

**Un poco de orden, estructura y nuevas ideas en el
cúmulo de invariantes asociados al algoritmo OIG en
aprendizaje multi-clase**

*(A bit of order, structure and new ideas in the cluster of
invariants associated to OIG algorithm in multi-class learning)*

Jiaze Zhang

Trabajo de Fin de Grado
para acceder al
Grado en Matemáticas
FACULTAD DE CIENCIAS
UNIVERSIDAD DE CANTABRIA

Directores: Daniel Sebastián San Martín & Luis Miguel Pardo Vasallo

Noviembre 2024

ABSTRACT. The aim of this paper is to introduce a bit of order, structure and contribute new ideas to the cluster of invariants in multiclass learning. In particular, we focus on the ones involved with the OIG algorithm that are present in [DS, 14] and [BCDMY, 22]. Moreover, we shall introduce and study three new invariants: $\text{outdeg}_{\text{av}}(H)$, $p\text{-dim}_{\text{DS}}(H)$ and $\max\text{-outdeg}_{\text{av}}(H)$ of a class H . On the other hand, we will also study in detail the concept of shifting (introduced in [RBR, 06] for the multiclass setting), obtaining relationships with closed downward classes and monomial basis for finitely generated Artin K -algebras (using a result from [Mes, 20]). Among other original results in this paper, we ought to highlight: a strict improvement for Corollary 28 of [BCDMY, 22] and a contribution to Lemma 3 from [DS, 14], providing a bound that is stable under shifting and allows to control $\text{outdeg}_{\text{max}}(\sigma)$, a key concept to analyze the OIG's error.

RESUMEN. El objetivo de esta memoria es poner un poco de orden, estructura y aportar nuevas ideas en el cúmulo de invariantes del aprendizaje multi-clase. En particular, nos centraremos en los invariantes involucrados en el algoritmo OIG que aparecen en [DS, 14] y [BCDMY, 22]. Además, introduciremos y estudiaremos tres nuevos invariantes: $\text{outdeg}_{\text{av}}(H)$, $p\text{-dim}_{\text{DS}}(H)$ y $\max\text{-outdeg}_{\text{av}}(H)$ de una clase H . Por otro lado, también realizaremos un estudio en detalle del *shifting* (introducido en [RBR, 06] para el caso multi-clase), obteniendo relaciones con las clases cerradas hacia abajo y las bases monomiales de K -álgebras de Artin finitamente generadas (usando un resultado de [Mes, 20]). Entre otros resultados originales de esta memoria, destacar: una mejora estricta del Corolario 28 de [BCDMY, 22] y una aportación al Lema 3 de [DS, 14], proporcionando una cota que es estable por *shifting* y que permite controlar el $\text{outdeg}_{\text{max}}(\sigma)$, que es clave para estudiar el error en el OIG.

Índice

Capítulo 0. Introducción y Resumen de Contenidos de la Memoria	i
0.1. Introducción, contexto y algunas notaciones básicas	i
0.2. Motivación de lo estudiado en la memoria, aunque no es objeto de la memoria: El algoritmo OIG bien escrito	iii
0.3. Objetivo del TFG: poner orden, estructura y aportar nuevas ideas en el cúmulo de invariantes sin estudio comparativo precedente	iv
0.4. Resumen de los principales resultados de la memoria	v
0.5. Otras consideraciones	x
0.5.1. Apéndice Final	x
0.5.2. Sobre el estilo y la ortografía usados en este TFG	x
Capítulo 1. Primeras nociones: Pre-orden de Dickson y bases monomiales	1
1.1. Pre-orden de Dickson y clases de conceptos cerradas hacia abajo: escalones y escaleras	1
1.2. Órdenes monomiales y el lenguaje de las Bases de Gröbner	3
1.2.1. Unas palabras sobre la existencia de algoritmos para el cálculo de las Bases de Gröner	7
1.2.2. El caso de los anillos de Artin y las escaleras	7
Capítulo 2. El hiper-grafo de una inclusión (OIG): orientaciones, grado, grado de salida	11
2.1. El hiper-grafo de una inclusión: orientaciones, grado, grado de salida	11
2.2. Grado y grado de salida promedios, densidad del OIG	13
2.3. Promedios de grado y grado de salida en conjuntos cerrados hacia abajo	21
2.3.1. Ejemplo: escalones \mathcal{E}_h : grado y grado de salida promedios	22
Capítulo 3. Pseudo-cubos, dimensiones DS, exponencial y de Natarajan	25
3.1. Introducción: Pseudo-cubos	25
3.2. Dimensiones	27
3.2.1. Pseudo-dimensión DS	29
3.2.2. Escalones \mathcal{E}_h y sus dimensiones	30
3.3. Pseudo-cubos cerrados hacia abajo.	31
Capítulo 4. Shifting	35
4.1. Introducción: la noción de shift	35
4.2. Shifting, cerrados hacia abajo y bases monomiales	37
4.2.1. Las ideas T. Mészáros en [Mes, 20]	41
4.3. Shifting e invariantes de clases de conceptos finitas: un par de mejoras estrictas de [BCDMY, 22]	42
Apéndice A. Pre-orden de Dickson, órdenes monomiales y Bases de Gröbner, con algo más de detalle en las pruebas	51
A.1. Pre-orden de Dickson y clases de conceptos cerradas hacia abajo: escalones y escaleras	51
A.2. Órdenes monomiales y el lenguaje de las Bases de Gröbner	53
A.2.1. Unas palabras sobre la existencia de algoritmos para el cálculo de las Bases de Gröner	58
A.2.2. El caso de los anillos de Artin y las escaleras	59
Apéndice. Bibliografía	65
Apéndice. Glosario de Términos	67
Apéndice. Glosario de teoremas y Resultados	69
Apéndice. Glosario de Símbolos y Abreviaturas	71

Introducción y Resumen de Contenidos de la Memoria

Índice

0.1.	Introducción, contexto y algunas notaciones básicas	i
0.2.	Motivación de lo estudiado en la memoria, aunque no es objeto de la memoria: El algoritmo OIG bien escrito	iii
0.3.	Objetivo del TFG: poner orden, estructura y aportar nuevas ideas en el cúmulo de invariantes sin estudio comparativo precedente	iv
0.4.	Resumen de los principales resultados de la memoria	v
0.5.	Otras consideraciones	x
0.5.1.	Apéndice Final	x
0.5.2.	Sobre el estilo y la ortografía usados en este TFG	x

0.1. Introducción, contexto y algunas notaciones básicas

Este manuscrito es una modesta contribución (los Capítulos 2 y 4 están constituidos de material original o demostraciones originales) a un estudio preliminar de elementos involucrados en el *Algoritmo OIG* (de *One Inclusion Graph* o grafo de una inclusión) dentro del contexto de *aprendizaje multi-clase*. Pero, *el manuscrito no estudia el algoritmo en sí, sino trata de analizar y estudiar toda una colección de invariantes ligados al grafo de una inclusión, como tarea preliminar al estudio (futuro) de ese algoritmo*. El OIG y su algoritmo asociado fue introducido en [Hau, 1995, HLW, 1994] para el aprendizaje binario. Posteriormente, fue extendido al caso de aprendizaje multi-clase en [RBR, 06] y re-discutido y analizado en [DS, 14]. Posteriormente fue re-analizado (creemos que con algún planteamiento incompleto) en [BCDMY, 22]. Seguiremos, con bastantes matices, corrigiendo o aclarando ciertas imprecisiones, las notaciones de [BCDMY, 22], [DS, 14] o [Aln, 1983].

El aprendizaje multi-clase difiere del aprendizaje binario (al estilo de L.G. Valiant, cf. [Val, 1984]) o la forma más precisa de V.N. Vapnik y A.Y. Chervonenkis (cf. [VC, 1971]) en varios aspectos. En el aprendizaje multi-clase disponemos de un conjunto \mathcal{X} de entradas (o *inputs*), y una segunda clase que extiende a la clase binaria (pero sigue siendo finita) y que denotaremos en todo el manuscrito mediante $\mathcal{Y} := \mathbb{Z}/p\mathbb{Z} = \{0, \dots, p-1\}$, que son los representantes canónicos al uso de las clases de restos módulo p , donde $p \geq 2$ es un entero positivo, no necesariamente primo. Nótese que el aprendizaje multi-clase contiene al aprendizaje binario (el caso binario sería el caso $p = 2$), pero mantiene diferencias relevantes que modifican su estudio. Como en todos los casos de aprendizaje, disponemos del conjunto $\mathcal{Z} := \mathcal{X} \times \mathcal{Y}$, dotado con una σ -álgebra $\mathcal{B}(\mathcal{Z})$ y una distribución de probabilidad $\nu : \mathcal{B}(\mathcal{Z}) \rightarrow [0, 1]$ “desconocida” (el uso del término “desconocida” quiere decir, en el contexto del aprendizaje, que los resultados deben probarse independientemente de la distribución elegida o, al menos, para todas las distribuciones de probabilidad sobre $\mathcal{B}(\mathcal{Z})$ que satisfagan propiedades muy generales). Denotaremos por $\mathcal{Y}^{\mathcal{X}}$ a la clase de todas las aplicaciones $f : \mathcal{X} \rightarrow \mathcal{Y}$. Nótese que, en el caso finito $\mathcal{X} := [n] := \{1, \dots, n\}$, la clase $\mathcal{Y}^{\mathcal{X}} = \mathcal{Y}^{[n]}$ puede identificarse mediante $\mathcal{Y}^{[n]} = \mathcal{Y}^n$ y preferiremos, usualmente, tratar con \mathcal{Y}^n . No obstante, en ocasiones, podemos considerar para un subconjunto $S \subseteq [n]$ de cardinal m la proyección $\pi_S : \mathcal{Y}^n \rightarrow \mathcal{Y}^m$ que “olvida” todas las coordenadas en $[n] \setminus S$. En ese caso, utilizaremos indistintamente las notaciones restricción $h|_S$ y $\pi_S(h)$ para cada $h \in \mathcal{Y}^n$. Aunque otros autores prefieren otras notaciones, en este manuscrito, dado un conjunto finito X , denotaremos por $\sharp(X)$ su cardinal.

En el aprendizaje multi-clase se dispone de una clase formada por las funciones aceptables de nuestro estudio $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ usualmente llamada *clase de conceptos* (por influencia del caso binario). El objetivo, “grosso modo”, consiste en inferir una función $h \in \mathcal{H}$ a partir de una lista finita (de longitud m) de observaciones (usualmente denominados *dataset*):

$$\mathcal{S} := \{(x_1, y_1), \dots, (x_m, y_m)\} \in (\mathcal{X} \times \mathcal{Y})^m.$$

No se trata de interpolar sino de hacer lo siguiente: asumiendo una extensión a la σ -álgebra producto $\mathcal{B}(\mathcal{Z}^m)$ de la distribución de probabilidad original $\nu^{\otimes m}$, tratar de hallar una función $h^* \in \mathcal{H}$ que minimice cierto error

(depende de la función de riesgo elegida) con alta probabilidad con respecto a la distribución “desconocida” $\nu^{\otimes m}$. Inferir h^* es una tarea compleja, por lo que se trata no tanto de inferir la función “completa” como de averiguar su valor en un punto $x \in \mathcal{X}$ dado.

Por tanto, nos enfrentamos a una cantidad de información finita y una clase de conceptos como la \mathcal{H} dada anteriormente, y el problema se reduce a conocer el valor de una candidata h_0 a partir de los objetos siguientes. En primer lugar, se considera la restricción de \mathcal{H} a los inputs $(x_1, \dots, x_m, x) \in \mathcal{Y}^{m+1}$:

$$H := \mathcal{H}|_{(x_1, \dots, x_m, x)} := \{(h(x_1), \dots, h(x_m), h(x)) \in \mathcal{Y}^{m+1} : h \in \mathcal{H}\}.$$

Usualmente, se asume una primera hipótesis:

HIPÓTESIS 0.1.1. Dados cualesquiera $(x_1, \dots, x_m, x_{m+1}) \in \mathcal{X}^{m+1}$ se puede “calcular” la clase $H := \mathcal{H}|_{(x_1, \dots, x_{m+1})} \subseteq \mathcal{Y}^{m+1}$.

Si bien esta hipótesis es un poco excesiva, subyace a todos los algoritmos de aprendizaje. Asumiendo esta Hipótesis 0.1.1, los algoritmos tratan fundamentalmente de analizar un problema de tipo finito es decir, asumiendo que $H := \mathcal{H}|_{(x_1, \dots, x_{m+1})}$ es dada como entrada al algoritmo. Por tanto, el modelo algorítmico pasa a tener la siguiente estructura:

INPUT:

- Un dataset:

$$\mathcal{S} := \{(x_1, y_1), \dots, (x_m, y_m)\} \in (\mathbb{Z})^m.$$

- Un punto adicional $x \in \mathcal{X}$
- Una clase $H \subseteq \mathcal{Y}^{m+1}$ % que sabemos que es restricción de \mathcal{H}

⋮

OUTPUT: Una función $h_0 \in H = \mathcal{H}|_{(x_1, \dots, x_m, x)}$ y el valor $h_0(x)$.

De la función h_0 se requieren ciertas condiciones adicionales: que aproxime la función buscada h^* de algún modo con pequeña probabilidad de error. Como el objetivo del trabajo no es el análisis de esos algoritmos no precisaremos con detalle el modo en que se procede a buscar esa aproximación. Pero sí observaremos que, en general, se trata de buscar una noción de función de crecimiento $\mathcal{G}(H, m)$, que dependa de H y de m , y que permita garantizar que la función h_0 obtenida por el algoritmo aproxima bien a la función buscada $h^* \in \mathcal{H}$ con alta probabilidad sobre los datasets elegidos aleatoriamente. Los valores a partir de los cuales existe aprendizaje son aquellos valores de m que satisfacen ciertas propiedades y cuyas cotas inferiores para su funcionamiento se denominan *dimensiones de \mathcal{H}* . Así, en el caso binario, la principal contribución de [VC, 1971] fue utilizar un resultado de Sauer, Shelah y Perles (cf. [Par, 24b] para una prueba alternativa a las muchas existentes), para determinar un invariante (hoy llamado dimensión de Vapnik-Chervonenkis y denotada como $\dim_{VC}(H)$) a partir del cual el aprendizaje era posible si la longitud de la muestra superaba esa dimensión, i.e., $m > \dim_{VC}(H)$.

El primero en darse cuenta de que los modelos de Vapnik y Chervonenkis no servían para el aprendizaje multi-clase fue B.K. Natarajan (cf. [Nat, 1988] o [Nat, 1989]). Natarajan propuso una noción de dimensión: la dimensión de Natarajan, denotada $\dim_N(H)$, que, como se prueba en [BCDMY, 22], es poco adecuada (véase la noción en la Observación 3.2.5 de este manuscrito). En todo caso, el aprendizaje multi-clase (i.e. en el caso $p \geq 2$) adolece de la falta de una buena noción de dimensión que pueda ser utilizada para el diseño de algoritmos razonables de aprendizaje multi-clase. En este contexto se enmarca este trabajo, para lo cual nos dedicamos a estudiar diversos invariantes relacionados con el *Algoritmo OIG* en su forma de ejemplo de Juguete (“*Toy Example*”) como se trata en [DS, 14] o [BCDMY, 22].

Para este nuevo estudio, tomamos como elemento de partida una clase de conceptos $H \subseteq \mathcal{Y}^{m+1}$ y se construye una estructura de hiper-grafo a partir de H y $m+1$ conocida como *grafo de una inclusión* (OIG) que denotaremos como $G(H)$. En la Sección 2.1 del el Capítulo 2 describiremos con detalle esa estructura. A partir de $G(H)$ se define el algoritmo de aprendizaje OIG. El objetivo de este Trabajo de Fin de Grado no es probar la correctitud del algoritmo OIG sino incidir en el análisis de su complejidad, estudiando nuevos invariantes del grafo que permitan controlarlo. En particular, introducimos los tres nuevos invariantes $\text{outdeg}_{\text{av}}(H)$, $p \dim_{DS}(H)$ y $\max - \text{outdeg}_{\text{av}}(H)$, hacemos un estudio de sus principales propiedades, sus relaciones con otros invariantes utilizados por la principal literatura sobre el tema y mostramos su relación con la complejidad del algoritmo OIG.

0.2. Motivación de lo estudiado en la memoria, aunque no es objeto de la memoria: El algoritmo OIG bien escrito

El algoritmo fue introducido en [Hau, 1995] y [HLW, 1994] para el aprendizaje binario. Posteriormente, fue extendido al caso de aprendizaje multi-clase en [RBR, 06] y discutido en [DS, 14] y [BCDMY, 22].

Trabajemos un poco las notaciones que son descritas con mayor detalle en la Sección 2.1. Dada $H \subseteq \mathcal{Y}^{m+1}$ una clase de conceptos finita, para $\mathcal{Y} := \{0, \dots, p-1\}$, consideremos las proyecciones $\pi_i : \mathcal{Y}^{m+1} \rightarrow \mathcal{Y}^m$ que “olvidan” la coordenada i -ésima de los elementos de \mathcal{Y}^{m+1} . Para cada $y \in \mathcal{Y}^m$, consideremos la fibra sobre y de π_i intersecada con H , es decir, el conjunto $e_{y,i} := \pi_i^{-1}(\{y\}) \cap H$. Si este conjunto es no vacío, definimos la arista $e_{y,i} \times \{i\} \subseteq H \times [m+1]$ y consideramos todas las aristas asociadas a la proyección π_i :

$$E_i(H) := \{e_{y,i} \times \{i\} : y \in \pi_i(H)\}.$$

Finalmente, consideramos el conjunto de aristas $E(H)$ del hiper-grafo como la siguiente unión disjunta:

$$E(H) := \bigsqcup_{i=1}^{m+1} E_i(H).$$

Para una arista $e \in E(H)$ y un elemento $h \in H$ escribimos $h \in e$ cuando $e = e_{y,i} \times \{i\}$ y $h \in e_{y,i}$. Asociada a esta estructura de hiper-grafo, está la noción de *orientación de $G(H)$* (véase la Definición 12 para una definición precisa), que no es sino una aplicación $\sigma : E(H) \rightarrow H$ que satisface la siguiente propiedad:

$$\forall e \in E(H), \sigma(e) \in e.$$

En cierto sentido, una orientación es una sección global (inversa a derecha) de todas las proyecciones π_i . Es decir, si $e = e'_{f,i} \in E(H)$ es una arista, $\sigma(e) \in H$ es un elemento de H tal que:

$$\pi_i(f_1, \dots, f_{i-1}, \sigma(e), f_i, \dots, f_m) = f.$$

Se pueden hacer otras consideraciones sobre el concepto de orientación, pero las dejamos para otros estudios que no inciden en el presente. En palabras de ciertos autores (como [RBR, 06], [DS, 14] o [BCDMY, 22]), el algoritmo de aprendizaje reposa fuertemente en el cálculo de una orientación adecuada. Por otro lado, un dataset $\mathcal{S} \in \mathcal{Y}^{[m+1]}$ pasa a ser simplemente una lista $S := \mathcal{S} \in H \subseteq \mathcal{Y}^{m+1}$. Pasemos a definir el algoritmo a la manera de [DS, 14]:

ALGORITMO 0.2.1 (ALGORITMO OIG, REVISADO, FORMATO TOY EXAMPLE).

- INPUT:
 - Un punto $(x_1, \dots, x_{m+1}) \in \mathcal{X}^{m+1}$
 - La clase de conceptos $H := \mathcal{H}|_{(x_1, \dots, x_{m+1})} \subseteq \mathcal{Y}^{m+1}$.
 - Un “dataset” $y := (y_1, \dots, y_m) \in \mathcal{Y}^m$.
- **guess** un índice $i \in [m+1]$
- **verify** que $y \in \pi_i(H)$
- **compute** una orientación $\sigma : E(H) \rightarrow H$ que minimice $\text{outdeg}_{\max}(\sigma)$ % Definiremos más tarde esta noción
- OUTPUT: $h := \sigma(e_{y,i} \times \{i\}) \in H$ % Es decir, $h(x_i) = h_i$, $h(x_j) = y_j$ si $1 \leq j \leq i-1$ y $h(x_j) = y_{j-1}$ si $i+1 \leq j \leq m+1$

La inclusión de la instrucción **guess** (con distribución de probabilidad uniforme en $[m+1]$) es un requerimiento necesario para que su Proposición 15 (clave para que este sea un algoritmo de aprendizaje) sea cierta. En otro caso la Proposición 15 de [BCDMY, 22] es falsa (está en estudio en [PSZ, 24]).

El objetivo de este Trabajo de Fin de Grado no consiste en estudiar la correctitud de este algoritmo. Aunque debemos señalar que la presentación en [BCDMY, 22] no parece correcta y sí lo es en [DS, 14]. Nuestro objetivo pasa por comenzar en el estudio de los elementos e invariantes que nacen en la complejidad de su ejecución. Claramente, el elemento clave alegado en esta presentación (que es la usual) es:

$$\min\{\text{outdeg}_{\max}(\sigma) \text{ t.q. } \sigma : E(H) \rightarrow H \text{ es una orientación}\}$$

Cantidad que denotaremos, a partir de ahora, como:

$$\min -\text{outdeg}_{\max}(H)$$

0.3. Objetivo del TFG: poner orden, estructura y aportar nuevas ideas en el cúmulo de invariantes sin estudio comparativo precedente

El tema que motiva este Trabajo Fin de Grado es que, en el análisis de este Algoritmo OIG (en particular del $\min\text{-outdeg}_{\max}(H)$), parece intervenir un CÚMULO DE NOCIONES, CASI NUNCA ESTUDIADAS CON PROFUNDIDAD NI COMPARADAS A FONDO. Éste es el verdadero *Objetivo de este Trabajo de Fin de Grado: empezar a poner un poco de orden entre este cúmulo de invariantes, en la medida de lo posible, aportando además nuevas ideas con potencial*. Vamos a hacer una rápida revisión de esas muchas nociones que acompañan a este humilde Algoritmo OIG:

- i) LAS DIVERSAS NOCIONES DE DIMENSIÓN (VÉASE EL CAPÍTULO 3). Además de la dimensión de Natarajan (i.e. $\dim_N(H)$) ya citada, aparecen la dimensión DS (denotada como $\dim_{DS}(H)$) y usada implícitamente en [DS, 14] (véase la Definición 17 para una definición precisa de esta noción de dimensión). La dimensión exponencial, introducida en [BCDMY, 22] y denotada mediante $\dim_E(H)$ (véase la Definición 18 correspondiente). *Adicionalmente, nuestro manuscrito introduce una nueva noción denominada pseudo-dimensión DS y denotada mediante $p\dim_{DS}(H)$ en la Subsección 3.2.1 y más concretamente en la Definición 19. Observamos que la dimensión DS es el máximo de la pseudo-dimensión de los subgrafos asociados a H en el Corolario 3.2.9.*
- ii) GRADOS: Para poder controlar las diversas dimensiones se usan también nociones de grados y grados de salida que tratamos de listar ahora. Dados $h \in H$ y $\sigma : E(H) \rightarrow H$ una orientación, en la Definición 13 se definen el grado de h en H y el grado de salida (*outdegree*) de σ en h con respecto a H mediante:

$$\deg_H(h) := \#\left(\{e \in E(H) : h \in e, e \notin E^{(1)}(H)\}\right),$$

donde $E^{(1)}(H)$ son las aristas unarias (i.e. de cardinal 1) en $E(H)$ y

$$\text{outdeg}_H(\sigma, h) := \#\left(\{e \in E(H) : h \in e, \sigma(e) \neq h\}\right).$$

Para no sobrecargar más aún las notaciones, sobreentendemos la existencia de un $H \subseteq \mathcal{Y}^{m+1}$ fijo y omitiremos H en los subíndices.

- iii) LOS VALORES PROMEDIO CON RESPECTO A LA DISTRIBUCIÓN UNIFORME EN H : A partir de aquí, [DS, 14] y [BCDMY, 22] consideran las nociones: grado promedio, $\text{avd}(H)$; densidad del hipergrafo, $\text{gd}(H)$; y la noción de grado de salida máximo de una orientación $\sigma : E(H) \rightarrow H$, denotado por $\text{outdeg}_{\max}(\sigma)$. *Nosotros añadimos una más que nos parece esencial en este estudio: el grado de salida promedio de una orientación σ denotado mediante $\text{outdeg}_{\text{av}}(\sigma)$.* Esas nociones se describen en la Definición 14:

$$(0.3.1) \quad \text{avd}(H) := \frac{1}{\#(H)} \sum_{h \in H} \deg(h).$$

$$(0.3.2) \quad \text{outdeg}_{\text{av}}(\sigma) := \frac{1}{\#(H)} \sum_{h \in H} \text{outdeg}(\sigma, h).$$

$$\text{gd}(H) := \frac{\#(E(H))}{\#(H)}.$$

$$\text{outdeg}_{\max}(H) := \max\{\text{outdeg}(\sigma, h) : h \in H\}.$$

- iv) DOS NOCIONES ESENCIALES: PSEUDO-CUBOS Y SHIFTING. Por si este cúmulo de invariantes no fuera suficiente, hay una noción nueva implícita en la definición de dimensión DS : la de *pseudo-cubo* (ver Definición 16). La noción es esencial para poder definir esa dimensión y aparece implícitamente en [DS, 14]. Hasta donde alcanza nuestro conocimiento, el término *pseudo-cubo* fue usado por vez primera en la literatura científica en [BCDMY, 22]. *No hay un estudio sistemático de la noción, que se hace por vez primera en este Trabajo Fin de Grado en el Capítulo 3.* El segundo elemento no es novedoso, se trata de la noción de *shifting* que se presenta en la Sección 4.1 del Capítulo 4. La noción es esencial en el tratamiento de las nociones involucradas hasta ahora. Hemos mantenido el término inglés por no encontrar un término en español apropiado para esta transformación de subconjuntos de \mathcal{Y}^n siguiendo las direcciones paralelas a los ejes y preservando la cardinalidad de las aristas. En [Hau, 1995], D. Haussler usó la técnica de “shifting” para producir una demostración del Lema de Sauer-Shelah-Perles y otros resultados en el caso de aprendizaje binario. Esta técnica aparece por vez primera en [Fra, 1983] e independientemente en [Aln, 1983] (véase también el papel del shifting en relación con la dualidad y la traza en [Par, 24b]). En [RBR, 06], los autores extienden este operador al caso del aprendizaje multi-clase y es una técnica analizada de nuevo en [BCDMY, 22]. Es esencial controlar el comportamiento de todos los invariantes antes descritos tras las operaciones

de shifting, que tampoco está tratado de manera sistemática y estructurada en la literatura sobre el tema, en especial, en el escenario multi-clase. *Lo hacemos en la Sección 4.3 del Capítulo 4. En esa sección tratamos de interpretar cómo los distintos invariantes de clases de conceptos descritos más arriba se comportan tras la operación de shifting. En particular, obtenemos una mejora estricta de algunos resultados de [BCDMY, 22].*

- v) POR SI FUERA POCO, EL PRE-ORDEN DE DICKSON (\preceq) Y SUS CERRADOS HACIA ABAJO (ESCALERAS) ESTÁN LIGADOS A BASES MONOMIALES DE K -ÁLGEBRAS DE ARTIN FINITAMENTE GENERADAS: *El objetivo de los procesos de shifting de una clase de conceptos finita es tratar de encontrar una clase cerrada hacia abajo con respecto al pre-orden de Dickson. Por eso dedicamos el Capítulo 1 y su Sección 1.1 a recordar el pre-orden de Dickson \preceq en el monoide $(\mathbb{N}^n, +)$ y las nociones de conjuntos cerrados hacia abajo (escaleras) y sus componentes (escalones). Escalones y escaleras están fuertemente ligados a las bases monomiales de anillos residuales de la forma $K[X_1, \dots, X_n]/\mathfrak{a}$ y expresamos, porque corresponde a nuestro contexto, en el caso finito donde las escaleras finitas son las bases monomiales con respecto a cualquier orden monomial \leq de anillos residuales de la forma $K[X_1, \dots, X_n]/\mathfrak{a}$, que son K -álgebras de Artin. Esto lo discutimos en la Sección 1.2 del Capítulo 1. Pero la relación entre conjuntos cerrados hacia abajo (escaleras finitas) y bases monomiales se puede precisar aún más (identificando un ideal de referencia) como en [Mes, 20]. El resultado de Mészáros es un corolario, casi inmediato, de un resultado fuerte debido a [FRR, 06]. Sin embargo, la demostración de [FRR, 06] reposa sobre una serie de ideas de Teoría de Juegos (en concreto el *lex game*) que creemos que no es necesaria. Por eso hemos desarrollado nuestra propia prueba original, totalmente alternativa e independiente, del resultado fundamental de [FRR, 06] en la Sección 4.2 del Capítulo 4.*

0.4. Resumen de los principales resultados de la memoria

Como suponemos que el lector ya debe estar un tanto saturado de tanta noción abreviada, pasemos a hacer una exposición más precisa de los principales resultados de esta memoria. Uno de los resultados con mayor contenido semántico del Capítulo 1 es el Corolario 1.2.10, que identifica los conjuntos cerrados hacia abajo finitos (i.e. escaleras finitas) con las bases monomiales de K -álgebras de Artin y que, creemos, debe ser conocido; aunque damos nuestra propia demostración. Recuérdese que el pre-orden de Dickson (véase la Definición 2) sobre el monoide $(\mathbb{N}^n, +)$ se define del modo siguiente: Dados $\alpha = (\alpha_1, \dots, \alpha_n), \beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$, diremos que α es menor que β con respecto al pre-orden de Dickson (y lo denotaremos mediante $\alpha \preceq \beta$) si se satisface

$$\exists \gamma \in \mathbb{N}^n, \alpha + \gamma = \beta.$$

Un subconjunto $H \subset \mathbb{N}^n$ se dice cerrado hacia abajo con respecto a \preceq (también denominado escalera) si verifica: $\forall \alpha, \beta \in \mathbb{N}^n$, si $\alpha \preceq \beta$ y $\beta \in H$, entonces $\alpha \in H$.

