad de un conjunto de generadores es fácil de ver cuando trabajamos con $\text{spte}(I)$ como ilustramos con el siguiente ejemplo.

Ejemplo 1.10. Tomemos $I = \langle y^4, x^3y^3, x^5y, x^6y^2 \rangle$. Nótese que si $x^{a_1}y^{a_2} \in I$, entonces para todo $(b_1, b_2) \in \mathbb{N}^2$ se cumple $x^{a_1+b_1}y^{a_2+b_2} \in I$, por lo tanto, en el soporte de I, los múltiplos del monomio $x^{a_1}y^{a_2}$ están codificados por una traslación de $\mathbb{N}^n$ que empieza en (a_1, a_2). De esta manera, podemos visualizar al ideal monomial I con el diagrama escalonado de la Figura 1.1, en donde la región sombreada representa a todos los monomios dentro de I. Notemos en

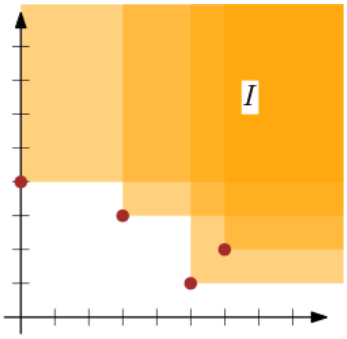


Figura 1.1: Diagrama escalonado para $I = \langle y^4, x^3y^3, x^5y, x^6y^2 \rangle$

este diagrama, el punto $(6, 2)$ y todos sus múltiplos caen dentro de la región formada por los múltiplos de $(5, 1)$, y por tanto es un generador redundante. De esta forma, podemos ver que I está generado mínimamente por $\{y^4, x^3y^3, x^5y\}$. △

Observemos que la región no sombreada de la Figura 1.1 corresponde a aquellos monomios que están fuera de I. Dichos monomios son un buen conjunto de representantes del cociente S/I y juegan un rol importante.

Definición 1.11. Llamamos *monomios estándar* a aquellos monomios que están fuera de un ideal monomial I.

Proposición 1.12. *Si $I \subseteq S$ es un ideal monomial, el cociente S/I es un espacio vectorial sobre $\mathbb{K}$, y los monomios estándar forman una base.*

Demostración. Se deja como ejercicio. hacer referencia al ejercicio corresp. al final del capítulo –Abraham Necesitamos incluir un recordatorio de lo que significa el cociente S/I en general. □

Teorema 1.13 (Lema de Dickson)**.** *Los ideales monomiales son finitamente generados.*

Demostración. Procederemos por inducción en el número de variables. Para $n = 1$, los monomios son de la forma x^d, con $d \in \mathbb{N}$. Por lo tanto, podemos considerar al exponente mínimo en I tal que $x^d \in I$, digamos x^a, y entonces $I = \langle x^a \rangle$.

Supongamos ahora que el teorema se cumple para alguna $n \geq 1$ y sea $I \subseteq \mathbb{K}[x_1, \ldots, x_n, y]$ un ideal monomial. Para $d \geq 0$, definamos

$$I_d := \{x^\alpha \mid x^\alpha y^d \in I\}.$$

Veamos a continuación que I_d es un ideal monomial. Para ello, tomemos $x^\alpha \in I_d$ y otro monomio x^β tal que $x^\alpha | x^\beta$ entonces, por definición de I_d, tenemos que $x^\alpha y^d \in I_d$; como I es ideal se tiene que $x^\beta y^d \in I$ y por lo tanto $x^\beta \in I_d$.

Ahora, definimos

$$I_\infty := \bigcup_{d \geq 0} I_d = \{x^\alpha \mid \exists\, d \geq 0 \text{ tal que } x^\alpha y^d \in I\}.$$

Notemos que I_∞ también es un ideal monomial (la demostración es análoga a la hecha con I_d). De la hipótesis de inducción, digamos que I_d está finitamente generada por G_d, esto para cada $d \geq 0$, y para $d = \infty$. Notemos además que tenemos una cadena creciente de ideales

$$I_0 \subseteq I_1 \subseteq I_2 \subseteq \cdots \subseteq I_\infty \subseteq \mathbb{K}[x_1, x_2, \ldots, x_n].$$

Por lo tanto, debemos tener que $I_\infty = I_d$ para alguna $d < \infty$. Veamos que esto se cumple, ya que para cada generador $x^\alpha \in G_\infty$ existe b_α tal que $x^\alpha y^{b_\alpha} \in I$. Sea $d' = \max_{x^\alpha \in G_\infty}\{b_\alpha\}$ entonces $I_\infty = I_d$ y también $I_b = I_{d'}$ para toda $b > d'$, y por tanto podemos considerar que $G_b = G_{d'}$. Definamos ahora el conjunto

$$G := \bigcup_{b=0}^{d'} \{x^\alpha y^b \mid x^\alpha \in G_b\}$$

Veamos que G genera a I mostrando que para cualquier monomio $x^\alpha y^b \in I$ podemos encontrar un monomio en G que divida a $x^\alpha y^b$. Como $x^\alpha \in I_b$ para esa $b \geq 0$, entonces existe un generador $x^\gamma \in G_b$ que divide a x^α, y por lo tanto $x^\gamma y^b$ divide a $x^\alpha y^b$. Solo nos resta ver que $x^\gamma y^b$ está en G. Cuando $b \geq d'$, tenemos que $x^\gamma \in G_b$ y por tanto $x^\gamma y^b \in G$ de la definición de G. Cuando $b > d'$ entonces $x^\gamma y^b \in G$ porque $G_b = G_{d'}$. □

Leonard E. Dickson probó este lema en 1913 en el contexto de teoría de números para probar un resultado sobre números perfectos, sin embargo, el resultado ya era conocido antes, inclusive fue probado por Paul Gordan en 1899 (referencias) como parte de una prueba del teorema de la base de Hilbert, el cual veremos más adelante. Por ahora, veamos que una consecuencia inmediata del Lema de Dickson es la siguiente.

Corolario 1.14. *Cualquier cadena creciente de ideales monomiales $I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$ es finita.*

1.2. Órdenes monomiales y bases de Gröbner

Para poder generalizar el algoritmo de Euclides de una a varias variables, necesitamos refinar el orden parcial en monomios a un orden total. Recordemos que un orden total, es un orden en el que cualesquiera dos elementos son comparables. Aunque el refinamiento no es único, de todos modos nos permitirá decidir quién es el término mayor de un polinomio.

Definición 1.15. Un *orden monomial* $\prec$ es un orden total de lo monomios de S tal que:

(i) El polinomio 1 es el mínimo.

(ii) El orden respeta la multiplicación por monomios:

$$\text{si } x^\alpha \prec x^\beta \Rightarrow x^\alpha \cdot x^\gamma \prec x^\beta \cdot x^\gamma, \text{ para cualquier monomio } x^\gamma.$$

Observación 1.16. *Las dos condiciones anteriores nos dicen que $\prec$ es una extensión lineal de la división por monomios, es decir: si $x^\alpha \mid x^\beta$ entonces $x^{\alpha-\beta}$ es un monomio, por (i) tenemos que $1 \preceq x^{\beta-\alpha}$, y de la propiedad (ii) tenemos que $x^\alpha \preceq x^\beta$.*

Lema 1.17. *Los órdenes monomiales de la Definición 1.15 son exáctamente aquellos órdenes $\prec$ en monomios, que satisfacen (ii) y son un buen orden.*

Demostración. Iniciaremos demostrando que si $\prec$ es un orden monomial como en la Definición 1.15, entonces también es un buen orden. Sea M un conjunto de monomios e I el ideal monomial que ellos generan. Por el lema de Dickson, existe $G \subseteq I$ un conjunto finito de monomios tal que $I = \langle G \rangle$. Podemos además considerar que G es un conjunto generador mínimo, es decir, que cualquier subconjunto de G no es generador de I.

Observemos que $G \subseteq M$, ya que si existiera $g \in G$ que no estuviera en M, entonces existiría un $f \in M$ tal que $f \mid g$, pues I está generado por M. Como G es un conjunto generador, también existe $h \in G$ tal que $h \mid f$. Por lo tanto, $h \mid g$ y son distintos, lo cual contradice la minimalidad de G.

Sea x^γ el elemento mínimo de G, el cual existe pues el orden $\prec$ es total y G es finito. Necesitamos probar que x^γ es un mínimo de M y no nada más de G. Para ello, consideremos un $x^\beta \in M$, entonces existe un $x^\alpha \in G$ tal que $x^\alpha \mid x^\beta$ y así $x^\alpha \prec x^\beta$; como x^γ es el mínimo de G, entonces $x^\gamma \prec x^\alpha \prec x^\beta$, y por lo tanto x^γ también es mínimo en M. De aquí se sigue que $\prec$ es un buen orden.

Ahora supongamos que $\prec$ es un buen orden que satisface (ii), pero que no es un orden monomial. Entonces existe $x^\alpha \prec 1$, y de la propiedad (ii) obtenemos una cadena descendiente $x^\alpha \succ x^{2\alpha} \succ x^{3\alpha} \succ \cdots$ que no es finita, lo cual contradice la hipótesis de que $\prec$ es un buen orden. Por lo tanto $\prec$ debe ser un orden monomial. □

Observación 1.18. *En la teoría de ideales monomiales, a los órdenes monomiales de la Definición 1.15 a veces se les llama órdenes globales, dejando el nombre de órdenes locales a aquellos órdenes que no necesariamente cumplen la condición (i), y que por el Lema 1.17, no son un buen orden.*

Veamos ahora algunos órdenes monomiales que son de gran importancia por sus propiedades, y que serán muy relevantes en desarrollo del resto de este capítulo.

Definición 1.19. Definimos el *grado* de un monomio x^α como $\mathrm{gr}(x^\alpha) := \alpha_1 + \ldots + \alpha_n$.

1. El *orden lexicográfico* (lex):

$$x^\beta \prec_{lex} x^\alpha \iff \text{la primera entrada distinta de cero de } \alpha - \beta \text{ es positiva.}$$

2. El *orden lexicográfico graduado* (glx):

$$x^\beta \prec_{glx} x^\alpha \iff \begin{cases} \mathrm{gr}(x^\beta) < \mathrm{gr}(x^\alpha) & \text{o bien ,} \\ \mathrm{gr}(x^\beta) = \mathrm{gr}(x^\alpha) & \text{y} \quad x^\beta \prec_{lex} x^\alpha. \end{cases}$$

3. El *orden lexicográfico inverso graduado* (lig):

$$x^\beta \prec_{lig} x^\alpha \iff \begin{cases} \mathrm{gr}(x^\beta) < \mathrm{gr}(x^\alpha) & \text{o bien ,} \\ \mathrm{gr}(x^\beta) = \mathrm{gr}(x^\alpha) & \text{y la última entrada distinta} \\ & \text{de cero de } \alpha - \beta \text{ es negativa.} \end{cases}$$

Ejemplo 1.20. Para ilustrar algunas diferencias entre estos órdenes, veamos cómo ordenan todos los monomios de grado dos en tres variables.

$$\begin{aligned} \text{lex, glx:} \quad & x^2 > xy > xz > y^2 > yz > z^2. \\ \text{lig:} \quad & x^2 > xy > y^2 > xz > yz > z^2. \end{aligned}$$

△

Por último, daremos la manera más general de definir un orden entre monomios.

4. Un *orden ponderado* $\prec_\omega$: cada vector $\omega \in \mathbb{R}^n_{\geq 0}$ define un orden de la siguiente manera

$$x^\beta \prec_\omega x^\alpha \iff \omega \cdot \beta \leq \omega \cdot \alpha.$$

Observación 1.21. *El requerimiento de que las entradas de ω no sean negativas es para garantizar que se cumpla la propiedad (i) de la Definición 1.15, y la linealidad del producto punto garantiza que se cumpla la propiedad (ii).*

Observación 1.22. *Para que $\prec_\omega$ sea un orden total, requerimos que las entradas de ω sean linealmente independientes sobre $\mathbb{Q}$, ya que de lo contrario, obtenemos un orden parcial. Sin embargo, si un vector de pesos $\omega \in \mathbb{R}^n_{\geq 0}$ produjera un orden parcial, podríamos usar otro orden $<_1$ para romper el empate, y así producir un nuevo orden $\prec_{\omega,1}$ tal que $x^\beta \prec_{\omega,1} x^\alpha$ si y solo si $\omega \cdot \beta < \omega \cdot \alpha$, o bien $\omega \cdot \beta = \omega \cdot \alpha$ pero $x^\beta <_1 x^\alpha$. De manera más general, se puede demostrar que todo orden monomial se puede escribir como un orden $\prec_{\omega_1,\omega_2,\ldots,\omega_r}$ definido de manera similar para ciertos pesos $\omega_1, \ldots, \omega_r \in \mathbb{R}^n_{\geq 0}$.*

Notemos que podemos extender un orden monomial $\prec$ a un orden entre los términos de un polinomio, si definimos $ax^\alpha \succ bx^\beta \iff x^\alpha \succ x^\beta$ cuando $a \cdot b \neq 0$.

Definición 1.23. Dado un $f \in S$ y un orden monomial $\prec$, definimos el *término inicial* de f como el término en f más grande con respecto a $\prec$, y lo denotamos por $in_\prec(f)$.

Ejemplo 1.24. Si $f = 3xyz - 7y^3 + 13z^2$, entonces $in_{<_{lex}}(f) = 3xyz$, pero $in_{<_{drl}}(f) = -7y^3$. △

Definición 1.25. Sea $I \subseteq S$ un ideal y $\prec$ un orden monomial. Definimos el *ideal inicial* de I como el ideal monomial generado por los términos iniciales de los elementos de I, es decir,

$$in_\prec(I) := \langle in_\prec(f) \mid f \in I \rangle.$$

Cuando el orden se entienda en el contexto, simplificaremos la notación escribiendo solamente $in(f)$ y $in(I)$.

Notemos también que para cualesquiera dos conjuntos de polinomios $G \subseteq G'$, tenemos una contención de ideales monomiales

$$\langle in_\prec(g) \mid g \in G\rangle \subseteq \langle in_\prec(g') \mid g' \in G'\rangle.$$

Sin embargo, es posible tener una igualdad de estos ideales monomiales aún cuando $G \subsetneq G'$. Esta situación especial aplicada al ideal inicial es lo que define una base de Gröbner.

Definición 1.26. Sean $I \subseteq S$ un ideal y $\prec$ un orden monomial. Decimos que un conjunto $G \subseteq I$ es una *base de Gröbner* de I si y solo si

$$in_\prec I = \langle in_\prec(g) \mid g \in G\rangle.$$

Notemos que si G es una base de Gröbner y $G \subseteq G'$, entonces también G' es una base de Gröbner, es decir, agregar elementos no cambia la propiedad de ser base de Gröbner. Una consecuencia del Lema de Dickson es que siempre podemos encontrar una G finita.

Lema 1.27. *Sea $\prec$ un orden monomial. Para todo polinomio $f \in S$ y todo monomio x^α se cumple que $x^\alpha \cdot in_\prec(f) = in_\prec(x^\alpha \cdot f)$.*

Demostración. Se deja como ejercicio (ver Ejercicio 3). □

Proposición 1.28. *Todo ideal tiene una base de Gröbner finita.*

Demostración. Sea $I \subseteq S$ un ideal y $\prec$ un orden monomial. Como $in_\prec(I)$ es monomial, el lema de Dickson garantiza la existencia de un número finito de elementos $m_1, \ldots, m_r \in in_\prec(I)$. Por definición, el ideal $in_\prec(I)$ está generado por todos los términos $in_\prec(g)$ con $g \in I$. Por lo tanto, existen $g_1, \ldots, g_r \in I$ tales que $in_\prec(g_i) \mid m_i$ para cada $i = 1, \ldots, r$. Por lo tanto, el conjunto finito $G = \{g_1, \ldots, g_l\} \subseteq I$ es una base de Gröbner. □

Teorema 1.29. *Si G es una base de Gröbner de un ideal I, entonces $\langle G\rangle = I$.*

Demostración. Sea G una base de Gröbner para un ideal I respecto a un orden monomial $\prec$. Como $G \subseteq I$ entonces $\langle G\rangle \subseteq I$. Supongamos que la igualdad no se cumple, por lo que los conjuntos $A := \{f \in I \mid f \notin \langle G\rangle\}$ y $in(A) := \{in_\prec(f) \mid f \in A\}$ no son vacíos. Sea entonces $h \in A$ tal que $in_\prec(h)$ es un elemento mínimo de $in(A)$ respecto a $\prec$.

Como $in_\prec(h) \in in_\prec(I)$ y G es base de Gröbner, entonces existe $g \in G$ tal que $in_\prec(g)$ divide a $in_\prec(h)$. Por lo tanto existe un término ax^α tal que $in_\prec(h) = ax^\alpha \cdot in_\prec(g)$, y del Lema 1.27 tenemos que $in_\prec(h) = in_\prec(ax^\alpha g)$. Consideremos el polinomio $f := h - ax^\alpha g$ y notemos que $in_\prec(f) \prec in_\prec(h)$.

Como $h \notin \langle G\rangle$ y $g \in G$ entonces $f \notin \langle G\rangle$, pues de otra forma $h = f + ax^\alpha g$ estaría en $\langle G\rangle$, contradiciendo la elección de h. Sin embargo, esto nos lleva a que $f \in A$, el cual tiene término inicial menor al de h, pero esto contradice la minimalidad de h. Por lo tanto, el conjunto A debe ser vacío y entonces $\langle G\rangle = I$. □

Estos dos últimos resultados nos muestra que todo ideal I tiene una base de Gröbner finita y ésta forma un conjunto generador. Esto constituye una prueba del teorema de la base de Hilbert, que nos dice que el anillo de polinomios S es un anillo noetheriano.

Corolario 1.30 (Teorema de la base de Hilbert)**.** *Todo ideal $I \subseteq S$ es finitamente generado.*

Otra aplicación importante de las propiedades de las bases de Gröbner, es que nos ayudan a decidir si dos ideales son el mismo o no.

Proposición 1.31. *Sean $J \subseteq I$ ideales y $\prec$ un orden. Si $in_\prec(J) = in_\prec(I)$, entonces $I = J$.*

Demostración. Sea $G \subseteq J$ una base de Gröbner, entonces $in_\prec(J) = \langle in_\prec(g) | g \in G \ \rangle = in_\prec(I)$. Como $G \subseteq I$, entonces es base de Gröbner para I. Por el teorema anterior, $J = \langle G \rangle = I$. □

Veremos ahora un primer ejemplo que ilustra las obstrucciones para que un conjunto generador cualquiera sea una base de Gröbner.

Ejemplo 1.32. Sea $I = \langle x^2, xy + y^2 \rangle \subseteq \mathbb{K}[x, y]$ y fijamos $\prec = <_{lex}$. Veremos que

$$\langle in_\prec(x^2),\ in_\prec(xy + y^2) \rangle = \langle x^2, xy \rangle \subsetneq in_\prec(I).$$

Observemos que podemos cancelar términos iniciales: $y \cdot (x^2) + (-x) \cdot (xy + y^2) = -xy^2 \in I$. Con esto tenemos que $y \cdot (xy + y^2) + (-xy^2) = y^3 \in I$, pero

$$y^3 \notin \langle in_\prec(x^2), in_\prec(xy + y^2) \rangle = \langle x^2, xy \rangle.$$

Por lo tanto, $\{x^2, xy + y^2\}$ es un conjunto generador de I pero no es una base de Gröbner. △

1.3. El algoritmo de la división y el teorema de Macaulay

En la introducción de este capítulo vimos que, de la definición de ideal, si $1 \in \langle g_1, \ldots, g_m \rangle$ entonces deben existir polinomios $h_1, \ldots, h_m$ tales que

$$1 = h_1 g_1 + \ldots + h_m g_m.$$

Por lo tanto, una solución del sistema $\{g_1, \ldots, g_m\}$ sería también un cero del polinomio constante 1. De este modo, vemos que el sistema no tiene solución si $1 \in \langle g_1, \ldots, g_m \rangle$.

Si tenemos una base de Gröbner, resolvemos esta interrogante de manera inmediata: Fija cualquier orden monomial $\prec$ y encuentra una base de Gröbner G del ideal $I = \langle g_1, \ldots, g_m \rangle$. Entonces, basta ver si $1 \in G$ o no, pues

$$1 \in I \iff 1 \in \langle in_\prec(G) \rangle \iff 1 \in in_\prec(G) \iff 1 \in G.$$

Sin embargo, las bases de Gröbner nos permiten contestar de manera más general, si un polinomio cae dentro de un ideal dado o no. Para el caso de una variable, esta interrogante se resuelve mediante el algoritmo de Euclides para dividir polinomios, y a continuación veremos cómo las bases de Gröbner extienden dicho algoritmo a varias variables.

Algoritmo 1 (División de polinomios en varias variables)

Entrada: Polinomios $g_1, \ldots, g_m, f$ y un orden monomial $<$
Salida: Polinomios $h_1, \ldots, h_m, r$ tales que:

(i) $f = h_1 g_1 + \ldots h_m g_m + r$,

(ii) $in(f) \geq in(h_i g_i)$ para todo $i = 1, \ldots, m$,

(iii) $in(f) \geq in(r)$,

(iv) ningún término de r es divisible por ningún $in(g_i)$.

1: Define $r = f$, y $h_1 = \cdots = h_m = 0$
2: **mientras** (algún término de r es divisible por algún $in(g_i)$) **hacer**
3: Sea ax^α el término más grande de r respecto a $<$ que cae en $\langle in(g_i)\rangle$.
4: Sea j el primer índice tal que $in(g_j) \mid ax^\alpha$.
5: Redefine las variables

$$r = r - \frac{ax^\alpha}{in(g_j)} g_j, \qquad h_j = h_j + \frac{ax^\alpha}{in(g_j)}.$$

6: **fin mientras**
7: **devolver** $r, h_1, \ldots, h_m$.

Prueba de validez. Cada iteración es una reducción de r por los polinomios $g_1, \ldots, g_m$. Con cada reducción, el término más grande en r divisible por algún $in(g_i)$ disminuye con respecto al orden monomial $<$, por lo que el algoritmo termina en un número finito de pasos.

Además, en cada iteración, la igualdad (i) se preserva. La propiedad (iii) también se preserva, pues se cumplía inicialmente (paso 1) y con cada reducción (paso 5), los nuevos términos de r son menores al término que se canceló. La propiedad (ii) también se preserva, ya que inicialmente $in(f) \geq in(h_i g_i)$ y los términos iniciales de $h_i g_i$ son siempre términos de r, por la propiedad (iii), la desigualdad se sigue cumpliendo hasta que el algoritmo termina. □

Ejemplo 1.33. Ahora ilustraremos el algoritmo de la división con el siguiente ejemplo. Dividamos el polinomio $f = x^2 y$ entre los polinomios $g_1 = x^2$ y $g_2 = xy + y^2\}$ usando el orden lexicográfico graduado $<_{lig}$.

Primero iniciamos las variables $r = x^2 y$ y $h_1 = h_2 = 0$. Para verificar el condicional en el paso 2, necesitamos calcular $in_<(g_i)$, que en este caso da $in_{<_{lig}}(g_1) = x^2$ y $in_{<_{lig}}(g_2) = xy$.

Notemos que un término de $r = x^2 y$ (el único) es divisible por ambos términos iniciales. Por lo tanto, el algoritmo define $ax^\alpha = x^2 y = r$, y como ambos polinomios dividen, el algoritmo toma el primero, haciendo $j = 1$. Redefinimos las variables

$$r = x^2 y - \frac{x^2 y}{x^2} x^2 = 0 \qquad h_1 = 0 + \frac{x^2 y}{x^2} = y.$$

Cuando el algoritmo vuelve a verificar el condicional del paso 2, la condición ya no se cumple y por tanto el algoritmo termina, devolviendo $r = 0, h_1 = y, h_2 = 0$, con lo que podemos escribir $f = y g_1 + 0 g_2 + 0$.

Si cambiamos el orden de g_1, g_2, el resultado es distinto, y vale la pena verlo. Si consideramos $g_1 = xy + y^2, g_2 = x^2$, entonces en la primer iteración el algoritmo define $ax^\alpha = x^2y$, y $j = 1$ de nuevo. Después, se redefinen las variables

$$r = x^2y - \frac{x^2y}{xy}(xy + y^2) = -xy^2 \qquad h_1 = 0 + \frac{x^2y}{xy} = x.$$

Como $-xy^2$ es divisible por $xy = in_{<_{lig}}(g_1)$, la condición del paso 2 se cumple, por lo que volvemos a definir $ax^\alpha = -xy^2$, y $j = 1$. Se redefinen las variables

$$r = -xy^2 - \frac{-xy^2}{xy}(xy + y^2) = y^3 \qquad h_1 = x + \frac{-xy^2}{xy} = x - y.$$

Ahora, la condición del paso 2 se cumple, regresando $r = y^3, h_1 = x - y$, y $h_2 = 0$. Por lo que tenemos que $f = (x - y)g_1 + 0g_2 + y^3$. △

Observación 1.34. *El residuo en el algoritmo no es único.* *considera el cambiar el orden del ejemplo con la definición del residuo –Abraham*

Definición 1.35. Dado un polinomio f, un orden monomial, y una lista ordenada de polinomios $G = (g_1, \ldots, g_m)$, al residuo r del Algoritmo 1 lo denotamos por f mód G.

Notemos que este residuo depende tanto del orden monomial, como el orden de la lista de polinomios. Además, en el Ejemplo 1.33 vemos que si f mód $G = 0$, entonces podemos concluir que $f \in \langle G \rangle$, pero si f mód $G \neq 0$ aún no podemos decidir si f está en el ideal. Esta ambigüedad es resuelta cuando G es una base de Gröbner, como veremos ahora.

Lema 1.36. *Sea G una base de Gröbner de un ideal I respecto a un orden $<$, y $f \in S$ un polinomio. Entonces $f \in I$ si y solo si f* mód $G = 0$

Demostración. Sea $r = f$ mód G. Si $r = 0$ sabemos que $f \in I$, por lo tanto, supongamos que $r \neq 0$. Como ningún término de r es divisible por ningún $in_<(g)$ para $g \in G$, en particular, $in(r) \notin in_< G = in_< I$, por lo tanto $r \notin I$, y así, $f \notin I$. □

Lema 1.37. *Sea G una base de Gröbner de I. Entonces, para todo $f \in S$, el residuo f* mód G *es único.*

Demostración. Supongamos que podemos escribir a f de dos maneras, $f = g_1 + r_1 = g_2 + r_2$ con $g_1, g_2 \in \langle G \rangle = I$. Entonces, $r_2 - r_1 = g_1 - g_2 \in I$. Como G es base de Gröbner, existe $g \in G$ tal que $in(g) \mid in(r_2 - r_1)$. Notemos que $in(r_2 - r_1)$ es un término de r_1 o de r_2, pero de la definición del residuo, ningún término de r_1 y r_2 es divisible por ningún $in(g)$, para todo $g \in G$. La única manera de que esto suceda es si $in(r_2 - r_1) = 0$ y así $r_2 - r_1 = 0$, por lo que el residuo es único. □

Las propiedades vistas hasta ahora caracterizan a las bases de Gröbner, por lo que en la literatura suelen presentarse varias definiciones equivalentes, entre ellas tenemos las siguientes (de las cuales hemos demostrado una implicación).

