TRABAJO MONOGRÁFICO

Ecuaciones polinomiales y polítopos: una aproximación al teorema de Bernstein.

Mathias Bourel
Orientador: Ángel Pereyra - Centro de Matemática

Febrero 2005

Licenciatura en Matemática
Facultad de Ciencias
Universidad de la República
Uruguay

Resumen

Bernstein probó -en [Ber]- que, genéricamente, la cantidad de raíces sobre el toro algebraico complejo de un sistema de ecuaciones polinomiales con exponentes prefijados depende únicamente de la geometría métrica de los polítopos de Newton asociados a los polinomios del sistema. En este trabajo estudiaremos en detalle el caso, analizado originalmente por Kushnirenko, en el que todos los polinomios tienen el mismo polítopo de Newton. Para ello nos basaremos en ciertos resultados relativos a los polítopos y a una clase particular de variedades algebraicas proyectivas, los cuales serán establecidos en los primeros capítulos.

Palabras claves: polítopos, volumen mixto, polinomio de Ehrhart, polinomio de Hilbert, resultante rala.

Abstract

Bernstein proved -in [Ber]- that given a generic polynomial system with fixed exponent the number of roots lying on the complex algebraic torus depends only on the geometry of the Newton Polytopes of the polynomials. In this work, we study in details the case, originally analized by Kushnirenko where all the polynomials have the same Newton Polytope. In order to accomplish this, we use some results about polytopes and a special kind of projective varieties; these are established on the firsts chapters.

Key Words: polytopes, mixed volume, Ehrhart's polynomial, Hilbert's polynomial, sparse resultant.

Índice general

Introducción

La obtención de información sobre las raíces de un polinomio fue y es un problema muy estudiado por los matemáticos. Además de que el problema es interesante de por sí, diversas situaciones, matemáticas o no, se modelan usando polinomios y tener información sobre ciertas raíces de los mismos suele ser de utilidad. Basta pensar en la teoría de Galois para tener una impresión de las dificultades que plantea la determinación de las raíces de una ecuación polinomial. Los sistemas de ecuaciones polinomiales en varias variables son más complejos todavía; no obstante, sus apariciones en las aplicaciones son muy frecuentes. Citemos por ejemplo: el estudio de extremos relativos condicionados por el método de los multiplicadores de Lagrange y la resolución de sistemas lineales con un elevado número de ecuaciones. La dificultad en la "determinación" de las soluciones de tales sistemas hace que sea muy importante el siguiente problema: ¿cuántas soluciones complejas tiene un sistema de ecuaciones polinomiales con exponentes prefijados? El conocido teorema de Bézout da una cota para el caso en que haya un número finito de soluciones, pero esta cota puede no ser buena si los polinomios de grados prefijados tienen pocos monomios. Tales sistemas se denominan *sistemas ralos*, señalando el interés por los sistemas con exponentes "dispersos". Acompañando el auge de las ciencias computacionales en las últimas décadas, se ha venido desarrollando el estudio de este tipo de sistemas de ecuaciones con el propósito de obtener algoritmos eficientes. El área no sólo es investigada por matemáticos de diversas orientaciones (algebristas, analistas numéricos, probabilistas, etc) sino también por ingenieros, informáticos, físicos, etc..

En la segunda mitad de la década de los años 70, D. Bernstein, A. Kushnirenko y A. Khovanskii obtuvieron una cota más precisa que la del teorema de Bézout para el número de soluciones complejas de un sistema polinomial ralo. Sin embargo esta cota es para soluciones que no tengan ninguna coordenada nula puesto que estudiaron la cantidad de ceros comunes de polinomios de Laurent. Más concretamente, probaron que la cantidad "esperada" de soluciones de los sistemas que tengan polinomios de Laurent con exponentes prefijados depende únicamente del "espacio ocupado" por éstos, el que se mide con el volumen mixto de las envolventes convexas de los exponentes de los polinomios.

En este trabajo estudiaremos el caso particular en que todos los polinomios tienen el mismo conjunto de exponentes prefijados. Este resultado se conoce como el teorema de Kushnirenko. La prueba, que sigue en líneas generales la de [**6**], es muy interesante pues combina nociones de geometría de los polítopos con nociones de geometría algebraica.

El trabajo está organizado de la manera que sigue. En el capítulo 1 definimos el polítopo de Newton de un polinomio de Laurent, que es el polítopo formado por la envolvente convexa de los exponentes de ese polinomio. Presentaremos por lo tanto algunos resultados de geometría convexa referidos a los polítopos en general. En el capítulo 2 damos un compendio de geometría algebraica específicamente referido a las variedades algebraicas afines y a las proyectivas, para luego estudiar un tipo particular de variedad proyectiva definida a partir de un subconjunto finito de vectores con coordenadas enteras. Para la demostración del teorema de Kushnirenko necesitaremos contar la cantidad de puntos con coordenadas enteras dentro de un polítopo y de sus homotetizados, es lo que hacemos en el capítulo 3, estudiando las propiedades del polinomio de Ehrhart de un polítopo. El capítulo 4 es el capítulo "medular" del trabajo; demostramos el teorema de Kushnirenko que relaciona la cantidad "esperada" de raíces de un sistema de polinomios de Laurent - en el cual todos los polinomios tienen los mismos exponentes prefijados - con el volumen del polítopo de

Newton asociado. En el capítulo 5 comentamos el teorema de Bernstein que generaliza al teorema de Kushnirenko. Finalmente hemos incluido un apéndice donde están probados ciertos resultados específicos de álgebra lineal y de geometría combinatoria que fueron utilizados.
Este trabajo surgió a raíz de la participación del autor y su orientador en la Escuela CIMPA "Sistemas de Ecuaciones Polinomiales: de la Geometría Algebraica a las Aplicaciones Industriales", Julio 2003, Universidad de Buenos Aires, Argentina.

Agradezco a A. Pereyra y a quienes incidieron directamente en la elaboración del trabajo:
los participantes de los seminarios de geometría algebraica y álgebra computacional del CMAT, en particular a V. Ferrer y A. Rittatore. Agradezco también A. Dickenstein por la ayuda brindada y la facilitación de materiales.

CAPíTULO 1

Polítopos

En este primer capítulo analizaremos la geometría de los conjuntos convexos, en particular, la de los conjuntos poliédricos y acotados, los polítopos, que nos serán de utilidad a la hora de estudiar los sistemas de ecuaciones polinomiales. El lector interesado en completar y ampliar el contenido del capítulo podrá recurrir a [**2**] o [**3**]. En todo el trabajo consideraremos $\mathbb{R}^n$ como $\mathbb{R}$-espacio vectorial, con la suma y el producto por un escalar habituales, y trabajaremos siempre con el producto interno usual y la norma inducida por él.

1. Generalidades

Decimos que un conjunto $C \subset \mathbb{R}^n$ es *convexo* si el segmento que une dos puntos cualesquiera de C está incluido en C, esto es

$$\forall\ A \text{ y } B \in C, \quad tA + (1-t)B \in C\ \forall\ t \in [0,1].$$

Dado un conjunto cualquiera $X \subset \mathbb{R}^n$ nos interesamos en el menor subconjunto convexo que lo contiene.

DEFINICIÓN 1.1. Sea X un subconjunto de $\mathbb{R}^n$. Definimos la *envolvente convexa de* X como el conjunto:

$$Conv(X) = \left\{ \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m :\ x_j \in X,\ \lambda_j \in \mathbb{R}_{\geq 0}\ \forall j = 1, \cdots, m, \sum_{j=1}^{m} \lambda_j = 1 \right\}$$

Las combinaciones lineales del tipo $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m$ con $x_j \in X,\ \lambda_j \geq 0$, $\forall j = 1, \cdots, m$ tales que $\sum_{j=1}^{m} \lambda_j = 1$ se llaman *combinaciones lineales convexas* de elementos de X.

Con esta definición tenemos los siguientes resultados:

1. AFIRMACIÓN: La envolvente convexa de dos puntos A y B es el segmento $[A, B]$.

   DEMOSTRACIÓN. Si $X = \{A, B\}$, entonces $Conv(X) = \{\lambda A + \mu B,\ \lambda \geq 0,\ \mu \geq 0,\ \lambda + \mu = 1\}$.
   Por lo tanto $Conv(X) = \{\lambda A + (1-\lambda)B,\ 0 \leq \lambda \leq 1\} = [A, B]$. □

2. AFIRMACIÓN: Si X es convexo entonces $X = Conv(X)$.

   DEMOSTRACIÓN. Todo conjunto X está contenido en su envolvente convexa o sea $Conv(X)$ ya $x = 1.x \in Conv(X)$.

   Supongamos ahora que X es convexo y probemos que $Conv(X) \subset X$. Razonaremos por inducción completa sobre la cantidad de puntos de las combinaciones lineales de X.

   Si $n = 2$, como X es convexo, el segmento que une dos puntos $x, x' \in X$ está contenido en X, esto es $tx + (1-t)x' \subset X, \forall\ t \in [0,1]$.

   Supongamos que $n \geq 2$ y que $\sum_{i=1}^{n} \lambda_i x_i \in Conv(X),\ x_i \in X, \lambda_i \in \mathbb{R}_{\geq 0}$, $\forall\ i = 1, \cdots, n$ donde $\sum_{i=1}^{n} \lambda_i = 1$.

Sea $\sum_{j=1}^{n+1} \mu_j y_j,\ y_j \in X,\ \mu_j \in \mathbb{R}_{\geq 0},\ \forall\ j = 1, \cdots, n+1$, y tal que $\sum_{j=1}^{n+1} \mu_j = 1$.
Deducimos que al menos uno de los coeficientes μ_j es menor estricto que 1.
Supongamos entonces que $0 \leq \mu_{n+1} < 1$. Por ser $\sum_{j=1}^{n+1} \mu_j = 1$ tenemos:
$\frac{\mu_1}{1-\mu_{n+1}} + \frac{\mu_2}{1-\mu_{n+1}} + \cdots + \frac{\mu_n}{1-\mu_{n+1}} = 1$ y $\frac{\mu_k}{1-\mu_{n+1}} \geq 0,\ \forall\ k = 1, \cdots, n$.
Por hipótesis de inducción, $z = \sum_{j=1}^{n} \frac{\mu_j}{1-\mu_{n+1}} y_j \in X$, luego

$$\sum_{j=1}^{n+1} \mu_j y_j = \sum_{j=1}^{n} \mu_j y_j + \mu_{n+1} y_{n+1} = (1-\mu_{n+1}) z + \mu_{n+1} y_{n+1} \in X$$

pues z e y_{n+1} son elementos de X y $(1-\mu_{n+1}) < 1,\ \mu_{n+1} < 1$. □

3. Afirmación: $Conv(X)$ es convexo.

Demostración. Sean $x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r \in Conv(X),\ x_i \in X,\ \sum \lambda_j = 1$ e $\mu_1 y_1 + \mu_2 y_2 + \cdots + \mu_s y_s \in Conv(X),\ y_i \in X,\ \sum \mu_j = 1$.
Si $t \in (0,1)$ entonces:

$$tx + (1-t)y = t\lambda_1 x_1 + t\lambda_2 x_2 + \cdots + t\lambda_r x_r + (1-t)\mu_1 y_1 + (1-t)\mu_2 y_2 + \cdots + (1-t)\mu_s y_s.$$

Tenemos que $t\lambda_i \geq 0\ \forall\ i = 1, \cdots, r$, $(1-t)\mu_j \geq 0\ \forall\ j = 1, \cdots, s$
y $\sum_{i=1}^{r} t\lambda_i + \sum_{j=1}^{s} (1-t)\mu_j = t + (1-t) = 1$, luego $tx + (1-t)y \in Conv(X)$. □

4. Afirmación: $Conv(X)$ es el mínimo conjunto convexo que contiene a X.

Demostración. El mapa $X \mapsto Conv(X)$ es creciente con la inclusión. Sea Y convexo tal que $X \subset Y \subset Conv(X)$.
Luego $Conv(X) \subset Conv(Y) = Y \subset Conv(Conv(X)) = Conv(X)$ porque Y es convexo (afirmación 3.). Entonces $Y = Conv(X)$ y $Conv(X)$ es el menor conjunto convexo que contiene al conjunto X. □

Observación 1.2. Se puede probar también que:

$$Conv(X) = \cap \{C/\ C \text{ convexo y } X \subset C\}.$$

Lema 1.3. *Sea $\mathcal{A} = \{x_1, x_2, \cdots, x_r\} \subset \mathbb{R}^n$ e $y \in \mathbb{R}^n$. Entonces $Conv(y + \mathcal{A}) = y + Conv(\mathcal{A})$,*

Demostración. $\mathcal{A} = \{x_1, x_2, \cdots, x_r\}$ e $y + \mathcal{A} = \{y + x_1, \cdots, y + x_r\}$.
Luego $\lambda_1(y + x_1) + \lambda_2(y + x_2) + \cdots + \lambda_r(y + x_r) = \sum_{i=1}^{r} \lambda_i y + \lambda_1 x_1 + \cdots + \lambda_r x_r$
$= y + \lambda_1 x_1 + \cdots + \lambda_r x_r \in y + Conv(\mathcal{A})$. □

Definición 1.4. Un *polítopo* Q es la envolvente convexa de un conjunto finito de puntos de $\mathbb{R}^n$ o sea si $\mathcal{A} = \{x_1, x_2, \cdots, x_t\} \subset \mathbb{R}^n$, el polítopo Q definido por $\mathcal{A}$ es el conjunto

$$Q = Conv(\mathcal{A}) = \left\{ \lambda_1 x_1 + \lambda_2 x_2 + \cdots, +\lambda_t x_t : \lambda_i \geq 0, \sum_{i=1}^{t} \lambda_i = 1 \right\}.$$

Ejemplos 1.5. Más adelante quedará plenamente justificado que:

- Un polítopo en $\mathbb{R}$ es un segmento o un punto.
- Un polítopo en $\mathbb{R}^2$ es un punto, un segmento o un polígono convexo.
- Un polítopo en $\mathbb{R}^3$ es un punto, un segmento o un polígono convexo incluido en un plano, o un poliedro tridimensional.

Definición 1.6. Sea Q un polítopo en $\mathbb{R}^n$.
La *dimensión* de Q es la dimensión del menor subespacio afín de $\mathbb{R}^n$ que contiene a Q y la notaremos $dim(Q)$.

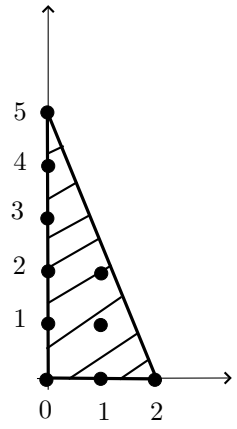


Figura 1. Envolvente convexa de $\mathcal{A} = \{(0,0),(2,0),(0,5),(1,1)\}$.

EJEMPLO 1.7. Sea $\mathcal{A} = \{(0,0),(2,0),(0,5),(1,1)\} \subset \mathbb{R}^2$. Es claro, como lo muestra la figura 1, que $Conv(\mathcal{A})$ es el triángulo determinado por los vértices $(0,0),(2,0)$ y $(0,5)$ ya que $(1,1) = \frac{3}{10}(0,0) + \frac{1}{2}(2,0) + \frac{1}{5}(0,5)$, con $\frac{3}{10} + \frac{1}{2} + \frac{1}{5} = 1$.

DEFINICIÓN 1.8.

1. Decimos que un polítopo es un *n-símplice* o *símplice de dimensión n* de $\mathbb{R}^n$ si es la envolvente convexa de $n+1$ puntos $x_1, x_2, \cdots, x_{n+1}$ tales que $\{x_2 - x_1, x_3 - x_1, \cdots, x_{n+1} - x_1\}$ es una base de $\mathbb{R}^n$ como espacio vectorial.
2. El *símplice elemental (o fundamental)* de $\mathbb{R}^n$ es la envolvente convexa de $0_{\mathbb{R}^n}, e_1, \cdots, e_n$, donde e_j es el j-ésimo vector canónico de $\mathbb{R}^n$.
3. En general decimos que un polítopo es un *r-símplice* o *símplice de dimensión r* de $\mathbb{R}^n$, con $r \leq n$, si es la envolvente convexa de $r+1$ puntos $x_1, \cdots x_{r+1}$ tal que $\{x_2 - x_1, \cdots, x_{r+1} - x_1\}$ es un conjunto linealmente independiente de $\mathbb{R}^n$.

   En este caso decimos que $x_1, \cdots x_{r+1}$ son *afinmente independientes*.

DEFINICIÓN 1.9. Dado un vector $m = (m_1, m_2, \cdots, m_n) \in \mathbb{Z}^n_{\geq 0}$ el peso de m se define como $|m| = \sum_{i=1}^n m_i$.
Sea d un número entero positivo. Definimos el conjunto:

$$\mathcal{A}_d = \left\{ m \in \mathbb{Z}^n_{\geq 0} : \ |m| \leq d \right\},$$

es decir $\mathcal{A}_d$ es el conjunto de los puntos con coordenadas enteras positivas o nulas de peso a lo sumo d.
Consideramos su envolvente convexa Q_d:

$$Q_d = Conv(\mathcal{A}_d).$$

En la figura 2 presentamos ejemplos de tales conjuntos $\mathcal{A}_d$ y Q_d.
Obsérvese que Q_1 es el símplice fundamental de $\mathbb{R}^n$.

AFIRMACIÓN: $\underline{Q_d = \{(a_1, a_2, \cdots, a_n) \in \mathbb{R}^n : \ a_i \geq 0, \sum_{i=1}^n a_i \leq d\}.}$

DEMOSTRACIÓN. Es claro que $\mathcal{A}_d \subset \{(a_1, a_2, \cdots, a_n) \in \mathbb{R}^n : \ a_i \geq 0, \sum_{i=1}^n a_i \leq d\}$, luego
$Q_d = Conv(\mathcal{A}_d) \subset Conv\left(\{(a_1, a_2, \cdots, a_n) \in \mathbb{R}^n : \ a_i \geq 0, \sum_{i=1}^n a_i \leq d\}\right)$
$= \{(a_1, a_2, \cdots, a_n) \in \mathbb{R}^n : \ a_i \geq 0, \sum_{i=1}^n a_i \leq d\}$ pues este conjunto es convexo.

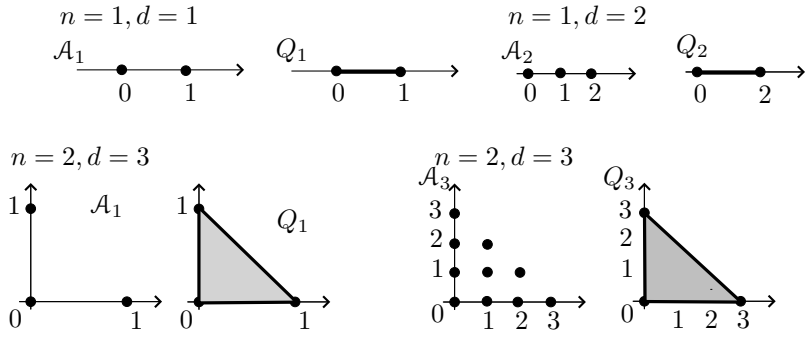


Figura 2. $\mathcal{A}_d$ y Q_d en casos particulares.

Por otro lado sea $(a_1, \cdots, a_n) \in \{(y_1, y_2, \cdots, y_n) \in \mathbb{R}^n : \ y_i \geq 0, \sum_{i=1}^n y_i \leq d\}$.
$(a_1, \cdots, a_n) = \frac{a_1}{d}(d, 0, \cdots, 0) + \frac{a_2}{d}(0, d, 0, \cdots, 0) + \cdots + \frac{a_n}{d}(0, \cdots, 0, d) + \frac{d - \sum a_j}{d}(0, \cdots, 0)$. Observemos que $0 \leq \frac{a_j}{d} \leq 1$ y que $\sum \frac{a_j}{d} + \frac{d - \sum a_j}{d} = 1$. Por lo cual $(a_1, \cdots, a_n)$ es una combinación lineal convexa de vectores de $\mathcal{A}_d$ por lo tanto $Q_d \subset Conv(\mathcal{A}_d)$. □

Es claro que Q_d es un n-símplice pues $Q_d = Conv\left((0, \cdots, 0), (d, 0, \cdots, 0), \cdots, (0, \cdots, 0, d)\right)$ donde $\{(d, 0, \cdots, 0), (0, d, 0, \cdots, 0), (0, \cdots, 0, d)\}$ es una base de $\mathbb{R}^n$. Más aún $Q_d = dQ_1$ donde Q_1 es el símplice elemental de $\mathbb{R}^n$.

A todo polítopo Q de $\mathbb{R}^n$ le podemos asignar un volumen $Vol_n(.)$, el volumen euclídeo de los conjuntos medibles de $\mathbb{R}^n$, definido por:

$$Vol_n(Q) = \int \int \cdots \int_Q 1 dx_1 dx_2 \cdots dx_n$$

donde $x_1, x_2, \cdots, x_n$ son las coordenadas en $\mathbb{R}^n$.

Observar que un polítopo Q de $\mathbb{R}^n$ tiene volumen positivo si y sólo si su dimensión es n. En la notación $Vol_n(.)$ el subíndice n indica que el volumen es n-dimensional. Por ejemplo, si consideramos un polígono Q en $\mathbb{R}^2$ su volumen $Vol_2(Q)$ es positivo pero si se mira el mismo polígono Q en el plano xOy de $\mathbb{R}^3$, $Vol_3(Q) = 0$.

AFIRMACIÓN: $Vol_n(Q_d) = \frac{d^n}{n!}$.

DEMOSTRACIÓN. Consideramos el mapa $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ dado por

$$\phi(x_1, x_2, \cdots, x_n) = (1 - x_1, x_1(1 - x_2), x_1 x_2(1 - x_3), \cdots, x_1 x_2 \cdots x_{n-1}(1 - x_n)).$$

Si C es el cubo unidad en $\mathbb{R}^n$ observamos que:
$\phi(C) \subset Q_1 = \{(a_1, a_2, \cdots, a_n) : \ a_i \geq 0, \sum a_i \leq 1\}$ ya que

$$(1 - x_1) + x_1(1 - x_2) + x_1 x_2(1 - x_3) + \cdots + x_1 \cdots x_{n-1}(1 - x_n) \leq 1.$$

Pero también $Q_1 \subset \phi(C)$. Si $(y_1, \cdots, y_n) \in Q_1$ nos preguntamos si existe un vector $(x_1, \cdots, x_n) \in C$ tal que $\phi(x_1, \cdots, x_n) = (y_1, \cdots, y_n)$. Esto equivale a estudiar la compatibilidad del sistema:

$$\begin{cases} y_1 = 1 - x_1 \\ y_2 = x_1(1 - x_2) \\ \vdots \\ y_n = x_1 x_2 \cdots x_{n-1}(1 - x_n) \end{cases}$$

Dejamos al lector interesado el estudio de este sistema y la conclusión de que siempre es compatible.

Por lo tanto $\phi(C) = Q_1$ y ϕ es un cambio de variable ya que si $\phi = (f_1, f_2, \cdots, f_n)$ la matriz jacobiana J_ϕ de ϕ es:

$$J_\phi = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 \\ 1-x_2 & -x_1 & 0 & \cdots & 0 \\ x_2(1-x_3) & x_1(1-x_3) & -x_1x_2 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ & & & & -x_1x_2\cdots x_{n-1} \end{pmatrix}$$

y $det(J_\phi) = (-1)^n x_1^{n-1} x_2^{n-2} \cdots x_{n-1}$ se anula sólo en $\cup_{i=1}^n x_i$.
Luego por la fórmula del cambio de variable:

$$Vol_n(Q_1) = \int \cdots \int_{Q_1} 1 dx_1 \cdots dx_n = \int \cdots \int_C det(J_\phi) dx_1 \cdots dx_n =$$

$$= \int \cdots \int_C (-1)^n x_1^{n-1} x_2^{n-2} \cdots x_{n-1} = \frac{1}{n!}.$$

Como $dQ_1 = Q_d$, se tiene que:

$$Vol_n(Q_d) = Vol_n(dQ_1) = d^n Vol_n(Q_1) = d^n \times \frac{1}{n!} = \frac{d^n}{n!}.$$

□

2. Caras de un polítopo

Si miramos un polítopo tridimensional de $\mathbb{R}^3$, el mismo tiene subconjuntos particulares que llamamos vértices (que son puntos), aristas (que son segmentos que conectan algunos pares de vértices) y muros (que son polígonos incluidos en planos). Todos estos conjuntos se llaman *caras* del polítopo Q. Veremos como podemos formalizar esta noción y mostraremos en particular que las caras de un polítopo son polítopos y que su número es finito.

LEMA 1.10. *Sea K un conjunto cerrado y convexo en $\mathbb{R}^n$. Para todo $x \in \mathbb{R}^n$, existe un único $x' \in K$ tal que*

$$||x - x'|| = inf_{y \in K} ||x - y|| = d(x, K).$$

DEMOSTRACIÓN. La existencia de tal elemento x' es clara ya que Q es un conjunto cerrado. Probemos la unicidad. Supongamos que existen x' y x'' en K, distintos, tales que $d(x, K) = ||x - x'|| = ||x - x''||$. Consideramos el triángulo isosceles $xx'x''$ y p el punto medio del segmento $[x'x'']$, definido como $p = \frac{1}{2}(x' + x'')$, al ser K convexo tenemos que $p \in K$ y satisface $||x - p|| < ||x - x'|| = inf_{y \in K} ||x - y||$ lo cual es absurdo. □

DEFINICIÓN 1.11. En el contexto del lema 1.10 definimos la función $p_K : \mathbb{R}^n \longrightarrow K$ dada por $p_K(x) = x'$.

OBSERVACIÓN 1.12. Es claro que $p_K(x) = x \Leftrightarrow x \in K$. En particular p_K es sobreyectiva.

DEFINICIÓN 1.13.

1. Un *hiperplano afín* H de $\mathbb{R}^n$ es un espacio $(n-1)$-dimensional definido por:
$$\{x \in \mathbb{R}^n / \langle x, u \rangle = a\}$$
donde $a \in \mathbb{R}$ y u es un vector no nulo de $\mathbb{R}^n$.
2. Un hiperplano de $\mathbb{R}^n$ se llama *hiperplano de apoyo o hiperplano de soporte* de un conjunto cerrado y convexo $K \subset \mathbb{R}^n$ si:
   - H es un hiperplano,
   - $K \cap H \neq \emptyset$ y $K \subset H^-$ o $K \subset H^+$ donde $H^- = \{x \in \mathbb{R}^n / \langle x, u \rangle \leq a\}$ y $H^+ = \{x \in \mathbb{R}^n / \langle x, u \rangle \geq a\}$.

   También se dice respectivamente que H^- o H^+ es un *semiespacio de apoyo* de K. Observemos que es posible que $K \subset H$.

En el correr del trabajo supondremos que, dado un conjunto cerrado y convexo K, si H es un hiperplano de apoyo de K entonces $K \subset H^-$, esto equivale a suponer que el vector u apunta hacia H^+. Diremos que u es un *vector normal saliente* de K.
Si K no está contenido en H, existe un único vector unitario u que es normal saliente de K respecto de H.

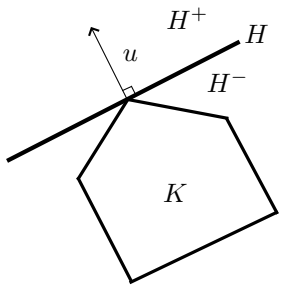


Figura 3.

Observación 1.14. Un hiperplano es el trasladado de un subespacio vectorial $(n-1)$-dimensional de $\mathbb{R}^n$ del tipo $\{x \in \mathbb{R}^n / \langle x, u\rangle = 0\} = [u]^\perp$ para $u \neq 0$. O sea si $a \in \mathbb{R}$, $\{x \in \mathbb{R}^n / \langle x, u\rangle = a\}$ $= [u]^\perp + w$ con w un cierto vector de $\mathbb{R}^n$ tal que $\langle w, u\rangle = a$

Demostración. $(\subset)$: Sea x tal que $\langle x, u\rangle = a$. Luego $\exists w \in \mathbb{R}^n / \langle x, u\rangle = \langle w, u\rangle$ pues $\langle \cdot, u\rangle : \mathbb{R}^n \longrightarrow \mathbb{R}$ es sobreyectivo.
Entonces $\langle x - w, u\rangle = 0$, luego $x - w \in [u]^\perp$ o sea $x \in [u]^\perp + w$.
$(\supset)$: Sea $x \in [u]^\perp + w$ con $\langle u, w\rangle = a$. Entonces:

$$\langle x, u\rangle = \langle y + w, u\rangle = \langle y, u\rangle + \langle w, u\rangle = 0 + \langle w, u\rangle = a.$$

□

Lema 1.15. *Suponemos que $\emptyset \neq K \subset \mathbb{R}^n$ es un conjunto cerrado y convexo.*
Para todo $x \in \mathbb{R}^n \setminus K$ el hiperplano H que contiene a $x' = p_K(x)$ y es perpendicular a la recta (xx') es un hiperplano de apoyo de K. Más aún $H = \{y \in \mathbb{R}^n / \langle y, u\rangle = 1\}$ donde $u = \frac{x-x'}{\langle x', x-x'\rangle}$.

Demostración. Es claro que el hiperplano $H = \{y \in \mathbb{R}^n / \langle y, u\rangle = 1\}$ es perpendicular a $x - x'$ y contiene a x'. Por otro lado, como $x \neq x'$ ya que $x' \notin K$ se tiene que $\langle x - x', x - x'\rangle > 0$ por lo tanto $\langle x, x - x'\rangle > \langle x', x - x'\rangle$, es decir $x \in H^+$.