COROLARIO (DESTACABLE) 0.4.1 (ESCALERAS FINITAS SON LAS BASES MONOMIALES DE K -ÁLGEBRAS DE ARTIN). *Sea K un cuerpo de característica distinta de 2. Sea $p \in \mathbb{N}$ un entero positivo con $p \geq 2$. Sea $\mathcal{Y} := \{0, \dots, p-1\}^n$ y sea $H \subseteq \mathcal{Y}^n$ un subconjunto finito. Entonces, son equivalentes:*

- i) H es una escalera finita (i.e. un conjunto finito cerrado hacia abajo con respecto a \preceq).
- ii) Para cualquier orden monomial \leq sobre el monoide $(\mathbb{N}^n, +)$, existe un ideal \mathfrak{a} de tal modo que:
 - $H := \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})$.
 - La K -álgebra $R := K[X_1, \dots, X_n]/\mathfrak{a}$ es una K -álgebra de Artin y la siguiente es una base monomial de R como K -espacio vectorial:

$$\mathcal{B}(H)/\mathfrak{a} := \{X_1^{\mu_1} \cdots X_n^{\mu_n} + \mathfrak{a} : \mu := (\mu_1, \dots, \mu_n) \in H\},$$

con lo que, además, se tiene la siguiente igualdad de dimensiones como K -espacios vectoriales:

$$\dim_K(K[X_1, \dots, X_n]/\mathfrak{a}) = \#(H).$$

Además, en ii) se puede elegir que el ideal \mathfrak{a} sea un ideal monomial de $K[X_1, \dots, X_n]$ si bien no es imprescindible.

En este Trabajo Fin de Grado no solamente se introduce una noción nueva de grado de salida promedio (con respecto a la distribución uniforme en H) de una orientación, i.e. $\text{outdeg}_{\text{av}}(\sigma)$, sino que se demuestra que es un invariante del OIG, $G(H)$, e independiente de la orientación elegida. Para ello se caracterizan bien las diversas nociones de $\text{avd}(H)$ y grado de salida promedio en el Teorema 2.2.1 (original de este Trabajo Fin de Grado) y que reproducimos a continuación:

TEOREMA (DESTACABLE) 0.4.2 (PRINCIPALES CARACTERIZACIONES DE avd , $\text{outdeg}_{\text{av}}$ Y $\text{gd}(H)$). *Con estas notaciones, suponiendo $n \geq 2$, se tiene:*

i) El grado promedio del hiper-grafo $G(H)$ satisface:

$$\text{avd}(H) = n - \frac{\#(E^{(1)}(H))}{\#(H)},$$

donde $E^{(1)}(H)$ son las aristas unarias de $G(H)$.

ii) El grado de salida promedio de cualquier orientación $\sigma : E(H) \rightarrow H$ satisface:

$$\text{outdeg}_{\text{av}}(\sigma) = n - \text{gd}(H) = n - \frac{\#(E(H))}{\#(H)}.$$

En particular, el grado de salida promedio de una orientación es independiente de la orientación y sólo depende de H . De hecho, sólo depende de la dimensión n del espacio ambiente y de la densidad del hiper-grafo $\text{gd}(H)$. Por ello, a partir de ahora hablaremos simplemente del grado de salida promedio del hiper-grafo y lo denotaremos mediante $\text{outdeg}_{\text{av}}(H)$. En particular, se tiene:

$$\text{avd}(H) - \text{outdeg}_{\text{av}}(H) - \text{gd}(H) = -\frac{\#(E^{(1)}(H))}{\#(H)} \leq 0.$$

iii) Lo siguiente también se verifica:

$$\text{outdeg}_{\text{av}}(H) \leq \text{avd}(H) \leq 2\text{outdeg}_{\text{av}}(H).$$

iv) Para cada orientación $\sigma : E(H) \rightarrow H$, se tiene:

$$\text{gd}(H) - \frac{\#(E^{(1)}(H))}{\#(H)} \leq \text{outdeg}_{\text{av}}(\sigma).$$

v) Sea $G \subseteq H$ un subconjunto de H y sea $\sigma : E(H) \rightarrow H$ una orientación. Entonces, existe una orientación $\sigma' : E(G) \rightarrow G$ tal que se verifica:

$$\forall g \in G, \text{outdeg}(\sigma', g) \leq \text{outdeg}(\sigma, g).$$

En particular, para toda orientación $\sigma : E(H) \rightarrow H$, existirá una orientación $\sigma' : E(G) \rightarrow G$ tal que:

$$\text{outdeg}_{\text{max}}(\sigma') \leq \text{outdeg}_{\text{max}}(H).$$

Finalmente, el grado de salida promedio raramente coincide con el grado promedio, porque se tiene:

- O bien

$$\text{outdeg}_{\text{av}}(H) < \text{avd}(H),$$

- o bien $E(H) = E^{(1)}(H)$ (i.e. $G(H)$ tiene solamente aristas unarias). En ese caso,

$$\text{outdeg}_{\text{av}}(H) = \text{avd}(H) = 0.$$

Aunque el grado es una función monótona creciente cuando afecta a elementos individuales (ver Lema 2.2.7) vistos desde subclases $G \subseteq H$, no ocurre lo mismo con las funciones promedio $\text{avd}(H)$ y $\text{outdeg}_{\text{av}}(H)$. En la Proposición 2.2.8 estudiamos las distintas relaciones que pueden darse entre $\text{avd}(G)$ y $\text{avd}(H)$, cuando $G \subseteq H$. Haremos lo mismo para el grado de salida promedio en la misma proposición. Además, daremos condiciones necesarias para poder caracterizar la condición de monotonía creciente de las funciones avd y $\text{outdeg}_{\text{av}}$ en relación con la inclusión de subclases. Desgraciadamente, observamos que ambos promedios no son funciones monótonas con respecto a la inclusión, lo que impide usar argumentos inductivos basados en el cardinal $\#(H)$ para estudiar $\text{avd}(H)$ o $\text{outdeg}_{\text{av}}(H)$.

Como ya hemos indicado anteriormente, hasta donde alcanza nuestro conocimiento, el término *pseudo-cubo* fue usado en la literatura científica por vez primera en 2022 (concretamente en [BCDMY, 22]). Lo usan para expresar una noción implícita de dimensión que aparece por vez primera en [DS, 14]. La noción es crucial para poder definir la dimensión DS implícita en [DS, 14]. Sin embargo, *no existe un estudio sistemático de esa noción de pseudo-cubo hasta el presente manuscrito, por vez primera en la literatura matemática*. Comenzaremos con una caracterización de la noción de pseudo-cubo basada en los invariantes introducidos en el Capítulo 2 precedente. Obviamente, todo el material, excepto la definición, es completamente original. Recuérdese que la noción se expone en la Definición 16, que reproducimos aquí para facilitar la lectura:

DEFINICIÓN 1 (PSEUDO-CUBO). Con las notaciones precedentes, dados $n \in \mathbb{N}$, $n \geq 2$, y una clase de conceptos $H \subseteq \mathcal{Y}^n$, diremos que el OIG asociado $G(H)$ es un pseudo-cubo (de dimensión n) si y solamente si se verifica la siguiente propiedad:

Para todo $i \in [n]$ y para todo $h = (h_1, \dots, h_n) \in H$, existe un i -vecino de h en H (i.e. existe $g = (g_1, \dots, g_n) \in H$ tal que $h_i \neq g_i$, mientras que $h_j = g_j$ para todo $j \in [n] \setminus \{i\}$).

En el caso $n = 1$ llamaremos pseudo-cubo de dimensión 1 a cualquier hiper-grafo $G(H)$ asociado a cualquier subconjunto $H \subseteq \mathcal{Y}$ de cardinal mayor o igual que 2. Llamaremos pseudo-cubo de dimensión 0 a $G(\emptyset)$.

Con esta terminología, podemos aportar la primera caracterización, completamente original, de esa noción, introducida en [BCDMY, 22] que se prueba en el Teorema 3.1.4 y que prueba que la noción es equivalente a la Identidad (0.4.1), que involucra los elementos $\text{avd}(H)$, $\text{outdeg}_{\text{av}}(H)$ y $\text{gd}(H)$. Reproducimos aquí, parcialmente, ese Teorema 3.1.4:

TEOREMA (DESTACABLE) 0.4.3 (CARACTERIZACIÓN DE PSEUDO-CUBOS DE DIMENSIÓN n). *Con las mismas notaciones de páginas precedentes, sea $n \in \mathbb{N}$, $n \geq 2$, sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita, con $\sharp(H) \geq 2$ y sea $G(H) = (H, E(H))$ el hiper-grafo de una inclusión (OIG) asociado a H . Las siguientes propiedades son equivalentes:*

- i) El hiper-grafo $G(H)$ es un pseudo-cubo de dimensión n .
- ii) Grado promedio, grado de salida promedio y densidad del grafo satisfacen la siguiente identidad

$$(0.4.1) \quad \text{avd}(H) - \text{outdeg}_{\text{av}}(H) - \text{gd}(H) = 0.$$

En particular, si H es un pseudo-cubo de dimensión n y $n \geq 2$, se verifica:

$$(0.4.2) \quad \frac{n}{2} \leq \text{outdeg}_{\text{av}}(H) < n.$$

ejemplos para los que la primera desigualdad es estricta son sencillos de obtener.

Los conjuntos cerrados hacia abajo finitos (escaleras finitas) para el pre-orden de Dickson, se caracterizan por una descomposición en escalones que denotamos como \mathcal{E}_h . Para cada $h \in \mathbb{N}^n$, se define su escalón asociado con respecto a \preceq como:

$$(0.4.3) \quad \mathcal{E}_h := \{g \in \mathbb{N}^n : g \preceq h\}.$$

Hacemos un estudio pormenorizado de todos los invariantes descritos anteriormente de los escalones. En la Subsección 2.3.1 estudiamos los valores de $\text{avd}(\mathcal{E}_n)$ y $\text{outdeg}_{\text{v}}(\mathcal{E}_h)$ y en la Subsección 3.2.2 las dimensiones $\text{dim}_{\text{DS}}(\mathcal{E}_h)$ y $\text{dim}_{\text{E}}(\mathcal{E}_h)$, todas ellas dependiendo únicamente de las coordenadas de h . También estudiamos las caracterizaciones de los pseudo-cubos cerrados hacia abajo en la Proposición 3.3.1, lo que nos permite concluir una caracterización alternativa de esos pseudo-cubos cerrados hacia abajo en el Corolario 3.3.2 que reproducimos aquí:

COROLARIO (DESTACABLE) 0.4.4. *Si $H \subseteq \mathcal{Y}^n$, $n \geq 1$, es un conjunto cerrado hacia abajo, las siguientes propiedades son equivalentes:*

- i) H es un pseudo-cubo de dimensión n .
- ii) Se verifica lo siguiente:

$$\text{outdeg}_{\text{av}}(H) < \text{avd}(H) = \text{dim}_{\text{DS}}(H) = \text{dim}_{\text{E}}(H) = n \leq 2\text{outdeg}_{\text{av}}(H).$$

En el estudio de estos objetos juega un papel esencial en la literatura la operación de “shifting” que definimos con detalle en la Sección 4.1 del Capítulo 4. Así, como en la Definición 22, para cada palabra $\omega \in [n]^*$ (i.e. palabra sobre el alfabeto finito $[n] = \{1, \dots, n\}$), se define una transformación:

$$\mathbb{S}_{\omega} : 2^{\mathcal{Y}^n} \longrightarrow 2^{\mathcal{Y}^n},$$

que respeta los cardinales, preserva \preceq , modifica las aristas en una dirección paralela a uno de los ejes y tiene una propiedad fundamental, probada en [Hau, 1995], aunque es un resultado evidente por inducción en la longitud de la palabra (al fin y al cabo, un conjunto es cerrado hacia abajo para el pre-orden de Dickson si y solamente si $\mathbb{S}_i(H) = H$ para todo $i \in [n]$):

Para cada $H \subseteq \mathcal{Y}^n$ finito existe una palabra $\omega \in [n]^*$ que define una secuencia de shifts \mathbb{S}_{ω} , tal que la clase

$$H_{\omega}^* := \mathbb{S}_{\omega}(H),$$

es cerrada hacia abajo con respecto al pre-orden de Dickson. Esta reducción ayuda a probar propiedades en el caso binario $p = 2$, pero tiene sus dificultades de uso cuando $p \geq 3$. Probaremos, entre otros, los siguientes resultados:

- La dimensión exponencial se “preserva” tras realizar varios shifts. En la Proposición 4.3.1 se prueba que para cada palabra $\omega \in [n]^*$ se tiene que :

$$\text{dim}_{\text{E}}(\mathbb{S}_{\omega}(H)) \leq \text{dim}_{\text{E}}(H).$$

- La dimensión DS puede crecer tras uno o varios shifts. Esto es conocido en [BCDMY, 22], pero en el Ejemplo 4.1.1 probamos con detalle este fenómeno, exhibiendo un ejemplo de una clase $H \subseteq \{0, 1, 2\}^2$ en el que cualquier shift en cualquier dirección puede hacer crecer la dimensión DS:

$$\dim_{\text{DS}}(\mathbb{S}_1(H)) = \dim_{\text{DS}}(\mathbb{S}_2(H)) = 2 > \dim_{\text{DS}}(H) = 1.$$

- El grado promedio tiene un buen comportamiento en el caso $p = 2$ (como ya se demostraba en [Hau, 1995], que probamos aquí con demostración propia en el Corolario 4.3.10). Sin embargo, en el caso $p \geq 3$, el grado promedio tiene un comportamiento errático al realizar uno o varios shifts, como probamos en el Ejemplo 4.3.11. En él se muestra un conjunto finito $H \subseteq \{0, 1, 2, 3\}^2$ tal que se tiene que:
 - Realizando un sólo shift con respecto a la dirección π_2 se tiene que:

$$\text{avd}(\mathbb{S}_2(H)) < \text{avd}(H).$$

- Por otro lado, si hacemos un shift en la dirección de la proyección π_1 obtenemos:

$$\text{avd}(\mathbb{S}_1(H)) > \text{avd}(H).$$

Este comportamiento errático de $\text{avd}(H)$ al aplicar shifts se sigue estudiando en [PSZ, 24], pero los resultados son aún muy primitivos.

Pero nosotros podemos probar la estabilidad del grado de salida promedio $\text{outdeg}_{\mathfrak{S}_{\text{av}}}(H)$, como se prueba en el Teorema 4.3.8 que reproducimos aquí:

TEOREMA (DESTACABLE) 0.4.5. *Con las notaciones anteriores, dado $H \subseteq \mathcal{Y}^n$ y dada $\omega \in [n]^*$, el grado de salida promedio no decrece tras el shifting, i.e.:*

$$\text{outdeg}_{\text{av}}(\mathfrak{S}_{\omega}(H)) \geq \text{outdeg}_{\mathfrak{S}_{\text{av}}}(H).$$

Dado que estamos insistiendo en transformar, a través de varios shifts, un conjunto finito $H \subseteq \mathcal{Y}^n$ en uno $\mathfrak{S}_{\omega}(H)$ cerrado hacia abajo para el pre-orden de Dickson, trataremos de entender mejor qué se obtiene tras unas secuencias finitas de shifts. Eso se expresa bien mediante un resultado de [Mes, 20] que se basa fuertemente en un teorema previo, muy fuerte, de [FRR, 06]. La demostración de [FRR, 06] reposa sobre el uso de un juego (el *lex game*) y nos pareció inapropiado que un resultado de este tipo dependiera de la Teoría de Juegos, por lo que hemos desarrollado nuestra propia demostración en la Sección 4.2. En esa sección escribimos una demostración alternativa y original del teorema principal de [FRR, 06]. Vamos a enunciar simplemente ese enunciado de [FRR, 06], para lo cual necesitaremos unas pocas nociones más.

Sea \leq_{lex} el orden monomial lexicográfico sobre el monoide $(\mathbb{N}^n, +)$ (ver definición en el Ejemplo 1.2.2). Para un orden monomial \leq sobre $(\mathbb{N}^n, +)$ y un ideal $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$, consideramos el conjunto $\text{exp}_{\leq}(\mathfrak{a}) \subseteq \mathbb{N}^n$ como el conjunto de los exponentes de los términos directores de los elementos del ideal \mathfrak{a} con respecto al orden monomial \leq (véase una definición precisa en la Definición 6). Fijamos \leq_{lex} el orden lexicográfico tal que

$$X_1 <_{\text{lex}} X_2 <_{\text{lex}} \cdots <_{\text{lex}} X_n,$$

denotemos para un ideal $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$:

$$\text{exp}_{\text{lex}}(\mathfrak{a}) := \text{exp}_{\leq_{\text{lex}}}(\mathfrak{a}),$$

y

$$\text{Nexp}_{\text{lex}}(\mathfrak{a}) := \mathbb{N}^n \setminus \text{exp}_{\leq_{\text{lex}}}(\mathfrak{a}).$$

Por supuesto, por no complicar las notaciones usaremos la misma notación, independientemente de n . Dado K un cuerpo y $H \subseteq K^n$, denotamos por $I(H)$ el ideal de los polinomios en $K[X_1, \dots, X_n]$ que se anulan en H . Y, con las mismas notaciones, denotamos por $K[H]$ al anillo cociente:

$$K[H] := K[X_1, \dots, X_n]/I(H).$$

Finalmente, dado $H \subseteq \mathcal{Y}^n$ y dado $b \in \mathcal{Y}$ denotamos mediante $H(b) \subseteq \mathcal{Y}^{n-1}$ a:

$$(0.4.4) \quad H(b) := \{(\theta_1, \dots, \theta_{n-1}) \in \mathcal{Y}^{n-1} : (\theta_1, \dots, \theta_{n-1}, b) \in H\}.$$

Probamos, de modo completamente original, el siguiente importante Teorema 4.2.4 en la Sección 4.2:

TEOREMA (DESTACABLE) 0.4.6 ([FRR, 06]). *Sea \mathcal{Y} un conjunto finito, K un cuerpo, supongamos $\mathcal{Y} \subseteq K$ (inmerso de alguna forma) y sea $H \subseteq \mathcal{Y}^n$. Son equivalentes para un exponente monomial $\omega := (\omega_1, \dots, \omega_n) \in \mathbb{N}^n$:*

- i) $\omega \in \text{Nexp}_{\text{lex}}(I(H))$.
- ii) Se verifica la siguiente desigualdad:

$$\#(\{b \in \mathcal{Y} : (\omega_1, \dots, \omega_{n-1}) \in \text{Nexp}_{\text{lex}}(I(H(b)))\}) \geq \omega_n + 1.$$

Este resultado nos permite probar, en la Subsección 4.2.1, el resultado que consideramos más destacable de [Mes, 20] y que expresamos aquí en la forma del Corolario 4.2.7:

COROLARIO (DESTACABLE) 0.4.7 (H ES CERRADO HACIA ABAJO $\Leftrightarrow H = \text{Nexp}_{1_{\text{ex}}}(I(H))$). *Con las notaciones precedentes, un subconjunto $H \subseteq \mathcal{Y}^n$ es cerrado hacia abajo si y solamente si $H = \text{Nexp}_{1_{\text{ex}}}(I(H))$, donde $I(H)$ es el ideal de todos los polinomios $p \in K[X_1, \dots, X_n]$ que sea anulados en H .*

En la sección final del presente manuscrito (Sección 4.3), presentamos una mejora estricta de la Proposición 27 y del Corolario 28 de [BCDMY, 22] a través de nuestro Teorema 4.3.12, que reproducimos aquí:

TEOREMA (DESTACABLE) 0.4.8. *Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Entonces, tenemos:*

$$\text{outdeg}_{\mathfrak{S}_{\text{av}}}(H) < \dim_{\mathbb{E}}(H).$$

Este teorema implica las siguientes consecuencias, que suponen una mejora estricta de la Proposición 27 de [BCDMY, 22]. Esto consiste en transformar $\text{avd}(H) \leq 4 \dim_{\mathbb{E}}(H)$ en la desigualdad que aparece en la Afirmación *ii*) ($\text{avd}(H) < 2 \dim_{\mathbb{E}}(H)$) del corolario siguiente. Evidentemente, el uso de $\text{outdeg}_{\mathfrak{S}_{\text{av}}}(H)$ es esencial en esta pequeña mejora.

COROLARIO (DESTACABLE) 0.4.9. *Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Entonces, tenemos:*

i) Las dimensiones satisfacen:

$$(0.4.5) \quad \dim_{\text{DS}}(H) < 2 \dim_{\mathbb{E}}(H).$$

ii) El grado promedio satisface:

$$(0.4.6) \quad \text{avd}(H) < 2 \dim_{\mathbb{E}}(H).$$

Estamos en condiciones de probar el segundo teorema relevante de esta sección final del manuscrito, que mejora estrictamente los resultados de [BCDMY, 22], y los pone en un contexto más amplio. Introduzcamos la noción de grado mínimo del hiper-grafo $G(H)$ definido por una clase de conceptos finita $H \subseteq \mathcal{Y}^n$ como:

$$\text{deg}_{\min}(H) := \min\{\text{deg}(h) : h \in H\}.$$

Como se observa en varios estudios, la función $\text{avd}(H)$ no es monótona para subgrafos. Es decir, dados $H' \subseteq H$, no es en general cierto que $\text{avd}(H') \leq \text{avd}(H)$. Para hacer que $\text{avd}(H)$ sea una función monótona con respecto a la inclusión, en [DS, 14], los autores introducen una variación, monotonizándola, del modo siguiente:

$$(0.4.7) \quad \max - \text{avd}(H) := \max\{\text{avd}(F) : F \subseteq H\}.$$

De modo similar a [DS, 14], podemos introducir el grado de salida promedio “maximal” mediante:

$$(0.4.8) \quad \max - \text{outdeg}_{\mathfrak{S}_{\text{av}}}(H) := \max\{\text{outdeg}_{\mathfrak{S}_{\text{av}}}(F) : F \subseteq H\}.$$

Como se observa de nuestros resultados, se tiene:

$$\max - \text{avd}(H) \leq 2 \max \text{outdeg}_{\mathfrak{S}_{\text{av}}}(H).$$

Esto nos permite demostrar una mejora estricta sobre la existencia de orientaciones con grado de salida maximal controlado que se expresa en el Corolario 28 de [BCDMY, 22]. En ese corolario, los autores probaron la existencia de una orientación $\sigma : E(H) \rightarrow H$ que satisface:

$$\text{outdeg}_{\max}(\sigma) \leq 4 \dim_{\mathbb{E}}(H).$$

Lo mejoramos en nuestro Teorema 4.3.14, que reproducimos a continuación, especialmente en la Afirmación *i*) y la Desigualdad (0.4.12):

TEOREMA (DESTACABLE) 0.4.10 (MEJORA ESTRICTA DEL COROLARIO 28 DE [BCDMY, 22] SOBRE EL GRADO DE SALIDA MÁXIMO DE ORIENTACIONES). *Con las mismas notaciones precedentes, sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Sea $\varphi : 2^H \rightarrow \mathbb{R}$ una función monótona con respecto a la inclusión en 2^H . Supongamos que la función φ también satisface la siguiente hipótesis:*

$$(0.4.9) \quad \forall F \in 2^H, \text{deg}_{\min}(F) \leq \varphi(F).$$

Entonces, existe una orientación $\sigma : E(H) \rightarrow H$ tal que se verifica:

$$(0.4.10) \quad \text{outdeg}_{\max}(\sigma) \leq \varphi(H).$$

Adicionalmente, tenemos:

i) La Hipótesis (0.4.9) se verifica siempre que la siguiente condición sea cierta:

$$(0.4.11) \quad \text{avd}(F) \leq \varphi(F), \quad \forall F \subseteq H.$$

En particular, la función monótona $\varphi(F) := 2 \dim_{\mathbb{E}}(F)$ satisface la Desigualdad (0.4.11) y, por tanto, existe una orientación $\sigma : E(H) \rightarrow H$ tal que

$$(0.4.12) \quad \text{outdeg}_{\max}(\sigma) \leq 2 \dim_{\mathbb{E}}(H).$$

ii) El grado promedio maximal también satisface la Hipótesis (0.4.11) y, por tanto, existe una orientación $\sigma : G(H) \rightarrow H$ tal que

$$\text{outdeg}_{\max}(\sigma) \leq \max -\text{avd}(H) \leq 2 \max -\text{outdeg}_{\text{av}}(H).$$

El siguiente corolario expresa la relación de nuestro estudio con la complejidad del algoritmo OIG a través de su invariante principal:

COROLARIO 0.4.11. *Con las notaciones precedentes, se tiene:*

$$\text{outdeg}_{\text{av}}(H) \leq \min -\text{outdeg}_{\max}(H) \leq 2 \max -\text{outdeg}_{\text{av}}(H).$$

OBSERVACIÓN 0.4.12 (CONSIDERACIONES FINALES). Observemos que la Afirmación *ii)* es el Lema 3 de [DS, 14], añadiendo el doble del grado promedio maximal, mientras que la Afirmación *i)* mejora estrictamente el Corolario 28 de [BCDMY, 22] por un factor estrictamente mayor que 2. Sin embargo, no hemos sido capaces de probar resultados similares para la función (naturalmente monótona) \dim_{DS} . El Ejemplo 4.3.16 final prueba que la Hipótesis (0.4.11) no es satisfecha por la dimensión DS. Más aún, este ejemplo prueba también que $\max -\text{avd}(H)$ puede ser mayor que $\dim_{\text{DS}}(H)$. Por tanto, el Lema 3 de [DS, 14] no se aplica para producir orientaciones cuyo grado de salida máximo esté acotado por $\dim_{\text{DS}}(H)$. En consecuencia, parece que $2 \max -\text{outdeg}_{\text{av}}(H)$ es la cantidad más fina que podemos encontrar tal que sea estable por shifting y que permita controlar el grado de salida máximo de alguna orientación en un OIG sobre H . Pero sigue siendo un objeto de estudio en [PSZ, 24].

OBSERVACIÓN 0.4.13 (ALGUNOS TEMAS DE FUTURAS INVESTIGACIONES). No está aún claro para este autor si $\text{outdeg}_{\text{av}}(H)$ es el invariante más adecuado para controlar la PAC *learnability*. En particular, no disponemos de su relación precisa con $\dim_{\text{DS}}(H)$, que parece, según [BCDMY, 22], la medida más adecuada para caracterizar “*learnability*”.

También convendría poder conocer mejor el comportamiento de $\text{outdeg}_{\text{av}}(H)$ bajo restricciones, asunto del que sólo disponemos de resultados incompletos. Por último, dados los resultados que conectan el shifting de clases finitas de monomios con la nulidad de polinomios sobre H , convendría disponer de un estudio que explique mejor la relación del hiper-grafo OIG (y sus invariantes) con el tratamiento y estudio de propiedades de polinomios con pocos coeficientes no nulos (“*fewnomials*” en la terminología de A. Khovanskiĭ).

Por último, no estamos seguros de haber logrado un orden suficiente ni una estructura óptima en nuestro estudio del cúmulo de nociones ligadas al algoritmo OIG, por lo que este trabajo es sólo una etapa inicial en la ardua tarea de poner orden, abstracción y síntesis en tal contexto.

0.5. Otras consideraciones

0.5.1. Apéndice Final. Las limitaciones auto-impuestas por la Facultad de Ciencias de la Universidad de Cantabria, exigen que el número máximo de páginas de un Trabajo Fin de Grado sea de 50 páginas. Esto nos obliga a dar una versión muy breve del Capítulo 1, cuya naturaleza esquemática lo hace hasta difícil de leer o asimilar. Por eso, hemos creado un único apéndice (el Apéndice A) donde, con un poco más de lugar, el Capítulo 1 aparece expandido. Nos excusamos por usar este subterfugio para respetar las condiciones auto-impuestas para los TFGs del Grado de Matemáticas.

0.5.2. Sobre el estilo y la ortografía usados en este TFG. En algún caso precedente se ha discutido el estilo y la ortografía de las memorias presentadas como Trabajo de Fin de Grado en Matemáticas. En evitación de intervenciones innecesarias, queremos clarificar algunos aspectos relativos al estilo elegido en este texto. Se ha elegido el formato de libro (book) de la American Mathematical Society (AMS). Aunque el idioma utilizado es el español, hemos tratado de seguir lo más fielmente posible las recomendaciones del Libro de Estilo de esta asociación ¹, conjuntamente con las reglas de estilo recomendadas por D. E. Knuth y co-autores para la Mathematical Association of America (MAA) ².

Específicamente, hemos tratado de seguir atentamente las siguientes dos reglas:

¹M. Letourneau, J. Wright Sharp, AMS Style Guide, Journals, October 2017, AMS, Providence, 2017

²D. E. Knuth, T. Larrabee, P. M. Roberts, Mathematical Writing, MAA, 1989

- “*Numbered theorems, lemmas, etc. are proper nouns and, thus, are capitalized: Theorem 2.3, Lemma 3.1, Figure 4.5*” (p. 79 del AMS Style Guide).
- “*Rule 19. Capitalize names like Theorem 1, Lemma 2, Algorithm 3, Method 4*” (en D. E. Knuth et al.).

Primeras nociones: Pre-orden de Dickson y bases monomiales

Índice

1.1. Pre-orden de Dickson y clases de conceptos cerradas hacia abajo: escalones y escaleras	1
1.2. Órdenes monomiales y el lenguaje de las Bases de Gröbner	3
1.2.1. Unas palabras sobre la existencia de algoritmos para el cálculo de las Bases de Gröner	7
1.2.2. El caso de los anillos de Artin y las escaleras	7

Como ya se ha indicado en la Introducción, este capítulo es sólo una breve introducción a nociones esencialmente conocidas por el graduado en Matemáticas por la Universidad de Cantabria, como son las nociones de *pre-orden de Dickson*, *orden monomial* o *Bases de Gröbner*. Al mismo tiempo hablaremos de conjuntos cerrados hacia abajo para el pre-orden de Dickson (conocidos popularmente como *escaleras*). Dedicaremos un tiempo mínimo al estudio de la relación de las escaleras finitas con las *bases monomiales* de K -álgebras de Artin. El resultado fundamental será el Corolario 1.2.10 que identifica escaleras finitas con bases monomiales de ciertas K -álgebras de Artin con respecto a cualquier orden monomial, dadas como los cocientes $K[X_1, \dots, X_n]/\mathfrak{a}$ con respecto a un ideal \mathfrak{a} apropiado. Más adelante demostraremos una versión refinada de esta idea, usando sólo el orden lexicográfico \leq_{lex} . Daremos una demostración nueva de un poderoso resultado de [FRR, 06] en la Sección 4.2 y podremos demostrar que los conjuntos cerrados hacia abajo son el resultado de aplicar la operación de “shifting” (cf. Capítulo 4), concluyendo un resultado de [Mes, 20] y le daremos una vuelta a esta idea mostrando que, en realidad $H \subseteq \mathcal{Y}^n$, con \mathcal{Y} finito, es una escalera finita si y solamente si las clases de sus monomios módulo $I(H)$ están en biyección con $\mathbb{N}^n \setminus \exp_{\leq_{\text{lex}}}(I(H))$, donde $I(H) \subseteq K[X_1, \dots, X_n]$ es el ideal de todos los polinomios en $K[X_1, \dots, X_n]$ que se anulan en H y $\exp_{\leq_{\text{lex}}}(I(H))$ son los exponentes directores de los polinomios en $I(H)$ con respecto al orden lexicográfico. Pero esto lo dejamos para el Teorema 4.2.6.