Proposición 1.38. *G es una base de Gröbner de I si y sólo si f* mód $G = 0$ *para todo $f \in I$.*

Demostración. Una implicación es el Lema 1.36 y la otra se deja como ejercicio (ver Ejercicio 4). □

Proposición 1.39. *G es una base de Gröbner de I si y sólo si G es finito y para todo $f \in I$ existe un $g \in G$ tal que $in(g) \mid in(f)$.*

Cuando G es una base de Gröbner de I, el residuo, f mód G se puede escribir como combinación lineal de monomios estándar, pues el cociente $S/in(I)$ es un espacio vectorial (cf. Proposición 1.12). Sin embargo, el cociente S/I es también un espacio vectorial sobre $\mathbb{K}$ sin importar que I no sea monomial, y Macaulay demostró algo aún más general acerca de los monomios estándar.

Teorema 1.40 (Macaulay)**.** *Sea $I \subseteq S$ un ideal y $<$ un orden monomial. Los representantes en S/I de los monomios estándar de $in_<(I)$ forman una base de espacio vectorial.*

Demostración. Sea G una base de Gröbner finita de I con respecto a $<$. Para un polinomio $f \in S$, tanto f como f mód G representan al mismo elemento en el cociente S/I. Como f mód G es combinación lineal de monomios estándar de $in_<(I)$, se sigue que el conjunto de monomios estándar generan.

Sea $h = \lambda_1 m_1 + \cdots + \lambda_s m_s$ una combinación lineal no trivial de monomios estándar $m_1, \ldots, m_s \notin in_<(I)$, con $\lambda_i \in \mathbb{K}$. Veamos que el representante de h es cero en S/I si y sólo si $h \in I$, que implica que $in_<(h) \in in_<(I)$; pero $in_<(h)$ debe ser un monomio estándar (pues es una combinación lineal de ellos), por lo que $in_<(h) \notin in_<(I)$, que nos lleva a una contradicción. Por lo tanto, h debe ser cero. □

Gracias al algoritmo de la división y el teorema de Macaulay, para todo polinomio f podemos calcular un único representante f mód G en S/I que involucra la base de monomios estándar, por lo que al residuo f mód G a veces se le conoce como la *forma normal* de f módulo I. De hecho, Hironaka y Gordan ya conocían las bases de Gröbner y las llamaban *bases estándar*.

1.4. El algoritmo de Buchberger

Ya hemos discutido un poco sobre la utilidad y la importancia de las bases de Gröbner, ahora veamos cómo calcularlas. Para saber si un conjunto G no es una base de Gröbner, basta ver que sus términos iniciales no generan al ideal inicial, es decir, que debemos encontrar al menos un polinomio en I cuyo término inicial no es divisible por los términos iniciales de G. En el Ejemplo 1.32 vimos un conjunto de generadores que no es base de Gröbner, y para demostrarlo hicimos una cancelación de términos iniciales para generar un nuevo polinomio con un término inicial diferente. Esta idea es la que ayudó a Bruno Buchberger a encontrar una caracterización de las bases de Gröbner, la cual veremos a continuación.

Definición 1.41. Dados dos términos ax^α, bx^β, el *mínimo común múltiplo* $\operatorname{mcm}(ax^\alpha, bx^\beta)$ es el mínimo monomio x^γ que es divisible por x^α y x^β, y que calculamos tomando $\gamma_i = \max(\alpha_i, \beta_i)$ para cada $i = 1, \ldots, n$.

Definición 1.42. Sean $f, g \in S$ polinomios distintos de cero y $<$ un orden monomial. Definimos el *S-polinomio* de f y g como

$$\text{Spol}(f,g) := \frac{\text{mcm}(in_<(f), in_<(g))}{in_<(f)} f - \frac{\text{mcm}(in_<(f), in_<(g))}{in_<(g)} g.$$

Teorema 1.43 (Criterio de Buchberger)**.** *Un conjunto G es una base de Gröbner del ideal que genera si y sólo si* $\text{Spol}(f,g)$ mód $G = 0$ *para todo* $f, g \in G$.

Bosquejo de prueba. Primero supongamos que G es una base de Gröbner de $I = \langle G \rangle$. Para cada $f, g \in G$ el S-polinomio $\text{Spol}(f,g) \in I$, y del Lema 1.36 concluimos que $\text{Spol}(f,g)$ mód $G = 0$.

Ahora supongamos que $G = \{g_1, ..., g_m\}$ satisface el criterio de Buchberger. Si $f \in \langle G \rangle$, se debe/puede demostrar que $in_<(f)$ es divisible por $in_<(g)$ para algún $g \in G$. □

Con este criterio, Buchberger enunció su algoritmo para encontrar bases de Gröbner, que se basa en añadir residuos de S-polinomios a una lista de generadores, y termina cuando esta lista satisface el criterio.

Algoritmo 2 (algoritmo de Buchberger)

Entrada: Polinomios $g_1, \ldots, g_m$ que generan un ideal I, y un orden monomial $<$.
Salida: Una base de Gröbner G de I respecto a $<$.

1: Definimos $G := \{g_1, \ldots, g_m\}$.
2: **para todo** $i < j$ **hacer**
3: Define $h_{ij} := \text{Spol}(g_i, g_j)$ mód G.
4: **fin para**
5: **mientras** algún $h_{ij} \neq 0$ **hacer**
6: Redefine $G = G \cup \{h_{i,j} \neq 0\}$.
7: Regresa al paso 2.
8: **fin mientras**
9: **devolver** G.

Prueba de validez. El algoritmo termina en un número finito de pasos, puesto que en cada paso, los términos iniciales de G generan un ideal monomial estrictamente más grande, por el Corolario 1.14, esta cadena creciente de ideales monomiales se detiene. El algoritmo termina cuando los polinomios de G satisfacen el criterio del Teorema 1.43. □

Ejemplo 1.44. Para ilustrar el algoritmo de Buchberger, retomemos el Ejemplo 1.32, y calculemos una base de Gröbner para $G = \{xy + y^2, x^2\}$ con respecto a $<_{lex}$.

Primero calculamos $\text{Spol}(g_1, g_2) = -xy^2$, y retomemos que en el Ejemplo 1.33 calculamos $h_{12} = -xy^2$ mód $G = y^3$. Por lo tanto, redefinimos $G = \{g_1, g_2, h_{12}\}$ y regresamos al paso 2.

Nótese que $\text{Spol}(g_2, h_{12}) = 0$, mientas que $\text{Spol}(g_1, h_{12}) = y^4$; sin embargo, y^4 mód $G = 0$. Como todos los residuos son cero, $\{xy + y^2, x^2, y^3\}$ es una base de Gröbner de $I = \langle G \rangle$. △

Lo primero que notamos en este algoritmo es que tenemos que calcular muchísimos S-polinomios y reducirlos, por lo que el algoritmo no es muy eficiente. De hecho, esta es una de

las razones que permiten ver que el cálculo de las bases de Gröbner no es paralelizable, y su complejidad no es polinomial. El predecir qué S-polinomios calcular, y reducir la complejidad aunque sea para algunos casos específicos, es aún una rama de estudio.

Las bases de Gröbner son muy útiles, como hemos visto, pero generalmente son mucho más grandes que el conjunto de polinomios con el que empezamos. Por lo tanto, cabe preguntarse si podemos definir una noción sensata de base mínima, y que esa base sea única.

Definición 1.45. Un conjunto G es una *base de Gröbner mínima* de un ideal I, si $in(G)$ genera a $in(I)$ minimamente, y cada $g \in G$ es mónico.

Ejemplo 1.46. Consideremos el ideal $I = \langle x^3 - 2xy, x^2y - 2y^2 + x, -x^2, -2xy, -2xy^2 + x\rangle$. Se puede verificar que el conjunto $\{x^2, xy, y^2 - \frac{1}{2}x\}$, es una base de Gröbner mínima con respecto al orden $<_{lig}$, y que para cualquier $a \in \mathbb{K}$, el conjunto $\{x^2 + ay, xy, y^2 - \frac{1}{2}x\}$ también lo es. △

Las bases de Gröbner mínimas no son únicas, sin embargo, todas tienen el mismo tamaño.

Lema 1.47. *Si G y H son bases de Gröbner mínimas para un ideal I, entonces $|G| = |H|$.*

Demostración. El resultado se sigue del hecho de que $in(G)$ e $in(H)$ forman una base mínimo del ideal monomial $in(I)$. Como bien señaló José Ángel, falta meter el lema de que los conjuntos generadores mínimos de un ideal monomial tienen la misma cardinalidad (después del Lema 1.9) –Abraham □

Definición 1.48. Una base de Gröbner se dice *reducida* si cada $g \in G$ es mónico, y ningún término de g es divisible por el término inicial de los otros polinomios en G.

Nótese que las bases de Gröbner reducidas son, en particular, bases mínimas; pero estas bases arreglan la pluralidad que mostramos en el Ejemplo 1.46.

Proposición 1.49. *Sean $I \subseteq S$ un ideal y $<$ un orden monomial. Existen una única base de Gröbner reducida de I con respecto a $\succ$.*

Demostración. Supongamos que $F = \{g_1, \ldots, g_s\}$ y $G = \{g_1, \ldots, g_t\}$ son bases de Gröbner reducidas de I. Por ser F y G bases mínimas, se sigue del Lema 1.47 que $s = t$. Ordenemos las bases de manera que $in_<(g_i) = in_<(g_i)$ para toda $i = 1, \ldots, n$. Ya que $g_i - g_i \in I$ para cada $i \in \{1, \ldots, n\}$, si $g_i - g_i \neq 0$ entonces $in_<(g_i - g_i)$ es un término de g_i o de g_i. Sin pérdida de generalidad supongamos que $in_<(g_i - g_i)$ es un término de g_i. Como $in_<(g_i - g_i) \in \langle in_<(F)\rangle$ entonces $in_<(g_j)$ divide a $in_<(g_i - g_i)$ para algún $j \in \{1, \ldots, n\}$ con $j \neq i$, pero esto contradice que F sea una base reducida. □

Ejercicios

1. Demuestra la equivalencia de las condiciones en la Definición 1.4 : Un ideal I es generado por monomios si y solo si $f \in I$ implica que cada monomio de f también está en I.

2. Ordena todos los monomios de grado menor o igual a tres en tres variables bajo los tres ordenes: $\prec_{lex}, \prec_{glx}, \prec_{lig}$.

3. *a*) Demuestra el Lema 1.27: El término inicial abre productos por monomios, es decir, si $\prec$ es un orden monomial, $f \in S$ un polinomio, y $ax^\alpha \in S$ un monomio, entonces $in_\prec(x^\alpha f) = x^\alpha \cdot in_\prec(f)$.

   b) Demuestra que en general, si $f, g \in S$ entonces $in_\prec(fg) = in_\prec(f) \cdot in_\prec(g)$.

   c) Demuestra que con la suma $in_<(f+g) \leq in_<(f)$.

4. Demuestra que si f mód $G = 0$ para toda $f \in \langle G \rangle$, entonces G es una base de Gröbner.

5. Demuestra la Proposición 1.39.

6. Supóngase que $I,\ J$ son ideales monomiales. Demuestra que

$$\dim(S/(I \cap J)) = \text{máx}\{\dim(S/I), \dim(S/J)\}$$

Capítulo 2

Diccionario algebro-geométrico

Nuestra motivación es resolver sistemas de ecuaciones, y para poder hablar más sobre la utilidad de las bases de Gröbner en ello, necesitamos entender un poco más sobre la geometría de las soluciones.

Esta sección debe ser elaborada con mucho más detalle, para que de verdad ayude a quienes lo están aprendiendo por primera vez, y mete muchos ejemplos.

2.1. Variedades afines y proyectivas

Definición 2.1. Dado un conjunto F de polinomios en S, definimos la *variedad* de F como

$$\mathcal{V}(F) := \{p \in \mathbb{K}^n \mid f(p) = 0,\ \forall f \in F\}.$$

Notemos que si $F = \{g_1, \ldots, g_r\}$ es un conjunto de polinomios, y consideramos un polinomio $f \in \langle F \rangle$, entonces podemos escribir $f = h_1 g_1 + \ldots + h_r g_r$. Así, $p \in \mathcal{V}(F)$ implica que $g_i(p) = 0$ para toda $i = 1, \ldots, r$, y por tanto $f(p) = 0$. Esto nos deja ver que si $I = \langle F \rangle$, entonces $\mathcal{V}(F) \subseteq \mathcal{V}(I)$, y resulta que esta es realmente una igualdad. agrega el argumento para la otra contención. –Abraham

Observación 2.2. *Si I es el ideal generado por un conjunto de polinomios F, entonces $\mathcal{V}(F) = \mathcal{V}(I)$.*

Las variedades son naturalmente un subconjunto del espacio vectorial $\mathbb{K}^n$, sin embargo, nuestro interés es entender la estructura geométrica de ellas, y no necesariamente su estructura como espacio vectorial. Al espacio $\mathbb{K}^n$ sin su estructura de espacio vectorial es el *espacio afín* $\mathbb{A}^n$. (el énfasis en hacer esta definición es que a $\mathbb{K}^n$ lo pensamos como un $\mathbb{K}$-espacio vectorial, mientras que a $\mathbb{A}^n$ lo pensamos como un objeto geométrico, en donde el 0 no juega ningún papel especial). A los subconjuntos de $\mathbb{A}^n$ de la forma $\mathcal{V}(F)$ para algún $F \in S$ los llamaremos *conjuntos algebraicos*.

Ejemplo 2.3. El círculo $X = \{\, (cos(t), sen(t)) \in \mathbb{R}^2 \mid t \in [0, 2\pi] \,\}$ es un conjunto algebraico, ya que X es el conjunto de soluciones del polinomio $x^2 + y^2 - 1 = 0$. una imagen y un ejemplo más sexy. –Abraham △

Notemos ahora que si buscamos entender los puntos $p \in \mathbb{K}^n$ que son soluciones de un polinomio $f \in S$, entonces, tanto $f(p) = 0$ como $a \cdot f(p) = 0$ para cualquier $a \in \mathbb{K}$, pero no siempre podemos deducir lo mismo de $f(a \cdot p)$ a partir de $f(p)$, a menos que f sea homogéneo. Decimos que un polinomio $f \in S$ es *homogéneo*, si todos sus términos son del mismo grado. Por ejemplo, el polinomio $g = x^2y + 3xy^2 + 4z^3 - 7xyz$ es homogéneo, mientras que $h = x^2y - xy + 2xyz + 5z - 7$ no. Observemos que siempre podemos escribir un polinomio f como suma de polinomios homogéneos

$$f = f_0 + f_1 + f_2 + \cdots + f_d, \tag{2.1}$$

donde f_i es un polinomio homogéneo de grado i. En los ejemplos anteriores, $g = g_3$ pues es homogéneo de grado 3, mientras que $h = h_0 + h_1 + h_2 + h_3$ con $h_0 = -7$, $h_1 = 5z$, $h_2 = -xy$ y $h_3 = x^2y + 2xyz$.

Notemos ahora que si f es un polinomio homogéneo, entonces $f(a \cdot p) = a^{\mathrm{gr}(f)} f(p)$, por lo que, en este caso $f(p)$ y $f(a \cdot p)$ son ambos 0 o ninguno lo es. Por ejemplo, el polinomio homogéneo g se anula en el punto $p = (2, 1, 1)$, y también en cualquier punto de la forma $ap = (2a, a, a)$ con $a \in \mathbb{K}$ distinto de cero, pues

$$g(2a, a, a) = (2a)^2(a) + 3(2a)(a)^2 + 4(a)^3 - 7(2a)(a)(a) = a^3 g(2, 1, 1) = 0.$$

Lo anterior induce la siguiente relación de equivalencia en $\mathbb{A}^n$, dada por $x \backsim y$ si y solo si $x = ay$, para algún escalar $a \neq 0$. A la clase de un punto $(x_1, \ldots, x_n) \in \mathbb{A}^n$ bajo esta relación la denotamos por *coordenadas homogéneas* $[x_1, \ldots, x_n]$. A partir de esta relación definimos el *espacio proyectivo* como el espacio cociente

$$\mathbb{P}^{n-1} := \mathbb{A}^n - \{0\} / \backsim .$$

Podemos pensar al espacio proyectivo $\mathbb{P}^n$ como el conjunto de subespacios de $\mathbb{K}^n$ de dimensión uno. Agregar aquí un ejemplo de cómo visualizar $\mathbb{P}^2$ por medio de rectas en $\mathbb{A}^3$. –Abraham Nótese que tenemos una inclusión de espacios $\mathbb{A}^n \subset \mathbb{P}^n \subset \mathbb{A}^{n+1}$.

Para definir las variedades en $\mathbb{P}^n$, necesitamos sistemas de polinomios cuyas soluciones están dadas en coordenadas homogéneas. Para ello, necesitamos considerar polinomios homogéneos, y los ideales que estos generan.

Definición 2.4. Un *ideal homogéneo* es un ideal $I \in \mathbb{K}[x_0, x_1, \ldots, x_n]$ que satisface cualquiera de las siguientes condiciones (que son equivalentes):

(i) I está generado por polinomios homogéneos.

(ii) Si $f \in I$, y escribimos $f = f_0 + \ldots + f_d$ como en (2.1), entonces $f_i \in I$.

Así, una variedad proyectiva es el conjunto de puntos en el espacio proyectivo en donde se anulan un conjunto de polinomios homogéneos dados. Los conjuntos algebraicos $\mathcal{V}(F) \subset \mathbb{A}^n$ definen los cerrados de una topología en $\mathbb{A}^n$, llamada la *topología de Zariski*. Esta topología no es Hausdorff, ya que los abiertos son muy grandes. Añade más propiedades de la topología de Zariski, como que los cerrados son o un conjunto finito, o unión de variedades. Que los abiertos fundamentales son densos. –Abraham

Definición 2.5. Dado un subconjunto $X \subset \mathbb{A}^n$ definimos el *ideal* de X como

$$\mathcal{I}(X) := \{f \in S : f(x) = 0 \text{ para cada } x \in X\}.$$

Veamos que hay una conexión entre variedades e ideales. Si $X \subset \mathbb{A}^n$ es un conjunto algebraico, se tiene que $X = \mathcal{V}(\mathcal{I}(X))$. Sin embargo, cuando X no es algebraico, entonces $X \subset \mathcal{V}(\mathcal{I}(X))$. De hecho, el conjunto $\overline{X} := \mathcal{V}(\mathcal{I}(X))$ es la *cerradura de Zariski* de X.

2.2. El Nullstellensatz

Al considerar ideales y variedades, las funciones $\mathcal{I}$ y $\mathcal{V}$ nos dan una correspondencia:

$$\{ \text{ ideales } I \subseteq S \} \underset{\mathcal{I}}{\overset{\mathcal{V}}{\rightleftarrows}} \{ \text{ variedades } X \subseteq \mathbb{A}^n \}.$$

Sin embargo, esta correspondencia **no** es una biyección. Por ejemplo, se tiene que $\mathcal{V}(x^2) = \mathcal{V}(x)$, pero $\mathcal{I}(\mathcal{V}(x^2)) = \langle x \rangle$. Notemos que una obstrucción es precisamente que los ceros del polinomio x y de x^2 son los mismos, pero la ecuación más pequeña que los define es x. Esto motiva la siguiente definición.

Definición 2.6. El *radical* de un ideal $I \subset S$ se define como

$$\sqrt{I} := \{f \in S : f^N \in I \text{ para algún } N \geq 1\}.$$

Si $I = \sqrt{I}$, decimos que I es un *ideal radical*.

Se puede ver de la definición que el conjunto $\sqrt{I}$ es de nuevo un ideal de S y es radical. Además, si $X \subset \mathbb{A}^n$ es cualquier subconjunto, el ideal $\mathcal{I}(X)$ es radical.

Teorema 2.7 (Nullstellensatz de Hilbert)**.** *Supongamos que $\mathbb{K}$ es algebraicamente cerrado, y sea $I \subseteq S$ un ideal. Se tiene que $\mathcal{I}(\mathcal{V}(I)) = \sqrt{I}$. En particular, $\mathcal{V}(I) = \varnothing$ si y sólo si $1 \in I$.*

En general es fácil ver que $\sqrt{I}$ contiene al ideal I, pero es difícil encontrar un método *efectivo* para calcular el radical $\sqrt{I}$ de un ideal I. Sin embargo, con esta definición podemos ya enunciar la correspondencia entre ideales y variedades.

Teorema 2.8. *Sobre cualquier campo $\mathbb{K}$ tenemos una correspondencia*

$$\{ \textit{ ideales radicales} I \subseteq S \} \underset{\mathcal{I}}{\overset{\mathcal{V}}{\rightleftarrows}} \{ \textit{ variedades } X \subseteq \mathbb{A}^n \}.$$

Esta correspondencia invierte inclusiones, y cumple que $\mathcal{V}(I(X)) = X$ para toda variedad $X \subset \mathbb{A}^n$. Si además $\mathbb{K} = \overline{\mathbb{K}}$, entonces esta correspondencia es biyectiva, en donde $\mathcal{I}$ y $\mathcal{V}$ son inversas.

Ahora veamos algunas de las propiedades que se pueden traducir a partir de esta correspondencia. Para ello, iniciamos definiendo la suma y el producto de dos ideales.

Definición 2.9. Dados dos ideales $I, J \subset S$, definimos la suma y el producto como

$$I + J := \{f + g \mid f \in I \text{ y } J \in J\}$$
$$IJ := \{f \cdot g \mid f \in I \text{ y } g \in J\}$$

Es un buen ejercicio demostrar que estos dos conjuntos son de nuevo ideales de S.

Proposición 2.10. *Sean $I, J \subset S$ dos ideales, y sean $X = \mathcal{V}(I), Y = \mathcal{V}(J)$ sus variedades correspondientes. Entonces*

i) $\mathcal{V}(I + J) = X \cap Y$

ii) $\mathcal{V}(IJ) = \mathcal{V}(I \cap J) = X \cup Y$

Si además $\mathbb{K} = \overline{\mathbb{K}}$ entonces

iii) $\mathcal{I}(X \cap Y) = \sqrt{I + J}$

iv) $\mathcal{I}(X \cup Y) = \sqrt{I \cap J} = \sqrt{IJ}$

Como resultado de esta última proposición, notamos que el ideal IJ y la intersección $I \cap J$ definen la misma variedad, y tienen el mismo radical. Sin embargo, estos dos ideales no coinciden, es decir, en general puede suceder que $IJ \neq I \cap J$. Para ver esto, consideremos $I = \langle xy - x^3 \rangle$ y $J = \langle y^2 - x^2 y \rangle$. Entonces, $I \cap J = \langle xy(y - x^2) \rangle$ mientras que $IJ = \langle xy(y - x^2)^2 \rangle$.

Observación 2.11. *Sea $m_0 = \langle x_0, x_1, \ldots, x_n \rangle \subset S = \mathbb{K}[x_0, x_1, \ldots, x_n]$. Entonces*

i) $\mathcal{V}(m_0) := \{x \in \mathbb{P}^n : x_i = 0 \text{ para cada } 1 \leq i \leq n\} = \varnothing$

ii) $\mathcal{V}(I \cdot m_0) = \mathcal{V}(I \cap m_0) = \mathcal{V}(I)$

A m_0 se le llama el ideal irrelevante de S. Para los demás ideales tenemos una correspondencia biyectiva

$$\{\text{ideales homogéneos radicales } I \subsetneq m_0\} \longleftrightarrow \{\text{Variedades proyectivas } X \subset \mathbb{P}^n\}$$

Proposición 2.12. *Sea $I = \langle f_1, f_2, \ldots, f_r \rangle \subset \mathbb{K}[x_1, \ldots, x_n]$ y $f \in \mathbb{K}[x_1, x_2, \ldots, x_n]$. Si $\langle f_1, \ldots, f_r, 1 - tf \rangle = \mathbb{K}[x_1, \ldots, x_n, t]$ entonces $f \in \sqrt{I}$.*

Demostración. Por hipótesis podemos escribir

$$1 = h_1 f_1 + h_2 f_2 + \ldots + h_r f_r + h_0 \cdot (1 - tf) \tag{2.2}$$

donde $h_1, \ldots, h_r, h_0 \in \mathbb{K}[x_1, \ldots, x_n, t]$. Para cada $i = 0, \ldots, r$, escribamos

$$h_i = \sum_{j=0}^{d_i} a_{ij}(x_1, \ldots, x_n) t^j,$$

donde $a_{ij}(x_1, \ldots, x_n) \in \mathbb{K}[x_1, \ldots, x_n]$, y los $d_0, \ldots, d_r \in \mathbb{N}$ son las potencias máximas en t de cada h_i. Consideremos un entero m que sea mayor a máx$(d_0, \ldots, d_r)$. En el campo de fracciones $\mathbb{K}(x_1, \ldots, x_n, t)$ podemos dividir la ecuación (2.2) por t^m para obtener

** Haz un pequeño paréntesis sobre el campo de fracciones, y quita la mención de aquí, para que quede más pura la idea de esta prueba –Abraham

$$\frac{1}{t^m} = \frac{h_1}{t^m} \cdot f_1 + \ldots + \frac{h_r}{t^m} \cdot f_r + \frac{h}{t^{m-1}} \left(\frac{1}{t} - f \right)$$

Sustituyendo $t = 1/f$ obtenemos

$$f^m = h_1' f_r + \ldots + h_r' f_r$$

con $h_i' = f^m h_i \left(x_1, \ldots, x_n, 1/f \right) \in \mathbb{K}[x_1, \ldots, x_n]$. Por lo tanto, $f^m \in \langle f_1, \ldots, f_r \rangle$.

□

Puedes terminar este capítulo con uno .ºpcional"sobre el anillo coordenado, su analogía con el espacio dual de un espacio vectorial, para terminar de cerrar la explicación sobre la geometría de los polinomios –Abraham

Ejercicios

Agrega aquí ejercicios buenos para agarrar callo en geometría.

1. Demuestra que es de equivalencia la relación en $\mathbb{K}^n$ definida por $x \backsim y$ si y solo si $x = ay$, para algún escalar $a \neq 0$.

2. Demuestra que cualesquiera dos rectas en $\mathbb{P}^2$ se intersectan.

3. Demuestra que los cambios de coordenadas en $\mathbb{P}^n$ son isormorfismos.

4. Demuestra la equivalencia de la definición de ideal homogéneo.

Capítulo 3

Teoría de eliminación

En Geometría Algebraica Computacional y en otras ramas de las matemáticas, una técnica común para simplificar argumentos en una demostración es el agregar variables temporales, y después tratar de regresar al anillo original, eliminando las variables que agregamos. Esta técnica se usa para calcular muchas cosas, como homogeneizaciones, parametrizaciones, y la usamos en la Proposición 2.12 para calcular el radical de un ideal.

En este capítulo desarrollaremos el marco teórico que justifica la eliminación de variables, además de desarrollar algunos algoritmos, y explicar la geometría detrás de ella. Iniciemos con el siguiente ejemplo.