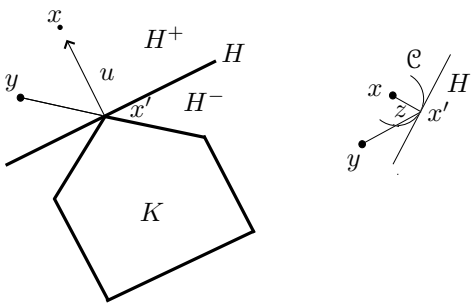


Figura 4.

Supongamos que H no es un hiperplano de apoyo de Q, o sea existe $y \in K \cap (H^+ \setminus H)$ con $y \neq x$. Si $\mathcal{C}$ es la bola cerrada de centro x y radio $||x - x'||$ existiría $z \in [y, x'] \cap \mathcal{C}$. Al ser K convexo el segmento $[y, x']$ está contenido en K, luego $||x - z|| < ||x - x'||$ lo cual es absurdo. □

Observación 1.16. El lema 1.15 se suele llamar "lema de separación".

Teorema 1.17. *Un subconjunto K propio, cerrado y convexo de $\mathbb{R}^n$ es la intersección de sus semi-espacios de apoyo.*

Demostración. Por el lema 1.15, sabemos que existe un hiperplano de apoyo para K, para el cual $K \subset H^-$. Sea $K' = \cap H^-$ donde la intersección se toma sobre todos los semi-espacios de apoyo. Es claro que $K \subset K'$. Supongamos que existe $x \in K' \setminus K$. Entonces $p_K(x) \neq x$. Pero sabemos que existe un hiperplano perpendicular al segmento $[x, p_K(x)]$, que pasa por $p_K(x)$ y separa x de K, luego x no pertenecería a $\cap H^-$, lo cual es absurdo. □

Veremos más adelante en el teorema 1.31, que en el caso de los polítopos, no es necesario tomar la intersección sobre todos los semi-espacios de apoyo del conjunto K: podemos reducirnos a una intersección finita.

Definición 1.18. Sea K un conjunto compacto y convexo de $\mathbb{R}^n$. La función:

$$\begin{array}{rccc} h_K: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & u & \mapsto & sup_{x \in K}\langle x, u\rangle \end{array}$$

se llama *función de soporte* del conjunto K.

En realidad al ser K un compacto, este supremo es un máximo.

Lema 1.19. *Sea K un subconjunto compacto y convexo de $\mathbb{R}^n$.*

1. *Para todo vector $u \neq 0_{\mathbb{R}^n}$, el hiperplano $H_K(u) = \{x \in \mathbb{R}^n / \ \langle x, u\rangle = h_K(u)\}$ es un hiperplano de apoyo de K.*
2. *Todo hiperplano de apoyo de K se puede escribir como en el item anterior.*

Demostración.
1. K es compacto y la función $\langle \cdot, u\rangle : \mathbb{R}^n \longrightarrow \mathbb{R}$ es continua por lo cual existe $x_0 \in K$ tal que $\langle x_0, u\rangle = h_K(u) = sup_{x \in K}\langle x, u\rangle$. En efecto si $y \in K$ entonces $\langle y, u\rangle \leq \langle x_0, u\rangle$, es decir, $K \subset H_K(u)^-$.
2. Sea $H = \{x \in \mathbb{R}^n / \ \langle x, u\rangle = \langle x_0, u\rangle\}$ un hiperplano de apoyo de K en $x_0 \in K$ donde u es un vector normal saliente de K, luego $\langle x_0, u\rangle = sup_{x \in K}\langle x, u\rangle = h_K(u)$.

□

Definición 1.20. Sea H un hiperplano de apoyo de un polítopo Q de $\mathbb{R}^n$.
Decimos que F es *una cara* de Q si $F = Q \cap H$.
En particular:
si $dim(F) = 0$ decimos que F es un *vértice* de Q,
si $dim(F) = 1$ decimos que F es una *arista* de Q,
y cuando Q tiene dimensión n, si $dim(F) = n - 1$, decimos que F es un *muro* de Q.

Notaremos a $Vert(P)$ al conjunto de vértices del polítopo P.

Veremos que los polítopos son compactos, y que por lo tanto sus caras también son compactas y convexas ya que son intersección de dos convexos y de un compacto con un cerrado. Nos encaminamos a formalizar este resultado.

Sea Q un polítopo de $\mathbb{R}^n$, u un vector normal saliente de Q.
Consideramos el conjunto $\{x \in \mathbb{R}^n / \ \langle x, u\rangle = h_Q(u)\}$. Entonces, del lema 1.19, deducimos que:

$$F_Q(u) = Q \cap \{x \in \mathbb{R}^n / \ \langle x, u\rangle = h_Q(u)\} \neq \emptyset \ \text{ y } \ Q \subset H_Q(u)^-.$$

Decimos que $F_Q(u)$ es la cara de Q determinada por el vector u.

OBSERVACIÓN 1.21. Por abuso de lenguaje, si $F_Q(u)$ es la cara de Q que determina el vector u, diremos también que u es un vector normal saliente de la cara $F_Q(u)$.

PROPOSICIÓN 1.22. *Sean F_0 y F_1 caras de un polítopo K tal que $F_0 \subset F_1$. Entonces F_0 es una cara de F_1.*

DEMOSTRACIÓN. Supongamos que $F_0 = K \cap H_0$, donde H_0 es un hiperplano de apoyo de K y por lo tanto $F_1 \cap H_0 \subset K \cap H_0 = F_0 \subset F_1 \cap H_0$. Por lo que $F_0 = F_1 \cap H_0$. □

EJEMPLO 1.23. Sea C el cuadrado unidad en $\mathbb{R}^2$. A partir de la figura 5, damos una manera de determinar algunas de las caras de C.

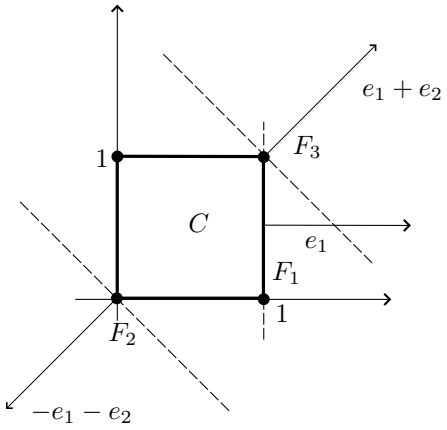


Figura 5. Ejemplo 1.23.

- La cara $F_1 = C \cap \{x \in \mathbb{R}^2 / \langle x, e_1 \rangle = 1\}$ corresponde a un muro de C y $C \subset \{x \in \mathbb{R}^2 / \langle x, e_1 \rangle \leq 1\}$.
- La cara $F_2 = C \cap \{x \in \mathbb{R}^2 / \langle x, e_1 + e_2 \rangle = 2\}$ corresponde a un vértice de C y $C \subset \{x \in \mathbb{R}^2 / \langle x, e_1 + e_2 \rangle \leq 2\}$.
- La cara $F_3 = C \cap \{x \in \mathbb{R}^2 / \langle x, -e_1 - e_2 \rangle = 0\}$ corresponde a un vértice de C y $C \subset \{x \in \mathbb{R}^2 / \langle x, -e_1 - e_2 \rangle \leq 0\}$.

TEOREMA 1.24. *Todo polítopo tiene una cantidad finita de caras. Más aún, estas caras son también polítopos.*

DEMOSTRACIÓN. Sea $P = Conv\{x_1, x_2, \cdots, x_r\}$ un polítopo, $H = \{x \in \mathbb{R}^n / \langle x, u \rangle = \alpha\}$ un hiperplano de apoyo de P y $F = P \cap H$ la cara asociada de manera que $P \subset H^-$. Sin pérdida de generalidad, podemos asumir que $x_1, \cdots, x_s$ son puntos de H y que $x_{s+1}, \cdots, x_r$ son puntos pertenecientes a $intH^-$, donde $intH^-$ es el conjunto de los puntos interiores a H^-. Luego:

- $\langle x_j, u \rangle = \alpha \ \forall \ j = 1, \cdots, s$,
- $\langle x_j, u \rangle < \alpha \ \forall \ j = s+1, \cdots, r$.

Si $x \in P$ entonces $x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r$ con $\sum_{i=1}^{r} \lambda_j = 1$, $\lambda_j \in \mathbb{R}_{\geq 0}$, $\forall j = 1, \cdots, r$. Por lo que, si definimos $\beta_j = \alpha - \langle x_j, u \rangle > 0$, $\forall \ j = s+1, \cdots, r$, tenemos:

$$\langle x, u \rangle = \sum_{j=1}^{r} \lambda_j \langle x_j, u \rangle = \sum_{j=1}^{s} \lambda_j \langle x_j, u \rangle + \sum_{j=s+1}^{r} \lambda_j \langle x_j, u \rangle = \sum_{j=1}^{s} \lambda_j \alpha + \sum_{j=s+1}^{r} \lambda_j (\alpha - \beta_j)$$

$$= \sum_{j=1}^{r} \lambda_j \alpha - \sum_{j=s+1}^{r} \lambda_j \beta_j = \alpha + \sum_{j=s+1}^{r} \lambda_j \beta_j.$$

Entonces:

$$x \in H \iff \sum_{j=s+1}^{r} \lambda_j \beta_j = 0 \iff \lambda_{s+1} = \lambda_{s+2} = \cdots = \lambda_r = 0. \tag{2.1}$$

ya que $\beta_j > 0, \lambda_j \geq 0,\ \forall j = s+1, \cdots, r$.

De la equivalencia (2.1) deducimos que x es una combinación lineal convexa de $x_1, \cdots, x_s$, o sea $H \cap P = Conv\{x_1, \cdots, x_s\}$; luego, $F = H \cap P$ es un polítopo.
Como sólo puede haber una cantidad finita de envolventes convexas de subconjuntos de $\{x_1, \cdots, x_r\}$, el polítopo P tiene una cantidad finita de caras, de donde se deduce la tesis. □

TEOREMA 1.25. *Todo polítopo P es la envolvente convexa de sus vértices.*

DEMOSTRACIÓN. Recordamos que $Vert(P)$ denota al conjunto de los vértices del polítopo P:

- Es claro que $Conv(Vert(P)) \subset P$.
- Si $P = Conv\{x_1, \cdots, x_r\}$, podemos asumir para cada $i = 1, \cdots, r$, el conjunto

$$P_i = Conv\,\{x_1, x_2 \cdots, x_{i-1}, x_{i+1}, \cdots, x_r\} \text{ es tal que } x_i \notin P_i.$$

Si $q_i = p_{P_i}(x_i)$, consideramos, por el lema 1.15, H_i el hiperplano de apoyo del polítopo P_i que pasa por q_i y es normal a $x_i - q_i$ (ver, por ejemplo, la figura 6).

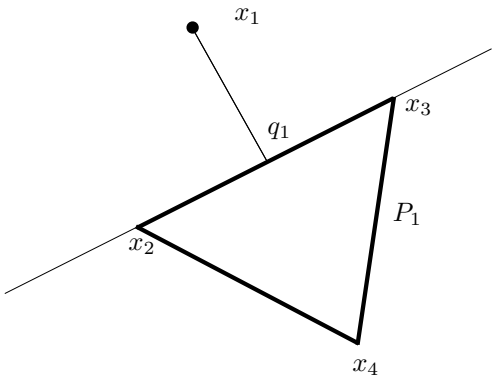


Figura 6.

Trasladando H_i por el vector $x_i - q_i$, obtenemos un hiperplano de apoyo de P, que notaremos H_i' para el cual $\{x_i\} = H_i' \cap P$. En efecto,

$$H_i = \{x \in \mathbb{R}^n / \ \langle x, x_i - q_i \rangle = \langle q_i, x_i - q_i \rangle\},$$

luego

$$H_i' = \{x \in \mathbb{R}^n / \ \langle x, x_i - q_i \rangle = \langle q_i, x_i - q_i \rangle + \langle x_i - q_i, x_i - q_i \rangle\}.$$

Entonces x_i es un vértice de P para cada $i = 1, \cdots, r$, es decir:

$$P = Conv\{x_1, \cdots, x_r\} \subset Conv(Vert(P)).$$

□

DEFINICIÓN 1.26. Si $x_1, \cdots, x_r$ son los vértices de un polítopo P decimos que el conjunto $\{x_1, \cdots, x_r\}$ es un *generador* de P.

OBSERVACIÓN 1.27. 1. Si $\{x_1, \cdots, x_s\}$ es un generador de un muro $\mathcal{F}$ de un polítopo P, entonces $\{x_2 - x_1, \cdots, x_s - x_1\}$ es un generador del trasladado de un hiperplano de apoyo que contiene a $\mathcal{F}$.

2. Si 0 pertenece al polítopo P, $\mathcal{F}$ es un muro de P y u es un vector normal saliente correspondiente a $\mathcal{F}$, entonces la distancia de 0 al muro $\mathcal{F}$ es:

$$d(0,\mathcal{F}) = \frac{h_P(u)}{||u||}.$$

En este trabajo nos interesaremos particularmente por aquellos polítopos cuyos vértices tienen coordenadas enteras.

DEFINICIÓN 1.28. Decimos que un polítopo P es *racional* (en inglés "lattice polytopes") cuando todos sus vértices tienen coordenadas enteras.

Los teoremas anteriores nos dicen que todo polítopo tiene una cantidad finita de caras y que se puede escribir como la envolvente convexa de sus vértices. Nuestro objetivo ahora es probar que podemos recuperar un polítopo P a partir de sus muros; es decir si $F_1, F_2, \cdots, F_N$ son los muros de P y $v_1, v_2, \cdots, v_N$ son los vectores normales salientes correspondientes, veremos que:

$$P = \bigcap_{j=1}^{N} \{x \in \mathbb{R}^n / \ \langle x, v_j \rangle \leq a_j\},$$

para ciertos $a_j \in \mathbb{R}$.

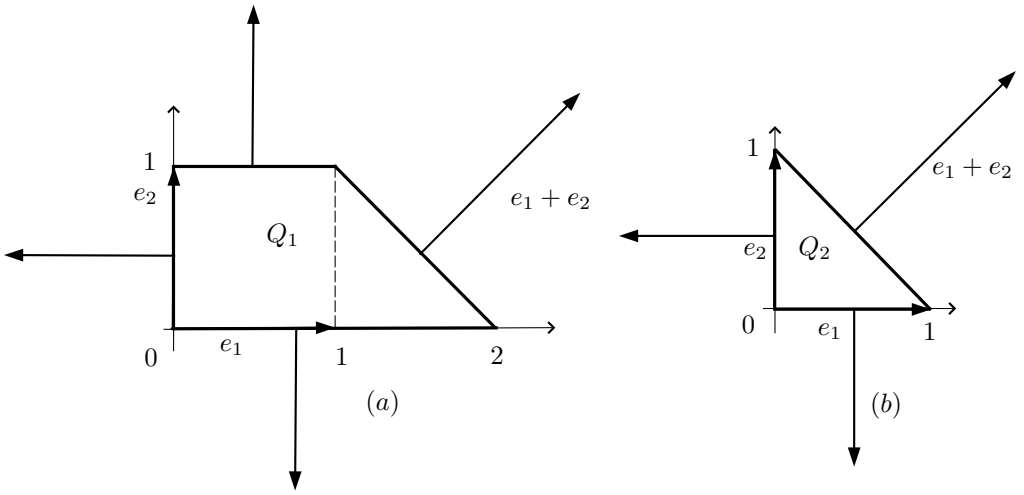


Figura 7.

EJEMPLOS 1.29. En la figura 7:

- en (a) tenemos que $Q_1 = Conv\{(0,0),(1,0),(0,1),(1,1),(2,0)\}$ y podemos escribir:

$Q_1 = \{x/ \ \langle x, e_1 \rangle \leq 0\} \cap \{x/ \ \langle x, -e_1 \rangle \leq 0\} \cap \{x/ \ \langle x, e_2 \rangle \leq 1\} \cap \{x/ \ \langle x, e_1 + e_2 \rangle \leq 2\}$.

- en (b) tenemos que $Q_2 = Conv\{(0,0),(1,0),(0,1)\}$ y podemos escribir:

$$Q_2 = \bigcap_{j=1}^{n} \{x/ \ \langle x, -e_j \rangle \leq 0\} \cap \{x/ \ \langle x, e_1 + \cdots + e_n \rangle \leq 1\}.$$

En la figura 8 tenemos que el cubo unidad $C = Conv\{(0,0),(1,0),(0,1),(1,1)\}$ se puede escribir como:

$$C = \{x/ \ \langle x, -e_2 \rangle \leq 0\} \cap \{x/ \ \langle x, -e_1 \rangle \leq 0\} \cap \{x/ \ \langle x, e_1 \rangle \leq 1\} \cap \{x/ \ \langle x, e_2 \rangle \leq 1\}.$$

Veamos como podemos generalizar lo observado en los ejemplos anteriores.

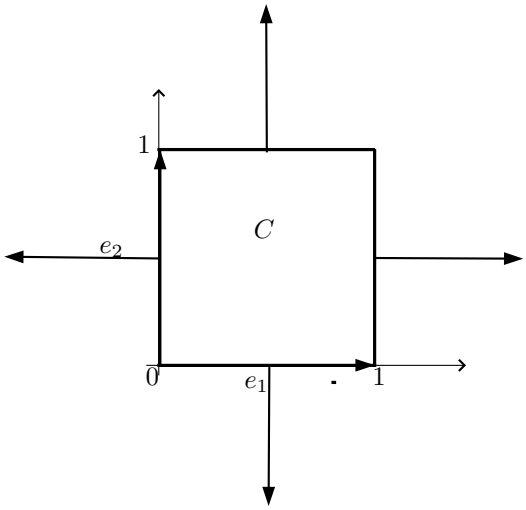


Figura 8. Ejemplo 1.29.

DEFINICIÓN 1.30. Un *conjunto poliédrico* es la intersección de una cantidad finita de semiespacios cerrados de $\mathbb{R}^n$.

TEOREMA 1.31. *Todo polítopo P es un conjunto poliédrico acotado.*

DEMOSTRACIÓN. Podemos suponer que el espacio afín generado por P es $\mathbb{R}^n$ de lo contrario, si la dimensión d es menor que n, trabajamos en el espacio euclídeo $\mathbb{R}^d$ correspondiente.

Sean $F_1, F_2, \cdots, F_s$, siendo $s \geq 0$, (ver teorema 1.24) los muros de P, y $H_1, H_2, \cdots, H_s$ los hiperplanos de apoyo asociados a cada uno de estos muros.

Es obvio que $P \subset \bigcap_{i=1}^{s} H_i^- = P'$.
Supongamos que existe $x_0 \in P' \setminus P$. Sea $\mathcal{A}$ la unión de todos los subespacios afines de $\mathbb{R}^n$ generados por x_0 y a lo sumo $n-1$ vértices de P.

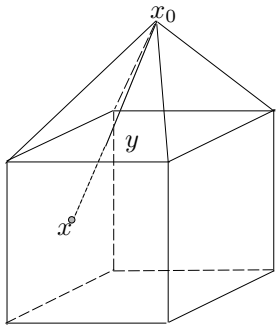


Figura 9. Ilustración de Proposición (1.31) suponiendo que P es un cubo en $\mathbb{R}^3$.

Entonces, por ser unión finita de subespacios, el conjunto $\mathcal{A}$ no tiene puntos interiores. Luego existe un punto $x \in int(P) \setminus \mathcal{A}$. El segmento $[x, x_0]$ no está incluido en $\mathcal{A}$ e intersecta a la frontera de P en un punto y.
Como el borde de P está contenido en la unión de todas las caras propias de P, el punto y se encuentra en una cara F de P. Si la dimensión de F fuese menor que $n-1$, está generada por a lo

sumo $n-1$ vértices y esto implicaría que x pertenece a $\mathcal{A}$, lo cual es una contradicción. Por lo tanto F es un muro de P, es decir es alguno de los F_{j_0} donde $j_0 \in \{1, \cdots, s\}$, y el punto y pertenece al interior relativo de F (nuevamente, en caso contrario, y estaría en una cara de dimensión menor y x tendría que estar en $\mathcal{A}$). Pero el espacio afín generado por F es uno de los hiperplanos H_{i_0} y de esa manera el punto x_0 no pertenecería a P'.
Dicho de otro modo si $x_0 \notin H_{i_0}$ entonces $x_0 \in H_{i_0}^-$, pero $y \in H_{i_0}$ lo cual implica forzosamente que $x_0 \in H_{i_0}$. Y si $x_0 \in H_{i_0}$, $[y, x_0] \in H_{i_0}$ por lo cual $x \in H_{i_0}$ y x no sería interior a P. Entonces necesariamente x_0 pertenece a P. □

Una consecuencia importante del teorema anterior y de su demostración es que todo polítopo es un conjunto compacto, por ser cerrado (intersección finita de cerrados) y acotado. Sin embargo no es cierto que todo compacto sea un polítopo: existen conjuntos compactos que no son convexos.

La proposición anterior nos asegura que todo polítopo es la intersección de los semi-espacios de apoyo definidos por sus muros y que, al tener una cantidad finita de muros, esta intersección es finita.

El recíproco también vale, por lo que, los polítopos son exactamente los conjuntos políedricos acotados.

Teorema 1.32. *Todo conjunto poliédrico y acotado es un polítopo.*

Demostración. Probaremos este resultado por inducción completa en la dimensión del conjunto poliédrico y acotado P. Si la dimensión de P es nula, el resultado es trivial.
Si la dimensión de P es positiva escribimos P como intersección de semi-espacios:
$P = H_1^- \cap H_2^- \cap \cdots \cap H_s^-$. Por hipótesis de inducción suponemos que cada cara propia de P, $F_j = H_j \cap P$ es un polítopo. Sin pérdida de generalidad, una vez más, podemos suponer que P tiene dimensión maximal n (nuevamente la dimensión de un conjunto poliédrico acotado es la del menor subespacio afín que lo contiene).

Es claro que $Conv\left(\bigcup_{j=1}^s F_j\right) \subset P$.

Falta probar la otra inclusión. Es suficiente probarla para el interior de P. Si $x \in int(P)$, consideramos σ una semirrecta de origen x que no sea paralela a ninguno de los hiperplanos H_j (esto es posible dado que hay una cantidad finita de hiperplanos).

Afirmación: $\sigma \cap \partial P$ es un único punto x_σ, donde ∂P denota el borde topológico de P.

Prueba: Sean x el punto en el interior de P, σ una semirrecta con origen en x y Δ la recta que contiene a σ. La intersección $\Delta \cap P$ es un compacto y convexo, por lo tanto es un segmento $[x_\sigma, x'_\sigma]$. Luego $\sigma \cap K$ es el punto x_σ o el punto x'_σ.

Como sabemos que $\partial P \subset \bigcup_{j=1}^s F_j$, el punto x_σ pertenece a alguna de las caras F_{j_σ}. La misma conclusión vale para τ, la semirrecta opuesta a σ.

Como $x \in [x_\tau, x_\sigma]$, entonces $x \in Conv(F_{j_\sigma} \cup F_{j_\tau})$, luego $int(P) \subset Conv(\bigcup_{j=1}^s F_j)$. Por lo tanto $Conv(int(P)) \subset Conv(\bigcup_{j=1}^s F_j)$, luego $P = \partial P \cup Conv(int(P)) \subset Conv(\bigcup_{j=1}^s F_j)$. □

Resumiendo, hemos probado que:

Proposición 1.33. *Si P es un polítopo de $\mathbb{R}^n$ entonces:*

1. *P es convexo y compacto.*
2. *P tiene una cantidad finita de caras.*
3. *P es la envolvente convexa de sus vértices.*
4. *P es un conjunto poliédrico acotado.*

Terminaremos esta parte acerca las propiedades de los polítopos probando dos resultados que permiten descomponer polítopos que pueden tener una forma compleja en unión de polítopos más "sencillos".

Teorema 1.34 (Carathéodory). *Sea M un subconjunto de $\mathbb{R}^n$. La envolvente convexa de M es la unión de todas las envolventes convexas de subconjuntos de M que contienen a lo sumo $n+1$ vectores.*

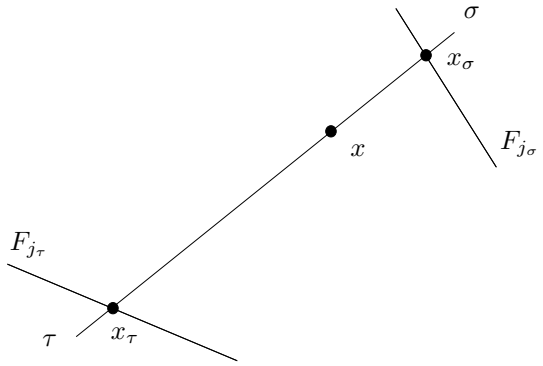


Figura 10. Demostración del teorema 1.32.

DEMOSTRACIÓN. Sea $x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r$, un elemento de la envolvente convexa de M, siendo r el entero más chico para el cual x es una combinación lineal convexa de elementos de M. Supongamos, por absurdo, que $r \geq n+2$. Entonces existen $\mu_1, \mu_2, \cdots, \mu_r$ no todos nulos tales que $\mu_1 x_1 + \cdots + \mu_r x_r = 0$. Además podemos tomar tal combinación lineal de manera que $\mu_1 + \mu_2 + \cdots + \mu_r = 0$. En efecto consideramos la transformación lineal

$$\varphi : \mathbb{R}^r \longrightarrow \mathbb{R}^n$$
$$(\mu_1, \mu_2, \cdots, \mu_r) \mapsto \mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_r x_r$$

Sabemos que se cumple que $dimKer(\varphi) + dimIm(\varphi) = dim\mathbb{R}^r = r$. Luego como $dimIm(\varphi) \leq n$ y $r \geq n+2$ se tiene que $dimKer(\varphi) \geq 2$.
Sea H el subespacio definido por:

$$H = \{(\mu_1, \mu_2, \cdots, \mu_r) \in \mathbb{R}^r / \ \mu_1 + \cdots + \mu_r = 0\}.$$

La dimensión de H es $r-1$. Por lo tanto $dim(Ker(\varphi \cap H)) \geq 1$.

Luego, para el subíndice j_0 tal que $\mu_{j_0} \neq 0$, obtenemos que:

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r = \left(\lambda_1 - \frac{\lambda_{j_0}}{\mu_{j_0}}\mu_1\right) x_1 + \cdots \left(\lambda_r - \frac{\lambda_{j_0}}{\mu_{j_0}}\mu_r\right) x_r$$

Si elegimos $\mu_{j_0} > 0$ de manera que para todos los escalares $\mu_k > 0$ se verifique $\frac{\lambda_{j_0}}{\mu_{j_0}} \leq \frac{\lambda_k}{\mu_k}$, entonces $\lambda_i - \frac{\lambda_{j_0}}{\mu_{j_0}}\mu_i \geq 0, \forall i = 1, \cdots, r$.

Pero como $\lambda_{j_0} - \frac{\lambda_{j_0}}{\mu_{j_0}}\mu_{j_0} = 0$, podríamos escribir x como combinación lineal con menos de r elementos, contradiciendo la minimalidad de r, por lo que $r \leq n+1$. □

Más aún, en el capítulo 3 nos interesará descomponer un polítopo en unión de símplices, por lo cual es importante la proposición siguiente:

PROPOSICIÓN 1.35. *Todo polítopo n-dimensional P de $\mathbb{R}^n$ es la unión de una cantidad finita de símplices cuyos vértices son vértices de P y tal que dos símplices de dimensión n no tienen puntos interiores en común.*

DEMOSTRACIÓN. Razonemos por inducción sobre n, la dimensión del polítopo
$P = Conv\{x_1, \cdots, x_r\}$.

Si $n = 1$, entonces el resultado es claro.

Si $n > 1$ sea $x_0 \in \{x_1, \cdots, x_r\}$ un vértice de P y sean $F_1, \cdots, F_t$ los muros de P que no contienen a x_0. Sabemos por hipótesis que cada F_i es un polítopo $(n-1)$-dimensional y se puede descomponer como unión finita de símplices $(n-1)$-dimensionales $F_i^1, \cdots, F_i^{n_i}$ sin puntos interiores en común. Sea $S_i^j = Conv\{x_0, F_i^j\}$ con $i = 1, \cdots, t, \ j = 1, \cdots, n_i$. Entonces:

- Para cada i, j, S_i^j es un n-símplice. Es claro pues para cada i y j, F_i^j es un $(n-1)$-símplice; por lo cual si $\{x_1^{i,j}, \cdots, x_{r_{i,j}}^{i,j}\}$ es el conjunto de vértices de F_i^j tenemos que $S_i^j = Conv\{x_0, F_i^j\} = Conv\{x_0, x_1^{i,j}, \cdots, x_{r_{i,j}}^{i,j}\}$ es un n-símplice pues $x_0 \notin F_i^j$.
- Es claro que $\bigcup_{i,j} S_i^j = P$.

Probamos entonces la descomposición en n-símplices de un polítopo n-dimensional. Falta ver que estos símplices dos a dos no tienen puntos interiores en común.