Por una cuestión de espacio, hemos reducido el capítulo hasta un contenido extremadamente estructural, por lo que lo hemos reproducido con algo más de detalle (especialmente los ejemplos y algunas demostraciones no esenciales en nuestro estudio) en el Apéndice A. Lamentamos esta forma de presentarlo, pero las limitaciones explícitas del número de páginas de los Trabajos Fin de Grado en la Facultad de Ciencias de la Universidad de Cantabria hacen que resulte, a veces, incómodo, poder expresar con suficiente elegancia una presentación de nociones e ideas y nos hemos visto obligados a este subterfugio.

1.1. Pre-orden de Dickson y clases de conceptos cerradas hacia abajo: escalones y escaleras

Comenzaremos con un sencillo resultado combinatorio sobre el monoide conmutativo $(\mathbb{N}^n, +)$ conocido como LEMA DE DICKSON (cf. [Dix, 1913]), aunque parece que también era conocido por P. Gourdan.

DEFINICIÓN 2 (PRE-ORDEN DE DICKSON). *Consideremos el monoide $(\mathbb{N}^n, +)$ y consideremos la relación (de orden parcial) siguiente:*

Dados $\alpha = (\alpha_1, \dots, \alpha_n), \beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$, diremos que α es menor que β con respecto al pre-orden de Dickson si se satisface

$$\exists \underline{\gamma} \in \mathbb{N}^n, \alpha + \underline{\gamma} = \beta.$$

Escribiremos $\alpha \preceq \beta$ para indicar esa relación. Se denomina pre-orden de Dickson

Algunas sencillas propiedades de \preceq se resumen en el siguiente enunciado, de fácil demostración:

PROPOSICIÓN 1.1.1. *Con las notaciones anteriores, se verifica:*

- i) En el caso $n = 1$ el orden \preceq es el orden usual en \mathbb{N} .*
- ii) La relación \preceq es una relación de orden parcial sobre \mathbb{N}^n .*
- iii) Para cada $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, denotemos por $\underline{X}^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n} \in K[X_1, \dots, X_n]$ al monomio cuyos exponentes vienen dados por las coordenadas de α , siendo K un cuerpo cualquiera. Entonces, para cada $\alpha, \beta \in \mathbb{N}^n$ se tiene que $\alpha \preceq \beta \iff \underline{X}^\alpha \mid \underline{X}^\beta$, donde \mid significa “divide” en $K[X_1, \dots, X_n]$.*

El resultado fundamental de las ideas de Dickson se resume en el siguiente resultado (conocido como *Lema de Dickson*) cuya prueba puede seguirse fácilmente de los pasos descritos en [Par, 24a].

TEOREMA 1.1.2 (LEMA DE DICKSON). *Con las notaciones precedentes, todo subconjunto no vacío $S \subseteq \mathbb{N}^n$ posee un número finito de elementos minimales para el orden \preceq .*

DEFINICIÓN 3 (CONJUNTOS CERRADOS HACIA ABAJO). *Con las notaciones precedentes, diremos que una clase de conceptos $H \subseteq \mathcal{Y}^n$ es cerrada hacia abajo (con respecto al pre-orden de Dickson \preceq) si satisface la siguiente propiedad:*

$$\forall h \in H, \forall g \in \mathcal{Y}^n, g \preceq h \implies g \in H.$$

EJEMPLO 1.1.3 (ESCALÓN). Un sencillo ejemplo de clase de conceptos cerrada hacia abajo es el ejemplo del *escalón* determinado por un exponente monomial. Así, dado $h \in \mathbb{N}^n$ definimos el escalón determinado por h como el conjunto siguiente:

$$(1.1.1) \quad \mathcal{E}_h := \{g \in \mathbb{N}^n : g \preceq h\}.$$

Obviamente, los escalones son conjuntos cerrados hacia abajo y para cada $h \in H$, si H es cerrado hacia abajo, se tiene $\mathcal{E}_h \subseteq H$.

La siguiente proposición explica el término de escalón y, en el caso finito, la “escalera” que determinan.

PROPOSICIÓN 1.1.4 (ESCALERA (FINITA)). *Una clase de conceptos finita $H \subseteq \mathcal{Y}^n$ es cerrada hacia abajo si y solamente si H admite una descomposición del tipo siguiente:*

Existen $h_1, \dots, h_s \in \mathbb{N}^n$ tales que las siguientes dos propiedades se verifican:

- *Cada elemento $h_i \in H$ no está en los escalones determinados por los demás elementos, i.e.:*

$$h_i \notin \bigcup_{j \neq i} \mathcal{E}_{h_j}.$$

- *La clase H es una unión finita de tales escalones:*

$$(1.1.2) \quad H := \bigcup_{i=1}^s \mathcal{E}_{h_i}.$$

Diremos que H es una escalera cuyos escalones son determinados por $\mathcal{E}_{h_1}, \dots, \mathcal{E}_{h_s}$.

Más aún, si $H \subseteq \mathbb{N}^n$ es una escalera, los escalones $\mathcal{E}_{h_1}, \dots, \mathcal{E}_{h_s}$ están determinados de manera única por H y los elementos h_1, \dots, h_s son los elementos maximales de H con respecto al pre-orden de Dickson \preceq .

Los escalones (únicamente determinados) $\mathcal{E}_{h_1}, \dots, \mathcal{E}_{h_s}$ serán denominados las componentes (de la escalera) del conjunto finito cerrado hacia abajo H .

DEMOSTRACIÓN. Ciertamente, si H es finito sólo puede tener un número finito de elementos maximales con respecto a \preceq . Como \mathcal{E}_h es siempre cerrado hacia abajo, tenemos que si $h \in H$, el escalón que determina satisface $\mathcal{E}_h \subseteq H$. Esto prueba que si H es cerrado hacia abajo, entonces es la unión finita de los escalones de sus elementos maximales con respecto a \preceq , lo que demuestra la Identidad (1.1.2). Más aún, elementos maximales de H con respecto a \preceq no serán comparables entre sí. Por tanto, para cada $i \in [s]$, $h_i \notin \bigcup_{j \neq i} \mathcal{E}_{h_j}$, lo que prueba la primera propiedad.

En consecuencia, cada unión finita de escalones es un conjunto cerrado hacia abajo, lo que prueba el recíproco. \square

OBSERVACIÓN 1.1.5. Dado un subconjunto $H \subseteq \mathbb{N}^n$ como una unión finita de escalones (y, por tanto, es también un conjunto finito):

$$H := \bigcup_{k=1}^t \mathcal{E}_{h_k},$$

concluimos que H es cerrado hacia abajo. Pero esa descomposición puede refinarse hasta obtener una descomposición que satisfaga las Afirmaciones *i)* y *ii)* de la proposición precedente. Basta con elegir los elementos maximales $S \subseteq H$ para \preceq de entre la clase $\{h_1, \dots, h_t\}$.

Por otro lado, podemos considerar el “anti-escalón” asociado a un elemento de \mathbb{N}^n . Es decir, podemos introducir la notación siguiente: Dado $h \in \mathbb{N}^n$, definamos el siguiente subconjunto de \mathbb{N}^n :

$$h + \mathbb{N}^n := \{g \in \mathbb{N}^n : \exists m \in \mathbb{N}^n, h + m = g\} = \{g \in \mathbb{N}^n : h \preceq g\},$$

donde $h + m$ denota la suma estándar en el monoide $(\mathbb{N}^n, +)$. Se tiene la siguiente caracterización alternativa:

PROPOSICIÓN 1.1.6. *Con las notaciones precedentes, para cada $H \subseteq \mathcal{Y}^n$, no necesariamente finito, H es cerrado hacia abajo si y solamente si existen $g_1, \dots, g_t \in \mathbb{N}^n$, no comparables entre ellos con respecto a \preceq , tales que:*

$$H := \bigcap_{i=1}^t (\mathbb{N}^n - (g_i + \mathbb{N}^n)).$$

Más aún, los elementos g_1, \dots, g_t están determinados de manera única por H por ser los elementos minimales para el pre-orden de Dickson \preceq del conjunto $\mathbb{N}^n \setminus H$.

DEMOSTRACIÓN. Se trata simplemente de considerar el conjunto $\mathbb{N}^n \setminus H$ y usar el Lema de Dickson (Teorema 1.1.2 anterior) para concluir que hay solamente un número finito de elementos minimales de $\mathbb{N}^n \setminus H$: $\{g_1, \dots, g_t\}$. Esos son los elementos que satisfacen las propiedades enunciadas. \square

Obsérvese que la condición de ser cerrada hacia abajo es hereditaria mediante restricciones.

LEMA 1.1.7. *Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Entonces, H es cerrada hacia abajo si y solamente si lo son todas sus restricciones $H|_S$ para cualquier $S \subseteq [n]$.*

DEMOSTRACIÓN. Supongamos que H es cerrada hacia abajo y sea $S \subseteq [n]$ un subconjunto. Dada $g \in H$ una función en la clase de conceptos y consideramos $g|_S \in H|_S$. Entonces, para cualquier $f : S \rightarrow \mathcal{Y}$ tal que $f \preceq g|_S$, definamos un “lifting”, $\tilde{f} : [n] \rightarrow \mathcal{Y}$, de f a $[n]$, mediante:

$$\tilde{f}(i) := \begin{cases} f(i), & \text{si } i \in S, \\ g(i), & \text{si } i \notin S. \end{cases}$$

Por tanto, se tiene que $\tilde{f}|_S = f$ y $\tilde{f} \preceq g$. Entonces, como H es cerrada hacia abajo, $\tilde{f} \in H$ y, por tanto, $f = \tilde{f}|_S \in H|_S$, lo que prueba que $H|_S$ es cerrada hacia abajo. El recíproco es inmediato con tomar $S = [n]$. \square

1.2. Órdenes monomiales y el lenguaje de las Bases de Gröbner

Uno de los elementos esenciales en el trabajo con polinomios univariados (especialmente en la División euclídea y en el desarrollo de la Teoría de la Eliminación basada en resultantes), es la existencia de un orden natural en el monoide $(\mathbb{N}, +)$. Para el trabajo con polinomios multivariados, la extensión natural de esta relación de orden son los *órdenes monomiales* sobre el monoide $(\mathbb{N}^n, +)$. Queremos introducir un modelo de división que se “parezca” a la División euclídea y que se denominará *División de Weierstarss-Hironaka*.

OBSERVACIÓN 1.2.1 (ORDENAR EL MONOIDE $(\mathbb{N}^n, +)$ MEDIANTE BIYECCIONES CON $(\mathbb{N}, +)$). Podemos identificar \mathbb{N}^n con \mathbb{N} de diversas formas. Cada biyección $\varphi : \mathbb{N}^n \rightarrow \mathbb{N}$ induce una relación de orden sobre \mathbb{N}^n del modo siguiente:

Dados $\mu, \theta \in \mathbb{N}^n$, $\mu \leq_\varphi \theta$ si y solamente si $\varphi(\mu) \leq \varphi(\theta)$. Nos interesa no solamente una biyección sino que, además, se comporte bien sobre la estructura del monoide $(\mathbb{N}^n, +)$. En este sentido, en lugar de tratar con biyecciones, interesan, sobre todo, los *órdenes monomiales*.

DEFINICIÓN 4 (ORDEN MONOMIAL). *Un orden monomial sobre $(\mathbb{N}^n, +)$ es un buen orden \leq sobre \mathbb{N}^n (i.e. todo subconjunto no vacío de \mathbb{N}^n posee mínimo para \leq) que verifica las siguientes propiedades:*

- i) $\forall \mu, \theta, \tau \in \mathbb{N}^n$, si $\mu \leq \tau$, entonces $\mu + \theta \leq \tau + \theta$.
- ii) El elemento $0 = (0, \dots, 0) \in \mathbb{N}^n$ es el mínimo de (\mathbb{N}^n, \leq) .

EJEMPLO 1.2.2 (ORDEN LEXICOGRÁFICO CANÓNICO). El orden lexicográfico \leq_{lex} es el orden dado por la regla siguiente:

Dados $\mu := (\mu_1, \dots, \mu_n) \in \mathbb{N}^n$ y $\nu := (\nu_1, \dots, \nu_n) \in \mathbb{N}^n$, diremos que $\mu \leq_{\text{lex}} \nu$ si $\mu = \nu$ ó $\mu <_{\text{lex}} \nu$, es decir, existe $k \in \{0, \dots, n-1\}$ tal que se verifican las dos propiedades siguientes:

- $\mu_i = \nu_i$, para $1 \leq i \leq k$. Obsérvese que si $k = 0$, ésta es la condición vacía.
- $\mu_{k+1} < \nu_{k+1}$.

Es un orden monomial en \mathbb{N}^n , pero no permite biyectar \mathbb{N}^n con \mathbb{N} para $n \geq 2$. Si el orden \leq_{lex} viniera de una biyección de \mathbb{N}^n con \mathbb{N} , dado $\mu \in \mathbb{N}^n$, el conjunto de los elementos $\{\theta \in \mathbb{N}^n : \theta \leq_{\text{lex}} \mu\}$ sería un conjunto finito porque así sucede en \mathbb{N} . Pero eso no es cierto para $n \geq 2$. Si elegimos $\mu = (1, 0, \dots, 0) \in \mathbb{N}^n$ observamos que este conjunto es infinito:

$$\{\theta \in \mathbb{N}^n : \theta \leq_{\text{lex}} (1, 0, \dots, 0)\} = (\{0\} \times \mathbb{N}^{n-1}) \cup \{(1, 0, \dots, 0)\}.$$

Otros ejemplos que definimos en el Apéndice A son el orden lexicográfico salvo permutación de los índices (Ejemplo A.2.3), Grado + lexicográfico (Ejemplo A.2.4) y otros muchos ejemplos que pueden verse en [CLO, 07] o [Mor, 16] y sus muchas referencias. Nos quedaremos con $\leq_{1\text{ex}}$ porque es el que más tratamos en este manuscrito. Otras extensiones de estos resultados están en fase de estudio en [PSZ, 24].

Los órdenes monomiales juegan un papel esencial en la ordenación de los monomios en varias variables. Así, consideremos un conjunto de variables $\{X_1, \dots, X_n\}$, algebraicamente independientes sobre un cuerpo K . Y consideremos la base monomial del K -espacio vectorial definida por esos monomios

$$\mathcal{M}(X_1, \dots, X_n) := \{X_1^{\mu_1} \cdots X_n^{\mu_n} : (\mu_1, \dots, \mu_n) \in \mathbb{N}^n\}.$$

Tenemos claramente un isomorfismo de monoides, donde en $\mathcal{M}(X_1, \dots, X_n)$ se considera la restricción de la operación producto natural del anillo de polinomios $K[X_1, \dots, X_n]$ dada del modo siguiente:

$$(1.2.1) \quad \begin{aligned} \exp : (\mathcal{M}(X_1, \dots, X_n), \cdot) &\longrightarrow (\mathbb{N}^n, +) \\ X_1^{\mu_1} \cdots X_n^{\mu_n} &\longmapsto (\mu_1, \dots, \mu_n). \end{aligned}$$

Dado un término $T = aX_1^{\alpha_1} \cdots X_n^{\alpha_n} \in K[X_1, \dots, X_n]$, con $a \neq 0$, llamaremos *exponente monomial* del término al exponente de su monomio, i.e.

$$\exp(T) := \exp(X_1^{\alpha_1} \cdots X_n^{\alpha_n}) := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n.$$

Los órdenes monomiales también tienen una relación particular con el pre-orden de Dickson:

PROPOSICIÓN 1.2.3. *Con las notaciones precedentes, se verifica:*

- i) *La relación \preceq es una relación de orden parcial sobre $(\mathbb{N}^n, +)$, que no es un orden monomial.*
- ii) *Si \leq es un orden monomial cualquiera en $(\mathbb{N}^n, +)$, dados $\alpha, \beta \in \mathbb{N}^n$, si $\alpha \preceq \beta$, entonces $\alpha \leq \beta$. El recíproco no es cierto, como lo prueban ejemplos muy elementales.*
- iii) *Dados $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ y $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$, son equivalentes:*

$$\alpha \preceq \beta \iff X_1^{\alpha_1} \cdots X_n^{\alpha_n} \mid X_1^{\beta_1} \cdots X_n^{\beta_n},$$

donde \mid significa “divide” en $K[X_1, \dots, X_n]$.

- iv) *Dado un orden monomial cualquiera \leq sobre el monoide $(\mathbb{N}^n, +)$ y cualquier subconjunto no vacío $S \subseteq \mathbb{N}^n$, existe un mínimo de S para \leq . Más aún, ese mínimo de S se encuentra entre los elementos minimales (en número finito) de S con respecto al pre-orden de Dickson \preceq .*

DEMOSTRACIÓN. Todas las afirmaciones son esencialmente inmediatas y pueden verse en [Par, 24a]. Las hemos incluido en el Apéndice A, en la prueba de la Proposición A.2.5 \square

Fijado un cuerpo K (o un anillo R) los polinomios en n variables con coeficientes en K son simplemente funciones $f \in K^{\mathbb{N}^n}$ con *soporte finito* finito, donde el soporte de f es el conjunto dado por la siguiente identidad:

$$(1.2.2) \quad \text{supp}(f) := \{\mu \in \mathbb{N}^n : f(\mu) \neq 0\}.$$

Dicho de otro modo, el anillo de polinomios $K[X_1, \dots, X_n]$ es simplemente el subanillo del anillo de las series de potencias formales $K[[X_1, \dots, X_n]]$ dado por la siguiente igualdad:

$$K[X_1, \dots, X_n] := \{f \in K[[X_1, \dots, X_n]] : \sharp(\text{supp}(f)) < \infty\}.$$

Así, un monomio $X_1^{\mu_1} \cdots X_n^{\mu_n} \in K[X_1, \dots, X_n]$, con $\mu := (\mu_1, \dots, \mu_n) \in \mathbb{N}^n$ es simplemente la función:

$$(1.2.3) \quad \begin{aligned} X_1^{\mu_1} \cdots X_n^{\mu_n} : \mathbb{N}^n &\longrightarrow K \\ \theta &\longmapsto \begin{cases} 1, & \text{si } \theta = \mu \\ 0, & \text{en otro caso.} \end{cases} \end{aligned}$$

Esto conduce a la representación usual de los polinomios como combinaciones lineales de monomios. A diferencia de la representación “densa”, nos concentraremos en la representación “dispersa” (conocida en inglés como “sparse”, “fewnomial” o, en francés, como “creuse”) en la que sólo destacaremos los monomios con coeficientes no nulos. Esto es, presentamos los polinomios como el siguiente tipo de combinación lineal:

$$(1.2.4) \quad f := \sum_{\mu \in \text{supp}(f)} a_\mu X_1^{\mu_1} \cdots X_n^{\mu_n} \in K[X_1, \dots, X_n],$$

donde, por definición, $a_\mu := f(\mu) \in K \setminus \{0\}$. Esta presentación es única, lo que recuerda que $\mathcal{M}(X_1, \dots, X_n)$ es una base de $K[X_1, \dots, X_n]$ como K -espacio vectorial. A los polinomios no nulos de la forma $T := a_\mu X_1^{\mu_1} \cdots X_n^{\mu_n} \in K[X_1, \dots, X_n] \setminus \{0\}$ se les denomina *términos* y al elemento $a_\mu \in K \setminus \{0\}$ se le denomina *coeficiente* del término no nulo. Al elemento $\mu = (\mu_1, \dots, \mu_n)$ se le denomina *exponente* del término $\exp(T) := \mu$, mientras que se denomina *grado* del término al valor $|\mu| = \mu_1 + \cdots + \mu_n \in \mathbb{N}$.

Dado $f \in K[X_1, \dots, X_n]$, su soporte $\text{supp}(f) \subseteq \mathbb{N}^n$ define una clase de conceptos finita que puede ser ordenada mediante dos estrategias:

- *Mediante el pre-orden de Dickson:* En este caso, dado $f \in K[X_1, \dots, X_n] \setminus \{0\}$, $\text{supp}(f)$ tendrá un número finito de elementos minimales y maximales. En el caso en que $\text{supp}(f)$ sea cerrado hacia abajo diremos que el polinomio es (genéricamente) denso. El conjunto $\text{supp}(f)$ tendrá un sólo mínimo y constituirá una escalera (finita) determinada por sus elementos maximales.
- *Mediante un orden monomial \leq :* Por otro lado, dado $f \in K[X_1, \dots, X_n] \setminus \{0\}$, el orden inducido sobre el conjunto finito $\text{supp}(f)$ por \leq será un orden total y un buen orden, con lo que tendrá un único máximo y un único mínimo, coincidentes si y solamente si f es un término no nulo. Así, fijado el orden monomial \leq sobre $(\mathbb{N}^n, +)$, y dado $f \in K[X_1, \dots, X_n]$, supongamos que $\mu := \max_{\leq}(\text{supp}(f))$ es el elemento maximal del soporte de f para \leq , llamaremos *término director de f* al término no nulo cuyo exponente sea, precisamente, ese máximo μ con respecto a \leq . Lo denotaremos mediante $\ell t_{\leq}(f)$ y se define formalmente como sigue:

$$\ell t_{\leq}(f) := a_{\mu} X_1^{\mu_1} \cdots X_n^{\mu_n} \in K[X_1, \dots, X_n],$$

donde $\mu := (\mu_1, \dots, \mu_n) = \max_{\leq}(\text{supp}(f))$. Como en el caso de monomios y términos, llamaremos a ese máximo y lo denotaremos mediante:

$$\exp_{\leq}(f) := \max_{\leq}(\text{supp}(f)).$$

Los exponentes y términos directores de cada polinomio dependen del orden monomial elegido.

Otra medida distinta del “tamaño” de la representación dispersa de un polinomio es su grado, que resulta más sencillo de definir por el único orden monomial en el monoide $(\mathbb{N}, +)$:

DEFINICIÓN 5 (GRADO Y GRADOS PARCIALES). *Sea $f \in K[X_1, \dots, X_n] \setminus \{0\}$ un polinomio no nulo.*

- i) *Llamaremos grado (total) de f a la cantidad dada por la siguiente igualdad:*

$$\deg(f) := \max\{|\mu| : \mu \in \text{supp}(f)\}.$$

- ii) *Fijado $i \in [n]$, consideremos el cuerpo de funciones racionales:*

$$K_i := K(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n),$$

que no contiene a la variable X_i . Como $f \in K_i[X_i]$, llamaremos grado (parcial) de f con respecto a la variable X_i y lo denotaremos mediante $\deg_{X_i}(f)$ al grado de f como elemento del anillo de polinomios $K_i[X_i]$.

Como la Matemática tiene esa tendencia a llamar con nombres distintos a las mismas cosas, dependiendo del contexto, vamos a tratar de ver que lo descrito en esta sección está fuertemente ligado a lo descrito en la Sección 1.1 precedente en una sencilla proposición que no requiere prueba por lo expuesto hasta ahora.

PROPOSICIÓN 1.2.4. *Con las notaciones precedentes, sea $\mathcal{Y} = \{0, \dots, p-1\}$. Entonces, se tiene:*

- i) *Dado cualquier polinomio no nulo $f \in K[X_1, \dots, X_n]$, tal que $\deg_{X_i}(f) \leq p-1$, para cada $i \in [n]$, su soporte $\text{supp}(f) \subseteq \mathcal{Y}^n$ es una clase de conceptos finita.*
- ii) *Recíprocamente, dada una clase de conceptos $H \subseteq \mathcal{Y}^n \subseteq \mathbb{N}^n$, entonces, existe un espacio vectorial $W(H) \subseteq K[X_1, \dots, X_n]$, de dimensión igual al cardinal $\sharp(H)$, formado por polinomios tales que $\deg_{X_i}(f) \leq p-1$, para cada $i \in [n]$, de tal que modo que para cada polinomio no nulo $f \in W(H)$ se tiene que:*

$$\text{supp}(f) \subseteq H.$$

- iii) *El conjunto $W(H)^* := W(H) \cap (K^{\times})^{\mathbb{N}^n}$ (tórico, i.e. con todos sus términos con coeficientes no nulos y, por tanto, un abierto Zariski en $W(H)$) satisface:*

$$\text{supp}(f) = H \iff f \in W(H)^*.$$

Más aún, la clase de conceptos $H \subseteq \mathcal{Y}^n$ es cerrada hacia abajo si y solamente si todos los polinomios $f \in W(H)^$ son (genéricamente) densos.*

La existencia de un orden es un elemento esencial en la División euclídea y, por tanto, en el desarrollo de algoritmos que trabajan con polinomios univariados. Del mismo modo, la existencia y uso de órdenes monomiales es esencial para concebir una algorítmica de división con polinomios multi-variados: la *División de Weierstarss-Hironaka* o, simplemente División de Hironaka (cf. sus trabajos [Hir, 1964a] y [Hir, 1964b], que le valieron la medalla Fields del año 1970), que discutimos a continuación:

TEOREMA 1.2.5 (DIVISIÓN DE WEIERSTARSS-HIRONAKA). *Sea \leq un orden monomial sobre el monoide $(\mathbb{N}^n, +)$. Sean $f_1, \dots, f_r \in K[X_1, \dots, X_n]$ y $g \in K[X_1, \dots, X_n]$. Supongamos fijados los exponentes (con respecto a \leq) de cada f_i mediante $\mu_i := \exp_{\leq}(f_i)$, $i \in [r]$.*

Definamos los siguientes conjuntos de exponentes monomiales en el monoide $(\mathbb{N}^n, +)$:

$$\begin{aligned} \Delta_1 &:= \mu_1 + \mathbb{N}^n, \\ \Delta_2 &:= (\mu_2 + \mathbb{N}^n) \setminus \Delta_1, \\ \Delta_3 &:= (\mu_3 + \mathbb{N}^n) \setminus (\Delta_1 \cup \Delta_2), \\ &\vdots \\ \Delta_r &:= (\mu_r + \mathbb{N}^n) \setminus (\cup_{i=1}^{r-1} \Delta_i), \\ \overline{\Delta} &:= \mathbb{N}^n \setminus (\cup_{i=1}^r \Delta_i). \end{aligned}$$

Entonces, existen polinomios $g_1, \dots, g_r, h \in K[X_1, \dots, X_n]$ tales que se satisfacen las siguientes propiedades:

- i) Para cada $i \in [r]$, los soportes satisfacen $\text{supp}(g_i X^{\mu_i}) \subseteq \Delta_i$.*
- ii) El soporte del polinomio h satisface: $\text{supp}(h) \subseteq \overline{\Delta}$.*
- iii) Se verifica la igualdad $g := g_1 f_1 + \dots + g_r f_r + h$.*
- iv) Los polinomios g_1, \dots, g_r verificando las anteriores propiedades son únicos y se denominan cocientes de la División de Hironaka.*
- v) El polinomio h es único y se denomina resto de la División de Hironaka.*

Además, si el orden monomial \leq es algoritmizable, existe un algoritmo que calcula los polinomios g_1, \dots, g_r, h en su forma dispersa (i.e. como en la Identidad (1.2.4)).

Ejemplos de algoritmos de los que se citan en el enunciado se puede ver en [Par, 24a] o en [CLO, 07], por citar algunos.

Del mismo modo que hemos introducido exponentes monomiales para monomios, términos y polinomios, podemos introducir los exponentes de un ideal:

DEFINICIÓN 6 (EXPONENTE DE UN IDEAL). *Sea K un cuerpo y $K[X_1, \dots, X_n]$ el anillo de polinomios en n variables con coeficientes en el cuerpo K . Sea $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$ un ideal de ese anillo. Sea \leq un orden monomial en el monoide $(\mathbb{N}^n, +)$. Llamaremos exponente del ideal \mathfrak{a} con respecto a \leq al conjunto formado por los exponentes de todos los elementos del ideal, i.e.*

$$(1.2.5) \quad \exp_{\leq}(\mathfrak{a}) := \{\exp_{\leq}(f) : f \in \mathfrak{a}\} \subseteq \mathbb{N}^n.$$

PROPOSICIÓN 1.2.6. *Sea \leq un orden monomial sobre el monoide $(\mathbb{N}^n, +)$ y sean K un cuerpo y $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$ un ideal de ese anillo de polinomios. Entonces, se tiene:*

- i) El conjunto $\exp_{\leq}(\mathfrak{a})$ es cerrado hacia arriba para el pre-orden de Dickson \preceq .*
- ii) Su complementario $\mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})$ es cerrado hacia abajo para el pre-orden de Dickson \preceq .*

DEMOSTRACIÓN. Es casi inmediato. Probemos i): si $\alpha \preceq \beta$ y $\alpha \in \exp_{\leq}(\mathfrak{a})$ es porque existen:

- $\gamma := (\gamma_1, \dots, \gamma_n) \in \mathbb{N}^n$ tal que $\beta = \alpha + \gamma$ en el monoide $(\mathbb{N}^n, +)$.
- Existe $f \in \mathfrak{a}$ tal que $lt_{\leq}(f) = \alpha$. Consideremos el polinomio:

$$g := X_1^{\gamma_1} \cdots X_n^{\gamma_n} f \in \mathfrak{a}.$$

Como los órdenes monomiales respetan la estructura del monoide $(\mathbb{N}^n, +)$ tendremos que:

$$lt(g) = \gamma + lt(f) = \gamma + \alpha = \beta \in \exp_{\leq}(\mathfrak{a}).$$

Esto prueba que $\exp_{\leq}(\mathfrak{a})$ es cerrado hacia arriba para \preceq . Por tanto, su complementario será cerrado hacia abajo para \preceq . \square

TEOREMA 1.2.7 (EXISTENCIA DE BASES DE GRÖBNER). *Sea K un cuerpo, y sea $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$ un ideal en el anillo de polinomios $K[X_1, \dots, X_n]$. Sea \leq un orden monomial en el monoide $(\mathbb{N}^n, +)$. Entonces, existe una familia finita $G := \{f_1, \dots, f_s\} \subseteq \mathfrak{a}$ de elementos, dos a dos distintos, en el ideal \mathfrak{a} de tal modo que se verifica:*

$$(1.2.6) \quad \exp_{\leq}(\mathfrak{a}) = \bigcup_{i=1}^s (\exp_{\leq}(f_i) + \mathbb{N}^n).$$

En particular, el ideal \mathfrak{a} es generado por el conjunto finito $\mathcal{G} := \{f_1, \dots, f_s\}$ (lo que implica el Basissatz de Hilbert para el anillo de polinomios $K[X_1, \dots, X_n]$).