Ejemplo 3.1. Sea $I = \langle ax + b, cx + d \rangle \subset \mathbb{K}[x, a, b, c, d]$ y consideremos el orden lexicográfico dado por $x > a > b > c > d$. Si calculamos el S-polinomio de estas dos ecuaciones, tenemos que

$$\text{Spol}(ax + b, cx + d) = c(ax + b) - a(cx + d) = bc - ad.$$

Por lo tanto, el polinomio $bc - ad$ está en el ideal I, y al no involucrar a la variable x, es también un elemento del anillo de polinomios $\mathbb{K}[a, b, c, d]$. Por lo tanto, tenemos una contención de ideales $\langle bc - ad \rangle \subseteq I \cap \mathbb{K}[a, b, c, d]$. Más adelante veremos que, de hecho, la igualdad se cumple para este caso. △

La teoría de eliminación trata de calcular ideales de la forma $I \cap \mathbb{K}[a, b, c, d]$ como los del ejemplo, que consiste de eliminar del ideal a todos aquellos polinomios que involucren a la variable x. Para desarrollar los algoritmos para calcular dicha intersección, primero tratemos de entender la geometría detrás de la eliminación.

3.1. Morfismos Inducidos

Supongamos que $\mathbb{K}$ es un campo infinito. Un polinomio $f \in \mathbb{K}[x_1, \ldots, x_n]$ define una función $\mathbb{A}^n \to \mathbb{K}$ dada por la evaluación $x \mapsto f(x)$. Así, tenemos que podemos interpretar a $\mathbb{K}[x_1, \ldots, x_n]$ como el anillo de todas las funciones polinomiales de $\mathbb{A}^n$ en $\mathbb{K}$

$$\mathbb{K}[x_1, \ldots, x_n] = \{f : \mathbb{A}^n \to \mathbb{K} \mid f \text{ es un polinomio}\}.$$

De esta manera, una colección $f_1, \ldots, f_m \in \mathbb{K}[x_1, \ldots, x_n]$ define una función polinomial

$$\begin{array}{rcl} \phi : \mathbb{A}^n & \longrightarrow & \mathbb{A}^m \\ a & \mapsto & (f_1(a), \ldots, f_m(a)). \end{array}$$

Nos gustaría entender o calcular la imagen $\operatorname{Im}\phi := \{\phi(a) : a \in \mathbb{A}^n\} \subseteq \mathbb{A}^m$. Si $\mathbb{K}[x_1, \ldots, x_n]$ y $\mathbb{K}[y_1, \ldots, y_m]$ son los anillos de funciones polinomiales de $\mathbb{A}^n$ y $\mathbb{A}^m$ en $\mathbb{K}$, respectivamente, entonces la función ϕ define una función

$$\begin{array}{rcl} \phi^* : \mathbb{K}[y_1, \ldots, y_m] & \longrightarrow & \mathbb{K}[x_1, \ldots, x_n] \\ y_i & \mapsto & f_i. \end{array}$$

Esta función ϕ^* es un homomorfismo de $\mathbb{K}$-álgebras, es decir, que preserva la estructura de $\mathbb{K}$-álgebra: por ejemplo, $\phi^*(2y_1 + y_2y_3)$ es igual a $2\phi^*(y_1) + \phi^*(y_2)\phi^*(y_3) = 2f_1 + f_2f_3$.

Lema 3.2. *Sea $\overline{\operatorname{Im}\phi}$ la cerradura de Zariski de la imagen $\operatorname{Im}\phi$. Entonces, $\mathcal{I}(\overline{\operatorname{Im}\phi}) = \ker\phi^*$.*

Demostración. Por definición, el ideal $\mathcal{I}(\overline{\operatorname{Im}\phi})$ consiste de todos los polinomios en $\mathbb{K}[y_1, \ldots, y_m]$ que se anulan a lo largo de $\operatorname{Im}\phi$. Si $g \in \mathbb{K}[y_1, \ldots, y_m]$, entonces $\phi^*g = g(f_1, \ldots, f_m)$. Por lo tanto, para todo $g \in \mathbb{K}[y_1, \ldots, y_m]$ y $a \in \mathbb{A}^m$, tenemos que

$$g(\phi(a)) = g(f_1(a), \ldots, f_m(a)) = (\phi^*g)(a).$$

Por lo tanto, si $g \in \ker\phi^*$ entonces g se anula a lo largo de $\operatorname{Im}\phi$ pues $\phi^*g = 0$. Conversamente, si g se anula en todo $\operatorname{Im}\phi$, entonces ϕ^*g se anula en cada punto $a \in \mathbb{A}^n$. Como $\mathbb{K}$ es un campo infinito, esto implica que $\phi^*g = 0$. □

3.2. Teoría de Eliminación

Proposición 3.3. *Sea $\alpha : \mathbb{K}[y_1, \ldots, y_m] \to \mathbb{K}[x_1, \ldots, x_n]$ el homomorfismo definido por $y_j \mapsto f_j$, con $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$. Definimos el ideal*

$$I := \langle y_1 - f_1, \ldots, y_m - f_m \rangle \subset \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m].$$

Entonces $\ker\alpha = I \cap \mathbb{K}[y_1, \ldots, y_m]$.

Observación 3.4. *Tenemos una proyección $\widetilde{\alpha} : \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m] \to \mathbb{K}[x_1, \ldots, x_n]$ dada por $y_j \mapsto f_j$ y $x_i \mapsto x_i$, la cual extiende a α en el sentido de que es un diagrama conmutativo. Además si $g(x_1, \ldots, x_n) \in \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$ entonces $\widetilde{\alpha}(g) = 0$ si y sólo si $g = 0$.*

Demostración de la Proposición 8.2. Si $f \in I \cap \mathbb{K}[y_1, \ldots, y_m]$ entonces podemos escribir

$$f = \sum_{i=1}^{m} h_i \cdot (y_i - f_i).$$

Por lo tanto, $\alpha(f) = \widetilde{\alpha}(f) = 0$; es decir, $f \in \ker \alpha$. Consideremos ahora el orden lexicográfico en $\mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$ dado por $y_1 > y_2 \ldots > y_m > x_1 > \ldots > x_n$. Sea $f \in \ker \alpha \subset \mathbb{K}[y_1, \ldots, y_m] \subset \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$. Si dividimos f entre los generadores de I, obtenemos una expresión de la forma

$$f = \sum_{i=1}^{m} h_i \cdot (y_i - f_i) + r,$$

donde $r = r(x_1, \ldots, x_n)$ no involucra a las y_i's. Por lo tanto

$$\begin{aligned} 0 &= \alpha(f) \\ &= \widetilde{\alpha}(f) \\ &= \widetilde{\alpha}(r) + \sum_{i=1}^{m} \widetilde{\alpha}(h_i) \cdot \widetilde{\alpha}(y_i - f_i) \\ &= \widetilde{\alpha}(r). \end{aligned}$$

De la observación vemos que $r = 0$, por lo que $f \in I$. □

Definición 3.5. Sea $I \subset S = \mathbb{K}[x_1, \ldots, x_n]$ un ideal de S. El k-ésimo *ideal de eliminación* de I es

$$I_k := I \cap \mathbb{K}[x_{k+1}, \ldots, x_n]$$

Más generalmente, si $R \subset S$ es un subanillo, el ideal $I \cap R$ se llama el *ideal de eliminación* de I con respecto a R.

Observación 3.6. *$I \cap R \subseteq R$ en efecto es un ideal.*

Teorema 3.7. *Sea G una base de Gröbner de I con respecto al orden lexicográfico en $\mathbb{K}[x_1, \ldots, x_n]$ dado por $x_1 > x_2 > \ldots > x_n$. Entonces $G_k := G \cap K[x_{k+1}, \ldots, x_n]$ es una base de Gröbner de I_k.*

Demostración. Por definición $G_k \subset I_k$. Sea $f \in I_k \subset I$. Entonces existe $g \in G$ tal que $in_< g$ divide a $in_< f$. Como $f \in I_k$, $in_< f \in \mathbb{K}[x_{k+1}, \ldots, x_n]$ y, por lo tanto, $in_< g \in \mathbb{K}[x_1, \ldots, x_n]$. Como el orden en $\mathbb{K}[x_1, \ldots, x_n]$ que hemos elegido es el lexicográfico, que $in_< g \in \mathbb{K}[x_{k+1}, \ldots, x_n]$ implica que todos los términos de g están en $\mathbb{K}[x_1, \ldots, x_n]$. Por lo tanto, $g \in G_k$. □

Ejemplo 3.8. Habíamos visto que si $I = \langle ax + b, cx + d \rangle \subset \mathbb{K}[x, a, b, c, d]$ entonces $G = \{ax + b, cx + d, ad - bc\}$ es una base de Gröbner de I con respecto al orden lexicográfico $x > a > b > c > d$. El teorema entonces implica que $\{ad - bc\}$ genera al ideal de eliminación $I \cap \mathbb{K}[a, b, c, d]$. △

Observación 3.9. *La propiedad del orden lexicográfico que usamos en la prueba del teorema fue: si $g \in S$ y $1 \leq k \leq n$ entonces $in_< g \in \mathbb{K}[x_{k+1}, \ldots, x_n]$ si y sólo si $g \in \mathbb{K}[x_{k+1}, \ldots, x_n]$.*

Definición 3.10. Sea $1 \leq k \leq n$. Un orden monomial en $\mathbb{K}[x_1, \ldots, x_n]$ se llama un *orden de eliminación* eliminando las primeras k variables, si cumple la propiedad de la observación anterior.

Algoritmo 3 (algoritmo de eliminación)

Entrada: $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ $i \in 1, 2, \ldots, n-1$.

Salida: Una base de Gröbner de I_i.

1: Sea G una base de Gröbner de I_i con respecto a el orden lexicográfico.

2: **devolver** $G \cap \mathbb{K}[x_{i+1}, \ldots, x_n]$.

Veamos ahora los algoritmos que están relacionado con los teoremas descritos en la clase anterior. Comenzaremos por el algoritmo que nos calcula una base de Gröbner del i-ésimo ideal de eliminación.

De esta forma podemos también generar un algoritmo para calcular el kernel de un homomorfismo entre $\mathbb{K}$-álgebras. Simplifiquemos la notación definiendo $S = \mathbb{K}[x_1, \ldots, x_n]$ y $R = \mathbb{K}[y_1, \ldots, y_m]$.

Algoritmo 4 (kernel de un homomorfismo de $\mathbb{K}-$álgebras)

Entrada: Una función de $\phi^* \colon R \to S$ dada for $f_1, f_2, \ldots, f_m \in S$.

Salida: Una base de Gröbner de $\ker \phi^*$.

1: Definamos $H := \langle y_1 - f_1, \ldots, y_m - f_m \rangle$ en $\mathbb{K}[y_1, y_2, \ldots, y_m, x_1, x_2, \ldots, x_n]$.

2: Calcule G la base de Gröbner de $H \cap R$ (utilizando el algoritmo 6).

3: **devolver** G.

Algoritmo 5 (imagen de un morfismo entre variedades)

Entrada: $f_1, \ldots, f_m \in S$ que definen un función $\phi \colon \mathbb{A}^n \to \mathbb{A}^m$.

Salida: Una base de Gröber para $\mathcal{I}(\overline{\mathrm{im}(\phi)})$.

1: Definamos $I_\phi := \{y_1 - f_1, \ldots, y_m - f_m\}$.

2: Calcule $G :=$ base de Gröbner de $I_\phi \cap R$.

3: **devolver** G.

Para $1 \leq i < n$ sea $\pi : \mathbb{A}^n \to \mathbb{A}^{n-i}$ la proyección que manda $(a_1, \ldots, a_n) \mapsto (a_{i+1}, \ldots, a_n)$.

Lema 3.11. *Si $I \subseteq \mathbb{K}[x_1, \ldots, x_m]$ es un ideal, entonces*

$$\pi(\mathcal{V}(I)) = \mathcal{V}(I_i)$$

donde $I_i := I \cap \mathbb{K}[x_{i+1}, \ldots, x_n]$.

Demostración. Sea $a = (a_1, \ldots, a_n) \in \mathcal{V}(I)$. Si $f \in I_i$ en particular $f \in I$. Entonces

$$0 = f(a) = f(a_{i+1}, \ldots, a_n) = f(\pi(a))$$

ya que f no contiene a ninguna de las $x_1, \ldots, x_i$. □

Anteriormente se vió que $\mathcal{V}(I \cup J) = \mathcal{V}(I) \cap \mathcal{V}(J)$ para cualesquiera I, J ideales. También vimos que

$$\mathcal{V}(IJ) = \mathcal{V}(I) \cup \mathcal{V}(J) = \mathcal{V}(I \cap J).$$

Y también vimos que IJ no necesariamente es igual a $I \cap J$. Entonces nos preguntamos: ¿cómo calculamos $I \cap J$?. El siguiente resultado nos ayudará a responder la pregunta anterior.
Sea t una nueva variable y considérese

$$N := \langle tI + (1-t)J \rangle \subseteq S[t].$$

Lema 3.12. *$I \cap J$ es el ideal de eliminación de $N \cap S$, es decir:*

$$I \cap J = (tI + (1-t)J) \cap S.$$

Demostración. Si $f \in I \cap J$, entonces $f = tf + (1-t)f \in N$. Inversamente, si $f \in N \cap S$, podemos escribir $f = f_I = f_J$, en donde $f_I = f_I(x,t) \in tI$ y $F_J = f_J(x,t) \in (1-t)J$. Observemos que f no involucra t pero f_I y f_J si lo hacen. Por lo tanto, podemos evalular f_I y f_J en cualquier valor y la igualdad $f = f_I + f_J$ se peserva, en particular, cuando $t = 0$,

$$f = f_I(x,0) + f_J(x,0) \in J.$$

y cuando $t = 1$, tenemos

$$f = f_I(x,1) + f_J(x,1) \in I.$$

Por lo tanto, $f \in I \cap J$. □

Definición 3.13. Un conjunto algebraico X es *irreducible* si cada vez que escribimos

$$X = V_1 \cup V_2, \quad V_1, V_2 \text{ son algebraicos.}$$

Entonces $X = V_1$ o $X = V_2$.

Definición 3.14. Sea R un anillo conmutativo con unidad. Un ideal $P \subseteq R$ es *primo* si, y sólo si, $P \neq R$ y cada $fg \in P$ cumplen que $f \in P$ o $g \in P$.

Ahora veamos una lema que nos relaciona las dos definiciones anteriores.

Lema 3.15. *X es irredusible si, y sólo si, $\mathcal{I}(X)$ es primo.*

Lema 3.16. *Si $X = X_1 \cup \cdots \cup X_r$ es una descomposición de X en irreducibles y sea*

$$P_i := \mathcal{I}(X_i)$$

entonces $\mathcal{I}(X) P_1 \cap \cdots \cap P_r$.

Problema. Dado un ideal I que define a X, y sea Z un subconjunto algebraico de X. Encuentre el ideal que define a $\overline{X \backslash Z}$.

Remover componentes está relacionado con la disivión: Si $gh^N \in \mathcal{I}(X)$, para cualquier N, entonces $g \in \mathcal{I}(X \setminus \mathcal{V}(h))$. ya que si $\forall n \geq 0$, $gn^N \in \mathcal{I}(X)$ entonces $g(x)h^N(x) = 0$, $\forall x \in X$.
Si consideramos $x \in X \setminus \mathcal{V}(h)$ entonces $g(x)h^N(x) = 0$ si, y sólo si, $g \in \mathcal{I}(X \setminus \mathcal{V}(h))$.

Definición 3.17. Sea $I, J \subseteq S$ ideales, $h \in S$. Definimos *el ideal cociente* (por h, se le llama "colon ideal" en Inglés). Se denota por:

$$(I : h) := \{g \in S \mid gh \in I\}.$$

Definimos también *la saturación* por h a el conjunto:

$$(I : h^\infty) := \{g \in S \mid gh^N \in I, \text{ para algún } N\}.$$

De la misma manera definimos el ideal conciente por J:

$$(I : J) := \{g \in S \mid gJ \in I\}.$$

La saturación con respecto a J:

$$(I : J^\infty) := \{g \in S \mid gJ^N \in I, \text{ para } N \text{ suficientemente grande}\}.$$

Ejemplo 3.18. Sea $S = \mathbb{K}[x, y, z]$. Por definición

$$\begin{aligned}(\langle xz, yz\rangle : \langle z\rangle) &= \{f \in S | fz \in \langle xz, yz\rangle\} \\ &= \{f \in S | fz = Axz + Byz\} \\ &= \{f \in S | f = Ax + By\} = \langle x, y\rangle.\end{aligned}$$

De aquí se puede ver que el por qué se le llama ideal cociente, pues se ha 'eliminado' la variable z. △

Observación 3.19. *Para ideales redicales, el cociente y la saturación coinciden, pero en general no es cierto.*

Ejemplo 3.20. $(\langle x^3, y\rangle : \langle x\rangle) = \langle x^2 y\rangle$ pero $(\langle x^3, y\rangle : \langle x\rangle^\infty) = \langle y\rangle$ △

Proposición 3.21. $(I : J)$ *es un ideal y contiene a* I

Demostración. Se deja como ejercicio para el lector. □

Teorema 3.22. *Si* I *y* J *son ideales en* S*, entonces*

$$\overline{\mathcal{V}(I) \setminus \mathcal{V}(J)} \subseteq \mathcal{V}(I : J)$$

si ademaás $\mathbb{K} = \overline{\mathbb{K}}$ *e* I *es radical, entonces se cumple la igualdad.*

Demostración. Notemos que $(I : J) \subset \mathcal{I}(\mathcal{V}(I) \setminus \mathcal{V}(J))$ ya que si $f \in (I : J)$ y $x \in \mathcal{V}(I) \setminus \mathcal{V}(J)$ por lo que $fg \in I, \forall g \in J$. Como $x \in \mathcal{V}(I)$, entonces $f(x)g(x) = 0, \forall g \in J$. Al tener que $x \in \mathcal{V}(J)$, $\exists h \in J$, tal que $h(x) \neq 0 \Rightarrow f(x) = 0$ lo que implica $\overline{\mathcal{V}(I) \setminus \mathcal{V}(J)} \subseteq \mathcal{V}(I : J)$.

Ahora supongamos que $\mathbb{K} = \overline{\mathbb{K}}$ y que $I = \sqrt{I}$. Sea $x \in \mathcal{V}(I : J)$, esto quiere decir que si $hg \in I$, $\forall g \in J$ entonces $h(x) = 0$. Sea $h \in \mathcal{I}(\mathcal{V}(I) \setminus \mathcal{V}(J))$, si $g \in J$, entonces hg se anula en $\mathcal{V}(I)$ y g se anula en $\mathcal{V}(J)$. Por el nullstellensatz, $hg \in \sqrt{I}$, pero por hipótesis $\sqrt{I} = I$, entonces $hg \in I, \forall g \in J$. Usando el mismo argunmento anterior tenemos que $h(x) = 0$. Por lo tanto $x \in \mathcal{V}(\mathcal{I}(\mathcal{V}(I) \setminus \mathcal{V}(J))) \Rightarrow \mathcal{V}(I : J) \subset \overline{\mathcal{V}(I) \setminus \mathcal{V}(J)}$. □

Lema 3.23. *Sean $P, Q \subset S$ ideales primos (diferentes). Para cualquier $f \in S$*

i) $(P : f) = \begin{cases} S & si\ f \in P \\ P & caso\ contrario \end{cases}$

ii) $(P \cap Q : f) = \begin{cases} S & si\ f \in P \cap Q \\ P & si\ f \in P \setminus Q \\ Q & si\ f \in Q \setminus P \\ I := P \cap Q & si\ f \in (P \cup Q)^c \end{cases}$

Demostración. **i)** Si $f \in P$, entonces $\forall g \in S$, tenemos que $fg \in P$, $\therefore (P : f) \subset P$. Sean $g \in (P : f)$, entonces $gf \in P$, y como P es ideal primo y $f \in P^c$ entonces $g \in P$. **ii)** el primer caso es es análogo al inciso i). Supongamos que $f \in Q \setminus P$, si $h \in P$, entonces $hf \in P$ y $hf \in Q$, lo que implica que $h \in (P \cap Q : f)$. Ahora supongamos que $h \in (P \cap Q : f) \Rightarrow hf \in P$ y $hf \in Q \Rightarrow h \in (P : f)$ y $h \in (Q : f)$, entonces por i) tenemos que

$$h \in (P : f) \cap (Q : f) = P \cap S = P$$

Por último, si $f \in (P \cup Q)^c$, sea $f \in (I : f)$, $\Rightarrow hf \in P$ y $hf \in Q$, por primalidad tenemos que $h \in P$ y $h \in Q$ por lo tanto $h \in I = P \cap Q$ □

Teorema 3.24. *Sea $X = \mathcal{V}(I)$ y $Y = \mathcal{V}(J)$, entonces $\mathcal{V}(I : J^\infty) = \overline{X \setminus Y}$*

Proposición 3.25. *Si $I \subset S$ es un ideal y $f \in S$. Definamos $J := \langle I, tf - 1\rangle \subset S[t]$, entonces $(I; f^\infty) = J \cap S$.*

Demostración. Sea $g \in J \cap S$, entonces podemos escribir $g = ph + q(1 - tf)$ con $h \in I$ y $p, q \in S[t]$. Sea m la máxima portencia de t en p. Usando el campo de fracciones $\mathcal{K}(x_1, \ldots, x_n, t)$, dividimos entre t^m y obtenemos

$$\frac{g}{t^m} = \frac{p}{t^m}h + \frac{q}{t^(m-1)}(\frac{1}{t} - f).$$

Si evaluamos en $t = \dfrac{1}{t}$ obtenemos $f^m g = f^m p' h$, donde $p' = p(x_1, \ldots, x_n, \dfrac{1}{f})$ y por elección de m, $f^m p' \in S$ entonces $f^m g \in I$ (pues $h \in S$) por lo que $g \in (I : f^\infty)$.

Ahora sea $g \in (I : f^\infty)$, entonces $\exists N \in \mathbb{N}$, tal que $gf^N \in I$. De la definición de J, podemos escribir $1 = tf + p$ para algún $p \in J$, entonces

$$1 = (tf)^N + p'$$

con $p' \in J$, por lo tanto, $g = (gt^N f^N + gp') \in J$ pues $gp' \in J$ y $gf^N \in J$, por lo tanto $g \in J \cap S$ □

Ejercicios

Agrega aquí ejercicios buenos para agarrar callo en geometría.

1. Demuestra algo

Capítulo 4

Resultantes

Para el caso de una variable, el teorema fundamental del álgebra, nos dice que $f \in \overline{\mathbb{K}}[x]$, entonces f tiene $d = deg(f)$ raíces y $f = \prod_{i=1}^{d}(x - \alpha_i)$. Si $f, g \in \overline{\mathbb{K}}[x]$ con $n = deg(g)$ y $m = deg(f)$, $n \geq m$. Podemos encontrar factores comunes entre f y g usando el algoritmo de Euclides (para encontrar MCD)

$$f = a_m x^m + a_{m-1}x^{m-1} + \ldots + a_0, \qquad g = b_n x^n + b_{n-1}x^{n-1} + \ldots + b_0$$

y remplazamos g por $g - \frac{b_n}{a_m}x^{n-m}f$ para poder escribir $MCD(f,g) = hf + kg$ OJO: n era el número de variables, ahora es el grado de g –Abraham

Ahora definamos $S_{<n}(x) = \{f \in \mathbb{K}[x] \mid deg(f) < n\}$, y de la misma manera $S_{\leqslant n}$ como los polinomios de grado a lo más n.

Observación 4.1. *$S_{\leqslant n} = S_{<n+1}$ es un espacio vectorial con una base ordenada canónica dada por los monomios $x^n,\ x^{n-1}, \ldots,\ x,\ 1$.*

Consideremos

$$S_{<n} \times S_{<m} := \{\, (h,k) \in \mathbb{K}[x] \times \mathbb{K}[x] \mid h = 0 \text{ o } \deg(h) < n \text{ y } k = 0 \text{ o } \deg(k) < m \,\}.$$

Dados f y g tales que $\deg(f) = m$ y $\deg(g) = n$ con $m \leq n$ podemos considerar la transformación lineal

$$\begin{aligned} \varphi_{f,g} : S_{<n} \times S_{<m} &\longrightarrow S_{<n+m} \\ (h,k) &\longmapsto h \cdot f + k \cdot g \end{aligned}$$

Notemos que la transformación está bien definida ya que $\deg(h \cdot f) < m + n$ y también $\deg(k \cdot g) < m + n$.

Observación 4.2. *Tenemos que $\dim_{\mathbb{K}} S_{<n+m} = \dim_{\mathbb{K}}(S_{<n} \times S_{<m}) = n + m$. Además, la transfomación $\varphi_{f,g}$ es un isomorfismo si y sólo si la matriz asociada a $\varphi_{f,g}$ es invertible.*

Lema 4.3. *Los polinomios f y g tienen un divisor común si y solo si $\ker \varphi_{f,g} \neq \{(0,0)\}$.*

Demostración. Supongamos que f y g tienen un divisor común p. Entonces existen polinomios h y k tales que $f = k \cdot p$ y $g = h \cdot p$.

Como p no es constante, entonces $\deg(k) < m = \deg(f)$ y $\deg(h) < n = \deg(g)$. De esto $(h, -k) \in S_{<n} \times S_{<m}$ y además

$$\varphi_{f,g}(h, -k) = f \cdot h - g \cdot k = (k \cdot p \cdot h) - (h \cdot p \cdot k) = 0$$

Por lo que $(h, -k) \neq (0, 0)$ y $(h, -k) \in \ker \varphi_{f,g}$.

Ahora supongamos que f y g son primos relativos, es decir $MCD(f, g) = 1$. Esto implica que existen polinomios p y q tales que $1 = f \cdot p + g \cdot q$, de modo qe $\langle f, g \rangle = \mathbb{K}[x]$.
Sean $(h, k) \in \ker \varphi_{f,g}$, entonces $0 = f \cdot h + g \cdot k$ y $g \cdot k = -f \cdot h$. Luego, podemos reescribir

$$\begin{aligned} k = k \cdot 1 &= f \cdot p \cdot k + g \cdot q \cdot k \\ &= f \cdot k \cdot p - f \cdot h \cdot q \\ &= f(k \cdot p - h \cdot q). \end{aligned}$$

De donde $f|k$. Como $\deg(k) \leq m - 1$ y $\deg(f) = m$ entonces k =0. De manera similar, podemos ver que $h = 0$, lo que implica que $\ker \varphi_{f,g} = \{(0, 0)\}$. □

Sean $B = (x^{n-1}, \ldots, x, 1, x^{m-1}, \ldots, x, 1)$ y $B' = \{x^{m+n-1}, \ldots 1\}$ las bases ordenadas canónicas de $S_{<n} \times S_{n<m}$ y $S_{<n+m}$ respectivamente. La matriz asociada de $\varphi_{f,g}$ respecto a las bases B y B' en vectores columna es

$$\left[\varphi_{f,g}\right]_B^{B'} = ([x^{n-1} \cdot f]_{B'}, [x^{n-2} \cdot f]_{B'}, \ldots, [f]_{B'}, [x^{m-1} \cdot g]_{B'}, \ldots, [g]_{B'}).$$

Si $f = a_m x^m + \ldots + a_1 x + a_0$ y $g = b_m x^n + \ldots + b_1 x + b_0$, se obtiene la siguiente matriz:

$$\left[\varphi_{f,g}\right]_B^{B'} = \left(\begin{array}{cccc|cccc} a_m & 0 & \ldots & 0 & b_n & 0 & \ldots & 0 \\ a_{m-1} & a_m & \ddots & \vdots & b_{n-1} & b_n & \ddots & 0 \\ \vdots & a_{m-1} & \ddots & 0 & \vdots & b_{n-1} & \ddots & 0 \\ \vdots & \vdots & \ddots & a_m & \vdots & \vdots & \ddots & b_n \\ a_0 & a_1 & \ddots & a_{m-1} & b_0 & b_1 & \ddots & b_{n-1} \\ 0 & a_0 & \ddots & \vdots & 0 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 & \vdots & \ddots & \ddots & b_1 \\ 0 & \ldots & 0 & a_0 & 0 & \ldots & 0 & b_0 \end{array}\right) := \mathrm{Syl}(f, g) \in \mathbb{K}^{(m+n)\times(m+n)}$$

La matriz $\mathrm{Syl}(f, g)$ suele llamarse la *matriz de Sylvester* de f y g, por medio de la cual se puede definir la resultante.