Sean $S_1 = Conv\{x_0, F_1\}$ y $S_2 = Conv\{x_0, F_2\}$ dos n-símplices de la descomposición anterior de P. Supongamos que F_1 y F_2 son dos $(n-1)$-símplices del mismo muro F de P. Sea y un punto interior a S_1 y sea σ la semirrecta con origen en x_0 y pasa por el punto y. Por la afirmación probada en el teorema 1.32, $\sigma \cap F_1 = \{y_1\} \subset int(F_1)$ ya que si y_1 perteneciera a alguna cara propia de F_1 tendríamos que y_1 no es punto interior de S_1. Al ser $int(F_1) \cap int(F_2) = \emptyset$ deducimos que y_1 no puede ser un punto interior de S_2.
Si F_1 y F_2 son $(n-1)$-símplices pertenecientes a muros distintos existen dos posibilidades: o son disjuntos, en cuyo caso es evidente que no puede haber puntos interiores en común, o se intersectan en una cara de dimensión menor por lo que tampoco pueden haber puntos interiores en común. □

Observación 1.36. La proposición anterior nos da un algoritmo para descomponer un polítopo en símplices donde dichos símplices son en realidad "pirámides" con vértice x_0 y bases que son $(n-1)$-símplices contenidos en los muros de P.

Ejemplos 1.37. 1. En la figura 11 ilustramos una descomposición en símplices del polítopo $P = Conv\{(0,0),(2,2),(3,2),(4,1)\}$.

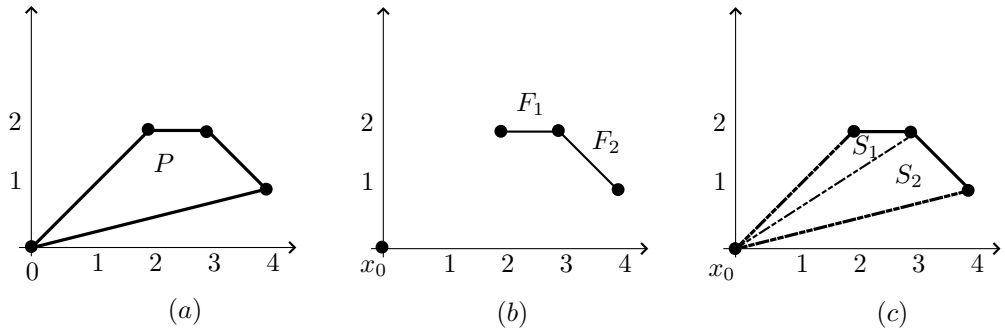


Figura 11.

- en (a) representamos al polítopo P;
- en (b) elegimos como x_0 al punto $(0,0)$ y dibujamos solamente los muros que no contienen a x_0;
- en (c) la descomposición buscada.

2. En la figura 12 ilustramos una descomposición en símplices del cubo unidad tridimensional $C = Conv\{(0,0,0),(1,0,0)),(0,1,0),(0,0,1),(1,1,0),(1,1,1),(1,0,1),(0,1,1)\}$.
   - en (a) representamos al cubo unidad;
   - en (b) elegimos como x_0 al punto $(1,0,1)$ y dibujamos solamente los muros que no contienen al vértice x_0;
   - en (c) la descomposición buscada.

3. Polítopo de Newton de un polinomio de Laurent

Sea k un cuerpo y f un polinomio en $k[x_1, x_2, \cdots, x_n]$.
Podemos escribir f como:

$$f(X) = \sum_{\alpha \in \mathbb{Z}^n_{\geq 0}} c_\alpha X^\alpha$$

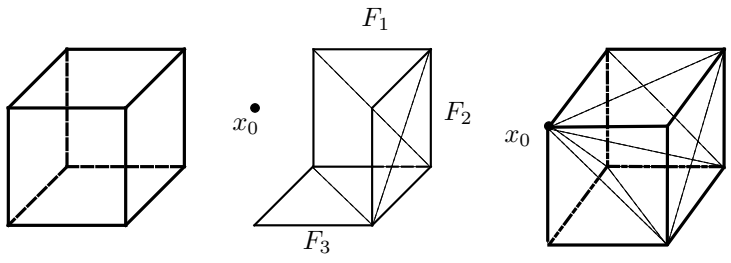


Figura 12.

siendo $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$, X^α es el *monomio* $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ y $c_\alpha \in k$ es nulo salvo para una cantidad finita de vectores α.

Notemos que:

$$X^\alpha X^\beta = X^{\alpha+\beta} \text{ y que } X^\alpha X^{-\alpha} = 1 \; \forall \; \alpha, \beta \in \mathbb{Z}^n.$$

Estaremos interesados en considerar *monomios de Laurent* que son monomios en n variables y sus inversas y también *polinomios de Laurent* que son combinaciones k-lineales finitas de monomios de Laurent.

Notamos al anillo de los polinomios de Laurent, $k[x_1^{\pm 1}, x_2^{\pm 1}, \cdots, x_n^{\pm 1}]$.

Es una k-álgebra conmutativa y además $k[x_1, \cdots, x_n] \subset k[x_1^{\pm 1}, x_2^{\pm 1}, \cdots, x_n^{\pm 1}]$: es decir, los polinomios usuales son polinomios de Laurent.

Definición 1.38. El *polítopo de Newton* del polinomio de Laurent f es el conjunto:

$$NP(f) = Conv\,(\mathcal{A})\,, \text{ donde } \mathcal{A} = \{\alpha \in \mathbb{Z}^n : c_\alpha \neq 0\}\,.$$

O sea el polítopo de Newton de un polinomio de Laurent en n variables es un polítopo racional cuyo conjunto de vértices es un subconjunto de los exponentes del polinomio.

Ejemplos 1.39. Los siguientes son ejemplos de polítopos de Newton ilustrados en la figura 13:

1. Sea $p(x) \in k[x], p(x) = \sum_{j=0}^m a_j x^j$ con $a_m \neq 0$. $NP(p) = [0, m]$.
2. Sea $f(x, y) = a + bx + cxy + dx^2y + ey^2 \in k[x, y]$, con $a, b, c, d, e \neq 0$. $NP(f) = Conv(\mathcal{A})$, siendo $\mathcal{A} = \{(0, 0), (1, 0), (1, 1), (2, 1), (0, 2)\}$, es el polítopo representado en (b).
3. Sea $g_1(x, y) = axy + bx^2 + cy^5 + d$ con $a, b, c, d \neq 0$. $NP(g) = Conv(\mathcal{A})$, siendo $\mathcal{A} = \{(0, 0), (2, 0), (0, 5), (1, 1)\}$, es el polítopo representado en (c).
   Observar que si $g_2(x, y) = bx^2 + cy^5 + d$ con $b, c, d \neq 0$ entonces $NP(g_1) = NP(g_2)$.

Observación 1.40. La asignación $NP : k[x_1^{\pm 1}, \cdots, x_n^{\pm 1}] \longrightarrow \{$ polítopos racionales de $\mathbb{R}^n\}$:

- no es inyectiva: dos polinomios distintos pueden tener el mismo polítopo de Newton. Efectivamente esta asignación no depende de los coeficientes de los polinomios. Pero tampoco es inyectiva al nivel de los monomios como lo muestra el ejemplo anterior.
- es sobreyectiva: dado un polítopo racional en $\mathbb{R}^n$, alcanza con tomar sus vértices, construirse los monomios correspondientes y tomar su suma.

Un corolario inmediato del lema 1.3 es el siguiente:

Corolario 1.41. *Sean $f \in k[x_1^{\pm 1}, x_2^{\pm 1}, \cdots, x_n^{\pm 1}]$, $NP(f)$ su polítopo de Newton y $\alpha \in \mathbb{Z}^n$. Entonces $NP(X^\alpha f) = \alpha + NP(f)$.*

Ejemplo 1.42. Si $f(x, y) = ax^2 + bxy + cy^2 + d$ y $\alpha = (1, 1) \in \mathbb{Z}^2$ entonces $X^\alpha f(X) = xy(ax^2 + bxy + cy^2 + d) = ayx^3 + bx^2y^2 + cxy^3 + dxy$ y $NP(X^\alpha f) = NP(f) + \alpha$. En la figura 14 observamos como se traslada el polítopo de Newton de f.

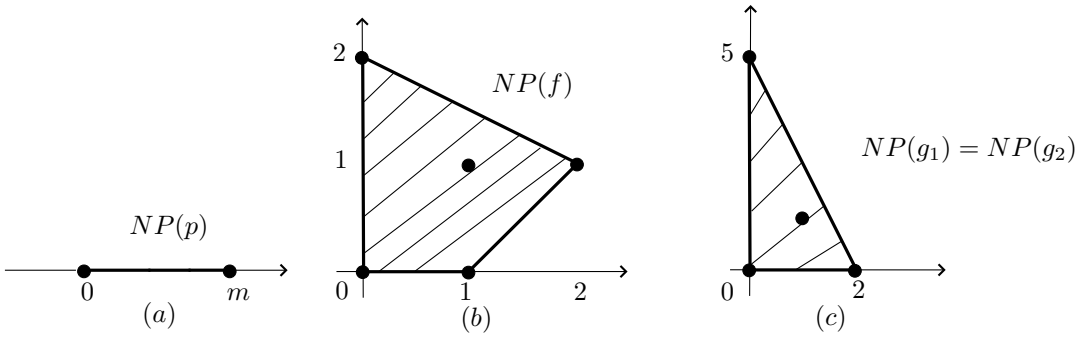


Figura 13.

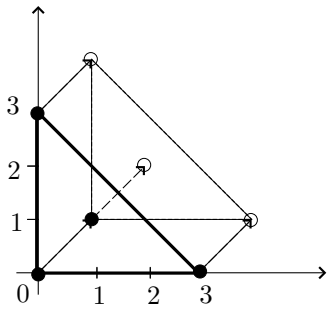


Figura 14.

Los polítopos de Newton reaparecerán en los teoremas de Kushnirenko y Bernstein contenidos en los capítulos 4 y 5.

CAPíTULO 2

Variedades algebraicas

Este capítulo se divide en dos grandes partes. En primer lugar presentaremos algunos resultados básicos de la teoría de las variedades algebraicas afines, de las variedades algebraicas proyectivas y de los grupos algebraicos afines, cuyo objetivo será el de asentar definiciones, notaciones y resultados que usaremos más adelante. Lejos de ser exhaustiva, esta lista cubre los conceptos esenciales necesarios para la comprensión del capítulo 4. No incluiremos las demostraciones de la mayoría de las afirmaciones citadas, señalando al lector interesado la posibilidad de consultar las mismas en los libros [**7**], [**8**] o [**9**]. En una segunda parte presentaremos un caso particular de variedad algebraica proyectiva, definida a partir de un conjunto finito de vectores de coordenadas enteras, la cual nos será de gran utilidad en la demostración del teorema de Kushnirenko.
En el correr de todo el capítulo trabajaremos sobre un cuerpo k algebraicamente cerrado.

1. Variedades algebraicas afines y proyectivas

1. Primeras definiciones.

DEFINICIÓN 2.1. Por $\mathbb{A}^n(k)$ o más simplemente $\mathbb{A}^n$ designaremos el producto cartesiano de k por sí mismo n veces; o sea $\mathbb{A}^n$ es el conjunto de las n-uplas de elementos en k, y se suele llamar *n-espacio afín*.

Recordamos que $k[x_1, x_2, \cdots, x_n]$ denota la k-álgebra de los polinomios en n variables con coeficientes en k.
Decimos que un punto $P = (p_1, p_2, \cdots, p_n)$ es *un cero* de un polinomio $f \in k[x_1, x_2, \cdots, x_n]$ si $f(P) = f(p_1, p_2, \cdots, p_n) = 0$.

DEFINICIÓN 2.2. Sea S un conjunto de polinomios de $k[x_1, \cdots, x_n]$.
Llamamos *el conjunto de ceros* de S al conjunto:

$$\mathcal{V}(S) = \{P \in \mathbb{A}^n / \ f(P) = 0, \ \forall f \in S\} \subset \mathbb{A}^n$$

DEFINICIÓN 2.3. Un subconjunto $X \subset \mathbb{A}^n$ es una *variedad algebraica afín* si $X = \mathcal{V}(S)$ para algún subconjunto S de polinomios en $k[x_1, x_2, \cdots, x_n]$.

OBSERVACIÓN 2.4. Si J es el ideal generado por S entonces $\mathcal{V}(J) = \mathcal{V}(S)$, por lo que toda variedad algebraica se puede expresar como el conjunto de ceros de una cantidad finita de polinomios. En efecto como $k[x_1, \cdots, x_n]$ es noetheriano todos sus ideales son finitamente generados.

Se prueba, ver [**8**], que las variedades algebraicas cumplen los axiomas de los cerrados de una topología que llamamos *topología de Zariski*. Una base de esta topología está dada por los *abiertos principales*:

$$\mathcal{V}(f)^c = \{P \in \mathbb{A}^n / \ f(P) \neq 0\}.$$

A su vez, dado un conjunto X de puntos de $\mathbb{A}^n$ podemos considerar el conjunto de los polinomios que se anulan en los puntos de X.

DEFINICIÓN 2.5. Dado X un subconjunto de $\mathbb{A}^n$, el *ideal de* X es el conjunto:

$$\mathcal{I}(X) = \{f \in k[x_1, \cdots, x_n] / \ f(P) = 0, \ \forall \ P \in X\}$$

Es fácil ver que $\mathcal{I}(X)$ es un ideal.

Así construimos dos mapas $\mathcal{V}$ e $\mathcal{I}$ que nos permiten pasar de conjuntos de polinomios de $k[x_1, \cdots, x_n]$ a subconjuntos de $\mathbb{A}^n$ con $\mathcal{V}$ y de subconjuntos de $\mathbb{A}^n$ a ideales de $k[x_1, \cdots, x_n]$ con $\mathcal{I}$.

$$\begin{array}{ccc} \text{subconjuntos de } k^n & & \text{ideales de } k[x_1, \cdots, x_n] \\ S & \xrightarrow{\mathcal{I}} & \mathcal{I}(S) \\ \mathcal{V}(J) & \xleftarrow{\mathcal{V}} & J \end{array}$$

Las propiedades de estos dos mapas y la correspondencia que definen están detalladas en [**4**]. En particular es fácil ver que $\mathcal{V}$ no es inyectivo.

2. Teorema de los ceros de Hilbert.

TEOREMA 2.6 (de los ceros de Hilbert (Nullstellensatz)). *Si k es un cuerpo algebraicamente cerrado y J un ideal de $k[x_1, \cdots, x_n]$, entonces $\mathcal{I}(\mathcal{V}(J)) = \sqrt{J}$.*

Una versión útil del teorema de Hilbert es también:

TEOREMA 2.7 (Versión débil del Nullstellensatz). *Si I es un ideal propio de $k[x_1, \cdots, x_n]$ entonces $V(I) \neq \emptyset$.*

La principal consecuencia del teorema de los ceros de Hilbert es que establece una biyección entre los ideales radicales y las variedades algebraicas afines (Ver [**4**]).

3. Irreducibilidad de variedades algebraicas afines.

DEFINICIÓN 2.8. Una variedad algebraica afín $X \subset \mathbb{A}^n$ es *reducible* si $X = X_1 \cup X_2$ donde X_1 y X_2 son variedades algebraicas propios de $\mathbb{A}^n$ tales que $X_i \neq V$, $i = 1, 2$.
En caso contrario se dice que la variedad es *irreducible*.

PROPOSICIÓN 2.9.
- *Una variedad algebraica afín X es irreducible si y sólo si $\mathcal{I}(X)$ es un ideal primo.*
- *Si $\varphi : X \longrightarrow X'$ es una función continua entre dos variedades algebraicas afines X y X' siendo X irreducible, entonces $\varphi(X)$ es irreducible.*

PROPOSICIÓN 2.10. *Si I es un ideal primo, entonces $\mathcal{V}(I)$ es irreducible. Luego existe una correspondencia biunívoca entre los ideales primos y las variedades algebraicas irreducibles.*

TEOREMA 2.11. *Si X es una variedad algebraica afín de $\mathbb{A}^n$ entonces existen variedades algebraicas afines irreducibles $X_1, \cdots, X_m$ tales que $X = X_1 \cup \cdots \cup X_m$. Si además imponemos la condición $X_i \nsubseteq X_j$ para todo $i \neq j$, esta descomposición es única a menos del orden de los conjuntos X_i.*

Descomponiendo la variedad X de esa manera decimos que los conjuntos X_i son las componentes irreducibles *de X.*

4. Variedad producto.

Si consideramos $X \subset \mathbb{A}^n$ e $Y \subset \mathbb{A}^n$ dos variedades algebraicas afines, decimos que el conjunto $X \times Y \subset \mathbb{A}^n \times \mathbb{A}^n \cong \mathbb{A}^{n+m}$ define la variedad producto de X por Y. La topología inducida en $X \times Y$ no es la topología producto de $\mathbb{A}^n \times \mathbb{A}^m$ sino la topología de Zariski en $\mathbb{A}^{n+m}$ restringida a $X \times Y$. Por ejemplo hay más cerrados en $\mathbb{A}^2$ que cerrados en $\mathbb{A}^1 \times \mathbb{A}^1$: el conjunto $\{(x, y) \in \mathbb{A}^2 : y = x^2\}$ es una variedad (cerrada) en $\mathbb{A}^2$ con la topología de Zariski, pero no es un cerrado en $\mathbb{A}^1 \times \mathbb{A}^1$ con la topología producto.

PROPOSICIÓN 2.12. *Si $X \subset \mathbb{A}^n$ e $Y \subset \mathbb{A}^m$ son variedades algebraicas afines e irreducibles entonces $X \times Y \subset \mathbb{A}^{n+m}$ es una variedad algebraica afín e irreducible.*

5. Anillo coordenado.

DEFINICIÓN 2.13. Sea X una variedad algebraica afín de $\mathbb{A}^n$. El conjunto

$$k[X] = \frac{k[x_1, \cdots, x_n]}{\mathcal{I}(X)}$$

se llama el *anillo coordenado de X* o *álgebra afín de X*.

OBSERVACIÓN 2.14.
- $k[X]$ es un álgebra conmutativa finitamente generada sobre k pues es la imagen de la proyección canónica $\varphi : k[x_1, \cdots, x_n] \longrightarrow \frac{k[x_1, \cdots, x_n]}{\mathcal{I}(X)}$.
- $k[X]$ no tiene elementos nilpotentes ya que $\mathcal{I}(X)$ es un ideal radical.
- Si f y g son dos polinomios de $k[x_1, \cdots, x_n]$ tales que $f(x) = g(x)\ \forall\ x \in X$, esto es equivalente a decir que el polinomio $f - g \in \mathcal{I}(X)$, o sea f y g coinciden como polinomios de $k[X]$, por lo que podemos interpretar a $k[X]$ como la restricción de los polinomios en n variables a X. En vista de esta observación es que a veces el anillo $k[X]$ se llama *anillo de funciones polinomiales sobre X*.

Es consecuencia del teorema de los ceros de Hilbert la siguiente:

PROPOSICIÓN 2.15. *Si X es una variedad algebraica afín con anillo coordenado $k[X]$, entonces existe una correspondencia biyectiva entre los puntos de X y los ideales maximales de $k[X]$.*

DEFINICIÓN 2.16. Sea $X \subset \mathbb{A}^n$ una variedad algebraica afín.
$Y \subset X$ es una *subvariedad* de X si también es una variedad afín.

TEOREMA 2.17. *Existe una correspondencia biunívoca entre las subvariedades de X y los ideales radicales de $k[X]$.*

La condición de irreducibilidad de una variedad afín se interpreta con respecto a su anillo coordenado:

PROPOSICIÓN 2.18. *Una variedad algebraica afín $X \subset \mathbb{A}^n$ es irreducible si y solamente si $k[X]$ es un dominio de integridad.*

6. Morfismos de variedades algebraicas afines.

DEFINICIÓN 2.19. Sean $X \subset \mathbb{A}^n$ e $Y \subset \mathbb{A}^m$ dos variedades algebraicas afines. Decimos que la función $\varphi : X \longrightarrow Y$ tal que:

$$\varphi(x_1, \cdots, x_n) = (\varphi_1(x_1, \cdots, x_n), \cdots, \varphi_m(x_1, \cdots, x_n))$$

es *un morfismo de variedades algebraicas afines* si $\varphi_i \in k[X]\ \forall\ i = 1, \cdots, m$.

OBSERVACIÓN 2.20. Todo morfismo de variedades es continuo: si $\varphi : X \longrightarrow Y$ es un morfismo de variedades y $C = \mathcal{V}(g_1, \cdots, g_t)$ es un cerrado en Y, entonces

$$\varphi^{-1}(C) = \{x \in X /\ (g_j \circ \varphi)(x) = 0\ \forall\ j = 1, \cdots, t\} = \mathcal{V}(g_1 \circ \varphi, \cdots, g_t \circ \varphi)$$

es un cerrado en X.

7. Variedades algebraicas proyectivas.

DEFINICIÓN 2.21 (El espacio proyectivo $\mathbb{P}^n$). En $\mathbb{A}^{n+1} \setminus \{0\}$ definimos la relación

$(a_0, a_1, \cdots, a_n) \sim (b_0, b_1, \cdots, b_n) \Leftrightarrow \exists \lambda \in k^* : \ a_i = \lambda b_i, \forall\ i = 0, 1, \cdots, n.$

$\sim$ es una relación de equivalencia. Luego definimos:

$$\mathbb{P}^n = \frac{\mathbb{A}^{n+1} \setminus \{0\}}{\sim}$$

el *espacio proyectivo n dimensional*.

Es decir $\mathbb{P}^n$ es el conjunto de los subespacios de $\mathbb{A}^{n+1}$ de dimensión uno sin el origen. La clase de $(a_0, a_1, \cdots, a_n)$ en $\mathbb{P}^n$ la representamos por $[a_0 : a_1 : \cdots : a_n]$; se dice que $(a_0, a_1, \cdots, a_n)$ es un vector de coordenadas homogéneas de $[a_0 : a_1 : \cdots : a_n]$, este vector es único a menos de multiplicación por un escalar no nulo.

Tal como lo hicimos en el caso afín queremos hablar del conjunto de ceros de un polinomio, pero esta vez en el espacio proyectivo $\mathbb{P}^n$. Dado $f \in k[x_0, x_1, \cdots, x_n]$, diremos que $f(P) = 0$ para $P = [a_0 : a_1 : \cdots : a_n]$ si $f(\lambda a_0, \lambda a_1, \cdots, \lambda a_n) = 0, \forall\ \lambda \in k^*$. Descomponiendo f en suma de componentes homogéneas, es claro que P es un cero de f si y sólo si es un cero de cada una de sus componentes homogéneas, ya que k es un cuerpo infinito.

DEFINICIÓN 2.22. Si $f \in k[x_0, x_1, \cdots, x_n]$ es un polinomio homogéneo entonces su *conjunto de ceros* es

$$\mathcal{V}_{\mathbb{P}}(f) = \{P \in \mathbb{P}^n / \ f(P) = 0\}.$$

De la misma manera definimos el conjunto de ceros de un subconjunto $S \subset k[x_0, x_1, \cdots, x_n]$ de polinomios homogéneos como $\mathcal{V}_{\mathbb{P}}(S) = \{P \in \mathbb{P}^n / \ f(P) = 0 \ \forall\ f \in S\}$ y si $J = \langle S \rangle$ entonces $\mathcal{V}_{\mathbb{P}}(S) = \mathcal{V}_{\mathbb{P}}(J)$.

DEFINICIÓN 2.23. Una *variedad algebraica proyectiva* en $\mathbb{P}^n$ es un subconjunto $X \subset \mathbb{P}^n$ tal que existen $f_1, \cdots, f_t$ polinomios homogéneos de $k[x_0, x_1, \cdots, x_n]$ tales que $X = \mathcal{V}_{\mathbb{P}}(f_1, \cdots, f_t)$.

De la misma forma que en el caso afín, estos conjuntos verifican los axiomas de los cerrados de una topología, que también llamaremos *topología de Zariski* de $\mathbb{P}^n$.

DEFINICIÓN 2.24. Una k-álgebra A es *graduada* sobre $\mathbb{Z}$ si $A = \bigoplus_{d \geq 0} A_d$ con $d \in \mathbb{Z}$ donde cada A_d es un grupo abeliano y se cumple que $A_0 = k$ y $A_d A_e \subset A_{d+e}$.

Los elementos de A_d se dicen *homogéneos* de grado d.

Con esta definición podemos escribir todo elemento de A como suma de elementos homogéneos y esta descomposición es única.

EJEMPLO 2.25. La k-álgebra $k[x_0, x_1, \cdots, x_n]$ es una k-álgebra graduada ya que

$$k[x_0, x_1, \cdots, x_n] = \bigoplus_{d \geq 0} (k[x_0, x_1, \cdots, x_n])_d$$

donde $(k[x_0, x_1, \cdots, x_n])_d$ denota el espacio vectorial de los polinomios homogéneos de grado exactamente d junto con el polinomio 0 (el polinomio 0 se considera homogéneo de todos los grados).

DEFINICIÓN 2.26. Un ideal $I \subset A$ es *homogéneo* si $I = \bigoplus_{d \geq 0} (I \cap A_d)$.

PROPOSICIÓN 2.27. *Un ideal I es homogéneo si y sólo si se lo puede generar a partir de elementos homogéneos.*

PROPOSICIÓN 2.28. *Son equivalentes las siguientes afirmaciones:*

- *I es un ideal homogéneo de $k[x_0, x_1, \cdots, x_n]$*
- *I es un ideal y si $f \in I, f = f_0 + f_1 + \cdots + f_d$ con f_i homogéneo de grado i entonces $f_i \in I, \forall\ i = 0, \cdots, d$.*

- $I = \langle g_1, \cdots, g_t \rangle$ *siendo* g_j *polinomio homogéneo* $\forall\, j = 1, \cdots, t$.

OBSERVACIÓN 2.29. Una variedad proyectiva en $\mathbb{P}^n$ es $\mathcal{V}_{\mathbb{P}}(I)$ donde I es un ideal homogéneo.

DEFINICIÓN 2.30. Si $X \subset \mathbb{P}^n$, entonces definimos $\mathcal{I}_{\mathbb{P}}(X)$, el ideal generado por:

$$\{f \in k[x_0, \cdots, x_n] / \; f \text{ es homogéneo y } f(P) = 0\,, \forall\, P \in X\}.$$

Es fácil ver que $\mathcal{I}_{\mathbb{P}}(X)$ es un ideal homogéneo.

Cuando trabajemos con variedades algebraicas afines y con variedades algebraicas proyectivas a la vez las diferenciaremos respectivamente por $\mathcal{V}_{\mathbb{A}}(.)$ y $\mathcal{V}_{\mathbb{P}}(.)$. Lo mismo con los ideales afines y proyectivos $\mathcal{I}_{\mathbb{A}}(.)$ y $\mathcal{I}_{\mathbb{P}}(.)$. Cuando el contexto esté suficientemente claro omitiremos estos subíndices.

También podemos precisar los conceptos de irreducibilidad y de anillo coordenado:

DEFINICIÓN 2.31. Una variedad proyectiva X es *irreducible* si y sólo si $\mathcal{I}(X)$ es un ideal primo.

DEFINICIÓN 2.32. Siendo $X \subset \mathbb{P}^n$ una variedad proyectiva, definimos el *anillo coordenado* sobre X al conjunto $S[X] = \frac{k[x_0, \cdots, x_n]}{\mathcal{I}(X)}$.

De la misma manera, al ser $k[x_0, \cdots, x_n]$ un anillo graduado y finitamente generado, $S[X]$ es un anillo graduado y finitamente generado. Tampoco contiene elementos nilpotentes.

DEFINICIÓN 2.33. Si $X \subset \mathbb{P}^n$ e $Y \subset \mathbb{P}^m$ son dos variedades proyectivas, una función $\varphi : X \longrightarrow Y$ es un *morfismo de variedades proyectivas* si para toda función $g \in S[Y]$ tenemos que $g \circ \varphi \in S[X]$.

OBSERVACIÓN 2.34.
- Todo morfismo de variedades proyectivas es continuo.
- Si X es una variedad proyectiva irreducible y si $\varphi : X \longrightarrow X'$ es un morfismo de variedades proyectivas entonces $\varphi(X)$ es irreducible también.

8. Dimensión.

DEFINICIÓN 2.35. Si X es un espacio topológico, definimos la *dimensión de* X como el supremo de los enteros n tal que existe una cadena de cerrados irreducibles $C_0 \subsetneq C_1 \subsetneq \cdots \subsetneq C_n = X$.

Sea X una variedad algebraica afín irreducible. La *dimensión de la variedad* X es su dimensión como espacio topológico y la denotamos $dim(X)$.

Si X es una variedad afín o proyectiva pero no necesariamente irreducible, sabemos que la podemos descomponer como $X = X_1 \cup \cdots \cup X_r$ donde cada X_j es irreducible. Definimos la dimensión de X como

$$dim(X) = max\{dim(X_j)\}_{j=1,\cdots,r}.$$

Equivalentemente podemos tomar supremos en las cadenas de ideales primos que contienen al ideal de X, esto es si X es una variedad algebraica afín irreducible, entonces sabemos que su ideal $\mathcal{I}_{\mathbb{A}}(X)$ es un ideal primo. Luego la dimensión de X es el supremo de las cadenas de ideales primos I_j tales que:

$$\mathcal{I}_{\mathbb{A}}(X) \subsetneq I_1 \subsetneq \cdots \subsetneq I_r \subsetneq k[x_1, \cdots, x_n].$$

Existen otras maneras de definir la dimensión: ver [**7**] o [**9**].

DEFINICIÓN 2.36. Si X es irreducible y $U \subset X$ es un abierto no vacío, definimos la dimensión de U, $dim(U) = dim(X)$.