A cualquier conjunto finito de elementos \mathcal{G} que satisface la Identidad (1.2.6) se le denomina base de Gröbner

del ideal \mathfrak{a} para el orden monomial \leq .

Más aún, una base como K -espacio vectorial del anillo cociente $K[X_1, \dots, X_n]/\mathfrak{a}$ es dada por el conjunto:

$$(1.2.7) \quad \mathcal{B}/\mathfrak{a} := \{X_1^{\mu_1} \cdots X_n^{\mu_n} + \mathfrak{a} : \mu := (\mu_1, \dots, \mu_n) \in \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})\}$$

Al conjunto \mathcal{B}/\mathfrak{a} se le denomina base monomial del anillo cociente $K[X_1, \dots, X_n]/\mathfrak{a}$.

DEMOSTRACIÓN. El enunciado es consecuencia casi inmediata de combinar el Lema de Dickson (cf. Teorema 1.1.2), con la División de Weierstarss-Hironaka (cf. Teorema 1.2.5). Dejamos el detalle de la demostración para el Apéndice A y la demostración del Teorema A.2.9. Véase también [CLO, 07], [Mor, 16] o [Par, 24a]. \square

Obsérvese que las Bases de Gröbner \mathcal{G} no son, en realidad, bases del ideal como K -espacio vectorial, sino sistemas finitos de generadores. En cambios, las clases \mathcal{B}/\mathfrak{a} , definidas como las clases módulo \mathfrak{a} del complementario de $\exp_{\leq}(\mathfrak{a})$, sí constituyen una base del anillo residual $K[X_1, \dots, X_n]/\mathfrak{a}$ como K -espacio vectorial. Es como si la terminología estuviera cambiada.

1.2.1. Unas palabras sobre la existencia de algoritmos para el cálculo de las Bases de Gröner.

Aunque este Trabajo Fin de Grado no trata de algoritmos sobre Bases de Gröbner, citaremos un poco de la historia y de las primeras contribuciones debidas a Bruno Buchberger. La historia, en los tiempos recientes, comienza con los trabajos de H. Hironaka sobre *resolución de singularidades en característica cero* (cf. [Hir, 1964a] y [Hir, 1964b]), aunque hay trazas de la misma idea en la obra de K. Weierstrass (cf. [Wei, 1854, Wei, 1876]). En estos trabajos ya aparece la idea de la División de Weierstarss-Hironaka y la idea de *bases estándar* que preluirá la obra de B. Buchberger. En [Buc, 1965] y en [Buc, 1985], B. Buchberger trata de interpretar la noción de *bases estándar* usada por Hironaka, pero adaptándola al anillo de polinomios $K[X_1, \dots, X_n]$. Las denomina *Bases de Gröbner* en honor de su director de tesis W. Gröbner quien, a su vez, hizo investigación bajo la dirección de E. Noether en Göttingen entre 1932 y la marcha de E. Noether a Suiza. Lo que hace Buchberger es desarrollar un algoritmo elemental (de ahí su popularidad) basado en *técnicas de reescritura en monooides*. No es un algoritmo eficiente ni teóricamente ni experimentalmente. Los algoritmos experimentalmente más eficientes hasta la fecha son los debidos a J.C. Faugère quien, sin embargo, abandonó el ámbito académico en la cumbre de su éxito (siendo ya, como poco, *Directeur de Recherche du CNRS*), para ganar más dinero en el mundo de la industria de software. Ha habido muchas variantes del algoritmo de Buchberger, muchos resultados que muestran que las técnicas de reescritura no son del todo las más naturales y existe toda una bibliografía extensísima con los pros y contras de este tipo de técnicas. Nos limitamos con dar al lector unas pocas referencias para que inicie su búsqueda en caso de tener interés: [CLO, 07], [Mor, 16], [Stur, 1993], [HHPS, 20], [HHPS, 21] y un extensísimo etcétera en el que no hemos incluido, precisamente, las alternativas mucho más eficientes para los propósitos de la Teoría de la Eliminación.

1.2.2. El caso de los anillos de Artin y las escaleras. El matemático austriaco Emil Artin se incorpora a la Universidad de Göttingen en 1921. Allí coincidirá con D. Hilbert y colaborará con E. Noether y H. Hasse, hasta su marcha, en 1923, a la Universidad de Hamburgo y su exilio en Princeton, por causa del nazismo, en 1937. Mientras se encuentra en Göttingen, colabora en el seminario de E. Noether. No fue E. Artin, sino E. Noether quien introdujo en su seminario la condición de cadena numerable descendente para ideales en anillos. Era una terminología enmarcada en sus conferencias sobre Teoría de Grupos, como noción dual de sus propias ideas sobre la condición de cadena ascendente numerable. El testimonio proviene del texto histórico [vdW, 1985], cuyo autor fue testigo privilegiado de aquel tiempo. En 1927, ya en Hamburgo, Artin generaliza el teorema de Wedderburn de 1908 sobre números hiper-complejos (es decir, una teoría de álgebras sobre cuerpos arbitrarios). La generalización de Artin en [Artin, 1927] reposa sobre esa condición de cadena descendente numerable de Noether (de nuevo, según testimonio de [vdW, 1985]). En nuestro contexto, nos interesan especialmente las K -álgebras de Artin (terminología más actualizada de los sistemas hiper-complejos) y, por extensión, los *anillos de Artin*, que serán el objeto de este capítulo. Todo el material que sigue puede verificarse en [Par, 24a].

DEFINICIÓN 7 (ANILLO DE ARTIN). *Un anillo R se dice anillo de Artin (o anillo artiniiano) si satisface la siguiente propiedad, conocida como la Condición de Cadena Numerable descendente de ideales :*

Toda cadena descendente numerable de ideales se estabiliza. Es decir, dada una cadena descendente numerable de ideales de R :

$$\mathfrak{a}_0 \supseteq \mathfrak{a}_1 \supseteq \cdots \supseteq \mathfrak{a}_n \supseteq \cdots,$$

existe $m \in \mathbb{N}$ tal que $\mathfrak{a}_n = \mathfrak{a}_m$, $\forall n \geq m$.

DEFINICIÓN 8 (ANILLOS DE DIMENSIÓN CERO). *Un anillo R se dice anillo de dimensión de Krull cero (o anillo cero-dimensional) si satisface $\text{Spec}(R) = \text{MaxSpec}(R)$, donde $\text{Spec}(R)$ es el espectro primo del anillo y $\text{MaxSpec}(R)$ es el espectro maximal de R .*

TEOREMA 1.2.8 (TEOREMA DE AKIZUKI). *Un anillo R es artiniano si y solamente si se verifican las dos propiedades siguientes:*

- i) R es noetheriano, es decir, todo ideal de R es finitamente generado.
- ii) Se verifica $\text{Spec}(R) = \text{MaxSpec}(R)$ (i.e. todo ideal \mathfrak{p} primo en R es un ideal maximal en R).

En otras palabras, un anillo de Artin es un anillo cero-dimensional en el que todo ideal es finitamente generado.

Debe recordarse que ser de dimensión cero no garantiza inmediatamente la condición de ser anillo de Artin (i.e. la condición noetheriana es esencial). Un ejemplo puede verse en [Par, 24a] (Capítulo 10) o en el Apéndice A como Ejemplo A.2.11. En [Par, 24a] se prueba una larga lista de equivalencias sobre la condición de ser K -álgebra de Artin. Nos conformaremos con mostrar algunas de ellas.

DEFINICIÓN 9 (VARIEDADES ALGEBRAICAS AFINES K -DEFINIBLES). *Sea K un cuerpo, \mathbb{K} una extensión algebraicamente cerrada y $\mathbb{A}^n(\mathbb{K})$ el espacio afín sobre \mathbb{K} de dimensión n . Consideremos el anillo de polinomios en n variables $K[X_1, \dots, X_n]$ y su extensión $\mathbb{K}[X_1, \dots, X_n]$ con coeficientes en \mathbb{K} .*

- i) Sea $f \in K[X_1, \dots, X_n]$, llamaremos hiper-superficie afín definida por f en $\mathbb{A}^n(\mathbb{K})$, que denotaremos por $V_{\mathbb{A}^n(\mathbb{K})}(f)$ al conjunto:

$$V_{\mathbb{A}^n(\mathbb{K})}(f) := \{x \in \mathbb{A}^n(\mathbb{K}) : f(x) = 0\} = f^{-1}(\{0\}).$$

- ii) Llamaremos intersección finita de hiper-superficies algebraicas K -definibles a la intersección de un número finito de tales hiper-superficies definidas por una familia finita $\{f_1, \dots, f_s\} \subseteq K[X_1, \dots, X_n]$ y lo denotaremos mediante:

$$V_{\mathbb{A}^n(\mathbb{K})}(f_1, \dots, f_s) := \bigcap_{i=1}^s V_{\mathbb{A}^n(\mathbb{K})}(f_i).$$

- iii) Dado un ideal $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$, llamaremos variedad algebraica K -definible asociada al ideal \mathfrak{a} al conjunto:

$$V_{\mathbb{A}^n(\mathbb{K})}(\mathfrak{a}) := \bigcup_{f \in \mathfrak{a}} V_{\mathbb{A}^n(\mathbb{K})}(f) = \{\zeta \in \mathbb{A}^n(\mathbb{K}) : f(\zeta) = 0, \forall f \in \mathfrak{a}\}.$$

En el caso de que no haya confusión entre las dimensiones, los cuerpos K y \mathbb{K} involucrados (especialmente cuando $K = \mathbb{K}$), nos conformaremos con usar las notaciones $V_{\mathbb{A}}(f)$, $V_{\mathbb{A}}(f_1, \dots, f_s)$ y $V_{\mathbb{A}}(\mathfrak{a})$.

Por supuesto, el famoso Basissatz de Hilbert implica que las variedades algebraicas K -definibles son exactamente las intersecciones finitas de hiper-superficies K -definibles. Y que $V_{\mathbb{A}}(\mathfrak{a})$ no depende del sistema generador finito del ideal elegido para esa intersección finita de hiper-superficies. En ocasiones no nos interesan solamente las soluciones de un sistema de ecuaciones polinomiales K -definibles en una extensión algebraica del cuerpo K , sino en el propio cuerpo de coeficientes (o, incluso, en un subcuerpo, aunque es un caso menos habitual).

DEFINICIÓN 10 (VARIEDADES Y PUNTOS L -RACIONALES). *Con notaciones similares a las de la anterior definición, sean $L \subseteq K$ una extensión de cuerpos, $\mathbb{A}^n(L)$ el espacio afín de dimensión n sobre L . Sea $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$ un ideal. Llamaremos variedad de puntos L -racionales asociada a \mathfrak{a} al conjunto siguiente:*

$$V_{\mathbb{A}^n(L)}(\mathfrak{a}) := V_{\mathbb{A}^n(\mathbb{K})}(\mathfrak{a}) \cap \mathbb{A}^n(L).$$

A los puntos de $V_{\mathbb{A}^n(L)}(\mathfrak{a})$ se les denomina denomina puntos L -racionales de la variedad algebraica $V_{\mathbb{A}^n(\mathbb{K})}(\mathfrak{a})$. De nuevo, cuando no haya confusión con los cuerpos L, K y \mathbb{K} escribiremos simplemente $V_{\mathbb{A}(L)}(\mathfrak{a})$.

En el caso de sistemas hiper-complejos (i.e. K -álgebras finitamente generadas de Artin), se tiene la siguiente caracterización que también puede verse en [Par, 24a]. El enunciado dice lo siguiente:

TEOREMA 1.2.9. *Sea K un cuerpo y \mathbb{K} su clausura algebraica. Sea $\{f_1, \dots, f_s\}$ un conjunto finito de polinomios en $K[X_1, \dots, X_n]$. Sea $\mathfrak{a} = (f_1, \dots, f_s)$ el ideal que genera ese conjunto en $K[X_1, \dots, X_n]$ y sea \mathfrak{a}^e la extensión de \mathfrak{a} al anillo en $\mathbb{K}[X_1, \dots, X_n]$. Sea $V_{\mathbb{A}}(\mathfrak{a}) = V_{\mathbb{A}}(f_1, \dots, f_s) = V_{\mathbb{A}}(\mathfrak{a}^e) \subseteq \mathbb{A}^n(\mathbb{K})$ la variedad algebraica K -definible de sus ceros comunes. Son equivalentes:*

- i) El sistema de ecuaciones siguiente:

$$(1.2.8) \quad \begin{cases} f_1(X_1, \dots, X_n) = 0, \\ f_2(X_1, \dots, X_n) = 0, \\ \vdots \\ f_s(X_1, \dots, X_n) = 0, \end{cases}$$

posee solamente un número finito de soluciones en $\mathbb{A}^n(\mathbb{K})$.

- ii) La variedad $V_{\mathbb{A}}(\mathfrak{a}) = V_{\mathbb{A}}(\mathfrak{a}^e)$ es un conjunto finito de puntos en $\mathbb{A}^n(\mathbb{K})$.

- iii) El anillo $K[X_1, \dots, X_n]/\mathfrak{a}$ es un anillo de Artin.
- iv) El anillo $\mathbb{K}[X_1, \dots, X_n]/\mathfrak{a}^e$ es un anillo de Artin.
- v) El anillo $K[X_1, \dots, X_n]/\mathfrak{a}$ es un K -espacio vectorial de dimensión finita.
- vi) El anillo $\mathbb{K}[X_1, \dots, X_n]/\mathfrak{a}^e$ es un \mathbb{K} -espacio vectorial de dimensión finita.

Además, en cualquiera de estos casos, se tienen las siguientes igualdades de dimensiones:

$$\dim_K(K[X_1, \dots, X_n]/\mathfrak{a}) = \dim_{\mathbb{K}}(\mathbb{K}[X_1, \dots, X_n]/\mathfrak{a}^e).$$

De toda esta discusión abstracta, extraemos simplemente la equivalencia entre la condición de ser una K -álgebra de Artin y las clases de conceptos finitas introducidas en la Sección 1.1 precedente. En realidad, el resultado no debe ser muy desconocido, por lo que lo probamos por falta de una referencia apropiada.

COROLARIO 1.2.10 (ESCALERAS FINITAS EQUIVALEN A BASES MONOMIALES DE K -ÁLGEBRAS DE ARTIN). *Sea K un cuerpo de característica distinta de 2. Sea $p \in \mathbb{N}$ un entero positivo con $p \geq 2$. Sea $\mathcal{Y} := \{0, \dots, p-1\}^n$ y sea $H \subseteq \mathcal{Y}^n$ un subconjunto finito. Entonces, son equivalentes:*

- i) H es una escalera finita (i.e. un conjunto finito cerrado hacia abajo con respecto a \preceq).
- ii) Para cualquier orden monomial \preceq sobre el monoide $(\mathbb{N}^n, +)$, existe un ideal \mathfrak{a} de tal modo que:
 - $H := \mathbb{N}^n \setminus \exp_{\preceq}(\mathfrak{a})$.
 - La K -álgebra $K[X_1, \dots, X_n]/\mathfrak{a}$ es una K -álgebra de Artin, que posee la siguiente base monomial como K -espacio vectorial:

$$\mathcal{B}(H)/\mathfrak{a} := \{X_1^{\mu_1} \cdots X_n^{\mu_n} + \mathfrak{a} : \mu := (\mu_1, \dots, \mu_n) \in H\},$$

con lo que, además, se tiene la siguiente igualdad de dimensiones como K -espacios vectoriales:

$$\dim_K(K[X_1, \dots, X_n]/\mathfrak{a}) = \#(H).$$

Además, en ii) se puede elegir que el ideal \mathfrak{a} sea un ideal monomial de $K[X_1, \dots, X_n]$ aunque no es imprescindible.

DEMOSTRACIÓN. Probemos ambas implicaciones.

- $i) \implies ii)$: Comencemos considerando el ideal monomial $\mathfrak{b} := (X_1^p, \dots, X_n^p)$ y, claramente, el anillo residual $A := K[X_1, \dots, X_n]/\mathfrak{b}$ es un anillo de Artin y un K -espacio vectorial de dimensión p^n . Seguidamente, consideremos el conjunto $\mathcal{Y}^n \setminus H$ y el conjunto finito de monomios asociado:

$$\mathcal{M}(\mathcal{Y}^n \setminus H) := \{X_1^{\mu_1} \cdots X_n^{\mu_n} : \mu := (\mu_1, \dots, \mu_n) \in \mathcal{Y}^n \setminus H\}.$$

Definamos los exponentes $\{p \cdot e_1, \dots, p \cdot e_n\} \subseteq \mathbb{N}^n$, donde $\{e_1, \dots, e_n\} \subseteq \{0, 1\}^n$ es la base “canónica” usual de \mathbb{R}^n , y consideremos el conjunto complementario de H en \mathbb{N}^n dado mediante:

$$H^c := \mathbb{N}^n \setminus H := (\mathcal{Y}^n \setminus H) \cup (\cup_{i=1}^n (p \cdot e_i + \mathbb{N}^n)).$$

Consideremos el conjunto de elementos minimales de H^c con respecto al pre-orden de Dickson y lo denotamos mediante $\mathfrak{m}(H^c)$. Tenemos que, por la propia definición de los elementos minimales de este conjunto y por ser ese conjunto cerrado hacia arriba, se tiene que:

$$(1.2.9) \quad H^c = \bigcup_{\mu \in \mathfrak{m}} (\mu + \mathbb{N}^n).$$

Escribamos $\mathfrak{m}(H^c) := \{\mu_1, \dots, \mu_t\}$, con la condición de que sean distintos dos a dos. Obviamente, se tiene el contenido siguiente: $\mathfrak{m}(H^c) \subseteq H^c$. Finalmente, definamos la familia de monomios siguiente:

$$\mathcal{M}_0(H^c) := \{X_1^{\mu_{i,1}} \cdots X_n^{\mu_{i,n}} : \mu_i := (\mu_{i,1}, \dots, \mu_{i,n}) \in \mathfrak{m}(H^c)\}.$$

Y consideremos el ideal monomial \mathfrak{a} generado por $\mathcal{M}_0(H^c)$, es decir, $\mathfrak{a} := (\mathcal{M}_0(H^c)) = (\mathcal{M}_0(H^c)) + \mathfrak{b}$, porque los exponentes $p \cdot e_i$ de los monomios de la forma X_i^p tienen que estar en algún $\mu_i + \mathbb{N}^n$, para algún $i \in [t]$ minimal en H^c para \preceq . Por tanto, $\mathfrak{b} \subseteq \mathfrak{a}$ y podemos considerar el ideal $\bar{\mathfrak{a}} := \mathfrak{a}/\mathfrak{b}$ de las clases de restos de elementos del ideal \mathfrak{a} módulo el ideal \mathfrak{b} . Además, tenemos un epimorfismo natural de K -álgebras que, combinado con el Segundo teorema de Isomorfía, nos da la siguiente cadena:

$$\pi : A \longrightarrow A/\bar{\mathfrak{a}} \cong K[X_1, \dots, X_n]/\mathfrak{a}.$$

En particular, π es un epimorfismo de K -espacios vectoriales. Por el Teorema 1.2.9, concluimos que $K[X_1, \dots, X_n]/\mathfrak{a}$ es un K -espacio vectorial de dimensión finita y una K -álgebra de Artin. Además, observamos que el conjunto siguiente no solamente es un sistema generador del ideal monomial \mathfrak{a} de $K[X_1, \dots, X_n]$, sino que es lo que en [\[CLO, 07\]](#) (Problem 8, Section 4, Chapter 2) se denomina una base minimal del ideal monomial \mathfrak{a} . La razón es que no puede haber relación de divisibilidad entre

ellos porque sus exponentes son los elementos minimales de H^c para \preceq y son distintos dos a dos. Finalmente, observamos que para cualquier orden monomial \leq se ha de tener:

$$(1.2.10) \quad \exp_{\leq}(\mathfrak{a}) = \bigcup_{i=1}^t (\mu_i + \mathbb{N}^n).$$

Un contenido es obvio dado que \mathfrak{a} está generado por los monomios $\mathfrak{m}(H^c) := \{\mu_1, \dots, \mu_t\}$. El otro contenido se sigue del hecho de que un orden monomial \leq induce una graduación sobre el anillo de polinomios $K[X_1, \dots, X_n]$ (i.e., $K[X_1, \dots, X_n] := \bigoplus_{\mu \in \mathbb{N}^n} K\langle X_1^{\mu_1} \dots X_n^{\mu_n} \rangle$), donde $K\langle X_1^{\mu_1} \dots X_n^{\mu_n} \rangle$ es el K -espacio vectorial de dimensión 1 generado por el monomio $X_1^{\mu_1} \dots X_n^{\mu_n}$. Los ideales monomiales son, justamente, ideales homogéneos con respecto a esa graduación y, por tanto, si $f \in \mathfrak{a}$, entonces todos sus términos están en \mathfrak{a} . En particular, $\ell_{t_{\leq}}(f) \in \mathfrak{a}$ y, por tanto, debe ser divisible por alguno de los monomios en $\mathcal{M}_o(H^c)$ (que es lo que viene a decir el Lemma 2, Section 4 of Chapter 2 de [CLO, 07]). Por tanto, para cada $f \in \mathfrak{a}$, se tiene:

$$\exp_{\leq}(f) \in \bigcup_{i=1}^t (\mu_i + \mathbb{N}^n) = \bigcup_{\mu \in \mathfrak{m}(H^c)} (\mu + \mathbb{N}^n).$$

Esto prueba la Identidad (1.2.10) que, gracias a la Igualdad (1.2.9), podemos escribir como:

$$\exp_{\leq}(\mathfrak{a}) = \bigcup_{\mu \in \mathfrak{m}(H^c)} (\mu + \mathbb{N}^n) = H^c.$$

Finalmente, el teorema sobre existencia de Bases de Gröbner (cf. Teorema 1.2.7) nos permite concluir que $\mathcal{B}(H)/\mathfrak{a}$ es una base de $K[X_1, \dots, X_n]/\mathfrak{a}$ como K -espacio vectorial, y el resto se sigue.

- *ii) \implies i):* Para probar esta implicación basta con demostrar que $\exp_{\leq}(\mathfrak{a})^c := \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})$ es siempre cerrado hacia abajo y finito, bajo las hipótesis de *ii*). Consideremos $\mathfrak{b} = (X_1^p, \dots, X_n^p)$ y comencemos observando que el anillo residual $A := K[X_1, \dots, X_n]/\mathfrak{b}$ es ya un anillo de Artin, porque es un espacio vectorial de dimensión p^n cuya base es sencilla de expresar como las clases residuales módulo \mathfrak{b} de los elementos de \mathcal{Y}^n y usando el Teorema 1.2.9 precedente. Como en la implicación precedente, definiendo $\bar{\mathfrak{a}} := \mathfrak{a}/\mathfrak{b}$ como las clases de restos de los elementos de $\mathfrak{a} \supseteq \mathfrak{b}$ módulo el ideal \mathfrak{b} , tenemos claramente un epimorfismo de K -álgebras (Segundo teorema de Isomorfía) $\pi : A \longrightarrow A/\bar{\mathfrak{a}} \cong K[X_1, \dots, X_n]/\mathfrak{a}$. Por tanto, usando el Teorema 1.2.9 concluimos que $\mathcal{B}(H)/\mathfrak{a}$ es base del anillo residual $K[X_1, \dots, X_n]/\mathfrak{a}$ como K -espacio vectorial. En consecuencia, $\exp_{\leq}(\mathfrak{a})^c$ es un conjunto finito (biyectable a $\mathcal{B}(H)/\mathfrak{a}$ para cualquier orden monomial \leq). Queda por ver que $\exp_{\leq}(\mathfrak{a})^c$ es cerrado hacia abajo. Pero eso es consecuencia casi inmediata de combinar la División de Weierstarss-Hironaka (Teorema 1.2.5) con el Teorema 1.2.7 sobre existencia de Bases de Gröbner. Por existir una base de Gröbner del ideal \mathfrak{a} , entonces existen $\{\mu_1, \dots, \mu_s\}$, todos ellos distintos dos a dos y tales que:

$$\exp_{\leq}(\mathfrak{a}) = \bigcup_{i=1}^s (\mu_i + \mathbb{N}^n).$$

Consideremos ahora el conjunto complementario $\exp_{\leq}(\mathfrak{a})^c = \bigcap_{i=1}^s (\mu_i + \mathbb{N}^n)^c$. Veamos que ese conjunto así descrito es cerrado hacia abajo, lo cual es del todo evidente por su forma. Para ello, sean γ, θ tales que $\theta \in \exp_{\leq}(\mathfrak{a})^c$ y $\gamma \preceq \theta$. Entonces, si $\gamma \in \exp_{\leq}(\mathfrak{a})$, como $\gamma \preceq \theta$ concluiríamos que $\mu_i \preceq \gamma \preceq \theta \in \mu_i + \mathbb{N}^n$ para algún $i \in [s]$. Y esto último es imposible por construcción. Por tanto, $\gamma \notin \exp_{\leq}(\mathfrak{a})$ y tendremos que $\exp_{\leq}(\mathfrak{a})^c$ es cerrado hacia abajo y finito (i.e. una escalera finita). \square

En conclusión, dada cualquier escalera finita $H \subseteq \mathcal{Y}^n$ y para cualquier orden monomial, existe un ideal $H := \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})$ y ese ideal se puede elegir que sea un ideal monomial. En la Sección 4.2 le daremos una vuelta a esta idea como se indicaba en la Introducción de este capítulo.

El hiper-grafo de una inclusión (OIG): orientaciones, grado, grado de salida de salida

Índice

2.1.	El hiper-grafo de una inclusión: orientaciones, grado, grado de salida	11
2.2.	Grado y grado de salida promedios, densidad del OIG	13
2.3.	Promedios de grado y grado de salida en conjuntos cerrados hacia abajo	21
2.3.1.	Ejemplo: escalones \mathcal{E}_h : grado y grado de salida promedios	22

2.1. El hiper-grafo de una inclusión: orientaciones, grado, grado de salida

Introducimos ahora la noción del hiper-grafo de una inclusión (OIG, por *One-Inclusion Graph*, que es la terminología al uso en la Bibliografía del tema), aunque es un “casi” hiper-grafo, como veremos. Fue introducido en [Hau, 1995, HLW, 1994] para el aprendizaje binario. Posteriormente, fue extendido al caso de aprendizaje multi-clase en [RBR, 06] y discutido en [DS, 14] y [BCDMY, 22]. Corrigiendo ciertas imprecisiones, seguiremos, sobre todo, las notaciones de [BCDMY, 22].

Con las mismas notaciones que en el capítulo precedente, sea $p \in \mathbb{N}$ un entero positivo, con $p \geq 2$, y consideremos el conjunto finito $\mathcal{Y} := \{0, 1, \dots, p-1\}$. Para cada $i \in \mathbb{N}$, sea $\pi_i : \mathcal{Y}^n \rightarrow \mathcal{Y}^{n-1}$ la proyección que “olvida” la i -ésima coordenada. Es decir,

$$(2.1.1) \quad \begin{aligned} \pi_i : \quad \mathcal{Y}^n &\longrightarrow \mathcal{Y}^{[n] \setminus \{i\}} = \mathcal{Y}^{n-1} \\ (h_1, \dots, h_n) &\longmapsto (h_1, \dots, h_{i-1}, h_{i+1}, \dots, h_n). \end{aligned}$$

Para cada clase de conceptos finita $H \subseteq \mathcal{Y}^n$ denotamos por $\pi_i(H)$ a la imagen de H en la “dirección” de π_i . Evidentemente, $\pi_i(H) = H|_{[n] \setminus \{i\}}$. Recíprocamente, para cada $f \in \mathcal{Y}^{n-1}$, consideramos la fibra sobre $\{f\}$ en la dirección de π_i , es decir, $\pi_i^{-1}(\{f\}) := \{h \in \mathcal{Y}^n : h|_{[n] \setminus \{i\}} = f\}$. Si intersecamos la fibra $\pi_i^{-1}(\{f\})$ con H (i.e. $\pi_i^{-1}(\{f\}) \cap H$), esa intersección es no vacía si y solamente si $f \in \pi_i(H)$. Así, para cada $f \in \pi_i(H)$, definimos los siguientes subconjuntos de H , que serán útiles para definir las aristas de nuestro OIG:

$$(2.1.2) \quad e_{f,i} := \pi_i^{-1}(\{f\}) \cap H := \left\{ h \in H : h|_{[n] \setminus \{i\}} = f \right\}.$$

Dada una dirección determinada por π_i , consideramos el conjunto de aristas en la dirección i -ésima:

$$E_i(H) := \{e_{f,i} : f \in \pi_i(H)\}.$$

Las fibras en la dirección i -ésima determinan una partición de H , lo que nos permite concluir:

$$(2.1.3) \quad \sharp(E_i(H)) = \sharp(\pi_i(H))$$

Finalmente, definimos el conjunto de aristas del grafo de una inclusión (OIG) determinado por H (y las proyecciones) como la unión disjunta de las clases $E_i(H)$:

$$(2.1.4) \quad E(H) := \bigsqcup_{i \in [n]} E_i(H),$$

donde el símbolo \sqcup denota unión disjunta. Recordemos que una forma clásica de representar esta unión disjunta es la usada en [BCDMY, 22], que reproducimos aquí. Para cada $i \in [n]$ y cada $f \in \pi_i(H)$, consideremos el producto cartesiano: $e_{f,i} \times \{i\} \subseteq H \times [n]$. Entonces, podemos identificar (poniendo siempre cuidado en nuestro uso de esa identificación):

$$E_i(H) \equiv \{e_{f,i} \times \{i\} : f \in \pi_i(H)\}, \quad \text{y} \quad E(H) \equiv \bigcup_{i \in [n]} \{e_{f,i} \times \{i\} : f \in \pi_i(H)\}.$$

Dado $e = e_{f,i} \times \{i\} \in E(H)$ y dado $h \in H$, escribiremos $h \in e$ en lugar de $h \in e_{f,i}$ y denotaremos el cardinal de una arista como el cardinal de su fibra subyacente: $\sharp(e) := \sharp(e_{f,i})$. Una arista $e \in E(H)$ se denomina unaria

(o, siendo un tanto anglosajones, diremos *singleton*) si $\#(e) = 1$. Denotaremos por $E^{(1)}(H)$ a la clase de aristas de $E(H)$ que son aristas unarias.