Definición 4.4. Sean $f = a_m x^m + \ldots + a_1 x + a_0$ y $g = b_m x^n + \ldots + b_1 x + b_0$, con $a_m \neq 0$ y $b_n \neq 0$. La *resultate* de Sylvester de f y g es

$$\mathrm{Res}(f, g) := \det(\mathrm{Syl}(f, g)).$$

Ejemplo 4.5. Sean $f = x - \alpha$ y $g = x - \beta$, entonces.

$$\mathrm{Res}(f,g) = \det \begin{pmatrix} 1 & 1 \\ -\alpha & -\beta \end{pmatrix} = \alpha - \beta.$$

Por lo tanto f y g tiene raices comunes si $\alpha = \beta$. △

Ejemplo 4.6. Sean $f = x - \alpha$ y $g = b_3x^3 + b_2x^2 + b_1x + b_0$. Se tiene que

$$\mathrm{Res}(f,g) = \det \left(\begin{array}{ccc|c} 1 & 0 & 0 & b_3 \\ -\alpha & 1 & 0 & b_2 \\ 0 & -\alpha & 1 & b_1 \\ 0 & 0 & -\alpha & b_0 \end{array} \right) = b_3\alpha^3 + b_2\alpha^2 + b_1\alpha - b_0.$$

△

Observación 4.7. *En general si $f = x - \alpha$ y $g = b_nx^n + \cdots + b_1x + b_0$ entonces*

$$\mathrm{Res}(f,g) = \det \left(\begin{array}{ccc|c} 1 & 0 & 0 & b_n \\ -\alpha & \ddots & 0 & \vdots \\ 0 & \ddots & 1 & b_1 \\ 0 & 0 & -\alpha & b_0 \end{array} \right) = g(\alpha)$$

y en todos estos casos $\mathrm{Res}(f,g) = 0$ si y sólo si α es una raíz de f y g.

Teorema 4.8. *Sean f y $g \in \mathbb{K}[x]$ entonces $\mathrm{Res}(f,g) = 0$ si y sólo si $\deg((MCD(f,g)) \geq 1$, si y sólo si existe $\alpha \in \overline{\mathbb{K}}$ tal que $f(\alpha) = g(\alpha) = 0$.*

Demostración. La segunda equivalencia se deriva del teorema fundamental del álgebra. La primera se deriva del lema anterior ya que $\mathrm{Res}(f,g) = 0$ si y sólo si $\varphi_{f,g}$ no es isomorfismo. □

Proposición 4.9. *Dados $f, g \in \mathbb{K}[x]$ existen polinomios $h, k \in \mathbb{K}[x]$, no ambos nulos con $\deg(h) < n$, $\deg(k) < m$, tales que $\mathrm{Res}(f,g) = f \cdot h + g \cdot k$ tienen coeficientes enteros en los coeficientes de f y g, esto es $\mathrm{Res}(f,g) = f \cdot h + g \cdot k \in \mathbb{Z}[a_m, \cdots, a_0, b_n, \cdots, b_0][x]$.*

Demostración. Cuando $\mathrm{Res}(f,g) = 0$, ya vimos que existen h, $k \in \mathbb{K}[x]$ tales que

$$\mathrm{Res}(f,g) = 0 = h \cdot f + g \cdot k,$$

que satisfacen las condiciones requeridas. Cuando $\mathrm{Res}(f,g) \neq 0$ entonces $\varphi_{f,g}$ es isomorfismo, por tanto epimorfismo, y de esto, existen h y k con las condiciones requeridas. Podemos

encontrar h y k explícitamente

$$\left(\begin{array}{cccc|cccc} a_m & 0 & \dots & 0 & b_n & 0 & \dots & 0 \\ a_{m-1} & a_m & \ddots & \vdots & b_{n-1} & b_n & \ddots & 0 \\ \vdots & a_{m-1} & \ddots & 0 & \vdots & b_{n-1} & \ddots & 0 \\ \vdots & \vdots & \ddots & a_m & \vdots & \vdots & \ddots & b_n \\ a_0 & a_1 & \ddots & a_{m-1} & b_0 & b_1 & \ddots & b_{n-1} \\ 0 & a_0 & \ddots & \vdots & 0 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 & \vdots & \ddots & \ddots & b_1 \\ x^{n-1}f & \dots & xf & f & x^{m-1}g & \dots & xg & g \end{array}\right) = \mathrm{Syl}(f,g)$$

(el determinante no cambia si al último renglón le sumamos x por el penúltimo, más x^2 por el antepenúltimo, así sucesivamente hasta x^{m+n-1} por el primero).

El último renglón lo podemos escribir como $(x^{n-1}f,\dots,f,0,\dots,0)+(0,\dots,0,xg^{m-1},\dots,g)$ con esto, si expandimos el determinante por el último renglón, lo podemos ver como la suma de dos determinantes

$$\det\left(\begin{array}{cccc|cccc} a_m & 0 & \dots & 0 & b_n & 0 & \dots & 0 \\ a_{m-1} & a_m & \ddots & \vdots & b_{n-1} & b_n & \ddots & 0 \\ \vdots & a_{m-1} & \ddots & 0 & \vdots & b_{n-1} & \ddots & 0 \\ \vdots & \vdots & \ddots & a_m & \vdots & \vdots & \ddots & b_n \\ a_0 & a_1 & \ddots & a_{m-1} & b_0 & b_1 & \ddots & b_{n-1} \\ 0 & a_0 & \ddots & \vdots & 0 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 & \vdots & \ddots & \ddots & b_1 \\ x^{n-1}f & \dots & xf & f & 0 & \dots & 0 & 0 \end{array}\right) = h$$

y análogamente para k. La igualdad con $\mathrm{Res}(f,g)$ se cumple cuando desarrollamos todos los determinantes por la última fila.

□

Proposición 4.10. *Supongamos que $\mathbb{K} = \overline{\mathbb{K}}$, escribimos a $f(x) = a_m\prod_{i=1}^{m}(x-\alpha_i)$, y $g(x) = b_n\prod_{j=1}^{n}(x-\beta_j)$, donde las α_i y β_j son las raíces de f y g respectivamente, entonces:*

$$\mathrm{Res}(f,g) = (-1)^{mn}a_m^n b_n^m \prod_{i=1}^{m}\prod_{j=1}^{n}(\alpha_i-\beta_j). \tag{4.1}$$

Un corolario inmediato es la fórmula de la Poisson:

$$\mathrm{Res}(f,g) \;=\; a_m^n\prod_{i=1}^{m} g(\alpha_i) \;=\; (-1)^{mn}b_n^m\prod_{j=1}^{n} f(\beta_j) \tag{4.2}$$

Demostración. Notemos que (4.1) se sigue de la fórmula de Poisson (4.2), por lo tanto sólo mostraremos la de Poisson. Mostraremos el caso en que f sólo tiene raíces simples, i.e. $\alpha_i \neq \alpha_j$ si $i \neq j$. Recordemos que la matriz de Vandermonde es

$$\mathrm{Vand}(\alpha_1, \ldots, \alpha_m) = \det \begin{pmatrix} \alpha_1^{m-1} & \alpha_2^{m-1} & \ldots & \alpha_m^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1 & \alpha_2 & \ldots & \alpha_m \\ 1 & 1 & \ldots & 1 \end{pmatrix} = \prod_{1 \leq i < j \leq m} (\alpha_i - \alpha_j) \tag{4.3}$$

que no es cero cuando $\alpha_i \neq \alpha_j$. Así, considerando el producto de matrices

$$A = \left[\begin{array}{cccc|cccc} & & I_{n\times n} & & & & 0_{n\times m} & \\ \hline \alpha_1^{m+n-1} & \alpha_1^{m+n-2} & \ldots & \alpha_1^m & \alpha_1^{m-1} & \alpha_1^{m-2} & \ldots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_m^{m+n-1} & \alpha_m^{m+n-2} & \ldots & \alpha_m^m & \alpha_m^{m-1} & \alpha_m^{m-2} & \ldots & 1 \end{array}\right] \left(\begin{array}{ccc|ccc} a_m & \ldots & 0 & b_n & \ldots & 0 \\ a_{m-1} & & \vdots & b_{n-1} & & \vdots \\ \vdots & & 0 & \vdots & & 0 \\ \vdots & \ddots & a_m & \vdots & \ddots & b_n \\ a_0 & & a_{m-1} & b_0 & & b_{n-1} \\ 0 & & \vdots & 0 & & \vdots \\ \vdots & & a_1 & \vdots & & b_1 \\ 0 & \ldots & a_0 & 0 & \ldots & b_0 \end{array}\right)$$

tenemos que

$$A = \left[\begin{array}{ccc|cccc} a_m & 0 & & & & & \\ \vdots & \ddots & & & \text{una matriz aquí} & & \\ * & \ldots & a_m & & & & \\ \hline & & & \alpha_1^{m-1} g(\alpha_1) & \alpha_1^{m-2} g(\alpha_1) & \ldots & g(\alpha_1) \\ & 0 & & \vdots & \vdots & \ddots & \vdots \\ & & & \alpha_m^{m-1} g(\alpha_m) & \alpha_m^{m-2} g(\alpha_m) & \ldots & g(\alpha_m) \end{array}\right]$$

Calculando determinantes, dado que tenemos matrices triangulares por bloques, nos queda

$$\mathrm{Vand}(\alpha_1, \ldots, \alpha_m)\,\mathrm{Res}(f, g) \;=\; a_m^n g(\alpha_1) \ldots g(\alpha_m) \mathrm{Vand}(\alpha_1, \ldots, \alpha_m) \tag{4.4}$$

como $\mathrm{Vand}(\alpha_1, \ldots, \alpha_m) \neq 0$ cuando $\alpha_i \neq \alpha_j$, podemos cancelar. Lo cual concluye la demostración. El caso general se obtiene haciendo la misma construcción con una matriz de Vandermonde generalizada. □

La fórmula de Poisson tiene una consecuencia inmediata respecto al algoritmo de la división: Si escribimos $f = qg + r$ con $l = \deg(r) < n = \deg(g)$ entonces

$$\mathrm{Res}(f, g) \;=\; (-1)^{mn} b_n^{m-l}\,\mathrm{Res}(g, r)$$

porque las β_i son raíces de g, con lo que $f(\beta_i) = r(\beta_i)\ \forall \beta_i$, lo que implica

$$\begin{aligned}\operatorname{Res}(f,g) &= (-1)^{mn} b_n^m \prod_{j=1}^{n} f(\beta_j) \\ &= (-1)^{mn} b_n^{m-l} b_n^l \prod_{j=1}^{n} r(\beta_j) \\ &= (-1)^{mn} b_n^{m-l} \operatorname{Res}(g,r)\end{aligned}$$

Recordemos que f tiene una raíz múltiple si y sólo si f y f' tienen una raíz común, es decir, cuando $\operatorname{Res}(f,f') = 0$.

Ejemplo 4.11. $f = ax^2 + bx + c \implies f' = 2ax + b$

$$\operatorname{Res}(f,f') = \det \begin{pmatrix} a & 2a & 0 \\ b & b & 2a \\ c & 0 & b \end{pmatrix} = ab^2 + 4a^2c - 2ab^2 = a(4ac - b^2),$$

por lo que, la ecuación cuadrática f tiene una raíz doble cuando $4ac - b^2 = 0$ (ya que si $a = 0$ entonces f no es cuadrática). △

Definición 4.12. Para $f = a_m x^m + \ldots + a_1 x + a_0 = a_m \prod_{i=1}^{m} (x - \alpha_i)$, el *discriminante* es

$$\operatorname{Disc}(f) := \frac{(-1)^{\frac{m(m-1)}{2}}}{a_m} \operatorname{Res}(f,f') = a_m^{2m-2} \prod_{1 \le i \le j \le m} (\alpha_i - \alpha_j)^2.$$

La última igualdad sale de la fórmula de Poisson, ya que

$$f' = a_m \sum_{1 \le i \le m} \prod_{j \ne i} (x - \alpha_j).$$

Observación 4.13. $\operatorname{Disc}(f) = 0 \iff f$ *tiene una raíz múltiple.*

Supongamos $1 \le i \le n$ y sea $\pi : \mathbb{A}^n \to \mathbb{A}^{n-i}$ la proyección $(a_1, \ldots, a_n) \mapsto (a_{i+1}, \ldots, a_n)$. Vimos:

Lema 4.14. *Si $I \subset \mathbb{K}[x_1, \ldots, x_n]$ es un ideal, entonces $\pi(\mathcal{V}V(I)) \subseteq \mathcal{V}(I_i)$ con I_i siendo el ideal de eliminación $I \cap \mathbb{K}[x_{i+1}, \ldots, x_n]$.*

La inclusión $\pi(\mathcal{V}(I)) \subseteq \mathcal{V}(I_i)$ puede ser estricta, como se muestra en el siguiente ejemplo.

Ejemplo 4.15. Sea $\pi : \mathbb{A}^2 \to \mathbb{A}^1$ la proyección $(x,y) \mapsto y$. Consideremos $\mathcal{V}(xy - 1)$, entonces $\pi(\mathcal{V}(xy-1)) = \mathbb{A}^1 - 0$ donde $0 = \langle xy - 1 \rangle \cap \mathbb{K}[y]$.

△

Ejemplo 4.16. Consideremos la curva $\varphi : \mathbb{A} \to \mathbb{A}^2$ dada por $t \mapsto (t^2 - 1, t^3 - t)$. ¿Quién es $im(\varphi)$? Consideremos $C = \mathcal{V}(x - t^2 + 1, y - t^3 + t)$ bajo la proyección $\pi(t,x,y) = (x,y)$.
Del lema anterior sabemos que la ecuación de esta curva es $\langle t^2 - 1 - x, t^3 - t - y \rangle \cap \mathbb{K}[x,y]$.

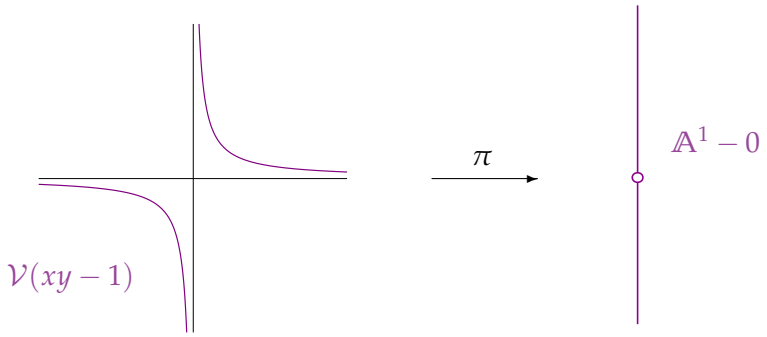


Figura 4.1: La proyección de $V(xy-1)$ en el eje y.

Sea $f(t) = t^2 - 1 - x$ y $g(t) = t^3 - t - y$.

$$\mathrm{Res}(f,g;t) = \begin{vmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ -1-x & 0 & 1 & -1 & 0 \\ 0 & -1-x & 0 & -y & -1 \\ 0 & 0 & -1-x & 0 & -y \end{vmatrix} = y^2 + x^2 - x^3,$$

que es la ecuación que define a $\pi(C)$. △

Observemos que $\mathbb{K}[x,y] \subset \mathbb{K}(y)[x]$, y para $f,g \in \mathbb{K}[x,y]$, la resultante $\mathrm{Res}(f,g;x)$ es un polinomio en y que es cero si y sólo si f y g tienen un factor común en $\mathbb{K}(y)[x]$.

Lema 4.17 (Gauss)**.** *Los polinomios $f,g \in \mathbb{K}[x,y]$ tienen un factor común de grado positivo en x si y sólo si tienen un factor común en $\mathbb{K}(y)[x]$.*

Demostración. $\Rightarrow$) Es claro, ya que $\mathbb{K}[x,y] \subseteq \mathbb{K}(y)[x]$.

$\Leftarrow$) Supongamos que $f = h\bar{f}$ y $g = h\bar{g}$ es una factorización en $\mathbb{K}(y)[x]$ con h de grado positivo en x. Sea d el mínimo común múltiplo de todos los denominadores en $h, \bar{f}, \bar{g}$, entonces $d \in \mathbb{K}(y)$ y podemos escribir $d^2 f = (dh)(d\bar{f})$ y $d^2 g = (dh)(d\bar{g})$ donde $dh, d\bar{f}, d\bar{g} \in \mathbb{K}[x,y]$. Sea $k(x,y)$ un factor irreducible de dh con grado positivo en $x \Rightarrow k | d^2 f$ y $k | d^2 g$, pero k no divide a d porque $d \in \mathbb{K}[y]$ y k tiene grado postivo en x. Por lo tanto $k | f$ y $k | g$. □

Sea $\pi : \mathbb{A}^2 \to \mathbb{A}$ la proyección $(x,y) \mapsto y$. Sea $I_{f,g} := \langle f,g \rangle \cap \mathbb{K}[y]$, habíamos visto:

Proposición 4.18. *Dados $f,g \in \mathbb{K}[x]$, existen $h,k \in \mathbb{K}[x]$ tales que $\mathrm{Res}(f,g) = f(x)h(x) + g(x)k(x)$ con $\mathrm{Res}(f,g) \in \mathbb{Z}[a_0,\ldots,b_n][x]$. Así, $\mathrm{Res}(f,g;x) \subseteq I_{f,g}$*

Lema 4.19. *Si I_i es el ideal de eliminación de I, entonces $\pi(\mathcal{V}(I)) \subseteq \mathcal{V}(I_i)$. Por lo tanto*

$$\pi(\mathcal{V}(f,g)) \subseteq \mathcal{V}(I_{f,g}) \subset \mathcal{V}(\mathrm{Res}(f,g;x)).$$

Ahora, supongamos $\mathbb{K} = \overline{\mathbb{K}}$ y escribamos para $f,g \in \mathbb{K}[x,y]$:

$$\begin{aligned} f &= a_m(y)x^m + a_{m-1}(y)x^{m-1} + \ldots + a_1(y)x + a_0(y), \\ g &= b_n(y)x^n + b_{n-1}(y)x^{n-1} + \ldots + b_1(y)x + b_0(y). \end{aligned}$$

Teorema 4.20 (Teorema de Extensión). *Si* $\mathbb{K} = \overline{\mathbb{K}}$ *y* $t \in \mathcal{V}(I_{f,g}) - \mathcal{V}(a_m(y), b_n(y))$, *entonces existe* $s \in \mathbb{K}$ *tal que* $(s,t) \in \mathcal{V}(f,g)$.

Observación 4.21. *Con esto demostramos que*

$$\mathcal{V}(I_{f,g}) \setminus \mathcal{V}(a_m, b_n) \subseteq \pi(\mathcal{V}(f,g)) \subseteq \mathcal{V}(I_{f,g}) \subseteq \mathcal{V}(\text{Res}(f,g;x)).$$

Demostración. Sea $t \in \mathcal{V}(I_{f,g}) \setminus \mathcal{V}(a_m, b_n)$ y supongamos que $a_m(t) \cdot b_n(t) \neq 0$. Entonces $f(x,t)$ y $g(x,t)$ son polinomios en x de grado m y n respectivamente. Nótese que $\text{Syl}(f(x,t), g(x,t))$ se puede obtener $\text{Syl}(f,g;x)$ substituyendo $y = t$. Esto implica que $\text{Res}(f(x,t), g(x,t))$ es igual a evaluar $\text{Res}(f,g;x)$ en $y = t$.

Como $\text{Res}(f,g;x) \in I_{f,g}$ y $t \in \mathcal{V}(I_{f,g})$, entonces esta evaluación es 0. Como $\mathbb{K} = \overline{\mathbb{K}}$, entonces $f(x,t)$ y $g(x,t)$ tienen una raíz común, digamos $s \in \mathbb{K}$. Pero esto implica que $(s,t) \in \mathcal{V}(f,g)$, lo que implica que $t \in \pi(\mathcal{V}(f,g))$.

Si $a_m(t) \neq 0$ pero $b_n(t) = 0$, como $\langle f, g \rangle = \langle f, g + x^l f \rangle$, si reemplazamos g por $g + x^l f$ con $l + m > n$, entonces caemos en el caso anterior. □

Corolario 4.22. *Si* $a_m(y), b_n(y)$ *son constantes, entonces* $\mathcal{V}(I_{f,g}) = \mathcal{V}(\text{Res}(f,g;x))$.

Lema 4.23. *Si* $\mathbb{K} = \overline{\mathbb{K}}$, *entoces el sistema de ecuaciónes*

$$f(x,y) = g(x,y) = 0$$

tiene un número finito de soluciónes en $\mathbb{A}^2$ *si y sólo si* f *y* g *no tienen un factor común que no sea constante.*

Demostración. Probaremos que $\mathcal{V}(f,g)$ es infinito si y sólo si f y g tienen un factor común que no es constante.

$\Leftarrow$) Si f, g tienen un factor común no constante, digamos $h(x,y)$, entonces los ceros $\mathcal{V}(f,g)$ incluyen a los ceros $\mathcal{V}(h)$ que es infinito, ya que h no es constante y $\mathbb{K} = \overline{\mathbb{K}}$.

$\Rightarrow$) Supongamos que $\mathcal{V}(f,g)$ es infinito. Entonces la proyección a alguno de los ejes coordenados es infinito.

Supongamos que la proyección π al eje y es infinito. Sea $I_{f,g} = \langle f, g \rangle \cap \mathbb{K}[y]$ el ideal de eliminación. Por el teorema de extensión, tenemos

$$\pi(\mathcal{V}(f,g)) \subset \mathcal{V}(I_{f,g}) \subset \mathcal{V}(\text{Res}(f,g;x))$$

y como $\pi(\mathcal{V}(f,g)) = \mathbb{A}^1$, entonces $\mathcal{V}(\text{Res}(f,g;x)) = \mathbb{A}^1$; por lo tanto, $\text{Res}(f,g;x) = 0$. Es decir, f, g tienen un factor común no constante. □

Con esta prueba, podemos destacar la siguiente observación.

Observación 4.24. *Si* $f, g \in \mathbb{K}[x,y]$ *cumplen que tanto* $\text{Res}(f,g;x)$, *como* $\text{Res}(f,g;y)$ *son distintas de* 0. *Entonces* f, g *no tienen factor comunes no constantes. Por lo que,* $\mathcal{V}(f,g)$ *consiste de un número finito de puntos.*

Si $\mathcal{V}(f,g)$ consiste de un número finito de puntos, entonces $\mathcal{V}(f,g)$ es una variedad de dimensión, como veremos a continuación. Para este caso, tenemos el siguiente algoritmo de eliminación. La idea de este algoritmo es reducir el resolver un sistema de ecuaciones en 2 variables, a encontrar las raíces de un polinomio en 1 variable. Una variante de este algoritmo puede usarse para encontrar todas las raíces de un sistema de ecuaciones polinomiales en cualquier número de variables, si es que el sistema tiene un número finito de soluciones.

Algoritmo 6 (algoritmo de eliminación)

Entrada: Polinomios $f, g \in \mathbb{K}[x,y]$ tales que $\mathcal{V}(f,g)$ es finito.
Salida: $\mathcal{V}(f,g)$

1: Calcula $\text{Res}(f,g;x)$ (que es un polinomio en y distinto de 0.
2: Calcula el conjunto B de todas las raíces de $\text{Res}(f,g;x)$.
3: Para cada $t \in B$ calcula las raíces comúnes s de $f(x,t)$ y $g(x,t)$.
4: **devolver** todas las parejas (s,t)

Ejercicios

Agrega aquí ejercicios buenos para agarrar callo en geometría.

1. Demuestra algo

Capítulo 5

Ideales de dimensión cero

Dado un sistema de ecuaciones polinomiales

$$f_1(x_1, \ldots, x_n) = \cdots = f_m(x_1, \ldots, x_n) = 0, \tag{5.1}$$

quisiéramos poder *entender* sus soluciones, es decir, quisiéramos contestar alguna de las siguientes preguntas:

Pregunta 5.1. Dado un sistema de ecuaciones polinomiales como en (5.1).

(i) ¿El sistema tiene un número finito de soluciones?

(ii) Si no, ¿podemos entender las soluciones aisladas?

(iii) ¿Podemos contar o dar una (buena) cota superior en el número de soluciones?

(iv) ¿Podemos resolver y encontrar todas las soluciones complejas?

(v) Si los polinomios tienen coeficientes reales, ¿podemos contar, acotar, o encontrar las soluciones reales?

Para poder contestar estas preguntas, nos enfocaremos ahora a aquellos sistemas de polinomios que tienen solo un número finito de soluciones, es decir, aquellos polinomios cuya variedad es un número finito de puntos, y les damos un nombre especial.

Definición 5.2. Un ideal $I \subseteq S = \mathbb{K}[x_1, \ldots, x_n]$ es de *dimensión cero* o *0-dimensional*, si sobre la cerradura $\mathbb{K}$, $\mathcal{V}(I)$ es una cantidad finita de puntos, i.e. la variedad $\mathcal{V}(I)$ es de dimensión 0.

Como la variedad de I es la misma que la de su radical, tenemos la siguiente observación.

Observación 5.3. *Notemos que I es 0-dimensional si y sólo si $\sqrt{I}$ es 0-dimensional.*

Pero, dado un ideal I, ¿cómo podemos identificar si I es 0-dimensional? Empecemos con el siguiente resultado.

Teorema 5.4. *Un ideal $I \subseteq S$ es 0-dimensional si y solo si S/I es un espacio vectorial sobre $\mathbb{K}$ de dimensión finita.*

Demostración. Podemos suponer que $\mathbb{K}$ es algebraicamente cerrado, ya que esto no cambia la dimensión del cociente S/I.