LEMA 2.37. *Si* X *es una variedad algebraica afín irreducible y* U *es un conjunto abierto en* X *entonces* $dim(X \setminus U) < dim(X)$.

PROPOSICIÓN 2.38.
- *Si X es irreducible e Y es un cerrado propio irreducible, entonces* $dim(Y) < dim(X)$.
- *Si X e Y son variedades afines entonces la dimensión de la variedad afín $X \times Y$ es* $dim(X) + dim(Y)$.

OBSERVACIÓN 2.39. Análogamente al caso afín tenemos la noción de dimensión para variedades algebraicas proyectivas.

DEFINICIÓN 2.40. Si X es una variedad afín o proyectivo (en $\mathbb{A}^n$ o en $\mathbb{P}^n$) y $dim(X) = r$ decimos que su *codimensión* es el número natural $n - r$.

9. Teorema de Chevalley.

DEFINICIÓN 2.41. Decimos que un conjunto $U \subset \mathbb{A}^n$ es *localmente cerrado* si U es abierto en su clausura $\overline{U}$. En este caso $dim(U) = dim(\overline{U})$.

Observar que la definición anterior es equivalente a que U sea intersección de un abierto y de un cerrado de $\mathbb{A}^n$

DEFINICIÓN 2.42. Si X es un espacio topológico, un conjunto se dice *constructible* si es una unión finita de conjuntos localmente cerrados.

TEOREMA 2.43 (Teorema de Chevalley). *Sea $\varphi : X \longrightarrow Y$ un morfismo de variedades afines o proyectivas, entonces $\varphi(X)$ contiene un abierto denso de $\overline{\varphi(X)}$.*

10. Cono sobre una variedad algebraica proyectiva.

DEFINICIÓN 2.44. Sea la proyección canónica asociada a $\mathbb{P}^n$:

$$\begin{array}{rccc} \Pi: & \mathbb{A}^{n+1} \setminus \{0\} & \longrightarrow & \mathbb{P}^n \\ & (a_0, a_1, \cdots, a_n) & \mapsto & [a_0 : a_1, \cdots : a_n] \end{array}$$

Si $X \subset \mathbb{P}^n$ es una variedad algebraica proyectiva definimos *el cono sobre X* como el conjunto

$$C(X) := \Pi^{-1}(X) \cup \{0\} \subset \mathbb{A}^{n+1}.$$

PROPOSICIÓN 2.45. *Si $X \subset \mathbb{P}^n$ es una variedad algebraica proyectiva, entonces:*
a) *$C(X)$ es una variedad algebraica afín en $\mathbb{A}^{n+1}$;*
b) *Si X es irreducible entonces $C(X)$ es irreducible;*
c) *$dim C(X) = dim X + 1$;*
d) *$\mathfrak{I}_{\mathbb{A}}(C(X)) = \mathfrak{I}_{\mathbb{P}}(X)$.*

11. Grado de una variedad proyectiva y polinomio de Hilbert.

DEFINICIÓN 2.46. Sea X una variedad proyectiva de dimensión r en $\mathbb{P}^n$. Definimos el *grado de la variedad X* como el cardinal del conjunto $\{y \in \mathbb{P}^n : \ y \in X \cap \mathcal{L}\}$ siendo $\mathcal{L}$ un subespacio genérico de codimensión r en $\mathbb{P}^n$.

OBSERVACIÓN 2.47. No definiremos aquí formalmente la noción de genericidad: lo haremos en el capítulo 4. Provisoriamente podemos pensar que la definición anterior exige al subespacio $\mathcal{L}$ que esté en "buena posición" respecto de X. Por ejemplo, en el plano lo "más común" es que una recta cualquiera corte a una cónica fija en dos puntos, sólo en casos especiales la recta será tangente. Este tipo de situaciones "más frecuentes" es lo que formaliza el concepto de genericidad.

Hay otras maneras de definir el grado de un variedad proyectiva. Recomendamos al lector interesado en el tema consultar [**9**].

DEFINICIÓN 2.48. Sea $A = \bigoplus_{d \geq 0} A_d$ una k-álgebra con graduación $\{A_d\}$. Un A-módulo M se dice graduado si admite una descomposición $M = \bigoplus_{d \geq 0} M_d$ tal que $A_d M_e \subset M_{d+e}$.

DEFINICIÓN 2.49. Con las notaciones anteriores, si M es finitamente generado como A-módulo, llamamos *función de Hilbert* de un A-módulo graduado M a la función:

$$\begin{array}{rccc} F_M: & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & d & \mapsto & dim_k M_d \end{array}$$

TEOREMA 2.50. *Si M es un A-módulo graduado finitamente generado, existe un único polinomio $P_M(t) \in \mathbb{Q}[t]$ tal que $P_M(d) = F_M(d)$ para valores de d suficientemente grandes. $P_M(t)$ se llama* polinomio de Hilbert *del módulo M.*

DEMOSTRACIÓN. ver [**9**] páginas 110-111. □

DEFINICIÓN 2.51. Si X es una variedad algebraica proyectiva en $\mathbb{P}^n(\mathbb{C})$, sabemos que *su anillo coordenado* $S[X] = \frac{\mathbb{C}[x_0,\cdots,x_n]}{I_{\mathbb{P}}(X)}$ es un $\mathbb{C}[x_0,\cdots,x_n]$-módulo graduado finitamente generado.

El *polinomio de Hilbert de la variedad proyectiva* X es el polinomio de Hilbert de $S[X]$ y se denotará P_X.

TEOREMA 2.52. *Si X es una variedad algebraica proyectiva en $\mathbb{P}^n(\mathbb{C})$ de dimensión r y de grado μ entonces :*

$$P_X(t) = \mu\frac{t^r}{r!} + \textit{ términos de orden menor.}$$

DEMOSTRACIÓN. Ver [**9**] p. 113. □

2. Grupos algebraicos afines

En esta parte definiremos e introduciremos, también muy rápidamente, algunos resultados elementales de la teoría de los grupos algebraicos. Recomendamos al lector interesado en completar la teoría y los resultados presentados consultar [**8**].
Supondremos, además de que k sea algebraicamente cerrado, que tiene característica cero.

1. Definiciones y ejemplos.

DEFINICIÓN 2.53. *Un grupo algebraico afín* G es un grupo abstracto (G, μ, e) donde G es una variedad algebraica afín tal que los mapas multiplicación

$$\begin{array}{rccc} \mu: & G\times G & \longrightarrow & G \\ & (x,y) & \mapsto & xy \end{array}$$

e inversión

$$\begin{array}{rccc} \iota: & G & \longrightarrow & G \\ & x & \mapsto & x^{-1} \end{array}$$

son morfismos de variedades algebraicas afines.

EJEMPLOS 2.54.
- El grupo aditivo $G_a = (\mathbb{A}^1, +)$. Las operaciones $(x,y) \mapsto x+y$ e $x \mapsto -x$ son morfismos de variedades (polinomiales).
- El grupo multiplicativo $G_m = (\mathbb{A}^1 \setminus \{0\}, \times)$.
  Sabemos que $\mathbb{A}^1 \setminus \{0\}$ es un abierto principal pero no es una variedad algebraica afín de $\mathbb{A}^1$. Sin embargo podemos mirar $\mathbb{A}^1 \setminus \{0\}$ en un espacio de dimensión mayor y considerarlo como cerrado de este espacio, mediante la siguiente identificación:

  $$k^* = \mathbb{A}^1 \setminus \{0\} \longleftrightarrow \{(x_1,x_2) \in \mathbb{A}^2 / \; x_1x_2 = 1\} \subset \mathbb{A}^2.$$

  Las operaciones son:

  $$((x_1,x_2)(y_1,y_2)) \mapsto (x_1y_1, x_2y_2) \text{ y } (x_1,x_2) \mapsto (x_1^{-1}, x_2^{-1}) = (x_2,x_1)$$

  y son claramente morfismos de variedades algebraicas afines, ya que son aplicaciones polinomiales.

- Otro ejemplo muy importante, que generaliza al anterior, es el de $GL_n(\mathbb{C})$, el grupo de las matrices invertibles, llamado *grupo lineal*. De la misma manera, si identificamos $M_n(k)$ con $\mathbb{A}^{n^2}$, GL_n es un abierto principal de $\mathbb{A}^{n^2}$ definido por:

$$\{M \in \mathbb{A}^{n^2} / \ det(M) \neq 0\}.$$

Pero también lo podemos identificar con el conjunto:

$$\{(a_{ij}, d) \in \mathbb{A}^{n^2+1} / \ det(a_{ij})d = 1\} = \mathcal{V}(\langle det(x_{ij})y - 1\rangle)$$

que es un cerrado de $\mathbb{A}^{n^2+1}$ (det es una función polinomial).
Las operaciones son:

$$((a_{ij}, d), (b_{ij}, d')) \mapsto ((c_{ij}), dd')$$

donde (c_{ij}) es el producto matricial de las matrices (a_{ij}) y (b_{ij}) y

$$((a_{ij}), d) \mapsto ((a_{ij})^{-1}, det(a_{ij})).$$

Estos dos mapas son morfismos de variedades pues son polinomiales en las entradas de las matrices y d. Entonces $GL_n(k)$ es también un grupo algebraico.
Observar que $GL_1(k)$ coincide con G_m.

- El grupo lineal SL_n de las matrices con determinante 1, es un subgrupo cerrado de GL_n pues $SL_n = \mathcal{V}(\langle det((x_{ij}) - 1)\rangle)$ y también es un grupo algebraico (un subgrupo cerrado de un grupo algebraico es también un grupo algebraico).
- Los grupos clásicos: ver [**8**] página 52.

Proposición 2.55. *Si $G_1, \cdots, G_n$ son grupos algebraicos afines entonces $G_1 \times G_2 \times \cdots \times G_n$ es un grupo algebraico afín.*

Ejemplo 2.56. En el ejemplo anterior, tomando k como $\mathbb{C}$ vimos que $\mathbb{C}^*$ es un grupo algebraico afín con la multiplicación. Luego $(\mathbb{C}^*)^n = \mathbb{C}^* \times \mathbb{C}^* \times \cdots \times \mathbb{C}^*$ es un grupo algebraico afín.

Definición 2.57. Si G es un grupo algebraico afín, definimos *el álgebra afín $k[G]$* de G como el álgebra afín de la variedad G.

Definición 2.58. Un *morfismo de grupos algebraicos afines* es un morfismo de grupos abstractos que es además un morfismo de variedades algebraicas afines.

Para concluir la sección, mencionamos un teorema muy importante y muy útil a la hora de trabajar con grupos algebraicos afines y la razón por la cual se le da tanta importancia al grupo lineal GL_n.

Teorema 2.59. *Todo grupo algebraico afín es isomorfo a un subgrupo cerrado de GL_n para algún entero n.*

Demostración. Ver [**8**] página 63. □

2. Acción de un grupo algebraico afín G sobre una variedad algebraica X.

Definición 2.60. Sea G un grupo algebraico afín y X una variedad algebraica afín o proyectiva. Se dice que G actúa en X si existe una función polinomial $\varphi : G \times X \longrightarrow X$ morfismo de variedades algebraicas tal que:

$$\varphi(g_1 g_2, x) = \varphi(g_1, \varphi(g_2, x)) \ \forall \ x \in X, \ \forall \ g_1, g_2 \in G$$

y

$$\varphi(e_G, x) = x, \ \forall \ x \in X.$$

donde e_G es el elemento neutro del grupo G.

Decimos también que X es una G-variedad.
Por comodidad en las notaciones notaremos $\varphi(g, x)$ como $g \cdot x$.

EJEMPLOS 2.61. *a)* GL_n actúa sobre las matrices M_n por conjugación vía:

$$g \cdot A = ((g_{ij}))((a_{ij}))((g_{ij}))^{-1}.$$

siendo $A = ((a_{ij})) \in \mathcal{M}_n$ y $g = ((g_{ij})) \in GL_n$

b) GL_n actúa sobre k^n por multiplicación vía:

$$g \cdot X = ((g_{ij}))X$$

siendo $X \in k^n$ y $g \in GL_n$.

Finalmente:

- Decimos que G actúa *transitivamente en* X si $\forall\, x \in X, G \cdot x = X$.
- $O_x = G \cdot x$ es la *órbita* del punto $x \in X$ y es G-estable, esto es que $g \cdot x \in O_x,\ \forall\, x \in O_x$. Es más la órbita de un punto x se puede ver como la imagen del mapa

$$\begin{array}{rccc} \varphi_x : & G & \longrightarrow & X \\ & g & \mapsto & \varphi_x(g) = g \cdot x \end{array}$$

  Las órbitas forman una partición de X y no son necesariamente cerradas.
- El *estabilizador* de $x \in X$, es el conjunto $G_x = \varphi_x^{-1}(\{x\}) = \{g \in G/\ g \cdot x = x\}$ y es un subgrupo cerrado de G.
- Si G actúa en X y en Y, también actúa en $X \times Y$: $g \cdot (x, y) = (g \cdot x, g \cdot y)$.

LEMA 2.62 (Lema de la órbita cerrada). *Sea G un grupo algebraico afín actuando sobre una variedad X afín o proyectiva. Entonces, para cada $x \in X$, O_x es una variedad localmente cerrada en X y su borde $\overline{O_x} \setminus O_x$ es unión de órbitas de dimensión menor. En particular las órbitas de dimensión minimal son cerradas.*

DEMOSTRACIÓN. La acción de un grupo algebraico G sobre una variedad X es un homeomorfismo.

- Se prueba que los grupos algebraicos son constructibles. Luego, si $x \in X$ entonces la órbita de x, $O_x = G \cdot x = \{g \cdot x : g \in G\}$ es constructible por ser la imagen de G por el mapa órbita (teorema de Chevalley). Entonces O_x contiene un abierto U denso en $\overline{O_x}$. Más aún, todo punto de O_x tiene un entorno abierto en $\overline{O_x}$, ya que si $x \in O_x$ y $u \in U$ entonces existe $g \in G$ tal que $g \cdot u = x$ pues la acción es transitiva en O_x. Eso significa que O_x es abierto en $\overline{O_x}$, luego por definición O_x es localmente cerrada en X.
- $\overline{O_x}$ es G-estable pues, por continuidad de la acción $g \cdot \overline{O_x} \subset \overline{g \cdot O_x} \subset \overline{O_x}\ \forall\, g \in G$.
- Probemos que $\overline{O_x} \setminus O_x$ es G-estable. Tenemos que demostrar que si $g \in G$ y $x \in \overline{O_x} \setminus O_x$ entonces $g \cdot x \in \overline{O_x} \setminus O_x$.
  Si $x \in \overline{O_x} \setminus O_x$, supongamos que $g \cdot x \in O_x$. Entonces $x \in g^{-1} \cdot O_x \subset O_x$ pues O_x es G-estable, lo cual contradice la hipótesis inicial.
  Por lo tanto estamos en condiciones de aplicar el lema 2.37:

$$dim(\overline{O_x} \setminus O_x) < dim(\overline{O_x}) = dim(O_x)$$

y, al ser $\overline{O_x} \setminus O_x$ G-estable, es unión de orbitas de dimensión menor.

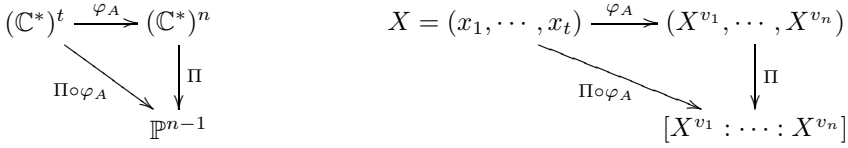
Si Z es una órbita de dimensión minimal entonces $\overline{Z} \setminus Z$ tiene que ser unión de órbitas de dimensión menor que Z. Por lo tanto, Z coincide con $\overline{Z}$ y es cerrada. □

3. La variedad proyectiva X_A

En esta sección le asociaremos a un subconjunto finito de vectores $A \subset \mathbb{Z}^t$ una variedad algebraica proyectiva, X_A que será de particular importancia en el capítulo 4.
Suponemos que $k = \mathbb{C}$.

Si $A = \{v_1, v_2, \cdots, v_n\} \subset \mathbb{Z}^t$, definimos el mapa:

$$\begin{array}{rccc} \varphi_A : & (\mathbb{C}^*)^t & \longrightarrow & (\mathbb{C}^*)^n \\ & X & \mapsto & (X^{v_1}, X^{v_2}, \cdots, X^{v_n}) \end{array}.$$



donde Π es la proyección canónica asociada al espacio proyectivo.
Definimos al conjunto:

$$X_A^0 := (\Pi \circ \varphi_A)((\mathbb{C}^*)^t) = \{[X^{v_1} : \cdots : X^{v_n}] : \ X = (x_1, x_2, \cdots, x_t) \in (\mathbb{C}^*)^t\} \subset \mathbb{P}^{n-1}$$

y su clausura de Zariski en $\mathbb{P}^{n-1}$, $X_A = \overline{X_A^0}$.

Proposición 2.63. *Si el $\mathbb{Z}$-módulo generado por A es $\mathbb{Z}^t$ entonces el conjunto X_A es una variedad algebraica proyectiva de dimensión t.*

Demostración. Definimos una acción del grupo algebraico $(\mathbb{C}^*)^t$ sobre $\mathbb{P}^{n-1}$ de la siguiente manera:

$$\begin{array}{rccc} \cdot : & (\mathbb{C}^*)^t \times \mathbb{P}^{n-1} & \longrightarrow & \mathbb{P}^{n-1} \\ & (X, [p_1 : p_2 : \cdots : p_n]) & \longmapsto & [X^{v_1} p_1 : X^{v_2} p_2 : \cdots : X^{v_n} p_n] \end{array} \tag{3.1}$$

Observamos que X_A^0 es igual al conjunto $O_{[1:1:\cdots:1]} = (\mathbb{C}^*)^t \cdot [1 : \cdots : 1]$, la órbita del punto $[1 : \cdots : 1]$.
Luego, por el lema de la órbita cerrada, X_A^0 es abierta en su clausura $X_A = \overline{O_{[1:1:\cdots:1]}}$.

X_A^0 es irreducible por ser la imagen por una acción del grupo algebraico irreducible $(\mathbb{C}^*)^t$. Por lo tanto X_A también es irreducible.

Como la acción de un grupo algebraico es continua tenemos que $(\mathbb{C}^*)^t \cdot X_A \subseteq X_A$ ya que:

$$(\mathbb{C}^*)^t \cdot X_A = (\mathbb{C}^*)^t \cdot \overline{X_A^0} \subset \overline{(\mathbb{C}^*)^t \cdot X_A^0} = \overline{X_A^0} = X_A.$$

Por otro lado existe una correspondencia uno a uno entre $(\mathbb{C}^*)^t$ y X_A^0.
Definimos el mapa:

$$\begin{array}{rccc} \varphi_{[1:\cdots:1]} : & (\mathbb{C}^*)^t & \longrightarrow & \mathbb{P}^{n-1} \\ & Y & \mapsto & Y \cdot [1 : \cdots : 1] = [Y^{v_1} : \cdots : Y^{v_n}] \end{array}$$

Tenemos que probar que si $Y = (y_1, \cdots, y_t)$ y $Z = (z_1, \cdots, z_t)$ son dos puntos de $(\mathbb{C}^*)^t$ tal que $\varphi(Y) = \varphi(Z)$ entonces $Y = Z$.
Puesto que $\{v_1, \cdots, v_n\}$ es una $\mathbb{Z}$-base de $\mathbb{Z}^t$, podemos escribir cada vector canónico e_i de $\mathbb{Z}^t$ como $e_i = \sum_{i=1}^n a_{ij} v_i$ donde $a_{ij} \in \mathbb{Z}$. Luego $Y^{v_j} = Z^{v_j}$ para todo $j = 1, \cdots, n$ implica que $y^i = Y^{e_i} = Z^{e_i} = z_i$ para todo $i = 1, \cdots, t$, es decir $Y = Z$. Hemos probado que el mapa $\varphi_{[1:\cdots:1]}$ es inyectivo.

Es claro, por el lema de la órbita cerrada, que en estas condiciones, X_A tiene dimensión t. □

Observación 2.64. En realidad la variedad proyectiva X_A es un tipo particular de variedad algebraica: es irreducible, contiene un abierto isomorfo a un toro algebraico y está equipada de una acción que extiende la acción del toro sobre sí mismo a toda la variedad. Estas variedades especiales se llaman variedades tóricas. Recomendamos al lector interesado en el tema consultar el capítulo 5 de [**6**].

Ejemplo 2.65.

- Sea $A = \{0, 1, 2, 3\} \subset \mathbb{Z}$ y $\varphi_A : \mathbb{C}^* \longrightarrow \mathbb{C}^4$ definida como $\varphi_A(t) = (1, t, t^2, t^3)$. Luego $X_A = \overline{\{(1 : t : t^2 : t^3), t \in \mathbb{C}^*\}} \subset \mathbb{P}^3$. X_A lleva el nombre de "twisted cubic".
- Sea $A = \{0, e_1, e_2, \cdots, e_n\} \subset \mathbb{Z}^n$ donde $\{e_j\}_{j=1}^n$ es la base canónica de $\mathbb{Z}^n$. Busquemos la variedad X_A asociada.
  $\varphi_A : (\mathbb{C}^*)^n \longrightarrow \mathbb{C}^{n+1}$ tal que $\varphi_A(x_1, x_2, \cdots, x_n) = (1, x_1, \cdots, x_n)$.
  Luego $X_A = \overline{\{(1 : x_1 : \cdots : x_n)\}} = \mathbb{P}^n$.

Estamos interesados ahora en considerar el cono sobre la variedad X_A.

Sea $X_A \subset \mathbb{P}^{n-1}$ una variedad algebraica proyectiva definida a partir de un subconjunto $A = \{v_1, \cdots, v_n\} \subset \mathbb{Z}^t$.

Recordamos que si Π es la proyección canónica de $\mathbb{C}^n \setminus \{0\}$ en $\mathbb{P}^{n-1}$, el *cono* sobre X_A es la variedad algebraica afín:

$$Y_A := C(X_A) = \Pi^{-1}(X_A) \cup \{0\} \subset \mathbb{C}^n.$$

Dado un subconjunto $A = \{v_1, \cdots, v_n\} \subset \mathbb{Z}^t$ definimos el semigrupo $S_A = \langle (a,1) : a \in A \rangle \subset \mathbb{Z}^{t+1}$. Por definición S_A es finitamente generado.

Por otro lado le podemos asociar una $\mathbb{C}-$álgebra $\mathbb{C}[S_A]$, con las siguientes características:

- $\mathbb{C}[S_A]$ es una $\mathbb{C}$-álgebra conmutativa generada como $\mathbb{C}$-espacio vectorial por χ^u si $u \in \mathbb{C}[S_A]$ y tiene a S_A como base.
- Los elementos de $\mathbb{C}[S_A]$ son combinaciones lineales formales $\sum_{u \in S_A} a_u \chi^u$ donde hay solamente una cantidad finita de coeficientes no nulos.
- El producto en $\mathbb{C}[S_A]$ se define como: $\chi^{u_1} \cdot \chi^{u_2} = \chi^{u_1+u_2}$ y $\chi^0 = 1$ es el elemento unidad. Si u es una unidad en S_A entonces χ^u es una unidad en $\mathbb{C}[S_A]$.
- Si $\{u_i : i \in I\}$ es un conjunto generador de S_A entonces $\{\chi^{u_i} : i \in I\}$ es un generador de $\mathbb{C}[S_A]$. Como consecuencia, como S_A es finitamente generado, $\mathbb{C}[S_A]$ es una $\mathbb{C}$-álgebra finitamente generada.

LEMA 2.66. *Sea $f \in \mathbb{C}[x_1^{\pm 1}, \cdots, x_n^{\pm 1}]$ no nulo. Entonces $f|_{(\mathbb{C}^*)^n} \neq 0$.*

DEMOSTRACIÓN. Multiplicando al polinomio f por un monomio adecuado, la demostración se reduce al hecho de que si $g \in \mathbb{C}[x_1, \cdots, x_n]$ y $g|_{(\mathbb{C}^*)^n} = 0$ resulta por continuidad que $g = 0$. □

TEOREMA 2.67. *Con las notaciones definidas anteriormente Y_A es una variedad algebraica afín de dimensión $t+1$ y $\mathbb{C}[Y_A] = \mathbb{C}[S_A]$.*

DEMOSTRACIÓN. Al ser X_A una variedad algebraica proyectiva de dimensión t en $\mathbb{P}^{n-1}$, $Y_A = C(X_A)$ es una variedad algebraica afín de dimensión $t+1$ en $\mathbb{C}^n$.

Primeramente vamos a probar que

$$I = \left\langle y_1^{a_1} y_2^{a_2} \cdots y_n^{a_n} - y_1^{b_1} y_2^{b_2} \cdots y_n^{b_n} / \; a_1 v_1 + \cdots + a_n v_n = b_1 v_1 + \cdots b_n v_n, \; a_i, b_i \in \mathbb{Z}_{\geq 0} \right\rangle$$

es el ideal de Y_A.

- Es evidente que $I \subset \mathcal{I}(Y_A)$.
- Por otro lado si $f \in \mathcal{I}(Y_A)$ se tiene que $f(X^{v_1}, X^{v_2}, \cdots, X^{v_n}) = 0, \; \forall \; X \in (\mathbb{C}^*)^t$ siendo $A = \{v_1, \cdots, v_n\} \subset \mathbb{Z}^t$. Reagrupando monomios conseguimos escribir

$$f(X^{v_1}, X^{v_2}, \cdots, X^{v_n}) = \sum_{i=1}^{s} c_i X^{\nu_i} = 0, \; \forall \; X \in (\mathbb{C}^*)^t$$

donde $c_i \in \mathbb{C}$ y los ν_i son vectores diferentes de $\mathbb{Z}^t$. Usando el lema anterior concluimos que $c_i = 0 \; \forall \; i = 1, \cdots, s$. Es decir es suficiente probar que $f \in I$ en el caso en que $f|_{Y_A} = 0$ y $f(X) = \sum_{i=1}^{p} c_i X^{\alpha_i}$ con $\sum a_{rj} v_j = \sum a_{sj} v_j$ si $\alpha_i = (a_{i1}, \cdots, a_{it}) \in (\mathbb{Z}_{\geq 0})^t$ y $\sum_{i=1}^{p} c_i = 0$. Ahora podemos escribir:

$$f(X) = c_1(X^{\alpha_1} - X^{\alpha_2}) + (c_1 + c_2)(X^{\alpha_2} - X^{\alpha_3}) + (c_1 + \cdots + c_{p-1})(X^{\alpha_{p-1}} - X^{\alpha_p}) \in I.$$

Finalmente probaremos que $\mathbb{C}[Y_A] = \frac{\mathbb{C}[y_1, \cdots, y_n]}{I}$ y $\mathbb{C}[S_A]$ son $\mathbb{C}$-álgebras isomorfas. Basta considerar el morfismo de $\mathbb{C}$-álgebras:

$$\begin{array}{rccc} \varphi: & \mathbb{C}[y_1, \cdots, y_n] & \longrightarrow & \mathbb{C}[S_A] = \mathbb{C}[\chi^{v_1}, \cdots, \chi^{v_n}] \\ & y_i & \mapsto & \chi^{v_i} \end{array}$$

Es evidente que $I \subset Ker(\varphi)$.

La inclusión $Ker(\varphi) \subset I$ se justifica con un argumento análogo al utilizado para probar que $\mathcal{I}(Y_A) \subset I$. □

CAPíTULO 3

Volúmenes y polinomio de Ehrhart

En 1960, Eugène Ehrhart probó que dado P un polítopo racional, la cantidad de puntos con coordenadas enteras del homotetizado de P, $dP = \{dp : p \in P\}$, $d \in \mathbb{Z}_{\geq 0}$, es un polinomio en d cuyo grado es la dimensión de P. A ese polinomio E_P se le llama *polinomio de Ehrhart* de P y también se usará en la demostración del teorema de Kushnirenko.

En una primera parte definiremos el concepto de retículo y veremos algunas propiedades relativas a los volúmenes. Luego procederemos al estudio del polinomio de Ehrhart de un polítopo racional P.

1. Retículos y volúmenes

DEFINICIÓN 3.1. El *rango* de un subconjunto de $\mathbb{R}^n$ es la dimensión del espacio afín de $\mathbb{R}^n$ generado por este subconjunto.

$A \subseteq \mathbb{R}^n$ es un *conjunto discreto* si la topología inducida por la topología usual de $\mathbb{R}^n$ sobre A es la topología discreta.

DEFINICIÓN 3.2. $L \subset \mathbb{R}^n$ es un retículo de rango n si es un subgrupo discreto de $(\mathbb{R}^n, +)$ de rango n.

TEOREMA 3.3. *Sea L un retículo de rango n de $\mathbb{R}^n$. Entonces existen $\{v_1, v_2, \cdots, v_n\} \subset \mathbb{R}^n$, linealmente independientes en $\mathbb{R}^n$ tales que $L = \sum_{i=1}^{n} \mathbb{Z}v_i$.*

Decimos que el conjunto $\{v_1, v_2, \cdots, v_n\}$ es una $\mathbb{Z}$-base de L.