DEFINICIÓN 11 (HIPER-GRAFO DE UNA INCLUSIÓN (OIG)). *Con estas notaciones, definimos el hiper-grafo de una inclusión (OIG) asociado a una clase de conceptos $H \subseteq \mathcal{Y}^n$ como el hiper-grafo $G(H) := (H, E(H))$, donde $E(H)$ es la clase de aristas definida en la Identidad (2.1.4).*

OBSERVACIÓN 2.1.1 (OIG NO ES PROPIAMENTE UN HIPER-GRAFO: LAS ARISTAS UNARIAS PUEDEN SER “CONTADAS” VARIAS VECES). Aunque hemos usado el término hiper-grafo, esta estructura $G(H)$ que acabamos de definir no es propiamente un hiper-grafo. Algunas de las aristas pueden ser consideradas como objetos distintos dependiendo de la dirección tenida en cuenta y, por tanto, contadas varias veces. Dos aristas $e := e_{f,i} \times \{i\} \in E(H)$ y $e' := e_{g,j} \times \{j\} \in E(H)$ con $i \neq j$ y la misma fibra subyacente, contenida en H (i.e. $e_{i,f} = e_{g,j}$), pueden ser aristas unarias y, sin embargo, se consideran como aristas distintas. Es decir, puede existir $h \in H$ tal que $e = \{h\} \times \{i\}$ y $e' = \{h\} \times \{j\}$, pero e y e' se considerarán aristas distintas de nuestro hiper-grafo por ser $i \neq j$. Los ejemplos son fácilmente visualizables simplemente tomando $H := \{(1,1)\} \subseteq \{0,1\}^2$ y considerando las dos aristas distintas $e := \{(1,1)\} \times \{1\}$ y $e' := \{(1,1)\} \times \{2\}$. Obviamente, este abuso de lenguaje se arreglaría considerando $\mathcal{G}(H) := (H \times [n], E(H))$ en lugar del modo en que lo acabamos de hacer. Pero sería un poco más engorroso y hemos tratado, con la mejor de las intenciones, seguir las notaciones confusas de [BCDMY, 22]. De ahí el abuso de lenguaje usado en la definición precedente por el que pedimos disculpas al lector.

OBSERVACIÓN 2.1.2 (ESTE OIG DIFIERE DE LA ESTRUCTURA DE GRAFO INTRODUCIDA EN [Hau, 1995]). En [Hau, 1995], una estructura de grafo distinta se asoció a clases de conceptos finitas $H \subseteq \{0,1\}^n$ (i.e. en el caso binario). La estructura introducida por D. Haussler difiere de la que acabamos de introducir, conforme a [BCDMY, 22]. Para ser completos, vamos a compararlas.

En [Hau, 1995], el autor considera una métrica proporcional a la distancia de Hamming. Es decir, dados $f, g \in H$, Haussler define la función:

$$\rho(f, g) := \frac{1}{n} \sum_{i \in [n]} |f(i) - g(i)| = \frac{d_{\text{Ham}}(f, g)}{n},$$

donde $d_{\text{Ham}}(f, g)$ es la distancia de Hamming o, equivalentemente, donde:

$$|f(i) - g(i)| := \begin{cases} 1, & \text{si } f(i) \neq g(i) \\ 0, & \text{n otro caso.} \end{cases}$$

Podríamos reproducir la misma métrica en el caso general tomando $H \subseteq \mathcal{Y}^n$, donde $\mathcal{Y} = \{0, \dots, p-1\}$, y considerando la métrica:

$$\rho : H \times H \longrightarrow \mathbb{R}_+ \\ (f, g) \longmapsto \frac{d_{\text{Ham}}(f, g)}{n}.$$

Podemos considerar, para un $h \in H$ dado, la bola cerrada de radio $1/n$ centrada en h con respecto a esta distancia ρ :

$$\overline{B}_\rho(h, 1/n) := \{g \in H : \rho(h, g) \leq 1/n\}.$$

Es un sencillo ejercicio el verificar que esta bola cerrada es dada por la siguiente igualdad:

$$\overline{B}_\rho(h, 1/n) = \bigcup_{i \in [n]} e_{\pi_i(h), i}.$$

Esto, obviamente, implica: $\#(\overline{B}_\rho(h, 1/n)) = (\sum_{i=1}^n \#(e_{\pi_i(h), i})) - (n-1)$ (recordemos que $h \in e_{\pi_i(h), i}, \forall i \in [n]$). En el caso $p = 2$, las fibras (y sus correspondientes aristas) de la forma $e_{\pi_i(h), i}$ sólo tienen a lo sumo dos elementos y [Hau, 1995] define una estructura de grafo sobre H definiendo las aristas en la forma siguiente: $\text{Haussl}(H) := \{(g, h) \in H^2 : \rho(g, h) = 1/n\}$. Nótese que un par $(h, g) \in \text{Haussl}(H)$ si y solamente si $g \in \overline{B}_\rho(h, 1/n) \setminus \{h\}$. En particular, en el caso $p = 2$ tenemos una biyección entre los conjuntos:

$$\text{Fan} - \text{Out}_{\text{Haussl}}(h) := \{g \in H : (h, g) \in \text{Haussl}(H)\},$$

y los conjuntos de aristas conteniendo a h y con exactamente dos elementos distintos, i.e.

$$\{e \in E(H) : h \in e, \#(e) = 2\}.$$

En particular, el cardinal del fan-out de h en $\text{Haussl}(H)$ coincide con lo que denominaremos más adelante el grado de h en $E(H)$: $\#(\text{Fan} - \text{Out}_{\text{Haussl}}(h)) = \deg(h)$. Esto también significa que, en el caso $p = 2$, cada arista $e_{f,i}$ es contada dos veces y ninguna es elemento de $E^{(1)}(H)$. En otras palabras, concluimos que :

$$(2.1.5) \quad \#(\text{Haussl}(H)) = 2 \left(\#(E(H) \setminus E^{(1)}(H)) \right).$$

Los siguientes resultados establecen las propiedades más elementales del hiper-grafo $G(H)$:

LEMA 2.1.3. *Con estas notaciones, las siguientes propiedades se verifican:*

- i) Para cada $h \in H$, denotemos por $E(H)_h$ al conjunto formado por las aristas de $E(H)$ que contienen a h . Es decir, $E(H)_h := \{e : h \in e\}$. Entonces, se verifica $\sharp(E(H)_h) = n$.
- ii) Para cada $i \in [n]$, $\sharp(E_i(H)) = \sharp(\pi_i(H))$. Luego también se tiene:

$$\sharp(E(H)) = \sum_{i \in [n]} \sharp(\pi_i(H)).$$

- iii) Para cada $i \in [n]$, la siguiente descomposición describe una partición de H en la dirección de π_i :

$$H := \bigsqcup_{f \in \pi_i(H)} e_{f,i}.$$

Como se trata de una unión disjunta, se tendrá también:

$$\sharp(H) = \sum_{f \in \pi_i(H)} \sharp(e_{f,i}).$$

DEMOSTRACIÓN. Son todos argumentos inmediatos a partir de nuestras definiciones y los omitiremos. \square

DEFINICIÓN 12 (ORIENTACIONES EN EL OIG). *Una orientación en el hiper-grafo $G(H)$ es una aplicación $\sigma : E(H) \rightarrow H$, tal que se verifica la siguiente propiedad: $\forall e \in E(H)$, $\sigma(e) \in e$.*

Introduzcamos otra partición del conjunto de aristas $E(H)$ del modo siguiente: para cada $k \in \mathbb{N}$, denotemos mediante $E^{(k)}(H)$ al conjunto de aristas cuyas fibras subyacentes tienen cardinal k , esto es:

$$(2.1.6) \quad E^{(k)}(H) := \{e \in E(H) : \sharp(e) = k\}.$$

Como cada arista tiene a lo sumo p elementos (i.e. $\sharp(e_{f,i}) \leq \sharp(\mathcal{Y}) = p$), la siguiente es una partición de $E(H)$ en función de los cardinales de las aristas:

$$E(H) := \bigsqcup_{k=1}^p E^{(k)}(H).$$

DEFINICIÓN 13 (GRADO, GRADO DE SALIDA). *Con estas notaciones definimos:*

- Para cada $h \in H$, el grado de h en $G(H)$ es dado mediante:

$$\deg_H(h) := \sharp(\{e \in E(H) : h \in e, e \notin E^{(1)}(H)\}) = n - \sharp(\{e \in E^{(1)}(H) : h \in e\}).$$

- Para cada $h \in H$ y para cada orientación $\sigma : E(H) \rightarrow H$, definimos el grado de salida de σ en h mediante: $\text{outdeg}_H(\sigma, h) := \sharp(\{e \in E(H) : h \in e, \sigma(e) \neq h\})$.

NOTACIÓN 2.1.4. Como es habitual, omitiremos el subíndice H de \deg_H y outdeg_H y simplemente escribiremos $\deg(h)$ y $\text{outdeg}(h)$, asumiendo que H y $G(H)$ son fijados por el contexto. Sin embargo, es obvio que ambas nociones dependen fuertemente del hiper-grafo $G(H)$ y, en particular, de H , como veremos en discusiones posteriores.

2.2. Grado y grado de salida promedios, densidad del OIG

En esta sección nos ocuparemos de las primeras propiedades de las nociones antes discutidas del OIG y sus elementos. Nos interesan especialmente los valores medios *con respecto a la distribución uniforme sobre H* .

DEFINICIÓN 14 (GRADO PROMEDIO, GRADO DE SALIDA PROMEDIO (UNIFORME) Y DENSIDAD DEL HIPER-GRAFO). *Con las notaciones de la sección precedente, introducimos nuestros primeros invariantes cuantitativos relacionados con el hiper-grafo $G(H)$, teniendo en cuenta siempre una distribución de probabilidad uniforme sobre el conjunto de clases de conceptos H .*

- i) El grado promedio del conjunto de los vértices $h \in H$ (con respecto a la distribución uniforme en H) es dado por la siguiente igualdad:

$$(2.2.1) \quad \text{avd}(H) := \frac{1}{\sharp(H)} \sum_{h \in H} \deg(h).$$

- ii) El grado de salida promedio (con respecto a la distribución uniforme en H) de una orientación $\sigma : E(H) \rightarrow H$ es dado por la siguiente identidad:

$$(2.2.2) \quad \text{outdeg}_{\text{av}}(\sigma) := \frac{1}{\sharp(H)} \sum_{h \in H} \text{outdeg}(\sigma, h).$$

iii) La densidad del hiper-grafo $G(H)$ (terminología tomada de [RBR, 06]) es definida por el siguiente número racional:

$$\text{gd}(H) := \frac{\#(E(H))}{\#(H)}.$$

iv) El máximo de los grados de salida de una orientación es definido mediante la siguiente expresión:

$$\text{outdeg}_{\max}(H) := \max\{\text{outdeg}(\sigma, h) : h \in H\}.$$

Nótese que $\text{avd}(H)$ sólo depende de $G(H)$, mientras que $\text{outdeg}_{\text{av}}(\sigma)$ parece depender de la orientación $\sigma : E(H) \rightarrow H$. Nuestra primera tarea será la de caracterizar todas estas expresiones en términos del grafo $G(H)$ y, en particular, probaremos que el grado de salida promedio no depende de la orientación elegida, sino que es un invariante de $G(H)$ que sólo depende de la dimensión del espacio “ambiente” (i.e. de n) y de la densidad del grafo $G(H)$, conforme a [RBR, 06]. Este es un resultado original que resulta ser uno de los puntos de partida del TFG.

TEOREMA 2.2.1 (PRINCIPALES CARACTERIZACIONES DE avd , $\text{outdeg}_{\text{av}}$ Y $\text{gd}(H)$). *Con estas notaciones, suponiendo $n \geq 2$, se tiene:*

i) El grado promedio del hiper-grafo $G(H)$ satisface:

$$\text{avd}(H) = n - \frac{\#(E^{(1)}(H))}{\#(H)}.$$

ii) El grado de salida promedio de cualquier orientación $\sigma : E(H) \rightarrow H$ satisface:

$$\text{outdeg}_{\text{av}}(\sigma) = n - \text{gd}(H) = n - \frac{\#(E(H))}{\#(H)}.$$

En particular, el grado de salida promedio de una orientación es independiente de la orientación y sólo depende de H . De hecho, sólo depende de la dimensión n del espacio ambiente y de la densidad del hiper-grafo $\text{gd}(H)$. Por ello, a partir de ahora hablaremos simplemente del grado de salida promedio del hiper-grafo y lo denotaremos mediante $\text{outdeg}_{\text{av}}(H)$. En particular, se tiene:

$$\text{avd}(H) - \text{outdeg}_{\text{av}}(H) - \text{gd}(H) = -\frac{\#(E^{(1)}(H))}{\#(H)} \leq 0.$$

iii) Lo siguiente también se verifica: $\text{outdeg}_{\text{av}}(\sigma) \leq \text{avd}(H) \leq 2 \text{outdeg}_{\text{av}}(\sigma)$.

iv) Para cada orientación $\sigma : E(H) \rightarrow H$, se tiene:

$$\text{gd}(H) - \frac{\#(E^{(1)}(H))}{\#(H)} \leq \text{outdeg}_{\text{av}}(\sigma).$$

v) Sea $G \subseteq H$ un subconjunto de H y sea $\sigma : E(H) \rightarrow H$ una orientación. Entonces, existe una orientación $\sigma' : E(G) \rightarrow G$ tal que se verifica: $\forall g \in G, \text{outdeg}(\sigma', g) \leq \text{outdeg}(\sigma, g)$. En particular, para toda orientación $\sigma : E(H) \rightarrow H$, existirá una orientación $\sigma' : E(G) \rightarrow G$ tal que:

$$\text{outdeg}_{\max}(\sigma') \leq \text{outdeg}_{\max}(H).$$

Finalmente, el grado de salida promedio raramente coincide con el grado promedio, porque se tiene:

- O bien $\text{outdeg}_{\text{av}}(H) < \text{avd}(H)$,
- o bien $E(H) = E^{(1)}(H)$ (i.e. $G(H)$ tiene solamente aristas unarias). En ese caso,

$$\text{outdeg}_{\text{av}}(H) = \text{avd}(H) = 0.$$

DEMOSTRACIÓN. Probamos cada afirmación separadamente:

- Afirmación i): observemos que, combinando la Definición 13 y la Definición 14, se tiene:

$$\text{avd}(H) = \frac{1}{\#(H)} \sum_{h \in H} \text{deg}(h) = \frac{1}{\#(H)} \sum_{h \in H} \left(n - \#(\{e \in E^{(1)}(H) : h \in e\}) \right),$$

y que, tras separar los términos del sumatorio, se tiene:

$$\text{avd}(H) = \frac{1}{\#(H)} \left(n\#(H) - \sum_{h \in H} \#(\{e \in E^{(1)}(H) : h \in e\}) \right).$$

Por lo que es sencillo concluir la Afirmación i), puesto que el sumatorio restante es equivalente a contar el número de aristas unarias:

$$\text{avd}(H) = n - \frac{\#(E^{(1)}(H))}{\#(H)}$$

- *Afirmación ii*): expresando el grado de salida a partir del conjunto complementario que lo define:

$$\begin{aligned}\text{outdeg}(\sigma, h) &= \#(\{e \in E(H)_h : \sigma(e) = h\}), \\ &= \#(E(H)_h \setminus \{e \in E(H)_h : \sigma(e) = h\})\end{aligned}$$

podemos simplificar la suma de los grados de salida, teniendo en cuenta que siempre se cumple que $\#(E(H)_h) = n$, y que la siguiente unión de fibras define una partición de $E(H)$:

$$E(H) = \bigcup_{h \in H} \sigma^{-1}(\{h\}) = \bigcup_{h \in H} \{e \in E(H)_h : \sigma(e) = h\},$$

es decir, se tiene:

$$\begin{aligned}\sum_{h \in H} \text{outdeg}(\sigma, h) &= \sum_{h \in H} (\#(E(H)_h) - \#(\sigma^{-1}(h))) = \sum_{h \in H} \#(E(H)_h) - \sum_{h \in H} \#(\sigma^{-1}(h)) = \\ &= n\#(H) - \#(E(H))\end{aligned}$$

Dividiendo la igualdad por $\#(H)$ se obtiene la igualdad de la Afirmación *ii*). El resto se sigue con aplicar la Afirmación *i*) y lo que acabamos de probar.

- *Afirmación iii*): el resultado es inmediato si probamos las siguientes igualdades:

$$(2.2.3) \quad \text{avd}(H) = \frac{1}{\#(H)} \sum_{e \in E_{>1}^{(1)}(H)} \#(e),$$

$$(2.2.4) \quad \text{outdeg}_{\text{av}}(H) = \frac{1}{\#(H)} \sum_{e \in E_{>1}^{(1)}(H)} (\#(e) - 1),$$

donde $E_{>1}^{(1)}(H)$ es el conjunto de aristas con cardinal estrictamente mayor que 1, es decir, las aristas no unarias. Para verlo, observemos que la primera desigualdad es inmediata con aplicar (2.2.3) y (2.2.4):

$$\text{outdeg}_{\text{av}}(H) = \frac{1}{\#(H)} \sum_{e \in E_{>1}^{(1)}(H)} (\#(e) - 1) \leq \frac{1}{\#(H)} \sum_{e \in E_{>1}^{(1)}(H)} \#(e) = \text{avd}(H),$$

y la segunda desigualdad se sigue con observar que $\forall e \in E_{>1}^{(1)}(H)$ tenemos que $2(\#(e) - 1) \geq \#(e)$, lo que, conjuntamente con (2.2.3) y (2.2.4), nos permite concluir:

$$2 \text{outdeg}_{\text{av}}(H) = \frac{1}{\#(H)} \sum_{e \in E_{>1}^{(1)}(H)} 2(\#(e) - 1) \geq \frac{1}{\#(H)} \sum_{e \in E_{>1}^{(1)}(H)} \#(e) = \text{avd}(H).$$

Finalmente, para completar la prueba, demostramos las igualdades (2.2.3) y (2.2.4). Para ello, consideramos la función característica¹ $\chi_e(h)$ de una arista e , obteniendo así una expresión equivalente de $\text{avd}(H)$:

$$\text{avd}(H) = \sum_{h \in H} \text{deg}(h) = \sum_{h \in H} \sum_{e \in E_{>1}^{(1)}(H)} \chi_e(h),$$

tras lo que alternamos el orden de los sumatorios para concluir la Igualdad (2.2.3):

$$\text{avd}(H) = \sum_{e \in E_{>1}^{(1)}(H)} \sum_{h \in H} \chi_e(h) = \sum_{e \in E_{>1}^{(1)}(H)} \#(e).$$

Para la Igualdad (2.2.4), también consideramos la función complementaria² a la Delta de Dirac $\bar{\delta}(h, h')$ para elementos de H . De esta manera, podemos obtener la siguiente expresión equivalente de $\text{outdeg}_{\text{av}}(H)$:

$$\text{outdeg}_{\text{av}}(H) = \sum_{h \in H} \#(\{e \in E(H)_h : h \neq \sigma(e)\}) = \sum_{h \in H} \sum_{e \in E_{>1}^{(1)}(H)} \bar{\delta}(h, \sigma(e)) \cdot \chi_e(h),$$

donde σ es una orientación arbitraria (recordemos que $\text{outdeg}_{\text{av}}(H)$ es independiente de la orientación considerada). Alternando el orden de los sumatorios, concluimos:

$$\text{outdeg}_{\text{av}}(H) = \sum_{e \in E_{>1}^{(1)}(H)} \sum_{h \in H} \bar{\delta}(h, \sigma(e)) \cdot \chi_e(h) = \sum_{e \in E_{>1}^{(1)}(H)} (\#(e) - 1)$$

¹Es decir, $\chi_e(h) = 1$ si $h \in e$, y $\chi_e(h) = 0$ si $h \notin e$.

²Es decir, $\bar{\delta}(h, h') = 0$ si $h = h'$ y $\bar{\delta}(h, h') = 1$ si $h \neq h'$.

- *Afirmación iv)*: a partir de las Afirmaciones *i)* y *ii)* concluimos:

$$(2.2.5) \quad \text{avd}(H) = \text{outdeg}_{\text{av}}(H) + \frac{\#(E(H)) - \#(E^{(1)}(H))}{\#(H)}.$$

Empleando la segunda desigualdad de la Afirmación *iii)*, concluimos:

$$\text{outdeg}_{\text{av}}(H) + \frac{\#(E(H)) - \#(E^{(1)}(H))}{\#(H)} \leq 2\text{outdeg}_{\text{av}}(H),$$

que, tras la simplificación pertinente, nos conduce a la desigualdad buscada.

- *Afirmación v)*: aunque sospechamos que se trata de un resultado conocido, demostraremos este resultado que tendrá algunas aplicaciones interesantes más adelante. Primero de todo, observemos la relación existente entre los conjuntos de aristas $E(H)$ y $E(G)$. Dada una arista $e := e_{f,i} \times \{i\} \in E(H)$, donde $e_{f,i}$ es el definido en la Identidad (2.1.2), denotamos mediante $e \cap G := (e_{i,f} \cap G) \times \{i\} \in E(G)$, siempre que $f \in \pi_i(G) \subseteq \pi_i(H)$ (o, equivalentemente, siempre que $e_{i,f} \cap G \neq \emptyset$). Entonces, tenemos: *Dado $e' \in E(G)$ hay una única arista $e \in E(H)$ tal que $e' = e \cap G \neq \emptyset$.*

Esto es una consecuencia inmediata de la construcción de los hiper-grafos de una inclusión: Dada $e' \in E(G)$ existe $i \in [n]$, $f \in \pi_i(G) \subseteq \pi_i(H) \subseteq \mathcal{Y}^n$ tal que $e' := e'_{i,f} \times \{i\}$, donde $e'_{i,f} := \{g \in G : g|_{[n] \setminus \{i\}} = f\}$. Por tanto, claramente existe un $e \in E(H)$ tal que $e' = e \cap G$ y esta arista $e \in E(H)$ tiene la forma única $e := e_{f,i} \times \{i\} \in E(H)$, porque $f \in \mathcal{Y}^{n-1}$ está predeterminado en todas sus coordenadas y es independiente de G y de H .

Por tanto, estamos en condiciones de definir $\sigma' : E(G) \rightarrow G$ en términos de σ . Distinguiremos los casos siguientes:

- Si $e' \in E(G)$ es tal que existe $e \in E(H)$, con $e' := e \cap G$ y $e \in \sigma^{-1}(G) \subseteq E(H)$. En ese caso definimos $\sigma'(e') := \sigma(e)$. Está bien definida porque e es única a partir de e' , mediante la identidad $e' := e \cap G$, como hemos discutido anteriormente.
- En otro caso, sucede lo siguiente: $e' \in E(G)$ es tal que existe $e \in E(H) \setminus \sigma^{-1}(G)$ y $e' := e \cap G$. En este caso, $e = e_{f,i} \times \{i\}$, $\sigma(e) \notin G$ y $e' := (e_{f,i} \cap G) \times \{i\} \in E(G)$. Pero, entonces $e_{f,i} \cap G \neq \emptyset$ y, por tanto, podemos elegir un punto $g_e \in e_{f,i} \cap G$. Como e es única para cada e' en este caso, podemos definir $\sigma'(e') := g_e$.

Tenemos finalmente las notaciones necesarias para expresar nuestro argumento. Supongamos $g \in G$ y definamos:

$$\begin{aligned} O(\sigma, g) &:= \{e \in E(H) : g \in e, g \neq \sigma(e)\}. \\ O(\sigma', g) &:= \{e' \in E(G) : g \in e', g \neq \sigma'(e')\}. \end{aligned}$$

Descompongamos entonces:

$$O(\sigma', g) = O_1(\sigma', g) \sqcup O_2(\sigma', g),$$

donde

$$\begin{aligned} O_1(\sigma', g) &= \{e' \in E(G) : \exists e \in \sigma^{-1}(G), g \in e' = e \cap G, g \neq \sigma'(e') = \sigma(e)\}, \\ O_2(\sigma', g) &= \{e' \in E(G) : \exists e \in E(H) \setminus \sigma^{-1}(G), g \in e' = e \cap G, g \neq g_e\}. \end{aligned}$$

Tenemos así una aplicación inyectiva: $\varphi : O(\sigma', g) \rightarrow O(\sigma, g)$, dada mediante $\varphi(e') := e$, donde $e' := e \cap G$. Ya hemos visto que e es única para cada e' y φ es claramente inyectiva por construcción. Sólo nos falta por demostrar que φ es una aplicación probando que su rango es el apropiado (i.e. $\varphi(O(\sigma', g)) \subseteq O(\sigma, g)$):

- Si $e' \in O_1(\sigma, g)$, entonces la única arista $e \in E(H)$ tal que $e' = e \cap G$, satisface $\sigma(e) \neq g$ y, por tanto, $\varphi(e') \in O(\sigma, g)$.
- Si $e' \in O_2(\sigma, g)$, entonces, obviamente, la única arista $e \in E(H)$ tal que $e' = e \cap G$, satisface que $\sigma(e) \notin G$ y, en particular, $\sigma(e) \neq g$, luego $\varphi(e') \in O(\sigma, g)$.

En conclusión, tenemos que se verifica: $\text{outdeg}(\sigma', g) = \#(O(\sigma', g)) \leq \#(O(\sigma, g)) = \text{outdeg}(\sigma, g)$, y la Afirmación *v)* queda demostrada.

En lo que respecta a la dicotomía final, se sigue inmediatamente de las Afirmaciones *i)* y *ii)*, puesto que $\text{outdeg}_{\text{av}}(H) = \text{avd}(H)$ si y solamente si $E(H) = E^{(1)}(H)$ y, en ese caso, $\#(E^{(1)}(H)) = n\#(H)$ (o, equivalentemente, $\text{avd}(H) = 0$). \square

OBSERVACIÓN 2.2.2 (EL CASO $n = 1$). En el caso $n = 1$ algunas de estas nociones carecen de significado apropiado. Si $n = 1$, no hay proyecciones que olviden la 1-ésima coordenada. Más aún, si $n = 1$ y $\#(H) = 1$ tenemos un cubo canónico de dimension 1 y ambos promedios (grado y grado de salida) son 0, mientras que la densidad del grafo es 1.

OBSERVACIÓN 2.2.3. Nótese que, en la demostración del último teorema, hemos usado la Identidad (2.2.4). Esta identidad indica que nuestro grado de salida promedio $\text{outdeg}_{\text{av}}(H)$ coincide con la cantidad avd' que aparece implícitamente en medio de una demostración (y sin ningún tipo de estudio o análisis) en [BCDMY, 22].

EJEMPLO 2.2.4 (LOS “PROMEDIOS” DEPENDEN DE LA DISTRIBUCIÓN UNIFORME SOBRE H). Las Afirmaciones i) y ii) del teorema previo dependen fuertemente de haber elegido la distribución uniforme en H :

EJEMPLO 1: En este ejemplo $\text{avd}(H)$ no cambia si modificamos la distribución de probabilidad sobre H . Por otro lado, el mismo ejemplo muestra que el grado de salida promedio de las orientaciones cambia cuando cambiamos la distribución de probabilidad. Adicionalmente, la independencia del valor promedio cambia cuando se cambian las orientaciones. El ejemplo es un cubo (discreto) de dimensión 2, esto es, tomamos $H := \{(0, 0), (1, 0), (0, 1), (1, 1)\} = \{0, 1\}^2$. Las aristas en $E(H)$ son aquellos subconjuntos de H de cardinal 2, $E(H) = \{e_1, e_2, e_3, e_4\}$, dados mediante las siguientes identidades:

$$e_1 = \{(0, 1), (1, 1)\}, e_2 = \{(1, 0), (0, 0)\}, e_3 = \{(0, 1), (0, 0)\}, e_4 = \{(1, 0), (1, 1)\}.$$

El grado en cada $h \in H$ es 2 puesto que cada vértice está en exactamente dos aristas del grafo. Por tanto, cualquiera que sea la distribución de probabilidad elegida sobre H , tendremos que $\text{avd}(H) = 2$. Por otro lado, podemos construir dos orientaciones de $G(H)$, a saber, $\sigma_1, \sigma_2 : E(H) \rightarrow H$, dadas de la forma siguiente:

$$\begin{array}{ccc|ccc} \sigma_1 : E(H) & \longrightarrow & H & & \sigma_2 : E(H) & \longrightarrow & H \\ e_1 & \mapsto & (0, 1) & & e_1 & \mapsto & (1, 1) \\ e_2 & \mapsto & (0, 0) & & e_2 & \mapsto & (1, 0) \\ e_3 & \mapsto & (0, 0) & & e_3 & \mapsto & (0, 1) \\ e_4 & \mapsto & (1, 0) & & e_4 & \mapsto & (1, 1) \end{array}$$

Obviamente tenemos que se verifica:

$$\text{outdeg}(\sigma_1, (0, 1)) = \text{outdeg}(\sigma_1, (1, 0)) = 1, \text{outdeg}(\sigma_1, (0, 0)) = 0, \text{outdeg}(\sigma_1, (1, 1)) = 2,$$

y también se verifica:

$$\text{outdeg}(\sigma_2, (0, 1)) = \text{outdeg}(\sigma_2, (1, 0)) = 1, \text{outdeg}(\sigma_2, (0, 0)) = 2, \text{outdeg}(\sigma_2, (1, 1)) = 0.$$

Como se prueba en el teorema precedente, suponiendo la distribución uniforme en H , tenemos:

$$\text{outdeg}_{\text{av}}(\sigma_1) = \text{outdeg}_{\text{av}}(\sigma_2) = \text{outdeg}_{\text{av}}(H) = 1.$$

Pero podemos elegir otra distribución de probabilidad sobre H cuyas frecuencias difieran en los puntos $(0, 0)$ y $(1, 1)$ como, por ejemplo, las siguientes:

$$f_{(0,0)} = \frac{1}{8}, f_{(1,0)} = f_{(0,1)} = \frac{1}{4}, f_{(1,1)} = \frac{3}{8}.$$

En ese caso tendremos que se verifica::

$$\frac{5}{4} = \sum_{h \in H} \text{outdeg}(\sigma_1, h) f_h \neq \sum_{h \in H} \text{outdeg}(\sigma_2, h) f_h = \frac{3}{4}.$$

En conclusión, con cambios en la distribución de probabilidad el grado de salida promedio puede variar con la orientación.