Supongamos primero que $I = \sqrt{I}$, y denotemos por $V = \mathcal{V}(I)$. Geométricamente, el cociente S/I es el anillo coordenado $\mathbb{K}[V]$, es decir, es el conjunto de todas las funciones que se obtienen de restringir polinomios a V:

$$\mathbb{K}[V] := \{f_{|_V} : V \to \mathbb{K} \mid f \in S\}$$

Si V es finito entonces $\mathbb{K}[V]$ es de dimensión finita (como espacio vectorial), pues el espacio de funciones de V tiene dimensión igual al número de puntos en V. De hecho, si $p_1, \ldots, p_r$ son los puntos de V, entonces una base de $\mathbb{K}[V]$ está dada por las funciones $x_i - p_{ji}$, es decir, por los ideales máximos definidos por cada punto. Agrega el apéndice de Anillo Coordenado después del diccionario algebro-geométrico, para incluir este tipo de observaciones y analogías con el espacio dual de un espacio vectorial. –Abraham

Si V es infinito, entonces existe una coordenada, digamos x_1, para la cual, la imagen de V bajo la proyección $\pi : (x_1, \ldots, x_n) \mapsto x_1$, consiste de una infinidad de puntos, y por lo tanto, es denso (complemento de un número finito de puntos). Restringiendo las funciones de $\mathbb{K}[x_1]$ a V, nos da una inyección

$$\mathbb{K}[x_1] \hookrightarrow \mathbb{K}[V], \qquad \text{definida por } f \mapsto f_{|_V}.$$

Esta inclusión muestra que $\mathbb{K}[V]$ es de dimensión infinita, pues $\mathbb{K}[x_1]$ lo es.

Ahora supongamos que I es cualquier ideal, no necesariamente radical. Si $\dim(S/I) < \infty$, entonces también se tiene que $\dim(S/\sqrt{I}) < \infty$. Esto es cierto, pues el hecho de que $I \subseteq \sqrt{I}$ implica que existe una función suprayectiva

$$S/I \to S/\sqrt{I} \qquad \text{definida por } \; s + I \mapsto s + \sqrt{I}.$$

Supongamos que $\dim(S/\sqrt{I}) < \infty$. De esto se sigue que para cada variable x_i, hay una combinación lineal de las potencias $1, x_i, x_i^2, \ldots$ que es 0 en $S/\sqrt{I}$ y por tanto está dentro de $\sqrt{I}$; o sea, existe un polinomio en una variable $g_i(x_i) \in \sqrt{I}$, y por tanto existe $M_i \in \mathbb{N}$ tal que $g_i^{M_i} \in I$ para cada x_i.

Por lo tanto $\langle g_1^{M_1}, \ldots, g_n^{M_n} \rangle \subseteq I$, la función $S/\langle g_1^{M_1}, \ldots, g_n^{M_n} \rangle \to S/I$ es suprayectiva. Pero el cociente $S/\langle g_1^{M_1}, \ldots, g_n^{M_n} \rangle$ tiene dimensión $M_1 \cdot \ldots \cdot M_n$. Esto implica que $\dim(S/I) < \infty$. □

En la prueba anterior construimos polinomios en una variable, que nos ayudaron a demostrar que S/I es de dimensión finita. El resultado general de esta construcción es el siguiente.

Corolario 5.5. *Un ideal $I \subseteq S$ es 0-dimensional si y sólo si para cada variable x_i, existe $g_i \in \mathbb{K}[x_i]$ tal que $g_i \in I$.*

Más aún, del Teorema 5.4 junto con el Teorema de Macaulay (Teorema 1.40), se desprende el siguiente resultado, que nos da un criterio para contestar la Pregunta 5.1.(i).

Corolario 5.6. *I es 0-dimensional si y sólo si para cualquier orden monomial $\succ$, el ideal $in_\succ I$ contiene una potencia de cada variable.*

Definición 5.7. Si I es 0-dimensional, definimos *el grado de I* como $\mathrm{gr}(I) := \dim(S/I)$. De otra manera, el grado de un ideal es el número de monomios estándar respecto de cualquier orden monomial.

Lema 5.8. *Sea I un ideal de dimensión 0, sea $G = \{g_1, \ldots, g_m\}$ su base de Gröbner reducida respecto de $\prec_{lex}$. Ordenamos variables y polinomios, de forma que $x_1 \succ x_2 \succ \cdots \succ x_n$, y que $in_{\prec_{lex}}(g_1) \succ in_{\prec_{lex}}(g_2) \succ \cdots \succ in_{\prec_{lex}}(g_m)$. Entonces, para cada $i \in \{1, \ldots, n\}$ existe $j \in \{1, \ldots, m\}$ tal que $in_{\prec}(g_j) = x_i^{d_i}$ para alguna $d_i > 0$.*

Demostración. El Corolario 5.5 nos garantiza la existencia de polinomios en una variable cuyo término inicial es una potencia pura de x_i. El resto se sigue de las propiedades de $\succ_{lex}$. □

Observemos que de este resultado, si encontramos esta base de Gröbner lexicográfica, podemos resolver el sistema de atrás para adelante, resolviendo primero el polinomio g_m, sustituyendo en g_{m-1} y resolver, y así sucesivamente hasta resolver g_1. Así, tenemos la siguiente cota.

Corolario 5.9. *Con la notación anterior, $|\mathcal{V}(I)| \leq d_1 \cdot \ldots \cdot d_n$*

Este resultado nos ayuda a contestar parcialmente la Pregunta 5.1.(iii), de encontrar cotas para el número de soluciones. Sin embargo, con lo desarrollado hasta ahora, podemos probar el siguiente caso, que es más general.

Teorema 5.10. *Si I es un ideal 0-dimensional, entonces $|\mathcal{V}(I)| \leq \mathrm{gr}(I)$. Si además se cumple que $I = \sqrt{I}$ y $\mathbb{K} = \overline{\mathbb{K}}$, entonces se satisface la igualdad.*

Agregar una definición o una explicación intuitiva de qué es un polinomio genérico. – Abraham

Teorema 5.11 (Teorema de Bézout; caso particular)**.** *Dos polinomios $f, g \in \mathbb{K}[x, y]$ o bien tienen un factor común, o $|\mathcal{V}(f, g)| \leq \mathrm{gr}(f) \cdot \mathrm{gr}(g)$. Esta cota se alcanza cuando $\mathbb{K} = \overline{\mathbb{K}}$ y los polinomios f y g son genéricos.*

Demostración. Supongamos que $m = \mathrm{gr}(f)$ y $n = \mathrm{gr}(g)$ y que f y g son primos relativos. Entonces $V(f, g)$ es finita. Si extendemos nuestro campo a $\overline{\mathbb{K}}$, podemos hacer un cambio de coordenadas remplazando a f por $f(A(x, y))$ y a g por $g(A(x, y))$ con A una transformación lineal invertible.

Escojamos A tal que $A(x, y) = (ax + by + c, \alpha x + \beta y + \gamma)$ con $a\beta - \alpha b \neq 0$ y con $a, b, c, \alpha, \beta, \gamma \in \overline{\mathbb{K}}$. Podemos escoger parámetros de tal manera que f y g tengan coeficientes no nulos en el término inicial y en el constante. En ese caso, la resultante $\mathrm{Res}(f, g; x)$ es un polinomio en y de grado a lo más mn y por tanto tiene a lo más mn raíces.

Si definimos $I_y := \langle f, g \rangle \cap \overline{\mathbb{K}}[y]$, entonces $\mathcal{V}(I) = \mathcal{V}(\mathrm{Res}(f, g; x))$. Podemos además escoger parámetros de A para que la proyección $\pi : (x, y) \to y$ sea uno a uno en $\mathcal{V}(f, g)$, que es finita. Por ende: $|\pi(\mathcal{V}(f, g))| = |\mathcal{V}(I)| = |\mathcal{V}(\mathrm{Res}(f, g))|$. Lo anterior resultante da la desigualdad del teorema, ya que $|\mathcal{V}(\mathrm{Res}(f, g; x)| \leq mn$.

Si $\overline{\mathbb{K}}$ es algebraicamente cerrado y $\text{Res}(f,g)$ tiene menos raíces que mn, entonces es porque tiene menor grado o porque tiene raíces múltiples. En el primer caso, su coeficiente inicial es cero; en el segundo, su determinante se anula. Por lo tanto el conjunto de parejas (f,g) de polinomios para los que $\mathcal{V}(f,g)$ tiene mn puntos en $\mathbb{A}^2$ es un abierto no vacío en el espacio $\mathbb{A}^{\binom{m+2}{2}+\binom{n+2}{2}}$ que consiste en las parejas de polinomios (f,g) de grados m y n. □

Definición 5.12. Sean $f_1,\dots,f_n$ polinomios en $\mathbb{K}[x_1,\dots,x_n]$, una solución a del sistema formado por tales polinomios es *no degenerada* si los vectores $\nabla f_i = \left(\frac{\partial f_i}{\partial x_1},\dots,\frac{\partial f_i}{\partial x_n}\right)$ son linealmente independientes en a.

Teorema 5.13 (Bézout)**.** *Dados $f_1,\dots,f_n \in \mathbb{K}[x_1,\dots,x_n]$ con $\text{gr}(f_i) = d_i$, el número de soluciones no degeneradas del sistema es a lo más $d_1\cdots d_n$. Cuando el campo es algebraicamente cerrado esta es una cota en el número de soluciones aisladas y se alcanza cuando los polinomios son genéricos.*

Agrega la explicación, al menos intuitiva, de qué es una solución aislada. Al producto de los grados $d_1\cdots d_n$ se le llama *cota de Bézout*.

5.1. Eliminantes

Definición 5.14. Sea $I \subset \mathbb{K}[x_1,\dots,x_n]$ un ideal. Un polinomio, $g(x_i)$, en una variable es un *eliminante* de I si

$$I \cap \mathbb{K}[x_i] = \langle g(x_i)\rangle$$

Teorema 5.15. *Supongamos que $g(x_i)$ es un eliminante de I, entonces $g(a_i) = 0$ para toda $a = (a_1,\dots,a_n) \in \mathcal{V}(I) \subset \mathbb{A}^n$. Cuando $\mathbb{K} = \overline{\mathbb{K}}$, todas las raíces de g son de esta forma.*

Demostración. Si $a = (a_1,\dots,a_n) \in \mathcal{V}(I)$, entonces $g(a_i) = 0$, pues ese es el valor de g en a. Supongamos que $\mathbb{K} = \overline{\mathbb{K}}$ y que ξ es una raíz de g, pero que no existe $a \in \mathcal{V}(I)$ cuya i-ésima coordenada es ξ.
Sea $h(x_i)$ un polinomio cuyas raíces son las raíces de g que sí provienen de $\mathcal{V}(I)$. Entonces $h(\xi) \neq 0$, pero h es cero en $\mathcal{V}(I)$, por lo que $h \in \sqrt{I}$ lo que implica que existe N tal que $h^N \in I$ y con ello

$$h^N \in I \cap \mathbb{K}[x_i] = \langle g(x_i)\rangle$$

lo cual es una contradicción ya que $h(\xi)^N \neq 0$ pero $g(\xi) = 0$ □

Observación 5.16. *La hipótesis de que $\mathbb{K}$ sea algebraicamente cerrado es necesaria. Consideremos $\mathbb{K} = \mathbb{R}$ y se $I = \langle x-1, y^2+1\rangle$, entonces $\mathcal{V}(I) = \emptyset$, pero $x-1$ es un eliminante.*

Teorema 5.17. *Si $g(x_i)$ es un eliminante mónico de I, entonces g es parte de cualquier base de Gröbner reducida de I con respecto a un orden de eliminación en el que x_i es la variable más pequeña.*

Demostración. Supongamos que $\succ$ es un orden de eliminación para el cual x_i es la mínima variable y sea $d = \text{gr}(g)$. Como $I \cap \mathbb{K}[x_i] = \langle g\rangle$, g es un polinomio mónico de grado mínimo en x_i que está en I. En particular x_i^d es uno de los generadores de $in(I)$, por lo tanto existe f en la base de Gröbner reducida cuyo término inicial es x_i^d y sus otros términos son monomios

estándar más chicos. Dado que el orden es de eliminación y x_i es la última variable, estos monomios estándar son de la forma x_i^l con $l < d$ por lo que $f \in I$ es un polinomio en x_i con grado d, de la unicidad de g, se sigue que $f = g$. □

Lema 5.18 (Lema de la forma)**.** *Supongamos que g es un eliminante de un ideal, I, de dimensión 0, con $\mathrm{gr}(g) = \mathrm{gr}(I)$. Entonces*

- *I es radical si y sólo si g no tiene raíces múltiples. Si además $g = g(x_n)$ y se tiene el orden lexicográfico $x_1 > \cdots > x_n$, entonces I tiene una base de Gröbner de la forma*

$$\{x_1 - g_1(x_n), \ldots, x_{n-1} - g_{n-1}(x_n), g(x_n)\}$$

  donde $\mathrm{gr}(g_i) < \mathrm{gr}(g)\ \forall i \in \{1, \ldots, n-1\}$.

- *Si I está generado por polinomios reales, el número de raíces reales de I es igual al de g.*

Demostración. Tenemos las siguientes desigualdades:

$$\#\ raices\ de\ g \leq \#\ de\ soluciones\ de\ I \leq \mathrm{gr}(I) = \mathrm{gr}(g)$$

Si las raíces de g son todas distintas, entonces g tiene $d = \mathrm{gr}(g)$ raíces, por lo que todas las desigualdades anteriores se convierten en igualdades, lo que implica que I es radical, gracias al Teorema 5.10. Recíprocamente, si $g = g(x_i)$ tiene raíces múltiples, entonces existe un polinomio h con las mismas raíces de g pero de grado menor. Como $I \cap \mathbb{K}[x_i] = \langle g(x_i) \rangle$, se tiene que $h \notin I$, pero h^d es divisible por g lo que implica que $h^d \in I$ y con ello I no es radical.

Si $g = g(x_n)$, entonces $1, x_n, \ldots, x_n^{d-1}$ son monomios estándar, como $d = \mathrm{gr}(I) = dim_{\mathbb{K}}(S/I)$, éstos son todos los monomios estándar, por lo que el ideal inicial con $\succ_{lex}$ es $\langle x_1, \ldots, x_{n-1}, x_n^d \rangle$. Cada elemento de la base de Gröbner reducida de I expresa a cada generador de I como una combinación lineal sobre $\mathbb{K}$ de monomios estándar, por lo que la base de Gröbner reducida tiene la forma deseada.

Finalmente, observemos que lo ceros en común de los polinomios son

$$\{(a_1, \ldots, a_n) | g(a_n) = 0, a_i = g_i(a_n)\ \forall i = 1, .., n-1\}$$

además los polinomios g_i son todos reales y una componente a_i es real si la raíz de $g(x_n)$ es real. □

Observación 5.19. *No todos los ideales tienen una base de Gröbner de este tipo, por ejemplo*

$$\langle x, y \rangle^2 = \langle x^2, xy, y^2 \rangle.$$

Sin embargo, la condición clave en el eliminante $(\mathrm{gr}(g) = \mathrm{gr}(I))$ generalmente se cumple después de un cambio de coordenadas genérico.

Algoritmo 7 (cálculo del eliminante)

Entrada: Ideal $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ y una variable x_i.
Salida: Un eliminante $g(x_i) \in I$ o un certificado de que no existe tal.

1: Calcular una base de Gröbner G con respecto a cualquier orden.
2: **si** ninguna potencia de x_i es término inicial de algún $g \in G$ **entonces**
3: **devolver** "No hay eliminante"
4: **si no**
5: Calcular $[1] := 1$ mód G
6: Define $indep :=$ **cierto** y $r := 0$
7: **mientras** $indep =$ **cierto** **hacer**
8: Calcular $[x_i^r] := x_i^r$ mód G
9: **si** $[1], [x_i], \ldots, [x_i^r]$ son linealmente dependientes **entonces**
10: $indep =$ **falso**
11: Calcular la combinación lineal

$$\sum_{j=0}^{r} a_j [x_j^i] = 0, \qquad \text{con } a_j \in \mathbb{K}$$

12: **si no**
13: Redefine $j = j + 1$
14: **fin si**
15: **fin mientras**
16: **devolver** El eliminante $g(x_i) = \sum_{j=0}^{r} a_j x_i^j$.
17: **fin si**

Prueba de validez. Si I no tiene un eliminante entonces $I \cap \mathbb{K}[x_i] = \{0\}$. Entonces $1, x_i, x_i^2, \ldots$ son todos monomios estándar y ninguna base de Gröbner G contiene un polinomio cuyo término inicial sea una potencia de x_i. Por lo tanto el algoritmo identifica correctamente cuando I no tiene eliminante.
Si I tiene un eliminante $g(x_i)$ con g mód $G = 0$ entonces G debe tener un polinomio cuyo término inicial divida a $\text{in}(g) = x_i^{\text{gr}(g)}$ y por tanto es una potencia pura de x_i.
Si $g = \sum_{j=0}^{N} b_j x_i^j$ entonces

$$\begin{aligned} 0 = g \text{ mód } G &= \left(\sum_{j=0}^{N} b_j x_i^j \right) \text{ mód } G \\ &= \sum_{j=0}^{N} b_j [x_i^j \text{ mód } G] \end{aligned}$$

Que es un dependencia lineal entre los elementos x_i^j mód G, por lo que el algoritmo termina. La minimalidad del grado de g implica $r = N$ y la unicidad de la combinación lineal mínima implica que los coeficientes b_j y a_j son proporcionales. Por lo tanto el algoritmo calcula un múltiplo escalar de g, que también es un eliminante. □

Ahora veamos un algoritmo que nos ayudará a realizar cambios de bases (de Gröbner).

Algoritmo 8 (FGLM)

Entrada: Una base de Gröbner G respecto a $<'$ de un ideal I de dimensión 0, y otro orden monomial $<$.
Salida: Una base de Gröbner H de I con respecto a el orden monomial $<$.
1: Declarar $H := \{\}$, $x^\alpha := 1$, $E := \{\}$, $inH := \{\}$.
2: Calcular $[x^\alpha] := x^\alpha$ mód G (una combinación lineal de monomios estándar de G).
3: **si** $[x^\alpha] \notin span(E)(\subseteq \mathbb{K}\{\text{monomios estándar}\})$ **entonces**
4: Redefine $E = E \cup \{[x^\alpha]\}$.
5: **si no**
6: Existe una combinación lineal (única) de elementos de E tales que

$$[x^\alpha] = \sum_{[x^\beta]\in E} c_\beta [x^\beta].$$

7: Redefine

$$H = H \cup \{x^\alpha - \sum_\beta c_\beta x^\beta\} \quad \text{y} \quad inH = inH \cup \{x^\alpha\}.$$

8: **fin si**
9: **si** $\{x^\gamma | x^\gamma > x^\alpha\} \subseteq \langle \text{in}(H)\rangle$ **entonces**
10: **devolver** H.
11: **si no**
12: Definir $x^\alpha = \text{mín}_<\{x^\gamma \notin \langle \text{in}(H)\rangle | x^\alpha < x^\gamma\}$.
13: **fin si**
14: Regresa al paso 2.

Prueba de validez. Por construcción, H siempre consiste de elementos de I, y los elementos de E son siempre linealmente independientes en S/I, por lo que $\text{in}(H) \subseteq \text{in}_<(I)$ y se cumplen las desigualdades

$$|E| \leq dim_{\mathbb{K}} S/I, \qquad \text{in}(H) \subseteq \text{in}_<(I).$$

Cuando regresamos al paso (2) o bien E creció, o bien H (y por lo tanto $\text{in}(H)$) creció. Como $|E|$ está acotado y los ideales monomiales generados por las distintas $\text{in}(H)$ forman una cadena estrictamente creciente, por lo que el algoritmo termina.

Cuando el algoritmo termina, todos lo monomios o bien están en $\langle \text{in}(H)\rangle$ o bien en $ME := \{x^\beta | [x^\beta] \in E\}$. Por nuestra elección de x^α en el paso (4), estos dos conjuntos son distintos, por lo que ME son los monomios estándar de $\text{in}(H)$. Como $\langle \text{in}(H)\rangle \subseteq \text{in}_<(H) \subseteq \text{in}_<(I)$ y los elementos de E son linealmente independientes modulo $\text{in}_<(I)$. Entonces

$$|E| \leq \dim_{\mathbb{K}} S/\text{in}_<(I) \leq \dim_{\mathbb{K}} S/\text{in}_<(H) \leq \dim S/\text{in}(H) = |E|.$$

Por lo tanto $\text{in}_<(I) = \text{in}(H)$, lo que prueba que H es una base de Gröbner respecto a $<$. Por la forma de los elementos de H, la base anterior es una base reducida. □

5.2. Métodos con autovalores

Supondremos $\mathbb{K}$ algebraicamente cerrado, e $I = \langle f_1, \ldots, f_N\rangle \subseteq \mathbb{K}[x_1, \ldots, x_n]$ un ideal de dimensión cero.

Definición 5.20. Dado un polinomio mónico en una variable

$$p = \sum_{i=0}^{d} c_i x^d \in \mathbb{K}[x], \text{ con } c_d = 1$$

la *matriz compañera de p* es

$$C_p := \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -c_0 & -c_1 & \cdots & -c_{d-1} \end{pmatrix} \in \mathbb{K}^{d\times d}$$

Si $p \in \mathbb{K}[x]$ es mónico y de grado d, entonces $\mathbb{K}[x]/\langle p\rangle$ es un espacio vectorial con base $\{1, x, x^2, \ldots, x^{d-1}\}$. La multiplicación por x nos da una transformación lineal

$$\begin{array}{rcl} M_x : \mathbb{K}[x]/\langle p\rangle & \longrightarrow & \mathbb{K}[x]/\langle p\rangle \\ [f(x)] & \mapsto & [x \cdot f(x)]. \end{array}$$

La matriz que representa M_x respecto a esta base es precisamente C_p. Ahora, dada una matriz $A \in \mathbb{K}^{d\times d}$, los eigenvalores de A son las raíces del polinomio característico $\chi_A(x) = \det(A - xI_d)$.

Teorema 5.21. *El polinomio característico de la matriz compañera C_p de un polinomio $p \in \mathbb{K}[x]$ de grado d es $\det(C_p - xI_d) = (-1)^d p(x)$ (i.e. las raíces de p son los eigenvalores de C_p).*

Demostración. Cuando $d = 1$, tenemos que $p = c_0 + x$, por lo que $C_p = [-c_0]$ y tenemos que

$$\begin{aligned} C_p - xI_d &= -c_0 - xI_d \\ &= (-1)(c_0 + x) = (-1)p. \end{aligned}$$

Para $d > 1$ expandamos el determinante $\det(C_p - xId)$ respecto de la primera columna:

$$\begin{aligned} \det(C_p - xI_d) &= \det \begin{pmatrix} -x & 1 & \cdots & 0 \\ 0 & -x & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -c_0 & -c_1 & \cdots & -c_{d-1} - x \end{pmatrix} \\ &= -x \cdot \det(C_q - xI_d) + (-1)^{d+1}(-c_0) \end{aligned}$$

con C_q siendo la matriz compañera de $q = x^{d-1} + \sum_{i=0}^{d-2} c_{i+1}x^i$. Por la hipótesis de inducción tenemos

$$\begin{aligned} \det(C_p - xId) &= -x(-1)^{d-1}q(x) + (-1)^{d+1}(-c_0) \\ &= (-1)^d xq(x) + (-1)^d c_0 = (-1)^d p(x). \end{aligned}$$

□

Corolario 5.22. *Las siguientes afirmaciones sobre p son equivalentes:*

(i) $p(x)$ *es libre de cuadrados (i.e. no tiene raíces múltiples en* $\overline{\mathbb{K}}$*).*

(ii) La matriz compañera es diagonalizable

(iii) el ideal $\langle p \rangle$ *es radical en* $\mathbb{K}[x]$

Demostración. Las raíces múltiples de $p(x)$ se pueden calcular con $q(x) = MCD(p(x), p'(x))$. Las tres condiciones son equivalentes a $q(x) = 1$, observando que $\sqrt{\langle p(x) \rangle} = \langle p(x)/q(x) \rangle$. □

Sea $I = \langle f_1, \ldots, f_N \rangle$ un ideal 0-dimensional. Tenemos que S/I es un espacio vectorial de dimensión finita y que $|V(I)|$ está acotado por $\dim_{\mathbb{K}} S/I$.

Si denotamos por $[f]$ a la clase de f en S/I, entonces para cualquier $i \in \{1, \ldots, n\}$, la multiplicación por x_i define un endomorfismo:

$$\begin{aligned} m_i : S/I &\longrightarrow S/I \\ [f] &\mapsto [x_i][f] = [x_i f]. \end{aligned}$$

Sea B el conjunto de monomios estándar de I y sean $M_1, \ldots, M_n \in \mathbb{K}^{|B| \times |B|}$ las matrices que representan los endomorfismos $m_1, .., m_n$ respecto de B.

Para $x^\alpha, x^\beta \in B$, la entrada de M_i en el renglón x^α y la columna x^β es el coeficiente de x^α en la forma normal del polinomio $x_i x^\beta \mod I$.

Lema 5.23. *Las matrices compañeras conmutan por pares, es decir,*

$$M_i \cdot M_j = M_j \cdot M_i \quad \textit{para } 1 \leq i < j \leq n.$$

Demostración. Las matrices $M_i M_j$ y $M_j M_i$ representan las composiciones $m_i \circ m_j$ y $m_j \circ m_i$ respectivamente. Como la multiplicación en S/I es conmutativa, se sigue el resultado. □

Observación 5.24. *Las matrices M_i generan una subálgebra conmutativa del anillo no conmutativo de matrices, considerando*

$$\mathbb{K}[M_i, \ldots, M_n] \simeq S/I \quad \textit{vía } M_i \mapsto [x_i].$$

Definición 5.25. Sea V un espacio vectorial y f un endomorfismo de V. Dado un polinomio

$$p(t) = \sum_{i=0}^{d} c_i t^i \in \mathbb{K}[t]$$

definimos el polinomio $p(f) := \sum_{i=0}^{d} c_i f^i$ con f^i denotando la composición $\overbrace{f \circ f \circ \cdots \circ f}^{i \text{ veces}}$.
Al ideal $I_f := \{p \in \mathbb{K}[t] | p(f) = 0\}$ se le llama el *ideal de f*. El *polinomio mínimo de f* es el único polinomio mónico h tal que $I_f = \langle h \rangle$ y lo denotamos por h_f.

Lema 5.26. *Sea V un espacio vectorial de dimensión finita sobre $\mathbb{K}$ un campo algebraicamente cerrado y f un endomorfismo de V. Para cada $\lambda \in \mathbb{K}$, λ es un eigenvalor de f si y sólo si λ es un cero de h_f.*

Demostración. Vamos a probar que el polinomio mínimo h_f divide al polinomio característico χ_f y que χ_f divide a h_f^k con $k = \dim V$ y esto implicará el lema. La primera parte se sigue del teorema de Cayley-Hamilton, que dice que cada endomorfismo es un cero de su polinomio característico.