DEFINICIÓN 3.4. Sea L un retículo de rango n en $\mathbb{R}^n$ y $B = \{v_1, v_2, \cdots, v_n\}$ una $\mathbb{Z}-$base de L. Definimos el *paralelotopo fundamental asociado a la base B*:

$$P_B = \left\{ \sum_{i=1}^{n} a_i v_i / \;\; a_i \in [0,1) \right\}$$

En virtud de la definición 3.4 diremos que si $S = Conv\{0, v_1, \cdots, v_n\}$ es un n-símplice de $\mathbb{R}^n$ el *paralelotopo definido por S* es:

$$P_S = \left\{ \sum_{i=1}^{n} a_i v_i / \;\; a_i \in [0,1) \right\}$$

OBSERVACIÓN 3.5. De la misma manera si $G = \{w_1, \cdots, w_m\}$ es un generador del retículo L se puede definir el *paralelotopo fundamental generado por G* como el conjunto:

$$P_B = \left\{ \sum_{i=1}^{m} a_i w_i | \;\; a_i \in [0,1) \right\}.$$

LEMA 3.6. *Sea $Q = Conv\{0, v_1, \cdots, v_n\}$ un n-símplice de $\mathbb{R}^n$. Entonces:*

$$Vol_{\mathbb{R}^n}(Q) = \frac{Vol_{\mathbb{R}^n}(P_Q)}{n!},$$

donde P_Q es el paralelotopo definido por el n-símplice Q.

DEMOSTRACIÓN. Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ la transformación lineal biyectiva definida por:

$$T(v_i) = e_i,$$

donde e_i es el i-ésimo vector de la base canónica de $\mathbb{R}^n$. Entonces $T(Q) = Q_1$ donde Q_1 es el n-símplice elemental de $\mathbb{R}^n$ y por la fórmula de cambio de variable:

$$Vol_{\mathbb{R}^n}(Q) = \int_Q 1 = |det(T^{-1})| \int_{Q_1} 1.$$

Pero $|det(T^{-1})| = Vol_{\mathbb{R}^n}(P_Q)$ (Ver Apéndice), por lo tanto:

$$Vol_{\mathbb{R}^n}(Q) = Vol_{\mathbb{R}^n}(P_Q) \int_{Q_1} 1 = \frac{Vol_{\mathbb{R}^n}(P_Q)}{n!}$$

ya que, como lo probamos en el capítulo 1, $\int_{Q_1} = Vol_n(Q_1) = \frac{1}{n!}$. □

Proposición 3.7. *Con las notaciones de la definición 3.4, resulta que todo vector $v \in \mathbb{Z}^n$ es congruente módulo L con un único vector de P_B.*

Definición 3.8. Sea L un retículo de $\mathbb{R}^n$. Decimos que L' es un *subretículo* de L si L' es un subgrupo de L y $\#(L/L')$ es finito.

Corolario 3.9. *Sea L un subretículo de $\mathbb{Z}^n$.*

$$[\mathbb{Z}^n : L] = \#\left(\{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n : \ 0 \leq \alpha_i < 1 \ \forall i = 1, \cdots, n\} \bigcap \mathbb{Z}^n\right),$$

donde $\{v_1, \cdots, v_n\}$ es una $\mathbb{Z}$-base de L.
Es decir $[\mathbb{Z}^n : L]$ es la cantidad de puntos con coordenadas enteras dentro del paralelotopo fundamental asociado a la base $\{v_1, \cdots, v_n\}$.

Definición 3.10. Dado un retículo $L \subset \mathbb{R}^n$ de rango n, los símplices fundamentales o elementales de L son aquellos cuyos vértices son el origen y los elementos de una $\mathbb{Z}$-base de L. El volumen en $\mathbb{R}^n$ normalizado respecto de L se denota por $Vol_L(\cdot)$ y es el que hace que los símplices fundamentales de L tengan volumen 1. En particular si $L = \mathbb{Z}^n$, a $Vol_{\mathbb{Z}^n}(\cdot)$ se le llama volumen entero.

Ejemplo 3.11. Si consideramos el cubo unidad I^n en $\mathbb{R}^n$:

$$I^n = \{(x_1, \cdots, x_n) \in \mathbb{R}^n : \ 0 \leq x_i \leq 1, \ \forall i = 1, \cdots, n\}.$$

En virtud del lema 3.6 tenemos que $Vol_{\mathbb{Z}^n}(I^n) = n!$.

Observación 3.12. Probaremos en el corolario 3.17 que, para todo polítopo racional P, su volumen entero es un número entero mayor o igual a cero. Sin embargo, a veces, resulta imposible descomponer un polítopo racional en símplices racionales con volumen entero igual a 1. Como se ve en la figura 1, el tetraedro $\mathcal{T} = EBGD$ inscripto en el cubo unidad $ABCDEFGH$ tiene volumen $Vol_{\mathbb{Z}^3}(\mathcal{T}) = 2$ pero no puede descomponerse en unión de símplices racionales más pequeños ya que no tiene puntos enteros en su interior.

Proposición 3.13. *Sea $A \subset \mathbb{Z}^n$ y L es el $\mathbb{Z}-$módulo generado por A con base $\{w_1, w_2, \cdots, w_n\}$. Entonces:*

$$[\mathbb{Z}^n : L] = |det(B)|$$

donde B es la matriz cuyas columnas son $w_1, w_2, \cdots, w_n$.

Demostración. Ver apéndice. □

Proposición 3.14.
- $Vol_{\mathbb{Z}^n}(P) = n! \, Vol_n(P) \ \ \forall$ *P polítopo racional en* $\mathbb{R}^n$.
- *Si L es un $\mathbb{Z}-$submódulo de $\mathbb{Z}^n$ de rango n entonces:*

$$Vol_{\mathbb{Z}^n}(P) = [\mathbb{Z}^n : L] \, Vol_L(P) \ \forall \ P \text{ polítopo.}$$

Estas fórmulas son consecuencias del teorema de cambio de variables y de la proposición anterior.

Observación 3.15. El resultado que establece la proposición anterior es independiente de la $\mathbb{Z}$-base de L elegida ya que se pasa de una $\mathbb{Z}$-base a otra multiplicando por una matriz de cambio de base Λ a coeficientes enteros e invertible en $\mathbb{Z}$. Luego $det(\Lambda)$ tiene que ser ± 1. Por lo que si B' es una matriz correspondiente a otra $\mathbb{Z}$-base de L:

$$[\mathbb{Z}^n : L] = |det(B)| = |det(\Lambda B')| = |det(\Lambda)||det(B')| = |det(B')|.$$

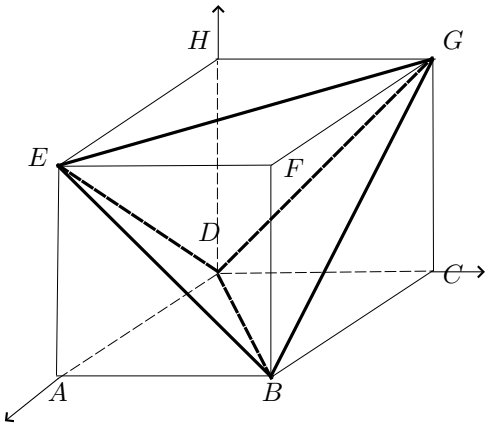


Figura 1.

Ejemplo 3.16. Consideramos Q el símplice de $\mathbb{R}^2$ tal que $Q = Conv\{(0,0),(2,0),(3,3)\}$ y sea P el paralelotopo fundamental definido por los vectores de Q.

Sea L el retículo generado por los vectores $(0,0),(2,0)$ y $(3,3)$.

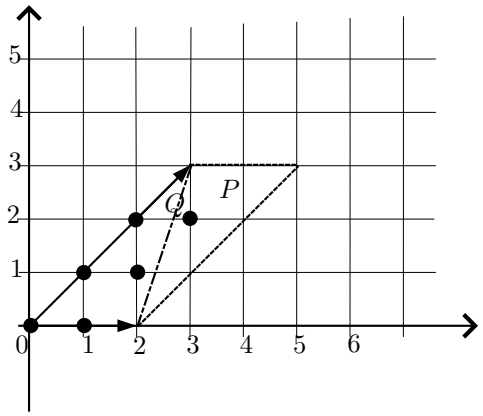


Figura 2.

Ayudándonos de la figura 2, es claro que $Vol_{\mathbb{Z}^2}(Q) = 6, Vol_{\mathbb{R}^2}(Q) = 3, Vol_{\mathbb{Z}^2}(P) = 12$ (P es unión de doce 2-símplices minimales) y $Vol_{\mathbb{R}^2}(P) = 6$.

Q es un símplice fundamental asociado a L por lo tanto $Vol_L(Q) = 1$ y se verifica la fórmula $Vol_{\mathbb{Z}^2}(Q) = [\mathbb{Z}^2 : L] \times Vol_L(Q)$ ya que $[\mathbb{Z}^2 : L] = 6$ es la cantidad de puntos enteros dentro del paralelotopo fundamental. También $Vol_L(P) = 2$ y verifica la fórmula
$Vol_{\mathbb{Z}^2}(P) = [\mathbb{Z}^2 : L] \times Vol_L(P)$ pues $Vol_{\mathbb{Z}^2}(P) = 12$.

Corolario 3.17. *Si P es un polítopo racional entonces $Vol_{\mathbb{Z}^n}(P)$ es un número entero.*

Demostración. Supongamos que $P = Conv\{x_1, \cdots, x_r\}$ entonces, por la proposición 1.35, podemos subdividir P en símplices con vértices en $\{x_1, \cdots, x_r\}$ tal que los símplices no tengan puntos interiores en común. Como para cada i, $x_i \in \mathbb{Z}^n$, los símplices de esta descomposición también son racionales. Sea Q uno de estos símplices, supongamos que uno sus vértices es $0_{\mathbb{R}^n}$ y que es n-dimensional (que son los únicos que contribuyen al volumen de P), es decir $Q = Conv\{0, x_1, \cdots, x_n\}$.

Si P_Q es el paralelotopo fundamental definido por el símplice Q entonces, por lo probado en el apéndice:

$$Vol_{\mathbb{Z}^n}(Q) = n! \times Vol_{\mathbb{R}^n}(Q) = n! \times \frac{Vol_{\mathbb{R}^n}(P_Q)}{n!} = Vol_{\mathbb{R}^n}(P_Q) = |det(A)|$$

donde $A = [x_1|x_2|\cdots|x_n]$ es una matriz con entradas enteras luego con determinante entero. Al tomar valor absoluto $|det(A)| \in \mathbb{Z}_{\geq 0}$ y se deduce la tesis. □

2. Polinomio de Ehrhart

Dado P un polítopo racional en $\mathbb{Z}^n$, es claro que el homotetizado de P de razón $d \in \mathbb{Z}$, $dP = \{dp : p \in P\}$, es también un polítopo racional.

Definición 3.18. Dado P un polítopo racional en $\mathbb{Z}^n$ definimos la función $E_P : \mathbb{N} \longrightarrow \mathbb{N}$ tal que

$$E_P(d) = \#(dP \cap \mathbb{Z}^n).$$

En 1960, Ehrhart probó que esta aplicación es un polinomio en d, que tiene actualmente el nombre de *polinomio de Ehrhart* del polítopo P.
En 1967, el propio Ehrhart probó que algunos de los coeficientes de este polinomio tienen un significado geométrico, más precisamente si P es un polítopo racional n-dimensional entonces:

$$E_P(d) = Vol_{\mathbb{R}^n}(P)d^n + \frac{1}{2}Vol_{\mathbb{R}^{n-1}}(\partial P)d^{n-1} + \cdots + \chi(P)$$

donde $Vol(\partial P)$ denota el "volumen del borde" de P y $\chi(P)$ es la característica de Euler de P.
Para los otros coeficientes, aún hoy en día, no se conoce ninguna interpretación similar.

En esta sección probaremos que dado un polítopo racional P la función E_P es un polinomio y que el coeficiente de su término principal es el volumen euclídeo de P.

Ejemplo 3.19. El polinomio de Ehrhart del segmento $I = [0, 1]$ es $E_I(d) = d + 1$ y el del cubo unitario C en el espacio es $E_C(d) = (d + 1)^3 = d^3 + 3d^2 + 3d + 1$.

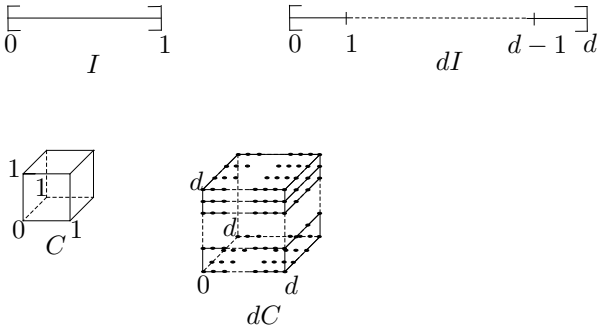


Figura 3. Ejemplos de homotetizados de polítopos y puntos enteros.

Lema 3.20. *Consideramos el símplice $S = Conv\{0, v_1, v_2, \cdots, v_n\}$ con $v_1, v_2, \cdots, v_n \in \mathbb{Z}^n$. Entonces existen constantes $\beta_1, \beta_2, \cdots, \beta_n \in \mathbb{Z}_{\geq 0}$ tales que para todo entero positivo d, la cantidad de puntos enteros contenidos en el polítopo dS, $E_S(d)$ es:*

$$E_S(d) = \begin{pmatrix} n+d \\ n \end{pmatrix} + \beta_1 \begin{pmatrix} n+d-1 \\ n \end{pmatrix} + \cdots + \beta_n \begin{pmatrix} d \\ n \end{pmatrix}.$$

En particular, $E_S(d)$ es un polinomio en d de grado n.

DEMOSTRACIÓN. La figura 4 ilustra los pasos de esta demostración en $\mathbb{Z}^2$ para $v_1 = (1,2)$ y $v_2 = (3,1)$.

Cada punto $y \in dS \cap \mathbb{Z}^n$ se escribe de manera única como

$$y = x + \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \quad (*)$$

con $x \in P_S \cap \mathbb{Z}^n$, siendo $P_S = \{\sum_{i=1}^n t_i v_i,\ 0 \leq t_i < 1,\ i = 1, \cdots, n\}$, el paralelotopo fundamental definido por el símplice S, $\alpha_i \in \mathbb{Z}_{\geq 0}$ y $\alpha_1 + \cdots + \alpha_n \leq d$. A los efectos de esta demostración diremos que y está asociado a x.

Llamamos H_j al hiperplano de $\mathbb{R}^n$ que contiene a los puntos $jv_1, \cdots, jv_n$ para $j = 0, \cdots, d$. Particionamos el conjunto $P_S \cap \mathbb{Z}^n$ con los hiperplanos $H_0, \cdots, H_d$: definimos $P_0 = \{0\}$, para $i = 1, \cdots, d$ llamamos P_i al conjunto de los $x \in P_S \cap \mathbb{Z}^n$ que están entre H_{i-1} y H_i o están en H_i. El número de puntos $y \in dS \cap \mathbb{Z}^n$ asociados a cierto $x \in P_i$ - según (*)- es exactamente la suma del número de soluciones $(\alpha_1, \cdots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$ de cada una de las ecuaciones $\alpha_1 + \cdots + \alpha_n = j$ para $j \in \mathbb{Z}$ tal que $i \leq j \leq d$. Para determinar este número procedemos en dos etapas.

1. AFIRMACIÓN: Para cada $j \in \{1, \cdots, d\}$ el número d de soluciones $(\alpha_1, \cdots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$ de la ecuación $\alpha_1 + \cdots + \alpha_n = j$ es $\begin{pmatrix} n+j-1 \\ n-1 \end{pmatrix}$.

   PRUEBA: Haremos la prueba por inducción completa en n la dimensión del espacio.

   Si $n = 1$ hay $1 = \begin{pmatrix} j \\ 0 \end{pmatrix}$ solución.

   Supongamos que la afirmación vale para $n-1$. Luego para la ecuación $\alpha_1 + \alpha_2 + \cdots + \alpha_{n-1} = j - \alpha_n$ sabemos que hay $\begin{pmatrix} n+j-\alpha_n-2 \\ n-2 \end{pmatrix}$ soluciones. Sumando sobre todos los valores posibles de α_n obtenemos:

   $$\sum_{\alpha_n=0}^{j} \begin{pmatrix} n+j-\alpha_n-2 \\ n-2 \end{pmatrix} = \begin{pmatrix} n+j-2 \\ n-2 \end{pmatrix} + \begin{pmatrix} n+j-3 \\ n-2 \end{pmatrix} + \cdots + \begin{pmatrix} n-2 \\ n-2 \end{pmatrix}.$$

   Usando inducción y la fórmula de Stieffel obtenemos:

   $$\sum_{\alpha_n=0}^{j} \begin{pmatrix} n+j-\alpha_n-2 \\ n-2 \end{pmatrix} = \begin{pmatrix} n+j-1 \\ n-1 \end{pmatrix}.$$

2. Es fácil verificar que el número de soluciones $(\alpha_1, \cdots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$ de las ecuaciones $\alpha_1 + \cdots \alpha_n = j$ con $i \leq j \leq d$ es $\sum_{j=0}^{d} \begin{pmatrix} n+j-1 \\ n-1 \end{pmatrix} = \begin{pmatrix} n+d-i \\ n \end{pmatrix}$.

   Ahora llamemos a β_i al número de elementos de P_i, este número no depende de d, además $\beta_0 = 1$ y $\beta_i = 0\ \forall\ i > n$ ($P_i = \emptyset$ si $i > n$). En particular si $d < n$ se tiene que $\beta_{d+1} = \cdots = \beta_n = 0$.

   Entonces el número de puntos de $dS \cap \mathbb{Z}^n$ es:

   $$E_S(d) = \begin{pmatrix} n+d \\ n \end{pmatrix} + \beta_1 \begin{pmatrix} n+d-1 \\ n \end{pmatrix} + \cdots + \beta_n \begin{pmatrix} d \\ n \end{pmatrix}.$$

Finalmente, como $\begin{pmatrix} n+d-j \\ n \end{pmatrix}$ es un polinomio en d de grado n y los β_i no dependen de d, tenemos que $E_S(d)$ es un polinomio en d de grado n. □

EJEMPLO 3.21. Si consideramos el símplice $S = Conv\{(0,0),(1,2),(3,1)\}$, y observando la figura 4, tenemos que $\beta_1 = 2$ y $\beta_2 = 2$.

Luego $E_S(d) = \begin{pmatrix} 2+d \\ 2 \end{pmatrix} + 2\begin{pmatrix} 1+d \\ 2 \end{pmatrix} + 2\begin{pmatrix} d \\ 2 \end{pmatrix} = \frac{5}{2}d^2 + \frac{3}{2}d + 1$.

Además:

- La cantidad de puntos enteros en $2S$ es $E_S(2)$ y $E_S(2) = \frac{5}{2}2^2 + \frac{3}{2}2 + 1 = 14$.
- La cantidad de puntos enteros en $3S$ es $E_S(3)$ y $E_S(3) = \frac{5}{2}3^2 + \frac{3}{2}3 + 1 = 28$.
- $Vol_{\mathbb{R}^2}(S) = \frac{5}{2}$ y $Vol_{\mathbb{Z}^2}(S) = 10 = 2! \times 5$.

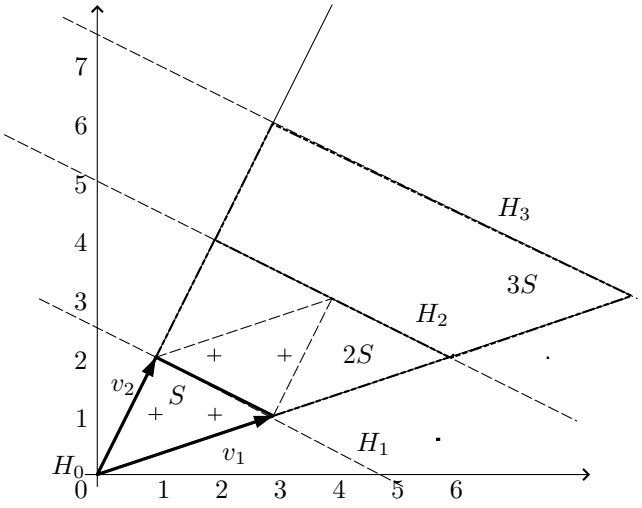


Figura 4.

- El perímetro de S es 3/2.
- La característica de Euler de S es 1 (# vértices - #aristas + # caras =3+1-3=1).

LEMA 3.22. *Sea $S = \{v_0, v_1, v_2, \cdots, v_r\}$ un r-símplice donde $v_1, v_2, \cdots, v_r \in \mathbb{Z}^n$ con $n \geq r$. Entonces $E_S(d) = \#(dS \cap \mathbb{Z}^n)$ es un polinomio en d de grado r.*

DEMOSTRACIÓN. Podemos suponer sin pérdida de generalidad, que $v_0 = 0_{\mathbb{R}^n}$. La prueba es análoga a la del lema 3.20. Sea V el subespacio de $\mathbb{R}^n$ generado por S. Cada punto y de $dS \cap \mathbb{Z}^n$ se escribe de manera única como $y = x + \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_r, \quad \alpha_i \in \mathbb{Z}_{\geq 0}$ donde x es un punto del paraleletopo fundamental definido por S. Repetimos la misma demostración que en el lema 3.20, considerando H_i el subespacio afín de V de codimensión 1 en S que contiene a los puntos $ia_1, ia_2, \cdots, ia_r$. □

DEFINICIÓN 3.23. Si $S \subset \mathbb{R}^n$ es un símplice, el *seudo-símplice* asociado a S se define como:

$$S_0 = S \backslash \partial_{rel}(S),$$

donde $\partial_{rel}(S)$ es el borde relativo de S.

OBSERVACIÓN 3.24.
- El seudo-símplice es el interior relativo del símplice.
- Un vértice se considera un seudo-símplice.

LEMA 3.25. *Sea S un símplice y S_0 el seudo-símplice asociado.*
Entonces si $d \in \mathbb{Z}_{\geq 0}$, $\#(dS_0 \cap \mathbb{Z}^n)$ es un polinomio en d de grado la dimensión de S.

DEMOSTRACIÓN. Procederemos por inducción completa en r la dimensión de S.

- Si $r = 1$ el símplice es un segmento y el seudo-símplice se obtiene sacando dos puntos. Entonces el resultado es cierto, dado que $\#(dS_0 \cap \mathbb{Z}^n) = \#(dS \cap \mathbb{Z}^n) - 2$.
- El conjunto $S \backslash S_0$ es el borde relativo del símplice S. Entonces $S \backslash S_0 = \bigcup_{F \in \mathcal{F}} F$ donde $\mathcal{F}$ es el conjunto de todas las caras propias de S. Es fácil verificar por inducción en la dimensión de S que sus caras son a su vez símplices.
  Ahora $S \backslash S_0 = \bigcup_{F \in \mathcal{F}} F_0$, donde F_0 es el seudo-símplice asociado a F, puesto que cada punto de $S \setminus S_0$ está en el interior relativo de alguna de las caras propias de S.
  Más aún $S \backslash S_0 = \biguplus_{F \in \mathcal{F}} F_0$ puesto que dos símplices diferentes de $\mathcal{F}$ tienen interiores relativos disjuntos. Los seudo-símplices F_0 con $F \in \mathcal{F}$ tienen dimensión menor que r, luego por hipótesis de inducción $\#(dF_0 \cap \mathbb{Z}^n)$ es un polinomio en d de grado menor que r. Como

$d(S \setminus S_0) = dS \setminus dS_0$, en virtud del lema 3.22, deducimos que $\#(dS_0 \cap \mathbb{Z}^n)$ es un polinomio en d de grado $r = dimS$.

□

TEOREMA 3.26 (Ehrhart). *Sea P un polítopo racional en $\mathbb{R}^n$ y $d \in \mathbb{Z}_{\geq 0}$.*
Entonces $\#(dP \cap \mathbb{Z}^n)$ es un polinomio en d de grado la dimensión del espacio generado por P.

DEMOSTRACIÓN. La proposición 1.35 nos permite descomponer P en una unión finita de símplices tal que dos a dos los símplices no tengan puntos interiores en común. Más aún, como en la demostración de 3.25, podemos descomponer P en una unión finita disjunta de seudo-símplices:

$$P = S_0^1 \uplus \cdots \uplus S_0^q.$$

La cantidad de puntos con coordenadas enteras en dP es igual a la suma de la cantidad de puntos con coordenadas enteras en cada uno de los dS_0^j, es decir:

$$\#(dP \cap \mathbb{Z}^n) = \sum_{j=1}^{q} \#(dS_0^j \cap \mathbb{Z}^n).$$

Cada $\#(dS_0^j \cap \mathbb{Z}^n)$ es un polinomio en $d \; \forall \; j = 1, \cdots, q$ de grado a lo sumo $dim(P)$ y puesto que algunos de estos polinomios tienen ese grado, se deduce la tesis. □

Ahora probaremos que el coeficiente del término principal del polinomio de Ehrhart de un polítopo racional P es el volumen euclídeo de P, es decir:

$$\lim_{d \longrightarrow \infty} \frac{\#(dP \cap \mathbb{Z}^n)}{d^n} = \frac{Vol_{\mathbb{Z}^n}(P)}{n!}.$$

LEMA 3.27. *Sea Q un n-símplice racional en $\mathbb{R}^n$. Entonces $\#(Q \cap \mathbb{Z}^n) \leq Vol_{\mathbb{Z}^n}(Q) + n$.*

DEMOSTRACIÓN. Sea $\{v_0, v_1, v_2, \cdots, v_n\}$ el conjunto de vértices del símplice Q. Suponemos, sin pérdida de generalidad, que $v_0 = 0$, en caso contrario trasladamos a Q. Sea S el $\mathbb{Z}-$submódulo generado por $\{v_1, v_2, \cdots v_n\}$.
$\#(Q \cap \mathbb{Z}^n) \leq [\mathbb{Z}^n : S] + n$, ya que $[\mathbb{Z}^n : S]$ es la cantidad de puntos con coordenadas enteras en P_Q, el paralelotopo fundamental definido por el símplice Q. Además, por lo que probamos en la proposición 3.13 y en la sección 4 del apéndice tenemos que:

$$[\mathbb{Z}^n : S] = Vol_{\mathbb{R}^n}(P_Q).$$

Entonces:

$$\#(Q \cap \mathbb{Z}^n) \leq Vol_{\mathbb{R}^n}(P_Q) + n = \frac{Vol_{\mathbb{Z}^n}(P_Q)}{n!} + n = Vol_{\mathbb{Z}^n}(Q) + n.$$

□

COROLARIO 3.28. *Si $P \subset \mathbb{Z}^n$ es un polítopo racional de dimensión n entonces*
$\#(P \cap \mathbb{Z}^n) \leq \frac{Vol_{\mathbb{Z}^n}(P)}{n!} + tn$ donde t es la cantidad de símplices en la descomposición de P, dada en la proposición 1.35.

DEMOSTRACIÓN. Recordamos que, por la proposición 1.35, todo polítopo se puede descomponer en unión finita de símplices sin puntos interiores en común. Por lo cual si $t \in \mathbb{Z}_{\geq 0}$ es la cantidad de símplices en la descomposición de P, aplicando a cada uno de los símplices el lema anterior se obtiene la tesis. □

TEOREMA 3.29. *Si P es un polítopo racional n-dimensional entonces:*

$$\lim_{d \longrightarrow \infty} \frac{\#(dP \cap \mathbb{Z}^n)}{d^n} = \frac{Vol_{\mathbb{Z}^n}(P)}{n!}.$$

DEMOSTRACIÓN. Consideramos el polítopo dP para d entero positivo. Al ser P racional recordamos que dP también es un polítopo racional.
Denotamos por $a(d)$ la cantidad de cubos unitarios $n-$dimensionales con vértices enteros incluidos totalmente en dP (recordar que dP es convexo por serlo P) y $c(d)$ denota la unión de tales cubos.

- Como $c(d) \subseteq dP$ entonces aplicando la función volumen entero se obtiene que $n!\, a(d) \leq Vol_{\mathbb{Z}^n}(dP)$ pues los cubos tienen volumen entero $n!$.
- Si a cada cubo n-dimensional le asociamos uno de sus vértices, podemos definir una inyección del conjunto de estos cubos en los puntos enteros de dP, obteniéndose la desigualdad $\#(dP \cap \mathbb{Z}^n) \geq a(d)$, luego:

$$0 \leq \#(dP \cap \mathbb{Z}^n) - a(d). \tag{2.1}$$

- Sea $B(d)$ el conjunto de puntos en dP cuya distancia del borde es menor o igual a $\sqrt{n}$, es decir

$$B(d) = \{x \in dP |\ d(x, \partial(dP)) \leq \sqrt{n}\}.$$

Al ser dP convexo, $B(d) \subseteq B_1 \cup B_2 \cup \cdots \cup B_r$ donde cada B_i es un paralelotopo recto de base en las muros de dP y altura $\sqrt{n}$.

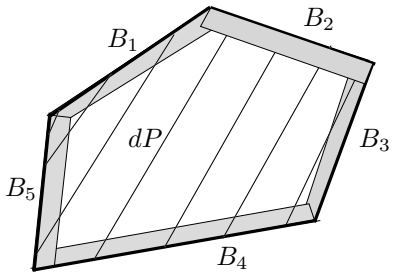


Figura 5. Ilustración de la demostración del teorema 3.29.