EJEMPLO 2: Tampoco el grado promedio (más común en los estudios al uso) es un invariante de la distribución de probabilidad elegida sobre H , como se prueba en el ejemplo siguiente.

Supongamos que H es dado por la siguiente igualdad $H := \{(0, 0), (1, 0), (0, 1)\}$, entonces, las aristas del hipergrafo son dadas por $E(H) = \{e_1, e_2, e_3, e_4\}$, donde:

$$e_1 = \{(0, 0), (0, 1)\}, e_2 = \{(0, 0), (1, 0)\}, e_3 = \{(0, 1)\}, e_4 = \{(1, 0)\}.$$

Los grados son, por tanto $\text{deg}((0, 0)) = 2, \text{deg}((0, 1)) = \text{deg}((1, 0)) = 1$. Suponiendo la distribución uniforme sobre H , tendremos:

$$\text{avd}(H) = \frac{4}{3}.$$

Si consideramos otras distribuciones de probabilidad sobre H , como, por ejemplo, la dada por las siguientes frecuencias:

$$f_{(0,0)} = \frac{1}{8}, f_{(1,0)} = \frac{3}{8}, f_{(0,1)} = \frac{4}{8},$$

tenemos que el grado promedio con respecto a esta distribución verifica:

$$\sum_{h \in H} \text{deg}(h) f_h = \frac{9}{8} \neq \frac{4}{3} = \text{avd}(H).$$

En suma, el grado promedio del hiper-grafo también cambia cuando modificamos la distribución de probabilidad sobre H . Y el grado de salida promedio pasa a depender de la orientación cuando se modifica la distribución de probabilidad sobre el conjunto H de vértices.

COROLARIO 2.2.5. *Con las notaciones precedentes, para cualquier hiper-grafo de una inclusión, $G(H)$, y para cada orientación $\sigma : E(H) \rightarrow H$ se verifican las siguientes desigualdades:*

$$\text{gd}(H) - \frac{\sharp(E^{(1)}(H))}{\sharp(H)} \leq \text{outdeg}_{\text{av}}(H) \leq \text{outdeg}_{\text{max}}(\sigma).$$

COROLARIO 2.2.6 (EL MODELO DE HAUSSLER EN EL CASO $p = 2$). *Con las notaciones precedentes, en el caso binario (i.e. el caso $p = 2$), se tiene:*

$$(2.2.6) \quad \text{avd}(H) = \text{outdeg}_{\text{av}}(H) + \frac{\sharp(\text{Haussl}(H))}{2\sharp(H)} \leq 2 \text{outdeg}_{\text{av}}(H),$$

donde $\text{Haussl}(H)$ es el conjunto de aristas de la estructura de grafo introducida por D. Haussler en [Hau, 1995]. En particular, en el caso $p = 2$, la diferencia entre el grado promedio de H y el grado de salida promedio de H es un medio de la densidad del grafo de Haussler en [Hau, 1995].

El siguiente resultado muestra cómo el grado promedio y el grado de salida promedio varían cuando se elimina un vértice $h \in H$, pero se preservan las distribuciones de probabilidad uniforme en H y $H' := H \setminus \{h\}$. En particular, probamos que ninguno de los dos invariantes es monótono con respecto a la inclusión de conjuntos.

LEMA 2.2.7. *Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita tal que $\sharp(H) \geq 2$ y $n \geq 2$. Sea $h \in H$ cualquier elemento. Definamos $H' := H \setminus \{h\}$ como la clase de los elementos de H distintos de h . Denotemos mediante $\text{deg}(g)$ el grado de un elemento $g \in H'$ en el hiper-grafo $G(H)$ y mediante $\text{deg}'(g)$ su correspondiente grado en el hiper-grafo $G(H')$. Entonces, se verifica:*

$$(2.2.7) \quad \text{deg}'(g) \leq \text{deg}(g), \quad \forall g \in H'.$$

DEMOSTRACIÓN. Aunque es una afirmación relativamente evidente, incluimos una prueba. Obsérvese que las aristas que contienen a $g \in H'$ en $G(H')$ son dadas mediante:

$$e'_{\pi_i(g),i} := \{t \in H' : t|_{[n] \setminus \{i\}} = \pi_i(g)\} \neq \emptyset.$$

Denotemos por $e_{\pi_i(g),i}$ las aristas que contienen a $g \in H' \subseteq H$, pero en el hiper-grafo $G(H)$, i.e.

$$e_{\pi_i(g),i} := \{t \in H : t|_{[n] \setminus \{i\}} = \pi_i(g)\} \neq \emptyset.$$

Entonces, tenemos varios casos posibles:

- *Caso 1:* Si $h \in e_{\pi_i(g),i}$ y $e'_{\pi_i(g),i} \neq \emptyset$, entonces $e'_{\pi_i(g),i} = e_{\pi_i(g),i} \setminus \{h\}$, y, por tanto,

$$\sharp(e'_{\pi_i(g),i}) = \sharp(e_{\pi_i(g),i}) - 1.$$

En este caso, $\sharp(e_{\pi_i(g),i}) \geq 2$. Obsérvese también que, en este caso, puede ocurrir que $\sharp(e'_{\pi_i(g),i}) = 1$, siempre que se tenga $\sharp(e_{\pi_i(g),i}) = 2$ o, equivalentemente, que $e_{\pi_i(g),i} = \{g, h\}$.

- *Caso 2:* Si $h \notin e_{\pi_i(g),i}$ y $\sharp(e'_{\pi_i(g),i}) \geq 2$, entonces $e'_{\pi_i(g),i} = e_{\pi_i(g),i}$, y, por tanto,

$$\sharp(e'_{\pi_i(g),i}) = \sharp(e_{\pi_i(g),i}).$$

En este caso tendremos: $\sharp(e_{\pi_i(g),i}) \geq 2$.

- *Caso 3:* Si $h \notin e_{\pi_i(g),i}$ y $\sharp(e'_{\pi_i(g),i}) = 1$, entonces

$$e'_{\pi_i(g),i} = e_{\pi_i(g),i} = \{g\},$$

y, por tanto, se tiene:

$$\sharp(e'_{\pi_i(g),i}) = \sharp(e_{\pi_i(g),i}) = 1.$$

En conclusión, solamente las aristas de los dos primeros casos cuentan para $\text{deg}'(g)$ y, por tanto, tenemos:

$$\text{deg}'(g) = \sharp(\{e \setminus \{h\} \in E(H') : e \in E(H), g \in e, \sharp(e \setminus \{h\}) \geq 2\}) + \sharp(\{e \in E(H') : e \in E(H), g \notin e, \sharp(e) \geq 2\}).$$

Es decir, tenemos

$$\text{deg}'(g) \leq \sharp(\{e \in E(H) : g \in e, \sharp(e) \geq 2\}) + \sharp(\{e \in E(H) : g \notin e, \sharp(e) \geq 2\}).$$

Y, finalmente, concluimos que $\text{deg}'(g) \leq \text{deg}(g)$, probando la Desigualdad (2.2.7). \square

PROPOSICIÓN 2.2.8 (LOS “PROMEDIOS” NO SON FUNCIONES MONÓTONAS CON RESPECTO A LA INCLUSIÓN CONJUNTISTA). *Con las mismas notaciones anteriores, sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos (finita) con $\#(H) \geq 2$ y sea $h \in H$. Definamos $H' := H \setminus \{h\}$ y sea $G(H')$ el hiper-grafo asociado a H' . Tenemos:*

- i) Si $\deg(h) \geq \text{avd}(H)$, entonces $\text{avd}(H') \leq \text{avd}(H)$.
- ii) Sin embargo, el grado promedio (avd) no es una función monótona con respecto a la inclusión (exhibimos algún ejemplo). Más generalmente, supongamos que existe $h \in H$ tal que $\deg(h) + e_2(h, H) \leq \text{avd}(H)$, donde:

$$e_2(h, H) := \#(\{e \in E(H) : \#(e) = 2, h \in e\}).$$

Entonces, si $H' := H \setminus \{h\}$, tenemos que $\text{avd}(H) \leq \text{avd}(H')$. Más aún, $\deg(h) + e_2(h, H) < \text{avd}(H)$ implica $\text{avd}(H) < \text{avd}(H')$.

- iii) Si $\deg(h) \geq \text{outdeg}_{\text{av}}(H)$, entonces $\text{outdeg}_{\text{av}}(H') \leq \text{outdeg}_{\text{av}}(H)$.
- iv) Si, por el contrario, $\deg(h) \leq \text{outdeg}_{\text{av}}(H)$, entonces $\text{outdeg}_{\text{av}}(H') \geq \text{outdeg}_{\text{av}}(H)$.
- v) Adicionalmente, $\text{outdeg}_{\text{av}}$ tampoco es monótona con respecto a la inclusión. Más aún, si $\deg(h) < \text{outdeg}_{\text{av}}(H)$, entonces $\text{outdeg}_{\text{av}}(H) < \text{outdeg}_{\text{av}}(H')$.

DEMOSTRACIÓN. Probemos cada afirmación separadamente:

- *Afirmación i):* Por la definición de $\text{avd}(H)$ sabemos que:

$$\text{avd}(H) = \frac{1}{\#(H)} \left(\sum_{g \in H} \deg(g) \right).$$

Por tanto, tenemos:

$$\#(H) \text{avd}(H) = \left(\sum_{g \in H'} \deg(g) \right) + \deg(h).$$

Entonces, también tenemos:

$$\#(H) \text{avd}(H) = (\#(H) - 1) \frac{\left(\sum_{g \in H'} \deg(g) \right)}{(\#(H) - 1)} + \deg(h).$$

Concluimos así:

$$(\#(H) - 1) \left(\text{avd}(H) - \frac{\left(\sum_{g \in H'} \deg(g) \right)}{(\#(H) - 1)} \right) = \deg(h) - \text{avd}(H).$$

Como la parte derecha de esta igualdad es positiva (i.e. $\deg(h) - \text{avd}(H) \geq 0$) concluimos:

$$\text{avd}(H) - \frac{\left(\sum_{g \in H'} \deg(g) \right)}{(\#(H) - 1)} \geq 0,$$

o, equivalentemente, concluimos:

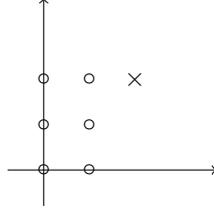
$$\text{avd}(H) \geq \frac{\left(\sum_{g \in H'} \deg(g) \right)}{(\#(H) - 1)}.$$

Por el Lema 2.2.7, sabemos que $\deg'(g) \leq \deg(g)$, para cada $g \in H'$ y, por tanto, concluimos:

$$\text{avd}(H) \geq \frac{\left(\sum_{g \in H'} \deg'(g) \right)}{(\#(H) - 1)} = \text{avd}(H').$$

- *Afirmación ii):* ejemplos donde $\text{avd}(H)$ no es una función monótona son sencillos de exhibir. Los incluimos por completitud:

Consideremos $n = 2$, $\mathcal{Y} := \{0, 1, 2\}$ y $H := \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 2)\}$. El hiper-grafo $G(H)$ tiene una arista que es un conjunto unario $\{(2, 2)\}$. Entonces, $\text{avd}(H) = 13/7 < 2$. Si eliminamos $h := (2, 2) \in H$, la clase $H' := H \setminus \{(2, 2)\}$ es un cubo con 6 vértices y se calcula obviamente que $\text{avd}(H') = 2 > \text{avd}(H)$, aunque $H' \subseteq H$. La imagen de más abajo muestra todos los vértices de H' con \circ y el elemento $h \in H$ eliminado con \times :



Para el segundo enunciado de esta afirmación, observemos la siguiente descomposición de los conjuntos de aristas de cardinal 1 en $G(H)$ y $G(H')$ en función de los cardinales:

$$E^{(1)}(H') = \{e \setminus \{h\} : e \in E(H), \#(e) = 2, h \in e\} \sqcup \{e \in E^{(1)}(H) : h \notin e\}.$$

$$E^{(1)}(H) = \{e \in E^{(1)}(H) : h \in e\} \sqcup \{e \in E^{(1)}(H) : h \notin e\}.$$

Entonces, $\#(E^{(1)}(H)) - \#(E^{(1)}(H')) = (n - \deg(h)) - e_2(h, H)$, porque:

$$e_2(h, H) = \#(\{e \setminus \{h\} : e \in E(H), \#(e) = 2, h \in e\}),$$

y $n - \deg(h) = \#(\{e \in E^{(1)}(H) : h \in e\})$. Por tanto, concluimos:

$$(\#(H)) \frac{\#(E^{(1)}(H))}{\#(H)} - (\#(H) - 1) \frac{\#(E^{(1)}(H'))}{\#(H) - 1} = (n - \deg(h)) - e_2(h, H).$$

Y, por tanto, concluimos:

$$(\#(H) - 1) \left(\frac{\#(E^{(1)}(H))}{\#(H)} - \frac{\#(E^{(1)}(H'))}{\#(H) - 1} \right) = (n - \deg(h)) - e_2(h, H) - \frac{\#(E^{(1)}(H))}{\#(H)}.$$

Por la Afirmación *i*) de la Proposición 2.2.1, concluimos:

$$(\#(H) - 1) \left(\frac{\#(E^{(1)}(H))}{\#(H)} - \frac{\#(E^{(1)}(H'))}{\#(H) - 1} \right) = \text{avd}(H) - (\deg(h)) + e_2(h, H).$$

Como $\#(H') = \#(H) - 1$, de nuestras hipótesis concluimos que:

$$\frac{\#(E^{(1)}(H))}{\#(H)} - \frac{\#(E^{(1)}(H'))}{\#(H')} \geq 0.$$

De nuevo, por la Afirmación *i*) de la Proposición 2.2.1 concluimos:

$$\text{avd}(H) = n - \frac{\#(E^{(1)}(H))}{\#(H)} \leq n - \frac{\#(E^{(1)}(H'))}{\#(H')} = \text{avd}(H').$$

- *Afirmación iii*): Siguiendo un argumento similar al del Lema 2.2.7 anterior, probamos que

$$(2.2.8) \quad \#(E(H)) = \#(E(H')) + (n - \deg(h)).$$

La razón es que podemos descomponer $E(H')$ como la siguiente unión disjunta de subconjuntos:

$$E(H') = \{e \setminus \{h\} : e \in E(H), h \in e, \#(e) \geq 2\} \sqcup \{e : e \in E(H), h \notin e\}.$$

Por otra parte, tenemos la partición como unión disjunta:

$$E(H) = \{e \in E(H) : h \in e, \#(e) \geq 2\} \sqcup \{e \in E(H) : h \notin e\} \sqcup \{e \in E(H) : h \in e, \#(e) = 1\}.$$

Por tanto, concluimos $E(H) - E(H') = \#(\{e \in E(H) : h \in e, \#(e) = 1\})$, y la Identidad (2.2.8) se sigue. Seguidamente, observemos que la Identidad (2.2.8) implica:

$$\#(H) \frac{\#(E(H))}{\#(H)} = (\#(H) - 1) \frac{\#(E(H'))}{\#(H) - 1} + (n - \deg(h)).$$

Esto implica:

$$(\#(H) - 1) \left(\frac{\#(E(H))}{\#(H)} - \frac{\#(E(H'))}{\#(H) - 1} \right) = n - \deg(h) - \frac{\#(E(H))}{\#(H)}.$$

Por la Afirmación *ii*) de la Proposición 2.2.1 concluimos:

$$(\#(H) - 1) \left(\frac{\#(E(H))}{\#(H)} - \frac{\#(E(H'))}{\#(H')} \right) = \text{outdeg}_{\text{av}}(H) - \deg(h).$$

Puesto que $\text{outdeg}_{\text{av}} \leq \text{deg}(h)$, la parte derecha de esta igualdad es negativa y, por tanto, concluimos:

$$\frac{\#(E(H))}{\#(H)} \leq \frac{\#(E(H'))}{\#(H')}.$$

De nuevo, por la Afirmación *ii*) de la Proposición 2.2.1 concluimos:

$$\text{outdeg}_{\text{av}}(H') = n - \frac{\#(E(H'))}{\#(H')} \leq n - \frac{\#(E(H))}{\#(H)} = \text{outdeg}_{\text{av}}(H).$$

- *Afirmación iv*): Por un argumento similar al usado en la Afirmación *iii*) concluimos:

$$(\#(H) - 1) \left(\frac{\#(E(H))}{\#(H)} - \frac{\#(E(H'))}{\#(H) - 1} \right) = n - \text{deg}(h) - \frac{\#(E(H))}{\#(H)}.$$

Por tanto, si $\text{deg}(h) \leq \text{outdeg}_{\text{av}}(H)$, entonces la expresión de la parte derecha es positiva. Esto implica:

$$\frac{\#(E(H))}{\#(H)} \geq \frac{\#(E(H'))}{\#(H')}.$$

Y hemos concluido la prueba de la Afirmación *iv*):

$$\text{outdeg}_{\text{av}}(H') = n - \frac{\#(E(H'))}{\#(H')} \geq n - \frac{\#(E(H))}{\#(H)} = \text{outdeg}_{\text{av}}(H).$$

- *Afirmación v*): Para la última afirmación, el mismo ejemplo usado en la Afirmación *ii*) para probar que $\text{avd}(H)$ no es monótona con respecto a la inclusión, prueba que el grado de salida promedio no es monótono con respecto a la inclusión. Es decir, consideremos: $n = 2$, $\mathcal{Y} := \{0, 1, 2\}$, $H := \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 2)\}$, $h = (2, 2)$ y $H' := H \setminus \{h\}$. Por la Proposición 2.2.1 concluimos fácilmente que $\text{outdeg}_{\text{av}}(H) = 2 - 6/7 < 2 - 5/6 = \text{outdeg}_{\text{av}}(H')$, y el grado de salida promedio no es monótono con respecto a la inclusión. □

OBSERVACIÓN 2.2.9 (INDUCCIÓN POR INCLUSIÓN NO ES USABLE PARA EL ESTUDIO DE LOS “PROMEDIOS”). Obviamente, esta última proposición dificulta hacer argumentos sobre avd o outdeg basados en inducción usando simplemente la inclusión de conjuntos, lo que hará mucho más difícil su estudio.

2.3. Promedios de grado y grado de salida en conjuntos cerrados hacia abajo

Para cada $h := (h_1, \dots, h_n) \in \mathbb{N}^n$, definamos el conjunto:

$$(2.3.1) \quad DS(h) := \{i \in [n] : h_i \geq 1\}.$$

LEMA 2.3.1. *Sea $H \subseteq \mathcal{Y}^n$ un subconjunto cerrado hacia abajo finito (escalera finita como en la Definición 3). Para cada $h \in H$ denotemos por $\text{deg}(h)$ al grado de h en el hiper-grafo $G(H)$. Entonces, tenemos:*

- i) Si $f \preceq g$ con respecto al pre-orden de Dickson, entonces $DS(f) \subseteq DS(g)$ y, por tanto, $\#(DS(f)) \leq \#(DS(g))$.*
- ii) $\text{deg}(h) \geq \#(DS(h))$. La igualdad se alcanza en el caso de que h sea un elemento maximal de H para el pre-orden de Dickson.*

DEMOSTRACIÓN. Probemos cada afirmación por separado.

- *Afirmación i*): es inmediata por la definición del pre-orden de Dickson.
- *Afirmación ii*): como H es cerrado hacia abajo, fijado un $h = (h_1, \dots, h_n) \in H$, sabemos que H contiene elementos de la forma $h_{\eta, i} = (h_1, \dots, h_{i-1}, \eta, h_{i+1}, \dots, h_n) \in H, \forall i \in [n]$ si $\eta \leq h_i$. En consecuencia, si $h_i \geq 1$, sabemos que la arista en la dirección i que contenga a $h_{\eta, i}$ tendrá al menos 2 elementos, a saber, $h_{0, i}$ y $h_{1, i}$.

Esto permite deducir que $\{e_{\pi_i(h), i} : h(i) \geq 1\} \subseteq \{e \in E(H)_h : \#(e) \geq 2\}$ y, por lo tanto, que $\text{deg}(h) \geq \#(DS(h))$.

Nótese que la inclusión en la otra dirección es sólo cierta en el caso de que h sea maximal en H para el pre-orden de Dickson ya que, al ser H cerrado hacia abajo, la condición $h_i \geq 1$, conjuntamente con la maximalidad de h , es necesaria y suficiente para que la arista arista en la dirección i tenga más de un elemento. □

2.3.1. Ejemplo: escalones \mathcal{E}_h : grado y grado de salida promedios. Recuperamos el ejemplo del escalón \mathcal{E}_h , determinado por algún $h \in \mathbb{N}^n$, introducido en la Identidad (1.1.1). Consideremos también el conjunto $DS(h)$ definido en la Identidad (2.3.1) anterior, esto es,

$$DS(h) := \{i \in [n] : h_i \geq 1\}.$$

PROPOSICIÓN 2.3.2. *Sea $n \in \mathbb{N}$, $n \geq 2$ y sea $h \in \mathbb{N}^n$ exponente monomial. Sea \mathcal{E}_h la clase de conceptos definida como el escalón determinado por h . Entonces, se verifican las siguientes propiedades:*

- i) *El grado promedio de \mathcal{E}_h , el grado de h en el hiper-grafo $G(\mathcal{E}_h)$ y el cardinal del conjunto $DS(h)$ coinciden. Es decir, tenemos:*

$$\deg(h) = \sharp(DS(h)) = \text{avd}(\mathcal{E}_h)$$

- ii) *El cardinal de \mathcal{E}_h es dado por la siguiente igualdad:*

$$\sharp(\mathcal{E}_h) = \prod_{i \in DS(h)} (1 + h_i)$$

- iii) *El número total de aristas en $E(\mathcal{E}_h)$ es dado por la siguiente igualdad:*

$$\sharp(E(\mathcal{E}_h)) = (n - m)\sharp(\mathcal{E}_h) + \sum_{i \in DS(h)} \prod_{j \in DS(h), j \neq i} (1 + h_j),$$

donde m es el cardinal de $DS(h)$.

- iv) *El grado de salida promedio de \mathcal{E}_h verifica:*

$$\text{outdeg}_{\text{av}}(\mathcal{E}_h) = \sharp(DS(h)) - \sum_{i \in DS(h)} \frac{1}{1 + h_i}.$$

DEMOSTRACIÓN. Probemos cada afirmación por separado:

- *Afirmación i):* la primera igualdad ya ha sido discutida en el Lema 3.2.10 anterior. En cuanto a la segunda igualdad, consideremos $DS(h)$ y denotemos por m al cardinal de este conjunto. Para cada $i \notin DS(h)$, se tiene que $h(i) = 0$ y, al ser \mathcal{E}_h cerrada hacia abajo y h maximal, tenemos que la arista en la dirección i ha de ser unaria. Es decir, podemos expresar el conjunto de aristas unarias de \mathcal{E}_h como:

$$E^{(1)}(\mathcal{E}_h) = \{e_{\pi_j(g), j} : g \in \mathcal{E}_h, j \notin DS(h)\},$$

que tiene $(n - m)\sharp(\mathcal{E}_h)$ elementos.

Como $\sharp(\mathcal{E}_h) = \sharp(\mathcal{E}_h|_{DS(h)})$, concluimos, a partir de la Afirmación i) de la Proposición 2.2.1, que se verifica lo siguiente:

$$\text{avd}(\mathcal{E}_h) = n - \frac{\sharp(E^{(1)}(\mathcal{E}_h))}{\sharp(\mathcal{E}_h)} = n - \frac{(n - m)\sharp(\mathcal{E}_h)}{\sharp(\mathcal{E}_h)} = m = \sharp(DS(h)).$$

- *Afirmación ii):* basta con observar que h es maximal y \mathcal{E}_h es cerrada hacia abajo para llegar a que:

$$\mathcal{E}_h = \{(\xi_1, \dots, \xi_n) : \xi_i \leq h_i, \forall i \in [n]\},$$

que tiene $\prod_{i \in DS(h)} (1 + h_i)$ elementos, donde hemos omitido las coordenadas donde $h_i = 0$ (i.e., $i \notin DS(h)$), ya que no afectan al producto.

- *Afirmación iii):* en virtud de la partición descrita en la Igualdad (2.1.4), el conjunto de las aristas se puede expresar a partir de los conjuntos $E_i(\mathcal{E}_h)$, es decir, la siguiente partición:

$$(2.3.2) \quad E(\mathcal{E}_h) = \bigsqcup_{i=1}^n E_i(\mathcal{E}_h)$$

Primero, observamos que el número de aristas, por la acción de la proyección i -ésima, depende del cardinal de $\pi_i(\mathcal{E}_h)$, es decir, cada elemento de la imagen determina una arista por acción de la anti-imagen. En consecuencia, el cardinal de $E_i(\mathcal{E}_h)$ se determina de forma similar a como hemos hecho en la prueba de la Afirmación ii), con la diferencia de que omitimos la coordenada i :

$$(2.3.3) \quad \sharp(E_i(\mathcal{E}_h)) = \prod_{j \in DS(h) \setminus \{i\}} (1 + h_j),$$

puesto que el valor de la coordenada i no es relevante para la proyección π_i . Además, no tenemos en cuenta las direcciones j tales que $j \notin DS(h)$, puesto que entonces se tiene $h_j = 0$, y, al ser h maximal en un cerrado hacia abajo, no afecta a la expresión del cardinal.

De forma similar como hemos hecho en la prueba de la Afirmación i), observemos que, si $i \in [n] \setminus DS(h)$,

la arista en la dirección i es unaria, por lo que los conjuntos $E_i(\mathcal{E}_h)$ son de aristas unarias. Por tanto, se tiene que:

$$(2.3.4) \quad (n - m)\sharp(\mathcal{E}_h) = \sharp\left(E^{(1)}(H)\right) = \sum_{i \in [n] \setminus DS(h)} \sharp(E_i(\mathcal{E}_h)),$$

donde, nuevamente, se emplea la notación $m = \sharp(DS(h))$. Ahora, expresando la partición de la Igualdad (2.3.2) en función de $DS(h)$, tomando cardinales y empleando la Igualdad (2.3.4), se tiene:

$$\begin{aligned} \sharp(E(\mathcal{E}_h)) &= \sum_{i \in [n] \setminus DS(h)} E_i(\mathcal{E}_h) + \sum_{i \in DS(h)} E_i(\mathcal{E}_h) = \\ &= (n - m)\sharp(\mathcal{E}_h) + \sum_{i \in DS(h)} \prod_{j \in DS(h) \setminus \{i\}} (1 + h_j). \end{aligned}$$

- *Afirmación iv)*: se sigue inmediatamente de la Afirmación *ii)* de la Proposición 2.2.1 y de la Afirmación *iii)* anterior. Es decir, por lo primero, se tiene:

$$\text{outdeg}_{\text{av}}(H) = n - \frac{\sharp(E(\mathcal{E}_h))}{\sharp(\mathcal{E}_h)}.$$

Ahora, empleando la Afirmación *iii)*, se tiene:

$$\text{outdeg}_{\text{av}}(\mathcal{E}_h) = n - \frac{(n - m)\sharp(\mathcal{E}_h) - \sum_{i \in DS(h)} \prod_{j \in DS(h) \setminus \{i\}} (1 + h_j)}{\sharp(\mathcal{E}_h)}.$$

donde, al igual que antes, $m = \sharp(DS(h))$. El primer término del numerador de la fracción se simplifica de forma inmediata. La segunda parte se puede simplificar empleando la Afirmación *ii)*. Es decir, nos queda:

$$\text{outdeg}_{\text{av}}(\mathcal{E}_h) = m - \frac{\sum_{i \in DS(h)} \prod_{j \in DS(h) \setminus \{i\}} (1 + h_j)}{\prod_{i \in DS(h)} (1 + h_i)},$$

que, tras simplificar la fracción, queda:

$$\text{outdeg}_{\text{av}}(\mathcal{E}_h) = m - \sum_{i \in DS(h)} \frac{1}{1 + h_i} = \sharp(DS(h)) - \sum_{i \in DS(h)} \frac{1}{1 + h_i}.$$

□

Pseudo-cubos, dimensiones DS, exponencial y de Natarajan

Índice

3.1. Introducción: Pseudo-cubos	25
3.2. Dimensiones	27
3.2.1. Pseudo-dimensión DS	29
3.2.2. Escalones \mathcal{E}_h y sus dimensiones	30
3.3. Pseudo-cubos cerrados hacia abajo.	31

3.1. Introducción: Pseudo-cubos

Hasta donde alcanza nuestro conocimiento, el término *pseudo-cubo* fue usado en la literatura científica por vez primera en 2022 (concretamente en [BCDMY, 22]). Lo usan para expresar una noción implícita de dimensión que aparece por vez primera en [DS, 14]. La noción es crucial para poder definir uno de los varios conceptos de dimensión con los que se trata de caracterizar las clases de funciones “aprendibles” en la teoría de funciones multi-clase; [BCDMY, 22] llamarán a esa dimensión la “*dimensión de Daniely-Shalev-Schwartz*” (o brevemente *dimensión DS*), que estudiaremos más adelante. Sin embargo, *no existe un estudio sistemático de esa noción de pseudo-cubo hasta el presente manuscrito, por vez primera en la literatura matemática*. Comenzaremos con una caracterización de la noción de pseudo-cubo basada en los invariantes introducidos en el Capítulo 2 precedente. Por tanto, todo lo que sigue es completamente original, como la mayoría del material en este TFG.