Para la segunda parte, observemos que χ_f y h_f se descomponen sobre $\mathbb{K}$ en factores lineales. Si A_f es la matriz que representa a f, entonces

$$\chi_f = \det(A_f - tI_n) = \pm(t-\lambda_1)^{d_i} \dots (t-\lambda_m)^{d_m}$$

con $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ y $d_1, \dots, d_m \in \mathbb{N}$. Como ya vimos que $h_f | \chi_f$, entonces el polinomio mínimo h_f debe ser de la forma

$$\prod_{i=1}^{m} (t-\lambda_i)^{e_i} \quad \text{con } 0 \leq e_i \leq d_i.$$

Basta mostrar que $e_i \geq 1$ para toda $i \in \{1, \dots, m\}$. Supongamos que $e_i = 0$ para alguna i, y sea v un eigenvector de λ_i. Para cada eigenvalor $\lambda_j \neq \lambda_i$ se tiene que

$$(A_f - \lambda_j I_n)v = (\lambda_i - \lambda_j)v \neq 0.$$

Entonces si aplicamos la matriz $h_f(A_f)$ al vector v, tenemos:

$$\begin{aligned} h_f(A_f)v &= \prod_{j \neq 1} (A_f - \lambda_j I_n)^{e_j} v \\ &= \prod_{j \neq i} (\lambda_i - \lambda_j)^{e_j} v \ \neq \ 0, \end{aligned}$$

pero esto contradice que h_f es el polinomio mínimo. De lo anterior se sigue el resultado. □

Observación 5.27 (Recordatorio)**.** *Sea V un espacio vectorial de dimensión finita sobre un campo algebraicamente cerrado $\mathbb{K}$, y sea $f : V \to V$ un $\mathbb{K}$-endomorfismo de V. Entonces $\lambda \in \mathbb{K}$ es un eigenvalor de f si y sólo si λ es raíz del polinomio mínimo h_f de f.*

Teorema 5.28 (Stickelberg)**.** *Sea $\mathbb{K}$ un campo algebraicamente cerrado, y sea $I \subset S = \mathbb{K}[x_1, \dots, x_n]$ un ideal 0-dimensional. Entonces, para cada $i \in \{1, 2, \dots, n\}$, $\lambda \in \mathbb{K}$ es un eigenvalor del endomorfismo*

$$\begin{aligned} m_i : S/I &\longrightarrow S/I \\ [f] &\mapsto [x_i \cdot f] \end{aligned}$$

si y sólo si existe $a = (a_1, a_2, \dots, a_n) \in \mathcal{V}(I)$ tal que $a_i = \lambda$.

Demostración. Sea λ un eigenvalor de $m_i : S/I \to S/I$ y sea $[v] \in S/I \setminus \{0\}$ un eigenvector de m_i asociado a λ. Entonces $[x_i v] = [\lambda v]$, así que $[(x_i - \lambda)v] = 0$ en S/I.

Supongamos, a modo de contradicción, que la segunda condición del teorema no se cumple; i.e., para cada $a \in \mathcal{V}(I)$, $a_i \neq \lambda$. Obtendremos una contradicción si demostramos que $[x_i - \lambda]$ tiene un inverso $[y]$ en S/I, pues de este modo se tendría que

$$0 = [y] \cdot [(x_i - \lambda)v] = [v] \neq 0$$

Asociamos a cada punto $a \in \mathcal{V}(I)$ un polinomio $g_a \in S$ con la propiedad de que, para cada $b \in \mathcal{V}(I)$, se cumple la identidad

$$g_a(b) = \begin{cases} 1 & \text{si } a = b \\ 0 & \text{si } a \neq b \end{cases}$$

Estos polinomios siempre existen. En efecto, supongamos que la primera coordenada de todos los puntos en $\mathcal{V}(I)$ es distinta. Entonces podemos definir explícitamente

$$g_a = g_a(x_1) = \frac{\prod_{b \in \mathcal{V}(I)\setminus\{a\}} x_1 - b_1}{\prod_{b \in \mathcal{V}(I)\setminus\{a\}} a_1 - b_1}$$

Como $\mathbb{K}$ es infinito, siempre podemos encontrar un cambio de coordenadas en el que esta condición suceda, y por lo tanto los g_a siempre existen. Sea

$$g^* = \sum_{a \in \mathcal{V}(I)} \frac{1}{a_i - \lambda} \cdot g_a$$

Entonces, para cada $a \in \mathcal{V}(I)$

$$\begin{aligned}(a_i - \lambda)g^*(a) &= (a_i - \lambda) \sum_{b \in \mathcal{V}(I)} \frac{1}{b_i - \lambda} \cdot g_b(a) \\ &= g_a(a) + (a_i - \lambda) \sum_{b \in \mathcal{V}(I)\setminus\{a\}} \frac{1}{b_i - \lambda} \cdot g_b(a) \\ &= 1\end{aligned}$$

En otras palabras, $1 - (x_i - \lambda)g^*$ se anula en todo $a \in \mathcal{V}(I)$. Por el Nullstellensatz, dado que $\mathbb{K}$ es algebraicamente cerrado, existe $l \geq 1$ tal que $(1 - (x_i - \lambda)g^*)^l \in I$. Expandiendo el binomio obtenemos

$$1 + \left(\sum_{k=1}^{l} (-1)^k \binom{l}{k} (x_i - \lambda)^k (g^*)^k \right) \in I.$$

Así, en el cociente S/I tenemos la siguiente identidad

$$1 = \left(\sum_{k=1}^{l} (-1)^{k+1} \binom{l}{k} (x_i - \lambda)^{k-1} (g^*)^k \right) \cdot (x_i - \lambda)$$

Por lo tanto, $x_i - \lambda$ tiene un inverso en S/I. Ésta es la contradicción que buscábamos.

Supongamos ahora que existe $a \in \mathcal{V}(I)$ tal que $a_i = \lambda$. Por el lema, basta demostrar que $h_i(\lambda) = 0$, donde h_i es el polinomio mínimo de m_i.

Por definición del polinomio mínimo de un endomorfismo, $h_i(m_i) = 0$ como endomorfismo de S/I. Evaluando en $[1] \in S/I$ vemos que

$$0 = h_i(m_i)[1] = h_i([x_i])$$

Por lo tanto, $h_i(x_i) \in I \subset S$, de modo que $h_i(x_i)$ se anula en todo $\mathcal{V}(I)$; en particular, $h_i(\lambda) = h_i(a_i) = 0$. □

Definición 5.29. Las matrices compañeras M_i de los endomorfismos m_i son *simultáneamente diagonalizables* si existen una matriz invertible $N \in \mathbb{K}^{n\times n}$ y matrices diagonales $D_1, D_2, \ldots, D_n \in \mathbb{K}^{n\times n}$ tales que $M_i N = N D_i$ para cada $i \in \{1, 2, \ldots, n\}$.

Teorema 5.30. *Supongamos que I es un ideal 0-dimensional radical. Entonces las matrices M_i son simultáneamente diagonalizables.*

Demostración. Sea $a \in \mathcal{V}(I)$ un punto. Como en la demostración del Teorema de Stickelberg, existe un polinomio $g_a \in S$ tal que $g_a(a) = 1$ y $g_a(b) = 0$ si $b \in \mathcal{V}(I) \setminus \{a\}$. Afirmamos que los $\{g_a\}_{a\in\mathcal{V}(I)}$ son linealmente independientes en S/I. En efecto, supongamos que existe una combinación lineal

$$0 = \sum_{a\in\mathcal{V}(I)} c_a \cdot [g_a]$$

en S/I. Entonces existe $F \in I$ tal que

$$0 = F + \sum_{a\in\mathcal{V}(i)} c_a \cdot g_a$$

Como $F \in I$, $F(b) = 0$ para cada $b \in \mathcal{V}(I)$, así que

$$c_b = F(b) + \sum_{a\in\mathcal{V}(I)} c_a \cdot g_a(b) = 0$$

Por lo tanto, los $\{g_a\}_{a\in\mathcal{V}(I)}$ son linealmente independientes en S/I. Como I es radical, S/I tiene dimensión $\#\mathcal{V}(I)$, así que los $\{g_a\}_{a\in\mathcal{V}(I)}$ forman una base de S/I como $\mathbb{K}$-espacio vectorial.

Notemos que $(x_i - a_i)g_a$ se anula en todo punto de $\mathcal{V}(I)$, para cada $i \in \{1, 2, \ldots, n\}$. Por el Nullstellensatz se tiene que $(x_i - a_i)g_a \in \sqrt{I} = I$ para cada i. Luego:

$$m_i([g_a]) = [x_i \cdot g_a] = [(x_i - a_i) \cdot g_a] + a_i \cdot [g_a] = a_i \cdot [g_a]$$

en S/I. Esto demuestra que $[g_a]$ es un autovector común de cada m_i, por lo tanto, el conjunto $\{[g_a]\}_{a\in\mathcal{V}(I)}$ forma una base de autovectores de S/I en donde cada m_i actúa diagonalmente. Es decir, las matrices compañeras M_i asociadas a los m_i se diagonalizan simultáneamente en esta base. □

Falta cerrar este tema de eigenvalores... ¿cómo es que esto de eigenvalores ayuda a resolver sistemas de ecuaciones y relacionarlo con las preguntas al inicio del capítulo? –Abraham

Ejercicios

Agrega aquí ejercicios buenos para agarrar callo en geometría.

1. Demuestra algo

Capítulo 6

Soluciones reales

Muchos de los métodos hasta ahora vistos, requieren de considerar las soluciones de un sistema de polinomios dentro de la cerradura algebraica, sin embargo, es muy común en aplicaciones, que estemos interesados por entender aquellas soluciones que son reales.

Para iniciar con nuestro análisis de soluciones reales, iniciaremos con la siguiente observación derivada del Lema 5.18, el lema de la forma. Gracias a este lema, para resolver un sistema de polinomios, podemos iniciar por calcular un eliminante y reconstruir a partir de él todas las soluciones. Más aún, si los polinomios que consideramos tienen coeficientes reales, entonces el número de soluciones reales del sistema son igual al número de soluciones del eliminante, pues cada solución real de éste último produce una solución real del sistema. Así, se deriva el siguiente algoritmo de conteo de soluciones reales, que funciona solo cuando el sistema cumpla las condiciones del Lema de la Forma.

Algoritmo 9 (conteo de raíces reales)

Entrada: $I = \langle f_1, f_2, \ldots, f_N \rangle \subseteq \mathbb{Q}[x_1, \ldots, x_n]$.
Salida: Número de soluciones reales de I.

1: Calcula una base de Gröbner G de I.
2: Usa G para calcular $d = \text{gr}(I)$.
3: Usa G para calcular un eliminante $g(x_i) \in I \cap \mathbb{Q}[x_i]$ para alguna x_i.
4: **si** $\text{gr}(g) = d$ y g es libre de cuadrados **entonces**
5: Cuenta el número de raíces reales r de g.
6: **devolver** "I tiene r raíces reales".
7: **si no**
8: Haz un cambio de variable y vuelve al paso 3.
9: **fin si**

Observación 6.1. *Este algoritmo no siempre termina. Por ejemplo, si consideramos el ideal $I = \langle x^2, xy, y^2 \rangle$, el eliminante sería x^2 o y^2, que nunca es libre de cuadrados; ni siquiera bajo un cambio de coordenadas $x = au + bv$, $y = cu + dv$. Sin embargo, el algoritmo 9 funciona siempre que I es radical y 0-dimensional. Equivalentemente, si $\mathcal{V}(I)$ es una variedad reducida (que el anillo coordenado $\mathbb{C}[\mathcal{V}(I)]$ no tenga nilpotentes) y de dimensión 0.*

6.1. Regla de Descartes

Quisiéramos poder contar el número de soluciones reales de un sistema, sin tener que resolverlo. Para ello, iniciaremos la sección con un resultado clásico en Geometría Algebraica Real.

Definición 6.2. Sea $f(x) = c_0 + c_1x + \cdot + c_dx^d$ un polinomio real en una variable, un *cambio de signos* de f es un par consecutivo de coeficientes tales que $c_i \cdot c_{i+1} < 0$.

Teorema 6.3 (Regla de signos de Descartes)**.** *El número de raíces reales positivas de un polinomio en una variable es a lo más el número de cambios en signo de sus coeficientes.*

Ejemplo 6.4. Consideremos el polinomio $f(x) = 5x^6 - 4x^5 - 27x^4 + 55x^2 - 6$. Vemos que f

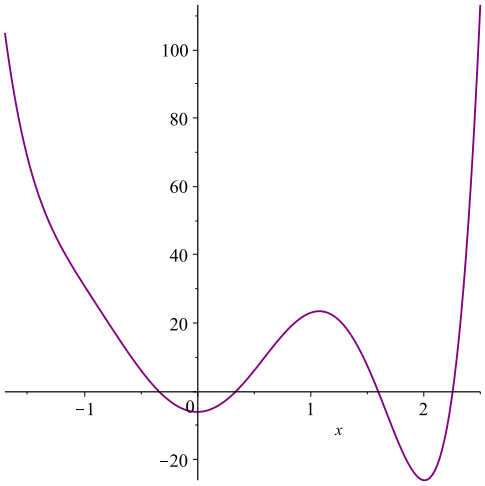


Figura 6.1: La curva $f(x) = 5x^6 - 4x^5 - 27x^4 + 55x^2 - 6$

tiene cuatro raíces reales, que son aproximadamente: -0.3393, 0.3404, 1.598 y 2.256. De estas, solo 3 son positivas. Notemos también que el número de cambios de signo en los coeficientes de f es 3, y este coincide con el número de soluciones reales positivas. △

Demostración. Sea $f(x) = c_0 + c_1x + \ldots + c_dx^d$ con $c_d \neq 0$ un polinomio real en una variable. La demostración será por inducción en $d = \text{gr}(f)$.

Para $d = 1$, el polinomio $f(x) = c_0 + c_1x$ tiene sólo una solución, y está dada por $x = -\frac{c_0}{c_1}$. Esta solución es positiva solo cuando c_0 y c_1 tienen signos opuestos.

Ahora, supongamos que el teorema se cumple para $d-1$, con $d > 1$, y sea $f(x) \in \mathbb{R}[x]$ un polinomio de grado d. Podemos suponer que x no divide a f, pues si $f(x) = c_kx^k + c_{k+1}x^{k+1} + \ldots + c_dx^d$, con $k > 0$, entonces $f(x) = x^k(c_k + c_{k+1}x + \ldots + c_dx^{d-k})$, donde tanto el número de raíces positivas como el número de cambios en signo de los coeficientes de $f(x)$ coinciden con los de $c_k + c_{k+1}x + \ldots + c_dx^{d-k}$. Escribimos entonces

$$f(x) = c_0 + \sum_{i=k}^{d} c_ix^i$$

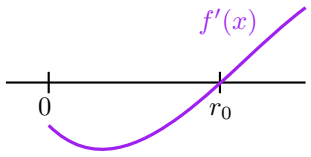
con $k \in \{1, 2, \ldots, d\}$ y $c_0, c_d \neq 0$. Si estos son los únicos términos distintos de cero, entonces $f(x) = c_0 + c_d x^d$ tiene a lo más una solución real positiva, que está dada por $x = \sqrt[d]{-c_0/c_d}$, cuando c_0 y c_d tienen signos opuestos.

Podemos ahora suponer que $c_0, c_k, c_d \neq 0$. Entonces:

$$f'(x) = \sum_{i=k}^{d} i \cdot c_i x^{i-1}.$$

Observemos que el signo de los coeficientes de f' es el mismo que el de los coeficientes de f, excepto por c_0. Por hipótesis de inducción, el número de cambios de signo en la sucesión $c_d, c_{d-1}, \ldots, c_k$ acota al número de raíces reales positivas de f'.

Sea r_0 la raíz positiva más pequeña de f' (hacemos $r_0 = \infty$ si no la hay). Entonces, como $f'(0) = c_k$, tenemos que f' tiene el mismo signo en el intervalo $(0, r_0)$.



Como $f(0) = c_0$ y $f' \neq 0$ en $(0, r_0)$, entonces f es monótona (creciente o decreciente) en ese intervalo, y por tanto, puede tener a lo más una raíz en $(0, r_0)$. Si $c_0 > 0$, para que f tenga una raíz en $(0, r_0)$, necesitamos que f sea decreciente, con lo cual $f'(0) = c_k < 0$. De manera análoga, si $c_0 < 0$ y f tiene una raíz en ese intervalo, necesitamos que $c_k > 0$. De esta forma, concluimos que para que f pueda tener una raíz en $(0, r_0)$, requerimos que $c_0 \cdot c_k < 0$. Este es justamente el caso en que el número de cambios de signo en f excede por 1 al número de cambios de signo de f'. Esto termina la demostración, puesto que por el teorema del valor medio entre cualesquiera dos raíces distitnas de f hay al menos una raíz de f'. □

Observación 6.5. *Al cambiar x por $-x$, la regla de Descartes para $f(-x)$ nos da una cota para el número de raíces reales negativas de $f(x)$.*

De esta manera, la regla de signos de Descartes junto con la última observación, proveen una cota en el número de raíces reales distintas de 0 de un polinomio real en una variable.

Corolario 6.6 (Cotas de Descartes)**.** *Un polinomio en una variable con $d+1$ términos tiene a lo más d raíces reales positivas reales, $2d$ raíces reales distintas de 0, y $2d+1$ raíces reales.*

Las cotas de Descartes se alcanzan. Por ejemplo, si $f(x) = x \cdot (x^2 - 1) \cdot (x^2 - 2) \cdot \ldots \cdot (x^2 - m)$.

6.2. Sucesión de Sturm

Veremos a continuación una generalización de la regla de signos de Descartes, y para ello necesitamos fijar un poco de notación. Comencemos escribiendo $f(x) = c_0 x^{a_0} + c_1 x^{a_1} + \ldots + c_m x^{a_m}$, con $c_i \neq 0$ para toda i, y con $0 \leq a_0 < a_1 < \ldots < a_m$.

Definición 6.7. Dada una sucesión $c = (c_1, c_2, \ldots, c_m)$ de números reales, su *variación* $var(c)$ es el número de veces que dos elementos (sin contar ceros) consecutivos de la sucesión tienen signos distintos.

Ejemplo 6.8. Si $c = (8, -4, -2, -1, 2, 3, -5, 7, 11, 12)$ entonces su variación es $var(c) = 4$. Mientras que para $c = (-1, 0, 1, 0, 1, -1, 1, 1, 0, 1)$, su variación es $var(c) = 3$. △

Definición 6.9. Dada una sucesión de polinomios en una variable $F = (f_0, f_1, \ldots, f_k)$ y un número real $a \in \mathbb{R}$, definimos $var(F, a)$ como la variación de $(f_0(a), f_1(a), \ldots, f_k(a))$. Si $a = \pm\infty$ entonces $var(F, \infty)$ es la variación de los términos iniciales de los polinomios $f_i(x)$, y $var(F, -\infty)$ es la variación de los coeficientes iniciales de los polinomios $f_i(-x)$.

La justificación para definir $var(F, \infty)$ como la diferencia de signos entre los coeficientes iniciales, es que cuando $a \to \infty$, el término dominante de $f_i(a)$ es el inicial, y por tanto, $f_i(a)$ tiene el mismo signo que su término inicial para a suficientemente grande.

Ejemplo 6.10. Sea $F = (f_1, f_2, f_3, f_4)$ con $f_1 = x^7 + 2x^4 - 5$, $f_2 = 8x^6 - 4x^4 - 2x - 1$, $f_3 = 2x^5 + 3x^2 - 5x + 7$, $f_4 = -x^5 + 5x^4 + 11x^3 - 1x + 2$. Entonces,

$$\begin{aligned} var(F, 1) &= var(-2, 1, 7, 16) = 1 \\ var(F, \infty) &= var(1, 8, 2, -1) = 1 \\ var(F, -\infty) &= var(-1, 8, -2, 1) = 3. \end{aligned}$$

△

Definición 6.11. Dado $f(x)$ un polinomio de grado d, definimos la siguiente sucesión

$$\delta(f) := (f(x), f'(x), \ldots, f^{(d)}(x)).$$

Definición 6.12. Para dos números distintos $a, b \in \mathbb{R} \cup \{\pm\infty\}$, denotamos con $r(f, a, b)$ al número de raíces reales de f dentro del intervalo $(a, b]$, contadas con multiplicidad.

Con la notación ya establecida, estamos en posición para enunciar el teorema principal de esta sección, que generaliza la regla de signos de Descartes.

Teorema 6.13 (Budan-Fourier). *Sea $f(x) \in \mathbb{R}[x]$ y sean $a < b$ en $\mathbb{R} \cup \{\pm\infty\}$. Entonces*

$$r(f, a, b) \leq var(\delta(f), a) - var(\delta(f), b),$$

y la diferencia entre los dos lados de la desigualdad es par.

Podemos recuperar la regla de Descartes de este teorema, tomando $a = 0$ y $b = \infty$, puesto que si escribimos $f(x) = c_0 + c_1 x + \cdots + c_d x^d$, entonces $var(\delta(f), 0) = var(c_0, c_1, \ldots, c_d)$ y $var(\delta(f), \infty) = 0$ pues los términos iniciales de $\delta(f)$ tienen todos el mismo signo.

Ejemplo 6.14. Sea $f(x) = 5x^6 - 4x^5 - 27x^4 + 55x^2 - 6$, cuya gráfica está ilustrada en el Ejemplo 6.4. Aunque f tiene cuatro raíces reales, solamente dos de ellas caen en el intervalo $[0, 2]$. Ahora, si evaluamos $\delta(f)$ en 0 obtenemos la sucesión

$$(-6,\ 0,\ 110,\ 0,\ -648,\ -480,\ 3600).$$

Por lo tanto, $var(\delta(f),0) = 3$. De manera similar, evaluando $\delta(f)$ en 2 obtenemos la sucesión

$$(-26, -4, 574, 2544, 5592, 6720, 3600),$$

por lo que $var(\delta(f),2) = 1$. El teorema de Budan-Fourier nos asegura que el polinomio $f(x)$ tiene 0 o 2 raíces reales en el intervalo $(0,2]$, como podemos verificar. △

El Teorema Teo:Budan-Fourier de Budan-Fourier, nos da una cota bastante buena para estimar el número de raíces reales de un polinomio, calculando la variación en los cambios de signo de una sucesión asociada. Esta idea se puede modificar, para producir otros métodos que puedan ser más efectivos (al menos parcialmente) para acotar el número de soluciones reales de un polinomio. Así, nos preguntamos si existen otras sucesiones de polinomios $F \neq \delta(f)$ que también determinen cotas. La sucesión de Sturm es un ejemplo afirmativo de esto.

Definición 6.15. Sea $f \in \mathbb{R}[x]$. La *sucesión de Sturm* de f es la siguiente sucesión de polinomios:

$$s_0(x) := f(x), \quad s_1(x) := f'(x), \quad s_i(x) := -res(s_{i-2}, s_{i-1}) \ \forall\ i \geq 2$$

donde $res(s_{i-2}, s_{i-1})$ es el residuo de dividir s_{i-2} entre s_{i-1} con el algoritmo de Euclides.

OJO: en la sección 1, el residuo $res(f,g)$ era denotado como f mód g –Abraham

Ejemplo 6.16. El polinomio $f(x) = 5x^6 - 4x^5 - 27x^4 + 55x^2 - 6$ tiene la siguiente sucesión de Sturm

$$\begin{aligned}
s_1 &= f' = 30x^5 - 20x^4 - 108x^3 - 110x \\
s_2 &= -res(s_0, s_1) = \frac{85}{9}x^4 + \frac{12}{3}x^3 + \frac{110}{3}x^2 - \frac{22}{9}x + 6 \\
s_3 &= -res(s_1, s_2) = -\tfrac{559584}{36125}\ x^3 + \tfrac{143748}{1445}x^2 - \tfrac{605394}{7225}x - \tfrac{126792}{7225} \\
s_4 &= -res(s_2, s_3) = -\tfrac{229905821875}{724847808}\ x^2 + \tfrac{1540527685625}{4349086848}\ x + \tfrac{7904908625}{120807968} \\
s_5 &= -res(s_3, s_4) = -\tfrac{280364022223059296}{58526435357253125}\ x + \tfrac{174201756039315072}{292632176786265625} \\
s_6 &= -res(s_4, s_5) = -\tfrac{17007035533771824564661037625}{162663080627869030112013128}
\end{aligned}$$

Notemos que $s_j = 0$ para $j \geq 7$, y por tanto, la sucesión la consideramos hasta s_6. △

Teorema 6.17 (Sturm). *Sea $f \in \mathbb{R}[x]$ y $a < b \in \mathbb{R} \cup \{\pm\infty\}$ con $f(a)$ y $f(b)$ distintos de 0. El número de raíces reales distintas de f, que caen en el intervalo $[a,b]$ está dado por*

$$var(\bar{s}, a) - var(\bar{s}, b)$$

con $\bar{s}$ la sucesión de Sturm de f.

Ejemplo 6.18. Corre el archivo `6SturmSeq.m2` en donde mostramos cómo calcular este ejemplo usando `Macaulay2`. Ilustraremos el teorema, utilizando el polinomio $f(x) = 5x^6 - 4x^5 - 27x^4 + 55x^2 - 6$ del ejemplo anterior. Sabemos del ejemplo 6.14 que el teorema de Budan-Fourier nos afirma que f tiene 0 o 2 raíces reales en el intervalo $(0,2)$. Utilizando la sucesión de Sturm, las evaluaciones en $x = 0$ y $x = 2$ son

$$\bar{s}(0) = \left\{-6, 0, 6, -\tfrac{126792}{7225}, \tfrac{7904908625}{120807968}, \tfrac{174201756039315072}{292632176786265625}, -\tfrac{17007035533771824564661037625}{162663080627869030112013128}\right\},$$
$$\bar{s}(2) = \left\{-26, -4, \tfrac{1114}{45}, \tfrac{3210228}{36125}, -\tfrac{1076053821625}{2174543424}, -\tfrac{2629438466191277888}{292632176786265625}, -\tfrac{17007035533771824564661037625}{162663080627869030112013128}\right\}.$$

De aquí tenemos que $var(\bar{s}, 0) = 4$ y $var(\bar{s}, 2) = 2$. El teorema nos ayuda a verificar que el número de soluciones reales de f en el intervalo $(0, 2)$ es $4 - 2 = 2$. △

Ya que ilustramos el teorema de Sturm, estamos listos para demostrarlo.

Demostración. Sea $f \in \mathbb{R}[x]$ y $\bar{s} = (s_0, s_1, \cdots, s_k)$ su sucesión de Sturm. Demostraremos el teorema analizando $var(\bar{s}, t)$ con t creciendo de a a b. Esta variación puede cambiar sólo cuando t pasa por algún $c \in (a, b)$ donde algún s_i se anula, ya que es en este caso que el signo de s_i puede cambiar. Demostraremos que si $i > 0$, entonces esto no tiene efecto en la variación de la sucesión, pero cuando c es una raíz de $s_0 = f$, la variación decrece exactamente en 1 cuando t pasa por c.

Observemos que si se satisface $s_i(c) = 0 = s_{i+1}(c)$ para alguna $c \in (a, b)$, entonces $s_{i-1}(c) = 0$, pues de la definición de s_{i+1}, podemos escribir

$$s_{i-1} = p_i s_i - s_{i+1} \tag{6.1}$$

para cierto polinomio $p_i \in \mathbb{R}[x]$. Por lo tanto, yendo hacia atrás vemos que $s_i(c) = 0 = s_{i+1}(c)$, implica que $f(c) = f'(c) = 0$, por lo que f tiene una raíz múltiple en c.