Aplicando la función $Vol_{\mathbb{Z}^n}(.)$ a la inclusión $dP \backslash c(d) \subseteq B(d)$ obtenemos:

$$0 \leq Vol_{\mathbb{Z}^n}(dP) - a(d)n! \leq Vol_{\mathbb{Z}^n}(B(d)).$$

De donde:

$$0 \leq Vol_{\mathbb{Z}^n}(dP) - a(d)n! \leq cVol_{\mathbb{Z}^{n-1}}(\partial(dP))\sqrt{n} = cd^{n-1}Vol_{\mathbb{Z}^{n-1}}(\partial P)\sqrt{n}. \tag{2.2}$$

donde c es un número real positivo.

Dividiendo por d^n:

$$0 \leq \frac{Vol_{\mathbb{Z}^n}(dP) - a(d)n!}{d^n} \leq \frac{cVol_{\mathbb{Z}^{n-1}}(\partial P)\sqrt{n}}{d} \longrightarrow 0. \tag{2.3}$$

cuando d tiende a ∞.

De donde:

$$\lim_{d \longrightarrow \infty} \frac{a(d)}{d^n} = \lim_{d \longrightarrow \infty} \frac{Vol_{\mathbb{Z}^n}(dP)}{n!\, d^n} = \frac{Vol_{\mathbb{Z}^n}(P)}{n!}. \tag{2.4}$$

Por otro lado se tiene que: $0 \leq \#(dP \cap \mathbb{Z}^n) - a(d) \leq \#(B(d) \cap \mathbb{Z}^n) \leq \sum_{i=1}^{r} \#(B_i \cap \mathbb{Z}^n)$.

Usando el corolario 3.28 se obtiene que:

$$\#(dP \cap \mathbb{Z}^n) - a(d) \leq \sum_{i=1}^{r} Vol_{\mathbb{Z}^n}(B_i) + t_i n. \tag{2.5}$$

donde t_i es la cantidad de símplices obtenidos en la descomposición de B_i de acuerdo a la proposición 1.35.

Luego, si $M = \sum_{i=1}^{r} t_i$:

$$\sum_{i=1}^{r} Vol_{\mathbb{Z}^n}(B_i) + t_i n = Mn + \sum_{i=1}^{r} Vol_{\mathbb{Z}^n}(B_i)$$

$$= Mn + \sqrt{n}\, Vol_{\mathbb{Z}^{n-1}}(\partial(dP)) = Mn + d^{n-1}\sqrt{n}\, Vol_{\mathbb{Z}^{n-1}}(P).$$

Dividiendo esta desigualdad por d^n y haciendo $d \longrightarrow \infty$ obtenemos que:

$$0 \leq \lim_{d\longrightarrow\infty} \frac{\#(dP \cap \mathbb{Z}^n) - a(d)}{d^n} \leq \lim_{d\longrightarrow\infty} \frac{Mn + d^{n-1}\sqrt{n}\, Vol_{\mathbb{Z}^{n-1}}(P)}{d^n} = 0.$$

Luego:

$$\lim_{d\longrightarrow\infty} \frac{\#(dP \cap \mathbb{Z}^n)}{d^n} = \lim_{d\longrightarrow\infty} \frac{a(d)}{d^n} = \frac{Vol_{\mathbb{Z}^n}(P)}{n!}$$

obteniéndose la tesis.

□

CAPíTULO 4

Teorema de Kushnirenko

Recordamos que nuestro propósito es determinar la cantidad de soluciones de un sistema de ecuaciones polinomiales con exponentes prefijados. En este capítulo estudiaremos el teorema probado por A. Kushnirenko en 1976 que nos da una respuesta para el caso de las soluciones en el toro algebraico $(\mathbb{C}^*)^t$ de un sistema de polinomios de Laurent. En una primera parte introduciremos muy brevemente el concepto de resultante rala de polinomios de Laurent. Luego enunciaremos y demostraremos el teorema de Kushnirenko y finalmente expondremos algunos ejemplos.

1. Resultante rala

La resultante generaliza la noción de compatibilidad de los sistemas lineales con coeficientes en un cuerpo k:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases}, \ a_{ij} \in k$$

Sabemos que existe una solución no trivial $X_0 = (x_1^0, x_2^0, \cdots, x_n^0)$ si y solamente si $detA = 0$, donde $A = (a_{ij})_{i,j}$. Observar que el determinante es un polinomio a coeficientes enteros en las entradas de la matriz, o sea $det(A) \in \mathbb{Z}[a_{ij} : i, j = 1, \cdots, n]$.

Ahora vamos a introducir el concepto de resultante rala para polinomios de Laurent.

Recordamos nuestras notaciones: si $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_t) \in \mathbb{Z}^n$ y $X \in (\mathbb{C}^*)^t$ entonces X^α corresponde al monomio de Laurent $x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_t^{\alpha_t}$ y que el anillo de polinomios de Laurent $\mathbb{C}[x_1^{\pm 1}, x_2^{\pm 1}, \cdots, x_t^{\pm 1}]$ se denota a veces $\mathbb{C}[X, X^{-1}]$.

Sea $A = \{v_1, \cdots, v_n\}$ un subconjunto finito de $\mathbb{Z}^t$ y sean $f_0, \cdots, f_t \in C^A$, donde

$$C^A = \left\{ f \in \mathbb{C}[X, X^{-1}] \ / \ f(X) = \sum_{v \in A} a_v X^v, \ X \in (\mathbb{C}^*)^t \right\}. \tag{1.1}$$

Nos preguntamos que condiciones deben verificar los coeficientes de los polinomios de Laurent $f_0, \cdots, f_t$ para que el sistema $f_0 = \cdots = f_t = 0$ tenga solución en $(\mathbb{C}^*)^t$. O sea queremos saber si el sistema:

$$\begin{cases} f_0(X) = c_{01}X^{v_1} + \cdots + c_{0n}X^{v_n} = 0 \\ f_1(X) = c_{11}X^{v_1} + \cdots + c_{1n}X^{v_n} = 0 \\ \vdots \\ f_t(X) = c_{t1}X^{v_1} + \cdots + c_{tn}X^{v_n} = 0 \end{cases} \quad c_{ij} \in \mathbb{C}, \ X \in (\mathbb{C}^*)^t. \tag{1.2}$$

tiene solución o no.

Convenimos que entendemos por "polinomio en los coeficientes de $f_0, \cdots, f_t$" lo siguiente: para cada coeficiente c_{ij} de f_j con $0 \leq j \leq n$ introducimos una variable u_{ij}. Si P es un polinomio en $\mathbb{C}[u_{ij}]$ entonces $P(f_0, \cdots, f_t)$ denota el elemento que se obtiene reemplazando cada variable u_{ij} por el coeficiente c_{ij} de los polinomios de Laurent $f_0, \cdots, f_t$.

Teorema 4.1. *Sea $A \subset \mathbb{Z}^t$ finito tal que $Conv(A) = Q$ es un polítopo t dimensional. Entonces existe un único (a menos de signo) polinomio irreducible en $\mathbb{Z}[u_{ij}]$, que notamos por Res_A, tal que*

si el sistema de polinomios de Laurent $f_0 = \cdots = f_t = 0$*, con* $f_i \in C^A \; \forall \; i = 0, \cdots, t$*, tiene solución en* $(\mathbb{C}^*)^t$ *entonces* $Res_A(f_0, \cdots, f_t) = 0$*.*

DEMOSTRACIÓN. Ver el capítulo 8 de [**6**] y el capítulo 7 de [**2**]. □

DEFINICIÓN 4.2. La *resultante rala* (en inglés sparse resultant) o *A-resultante* de los polinomios de Laurent $f_0, \cdots, f_t \in C^A$ es un polinomio irreducible P en los coeficientes de $f_0, \cdots, f_t$, que se anula cuando existe una solución del sistema 1.2.
La notamos por $Res_A(f_0, \cdots, f_t)$.

A modo de ejemplo, fijado un conjunto $A = \{0, 1, \cdots, n\}$ supongamos que tenemos dos polinomios f y g en C^A (en una variable) y de grado n, es decir:
$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ y $g(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$ son polinomios de $k[x]$ con $n \geq 0,\ a_n \neq 0, b_n \neq 0$.

Entonces, la resultante rala coincide con la resultante de Silvester de los polinomios f y g. Esto es consecuencia del teorema anterior y de la unicidad de la resultante de Silvester.
En general, la *resultante de Silvester* de dos polinomios f y g, de grados n y m respectivamente se define como:

$$Res(f,g,x) = det \begin{pmatrix} a_n & 0 & \cdots & \cdots & 0 & b_m & 0 & \cdots & 0 \\ a_{n-1} & a_n & \ddots & & \vdots & b_{m-1} & b_m & \ddots & \vdots \\ a_{n-2} & a_{n-1} & \ddots & \ddots & \vdots & b_{m-2} & b_{m-1} & \ddots & 0 \\ \vdots & a_{n-2} & \ddots & \ddots & 0 & \vdots & b_{m-2} & \ddots & b_m \\ a_0 & \vdots & \ddots & \ddots & a_n & \vdots & \vdots & \ddots & b_{m-1} \\ 0 & a_0 & & \ddots & a_{n-1} & b_0 & \vdots & & b_{m-2} \\ \vdots & \ddots & \ddots & & a_{n-2} & 0 & b_0 & & \vdots \\ \vdots & & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & a_0 & 0 & \cdots & 0 & b_0 \end{pmatrix}$$

EJEMPLO 4.3. Supongamos que $f(x) = x^3 + 4x - 1$ y $g(x) = 2x^2 + 3x + 7$. Entonces:

$$Res(f,g,x) = det \begin{pmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 3 & 2 & 0 \\ 4 & 0 & 7 & 3 & 2 \\ -1 & 4 & 0 & 7 & 3 \\ 0 & -1 & 0 & 0 & 7 \end{pmatrix} = 159$$

.

PROPOSICIÓN 4.4. *Supongamos que* $f, g \in k[x]$ *son dos polinomios de grado positivo.*

1. $Res(f,g,x)$ *es un polinomio en los coeficientes de* f *y de* g*.*
   Es decir $Res(f,g,x) = Res_{n,m}(a_0, \cdots, a_n, b_0, \cdots, b_m) \in \mathbb{Z}[u_0, \cdots, u_n, v_0, \cdots, v_m]$*, donde* n *y* m *son respectivamente los grados de* f *y de* g*.*
2. $Res(f,g,x) = 0$ *si y solamente si* f *y* g *tienen algún factor en común.*
3. *Existen polinomios* p *y* q *en* $k[x]$ *tales que* $pf + qg = Res(f,g,x)$*.*

DEMOSTRACIÓN. : Ver capítulo 3 de [**2**]. □

OBSERVACIÓN 4.5. Si $k = \mathbb{C}$, la propiedad 2. implica en particular que la resultante de dos polinomios es nula si y solamente si tienen una raíz en común. En el ejemplo 4.3 los polinomios f y g no tienen raíces comunes.

Por más información sobre la resultante y sus propiedades, recomendamos al lector consultar [**6**] y [**2**].

2. Enunciado y demostración del teorema de Kushnirenko

Nos proponemos determinar el número de raíces comunes en el toro algebraico $(\mathbb{C}^*)^t$ de una cantidad finita de polinomios de Laurent en t variables con exponentes prefijados en un mismo conjunto de vectores, es decir buscamos la cantidad de soluciones en $(\mathbb{C}^*)^t$ del sistema:

$$\begin{cases} f_1(x_1, x_2, \cdots, x_t) = \sum_{v\in A} c_v^1 X^v = 0 \\ f_2(x_1, x_2, \cdots, x_t) = \sum_{v\in A} c_v^2 X^v = 0 \\ \vdots \\ f_m(x_1, x_2, \cdots, x_t) = \sum_{v\in A} c_v^t X^v = 0 \end{cases} \tag{2.1}$$

donde $f_i \in C^A$, $\forall\, i = 1, \cdots, m$.

Por ejemplo, si $A = \{s, \cdots, n\} \subset \mathbb{Z}$ con $s \leq n$, el polinomio en una variable $f(x) = a_n x^n + \cdots + a_s x^s$ con exponentes prefijados en A tiene, en general, $n - s$ raíces distintas en $\mathbb{C}^*$. Observamos que el polítopo de Newton de f es el segmento $[n, s]$ cuya longitud es también $n - s$. Por lo que, en general, la cantidad de raíces distintas en $\mathbb{C}^*$ de un polinomio en una variable coincide con la longitud (el volumen unidimensional) del polítopo de Newton del polinomio.

El teorema de Kushnirenko generaliza este resultado.

Teorema 4.6 (Kushnirenko (1976)). *Sea $A = \{v_1, v_2, \cdots, v_n\}$ un subconjunto finito de vectores de $\mathbb{Z}^t$ y $Q = Conv(A) \subset \mathbb{R}^t$.*

Para una elección genérica de t polinomios de Laurent $f_1, f_2, \cdots, f_t \in C^A$ la cantidad de raíces comunes es $Vol_{\mathbb{Z}^t}(Q)$.

Es decir la cantidad de raíces comunes sobre el toro algebraico depende, genéricamente, únicamente del "tamaño" del polítopo de Newton de los polinomios.

Antes de pasar a la demostración del teorema, haremos algunos comentarios sobre el significado y la pertinencia de las hipótesis.

1. Explicitemos la noción de genericidad que corresponde a "la cantidad de raíces comunes esperadas".

   Definición 4.7. Se dice que una propiedad *vale genéricamente* para los polinomios de Laurent $f_1, f_2, \cdots, f_t \in C^A$ si existe un polinomio P no nulo en los coeficientes de estos t polinomios de manera que si $P(f_0, \cdots, f_t) \neq 0$ la propiedad se cumple.

   Geométricamente esta noción se corresponde con el hecho que los coeficientes c_{ij} de los polinomios de Laurent considerados pertenezcan a un abierto Zariski de $\mathbb{A}^{nt}$.

   O más intuitivamente estamos diciendo que esta propiedad vale para la gran mayoría de las elecciones de estos t polinomios, pero no necesariamente para toda elección.
   Por ejemplo:
   a) El polinomio $ax^2 + bx + c$ tiene, genéricamente, dos raíces distintas: esto pasa si $a(b^2 - 4ac) \neq 0$.
   Un polinomio que sirve para asegurar la validez de la propiedad es $P(\alpha, \beta, \gamma) = \alpha(\beta^2 - 4\alpha\gamma)$.
   b) Genéricamente, dos rectas se cortan en un solo punto. Esto equivale a considerar las raíces comunes de los polinomios $f(x, y) = ax + by + c$ y $g(x, y) = a'x + b'y + c'$.
   Un polinomio que sirve es $P(\alpha, \beta, \gamma, \delta) = \alpha\delta - \beta\gamma$, o sea el "determinante".

   Por último observamos que, genéricamente, cada uno de los polinomios $f_1, \cdots, f_t \in C^A$ tienen exponentes cuya envolvente convexa es $Conv(A)$: basta poner como factores del polinomio P de la definición 4.7 a las indeterminadas correspondientes a los coeficientes de $f_1, \cdots, f_t$ asociados a los vértices de $Conv(A)$.
2. Podemos suponer que la cantidad de polinomios coincide con la cantidad de variables.
   a) Si en el sistema (2.1) hay más variables que polinomios, o sea si $t > m$, genéricamente hay infinitas soluciones. Efectivamente, como queremos hallar la cantidad de soluciones sobre el toro algebraico, podemos multiplicar cada uno de los polinomios

de Laurent por monomios adecuados, sin agregar raíces, de tal manera que los polinomios tengan todos exponentes positivos. Cada una de las m ecuaciones puede ser considerada como la ecuación de una hipersuperficie y las soluciones del sistema (2.1) como la intersección de todas. La proposición 7.1, página 48 de [**7**] nos indica que dicha intersección es una variedad de dimensión mayor o igual a 1, lo cual implica que los polinomios tienen infinitas raíces en común.

b) Si en el sistema (2.1) hay más polinomios que variables, o sea si $m > t$, genéricamente, no hay soluciones en $(\mathbb{C}^*)^t$.
Al considerar el sistema

$$\begin{cases} f_1(x_1, \cdots, x_t) = 0 \\ f_2(x_1, \cdots, x_t) = 0 \\ \vdots \\ f_m(x_1, \cdots, x_t) = 0 \end{cases} \tag{2.2}$$

Tenemos m polinomios de Laurent $f_1, f_2, \cdots, f_t, f_{t+1}, \cdots, f_m$ con variables $x_1, \cdots, x_t$. Luego por el teorema 4.1, para una elección genérica de $t+1$ polinomios $f_1, f_2, \cdots, f_{t+1}$ en C^A se tiene que $Res_A(f_1, f_2, \cdots, f_{t+1}) \neq 0$, es decir el sistema

$$\begin{cases} f_1(x_1, \cdots, x_t) = 0 \\ f_2(x_1, \cdots, x_t) = 0 \\ \vdots \\ f_{t+1}(x_1, \cdots, x_t) = 0 \end{cases} \tag{2.3}$$

no tiene solución en $(\mathbb{C}^*)^t$ por lo que tampoco el sistema (2.2).

Seguiremos los pasos de la demostración del teorema de Kushnirenko dada en [**6**].

Demostración. ***Primer paso.*** Veamos algunas observaciones que nos permiten simplificar el problema:

1. Podemos suponer que A contiene al vector nulo $0_{\mathbb{R}^t}$ de $\mathbb{R}^t$.
   Esto corresponde en el sistema:

$$\begin{cases} f_1(X) = \sum_{j=1}^{n} c_{1j} X^{v_j} = 0 \\ f_2(X) = \sum_{j=1}^{n} c_{2j} X^{v_j} = 0 \\ \vdots \\ f_t(X) = \sum_{j=1}^{n} c_{tj} X^{v_j} = 0 \end{cases} \tag{2.4}$$

   a multiplicar por un mismo monomio. Esta operación no modifica la cantidad de raíces comunes de los polinomios y corresponde a trasladar los vectores de A; por lo tanto el volumen de la envolvente convexa de A no cambia.
2. Podemos suponer que A sea de rango t, o sea genera a $\mathbb{R}^t$ como espacio vectorial. Si el rango de A es menor que t, haciendo el mismo cálculo que haremos en la parte siguiente y usando el mismo argumento que en "2.b)", se prueba que genéricamente, el sistema no tiene soluciones en $(\mathbb{C}^*)^t$.
   Por otro lado si A no genera a $\mathbb{R}^t$ entonces Q está contenido en un subespacio de codimensión ≥ 1, luego $Vol_{\mathbb{Z}^t}(Q) = 0$, lo cual es consistente con que no haya raíces comunes.
3. Vamos a probar que no se pierde generalidad si se supone que A genera a $\mathbb{Z}^t$ como $\mathbb{Z}$-módulo.

   Sea $B = \{w_1, w_2, \cdots, w_t\}$ una base de S, el $\mathbb{Z}-$módulo generado por A. Notaremos $S =< A >_{\mathbb{Z}}$ o $S = \mathbb{Z}A$. Esta base no tiene porque tener a los vectores de A, pero éstos se pueden escribir como combinaciones lineales de vectores de B, o sea $v_i = \sum_{j=1}^{t} \beta_j^i w_j$ $\forall\ i = 1, \cdots, n$.

   Si hacemos el cambio de variable $y_j = X^{w_j}$ obtenemos que
   $f(X) = g(y_1, y_2, \cdots, y_t) = g(Y)$.

Es decir pasamos del sistema:

$$\begin{cases} f_1(X) = 0 \\ f_2(X) = 0 \\ \vdots \\ f_t(X) = 0 \end{cases} \tag{2.5}$$

al sistema:

$$\begin{cases} g_1(Y) = 0 \\ g_2(Y) = 0 \\ \vdots \\ g_t(Y) = 0 \end{cases} \tag{2.6}$$

Vamos a ver que conociendo la cantidad de soluciones del sistema (2.6) podemos deducir la cantidad de soluciones del sistema (2.5).

Sea $Y = (y_1, y_2, \cdots, y_t)$ una solución del sistema (2.6), entonces si queremos ver que solución(es) le corresponde(n) en el sistema (2.5) tendremos que resolver el sistema:

$$\begin{cases} x_1^{\lambda_{11}} x_2^{\lambda_{12}} \cdots x_t^{\lambda_{1t}} = y_1 \\ x_1^{\lambda_{21}} x_2^{\lambda_{22}} \cdots x_t^{\lambda_{2t}} = y_2 \\ x_1^{\lambda_{31}} x_2^{\lambda_{32}} \cdots x_t^{\lambda_{3t}} = y_3 \\ \vdots \\ \vdots \\ x_1^{\lambda_{t1}} x_2^{\lambda_{t2}} \cdots x_t^{\lambda_{tt}} = y_t \end{cases} \tag{2.7}$$

donde $(\lambda_{11}, \lambda_{12}, \cdots, \lambda_{1t}) = w_1$, $(\lambda_{21}, \lambda_{22}, \cdots, \lambda_{2t}) = w_2, \cdots$, $(\lambda_{t1}, \lambda_{t2}, \cdots, \lambda_{tt}) = w_t$ son los vectores de la base B de S.

En el apéndice se prueba que si $\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \cdots & \lambda_{1t} \\ \lambda_{21} & \lambda_{22} & \cdots & \cdots & \lambda_{2t} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ \lambda_{t1} & \lambda_{t2} & \cdots & \cdots & \lambda_{tt} \end{pmatrix}$ entonces el sistema (2.7) tiene $m = |det(\Lambda)|$ soluciones. Luego para cada raíz común Y de los polinomios $g_1, g_2, \cdots, g_t$ obtenemos m raíces comunes a los polinomios $f_1, f_2, \cdots, f_t$. Por lo cual:

$$m\left(\#\{Y \in (\mathbb{C}^*)^t \mid g_j(Y) = 0 \; \forall \; j = 1, \cdots, t\}\right) = \#\{X \in (\mathbb{C}^*)^t \mid f_j(X) = 0 \; \forall \; j = 1, \cdots, t\}.$$

Finalmente veremos que alcanza probar el teorema cuando el $\mathbb{Z}$-módulo generado por el conjunto $A = \{v_1, v_2, \cdots, v_n\}$ es $\mathbb{Z}^t$, o sea:

$$S =< A >_{\mathbb{Z}} = \{\alpha_0 0_{\mathbb{R}^t} + \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \mid \; \alpha_i \in \mathbb{Z} \; \forall \; i = 0, 1, \cdots, n\} = \mathbb{Z}^t.$$

Nuevamente si $S =< A >_{\mathbb{Z}}$ no coincide con $\mathbb{Z}^t$, con el mismo cambio de variable usado anteriormente, pasamos del sistema (2.5) al sistema (2.6) a través del mapa:

$$T : \mathbb{R}^t \longrightarrow \mathbb{R}^t / \; w_i \mapsto e_i = (0, 0, \cdots, 0, 1, 0, \cdots, 0).$$

Es claro que T es una transformación lineal biyectiva y lleva una base de $\mathbb{R}^t$ en la base canónica de $\mathbb{R}^t$. Entonces:

- T lleva el polítopo Q en el polítopo $T(Q)$,
- los polinomios $g_j's$ tienen exponentes en $T(A)$,
- $T(S) =< T(A) >_{\mathbb{Z}} = \mathbb{Z}^t$.

Luego si probamos que para el sistema (2.6), la cantidad de raíces comunes es $Vol_{\mathbb{Z}^t}(T(Q))$, la tesis se deduce automáticamente, pues:

$$\#\{X \in (\mathbb{C}^*)^t \mid f_j(X) = 0 \ \forall\ j = 1, \cdots, t\} = m\left(\#\{Y \in (\mathbb{C}^*)^t \mid g_j(Y) = 0 \ \forall\ j = 1, \cdots, t\}\right)$$

$$= m \times Vol_{Z^t}(T(Q)) = m \times t! \times Vol_{\mathbb{R}^t}(T(Q)).$$

Como:

$$m \times Vol_{\mathbb{R}^t}(T(Q)) = |det(T^{-1})| \int_{T(Q)} 1 = |det(_C((T^{-1}))_C)| \int_{T(Q)} 1$$

$$= |det([w_1|w_2|\cdots|w_t])| \int_{T(Q)} 1 = Vol_{\mathbb{R}^t}(Q),$$

entonces:

$$\#\{X \in (\mathbb{C}^*)^t \mid f_j(X) = 0 \ \forall\ j = 1, \cdots, t\} = m \times t! \times Vol_{\mathbb{R}^t}(T(Q)) = t! \times Vol_{\mathbb{R}^t}(Q) = Vol_{\mathbb{Z}^t}(Q).$$

Resumiendo, podremos suponer en la demostración que si $A = \{v_1, \cdots, v_n\} \subset \mathbb{Z}^t$:

1. $0_{\mathbb{R}^t}$ pertenece al conjunto A.
2. A genera a $\mathbb{Z}^t$ como $\mathbb{Z}$-módulo (en particular A genera $\mathbb{R}^t$ como espacio vectorial).

Segundo paso.

Sea $X_A \subset \mathbb{P}^{n-1}$ la variedad algebraica proyectiva definida en el capítulo 2. Recordamos que:

1. El mapa:

$$\begin{array}{rccc} \varphi: & (\mathbb{C}^*)^t \times \mathbb{P}^{n-1} & \longrightarrow & \mathbb{P}^{n-1} \\ & (X, [p_1 : p_2 : \cdots : p_n]) & \longmapsto & [X^{v_1}p_1 : X^{v_2}p_2 : \cdots : X^{v_n}p_n] \end{array}$$

   define una acción del toro algebraico, como grupo algebraico, $(\mathbb{C}^*)^t$ sobre $\mathbb{P}^{n-1}$.
2. El conjunto X_A^0 definido como:

$$X_A^0 = \{[X^{v_1} : X^{v_2} : \cdots : X^{v_n}] : \ \ X = (x_1, x_2, \cdots, x_t) \in (\mathbb{C}^*)^t\}$$

   es la órbita del punto $[1 : \cdots : 1]$ y X_A es su clausura con la topología de Zariski.
3. Por el lema de la órbita cerrada, X_A es una variedad proyectiva de dimensión t, ${X_A}^0$ es localmente cerrado en $\mathbb{P}^{n-1}$ y $X_A \setminus {X_A}^0$ es unión de órbitas de dimensión menor.

Una forma lineal $L(y)$ en $\mathbb{P}^{n-1}$ se escribe como:

$$L(y) = \sum_{i=1}^{n} \alpha_i y_i$$

donde $y = [y_1 : y_2 : \cdots : y_n] \in \mathbb{P}^{n-1}$.
Sean $L_1, L_2, \cdots, L_t$, t formas lineales genéricas en $\mathbb{P}^{n-1}$ y consideramos el subespacio $\mathcal{L}$ definido como

$$\mathcal{L} = \left\{y \in \mathbb{P}^{n-1} / \ L_1(y) = 0,\ L_2(y) = 0,\ \ldots,\ L_t(y) = 0\right\} \subset \mathbb{P}^{n-1}.$$

$\mathcal{L}$ es un subespacio genérico de $\mathbb{P}^{n-1}$ de codimensión t.

Recordamos que el mapa $\varphi_{[1:\cdots:1]} = \varphi(\ \cdot\ , [1 : \cdots : 1]) : (\mathbb{C}^*)^t \longrightarrow \mathbb{P}^{n-1}$ donde

$$\varphi_{[1:\cdots:1]}(X) = [X^{v_1} : \cdots : X^{v_n}].$$

es inyectivo.

Si componemos las formas lineales genéricas $L_1, \cdots, L_t$ con la inyección del toro $(\mathbb{C}^*)^t$ en X_A obtenemos polinomios de Laurent genéricos en C^A, es decir tenemos :

$$L_j \circ \varphi_{[1:\cdots:1]}(X) = f_j(X) = \sum_{k=1}^{n} c_{jk} X^{v_k}.$$

Por lo cual deducimos que, para una elección genérica:

$$\#\{y \in \mathbb{P}^{n-1} |\ y \in X_A^0 \cap \mathcal{L}\} = \#\{X \in (\mathbb{C}^*)^t |\ f_1(X) = f_2(X) = \cdots = f_t(X) = 0\}$$

Finalmente al ser $\mathcal{L}$ un subespacio genérico de $\mathbb{P}^{n-1}$ de codimensión t, podemos "afinar" la genericidad de $\mathcal{L}$ para que valga que $\mathcal{L} \cap (X_A \backslash X_A^0) = \emptyset$, puesto que $X_A \backslash X_A^0$ es unión de órbitas de dimensión menor a t. Esto quiere decir que, genéricamente, si miramos la intersección de $\mathcal{L}$ con X_A no añadimos nuevos puntos. O sea:

$$\#\{y \in \mathbb{P}^{n-1} |\ y \in X_A^0 \cap \mathcal{L}\} = \#\{y \in \mathbb{P}^{n-1} |\ x \in X_A \cap \mathcal{L}\} = deg(X_A).$$

donde $deg(X_A)$ es el grado de la variedad proyectiva X_A definido en el capítulo 2.

Por lo que concluímos que el número que estamos buscando es igual al grado de la variedad proyectiva X_A, es decir:

$$\#\{X \in (\mathbb{C}^*)^t | f_1(X) = f_2(X) = \cdots = f_t(X) = 0\} = deg(X_A)$$

Tercer paso.

En esta parte probaremos que $deg(X_A) = Vol_{\mathbb{Z}^t}(Q)$ lo cual termina de probar el teorema de Kushnirenko.

$$S[X_A] = \frac{\mathbb{C}[y_1, y_2, \cdots, y_n]}{\mathfrak{I}_{\mathbb{P}}(X_A)} = \frac{\mathbb{C}[y_1, y_2, \cdots, y_n]}{\mathfrak{I}_{\mathbb{A}}(Y_A)}$$

ya que $\mathfrak{I}_{\mathbb{A}}(Y_A) = \mathfrak{I}_{\mathbb{P}}(X_A)$ por la proposición 2.45, donde Y_A es el cono sobre la variedad proyectiva X_A.