DEFINICIÓN 15 (CUBOS). *Un cubo $H \subseteq \mathcal{Y}^n$ de dimensión n es cualquier subconjunto tal que existe $\mathcal{Z} \subseteq \mathcal{Y}$ de modo que $H = \mathcal{Z}^n$.*

OBSERVACIÓN 3.1.1 (POLI-CUBOS). Dado que estamos hablando de conjuntos finitos nos resulta más complicado usar los términos politopo, poldisco o rectángulo, entre otros. Por lo tanto, hablaremos de *poli-cubos* para hablar de productos cartesianos de cubos, esto es, $H \subseteq \mathcal{Y}^n$ es un poli-cubo si existen $\mathcal{Z}_1, \dots, \mathcal{Z}_n \subseteq \mathcal{Y}$ tales que:

$$(3.1.1) \quad H := \prod_{i=1}^n \mathcal{Z}_i.$$

En cambio, en [BCDMY, 22], inspirados en [DS, 14], están más interesados en la noción siguiente:

DEFINICIÓN 16 (PSEUDO-CUBO). *Con las mismas notaciones que en capítulos precedentes, dados $n \in \mathbb{N}$, $n \geq 2$, y una clase de conceptos $H \subseteq \mathcal{Y}^n$, diremos que el OIG asociado $G(H)$ es un pseudo-cubo (de dimensión n) si y solamente si se verifica la siguiente propiedad:*

Para todo $i \in [n]$ y para todo $h = (h_1, \dots, h_n) \in H$, existe un i -vecino de h en H (i.e. existe $g = (g_1, \dots, g_n) \in H$ tal que $h_i \neq g_i$, mientras que $h_j = g_j$ para todo $j \in [n] \setminus \{i\}$).

En el caso $n = 1$ llamaremos pseudo-cubo de dimensión 1 a cualquier hiper-grafo $G(H)$ asociado a cualquier subconjunto $H \subseteq \mathcal{Y}$ de cardinal mayor o igual que 2. Llamaremos pseudo-cubo de dimensión 0 a $G(\emptyset)$.

OBSERVACIÓN 3.1.2 (VECINOS, DISTANCIA DE HAMMING Y ARISTAS DEL OIG). En la definición precedente hemos usado la noción de i -vecino y la hemos definido. Podemos dar definiciones equivalentes usando elementos de estudio ya descritos anteriormente. Así, dos elementos $g, h \in H$ son i -vecinos si y solamente si se verifica cualquiera de las dos propiedades siguientes:

- i) Ambos elementos son *vecinos*, i.e. la distancia de Hamming entre f y g es 1 (i.e. $d_{\text{Ham}}(f, g) = 1$) o, en la terminología de [Hau, 1995] (cf. Observación 2.1.2) la distancia de Haussler ρ entre f y g satisface $\rho(f, g) = 1/n$.
- ii) Se tiene $f \neq g$ y existe $i \in [n]$ tal que, considerando la proyección que “olvida” la coordenada i -ésima de $\pi_i : \mathcal{Y}^n \rightarrow \mathcal{Y}^{n-1}$, se tiene que $\pi_i(f) = \pi_i(g)$.

En términos de las aristas del hiper-grafo $G(H)$, dos elementos $g, h \in H$ distintos son i -vecinos si y solamente si están en la misma arista i -ésima de $G(H)$ para algún $i \in [n]$ (i.e. si y solamente si $g, h \in e_{f,i}$ para algún $f = \pi_i(h) = \pi_i(g)$, con $g \neq h$).

OBSERVACIÓN 3.1.3 (LOS CASOS DEFECTIVOS). Obsérvese que los casos defectivos $n = 1$ y $H = \emptyset$ han sido definidos por las siguientes razones:

- En el caso $n = 1$ y $H \subseteq \mathcal{Y}$, solamente hay una arista y, en ese caso, todos los elementos poseen, al menos, un vecino distinto si y solamente si $\sharp(H) \geq 2$. Además, en este caso, tenemos que $\deg(h) = 1$ para todo $h \in H$ y, por tanto, tenemos:

$$\text{avd}(H) = 1, \text{gd}(H) = \frac{1}{\sharp(H)},$$

mientras que el grado de salida de cualquier orientación $\sigma : E(H) = \{e\} \rightarrow H$ verifica:

$$\text{outdeg}(\sigma, h) := \begin{cases} 1, & \text{si } h \neq \sigma(e) \\ 0, & \text{en otro caso.} \end{cases}$$

Por tanto, el grado de salida promedio en el caso $n = 1$ verifica:

$$\text{outdeg}_{\text{av}}(H) = \frac{1}{\sharp(H)} \sum_{h \in H} \text{outdeg}(\sigma, h) = \frac{\sharp(H) - 1}{\sharp(H)},$$

y, por tanto, en este caso $n = 1$, tenemos $\text{avd}(H) - \text{outdeg}_{\text{av}}(H) - \text{gd}(H) = 0$. Nótese que, en el caso $n = 1$, si $\sharp(H) = 1$, es un cubo, pero no es un pseudo-cubo porque el único elemento no posee ningún vecino distinto en la única dirección de proyección posible. De ahí lo “defectivo” del caso.

- El caso $H = \emptyset$ es usado como caso defectivo del caso defectivo, si se admite la redundancia. Si no hay ningún punto, todos los puntos poseen vecinos distintos, i.e., una verdad vacía.

TEOREMA 3.1.4 (CARACTERIZACIÓN DE PSEUDO-CUBOS DE DIMENSIÓN n). *Con las mismas notaciones de páginas precedentes, sea $n \in \mathbb{N}$, $n \geq 1$, sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita, con $\sharp(H) \geq 2$ y sea $G(H) = (H, E(H))$ el hiper-grafo de una inclusión (OIG) asociado a H . Las siguientes propiedades son equivalentes:*

- El hiper-grafo $G(H)$ es un pseudo-cubo de dimensión n .
- Para cada arista $e \in E(H)$, $\sharp(e) \geq 2$.
- El grado promedio satisface $\text{avd}(H) = n$.
- Para cada $h \in H$, $\deg(h) = n$.
- Grado promedio, grado de salida promedio y densidad del grafo satisfacen la siguiente identidad

$$(3.1.2) \quad \text{avd}(H) - \text{outdeg}_{\text{av}}(H) - \text{gd}(H) = 0.$$

En particular, si H es un pseudo-cubo de dimensión n y $n \geq 2$, se verifica:

$$(3.1.3) \quad \frac{n}{2} \leq \text{outdeg}_{\text{av}}(H) < n.$$

DEMOSTRACIÓN. El caso $n = 1$ ya ha sido discutido en la Observación 3.1.3, por tanto supondremos $n \geq 2$. Nuestra Observación 3.1.2 anterior muestra que la Afirmación *i*) implica la Afirmación *ii*). También resulta obvia la implicación Afirmación *ii*) \implies Afirmación *i*). Sólo por ser completos, como en la Observación 3.1.2, dados $i \in [n]$ y $h \in H$, la arista $e := e_{\pi_i(h), i}$ contiene a todos los i -vecinos de h . Por tanto, dados $\sharp(e) \geq 2$ y $h \in e$, siempre habrá un i -vecino de h en H . Por la Afirmación *i*) del Teorema 2.2.1, tenemos inmediatamente que las Afirmaciones *ii*) y *iii*) son equivalentes. La Afirmación *iv*) implica la Afirmación *iii*), mientras que la Afirmación *ii*) implica inmediatamente la Afirmación *iv*).

La equivalencia entre *iii*) y *v*) se sigue de las Afirmaciones *i*) y *ii*) del Teorema 2.2.1.

Las últimas desigualdades se siguen del Teorema 2.2.1:

$$\text{outdeg}_{\text{av}}(H) \leq \text{avd}(H) = n \leq 2 \text{outdeg}_{\text{av}}(H),$$

lo que implica:

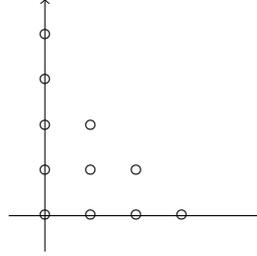
$$\frac{n}{2} = \frac{\text{avd}(H)}{2} \leq \text{outdeg}_{\text{av}}(H) \leq n = \text{avd}(H).$$

Por la última dicotomía del Teorema 2.2.1, sabemos que si $n \geq 1$, $\text{outdeg}_{\text{av}} < \text{avd}(H)$ y se concluye la prueba. \square

EJEMPLO 3.1.5 (CUBOS, POLI-CUBOS Y CERRADOS HACIA ABAJO EN RELACIÓN CON PSEUDO-CUBOS). Trate-mos cada ejemplo para observar el fenómeno. Nos ocupamos de casos con $n \geq 2$:

- Todo cubo no vacío $H := \mathcal{Z}^n$, con $n \geq 2$, es pseudo-cubo si y solamente si $\sharp(\mathcal{Z}) \geq 2$.
- Todo poli-cubo no vacío $H := \prod_{i=1}^n \mathcal{Z}_i$ es pseudo-cubo si y solamente si $\sharp(\mathcal{Z}_i) \geq 2$. Nótese que si todos los \mathcal{Z}_i 's tienen cardinal al menos 2, entonces, en la dirección de proyección π_i todo punto $(x_1, \dots, x_n) \in H$ posee un i -vecino reemplazando $x_i \in \mathcal{Z}_i$ por el otro elemento de \mathcal{Z}_i distinto de x_i , que debe existir en \mathcal{Z}_i por ser $\sharp(\mathcal{Z}_i) \geq 2$.

- iii) Un subconjunto $H \subseteq \mathcal{Y}^n$ cerrado hacia abajo es pseudo-cubo si y solamente si ninguno de sus elementos maximales (con respecto al pre-orden de Dickson \preceq) de H están en alguno de los ejes. Es decir, si no hay elementos maximales de H en $\cup_{i=1}^n \{\mu \in \mathbb{N}_i : x_i = 0\}$. Claramente, si un elemento maximal posee un coordenada nula (por ejemplo $h = (0, x_2, \dots, x_n) \in H$ elemento maximal) y H es cerrado hacia abajo, no puede haber dos puntos en la arista $e_{\pi_1(h),1} = \{h\}$. Porque si hubiera un segundo punto en esa arista de H tendría que tener la forma (a, x_2, \dots, x_n) con $a > 0$, lo que impediría que h fuera un elemento maximal. Un ejemplo dibujado puede ser el siguiente en el que “fallan” los dos ejes:



Tomando $\mathcal{Y} := \{0, 1, 2, 3, 4, \}$ vemos que hay dos elementos maximales de H $\{(0, 4), (3, 0)\}$ (con respecto al pre-orden de Dickson) que no tienen vecinos en alguna de las dos direcciones: $(0, 4)$ no posee 1-vecinos (en la dirección π_1 que “olvida” la primera coordenada), mientras que $(3, 0)$ no posee 2-vecinos.

OBSERVACIÓN 3.1.6. Si H es un pseudo-cubo de dimensión n , entonces, para una orientación cualquiera $\sigma : G(H) \rightarrow H$ se tiene:

$$\frac{n}{2} \leq \text{outdeg}_{\text{av}}(H) \leq \text{outdeg}_{\text{max}}(\sigma).$$

La desigualdad simplificada (i.e. $n/2 \leq \text{outdeg}_{\text{max}}(\sigma)$ para orientaciones de pseudo-cubos de dimensión n) es “Lemma 12” de [BCDMY, 22], sin percatarse del papel esencial que juega $\text{outdeg}_{\text{av}}(H)$ en esa desigualdad (porque no tienen la noción $\text{outdeg}_{\text{av}}(H)$ como noción distinguida, ni tampoco son conscientes de que esa cantidad sea independiente de σ).

3.2. Dimensiones

Comencemos nuestro recorrido hacia ciertas nociones de dimensión relacionadas con el aprendizaje multi-clase. En esencia, la dimensión mide la longitud mínima que debe tener la familia de muestreo (longitud del dataset) para tener un aprendizaje aproximadamente correcto. Aunque a primera vista parece una descripción compleja, dejaremos la explicación detallada para otro momento y nos centraremos primero en algunas nociones de dimensión de manera directa y simplificada. En primer lugar, exponemos la noción de dimensión usada en [DS, 14], parafraseada aquí con la noción de pseudo-cubo discutida en la sección precedente.

DEFINICIÓN 17 (DIMENSIÓN DS, [DS, 14]). Con las notaciones precedentes, dados $H \subseteq \mathcal{Y}^n$ y dado $S \subseteq [n]$, $\sharp(S) = d$, diremos que S es desmenuzado (roto en pedazos, “shattered”) “à la DS” (o DS-desmenuzado) por H si $H|_S$ contiene un pseudo-cubo de dimensión d . Definimos la dimensión DS de H como el máximo d tal que existe $S \subseteq [n]$, de cardinal d (i.e. $\sharp(S) = d$), tal que S es desmenuzado “à la DS” por H . Denotaremos la dimensión DS de H mediante $\text{dim}_{\text{DS}}(H)$.

OBSERVACIÓN 3.2.1 (LA DIMENSIÓN DS ES MONÓTONA CON RESPECTO A LA INCLUSIÓN). De la definición anterior, se infiere fácilmente que la dimensión DS es una función monótona con respecto a la inclusión de conjuntos, i.e. dados $G \subseteq H \subseteq \mathcal{Y}^n$, se tiene $\text{dim}_{\text{DS}}(G) \leq \text{dim}_{\text{DS}}(H)$, porque si $S \subseteq [n]$ es un conjunto de cardinal máximo tal que la restricción $G|_S$ “contiene” un pseudo-cubo de dimensión $\sharp(S)$, entonces $H|_S \supseteq G|_S$ también contiene un pseudo-cubo de dimensión $\sharp(S)$.

En [BCDMY, 22], se introduce otra noción de dimensión de una clase de conceptos finita:

DEFINICIÓN 18 (DIMENSIÓN EXPONENCIAL, [BCDMY, 22]). Dado $H \subseteq \mathcal{Y}^n$, definimos la dimensión exponencial de H como la cantidad:

$$\text{dim}_{\text{E}}(H) := \max\{\sharp(S) : \sharp(H|_S) \geq 2^{\sharp(S)}\}.$$

OBSERVACIÓN 3.2.2 (LA DIMENSIÓN EXPONENCIAL ES HEREDITARIA). Observemos que la dimensión exponencial es hereditaria para subconjuntos y restricciones. Es decir,

- Dados $F \subseteq H \subseteq \mathcal{Y}^n$, tenemos: $\text{dim}_{\text{E}}(F) \leq \text{dim}_{\text{E}}(H) \leq n$.
- Dado $S \subseteq [n]$ y dado $H \subseteq \mathcal{Y}^n$, también tenemos: $\text{dim}_{\text{E}}(H|_S) \leq \text{dim}_{\text{E}}(H) \leq n$.

El Lema 12 de [BCDMY, 22] es el siguiente corolario de nuestros resultados previos:

COROLARIO 3.2.3 (LEMA 12 DE [BCDMY, 22]). *Sea $H \subseteq \mathcal{Y}^n$ una clase finita de conceptos de dimensión DS igual a n . Entonces, para cualquier orientación $\sigma : E(H) \rightarrow H$ tenemos:*

$$\frac{n}{2} = \frac{\dim_{\text{DS}}(H)}{2} \leq \text{outdeg}_{\text{max}}(\sigma).$$

DEMOSTRACIÓN. Como $\dim_{\text{DS}}(H) = n$, hay un pseudo-cubo $\mathcal{C} \subseteq H$ con cardinal maximal de dimensión n . Por la Afirmación *v*) de la Proposición 2.2.1, para cada orientación $\sigma : E(H) \rightarrow H$ hay una orientación $\sigma' : E(\mathcal{C}) \rightarrow \mathcal{C}$ tal que $\text{outdeg}_{\text{max}}(\sigma') \leq \text{outdeg}_{\text{max}}(\sigma)$. Por la Identidad (3.1.3) del Teorema 3.1.4, también concluimos lo siguiente:

$$\frac{n}{2} \leq \text{outdeg}_{\text{av}}(\mathcal{C}) \leq \text{outdeg}_{\text{max}}(\sigma') \leq \text{outdeg}_{\text{max}}(\sigma),$$

y la afirmación se sigue. \square

Se podría cuestionar si la Desigualdad (3.1.3) sigue siendo cierta si H no tiene dimensión DS igual a n . El siguiente ejemplo prueba que esa desigualdad no es cierta en general (en particular, no es cierta si H no es un pseudo-cubo de dimensión n).

EJEMPLO 3.2.4. Como se ha indicado, nos preguntamos si la cota inferior del Lema 12 de [BCDMY, 22] sigue siendo cierta si reemplazamos $\text{outdeg}_{\text{max}}(H)$ por $\text{outdeg}_{\text{av}}(H)$ cuando $H \subseteq \mathcal{Y}^n$ tiene dimensión DS igual a n . El siguiente ejemplo prueba que esta generalización no es cierta en general: Dado $\mathcal{Y}_p = \{0, \dots, p-1\}$, hay una secuencia de clases asociadas $H_p \subseteq \mathcal{Y}_p^2$ tales que se verifica lo siguiente:

$$(3.2.1) \quad \lim_{p \rightarrow \infty} \text{outdeg}_{\text{av}}(H_p) = \lim_{p \rightarrow \infty} \text{avd}(H_p) = 0 < 2 = \dim_{\text{DS}}(H_p) = \dim_{\text{E}}(H_p).$$

Consideremos los conjuntos $H_p := \{(0,0), (0,1), (1,0), (1,1)\} \cup \{(k,k) : 2 \leq k \leq p-1\}$. Observemos que $\sharp(H_p) = p+2$. Obviamente, H_p contiene un cubo de dimensión 2 y su dimensión exponencial no puede ser mayor que 2. Por tanto, $\dim_{\text{DS}}(H_p) = \dim_{\text{E}}(H_p) = 2$. El número total de aristas satisface $\sharp(E(H_p)) = 4+2(p-2) = 2p$, mientras que el número total de aristas unarias satisface $\sharp(E^{(1)}(H_p)) = 2(p-2)$.

Entonces,

$$\text{outdeg}_{\text{av}}(H_p) = 2 - \frac{2p}{p+2} = \frac{4}{p+2},$$

mientras que el grado promedio satisface:

$$\text{avd}(H_p) = 2 - \frac{2p-4}{p+2} = \frac{8}{p+2}.$$

Concluimos así la Desigualdad (3.2.1) tomando límites del modo más elemental.

La noción de dimensión usada tradicionalmente en aprendizaje multi-clase es la noción de dimensión de Natarajan (que, como se observó en [BCDMY, 22] es insuficiente para caracterizar completamente el aprendizaje multi-clase en el caso de que $\sharp(\mathcal{Y})$ no esté acotado). Escribamos su relación con la dimensión exponencial.

OBSERVACIÓN 3.2.5 (DIMENSIÓN DE NATARAJAN). En [Nat, 1988] y [Nat, 1989], B.K. Natarajan introdujo una primera noción de dimensión aplicable a aprendizaje multi-clase. Una noción un tanto artificial que no será central en las discusiones de esta memoria, aunque merece un cierto reconocimiento histórico. Seguimos la aproximación de [BCDMY, 22].

Con las notaciones precedentes, sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Dado $q \in \mathbb{N}$, $1 \leq q \leq n$, y dado $S = (s_1, \dots, s_q) \in [n]^q$, decimos que S es desmenuzado “à la Natarajan” (N -desmenuzado o N -shattered) por H si existen $f, g \in H$ tales que $f(s_i) \neq g(s_i)$ para todo $i \in [q]$, y tales que

$$H(S) \supseteq \{f(s_1), g(s_1)\} \times \dots \times \{f(s_q), g(s_q)\},$$

donde

$$H(S) := \{(h(s_1), \dots, h(s_q)) \in \mathcal{Y}^q : h \in H\}.$$

Nótese que se trata de que $H(S)$ contenga un poli-cubo, dado como el producto cartesiano de conjuntos de cardinal 2, aunque sus elementos vengan de distintas funciones de H restringidas a S . La dimensión de Natarajan de H es el máximo de los q 's tales que existe $S \in [n]^q$ que es N -desmenuzado por H . Denotamos la dimensión de Natarajan mediante $\dim_{\text{N}}(H)$.

Con estas notaciones, sea $S \in [n]^q$ una lista de longitud q , N -desmenuzada por H .

Definamos $\tilde{S} := \{s_i : i \in [q]\}$, i.e., el conjunto formado por las diferentes coordenadas de una lista $S = (s_1, \dots, s_q) \in [n]^q$. Entonces, $\sharp(\tilde{S}) = q$. Es decir, si S es N -desmenuzado por H , entonces $s_i \neq s_j$, para

todo $i \neq j$. Observemos que si $s_i = s_j$, con $i \neq j$, entonces existe $h \in H$ tal que se verifican las siguientes propiedades:

$$h(s_i) = f(s_i), h(s_j) = g(s_j), h(s_k) \in \{f(s_k), g(s_k)\}, \text{ para } k \notin \{i, j\}.$$

Pero, si $s_i = s_j$, entonces $f(s_i) = f(s_j) = h(s_i)$ y también $g(s_i) = g(s_j) = h(s_j) = h(s_i)$, lo que implica $f(s_i) = g(s_i)$. Y eso no es posible si S es N -desmenuzado por H .

Entonces, si $S \in [n]^q$ es N -desmenuzado por H , entonces:

$$(3.2.2) \quad \sharp(H|_{\tilde{S}}) \geq 2^{\sharp(\tilde{S})} = 2^q.$$

Lo que conduce al siguiente sencillo resultado.

COROLARIO 3.2.6 (DIMENSIÓN DE NATARAJAN INFERIOR A DIMENSIÓN EXPONENCIAL).

$$(3.2.3) \quad \dim_N(H) \leq \dim_E(H).$$

3.2.1. Pseudo-dimensión DS. Esta es una noción original del manuscrito, cuyo alcance es aún tema de trabajo conjunto. Pero creemos que, al menos, unas pocas ideas pueden formar parte de este TFG.

Mientras que la dimensión exponencial depende solamente de sus proyecciones, la dimensión DS ha sido definida para ser una función monótona creciente con respecto a la inclusión (y a restricciones), a través del uso del término “*contiene*” en la propia definición. Esto conduce a nuestra noción de pseudo-dimensión DS.

DEFINICIÓN 19 (PSEUDO-DIMENSIÓN DS). *Con las notaciones anteriores, dado $H \subseteq \mathcal{Y}^n$ definimos su pseudo-dimensión DS como el máximo de los cardinales de cualquier subconjunto $S \subseteq [n]$ tal que $H|_S$ sea un pseudo-cubo de dimensión $\sharp(S)$. Es decir, definimos*

$$p \dim_{DS}(H) := \max\{\sharp(S) : S \subseteq [n], H|_S \text{ es un pseudo-cubo de dimensión } \sharp(S)\}.$$

Si no hay ningún subconjunto $S \subseteq [n]$ tal que $H|_S$ es un pseudo-cubo, diremos simplemente que su pseudo-dimensión DS es 0.

La pseudo-dimensión DS se caracteriza por el cardinal de un conjunto que maximiza cierta función (característica) en el conjunto $2^{[n]}$.

PROPOSICIÓN 3.2.7 (PSEUDO-DIMENSIÓN DS ES UN MÁXIMO DE CIERTA FUNCIÓN). *Con las mismas notaciones precedentes, sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Definamos la siguiente función:*

$$\psi_H : 2^{[n]} \longrightarrow \mathbb{R},$$

$$S \longmapsto \begin{cases} \text{avd}(H|_S) - \text{outdeg}_{\text{av}}(H|_S) - \text{gd}(H|_S), & \text{si } n \geq 2 \text{ o } \sharp(H) \geq 2 \\ 0 & \text{si } n = 1 \text{ y } \sharp(H) = 1 \\ 0 & \text{si } H = \emptyset. \end{cases}$$

Entonces, tenemos:

$$p \dim_{DS}(H) = \max\{\sharp(S) : S \subseteq [n], S \text{ maximiza } \psi_H\}.$$

DEMOSTRACIÓN. Omitiremos los casos defectivos $n = 1$ y $\sharp(H) = 1$ o el caso $\sharp(H) = 0$.

Conforme a nuestros resultados previos, tenemos que para cada $S \subseteq [n]$ se verifica:

$$\psi_H(S) = -\frac{\sharp(E^{(1)}(H|_S))}{\sharp(H|_S)} \leq 0.$$

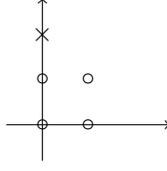
Y $\psi_H(S) = 0$ si y solamente si $H|_S$ es un pseudo-cubo de dimensión $\sharp(S)$. Entonces, S maximiza ψ_H si y solamente si $H|_S$ es un pseudo-cubo de dimensión $\sharp(S)$. \square

EJEMPLO 3.2.8 (LA PSEUDO-DIMENSIÓN NO ES MONÓTONA CON RESPECTO A LA INCLUSIÓN). El problema con la pseudo-dimensión es que no es una función monótona con respecto a la inclusión. Lo observaremos con el siguiente ejemplo.

Consideremos $H := \{(0, 0), (0, 1), (1, 0), (1, 1), (0, 2)\} \subseteq \{0, 1, 2\}^{[2]}$ y el subconjunto $G := H \setminus \{(0, 2)\}$.

G es un cubo de dimensión 2 y, por tanto, un pseudo-cubo de dimensión 2 (ver Ejemplo 3.1.5, donde se muestra que todo cubo es un pseudo-cubo). Por otro lado, H no es un pseudo-cubo puesto que tiene la arista unaria $\{(0, 2)\}$ (ver el Teorema 3.1.4) y, en consecuencia, se tiene $p \dim_{DS}(H) < 2$. Si se toma $S = \{1\}$, se tiene que $H|_S = \{0, 1\}$, por lo que es un pseudo-cubo, puesto que tiene más de un elemento. De lo anterior, se deduce que $p \dim_{DS}(H) = 1$. Esto es, $1 = p \dim_{DS}(H) < p \dim_{DS}(G) = 2$, pero $G \subseteq H$.

El dibujo siguiente, donde H son todos los puntos y \times representa el punto que se elimina para obtener G , lo aclara:



COROLARIO 3.2.9 (LA DIMENSIÓN DS SE SIGUE DE LA PSEUDO-DIMENSIÓN COMO UN MÁXIMO). *La pseudo-dimensión DS devuelve la dimensión DS como un máximo. Es decir, tenemos:*

$$(3.2.4) \quad \dim_{DS}(H) = \max\{p \dim_{DS}(G) : G \subseteq H\}.$$

3.2.2. Escalones \mathcal{E}_h y sus dimensiones. Recuperamos la noción de escalón introducida en la Identidad (1.1.1). La siguiente proposición complementa (en términos de dimensiones) el análisis de los invariantes de de \mathcal{E}_h ya iniciados en la Proposición 2.3.2. En el enunciado usaremos el conjunto $DS(h)$, asociado a cada $h \in \mathbb{N}^n$ y definido en la Identidad (2.3.1):

$$DS(h) := \{i \in [n] : h_i \geq 1\} = \{i \in [n] : h_i \neq 0\}.$$

PROPOSICIÓN 3.2.10. *Sea $n \in \mathbb{N}$ un entero positivo, $n \geq 2$, sea $h \in \mathbb{N}^n$ un exponente monomial y sea \mathcal{E}_h la clase de conceptos de todos los elementos de \mathbb{N}^n menores o iguales que h con respecto al pre-orden de Dickson. Se verifican las siguientes propiedades:*

- i) *El grado promedio de \mathcal{E}_h , el grado de h en $G(\mathcal{E}_h)$ y la dimensión DS de \mathcal{E}_h coinciden. Es decir, tenemos: $\deg(h) = \sharp(DS(h)) = \text{avd}(\mathcal{E}_h) = \dim_{DS}(\mathcal{E}_h)$.*
- ii) *La dimensión exponencial de \mathcal{E}_h verifica:*

$$\dim_E(\mathcal{E}_h) := \min \left\{ n, \left\lfloor \left(\sum_{i \in DS(h)} \log_2(1 + h_i) \right) \right\rfloor \right\}.$$

En particular, se tienen las siguientes desigualdades: $\text{outdeg}_{\text{av}}(\mathcal{E}_h) < \text{avd}(\mathcal{E}_h) = \dim_{DS}(\mathcal{E}_h) \leq \dim_E(\mathcal{E}_h)$.

La última desigualdad es una igualdad si y solamente si \mathcal{E}_h es un pseudo-cubo de dimensión n (o, equivalentemente, si $h \in (\mathbb{Z}^\times)^n$, donde $\mathbb{Z}^\times := \mathbb{Z} \setminus \{0\}$ o, también equivalentemente, si $DS(h) = [n]$). Ejemplos para los cuales la última desigualdad es estricta no son difíciles de obtener con estas últimas afirmaciones.

DEMOSTRACIÓN. Probamos las afirmaciones del modo siguiente:

- *Afirmación i):* ya hemos probado las 2 primeras igualdades en la Proposición 2.3.2. Antes de proceder con la prueba de la igualdad restante, recordamos algunos elementos de esa proposición, que emplearemos en esta demostración. Sea $DS(h) = m$ y el conjunto de las aristas unarias del escalón \mathcal{E}_h , que se puede expresar como:

$$E^{(1)}(\mathcal{E}_h) = \{e_{\pi_i(f),i} : f \in \mathcal{E}_h, i \notin DS(h)\}.$$

Por tanto, si restringimos el escalón acorde a $S = DS(h)$, tenemos que $\mathcal{E}_h|_{DS(h)}$ es un pseudo-cubo cerrado hacia abajo de dimensión DS igual a m .

Además, dado cualquier otro conjunto $S \subseteq [n]$, si existe algún $i \in S$ tal que $i \notin DS(h)$, entonces la proyección resultante $\mathcal{E}_h|_S$ tendrá al menos una arista unaria de la forma $e_{\pi_i(g),i}$, con $g \in \mathcal{E}_h$ cualquiera. En consecuencia, $DS(h)$ es maximal en cuanto a la dimensión del pseudo-cubo inducido, concluyendo así que $\deg(h) = m = \dim_{DS}(\mathcal{E}_h)$, donde la primera igualdad es debida a la condición de pseudo-cubo.