Supongamos primero que esto no pasa, o sea $f(c) \neq 0$ o c es una raíz simple de f. Si $s_i(c) = 0$ para algún $i > 0$, entonces (6.1) implica que $s_{i+1}(c)$ y $s_{i-1}(c)$ tienen signos opuestos. Sin pérdida de generalidad, supongamos que $s_{i-1}(c) > 0$. En este caso, tendríamos que $s_{i-1}(c) > 0$, $s_i(c) = 0$, y $s_{i+1}(c) < 0$. Veamos con la siguiente tabla, cómo es el comportamiento de los signos de estas tres funciones en una vecindad pequeña al rededor de c.

	$c - \epsilon$	c	$c + \epsilon$
s_{i-1}	$+$	$+$	$+$
s_i	?	0	?
s_{i+1}	$-$	$-$	$-$

De aquí vemos que sólo hay una variación en la subsucesión $s_{i-1}(t), s_i(t), s_{i+1}(t)$, sin importar el signo de s_i cerca de c; por lo que $s_i(c) = 0$ no afecta a la variación $var(\bar{s}, a) - var(\bar{s}, b)$.

Supongamos ahora que c es una raíz simple de f. Entonces $f'(c) \neq 0$, por lo que podemos suponer que $f'(c) > 0$.

Insertar imagen de f pasando de negativo a positivo por c y $f' > 0$ decreciendo.

Entonces f es creciente, y por tanto $var(\bar{s}, c - \epsilon) = (-, +, \cdots)$, mientras que $var(\bar{s}, c + \epsilon) = (+, +, \ldots)$. Por lo tanto, $var(\bar{s}, t)$ decrece por 1 cuando t para por una raíz simple (análogamente si $f'(c) < 0$) y no cambia cuando f se anula.

Nos resta solamente el caso cuando c es una raíz múltiple de f. Supongamos que c es una raíz de f con multiplicidad $m+1$. Entonces $(x-c)^m$ divide a f y f', pero por (6.1) también divide a todos los polinomios de la sucesión $\bar{s}$. Consideremos entonces la sucesión

$$\bar{q} = (q_0, q_1, \ldots, q_k) := \left(\frac{s_0}{(x-c)^m}, \frac{s_1}{(x-c)^m}, \ldots, \frac{s_k}{(x-c)^m} \right),$$

y notemos que $var(\bar{q}, t) = var(\bar{s}, t)$ cuando $t \neq c$, ya que al multiplicar la sucesión por un número distinto de 0 no cambia su variación. También notemos que $q_{i+1} = -res(q_i, q_{i-1})$, para toda $i \geq 1$.

Como $q_1 = \frac{f'}{(x-c)^m}$, entonces $q_1(c) \neq 0$ y por tanto no todos los q_i son 0 en c. Además, ya hemos demostrado que si $q_i(c) = 0$ para alguna $i > 0$, esto no cambia la variación cerca de c. Solo resta examinar la contribución de q_0 a la variación, cuando estamos cerca de c.

Si escribimos $f(x) = (x-c)^{m+1}g(x)$, con $g(c) \neq 0$, entonces

$$f'(x) = (m+1)(x-c)^m g(x) + (x-c)^{m+1}g'(x).$$

En particular, podemos escribir $q_0(x) = (x-c)g(x)$, y $q_1(x) = (m+1)g(x) + (x-c)g'(x)$. Si suponemos que $g(c) > 0$, entonces $q_1(c) > 0$ y q_0 cambia signo de $-$ a $+$ cuando t pasa por c. Por tanto $var(\bar{s}, t)$ decrece por 1 cuando t pasa por una raíz de f, como queríamos mostrar. □

Corolario 6.19. *El número de raíces reales de $f \in \mathbb{R}[x]$ es igual a la diferencia*

$$var(\bar{s}, -\infty) - var(\bar{s}, +\infty).$$

6.3. Criterio de Hermite-Sylvester

A continuación veremos otro método clásico para contar el número de soluciones reales, que es el criterio de Hermite-Sylvester.

Definición 6.20. Sea $f \in \mathbb{R}[x]$ con $\mathrm{gr}(f) = d$. Para un polinomio $g \in \mathbb{R}[x]$ fijo, la *matriz de Hankel $H_g(f)$* es la matriz simétrica de $d \times d$ definida por

$$H_g(f)_{ij} := \sum_{k=1}^{d} g(\xi_k)\xi_k^{i+j-2},$$

donde $\xi_1, \ldots, \xi_d$ son las raíces de f (en $\mathbb{C}$).

Ejemplo 6.21. Si $f \in \mathbb{R}[x]$ es de grado d, con $\xi_1, \ldots, \xi_d$ todas sus raíces. Para $q(x) = 1$, la matriz de Hankel es

$$H_1(f) = \begin{pmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & \ddots & \vdots \\ h_{d-1} & h_d & \cdots & h_{2d-2} \end{pmatrix}$$

donde $h_i = \sum_{k=1}^{d} \xi_k^i$. Por ejemplo,

$$h_0 = \sum_{k=1}^{d} 1 = d, \qquad h_1 = \sum_{k=1}^{d} \xi_k, \qquad h_2 = \sum_{k=1}^{d} \xi_k^2.$$

Las h_i's son llamadas las *i-ésimas sumas de Newton de f*. △

Los coeficientes de un polinomio en una variable pueden ser escrito como funciones simétricas. Si escribimos $f = x^d + \sum_{i=0}^{d-1} c_i x^i$, y $\xi_1, \ldots, \xi_d$ son todas sus raíces, entonces

$$f = x^d + \sum_{i=0}^{d-1} c_i x^i = (x - \xi_1) \cdots (x - \xi_d).$$

Desarrollando el producto en la última igualdad, obtenemos las identidades de Vieta, que igualan los coeficientes de f con funciones simétricas elementales en sus raíces, es decir,

$$\begin{aligned}
c_0 &= (-1)^d \xi_1 \cdots \xi_d \\
c_1 &= (-1)^{d-1}(\xi_2\xi_3 \cdots \xi_d + \xi_1\xi_3 \cdots \xi_d + \cdots + \xi_1\xi_2 \cdots \xi_{d-1}) = (-1)^{d-1} \sum_{i=1}^{d} \prod_{j \neq i} \xi_j. \\
&\vdots \\
c_{d-2} &= \xi_1\xi_2 + \xi_1\xi_3 + \cdots + \xi_{d-1}\xi_d = (-1)^2 \sum_{1 \leq i < j \leq d} \xi_i \xi_j. \\
c_{d-1} &= (-1)(\xi_1 + \xi_2 + \cdots + \xi_d) = (-1) \sum_{i=1}^{d} \xi_j.
\end{aligned}$$

Las funciones simétricas elementales, junto con las funciones simétricas completamente homogéneas, son muy importantes y muy estudiadas en combinatoria algebraica. De esta manera, utilizando la relación entre estas funciones simétricas, obtenemos las *identidades de Newton*:

$$\begin{aligned}
h_k + \sum_{i=1}^{k-1} c_{d-k+i} h_i &= -k c_{d-k} \qquad (1 \leq k < d). \\
h_k + \sum_{i=k-d}^{k-1} c_{d-k+i} h_i &= 0 \qquad (\forall\, k \geq d),
\end{aligned}$$

Escribiendo estas identidades recursivamente, podemos escribir las sumas de Newton como polinomios en los coeficientes de f. Ilustramos esto para las primeras sumas de Newton:

$$\begin{aligned}
h_0 &= d, \\
h_1 &= -c_{d-1}, \\
h_2 &= -c_{d-1}s_1 - 2c_{d-2} = c_{d-1}^2 - 2c_{d-2}, \\
h_3 &= -c_{d-1}s_2 - c_{d-2}s_1 - 3c_{d-3} = -c_{d-1}^3 + 3c_{d-1}c_{d-2} - 3c_{d-3}.
\end{aligned}$$

Para continuar con los preámbulos necesarios para enunciar el criterio de Hermite-Sylvester, recordemos que una matriz simétrica Q define naturalmente una forma cuadrática $\mathbf{z}^t Q \mathbf{z}$. Por ejemplo, si $Q = \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix}$, entonces $\begin{pmatrix} x & y \end{pmatrix} Q \begin{pmatrix} x \\ y \end{pmatrix} = Ax^2 + Bxy + Cy^2$. Ahora, si f es un polinomio de grado d, y $\xi_1, \ldots, \xi_d$ sus raíces, la matriz de Henkel de f es una matriz simétrica

$H_g(f)$, y su forma cuadrática es

$$\mathbf{z}^t H_g \mathbf{z} = \begin{pmatrix} z_0 & z_1 & \cdots & z_{d-1} \end{pmatrix} \begin{bmatrix} \sum g(\xi_k) & \sum g(\xi_k)\xi_k & \cdots & \sum g(\xi_k)\xi_k^{d-1} \\ \sum g(\xi_k)\xi_k & \sum g(\xi_k)\xi_k^2 & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ \sum g(\xi_k)\xi_k^{d-1} & \sum g(\xi_k)\xi_k^d & \cdots & \sum g(\xi_k)\xi_k^{2d-2} \end{bmatrix} \begin{pmatrix} z_0 \\ z_1 \\ \vdots \\ z_{d-1} \end{pmatrix}$$

$$= \sum_{k=1}^{d} g(\xi_k)(z_0 + z_1\xi_k + \ldots + z_{d-1}\xi_k^{d-1})^2$$

A lo mejor es mejor simplificar las matrices de Hankel, dejando de lado el caso general g, y quedándonos solamente el caso $g = 1$ –Abraham

Ahora, si denotamos por V a la matriz de Vandermonde

$$V := \begin{pmatrix} 1 & \xi_1 & \cdots & \xi_1^{d-1} \\ 1 & \xi_2 & \cdots & \xi_2^{d-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \xi_d & \cdots & \xi_d^{d-1} \end{pmatrix}$$

entonces

$$H_g(f) = V^t \begin{bmatrix} g(\xi_1) & 0 & \cdots & 0 \\ 0 & g(\xi_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g(\xi_d) \end{bmatrix} V$$

Definición 6.22. Dada una matriz simétrica A no singular de $d \times d$, existe una matriz C (no singular) tal que $C^t AC = D$ es una matriz diagonal con r 1's y s -1's. La *signatura* de A la definimos como $r - s$.

Teorema 6.23. *El rango de $H_g(p)$ es igual al número de raíces distintas ξ_j de p tales que $g(\xi_j) \neq 0$. Más aún, la signatura de $H_g(p)$ es el número de raíces ξ_k reales y diferentes para las que $g(\xi_k) > 0$ menos aquellas tales que $g(\xi_k) < 0$*

Demostración. Primero consideremos el caso en el que todas las raíces $\xi_1, \ldots, \xi_d$ son distintas. Si definimos $\zeta(\xi_k) := \sum_{i=0}^{d-1} z_i \xi_k^i$, entonces

$$\mathbf{z}^t H_g(p) \mathbf{z} = \sum_{k=1}^{d} g(\xi_k)\zeta(\xi_k)^2.$$

Escribamos esta forma cuadrática en ξ como

$$\mathbf{z}^t H_g(p) \mathbf{z} = \sum_{\xi_k \in \mathbb{R}}^{d} g(\xi_k)\zeta(\xi_k)^2 + \frac{1}{2} \sum_{\xi_k \in \mathbb{C}\setminus\mathbb{R}}^{d} g(\xi_k)\zeta(\xi_k)^2 + g(\overline{\xi_k})\zeta(\overline{\xi_k})^2. \tag{6.2}$$

Si escribimos $g(\xi_k) = a + bi$ y $\zeta(\xi_k) = c + di$ entonces

$$g(\xi_k)\zeta(\xi_k)^2 + g(\overline{\xi_k})\zeta(\overline{\xi_k})^2 = (a + bi)(c + di)^2 + (a - bi)(c - di)^2 = 2a(c^2 - d^2) - 4bcd$$

por lo que (6.2) se reescribe como

$$= \sum_{\xi_k \in \mathbb{R}}^{d} g(\xi_k)\zeta(\xi_k)^2 + \sum_{\xi_k \in \mathbb{C}\setminus\mathbb{R}} \begin{bmatrix} Re(\zeta(\xi_k)) \\ Im(\zeta(\xi_k)) \end{bmatrix}^t \begin{bmatrix} Re(g(\xi_k)) & -Im(g(\xi_k)) \\ -Im(g(\xi_k)) & -Re(g(\xi_k)) \end{bmatrix} \begin{bmatrix} Re(\zeta(\xi_k)) \\ Im(\zeta(\xi_k)) \end{bmatrix}.$$

Ojo: Para que $\zeta(\overline{\xi_k}) = \overline{\zeta(\xi_k)}$, estás asumiendo que los z_i solo pueden tomar valores reales. Como los ceros ξ_k son distintos, los polinomios $\zeta(\xi_k)$ son linealmente independientes y por lo tanto, también lo son

$$\{\zeta(\xi_k)\}_{\xi_k \in \mathbb{R}} \cup \{Re(\zeta(\xi_k)),\ Im(\zeta(\xi_k))\}_{\xi_k \in \mathbb{C}\setminus\mathbb{R}}$$

que corresponden a formas lineales en $z_0, \ldots, z_{d-1}$, por lo tanto hemos representado la forma cuadrática de $H_g(p)$ con respecto a otra base.

Dado que la signatura es invariante bajo cambio de base, podemos calcular la signatura de los escalares $g(\xi_k)$ y de los bloques de tamaño de 2×2, pero estos últimos tienen signatura 0 (porque la traza de bloques es 0) lo que implica el resultado.

Para el caso general, si $\xi_1, \ldots, \xi_s$ son raices distintas, con multiplicidad $\mu(\xi_k)$ entonces $\mathbf{z}^t H_g(p)\mathbf{z} = \sum_{k=1}^{s} \mu(\xi_k) g(\xi_k)\zeta(\xi_k)^2$ y el resultado es análogo. □

Observación 6.24. *Podemos contar las raíces reales de p considerando $g(\xi) = 1$.*

Corolario 6.25. *La signatura de $H_1(p)$ = número de raíces reales distintas de p.*

Corolario 6.26. *Todas las raíces de p son reales $\Leftrightarrow$ la matriz $H_1(p)$ es positiva semidefinida.*

Definición 6.27. Una matriz A es *positiva semidefinida* si $\mathbf{x}^t A\mathbf{x} \geq 0\ \forall \mathbf{x} \neq 0$.

6.4. Forma de la Traza

Definición 6.28. Dada una forma cuadrática (real) F, *la signatura* $\sigma(F)$ es el número de eigenvalores positivos menos los eigenvalores negativos de la matriz que representa a F. El *rango* $\rho(F)$ es el rango de la matriz.

Sea $I \subset S := \mathbb{C}[x_1, \ldots, x_n]$ un ideal de dimensión cero, y $\mathcal{B}$ una base monomial de S/I. Para $g \in S$ definimos la *operación multiplicación por g* como

$$\begin{aligned} m_g : S/I &\to S/I \\ [f] &\mapsto [g][f] = [gf] \end{aligned}$$

Si fijamos $q \in \mathbb{R}[x_1, \ldots, x_n]$, construimos una forma bilineal T_q de la siguiente manera:

$$\begin{aligned} T_q : S/I \times S/I &\to \mathbb{R} \\ ([g], [f]) &\mapsto Tr(m_{qgf}). \end{aligned}$$

La forma T_q es la *forma de la traza de q*.

En la definición, $T_q(g, f) = Tr(m_{qgf})$ es la traza de la matriz que representa a m_{qgf} respecto a la base $\mathcal{B}$.

Lema 6.29. *Sea A una matriz y $\lambda_1, \ldots, \lambda_n$ sus eigenvalores, entonces*

$$Tr(A) = \sum_{i=1}^{n} \lambda_i \quad y \quad \det(A) = \prod_{i=1}^{n} \lambda_i.$$

Bosquejo de prueba. Mostraremos los elementos necesarios para la demostración del teorema analizando los primeros dos casos. Para

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

el polinomio característico se obtiene de la siguiente manera

$$\det(A - tI) = \det \begin{bmatrix} a-t & b \\ c & d-t \end{bmatrix} = t^2 - (a+d)t + (ad - bc) = t^2 - Tr(A) \cdot t + \det(A).$$

Para

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix},$$

su polinomio característico es $\det(A - tI) = (-1)^3(t^3 - (a+e+k)t^2 + \text{ “algo” } \cdot t + \det(A))$. En general, en el polinomio característico, el coeficiente más grande de t es 1, el siguiente está dado por la traza, y por último, el término constante está dado por el determinante. □

Teorema 6.30. *Para $q \in \mathbb{R}[x_1, \ldots, x_n]$, la signatura y el rango de la forma de la traza T_q satisfacen:*

$$\sigma(T_q) = |\{a \in V(I) \cap \mathbb{R}^n | q(a) > 0\}| - |\{a \in V(I) \cap \mathbb{R}^n | q(a) < 0\}|,$$
$$\rho(T_q) = |\{a \in V(I) | q(a) \neq 0\}|.$$

Para la demostración, usaremos el teorema de Stickelberger (que ya hemos visto):

Lema 6.31 (Teorema de Stickelberger)**.** *Si $I \subset S$ es un ideal 0-dimensional, para cualquier $i \in \{1, \ldots, n\}$, $\lambda \in \mathbb{C}$ es un eigenvalor de $m_i \iff \exists\, a \in V(I)$ tal que $a_i = \lambda$.*

Demostración. Por simplicidad, supongamos que todos los puntos tienen multiplicidad 1. Sea $\mathfrak{B} = \{x^{\alpha(1)}, \ldots, x^{\alpha(d)}\}$ una base monomial de S/I. La matriz M_q que representa a T_q respecto a $\mathfrak{B}$ tiene la entrada (i, j) dada por $Tr(m_q, x^{\alpha(i)}, x^{\alpha(j)})$. Vemos donde van los d^2 elementos de la base, es decir, la imágen de los elementos $x^{\alpha(i)} \times x^{\alpha(j)}$, $\forall i, j$.

Sea $f \in S$, del teorema de Stickelberger (Lema 6.31) sabemos que el conjunto de eigenvalores de m_f coincide con el conjunto de valores $f(p)$ tal que $p \in V(I)$.

Sean $p_1, \ldots, p_d$ los puntos de $V(I)$ (que estamos suponiendo distintos). Entonces, la suma de los eigenvalores de $m_{qx^{\alpha(i)}x^{\alpha(j)}}$ es

$$\sum_{p \in V(I)} q(p) p^{\alpha(i)} p^{\alpha(j)} \tag{6.3}$$

donde $p^{\alpha(i)}$ denota el valor del monomio $x^{\alpha(i)}$ evaluado en p. Sean

$$C = \begin{bmatrix} p_1^{\alpha(1)} & \cdots & p_d^{\alpha(1)} \\ \vdots & \ddots & \vdots \\ p_1^{\alpha(d)} & \cdots & p_d^{\alpha(d)} \end{bmatrix}, D = \begin{bmatrix} q(p_1) & 0 & \dots & 0 \\ 0 & q(p_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & q(p_d) \end{bmatrix}.$$

Entonces la expresión (6.3) implica que podemos descomponer M_q como: $M_q = CDC^T$.

En general, C, D, son matrices con entradas complejas. Sin embargo, los puntos no reales vienen en pares conjugados, y se puede usar un argumento similar a la prueba de la signatura de la matriz de Hankel para omitir estos pares conjugados a la hora de calcular la signatura de M_q. Por lo tanto, los eigenvalores de T_q correspondientes a los puntos reales son los $q(p)$ tales que $p \in V(I) \cap \mathbb{R}^n$. El teorema se sigue de aquí. □

Corolario 6.32. *La signatura de T_1 (i.e., $q(x) = 1$) es igual al número de raíces reales distintas de I.*

Para el caso especial $q(x) = 1$ y $n = 1$, el ideal I es principal; es decir, $I = \langle f \rangle$ con $f \in \mathbb{R}[x]$ de grado d. Si $\mathfrak{B} = \{1, x, \dots, x^{d-1}\}$, entonces

$$\sum_{p \in V(I)} q(p) p^{\alpha(i)} p^{\alpha(j)} \quad \mapsto \quad (M_1)_{ij} = \sum_{p \in V(I)} p^{i-1} p^{j-1} = VV^T$$

donde V es la matriz de Vandermonde. La matriz de Hankel $H_1(f)$ es M_1, lo que muestra que esta técnica generaliza lo visto en la clase anterior.

Teorema 6.33. *Sea A una matriz real simétrica, el número de eigenvalores positivos de A es igual al número de cambios de signo de su polinomio característico $\chi_A(t) = \det(A - tI_d)$.*

Ejercicios

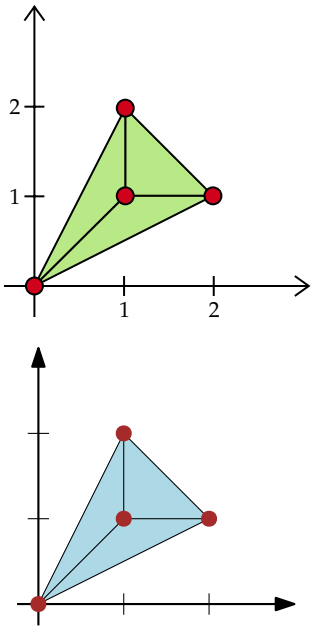
Agrega aquí ejercicios buenos para agarrar callo en geometría.

1. Demuestra algo

Capítulo 7

Polinomios Ralos y variedades Tóricas

Ejemplo 7.1. Sean $f(x) = x^2y + 2xy^2 + xy - 1$ y $g(x) = x^2y + xy^2 - xy + 2$. Los exponentes son $(2,1), (1,2), (1,1), (0,0)$. Gráficamente:



△

Definición 7.2. El *anillo de polinomios de Laurent* se define por

$$\mathbb{C}[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}]$$

Observación 7.3. *Consideremos*

$$\begin{aligned} \varphi : \mathbb{C}[a,b] &\longrightarrow \mathbb{C}[x, x^{-1}] \\ a &\mapsto x \\ b &\mapsto x^{-1} \end{aligned}$$

se tiene que $1-ab \in \ker(\varphi)$, *de hecho* $\langle 1-ab \rangle = \ker(\varphi)$, *así que* $\mathbb{C}[x, x^{-1}] \simeq \frac{\mathbb{C}[a,b]}{\langle 1-ab \rangle}$.

Definición 7.4. Sea $\mathbb{C}^\times := \mathbb{C}\backslash\{0\}$ el grupo multiplicativo de $\mathbb{C}$, al conjunto

$$(\mathbb{C}^\times)^n := \{(a_1, \ldots, a_n) \in \mathbb{C}^n \mid a_i \neq 0 \ \forall i\}$$

se le llama *toro algebraico complejo*.

Resulta que $(\mathbb{C}^\times)^n$ es una variedad algebraica y su anillo de coordenadas es el de Laurent:

$$\mathbb{C}[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}] = \mathbb{C}[x_1^{\pm}, \ldots, x_n^{\pm}] \simeq \frac{\mathbb{C}[a_1, \ldots, a_n, b]}{\langle ba_1 \ldots a_n - 1\rangle}$$

Definición 7.5. Dado un conjunto $\mathcal{A} = [\alpha_1, \ldots, \alpha_n] \in \mathbb{Z}^n$ finito. Un *polinomio ralo con soporte en* $\mathcal{A}$ es

$$f = \sum_{\alpha \in \mathcal{A}} c_\alpha \, x^\alpha \in \mathbb{C}[x_1^{\pm}, \ldots, x_n^{\pm}] \tag{7.1}$$

Nuestro objetivo es estudiar sistemas de polinomios ralos

$$f_1(p) = \cdots = f_n(p) = 0$$

Definición 7.6. Sea $\mathcal{A} \subset \mathbb{Z}^n$ finito, una *combinación convexa* de puntos de $\mathcal{A}$ es $\sum_{\alpha\in\mathcal{A}} \lambda_\alpha \alpha$ con $\lambda_\alpha \geq 0$ y $\sum_{\alpha\in\mathcal{A}} \lambda_\alpha = 1$.

Definición 7.7. La *envolvente convexa* de $\mathcal{A}$ se define por

$$\Delta_{\mathcal{A}} = \text{conv}(\mathcal{A}) := \left\{ \sum_{\alpha\in\mathcal{A}} \lambda_\alpha \alpha \mid \lambda_\alpha \geq 0, \sum_{\alpha\in\mathcal{A}} \lambda_\alpha = 1 \right\}$$

Definición 7.8. Un *politopo* es la envolvente convexa de un número finito de puntos en $\mathbb{R}^n$.

Definición 7.9. La *dimensión* de un politopo $\Delta = \text{conv}(a_1, \ldots, a_m)$ es la dimensión del espacio vectorial dado por $span\{a_2 - a_1, \ldots, a_m - a_1\}$.

Lema 7.10. *Sea $\Delta = \text{conv}(\mathcal{A}) \subset \mathbb{R}^n$ un politopo. Para todo $\omega \in \mathbb{R}^n$ existe $\alpha \in \Delta$ tal que $\omega \cdot \alpha = \sum \omega_i \alpha_i$ es mínimo.*

Demostración. Define

$$\begin{aligned} \varphi \colon \Delta &\to \mathbb{R} \\ \alpha &\mapsto \omega \cdot \alpha \end{aligned}$$

Dado que Δ es compacto, φ tiene un mínimo. □

Definición 7.11. Sea $\Delta = \text{conv}(\mathcal{A})$ un *politopo* y $\omega \in \mathbb{R}^n$. La *cara* de Δ con soporte en ω es:

$$F_w = \text{conv}\{v \in \mathcal{A} \mid \omega \cdot v \ \text{ es mínimo}\}$$

Definición 7.12. Un punto $\alpha \in \Delta = \text{conv}(\mathcal{A})$ es un *vértice* si no se puede escribir como combinación lineal convexa de otros puntos.

Lema 7.13. *Sea $\Delta = \text{conv}(\mathcal{A}) \subseteq \mathbb{R}^n$ un politopo. Para todo w en $\mathbb{R}^n$ existe un único $\alpha \in \Delta$ tal que $w \cdot \alpha = w^T\alpha = w_1\alpha_1 + \ldots + w_n\alpha_n$ es mínimo.*

Definición 7.14. Sea $\Delta = \text{conv}(\mathcal{A})$ un politopo y $w \in \mathbb{R}^n$. Definimos la *cara* de Δ *con soporte en* w como

$$F_w := \text{conv}\{v \in \mathcal{A} \mid w \cdot v \text{ es mínimo}\}.$$

Definición 7.15. Un punto $\alpha \in \Delta = \text{conv}(\mathcal{A})$ es un *vértice* si no se puede escribir como combinación convexa de otros puntos. Una *arista* es una cara de la forma $\text{conv}(v_1, v_2)$, donde v_1, v_2 son vértices. En general, las caras de Δ son politopos en si, y las caras de dimensión $d-1$ (las caras de dimensión máxima) son llamadas *facetas*.