Por lo tanto $\mathfrak{I}_{\mathbb{A}}(Y_A)$ es un ideal homogéneo lo cual implica que $\frac{\mathbb{C}[y_1, y_2, \cdots, y_n]}{\mathfrak{I}_{\mathbb{A}}(Y_A)}$ es un anillo graduado o sea:

$$\frac{\mathbb{C}[y_1, y_2, \cdots, y_n]}{\mathfrak{I}_{\mathbb{A}}(Y_A)} = \bigoplus_{p \geq 0} A_p$$

donde para cada d fijo $A_d = \left(\frac{\mathbb{C}[y_1, y_2, \cdots, y_n]}{\mathfrak{I}_{\mathbb{A}}(Y_A)}\right)_d$ es el espacio vectorial de las clases de las combinaciones lineales de monomios de grado total d en $y_1, y_2, \cdots, y_n$ (el polinomio 0 se considera como polinomio homogéneo de todos los grados).

Consideramos M_{0,Y_A} el ideal de los polinomios de $\mathbb{C}[Y_A]$ que se anulan en 0 , es decir

$$M_{0,Y_A} = \{f \in \mathbb{C}[Y_A] :\ f(0) = 0\}$$

Es claro que: $M_{0,Y_A} = \bigoplus_{p \geq 1} \left(\frac{\mathbb{C}[y_1, y_2, \cdots, y_n]}{\mathfrak{I}_{\mathbb{A}}(Y_A)}\right)_p$, $\cdots$, $M_{0,Y_A}^d = \bigoplus_{p \geq d} \left(\frac{\mathbb{C}[y_1, y_2, \cdots, y_n]}{\mathfrak{I}_{\mathbb{A}}(Y_A)}\right)_p$.

Si consideramos el cociente $\frac{M_{0,Y_A}^d}{M_{0,Y_A}^{d+1}}$ éste coincide con el espacio vectorial de las polinomios de grado d (siempre con el polinomio cero). Luego:

$$\frac{M_{0,Y_A}^d}{M_{0,Y_A}^{d+1}} = \frac{\bigoplus_{p \geq d} A_p}{\bigoplus_{p \geq d+1} A_p} = \frac{A_d \oplus \bigoplus_{p \geq d+1} A_p}{\bigoplus_{p \geq d+1} A_p} \cong A_d = \left(\frac{\mathbb{C}[y_1, y_2, \cdots, y_n]}{\mathfrak{I}_{\mathbb{A}}(Y_A)}\right)_d$$

Por lo tanto:

$$dim_{\mathbb{C}} \frac{M_{0,Y_A}^d}{M_{0,Y_A}^{d+1}} = dim_{\mathbb{C}} \left(\frac{\mathbb{C}[y_1, y_2, \cdots, y_n]}{\mathfrak{I}_{\mathbb{A}}(Y_A)}\right)_d$$

Recordamos que $\mathbb{N}A$ y $\mathbb{Z}A$ denotan respectivamente a las combinaciones lineales enteras positivas y a las combinaciones lineales enteras de elementos de A.

Definimos:

- $\tilde{A} = \{\tilde{v} = (v, 1) : v \in A\} = \{(v_1, 1), (v_2, 1), \cdots, (v_n, 1)\} \subset \mathbb{Z}^{t+1}$ y llamemos S_A al monoide $(0 \in S_A)$ generado por $\tilde{A}$. Además al suponer que $S = \mathbb{Z}A = \mathbb{Z}^t$ entonces $\mathbb{Z}\tilde{A} = \mathbb{Z}^{t+1}$.

- $\tilde{Q} = Conv(\tilde{A}) = Conv\{(v_1, 1), \cdots, (v_n, 1)\}$ y luego $d\tilde{Q}$ es el homotetizado de Q de razón d a "nivel" d.

Los argumentos para esta parte de la demostración del teorema de Kushnirenko son de [**11**].

Probamos, en la proposición 2.67, que tenemos un isomorfismo de $\mathbb{C}$-álgebras entre $\mathbb{C}[S_A]$ y $\mathbb{C}[Y_A]$. Luego el ideal $M_{0,Y_A} \subset \mathbb{C}[Y_A]$ se puede identificar con el ideal $M_0 \subset \mathbb{C}[S_A]$.
Nos interesa en particular mirar el ideal:

$$M_{0,Y_A}^d = \left\langle \overline{y_1}^{i_1} \overline{y_2}^{i_2} \cdots \overline{y_n}^{i_n} : i_1 + \cdots + i_n \geq d \right\rangle \subset \mathbb{C}[Y_A],$$

que se identifica, usando las notaciones de la proposición 2.67 con:

$$M_0^d = \left\langle \chi^{\tilde{v_1}^{i_1}} \chi^{\tilde{v_2}^{i_2}} \cdots \chi^{\tilde{v_n}^{i_n}} : i_1 + \cdots + i_n \geq d \right\rangle \subset \mathbb{C}[S_A].$$

Luego el cociente $\frac{M_0^d}{M_0^{d+1}}$ es igual a $\left\langle \chi^{i_1\tilde{v_1} + \cdots + i_n\tilde{v_n}} : i_1 + \cdots + i_n = d \right\rangle$, quedándonos entonces con sumas de "largo" exactamente d.

Entonces $dim_{\mathbb{C}} \frac{M_0^d}{M_0^{d+1}} = dim_{\mathbb{C}} \left\langle \chi^{i_1\tilde{v_1} + \cdots + i_n\tilde{v_n}} : i_1 + \cdots + i_n = d \right\rangle = \#(d\tilde{Q} \cap \mathbb{N}\tilde{A})$.

Además $d\tilde{Q} \cap \mathbb{N}\tilde{A} = dQ \cap \mathbb{N}A$ via la aplicación biyectiva:

$$\begin{array}{rccc} \psi: & dQ \cap \mathbb{N}A & \longrightarrow & d\tilde{Q} \cap \mathbb{N}\tilde{A} \\ & w = \sum_{i=1}^n (da_i)v_i & \mapsto & \sum_{i=1}^n da_i(v_i, 1) \end{array}, \quad \sum_{i=1}^n a_i = 1, a_i \geq 0 \; \forall \; i = 1, \cdots, n.$$

Por un lado sabemos que para valores de d suficientemente grandes,

$$P_{X_A}(d) = dim_{\mathbb{C}} \left(\frac{\mathbb{C}[x_1, \cdots, x_n]}{\mathfrak{I}_{\mathbb{P}}(X_A)} \right)_d$$

donde P_{X_A} es el polinomio de Hilbert de la variedad X_A.
Luego

$$P_{X_A}(d) = \#(dQ \cap \mathbb{N}A).$$

Por la proposición 2.52, sabemos que P_{X_A} es un polinomio en d de grado t cuyo coeficiente principal es $\frac{deg(X_A)}{t!}$.

Por otro lado, sabemos también que el polinomio de Ehrhart del polítopo Q,
$E_Q(d) = \#(dQ \cap \mathbb{Z}^t)$, es un polinomio en d de grado t cuyo coeficiente principal es $Vol_t(Q)$.

Afirmamos que

existe $m \in \mathbb{Z}_{\geq 0}$, que hallaremos, tal que $\forall \; d \in \mathbb{Z}_{\geq 0}$ suficientemente grande:

$$\#((d-m)Q \cap \mathbb{Z}A) \leq P_{X_A}(d) \leq \#(dQ \cap \mathbb{Z}A)$$

Siendo $\mathbb{Z}A = \mathbb{Z}^t$ entonces $\#(dQ \cap \mathbb{Z}A) = \#(dQ \cap \mathbb{Z}^t) = E_Q(d)$ es el polinomio de Ehrhart del polítopo Q, por lo cual la desigualdad se puede reescribir como:

$$E_Q(d-m) \leq P_{X_A}(d) \leq E_Q(d).$$

Luego dividiendo la desigualdad por d^t y haciendo tender d a infinito obtenemos la igualdad entre los coeficientes principales, por lo cual $Vol_{\mathbb{Z}^t}(Q) = deg(X_A)$.

Probemos entonces la afirmación anterior.

Afirmación: existe $m \in \mathbb{Z}_{\geq 0}$ fijo tal que $\forall \; d \in \mathbb{Z}_{\geq 0}$ suficientemente grande :

$$E_Q(d-m) \leq P_{X_A}(d) \leq E_Q(d).$$

Prueba:

- $P_{X_A}(d) \leq E_Q(d)$ es evidente pues para valores de d suficientemente grandes $P_{X_A}(d) = \#(dQ \cap \mathbb{N}A) \leq \#(dQ \cap \mathbb{Z}A) = E_Q(d)$.
- Un punto cualquiera w de $(d-m)Q \cap \mathbb{Z}^t$ es de la forma $w = \sum_{i=1}^n (d-m)a_i v_i$ con $a_i \geq 0$ y $\sum_{i=1}^n a_i = 1$. Para cada i podemos escribir $(d-m)a_i = m_i + r_i$ con $m_i \in \mathbb{Z}_{\geq 0}$ y $0 \leq r_i < 1$, o sea:

$$w = \sum_{i=1}^n (d-m)v_i = \sum_{i=1}^n m_i v_i + \sum_{i=1}^n r_i v_i$$

donde $\sum_{i=1}^n r_i v_i \in \mathbb{Z}^t \cap \{\sum_{i=1}^n \alpha_i v_i / \ 0 \leq \alpha_i \leq 1\}$. Es decir w se escribe como suma de un punto a coordenada entera de $(d-m)Q$ y de un punto, también con coordenadas enteras, del paralelotopo fundamental asociado a los vectores $\{v_1, \cdots, v_n\}$. Como dentro de este paralelotopo hay una cantidad finita de tales puntos y dado que $\mathbb{Z}A = \mathbb{Z}^t$ tenemos, para cada punto, que

$$\sum_{i=1}^n r_i v_i = \sum_{i=1}^n n_i v_i,$$

donde $n_i \in \mathbb{Z} \ \forall \ i = 1, \cdots, n$.
Luego existe $R \in \mathbb{Z}_{\geq 0}$ tal que $n_i + R \geq 0$ para todo $i = 1, \cdots, n$.

Además $w + R(v_1 + \cdots + v_n) = \sum_{i=1}^n m_i v_i + \sum_{i=1}^n (n_i + R) v_i \in \mathbb{N}A$ pues cada uno de los sumandos $\sum_{i=1}^n m_i v_i$ y $\sum_{i=1}^n (n_i + R) v_i$ pertenecen a $\mathbb{N}A$.

Si tomamos $m = nR$ y consideramos la aplicación:

$$\begin{array}{rccc} \psi : (d-nR)Q \cap \mathbb{Z}^t & \longrightarrow & dQ \cap \mathbb{N}A \\ w & \mapsto & w + R(v_1 + \cdots v_n) \end{array},$$

entonces ψ está correctamente definida: ya vimos que si $w \in (d-nR)Q$ entonces $w + R(v_1 + \cdots v_n)$ es un vector de $\mathbb{N}A$.
Falta probar que pertenece también a dQ. Pero esto es cierto pues:
$w + R(v_1 + \cdots v_n) = \sum_{i=1}^n (d-nR) a_i v_i + R(v_1 + \cdots + v_n)$ y si sumamos los coeficientes

$$\sum_{i=1}^n (d-nR)a_i + nR = (d-nR)\left(\sum_{i=1}^n a_i\right) + nR = (d-nR)1 + nR = d.$$

Luego $w \in (d-nR)Q \in dQ$.

ψ es inyectiva: es claro.

Por lo tanto $E_Q(d-nR) = \#((d-nR)Q \cap \mathbb{Z}^t) \leq \#(dQ \cap \mathbb{N}A) = P_{X_A}(d)$ y se termina de probar la desigualdad.

Queda probada entonces la afirmación y por lo tanto el teorema de Kushnirenko. □

Observemos que la tesis del teorema de Kushnirenko implica que, genéricamente, la cantidad de raíces en $(\mathbb{C}^*)^t$ de un sistema de t ecuaciones polinomiales es finita.

En los ejemplos que siguen veremos que el teorema de Kushnirenko es más fuerte que el conocido teorema de Bézout a la hora de calcular la cantidad de raíces sobre el toro algebraico de un sistema de ecuaciones polinomiales.

Teorema 4.8 (Bézout). *La cantidad de raíces comunes en $(\mathbb{C}^t)$ de t polinomios genéricos $f_1, \cdots, f_t \in \mathbb{C}[x_1, \cdots, x_t]$ de grados respectivos $d_1, \cdots, d_t$ es $d_1 \cdots d_t$.*

Observamos también que, si consideramos una elección genérica de polinomios $f_1, \cdots, f_t \in \mathbb{C}[x_1, \cdots, x_t]$ tal que $NP(f_i) = Q_d, \ d \geq 0, \ \forall \ i = 1, \cdots, t$ - donde Q_d es el polítopo definido en el capítulo 1 - tal que se aplique el teorema de Kushnirenko, la cantidad de raíces comunes a $f_1, \cdots, f_t$ en $(\mathbb{C}^*)^t$ es $Vol_{\mathbb{Z}^t}(Q_d) = d^t Vol_{\mathbb{Z}^t}(Q_1) = d^t$ obteniéndose entonces el número de raíces dado por el Teorema de Bézout.

3. Ejemplos

Veamos ahora algunos ejemplos que ilustran el teorema de Kushnirenko. Primero veamos dos aplicaciones geométricas:

1. Supongamos que tenemos dos rectas en el plano complejo. Estudiar la intersección de dos rectas en el plano equivale a estudiar las raíces comunes de los polinomios:
   $\begin{cases} f(x,y) = ax + by + c = 0 \\ g(x,y) = a'x + b'y + c' = 0 \end{cases}$
   donde a, b y $c \in \mathbb{C}$.
   f y g tienen el mismo polítopo de Newton: el triángulo $T = Conv\{(0,0),(1,0),(0,1)\}$, como lo muestra la figura 1. Por lo que, al aplicar el teorema de Kushnirenko y al ser

$Vol_{\mathbb{Z}^2}(T) = 1$, genéricamente f y g tienen una sola raíz común en $(\mathbb{C}^*)^2$ (¡lo que era de esperar en el plano real!).

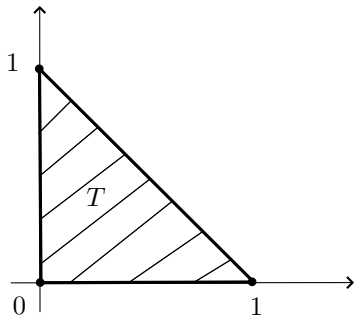


Figura 1.

2. $\begin{cases} f(x,y) = axy + bx + c = 0 \\ g(x,y) = a'xy + b'x + c' = 0 \end{cases}$ f y g tienen el mismo polítopo de Newton: el triángulo $T = Conv\{(0,0),(1,0),(1,1)\}$, como lo muestra la figura 2. Por lo que, al aplicar el teorema de Kushnirenko y al ser $Vol_{\mathbb{Z}^2}(T) = 1$, genéricamente f y g tienen una raíz común en $(\mathbb{C}^*)^2$. Observar que la cota dada por el teorema de Bézout es 4.

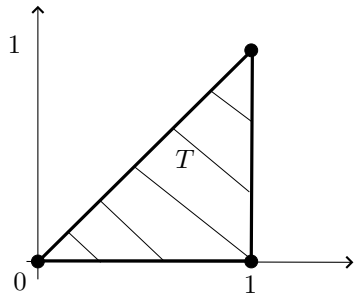


Figura 2.

3. $\begin{cases} f(x,y) = ax^2y + bxy^2 + cx + dy = 0 \\ g(x,y) = a'x^2y + b'xy^2 + c'x + d'y = 0 \end{cases}$
Tenemos que $NP(f) = NP(g) = Q = Conv\{(2,1),(1,2),(1,0),(0,1)\}$ como lo muestra la figura 4. Si aplicamos el teorema de Kushnirenko, genéricamente, estos dos polinomios tienen $Vol_{\mathbb{Z}^2}(Q) = 4$ raíces comunes. Observar que la cota dada por el teorema de Bézout es 5.
4. Observemos que los sistemas:
$$\begin{cases} f_1(x,y) = ax^2y^2 + bx^3y^2 + cx^4y + d = 0 \\ g_1(x,y) = a'x^2y^2 + b'x^3y^2 + c'x^4y + d' = 0 \end{cases}$$
y
$$\begin{cases} f_2(x,y) = axy + bx^2y + cx^2y^2 + dx^3y + ex^3y^2 + fx^4y + g = 0 \\ g_2(x,y) = a'xy + b'x^2y + c'x^2y^2 + d'x^3y + e'x^3y^2 + f'x^4y + g' = 0 \end{cases}$$
tienen, genéricamente, la misma cantidad de soluciones ya que $NP(f_1) = NP(g_1) = NP(f_2) = NP(g_2) = Q$ como lo muestra la figura 4.

Esa cantidad de soluciones es $Vol_{\mathbb{Z}^2}(Q) = 9$.

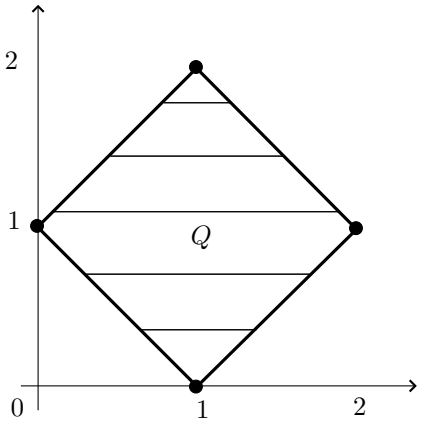


Figura 3.

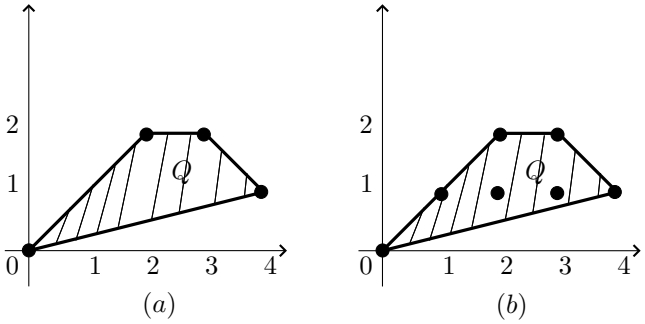


Figura 4.

Una vez más: la cantidad de soluciones de un sistema de ecuaciones polinomiales en las condiciones del teorema de Kushnirenko depende únicamente de los vértices de los polítopos de Newton que se consideran.

5. Consideramos el sistema $\begin{cases} f(x,y) = a + bx + cy + dxy = 0 \\ g(x,y) = a + b'x + c'y + d'xy = 0 \end{cases}$.

Como lo muestra la figura 5, $Q = Conv\{(0,0),(1,0),(0,1),(1,1)\}$. Afirmamos, aplicando el teorema de Kushnirenko, que genéricamente la cantidad de raíces comunes a f y a g es $Vol_{\mathbb{Z}^2}(Q) = 2$ mientras la cota dada por el teorema de Bézout es 4.

Basémosnos en los argumentos utilizados en la demostración del teorema.
Como $A = \{(0,0),(1,0),(0,1),(1,1)\}$, La variedad X_A es:

$$X_A = \overline{\{[X^{(0,0)} : X^{(1,0)} : X^{(0,1)} : X^{(1,1)}] :\ X \in (\mathbb{C}^*)^t\}} \subset \mathbb{P}^{4-1} = \mathbb{P}^3,$$

su dimensión en $\mathbb{P}^3$ es 2, que el subespacio $\mathcal{L}$ correspondiente al sistema tenga dimensión 2 equivale a que los vectores (a,b,c,d) y (a',b',c',d') sean linealmente independientes.

Se prueba en [**6**], proposición 1.9, que las órbitas por la acción del toro en X_A están en biyección con las caras del polítopo $Q = Conv(A)$.
Esta biyección se establece a partir de la construcción siguiente:
si P es una cara de Q definimos el vector $e_P = [e_{P,0} : e_{P,1} : \cdots : e_{P,n}] \in \mathbb{P}^n$ tal que:

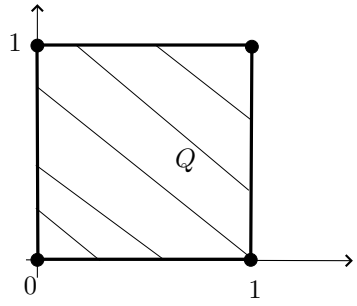


Figura 5.

$$e_{P,j} = \begin{cases} 1, & v_j \in P \\ 0, & v_j \notin P \end{cases}$$

Por lo cual la asociación es: $P \mapsto \chi(P) := (\mathbb{C}^*)^t.\ e_P \in \mathbb{P}^{n-1}$.

Luego $X_A = \biguplus_{P \in F(Q)} \chi(P)$ donde $F(Q)$ es el conjunto formado por las caras de Q.

En la demostración del teorema de Kushnirenko pedíamos que $\mathcal{L}$ evite las caras de Q de dimensión menor que 2, o sea considerando la biyección anterior estas órbitas son: $O_{[1:0:0:0]}, O_{[0:1:0:0]}, O_{[0:0:1:0]}, O_{[0:0:0:1]}, O_{[1:1:0:0]}, O_{[0:0:1:1]}, O_{[1:0:1:0]}$ y $O_{[0:1:0:1]}$.
Por ejemplo:

Para que $\mathcal{L} \cap O_{[1:1:0:0]} = \emptyset$ es suficiente que: $\begin{cases} a + bx = 0 \\ a' + b'x = 0 \end{cases}$ no tenga solución no trivial, es decir que $det \begin{pmatrix} a & b \\ a' & b' \end{pmatrix} \neq 0$,

La condición

$$det \left[\begin{pmatrix} a & b \\ a' & b' \end{pmatrix} \begin{pmatrix} a & c \\ a' & c' \end{pmatrix} \begin{pmatrix} c & d \\ c' & d' \end{pmatrix} \begin{pmatrix} b & d \\ b' & d' \end{pmatrix} \right] \neq 0.$$

asegura que $\mathcal{L}$ evita las órbitas de X_A de dimensión menor que 2.

CAPíTULO 5

Teorema de Bernstein

En este último capítulo, presentaremos sin demostración, una extensión del teorema de Kushnirenko al caso en que los polinomios de Laurent que conforman el sistema tengan eventualmente distintos conjuntos prefijados de exponentes.

1. El volumen mixto

DEFINICIÓN 5.1. Sean P y Q dos polítopos en $\mathbb{R}^n$ y sea $\lambda \in \mathbb{R}$. Definimos los conjuntos:

$$P + Q = \{p + q : p \in P, q \in Q\} \text{ y } \lambda P = \{\lambda p : p \in P\}.$$

- La suma de dos polítopos P y Q lleva el nombre de *suma de Minkowski* de P y de Q.
- Como ya lo hemos mencionado en capítulos anteriores, el conjunto λP es el homotetizado de razón λ del polítopo P. Es claro que λP es un polítopo.

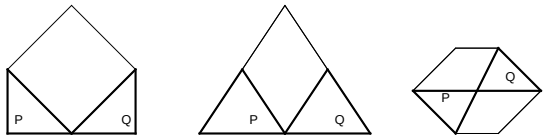


Figura 1. Suma de Minkowski.

PROPOSICIÓN 5.2.
1. *Si P y Q son polítopos entonces $P + Q$ es un polítopo. Más aún si P y Q son polítopos racionales entonces $P + Q$ es un polítopo racional.*
2. *Si P es un polítopo racional y $d \in \mathbb{Z}_{\geq 0}$ entonces dP es un polítopo racional.*
3. *Si $P_1, P_2, \cdots, P_r$ son polítopos y $\lambda_1, \lambda_2, \cdots, \lambda_r \in \mathbb{R}$ entonces $\lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_r P_r$ es un polítopo.*

DEMOSTRACIÓN. Ver [**2**], capítulo 7, páginas 317-318 y [**3**], capítulo 4, lema 1.2 página 103-104 y teorema 1.5 página 105. □

TEOREMA 5.3. *Sean $P_1, P_2, \cdots, P_r$ polítopos n-dimensionales en $\mathbb{R}^n$ y $\lambda_i \geq 0, \forall i = 1, \cdots, r$. Entonces $Vol_n(\lambda_1 P_1 + \cdots + \lambda_r P_r)$ es un polinomio homogéneo de grado n en $\lambda_1, \cdots, \lambda_r$ donde $Vol_n(.)$ denota al volumen n-dimensional en $\mathbb{R}^n$.*

DEMOSTRACIÓN. Ver [**3**], capítulo 4, página 116. □

Volviendo a los polítopos de Newton asociados a polinomios de Laurent, la proposición siguiente nos da una interesante relación entre la suma de Minkowski de dos polítopos de Newton y el polítopo de Newton del producto de los polinomios:

PROPOSICIÓN 5.4. *Si $f_1, f_2 \in \mathbb{C}[x_1^{\pm 1}, x_2^{\pm 1}, \cdots, x_n^{\pm 1}]$ con $P = NP(f_1)$ y $Q = NP(f_2)$ entonces $NP(f_1 \cdot f_2) = NP(f_1) + NP(f_2)$ siendo $+$ la suma de Minkowski en $\mathbb{R}^n$.*

DEMOSTRACIÓN. ($\subset$): Esta inclusión es clara ya que cada exponente de $f_1 \cdot f_2$ aparece en $NP(f_1) + NP(f_2)$.

($\supset$): Consideramos un vértice de $NP(f_1) + NP(f_2)$. Este vértice no está en el interior relativo de ninguna arista de $NP(f_1) + NP(f_2)$.

Sea $h = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ el monomio correspondiente a dicho vértice. El grado de h es $gr(h) = (\alpha_1, \alpha_2, \cdots, \alpha_n) \in \mathbb{Z}^n$.

AFIRMACIÓN: El monomio h se escribe de manera única como $m \cdot n$ con m monomio en f_1 y n monomio en f_2 (los monomios son mónicos).

PRUEBA: : Suponemos que h se escribe como mn y $m'n'$ con m ym' monomios en f_1 tal que $m \neq m'$ y n y n' monomios en f_2 tal que $n \neq n'$. Entonces $gr(h) = gr(m) + gr(n) = gr(m') + gr(n')$. Luego

$$2gr(h) = gr(m) + gr(n) + gr(m') + gr(n')$$

y

$$gr(h) = \frac{1}{2}\left(gr(m) + gr(n')\right) + \frac{1}{2}\left(gr(m) + gr(n)\right).$$

Por lo que $gr(h)$ es el punto medio del segmento que une $gr(m) + gr(n')$ y $gr(m) + gr(n)$. Como cada uno de los puntos $gr(m) + gr(n)$ y $gr(m) + gr(n)$ está en $NP(f_1) + NP(f_2)$, llegamos a un absurdo: $gr(h)$ no sería un vértice de $NP(f_1) + NP(f_2)$.

Por lo que h se escribe de manera única como producto de un monomio en f_1 y de un monomio en f_2. Consecuentemente solamente uno de los términos de $f_1 \cdot f_2$ corresponde al coeficiente de h y ese término es el producto de coeficientes no nulos. Por lo que el coeficiente de h es no nulo. Por lo tanto $gr(h) \in NP(f_1 \cdot f_2)$, luego $NP(f_1) + NP(f_2) \subset NP(f_1 \cdot f_2)$. □

El corolario siguiente es una consecuencia inmediata de la proposición anterior:

COROLARIO 5.5. *Si $f \in \mathbb{C}[x_1^{\pm 1}, x_2^{\pm 1}, \cdots, x_n^{\pm 1}]$ y $q \in \mathbb{Z}_{\geq 0}$ entonces $NP(f^q) = qNP(f)$.*

EJEMPLO 5.6. Sean $f_1(x, y) = ax^3y^2 + bx + cy^2 + d$ y $f_2(x, y) = exy^4 + fx^3 + gy$ con polítopos de Newton respectivos $P_1 = NP(f_1)$ y $P_2 = NP(f_2)$ y donde $a, b, c, d, e, f, g \in \mathbb{C}^*$. Luego
$P_1 = NP(f_1) = conv\{(3, 2), (1, 0), (0, 2), (0, 0)\}$ y $P_2 = NP(f_2) = Conv\{(1, 4), (3, 0), (0, 1)\}$

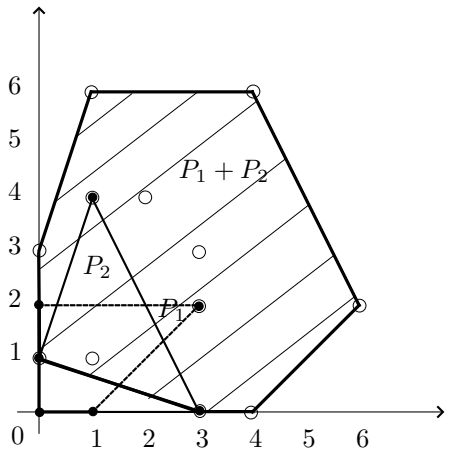


Figura 2.

$f_1(x, y) \cdot f_2(x, y) = (ax^3y^2 + bx + cy^2 + d) \cdot (exy^4 + fx^3 + gy) =$
$= aex^4y^6 + afx^6y^2 + agx^3y^3 + bex^2y^4 + bfx^4 + bgxy + cexy^6 + cfx^3y^2 + cgy^3 + dexy^4 + dfx^3 + dgy.$

En la figura 2 hemos representados los polítopos P_1, P_2 y $P_1 + P_2$ así como los vértices de P_1 y P_2 y los grados de los monomios del producto.