Definición 7.16. Un *hiperplano* $H \subseteq \mathbb{R}^n$ es *de soporte* de $P = \text{conv}(\mathcal{A})$ si $P \cap H \neq \emptyset$ y $P \leq H^-$ o $P \leq H^+$.

Teorema 7.17 (*Kushnirenko*)**.** *Un sistema de n polinomios ralos en n variables con soporto en $\mathcal{A}$ tiene a lo más $n! \cdot \text{Vol}(\text{conv}(\mathcal{A}))$ soluciones en $(\mathbb{C}^\times)^n$ y es exactamente este número si los politopos son genéricos.*

Notación. $d(\mathcal{A})$ denotará el número de soluciones de un sistema ralo soportado en $\mathcal{A}$.

Definición 7.18. Dado un polinomio $f = \sum_{i=1}^m c_i x^{\alpha_i}$, el *polinomio de Newton de* f se define por

$$New(f) := \text{conv}(\alpha_1, \ldots, \alpha_m).$$

Definición 7.19. Dados P y Q dos politopo, definimos su *suma de Minkowski* como

$$P + Q = \{p + q \mid p \in P,\ q \in Q\}.$$

Observación 7.20. *Si $w \in \mathbb{R}^n$, entonces $F_w(P+Q) = F_w(P) + F_w(Q)$, por lo que la suma es también aditiva para las caras. Además, para todo vértice v de $P+Q$ existen únicos vértices $a \in P$ y $b \in Q$ tales que $a + b = v$.*

Esta observación nos ayuda para calcular los vértices de la suma de Minkowski, la cual ilustramos en la Figura 7.1.

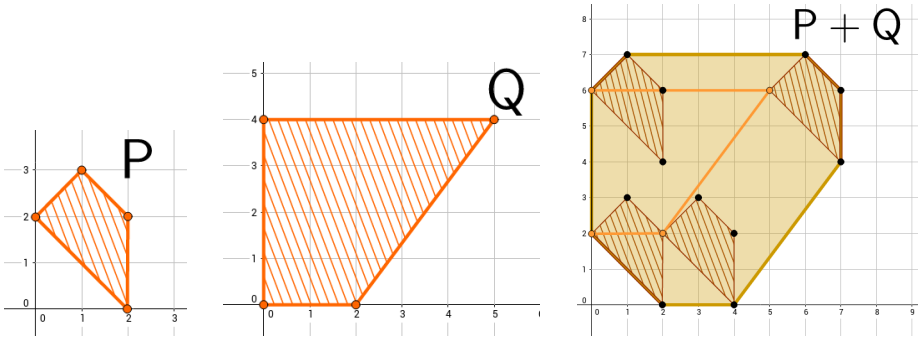


Figura 7.1: Suma de Minkowski de dos politopos

Observación 7.21. *Si f y g son polinomios entonces*

$$New(f \cdot g) = New(f) + New(g).$$

Una interpretación del politopo de Newton, es que dicho politopo es una generalización geométrica del grado. En una variable, sabemos que el soporte de un polinomio de grado d es un politopo correspondiente al segmento $[0, d]$, y que el producto de dos polinomios nos dará un nuevo polinomio cuyo grado es igual a la suma de los grados. En este sentido, el politopo de Newton y la suma de Minkowski nos generaliza estas ideas.

Proposición 7.22. *Sean $k_1, \ldots, K_m$ politopos y $\lambda_1, \ldots, \lambda_m \in \mathbb{R}_{>0}$ entonces,*

$$\mathrm{Vol}(\lambda_1 K_1 + \ldots + \lambda_m K_m) = \sum_{c_1,\ldots,c_t \in [m]} \mathrm{Vol}(K_{c_1} + \ldots + K_{c_t})\lambda_{c_1} \cdots \lambda_{c_t}.$$

Proposición 7.23. *Sean $f_1, \ldots, f_n \in \mathbb{C}[x_1, \ldots, x_n]$ genéricos y $\mathcal{A}_1, \ldots, \mathcal{A}_n$ sus soportes. Si $d(\mathcal{A}_1, \ldots, \mathcal{A}_n)$ denota el número de soluciones del sistema, entonces*

$$d(\mathcal{A}_1 + B, \mathcal{A}_2, \ldots, \mathcal{A}_n) = d(\mathcal{A}_1, \ldots, \mathcal{A}_n) + d(B, \mathcal{A}_2, \ldots, \mathcal{A}_n),$$

si $\mathcal{A}_i = \mathcal{A}$ para toda $i = 1, \ldots, n$, entonces $d(\mathcal{A}) = n!\,\mathrm{Vol}(\Delta_{\mathcal{A}})$.

Proposición 7.24 (Minkowski)**.** *Si $K_1, \ldots, K_m \subset \mathbb{R}^n$ son politopos convexos, y $\lambda_1, \ldots, \lambda_n$ son números positivos, entonces $\mathrm{Vol}(\lambda_1 K_1 + \cdots + \lambda_n K_n)$ es un polinomio homogéneo de grado n en las variables $\lambda_1, \ldots, \lambda_n$. Podemos escribir*

$$\mathrm{Vol}(\lambda_1 K_1 + \cdots + \lambda_n K_n) = \sum_{i_1,\ldots,i_n \in [n]} \mathrm{Vol}(K_{i_1}, \ldots, K_{i_n})\lambda_{i_1} \cdot \ldots \cdot \lambda_{i_n}.$$

Definición 7.25. Si $K_1, \ldots, K_m \subset \mathbb{R}^n$ son politopos convexos, se define su *volumen mixto* por

$$\mathrm{MixVol}(K_1, \ldots, K_m) = \sum_{k=1}^{m} (-1)^{m-k} \sum_{1 \le i_1 \le \ldots \le i_k \le m} \mathrm{Vol}(K_{i_1} + \ldots + K_{i_k}).$$

Teorema 7.26 (*Bernstein*)**.** *Sean $\mathcal{A}_1, \ldots, \mathcal{A}_n$ subconjuntos finitos de vectores de $\mathbb{Z}^n$. Un sistema ralo de polinomios en n variables con soporte en $\mathcal{A}_1, \ldots, \mathcal{A}_n$ y con un número finito de soluciones tiene a lo más* $\mathrm{MixVol}(\mathrm{conv}(\mathcal{A}_1), \ldots, \mathrm{conv}(\mathcal{A}_n))$ *soluciones en $(\mathbb{C}^\times)^n$ y alcaza este número cuando los polinomios son genéricos.*

Teorema 7.27 (Bernstein)**.** *Un sistema de n polinomios en n variables que tiene soporte $A_1, \ldots, A_n$ tienen a lo más* $\mathrm{MixVol}(A_1, \ldots, A_n)$ *soluciones en $(\mathbb{C}^\times)^n$.*

Ahora vamos a probar el teorema anterior. Sólo se hará para sistemas binomiales. Supongamos que tenemos

$$f = x^{a_1} y^{b_1} - c_1, \quad g = x^{a_2} y^{b_2} - c_2$$

con $a_i,\ b_i \in \mathbb{Z}$, y $c_i \in \mathbb{C}^\times$. Recordando que

$$SL_2(\mathbb{Z}) = \left\{ u = \begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix} \in \mathbb{Z}^{2\times 2} \mid \det(u) = 1 \right\}$$

y dada $A \in \mathbb{Z}^{n\times n}$, $\exists!\ U \in SL_n(\mathbb{Z})$ tal que $UA = H$, con H matriz triangular $\therefore \exists!\ U = \begin{pmatrix} u_{1,1} & u_{1,2} \\ u_{2,1} & u_{2,2} \end{pmatrix}$ tal que

$$\begin{pmatrix} u_{1,1} & u_{1,2} \\ u_{2,1} & u_{2,2} \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \begin{pmatrix} r_1 & r_3 \\ 0 & r_2 \end{pmatrix}$$

entonces si queremos que $f = 0 = g$, $\Rightarrow\ x^{a_1}y^{b_1} = c_1$ y $x^{a_2}y^{b_2} = c_2$

$$\Leftrightarrow\quad (x^{a_1}y^{b_1})^{u_{1,1}}(x^{a_2}y^{b_2})^{u_{1,2}} = c_1^{u_{1,1}}c_2^{u_{1,2}},$$

$$(x^{a_1}y^{b_1})^{u_{2,1}}(x^{a_2}y^{b_2})^{u_{2,2}} = c_1^{u_{2,1}}c_2^{u_{2,2}}$$

$\Leftrightarrow\quad x^{r_1}y^{r_3} = c_1^{u_{1,1}}c_2^{u_{1,2}}$ y $y^{r_2} = c_1^{u_{2,1}}c_2^{u_{2,2}}$.

Este último sistema tiene $r_1 \cdot r_2$ soluciones. Ahora calculemos MixVol(P), donde P es la suma de Minkowski de los polígonos de Newton de f y g:

$$\begin{aligned} \text{MixVol}(P) &= \text{Vol}(New(f) + New(g)) - \text{Vol}(f) - \text{Vol}(g) \\ &= \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \det \begin{pmatrix} r_1 & r_3 \\ 0 & r_2 \end{pmatrix} = r_1 r_2. \end{aligned}$$

Ahora veamos como demostrar lo anterior para polinomios ralos en dos variables. Para esto, es necesario introducir la noción de *deformacones tóricas*:

Idea: Queremos agregar una nueva variable t, y mutiplicar los términos de f y g por potencias de t de tal forma que podamos analizar las soluciones cuando $t = 1$ a partir de las soluciones cuando $t = 0$.

Consideremos el sistema

$$\begin{aligned} f(x,y) &= a_1 + a_2x + a_3xy + c_4y \\ g(x,y) &= b_1 + b_2x^2y + b_3xy^2. \end{aligned}$$

Agregando potencias de t obtenemos

$$f_t(x,y) = a_1t^{v_1} + a_2xt^{v_2} + a_3xyt^{v_3} + c_4yt^{v_4} \tag{7.2}$$

$$g_t(x,y) = b_1t^{w_1} + b_2x^2yt^{w_2} + b_3xy^2t^{w_3}. \tag{7.3}$$

Necesitamos unas $v_i,\ w_i$ que sean *suficientemente genéricas* (por definir).

Observación 7.28. *Podemos interpretar el sistema $f_t = 0 = g_t$ como un sistema en 2 variables cuyas soluciones dependen del parámetro t, es decir, la función $t \longmapsto (x(t), y(t))$ se ramifica, y queremos identificar las ramas.*

Si consideramos un polinomio

$$p(t;x) = a_d(t)x^d + a_{d-1}(t)xd - 1 + \ldots + a_1(t)x + a_0$$

en el anillo $\mathbb{K}[x]$ con $\mathbb{K} = \mathbb{Q}(t)$, entonces todas sus soluciones viven en el campo de series de Puiseux $\mathbb{C}\{\{t\}\}$ que es algebraicamente cerrado, donde $\mathbb{C}\{\{t\}\}$ es el conjunto de series de potencias formales en t, con coeficientes en $\mathbb{C}$, con exponentes racionales acotadas por abajo:

$$\mathbb{C}\{\{t\}\} = \bigcup_{N=1}^{\infty} \mathbb{C}((t^{1/N}))$$

donde $\mathbb{C}((t^{1/N}))$ es el campo de series de Laurent en $t^{1/n}$, es decir,

$$\mathbb{C}((t^{1/N})) = \left\{ \sum_{n \geq M} a_n (t^{1/N})^n \mid a_n \in \mathbb{C}, \text{ para alguna } M \in \mathbb{Z} \right\}$$

Teorema 7.29 (Puiseux)**.** *El polinomio $p(t; x)$ tiene d raíces, contadas con multiplicidad, dentro del campo de series de Puiseux $\mathbb{C}\{\{t\}\}$*

Hecho: En una vecindad del origen, cada rama de la función $t \mapsto (x(t), y(t))$ se puede escribir como

$$x(t) = x_0 t^u + \text{términos mayores en } t$$

$$y(t) = y_0 t^v + \text{términos mayores en } t$$

con $x_0, y_0 \in \mathbb{C}^\times$ y $u, v \in \mathbb{Q}$.

Para encontrar u, v sustituimos $x = x(t)$, $y = y(t)$ en (27.1) y (27.2).

$$f_t(x(t), y(t)) = a_1 t^{v_1} + a_2 x_0 t^{u+v_2} + a_3 x_0 y_0 t^{u+v+v_3} + a_4 y_0 t^{v+v_4} + \dots$$

$$g_t(x(t), y(t)) = b_1 t^{w_1} + b_2 x_0^2 y_0 t^{2u+v+w_2} + b_3 x_0 y_0^2 t^{u+2v+w_3} + \dots$$

Observación 7.30. *Para que $(x(t), y(t))$ sea una raíz del sistema, necesitamos que el término menor se anule en cada una de estas ecuaciones $\Rightarrow$ esto es posible si el término menor en t se alcanza en al menos dos términos. Entonces necesitamos encontrar $(u, v) \in \mathbb{Q}^2$ tal que*

$$\text{mín}\{v_1,\ u + v_2,\ u + v + v_3,\ v + v_4\}, \quad \text{mín}\{w_1,\ 2u + v + w_2,\ u + 2v + w_3\} \tag{7.4}$$

se alcance al menos dos veces. Entonces necesitamos encontrar soluciones de un sistema lineal de ecuaciones y desigualdades.

Ejemplo 7.31.

$$w_1 = 2u + v + w_2 \leq u + 2v + w_3, \quad \textit{o bien}$$

$$w_1 = u + 2v + w_3 \leq 2u + v + w_2, \quad \textit{o bien}$$

$$2u + v + w_2 = u + 2v + w_3 \leq w_1.$$

Entonces v_i, w_i son suficientemente genéricos si el mínimo en (7.4) se alcanza 2 veces pero no 3. $\triangle$

Retomemos el sistema:

$$f(x,y) = a_1 + a_2x + a_3xy + a_4y$$
$$g(x,y) = b_1 + b_2x^2y + b_3xy^2$$

el cual, al agregar potencias de t, se convierte en:

$$f_t = a_1t^{v_1} + a_2xt^{v_2} + a_3xyt^{v_3} + a_4yt^{v_4} \qquad g_t = b_1t^{w_1} + b_2x^2yt^{w_2} + b_3xy^2t^{w_3}, \tag{7.5}$$

para el cual necesitamos encontrar $(u,v) \in \mathbb{Q}^2$ tal que,

$$\text{mín}\{v_1,\ u+v_2,\ u+v+v_3,\ v+v_4\} \quad \text{mín}\{w_1,\ 2u+v+w_2,\ u+2v+w_3\} \tag{7.6}$$

se alcance al menos dos veces.

Observación 7.32. *Un problema que tiene este método es que necesitamos escoger potencias de t que sean pequeñas y genéricas, pero no existe un forma global de hacer esto.*

Si en nuestro ejemplo escogemos

$$v_1 = v_2 = v_3 = v_4 = w_3 = 0, \qquad w_1 = w_2 = 1$$

el sistema (7.5) nos queda:

$$f_t = a_1 + a_2xt + a_3xyt + a_4y$$
$$g_t = b_1t + b_2x^2yt + b_3xy^2,$$

mientras que las ecuaciones en (7.6) se convierten en:

$$\text{mín}\{0,\ u,\ u+v,\ v\}, \qquad \text{mín}\{0,\ 2u+v+1,\ u+2v\}.$$

Para que el mínimo se alcance dos veces en las ecuaciones anteriores se debe de tener que:

(i) Si $u = 0$, entonces $v = \frac{1}{2}$, lo que nos da la solución $(0, \frac{1}{2})$.

(ii) Si $v = 0$, entonces $u = 1$ ó $u = -1$, lo que nos da las soluciones $(1,0)$ y $(-1,0)$.

Las soluciones son $\{(0,\frac{1}{2}),\ (1,0),\ (-1,0)\}$. Para cada solución obtenemos un sistema binomial:

$$\hat{f}(x_0,y_0) = 0 = \hat{g}(x_0,y_0).$$

los cuales están dados en el Cuadro 7.1 abajo.

Estos sistemas binomiales tiene 1, 2, y 1 soluciones respectivamente. Por ejemplo, para $(u,v) = (1,0)$, la solución es:

$$x_0 = \frac{-a_4^2b_1}{a_1^2b_3}, \quad y_0 = \frac{-a_1}{a_4}$$

por lo que la serie de Poisson tiene la forma:

$$x(t) = \frac{-a_4^2b_1}{a_1^2b_3}t + \mathcal{O}(t^2)$$
$$y(t) = \frac{-a_1}{a_4} + \mathcal{O}(t).$$

Solución	$\hat{f}(x_0, y_0)$	$\hat{g}(x_0, y_0)$
$(1, 0)$	$a_1 + a_4 y_0$	$b_1 + b_3 x_0 y_0^2$
$(0, \frac{1}{2})$	$a_1 + a_2 x_0$	$b_1 + b_3 x_0 y_0^2$
$(-1, 0)$	$a_2 x_0 + a_3 x_0 y_0$	$b_2 x_0^2 y_0 + b_3 x_0 y_0^2$

Cuadro 7.1: Sistemas binomiales asociados

Observación 7.33. *Existen métodos para calcular mas términos de la serie de Poisson y obtener más información de las ramas.*

Podemos pensar a f_t y g_t en (7.5) como un sistema de Q ecuaciones en 3 variables, cuya variedad es una curva en $(\mathbb{C}^\times)^3$. El politopo de Newton de estos polinomios:

$$P := \text{conv}\{(0,0,v_1),\ (1,0,v_2),\ (1,1,v_3),\ (0,1,v_4)\},$$
$$Q := \text{conv}\{(0,0,w_1),\ (2,1,w_2),\ (1,2,w_3)\}.$$

La suma de Minkowski, $P + Q$, es un politopo en $\mathbb{R}^3$. Decimos que una faceta F de $P + Q$ es *inferior*, si existe $(u, v) \in \mathbb{R}^2$ tal que $(u, v, 1)$ es un vector normal de $P + Q$ en F que apunta hacia adentro.

Las condiciones de genericidad en $v_i's$ y $w_i's$ son equivalentes a:

1. El politopo $P + Q$ es 3-dimensional.
2. Cada faceta inferior de $P + Q$ es de la forma $F_1 + F_2$, donde:
   (a) F_1 es un vértice de P y F_2 es una faceta de Q.
   (b) F_2 es un vértice de Q y F_1 es una faceta de P.
   (c) F_1 y F_2 son aristas.

La *envolvente inferior* es la unión de las facetas inferiores. Si $\pi\colon \mathbb{R}^3 \to \mathbb{R}^2$ es la proyección en las primeras dos coordenadas, entonces:

$$\pi(P) = New(f), \quad \pi(Q) = New(g) \quad \text{y} \quad \pi(P + Q) = New(f) + New(g).$$

Así, el conjunto de polígonos

$$\{\pi(F) \mid F \text{ es faceta inferior de } P + Q\}$$

nos da un subdivisión de $New(f) + New(g)$ que se llamamos *subdivisión mixta*.

Observación 7.34. *Cada faceta inferior de $P + Q$ es de la forma $F_1 + F_2$*

(a). F_1 es vertice de P y F_2 es una faceta de Q.

(b). F_1 y F_2 son aristas de P y Q, respectivamente.

(c). F_1 es faceta de P y F_2 es vertice de Q.

Si tomamos $\pi\colon \mathbb{R}^3 \to \mathbb{R}^2$, la proyección a las primeras dos coordenadas, entonces

$$\pi(P) = New(f), \;\; \pi(Q) = New(g)$$

$$\pi(P+Q) = New(f) + New(g).$$

Definición 7.35. Definimos a la *subdivisión mixta* como el conjunto:

$$\Delta = \{\pi(F) \mid F \text{ es faceta inferior de } P+Q\}.$$

Lema 7.36. *Sea Δ una subdivisión mixta para f y g. La suma de los volumenes de las celdas mixtas de Δ es igual al volumen mixto:*

$$\text{MixVol}(New(f), New(g)).$$

Demostración. Sean γ y δ dos números reales positivos y consideremos el politopo $\gamma P + \delta Q$ en $\mathbb{R}^3$. Su proyección al plano en $\mathbb{R}^2$ es

$$\begin{aligned} \pi(\gamma P + \delta Q) &= \gamma\pi(P) + \delta\pi(Q) \\ &= \gamma New(f) + \delta New(g). \end{aligned}$$

Denotemos con $V(,\gamma\delta)$ el volumen de este politopo. Este poligono se puede dividir en celdas mixtas $\delta F_1 + \gamma F_2$ donde $F_1 + F_2$ varia entre todas las celdas de Δ. Notemos que $\text{Vol}(\delta F_1 + \gamma F_2) = \delta^2 \text{Vol}(F_1 + F_2)$ si $F_1 + F_2$ es una celda del tipo (a), es igual a $\text{Vol}(\delta F_1 + \gamma F_2) = \delta\gamma \text{Vol}(F_1 + F_2)$ si es celda mixta de tipo (b), e igual a $\text{Vol}(\delta F_1 + \gamma F_2) = \gamma^2 \text{Vol}(F_1 + F_2)$ si es igual a una de tipo (c).

La suma de todos estos volumenes es igual a $V(\gamma, \delta)$. Por lo tanto,

$$V(\delta, \gamma) = \delta^2 V_{(a)} + \delta\gamma V_{(b)} + \gamma^2 V_{(c)}.$$

Con $V_{(b)}$ la suma de los volumenes de las celdas mixtas en Δ. Concluimos que

$$\begin{aligned} V_{(b)} &= V(1,1) - V(1,0) - V(0,1) \\ &= \text{MixVol}(New(f), New(g)). \end{aligned}$$

□

Lema 7.37. *Una pareja $(u, v) \in \mathbb{Q}^2$ es solución de las ecuaciones lineales de minimalidad sí, y sólo sí, $(u, v, 1)$ es el vector normal de una faceta interior mixta de $P + Q$.*

Observación 7.38. *Las celdas mixtas de la subdivisión se expresan de manera única como suma de Minkowski de un segmento de $New(\overline{f})$ y otro $New(\overline{g})$. Donde $\overline{f}$ y $\overline{g}$ son binomios que consisten en dos términos de f y g. Por lo tanto, cada celda mixta de Δ se puede identificar con un sistema de dos binomios*

$$\overline{f} = 0 = \overline{g}.$$

En este caso, podemos reescribir el sistema:

$$f_t(x(t), y(t)) = \overline{f}(x_0, y_0)t^a + \textit{términos mayores.}$$

$$g_t(x(t), y(t)) = \overline{g}(x_0, y_0)t^b + \textit{términos mayores.}$$

Lema 7.39. *Sean (u, v) como en el Lema 2. Las soluciones de $(x_0, y_0) \in (\mathbb{C}^\times)^2$ son las soluciones distintas de cero del sistema binomial*

$$\overline{f} = 0 = \overline{g}.$$

Demostración. (Teorema de Bernstein.)
Mostraremos que $f_t(x, y) = g_t(x, y) = 0$ tiene exáctamente MixVol$(New(f), New(g))$ soluciones en $(\mathbb{K}^\times)^2$, donde $\mathbb{K} = \mathbb{C}\{\{t\}\}$ que es algebraicamente cerrado.

Vimos que en sistemas binomiales, el número de raíces $(x_0, y_0) \in (\mathbb{C}^\times)^2$ del lema 3 coincide con el área de las celdas mixtas de $New(\overline{f}) + New(\overline{g})$.

Cada una de estas raíces define los coeficientes iniciales de una solución $(x(t), y(t))$ de series de Puiseux $f_t = g_t = 0$ de nuestras ecuaciones. Inversamente, del lema 2, cada solución de serie $(x(t), y(t))$ proviene de algunas celdas mixtas de la subdivisión Δ. Concluimos que el número de soluciones de la serie es igual a la suma de área de las celdas mixtas de Δ.

(Cada celda mixta es una paralelogramo y el área es la longitud del vector normal). Del Lema 1, esta cantidad coincide con el MixVol$(New(f), New(g))$.

Argumentos generales de geometría algebráica garantiza que el mismo número de raíces se alcanzan para casi toda selección de coeficientes y que también podemos pasar del campo $\mathbb{K}$ a $\mathbb{C}$ con la substitución $t = 1$. □

Recordemos que $\mathcal{A} \subseteq \mathbb{Z}^n$ finito nos definen exponentes de monomios de Laurent. e.g. $\alpha = (2, -3, -5, 1)$ nos da el monomio: $x_1^2 x_2^{-3} x_3^{-3} x_4^1$. Un polinomio de Laurent con soporte en $\mathcal{A}$ es

$$f = \sum_{\alpha \in \mathcal{A}} c_\alpha x^\alpha.$$

Si $\mathcal{A} = \begin{pmatrix} 0 & 1 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & -1 \end{pmatrix}$, un polinomio de Laurent es

$$f = 1 + 2x + 3xy + 5y + 7^{-1} + 11x^{-1}y^{-1} + 13y^{-1}.$$

El politopo $\Delta_{\mathcal{A}} = \text{conv}(\mathcal{A})$ son los puntos enteros de un hexagono.

Figura 7.2: HAY QUE PONER UNA IMAGEN DEL POLITOPO $\mathcal{A}$.

El polinomio f vive en el anillo de polinomios de Laurent. $\mathbb{C}[x_1^\pm, x_2^\pm, \ldots, x_n^\pm]$ que es el anillo de coordenadas del toro algebraico $(\mathbb{C}^\times)^n$ donde

$$(\mathbb{C}^\times)^n = \{v \in \mathbb{C}^n \mid x_1 \ldots x_n \neq 0\} = \mathcal{V}(x_1 \ldots x_{n+1} - 1) \subseteq \mathbb{A}^{n+1}.$$

El conjunto $\mathcal{A} \subseteq \mathbb{Z}^n$ también nos define una función:

$$\begin{aligned} \varphi_{\mathbb{A}} \colon (\mathbb{C}^\times)^n &\to \mathbb{P}^{\mathcal{A}} := [y^\alpha \mid \alpha \in \mathcal{A}] \\ (x, y) &\mapsto [x^\alpha \mid \alpha \in \mathcal{A}]. \end{aligned}$$

En el ejemplo:

$$\begin{aligned} \varphi_{\mathbb{A}} \colon (\mathbb{C}^\times)^2 &\to \mathbb{P}^{\mathcal{A}} = \mathbb{P}^6 \\ x &\mapsto [1, x, xy, y, x^{-1}, x^{-1}y^{-1}, y^{-1}]. \end{aligned}$$

Si $\Lambda := \sum_{\alpha \in \mathcal{A}} c_\alpha y^\alpha$, con $c_\alpha \in \mathbb{C}$; esto es una forma lineal en $\mathbb{P}^{\mathcal{A}}$. El *pullback* (regrediente)

$$\varphi_{\mathcal{A}}^*(\Lambda) = \sum_{\alpha \in \mathcal{A}} c_\alpha x^\alpha$$

es un polinomio con soporte en $\mathcal{A}$.

Ejercicios

Agrega aquí ejercicios buenos para agarrar callo en geometría.

1. Demuestra algo

Capítulo 8

Geometría Algebraica Numérica

Ejercicios

Agrega aquí ejercicios buenos para agarrar callo en geometría.

1. Demuestra algo

Apéndice

.1. Ordenes parciales en los Naturales

Explicar aqu'i algo de ese rollo