PROPOSICIÓN 5.7.
1. $MV_n(P, P, \cdots, P) = n!Vol_n(P) = Vol_{\mathbb{Z}^n}(P)$.
2. $MV_n(P_1, \cdots, P_i, \cdots, P_n) = MV_n(P_1, \cdots, P_i + \alpha, \cdots, P_n), \forall\ \alpha \in \mathbb{R}^n$.
3. $MV_n(P_1, \cdots, P_n) = MV_n(P_{\pi(1)}, \cdots, P_{\pi(n)}), \forall \pi$ *permutación de* $\{1, \cdots, n\}$.
4. $MV_n(\alpha P_1, \cdots, P_n) = \alpha MV_n(P_1, \cdots, P_n)$
5. $MV_n(P + Q, P_2, \cdots, P_n) = MV_n(P, P_2, \cdots, P_n) + MV_n(Q, P_2, \cdots, P_n)$
6. *Sean* $P_1, P_2, \cdots, P_n$ *polítopos en* $\mathbb{R}^n$. *Entonces:*

$$MV_n(P_1, \cdots, P_n) = \sum_{k=1}^{n} (-1)^{n-k} \sum_{I \subset 1, \cdots, n, |I|=k} Vol_n(\sum_{i \in I} P_i).$$

En el caso que P *y* Q *sean dos polítopos en* $\mathbb{R}^2$, *el volumen mixto de* P *y de* Q *se calcula como:*

$$MV_2(P, Q) = \text{área}(P + Q) - \text{área}(P) - \text{área}(Q)$$

ya que $Vol_2(.)$ *es el área en* $\mathbb{R}^2$.

DEMOSTRACIÓN. Ver [**3**] capítulo 4, lema 3.4 y 3.5 página 117, teorema 3.7 página 117 y lema 3.6 página 118. □

Mencionamos que en el capítulo 7 de [**2**] se expone una manera más eficaz y explícita de calcular el volumen mixto de polítopos.

2. El teorema de Bernstein

TEOREMA 5.8 (Teorema de Bernstein (1975)). *Sean* $A_1, \cdots, A_n$ *subconjuntos finitos de vectores de* $\mathbb{Z}^n$ *tales que* $A_1 \cup \cdots \cup A_n$ *tiene rango n. Denotamos por* $P_1, \cdots, P_n$ *a las envolventes convexas de* $A_1, \cdots, A_n$ *respectivamente.*
Entonces para una elección genérica de polinomios de Laurent $f_1, \cdots, f_n$ *en* $\mathbb{C}[x_1^{\pm 1}, \cdots, x_n^{\pm 1}]$ *tales que* $f_i \in C^{A_i}, \forall\ i = 1, \cdots, n$ *se verifica que la cantidad de raíces comunes (contadas con sus multiplicidades) de* $f_1, \cdots, f_n$ *coincide con* $MV_n(P_1, \cdots, P_n)$.

DEMOSTRACIÓN. : Ver [**1**] para una demostración que usa series de Puiseux o [**5**], páginas 122-123, para una demostración que usa variedades tóricas y divisores. □

Algunas observaciones:

1. Si todos los polinomios tienen el mismo polítopo de Newton Q, dado que $MV_n(Q, \cdots, Q) = n!Vol_n(Q) = Vol_{\mathbb{Z}^n}(Q)$, es decir el teorema de Bernstein generaliza el teorema de Kushnirenko.
2. Al teorema de Bernstein muchas veces se le llama "teorema BKK" en reconocimiento a Bernstein, Kushnirenko y Khovanskii quienes escribieron en la segunda mitad de la década de los 70 varios artículos acerca de los sistemas de ecuaciones polinomiales con un número finito de soluciones.
3. Aplicando el teorema de Bernstein y la definición del volumen mixto, podemos afirmar entonces que si f y g son dos polinomios de Laurent en dos variables genéricos, la cantidad de raíces comunes a f y a g sobre $(\mathbb{C}^*)^2$ es igual a:

$$\text{área}(NP(f) + NP(g)) - \text{área}(NP(f)) - \text{área}(NP(g)).$$

EJEMPLO 5.9. Consideramos el sistema de ecuaciones polinomiales dados por los polinomios del ejemplo 5.6 para una elección genérica de los coeficientes.

$$\begin{cases} f_1(x, y) = ax^3y^2 + bx + cy^2 + d = 0 \\ f_2(x, y) = exy^4 + fx^3 + gy = 0 \end{cases}$$

Entonces $NP(f_1) = P_1$ y $NP(f_2) = P_2$.

Es fácil calcular que $MV(P_1, P_2) = 18$, por lo cual genéricamente, el sistema anterior tiene 18 soluciones en $(\mathbb{C}^*)^2$. Observamos que la cota dada por el teorema de Bézout es 25.

Observamos que si consideramos $f_1, \cdots, f_t$ de grado respectivos $d_1, \cdots, d_t$, la cota dada por el Teorema de Bézout coincide con la que da el Teorema de Bernstein para el caso en que cada polinomio f_i tenga como polítopo de Newton asociada a Q_{d_i} definido en el capítulo 1, puesto que:

$$MV_n(Q_{d_1}, Q_{d_2}, \cdots, Q_{d_n}) = MV_n(d_1Q_1, d_2Q_2, \cdots, d_nQ_n) = \prod_{i=1}^{n} d_i MV_n(Q_1, \cdots, Q_1) = \prod_{i=1}^{n} d_i$$

ya que $Vol_n(Q_1) = \frac{1}{n!}$.

Apéndice

En este apéndice justificaremos algunos resultados específicos de álgebra lineal y álgebra combinatoria que usamos en el trabajo. El primero es un procedimiento para triangularizar una matriz con coeficientes enteros sin alterar su determinante. Este algoritmo que presentamos, a diferencia del de Hermite y del de Schmidt, no asegura la unicidad de la matriz triangular. El lector interesado en ampliar podrá consultar, por ejemplo, [**10**]. A continuación estudiamos un tipo muy particular de sistema polinomial. Luego justificamos que el índice del $\mathbb{Z}$-módulo generado por un subconjunto finito de vectores de rango n en $\mathbb{Z}^n$ es igual al valor absoluto del determinante de la matriz cuyas columnas son los vectores de una base de ese módulo. Finalmente probamos la conocida fórmula que relaciona el volumen de un paralelotopo con la matriz de los vectores que lo definen.

1. **Triangularización de matrices.**

Veamos primero el caso de una matriz 2×2. Empezamos recordando el siguiente resultado.

Propiedad: Sean m y n dos números enteros. Entonces:
m y n son primos entre sí $\Leftrightarrow$ existen α y $\beta \in \mathbb{Z}$ tales que $\alpha m + \beta n = 1$.

Nuestro propósito es transformar la matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ donde a, b, c y d son números enteros en una matriz con el mismo determinante, triangular y con coeficientes enteros.

O sea vamos a ver que podemos multiplicar la matriz A por una matriz $Q = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ de determinante 1 tal que $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \times \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} r_1 & r_2 \\ 0 & r_3 \end{pmatrix}$ siendo r_1, r_2 y r_3 números enteros.

Esto equivale a buscar soluciones del sistema: $\begin{cases} \gamma a + \delta c = 0 \\ \alpha\delta - \gamma\beta = 1 \end{cases}$

Si elegimos $\gamma = c$ y $\delta = -a$, con el fin de producir un 0 en la entrada que se encuentra en la intersección de la segunda fila y primer columna, se verifica la primera ecuación y la segunda ecuación se transforma en $\alpha(-a) - \beta(c) = 1$ o sea $\alpha' a + \beta' c = 1$. Si a y c son primos entre sí entonces por la propiedad anterior sabemos que esta ecuación tiene solución (no necesariamente única).

Si a y c no son primos entre sí, sabemos que existen a' y c' tales que $a = mcd(a, c)a'$ y $c = mcd(a, b)c'$ con a' y c' primos entre sí.

Resumiendo, si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con a, b, c y d enteros:

- Si a y c son primos entre sí entonces multiplicamos A por $Q_1 = \begin{pmatrix} \alpha & \beta \\ c & -a \end{pmatrix}$ o sea

$$Q_1 A = \begin{pmatrix} \alpha & \beta \\ c & -a \end{pmatrix} \times \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} r_1 & r_2 \\ 0 & r_3 \end{pmatrix} = R$$

donde α y β son soluciones enteras de la ecuación $(-\alpha)a + (-\beta)c = 1$.

- Si a y c no son primos entre sí multiplicamos A por $Q_2 = \begin{pmatrix} \alpha & \beta \\ \frac{c}{mcd(a,c)} & \frac{-a}{mcd(a,c)} \end{pmatrix}$ o sea

$$Q_2 A = \begin{pmatrix} \alpha & \beta \\ \frac{c}{mcd(a,c)} & \frac{-a}{mcd(a,c)} \end{pmatrix} \times \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} r_1 & r_2 \\ 0 & r_3 \end{pmatrix} = R$$

donde α y β son soluciones enteras de la ecuación $(-\alpha)\frac{c}{mcd(a,c)} + (-\beta)\frac{a}{mcd(a,c)} = 1$. Observar que la matriz R obtenida sigue siendo una matriz con entradas enteras.

Ejemplo 1: Sea $A = \begin{pmatrix} 2 & 4 \\ 5 & -3 \end{pmatrix}$. Entonces una matriz Q que nos sirve es $\begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix}$ ya que acá 2 y 5 son primos entre sí y $(-2)(2)+(1)(5) = 1$ y $\begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix} \times \begin{pmatrix} 2 & 4 \\ 5 & -3 \end{pmatrix} = \begin{pmatrix} -1 & 11 \\ 0 & 26 \end{pmatrix}$. Observamos que la matriz resultante y la matriz original tienen el mismo determinante (-26).

Ejemplo 2: Sea $A = \begin{pmatrix} 12 & 2 \\ 4 & 1 \end{pmatrix}$. Entonces una matriz Q que nos sirve es $\begin{pmatrix} -1 & 2 \\ 1 & -3 \end{pmatrix}$ ya que acá 12 y 4 no son primos entre sí y $(-1)(3) + (2)(1) = 1$ y $\begin{pmatrix} -1 & 2 \\ 1 & -3 \end{pmatrix} \times \begin{pmatrix} 12 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$. Observamos que la matriz resultante y la matriz original tienen el mismo determinante (-4).

Proposición A: Si $A \in \mathcal{M}_{n\times n}(\mathbb{Z})$ entonces existe una matriz $Q \in \mathcal{M}_{n\times n}(\mathbb{Z})$ invertible y con determinante igual a 1, tal que la matriz $R = QA$ es triangular superior.

Demostración. Consideramos $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-11} & a_{n-12} & a_{n-13} & \cdots & a_{n-1n} \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$ con $a_{ij} \in \mathbb{Z}\ \forall\ i, j = 1, \cdots, n$.

Probaremos que podemos multiplicar a la izquierda por una matriz $Q \in SL_n(\mathbb{Z})$ y obtener de esa manera una matriz triangular superior

$$R = QA = \begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & r_{n-1n-1} & r_{n-1n} \\ 0 & 0 & 0 & \cdots & r_{nn} \end{pmatrix}$$

En un primer momento vamos a querer producir entradas nulas en la primer columna, salvo en la primer entrada, o sea pasar de la matriz A a una matriz A' tal que :

$$A' = \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} & \cdots & a'_{1n} \\ 0 & a'_{22} & a'_{23} & \cdots & a'_{2n} \\ 0 & a'_{32} & a'_{33} & \cdots & a'_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a'_{n-12} & a'_{n-13} & \cdots & a'_{n-1n} \\ 0 & a'_{n2} & a'_{n3} & \cdots & a'_{nn} \end{pmatrix}.$$

La idea consiste en aprovechar el trabajo hecho en el caso 2×2. Multiplicando A por la matriz $Q_1 = \begin{pmatrix} \alpha & \beta & 0 & \cdots & 0 \\ \gamma & \delta & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} \alpha & \beta & \\ \gamma & \delta & 0 \\ 0 & & I_{n-2} \end{pmatrix}$ donde la matriz $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ es una matriz que transforma la matriz $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ en una matriz triangular superior $\begin{pmatrix} a'_{11} & a'_{12} \\ 0 & a'_{22} \end{pmatrix}$ con igual determinante que $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

Obtenemos entonces la matriz $Q_1A = \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} & \cdots & a'_{1n} \\ 0 & a'_{22} & a'_{23} & \cdots & a'_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-11} & a_{n-12} & a_{n-13} & \cdots & a_{n-1n} \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$ con igual determinante que A pues Q_1 tiene determinante igual a 1.

Para producir otro cero en la tercera fila de la primera columna, sin cambiar el cero obtenido en la segunda fila, multiplicamos por la matriz $Q_2^1 = \begin{pmatrix} \alpha & 0 & \beta & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \gamma & 0 & \delta & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} =$

$\begin{pmatrix} \alpha & 0 & \beta & \\ 0 & 1 & 0 & 0 \\ \gamma & 0 & \delta & \\ & 0 & & I_{n-3} \end{pmatrix}$ donde la matriz $\begin{pmatrix} \alpha & 0 & \beta \\ 0 & 1 & 0 \\ \gamma & 0 & \delta \end{pmatrix}$ es la que transforma la matriz $\begin{pmatrix} a'_{11} & a'_{12} & a'_{13} \\ 0 & a'_{22} & a'_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ en la matriz $\begin{pmatrix} a''_{11} & a''_{12} & a''_{13} \\ 0 & a''_{22} & a''_{23} \\ 0 & a''_{32} & a''_{33} \end{pmatrix}$ tal que:

$$\begin{pmatrix} \alpha & 0 & \beta \\ 0 & 1 & 0 \\ \gamma & 0 & \delta \end{pmatrix} \times \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} \\ 0 & a'_{22} & a'_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a''_{11} & a''_{12} & a''_{13} \\ 0 & a''_{22} & a''_{23} \\ 0 & a''_{32} & a''_{33} \end{pmatrix}.$$

De la misma manera, vamos a querer producir entradas nulas en $a_{41}, a_{51}, \cdots, a_{n1}$. Luego multiplicamos sucesivamente por matrices del tipo:

$\begin{pmatrix} \alpha & 0 & 0 & \beta & \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & 0 \\ \gamma & 0 & 0 & \delta & \\ & & 0 & & I_{n-4} \end{pmatrix}, \cdots, \begin{pmatrix} \alpha & 0 & \cdots & 0 & \beta \\ 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & & 1 & 0 \\ \gamma & 0 & \cdots & 0 & \delta \end{pmatrix}$ obteniéndose así una matriz A' cuya primer columna tiene todos sus elementos nulos salvo quizás, la primer entrada, siendo:

$$A' = \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} & \cdots & a'_{1n} \\ 0 & a'_{22} & a'_{23} & \cdots & a'_{2n} \\ 0 & a'_{32} & a'_{33} & \cdots & a'_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a'_{n-12} & a'_{n-13} & \cdots & a'_{n-1n} \\ 0 & a'_{n2} & a'_{n3} & \cdots & a'_{nn} \end{pmatrix}.$$

Además A' tiene igual determinante que A.

Ahora el paso siguiente consiste en producir ceros en la segunda columna, debajo de la entrada segunda fila. Procederemos de igual manera multiplicando sucesivamente por las matrices:

$$Q_2^2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \alpha & \beta & 0 & \cdots & 0 \\ 0 & \delta & \gamma & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \ddots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, Q_3^2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \alpha & 0 & \beta & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & \delta & 0 & \gamma & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \cdots.$$

Llegamos entonces a una matriz A'' con entradas nulas en la primera columna debajo de la primera fila y en la segunda columna debajo de la segunda fila.
Reiteramos así el procedimiento hasta obtener una matriz R de la forma:

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} & r_{14} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & r_{24} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & r_{34} & \cdots & r_{3n} \\ 0 & 0 & 0 & r_{44} & \cdots & r_{4n} \\ \vdots & \vdots & \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{n-1n-1} & r_{n-1n} \\ 0 & 0 & 0 & 0 & \cdots & r_{nn} \end{pmatrix}.$$

Luego $Q = (Q_1^1 Q_2^1 \cdots Q_{n-1}^1)(Q_1^2 Q_2^2 \cdots Q_{n-2}^2) \cdots (Q_1^{n-2} Q_2^{n-2})(Q_1^{n-1})$ y $det(Q) = 1$ pues $det(Q_i^j) = 1$, $\forall j = 1, \cdots, i = j, \cdots, n$ y $det(R) = det(A)$. □

2. Resolución de un sistema exponencial.

Proposición B: El número de soluciones del sistema

$$\begin{cases} x_1^{\lambda_{11}} x_2^{\lambda_{12}} \cdots x_t^{\lambda_{1t}} = y_1 \\ x_1^{\lambda_{21}} x_2^{\lambda_{22}} \cdots x_t^{\lambda_{2t}} = y_2 \\ x_1^{\lambda_{31}} x_2^{\lambda_{32}} \cdots x_t^{\lambda_{3t}} = y_3 \\ \qquad\qquad\qquad\qquad \vdots \\ \qquad\qquad\qquad\qquad \vdots \\ x_1^{\lambda_{t1}} x_2^{\lambda_{t2}} \cdots x_t^{\lambda_{tt}} = y_t \end{cases} \tag{2.1}$$

donde $\lambda_{ij} \in \mathbb{Z}, \forall\, i, j,\ y_j \in \mathbb{C}^* \forall\, j$ es $|det((\lambda_{ij})_{ij}|$.

Demostración. Consideramos ahora la matriz de exponentes del sistema (2.1) es decir la matriz

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \cdots & \lambda_{1t} \\ \lambda_{21} & \lambda_{22} & \cdots & \cdots & \lambda_{2t} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ \lambda_{t1} & \lambda_{t2} & \cdots & \cdots & \lambda_{tt} \end{pmatrix}.$$

Como A tiene entradas enteras, sabemos que existe una matriz $Q \in SL_n(\mathbb{Z})$ tal que:

$$Q\Lambda = T = \begin{pmatrix} r_{11} & r_{12} & \cdots & \cdots & r_{1t} \\ 0 & r_{22} & \cdots & \cdots & r_{2t} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & r_{tt} \end{pmatrix},$$

donde T es una matriz triangular superior, con entradas enteras y con el mismo determinante que Λ.

Por lo que el sistema (2.1) pasa a ser ahora:

$$\left\{ \begin{array}{r} x_1^{r_{11}} x_2^{r_{12}} \cdots x_t^{r_{1t}} = z_1 \\ x_2^{r_{22}} \cdots x_t^{r_{2t}} = z_2 \\ \vdots \\ \vdots \\ x_t^{r_{tt}} = z_t \end{array} \right. \tag{2.2}$$

Una vez más explicitemos el cálculo que permite llegar al sistema anterior, y en particular de donde provienen los z_j's.

Suponemos que tenemos: $\left\{ \begin{array}{l} x_1^{\lambda_{11}} x_2^{\lambda_{12}} = y_1 \\ x_1^{\lambda_{21}} x_2^{\lambda_{22}} = y_2 \end{array} \right.$

Si $Q = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}$ una matriz en $SL_n(\mathbb{Z})$ tal que

$$Q \times \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}$$

entonces:

$$\left\{ \begin{array}{l} x_1^{\lambda_{11}} x_2^{\lambda_{12}} = y_1 \\ x_1^{\lambda_{21}} x_2^{\lambda_{22}} = y_2 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} (x_1^{\lambda_{11}} x_2^{\lambda_{12}})^{u_{11}} (x_1^{\lambda_{21}} x_2^{\lambda_{22}})^{u_{12}} = y_1^{u_{11}} y_2^{u_{12}} \\ (x_1^{\lambda_{11}} x_2^{\lambda_{12}})^{u_{21}} (x_1^{\lambda_{21}} x_2^{\lambda_{22}})^{u_{22}} = y_1^{u_{21}} y_2^{u_{22}} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{r} x_1^{r_1} x_2^{r_3} = y_1^{u_{11}} y_2^{u_{12}} = z_1 \\ x_2^{r_2} = y_1^{u_{21}} y_2^{u_{22}} = z_2 \end{array} \right.$$

Por lo que fijados y_1 e y_2 el último sistema tiene $|r_2||r_1|$ soluciones.

Generalizando, el sistema (2.2) y por lo tanto el sistema (2.1) tienen:

$$|r_{tt}||r_{t-1t-1}| \cdots |r_{22}||r_{11}| = m = |det(R)| = |det(\Lambda)|$$

soluciones. □

3. Índice y determinante.

Recordamos que si L es un retículo de rango n de $\mathbb{Z}^n$ entonces:

a) todo vector $v \in \mathbb{Z}^n$ es congruente módulo L con un único vector del paralelotopo fundamental asociado a una $\mathbb{Z}$-base de L,

b) $[\mathbb{Z}^n : L]$ es la cantidad de puntos con coordenadas enteras en el paralelotopo fundamental asociado a una $\mathbb{Z}$-base de L.

Proposición C: Sea $A \subset \mathbb{Z}^n$ y L el $\mathbb{Z}-$módulo generado por A con base $\{w_1, w_2, \cdots, w_n\}$. Entonces:

$$[\mathbb{Z}^n : L] = |det(B)|$$

donde B es la matriz cuyas columnas son $w_1, w_2, \cdots, w_n$.

Demostración. Sabemos, por la parte anterior, que existe una matriz $Q \in SL_n(\mathbb{Z})$, por lo tanto invertible y de determinante 1, tal que $QB = T$ donde $T = [u_1|u_2|\cdots|u_n]$ es una matriz triangular superior. Por lo tanto B y T tienen igual determinante. Sabemos también que:

$$[\mathbb{Z}^n : L] = \#\left(\{\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n :\ 0 \leq \alpha_i < 1\ \forall i = 1, \cdots, n\} \bigcap \mathbb{Z}^n\right).$$

Sea ahora $Q \in \mathcal{M}_{n\times n}(\mathbb{Z})$ una matriz invertible que verifica $QB = T$.

Entonces $Qv = Q(\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n) = \alpha_1 Q w_1 + \alpha_2 Q w_2 + \cdots + \alpha_n Q w_n$ $= \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n$ donde $\{u_1, u_2, \cdots, u_n\}$ es otra base de L.
Por otro lado si $u = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n$ entonces de la misma manera:
$Q^{-1}u = \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_n w_n$.

Luego tenemos:

$$[\mathbb{Z}^n : L] = \#\left(\{\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n :\ 0 \leq \alpha_i < 1\ \forall i = 1, \cdots, n\} \bigcap \mathbb{Z}^n\right)$$

$$= \#\left(\{\beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n :\ 0 \leq \beta_i < 1\ \forall i = 1, \cdots, n\} \bigcap \mathbb{Z}^n\right) = \#\left(P_U \cap \mathbb{Z}^n\right).$$

donde P_U es el paralelotopo fundamental asociado a $\{u_1, \cdots, u_n\}$.

Entonces es suficiente probar que $\#\left(P_U \cap \mathbb{Z}^n\right) = |det(T)|$ ya que $det(B) = det(T)$.
Haremos esta demostración por inducción completa sobre n.
Para $n = 1$, el resultado es evidente.
Supongamos que $T = \begin{pmatrix} u_{11} & \cdots & \cdots & u_{1n} \\ 0 & u_{22} & & u_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{nn} \end{pmatrix}$.
Para cada $z \in [-u_{nn} + 1, 0] \cap \mathbb{Z}$ o $z \in [0, u_{nn} - 1] \cap \mathbb{Z}$, según corresponda, consideramos un hiperplano $H_z = \{(x_1, \cdots, x_{n-1}, z) \in \mathbb{R}^n /\ x_1, \cdots, x_{n-1} \in \mathbb{R}\}$. Luego:

$$\#\left(P_U \cap \mathbb{Z}^n\right) = \sum_z \#\left(P_U \cap H_z \cap \mathbb{Z}^n\right) = |u_{nn}| \#\left(P_U \cap H_0 \cap \mathbb{Z}^n\right).$$

La hipótesis inductiva nos asegura que

$$\#\left(P_U \cap H_0 \cap \mathbb{Z}^n\right) = \left| det \begin{pmatrix} u_{11} & \cdots & \cdots & u_{1n-1} \\ 0 & u_{22} & & u_{2n-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{n-1n-1} \end{pmatrix} \right|,$$

luego $\#\left(P_U \cap \mathbb{Z}^n\right) = |det(T)|$. □

4. Matrices y volúmenes.

En esta parte justificaremos la fórmula que permite calcular el volumen del paralelotopo definido por n vectores en $\mathbb{R}^n$.

Proposición D: Sea $A = [v_1|v_2|\cdots|v_n] \in \mathcal{M}_{n\times n}(\mathbb{R})$ con $det(A) \neq 0$. Entonces existen una matriz $Q \in SL_n(\mathbb{R})$ ortogonal y una matriz R triangular superior tales que $A = QR$.

Demostración. La demostración se basa en aplicar el proceso de ortonormalización de Gramm-Schmidt al conjunto linealmente independiente $\{v_1, v_2, \cdots, v_n\}$.

Sea

$$w_1 = \frac{v_1}{||v_1||} \rightarrow v_1 = ||v_1|| w_1 = r_{11} w_1$$
$$\vdots$$
$$v_i = < w_1, v_i > w_1 + \cdots + < w_{i-1}, v_i > w_{i-1} + ||v_i|| w_i$$
$$\vdots$$
$$v_n = < w_1, v_n > w_1 + \cdots + < w_{n-1}, v_n > w_{n-1} = r_{1n} w_1 + r_{nn} w_n$$

O sea

$$A = [v_1|v_2|\cdots|v_n] = [w_1|w_2|\cdots|w_n]\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{pmatrix} = QR$$

Es claro que $Q \in SL_n(\mathbb{R})$ es una matriz ortogonal. □

Proposición E: Sea $A = [v_1|v_2|\cdots|v_n] \in \mathcal{M}_{n\times n}(\mathbb{R})$ entonces $|det(A)| = Vol_n(P_n)$ donde P_n es el paralelotopo n-dimensional generado por los vectores $\{v_1, v_2, \cdots, v_n\}$.

Demostración. Por la afirmación anterior sabemos que:

$$A = [v_1|v_2|\cdots|v_n] = [w_1|w_2|\cdots|w_n]\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{pmatrix} = QR$$

donde Q es una matriz ortogonal tal que $|det(Q)| = 1$.

Por lo tanto $|det(A)| = det(R)$.

Por construcción $r_{ii} = ||v_i - P_{<v_1,\cdots,v_{i-1}>}(v_i)||$ donde $P_{<v_1,\cdots,v_{i-1}>}(v_i)$ indica la proyección ortogonal del vector v_i sobre el subespacio generado por los vectores $\{v_1, \cdots, v_{i-1}\}$.

Entonces:

$$det(R) = \prod_{i=1}^{n} ||v_i - P_{<v_1,\cdots,v_{i-1}>}(v_i)|| = \prod_{i=1}^{n} d(v_i, < v_1, \cdots, v_{i-1} >).$$

Denotamos por P_j el paralelotopo que definen los vectores $\{v_1, \cdots, v_{j-1}\}$; este paralelotopo tiene a P_{j-1} como base y por altura $||v_j - P_{<v_1,\cdots,v_{j-1}>}(v_j)||$. Entonces:

$$Vol_j(P_j) = Vol_{j-1}(P_{j-1}) \times ||v_j - P_{<v_1,\cdots,v_{j-1}>}(v_j)||.$$

Por lo tanto $Vol_j(P_j) = r_{jj}Vol_{j-1}(P_{j-1})$ y finalmente $Vol_n(P_n) = \prod_{i=1}^{n} r_{ii} = det(R) = |det(A)|$. □

Bibliografía

[1] Bernstein, D.N. *The number of roots of a system of equations*. Funct. Anal. Appl. **9**, (1975), p. 183-185.

[2] Cox, D.; Little, J.; O'Shea, D. *Using algebraic geometry*. Graduate Texts in Mathematics, **185**. New York, Springer, 1998.

[3] Ewald, G. *Combinatorial convexity and algebraic geometry*. Graduate Texts in Mathematics, **168**. New York, Springer, 1996.

[4] Fulton, W. *Curvas algebraicas : introducción a la geometría algebraica*. Barcelona, Reverté, 1971.

[5] Fulton, W. *Introduction to toric varieties*. Annals of Mathematics Studies, **131**. Princeton, Princeton University Press, 1993.

[6] Gelfand, I.M.; Kapranov, M.M.; Zelevinsky, A.V. *Discriminants, resultants, and multidimensional determinants*. Boston, Birkhäuser, 1994.

[7] Hartshorne, R. *Algebraic geometry*. Graduate Texts in Mathematics, **52**. New York, Springer, 1977.

[8] Humphreys, J. E. *Linear algebraic groups*. Graduate Texts in Mathematics, **21**. New York, Springer, 1981.

[9] Mumford, D. *Algebraic geometry I. Complex projective varieties*. Grundlehren der mathematischen Wissenschaften, **221**. Berlin, Springer, 1976.

[10] Nering, E. D. *Linear algebra and matrix theory*. 2nd ed., New York, Wiley, 1970.

[11] Sturmfels, B. *Gröbner bases and convex polytopes*. University Lecture Series, **8**. Providece, RI, American Mathematical Society, 1996.

[12] Sturmfels, B. *Polynomial equations and convex polytopes*. The Am. Math. Mon., **105**, No.10, (1998), p. 907-922.