- *Afirmación ii):* sea S un subconjunto arbitrario de $[n]$ que contiene a $DS(h)$, es decir, $DS(h) \subseteq S \subseteq [n]$. Empleando la Afirmación ii) de la Proposición 2.3.2, tenemos que:

$$\sharp(\mathcal{E}_h|_S) = \prod_{i \in DS(h)} (1 + h_i).$$

La siguiente desigualdad siempre se cumple, puesto que $1 + h_i \geq 2$, para todo $i \in DS(h)$:

$$\prod_{i \in DS(h)} (1 + h_i) \geq 2^{\sharp(S)},$$

por lo que la dimensión exponencial es un valor entre m y n , donde recordamos que $m = DS(h)$. Aplicando el logaritmo de base 2 a ambos lados de la desigualdad, se tiene:

$$\sum_{i \in DS(h)} \log_2(1 + h_i) \geq \sharp(S).$$

Teniendo en cuenta que la dimensión exponencial es un número entero y que está acotado por n , concluimos la prueba de esta afirmación.

Las últimas desigualdades se siguen inmediatamente de la Proposición 2.3.2 y las afirmaciones previas. Los ejemplos citados son inmediatos. \square

3.3. Pseudo-cubos cerrados hacia abajo.

Comenzamos con la siguiente caracterización de los pseudo-cubos cerrados hacia abajo:

PROPOSICIÓN 3.3.1 (PSEUDO-CUBOS CERRADOS HACIA ABAJO). *Sea $H \subseteq \mathcal{Y}^n$, $n \geq 1$, una clase de conceptos finita (escalera finita). Las siguientes dos propiedades son equivalentes:*

- i) H es un pseudo-cubo cerrado hacia abajo de dimensión n .
- ii) H es una unión finita de escalones de dimensión DS igual a n . Es decir, existen $h_1, \dots, h_s \in \mathbb{N}^n$ una familia finita de exponentes monomiales tales que:
 - (a) El grado promedio, la dimensión DS y la dimensión exponencial de cada escalón determinado por h_k es n , i.e. $\forall k \in \{1, \dots, s\}$, $\text{avd}(\mathcal{E}_{h_k}) = \dim_{DS}(\mathcal{E}_{h_k}) = \dim_E(\mathcal{E}_{h_k}) = n$.
 - (b) H verifica:

$$(3.3.1) \quad H := \bigcup_{i=1}^s \mathcal{E}_{h_i}.$$

DEMOSTRACIÓN. Probemos las dos implicaciones para enfatizar el significado de esta equivalencia:

- \implies : Si H es cerrado hacia abajo, la Proposición 1.1.4 implica que H admite una descomposición como en la Identidad (3.3.1), donde h_1, \dots, h_s están determinados de modo único y son los elementos maximales (distintos) de H con respecto al pre-orden de Dickson \preceq . Supongamos que tenemos esa descomposición y que H es cerrado hacia abajo, entonces $\sharp(DS(h_i)) = n$ para cada $i \in [s]$. Sin pérdida de la generalidad, supongamos que $\dim_{DS}(\mathcal{E}_{h_1}) < n$. Entonces, conforme a la Proposición 2.3.2, $\sharp(DS(h_1)) < n$. Equivalentemente, suponiendo que $h_1 = (h_{1,1}, \dots, h_{1,n})$ debería existir algún $i \in [n]$ tal que $h_{1,i} = 0$. Consideremos entonces la proyección $\pi_i : H \rightarrow \mathcal{Y}^{n-1}$, supongamos $f = \pi_i(h_1)$ y sea $e_{f,i}$ la fibra en H asociada. Obviamente $h_1 \in e_{f,i}$. Probemos que $e_{f,i} = \{h_1\}$. Supongamos $g = (g_1, \dots, g_n) \in e_{f,i}$. Entonces, $\pi_i(g) = \pi_i(h_1)$ y, por tanto, tenemos $g_j = h_{1,j}$, $\forall j \neq i$. Como $h_1 = (g_1, \dots, g_{i-1}, 0, g_{i+1}, \dots, g_n)$ y $g_i \geq 0$, concluimos que $h_1 \preceq g$. Como h_1 es un elemento maximal para \preceq en H , entonces $h = h_1$ y $e_{f,i} = \{h_1\}$. Por tanto, H no puede ser un pseudo-cubo de dimensión n . Es decir, si H es un pseudo-cubo cerrado hacia abajo, entonces, por la Proposición 3.2.10 concluimos: $\text{avd}(\mathcal{E}_{h_k}) = \dim_{DS}(\mathcal{E}_{h_k}) = n, \forall k \in \{1, \dots, s\}$. Como la dimensión exponencial no puede ser estrictamente mayor que n , también tendremos que $\dim_E(\mathcal{E}_{h_k}) = n$ para cada k y las afirmaciones (a) y (b) se siguen.
- \impliedby : Como H es una unión finita de escalones, es inmediatamente cerrado hacia abajo. Por la Observación 1.1.5 cada descomposición de un conjunto cerrado hacia abajo H como en la Identidad (3.3.1) puede ser refinada hasta obtener una en la cual h_1, \dots, h_s sean los elementos maximales (distintos dos a dos) de H con respecto al pre-orden de Dickson. Supongamos que la descomposición descrita en la Identidad (3.3.1) es esa descomposición única de H . Veamos que H es un pseudo-cubo de dimensión DS, igual a n . Para ello, consideremos $i \in [n]$ y $h \in H$. Debe haber un elemento maximal h_k de H tal que $h \in \mathcal{E}_{h_k}$. Consideremos entonces la proyección $\pi_i : H \rightarrow \mathcal{Y}^{n-1}$. Sea $f := \pi_i(h)$ la proyección de h y sea $e_{f,i} \in E(H)$ la arista determinada por esa proyección. Observemos que:

$$h \in \mathcal{E}_{h_k} \cap e_{f,i} \subseteq e_{f,i},$$

y, aplicando (a), sabemos que \mathcal{E}_{h_k} es un pseudo-cubo de dimensión n . Entonces, tenemos:

$$\sharp(e_{f,i}) \geq \sharp(\mathcal{E}_{h_k} \cap e_{f,i}) \geq 2.$$

En conclusión, H es un pseudo-cubo de dimensión n . \square

COROLARIO 3.3.2. *Si $H \subseteq \mathcal{Y}^n$, $n \geq 1$, es un conjunto cerrado hacia abajo, las siguientes propiedades son equivalentes:*

- i) H es un pseudo-cubo de dimensión n .
- ii) Se verifica lo siguiente:

$$\text{outdeg}_{\text{av}}(H) < \text{avd}(H) = \dim_{DS}(H) = \dim_E(H) = n \leq 2\text{outdeg}_{\text{av}}(H).$$

iii) Dada una descomposición de H en sus componentes (escalones):

$$H =: \bigcup_{i=1}^s \mathcal{E}_{h_i},$$

entonces, se verifica lo siguiente

$$\text{avd}(H) = \frac{1}{s} \sum_{i=1}^s \deg(h_i) = \max\{\deg(h_1), \dots, \deg(h_s)\} = \dim_{\text{DS}}(H) = \dim_{\text{E}}(H) = n.$$

iv) Dada una descomposición de H en sus componentes (escalones):

$$H =: \bigcup_{i=1}^s \mathcal{E}_{h_i},$$

entonces, para cada i , $1 \leq i \leq s$, se verifica:

$$\text{avd}(\mathcal{E}_h) = \deg(h_i) = \dim_{\text{DS}}(\mathcal{E}_{h_i}) = \dim_{\text{E}}(\mathcal{E}_{h_i}) = \dim_{\text{DS}}(H) = \dim_{\text{E}}(H) = \text{avd}(H) = n.$$

DEMOSTRACIÓN. Probemos las equivalencias del modo siguiente:

- *Afirmación i*) \iff *Afirmación ii*): por la Observación 3.2.2, la dimensión exponencial de H es menor o igual que n y es mayor o igual que la dimensión exponencial de cualquier subconjunto (es hereditaria). Entonces, usando el Teorema 3.1.4 y la Afirmación ii) de la Proposición 3.3.1 anteriores, concluimos inmediatamente que si H es un pseudo-cubo de dimensión n , se verifica la siguiente igualdad:

$$\text{avd}(H) = \dim_{\text{DS}}(H) = \dim_{\text{E}}(H).$$

Como $n \geq 1$, la última dicotomía de la Proposición 2.2.1 implica que $\text{outdeg}_{\text{av}}(H) < \text{avd}(H)$. La otra desigualdad se sigue del Teorema 2.2.1. Para el recíproco, si $\text{avd}(H) = \dim_{\text{DS}}(H) = n$, entonces el Teorema 3.1.4 implica que H es un pseudo-cubo.

- *Afirmación iii*) \iff *Afirmación iv*): es inmediata.
- *Afirmación ii*) \implies *Afirmación iv*): como la dimensión DS es monótona respecto a la inclusión (véase Observación 3.2.1), tenemos:

$$(3.3.2) \quad \deg(h_i) = \dim_{\text{DS}}(\mathcal{E}_{h_i}) \leq \max\{\dim_{\text{DS}}(\mathcal{E}_{h_j}) : 1 \leq j \leq s\} \leq \dim_{\text{DS}}(H) = \dim_{\text{E}}(H) = \text{avd}(H) = n,$$

donde la primera igualdad es consecuencia de la Afirmación i) de la Proposición 3.2.10. Además, por la misma afirmación, se tiene que la dimensión DS del escalón \mathcal{E}_{h_i} es equivalente a $\sharp(\text{DS}(h_i))$. De esta forma, podemos concluir que las desigualdades de la Identidad (3.3.2) son en realidad igualdades, empleando lo que hemos discutido en el tercer ejemplo del Ejemplo 3.1.5, donde se observa que los elementos maximales no pueden estar en los ejes, es decir, todas las coordenadas son no nulas y, por tanto, se tiene que $\dim_{\text{DS}}(\mathcal{E}_{h_i}) = \sharp(\text{DS}(h_i)) = n$.

La igualdad restante $\dim_{\text{DS}}(\mathcal{E}_{h_i}) = \dim_{\text{E}}(\mathcal{E}_{h_i})$ es consecuencia directa del último comentario de la Proposición 3.2.10.

- *Afirmación iv*) \implies *Afirmación ii*): es una consecuencia inmediata de la Proposición 3.3.1. \square

EJEMPLO 3.3.3. Los resultados previos demuestran que tanto $\text{avd}(H)$ como $\text{outdeg}_{\text{av}}(H)$ son equivalentes a la dimensión DS y a la dimensión exponencial en el caso de clases que son cerradas hacia abajo y pseudo-cubos. Sin embargo, esto no es cierto en general cuando H no es un pseudo-cubo. Exhibimos el siguiente contra-ejemplo. Dado $\mathcal{Y}_p := \{0, 1, \dots, p-1\}$, entonces existe una sucesión de clases $F_p \subseteq \mathcal{Y}_p^2$ cerradas hacia abajo que satisface:

$$1 = \text{outdeg}_{\text{av}}(F_p) = \lim_{p \rightarrow \infty} \text{avd}(F_p) < \dim_{\text{DS}}(F_p) = \dim_{\text{E}}(F_p) \leq 2\text{outdeg}_{\text{av}}(F_p) < 2\text{avd}(F_p).$$

Consideremos las clases siguientes: $F_p := \{(0, 0), (0, 1), (1, 0), (1, 1)\} \cup \{(0, k) : 2 \leq k \leq p-2\}$. Estos conjuntos son cerrados hacia abajo, pero no son pseudo-cubos para $p \geq 1$. Como en el Ejemplo 3.2.4, contienen un pseudo-cubo de dimensión 2 y, por tanto, $\dim_{\text{DS}}(F_p) = \dim_{\text{E}}(F_p) = 2$. En número de vértices también satisface: $\sharp(F_p) = p+2$. El número de aristas satisface: $\sharp(E(F_p)) = p+2$, $\sharp(E^{(1)}(F_p)) = p-2$. Entonces, tenemos:

$$\text{outdeg}_{\text{av}}(F_p) = 2 - \frac{p+2}{p+2} = 1, \quad \text{avd}(F_p) = 2 - \frac{p-2}{p+2} = 1 + \frac{4}{p+2},$$

con lo que concluimos $1 = \text{outdeg}_{\text{av}}(F_p) = \lim_{p \rightarrow \infty} \text{avd}(F_p) < \dim_{\text{DS}}(F_p) = \dim_{\text{E}}(F_p) = 2 \leq 2\text{outdeg}_{\text{av}}(F_p)$.

LEMA 3.3.4. Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Sea $S \subseteq [n]$ un subconjunto no vacío. Sea $f \in H$ uno de esos vértices. Denotemos por $f' := f|_S$ la restricción de f vista como función $f : [n] \rightarrow \mathcal{Y}$. Entonces:

$$\mathcal{E}_f|_S = \{g|_S : g \in \mathcal{E}_f\} = \{g : S \rightarrow \mathcal{Y} : g \preceq f'\} = \mathcal{E}_{f'} \subseteq H|_S.$$

DEMOSTRACIÓN. La inclusión \subseteq es inmediata por la definición del pre-orden de Dickson. Si $h \preceq f$, entonces $h|_S \preceq f|_S$. Para la otra inclusión, supongamos, sin pérdida de la generalidad, que $S = [m]$ para algún $m \leq n$, $m \geq 1$. Supongamos que $h' \in \mathcal{E}_{f'}$. Entonces, tenemos:

$$h' = (h'(1), \dots, h'(m)) \preceq (f(1), \dots, f(m)) = f'.$$

Consideremos $h \in \mathcal{Y}^n$ dado mediante: $h := (h'(1), \dots, h'(m), 0, \dots, 0)$. Observemos que:

$$h \preceq (f(1), \dots, f(m), f(m+1), \dots, f(n)) \leftrightarrow f,$$

puesto que $f(j) \geq 0$ para cualquier $j \in [n]$ y $f(i) = f'(i)$ para $i \in [m]$. Como \mathcal{E}_f es cerrado hacia abajo, entonces $h \in \mathcal{E}_f$ y $h|_S = h'$. Lo que prueba la inclusión \supseteq y concluye la prueba del Lema. \square

Antes de continuar, introducimos un resultado conocido para ayudarnos:

LEMA 3.3.5. Sean $a_1, \dots, a_r, b_1, \dots, b_m \in \mathbb{R}$ números reales no negativos verificando las siguientes desigualdades:

$$b_1 \leq b_2 \leq \dots \leq b_m \leq a_1 \leq \dots \leq a_r.$$

Entonces, tenemos:

$$\frac{a_1 + a_2 + \dots + a_r + b_1 + \dots + b_m}{m+r} \leq \frac{a_1 + \dots + a_r}{r}.$$

DEMOSTRACIÓN. Aunque es elemental, incluimos una prueba por completitud del texto. Bajo nuestras hipótesis, tenemos:

$$\frac{a_1 + \dots + a_r}{r} \geq \frac{ra_1}{r} = \frac{ma_1}{m} \geq \frac{b_1 + \dots + b_m}{m}.$$

Por tanto, tenemos $b_1 + \dots + b_m \leq m \left(\frac{a_1 + \dots + a_r}{r} \right)$. Luego concluimos:

$$b_1 + \dots + b_m + a_1 + \dots + a_r \leq m \left(\frac{a_1 + \dots + a_r}{r} \right) + r \left(\frac{a_1 + \dots + a_r}{r} \right) = (m+r) \left(\frac{a_1 + \dots + a_r}{r} \right),$$

y se sigue el resultado anunciado. \square

El siguiente resultado es una mejora de un resultado de [BCDMY, 22]:

PROPOSICIÓN 3.3.6. Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos cerrada hacia abajo. Entonces, tenemos:

- i) Si $\dim_{\text{DS}}(H) \geq 1$, entonces $\text{outdeg}_{\text{av}}(H) < \text{avd}(H) \leq \dim_{\text{DS}}(H)$.
- ii) Si $\dim_{\text{E}}(H) \geq 1$, entonces $\text{outdeg}_{\text{av}}(H) < \text{avd}(H) \leq \dim_{\text{E}}(H)$.

DEMOSTRACIÓN. Supongamos una descomposición de H en sus componentes como en la Proposición 1.1.4:

$$H := \bigcup_{i=1}^s \mathcal{E}_{h_i},$$

donde $\{h_1, \dots, h_s\}$ son los elementos maximales (distintos dos a dos) de H con respecto al pre-orden de Dickson. Salvo alguna reordenación de los índices, supongamos $r, 1 \leq r \leq s$, tal que:

$$T := \sharp(DS(h_i)) = \max\{\sharp(DS(h_i)) : 1 \leq i \leq s\} \Leftrightarrow 1 \leq i \leq r.$$

Y $\sharp(DS(h_j)) < T$ para $r+1 \leq j \leq s$. Consideremos $S_i := DS(h_i)$, donde $1 \leq i \leq r$ y $\sharp(S_i) = T$. La restricción de H a S_i satisface:

$$H|_{S_i} := \bigcup_{k=1}^s \mathcal{E}_{h_k}|_{S_i} \supseteq \mathcal{E}_{h_i}|_{DS(h_i)}.$$

Consideremos $h'_i := h_i|_{DS(h_i)}$ y tendremos:

$$H|_{S_i} \supseteq \mathcal{E}_{h'_i},$$

y $h'_i(j) \geq 1$ para cada $j \in DS(h_i)$. Más aún, $\text{deg}(h'_i)$ en $G(\mathcal{E}_{h'_i})$ satisface, $T = \text{deg}(h'_i) = \text{deg}(h_i)$. Entonces, $\mathcal{E}_{h'_i}$ es un pseudo-cubo cerrado hacia abajo verificando: $\text{deg}(h'_i) = \sharp(DS(h_i)) = T = \dim_{\text{DS}}(\mathcal{E}_{h'_i})$. Entonces, $T \leq \dim_{\text{DS}}(H)$. Adicionalmente, para cada $h \in H$, tenemos: $\text{deg}(h) \leq \max\{\text{deg}(h_1), \dots, \text{deg}(h_s)\}$. Por tanto, tenemos:

$$\text{deg}(h) \leq \min\{\text{deg}(h_1), \dots, \text{deg}(h_r)\} = T = \max\{\text{deg}(h_1), \dots, \text{deg}(h_r)\} = \frac{\text{deg}(h_1) + \dots + \text{deg}(h_r)}{r}.$$

Por el Lema 3.3.5, salvo re-ordenación de los índices, concluimos la Afirmación i):

$$\text{avd}(H) = \frac{1}{\sharp(H)} \left(\sum_{h \in H} \text{deg}(h) \right) \leq \frac{\text{deg}(h_1) + \dots + \text{deg}(h_r)}{r} = T \leq \dim_{\text{DS}}(H).$$

Para la Afirmación *ii*) procedemos del mismo modo y concluiremos que para cada $i \in [r]$ tenemos:

$$2^T \leq \sharp(\mathcal{E}_{h'_i}) = \prod_{k \in DS(h_i)} (h_i(k) + 1),$$

y

$$\mathcal{E}_{h'_i} = \mathcal{E}_{h_i} \big|_{DS(h_i)} \subseteq H \big|_{DS(h_i)}.$$

Como $T = \sharp(DS(h_i))$, tenemos $2^T = 2^{\sharp(DS(h_i))} \leq \sharp(H \big|_{DS(h_i)})$. Finalmente, $T \leq \dim_{\mathbb{E}}(H)$ como se pretendía. \square

COROLARIO 3.3.7. *Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita, cerrada hacia abajo, y sea $m := \dim_{DS}(H)$. Entonces, para cada subconjunto $S \subseteq [n]$, tal que $\sharp(S) = m$ y tal que $H \big|_S$ contiene un pseudo-cubo de dimensión m , podemos suponer que existe un pseudo-cubo $\mathcal{C} \subseteq H \big|_S$ tal que:*

- \mathcal{C} es un pseudo-cubo cerrado hacia abajo.
- Se verifica $\text{avd}(\mathcal{C}) = \dim_{DS}(H) = \dim_{\mathbb{E}}(\mathcal{C}) \leq \dim_{\mathbb{E}}(H)$.
- \mathcal{C} es maximal con respecto a la inclusión (\subseteq) satisfaciendo las dos propiedades anteriores.

Más aún, $H \big|_S$ es él mismo un pseudo-cubo si y solamente si $\text{avd}(H \big|_S) = \dim_{DS}(H)$.

DEMOSTRACIÓN. Por el Lema 1.1.7, si H es cerrado hacia abajo y $S \subseteq [n]$ es un subconjunto, entonces $H \big|_S$ es también cerrado hacia abajo. Por la Proposición 1.1.4, $H \big|_S$ admite una descomposición en sus escalones:

$$H \big|_S := \bigcup_{k=1}^s \mathcal{E}_{h_k},$$

de tal modo que para cada $k \in [s]$:

$$h_k \notin \bigcup_{j \neq k} \mathcal{E}_{h_j}.$$

Re-ordenando los sub-índices si fuera necesario, podemos suponer $t \in [s]$ tal que:

$$\dim_{DS}(\mathcal{E}_k) = \sharp(S) = \dim_{DS}(H) \iff k \leq t.$$

Definamos $\mathcal{C} \subseteq H \big|_S$ mediante:

$$\mathcal{C} := \bigcup_{k=1}^t \mathcal{E}_{h_k}.$$

La Proposición 3.3.1 implica que \mathcal{C} es un pseudo-cubo cerrado hacia abajo que satisface lo requerido. \square

OBSERVACIÓN 3.3.8. Esto significa que para subconjuntos cerrados hacia abajo $H \subseteq \mathcal{Y}^n$, la dimensión DS puede caracterizarse del modo siguiente:

$$\dim_{DS}(H) := \max\{\sharp(S) : S \subseteq [n], \exists \mathcal{C} \subseteq H \big|_S, \text{avd}(\mathcal{C}) = \sharp(S)\}.$$

El corolario precedente prueba que si H es cerrado hacia abajo, entonces \leq se verifica. Recíprocamente, si $S \subseteq [n]$ es tal que $\exists \mathcal{C} \subseteq H \big|_S$, $\text{avd}(\mathcal{C}) = \sharp(S)$, entonces \mathcal{C} es un pseudo-cubo porque no hay ninguna arista de cardinal uno en $G(\mathcal{C})$. Entonces, $H \big|_S$ contiene un pseudo-cubo y $\sharp(S) \leq \dim_{DS}(H)$.

Shifting

Índice

4.1. Introducción: la noción de shift	35
4.2. Shifting, cerrados hacia abajo y bases monomiales	37
4.2.1. Las ideas T. Mészáros en [Mes, 20]	41
4.3. Shifting e invariantes de clases de conceptos finitas: un par de mejoras estrictas de [BCDMY, 22]	42

4.1. Introducción: la noción de shift

En [Hau, 1995], D. Haussler usó la técnica de “shifting” para producir una demostración del Lema de Sauer-Shelah-Perles y otros resultados en el caso de aprendizaje binario. Esta técnica aparece por vez primera en [Fra, 1983] e independientemente en [Aln, 1983] (véase también el papel del shifting en relación con la dualidad y la traza en [Par, 24b]). En [RBR, 06], los autores extienden este operador al caso del aprendizaje multi-clase y es una técnica analizada de nuevo en [BCDMY, 22]. Seguiremos esta última referencia para las notaciones, aunque las remozaremos para darles un buen sentido.

DEFINICIÓN 20 (SHIFTING DE UNA ARISTA). Como en capítulos precedentes, sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita y sea $G(H)$ el hiper-grafo de una inclusión sobre H . Para cada $i \in [n]$, sea $\pi_i : \mathcal{Y}^n \rightarrow \mathcal{Y}^{n-1}$ la proyección que “olvida” la i -ésima coordenada. Y, para cada $f \in \mathcal{Y}^{n-1}$, sea $e_{f,i} := \pi_i^{-1}(\{f\}) \cap H$ la fibra sobre f en la dirección determinada por π_i , que es el sustrato como subconjunto de H de la correspondiente arista en $E(H)$. Para cada $i \in [n]$, definimos el shift de la arista $e_{f,i}$ en la dirección de i mediante: $\mathbb{S}_i(e_{f,i}) \subseteq \mathcal{Y}^n$:

$$\mathbb{S}(e_{f,i}) := \{g \in \mathcal{Y}^n : \pi_i(g) = f, g(i) \in \{0, \dots, \#(e_{f,i}) - 1\}\}.$$

En otras palabras, cambiamos la arista $e_{f,i}$ para obtener otra arista del mismo cardinal, en la misma dirección y cerrada hacia abajo, preservando el orden \preceq sobre sus elementos. Esto se obtiene “empujando hacia abajo” los elementos de $e_{f,i}$ a través de la coordenada i -ésima de los elementos de $e_{f,i}$, preservando el orden de esa coordenada en el monoide $(\mathbb{N}, +)$ y el resto de coordenadas (i.e., preservamos f). En particular, $\mathbb{S}_i(e_{f,i})$ y $e_{f,i}$ son biyectables mediante una biyección que preserva el pre-orden de Dickson \preceq , pero hace que la columna i -ésima sea cerrada hacia abajo. Denotemos por $\widetilde{\mathbb{S}}_{f,i} : e_{f,i} \rightarrow \mathbb{S}_i(e_{f,i})$ a esa biyección entre aristas que preserva \preceq .

DEFINICIÓN 21 (SHIFTING DE UNA CLASE). Con las notaciones precedentes, definimos el shift de H en la dirección de π_i como el conjunto dado por la igualdad siguiente:

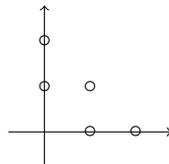
$$\mathbb{S}_i(H) := \bigcup_{f \in \pi_i(H)} \mathbb{S}_i(e_{f,i}).$$

Es decir, tenemos una aplicación:

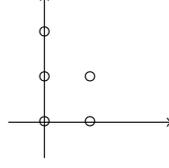
$$\mathbb{S}_i : 2^{\mathcal{Y}^n} \rightarrow 2^{\mathcal{Y}^n},$$

donde $2^{\mathcal{Y}^n}$ es la clase de los subconjuntos de \mathcal{Y}^n .

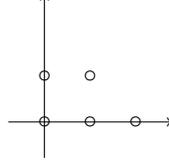
EJEMPLO 4.1.1 (EJEMPLO 19 DE [BCDMY, 22]). La dimensión DS puede crecer tras un shift como se prueba en el Ejemplo 19 de [BCDMY, 22]. Lo reproducimos por ilustrar el proceso de shift. Sea $\mathcal{Y} := \{0, 1, 2\}$ y sea $H \subseteq \mathcal{Y}^2$ dado por: $H := \{(0, 1), (0, 2), (1, 0), (1, 1), (2, 0)\}$, cuya imagen sería:



Aplicamos un shift en la dirección de π_1 (que “olvida” la primera coordenada) y tendremos: $H' := \mathbb{S}_1(H) := \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1)\}$, cuya imagen resulta ser:



Algo similar sucede si hacemos un shift en la dirección de π_2 : $H'' := \mathbb{S}_2(H) := \{(0, 0), (0, 1), (1, 0), (1, 1), (2, 0)\}$, cuya imagen sería:



En ambos casos recuperamos conjuntos cerrados hacia abajo que no son pseudo-cubos. Más aún, ambas transformaciones contienen un pseudo-cubo de dimensión 2. Así que tenemos la siguiente igualdad:

$$\dim_{\text{DS}}(H') = \dim_{\text{E}}(H') = \dim_{\text{DS}}(H'') = \dim_{\text{E}}(H'') = 2 > \dim_{\text{DS}}(H) = 1.$$

En ambos casos el número de aristas decrece (esto es estándar con el shifting, como veremos más adelante), pero también en número de aristas unarias disminuye. Es decir, tenemos:

$$1 = 2 - 5/5 = \text{outdeg}_{\text{av}}(H'') = \text{outdeg}_{\text{av}}(H') > \text{outdeg}_{\text{av}}(H) = 2 - 6/5 = 4/5,$$

y

$$9/5 = 2 - 1/5 = \text{avd}(H'') = \text{avd}(H') > \text{avd}(H) = 2 - 2/5 = 8/5.$$

OBSERVACIÓN 4.1.2. En primer lugar, es claro que el shift no modifica las proyecciones de los conjuntos en la dirección determinada por la proyección i -ésima. Esto es, tenemos $\pi_i(H) = \pi_i(\mathbb{S}_i(H))$. Como, además, fijada una dirección de proyección π_i , tenemos una partición de cada clase de conceptos en función de las fibras de esa proyección:

$$H := \bigsqcup_{f \in \pi_i(H)} e_{f,i},$$

podemos generar una biyección natural a partir de las biyecciones sobre cada una de las fibras de esa partición. Denotaremos a esa biyección mediante:

$$(4.1.1) \quad \widetilde{\mathbb{S}}_{H,i} : H \longrightarrow \mathbb{S}_i(H),$$

dada por la propiedad:

$$\widetilde{\mathbb{S}}_{H,i} \Big|_{e_{f,i}} = \widetilde{\mathbb{S}}_{f,i},$$

para cada $f \in \pi_i(H) = \pi_i(\mathbb{S}_i(H))$. Obsérvese que $\widetilde{\mathbb{S}}_{H,i}$ es una biyección que preserva el pre-orden de Dickson \preceq y, por tanto, conservamos los cardinales, i.e. $\sharp(H) = \sharp(\mathbb{S}_i(H))$. Además, tenemos una biyección entre las aristas asociadas a la dirección π_i de cada uno de los OIG's asociados a cada uno de los conjuntos, i.e., $\widetilde{\mathbb{S}}_{H,i}$ define una biyección entre $E_i(H)$ y $E_i(\mathbb{S}_i(H))$. En particular, tenemos $\sharp(E_i(H)) = \sharp(E_i(\mathbb{S}_i(H)))$. Finalmente, observemos que $\mathbb{S}_i(H)$ es cerrado hacia abajo en la dirección asociada a la proyección π_i . Pero no debe olvidarse que la biyección entre H y $\mathbb{S}_i(H)$ depende del conjunto H involucrado.

En [Hau, 1995] no sólo se considera la aplicación de un shifting, sino que podemos realizar una secuencia de shiftings, definida del modo siguiente:

DEFINICIÓN 22 (SECUENCIA DE SHIFTINGS). Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita y sea $\omega := \omega_1, \dots, \omega_k \in [n]^*$ una palabra sobre el alfabeto $[n]$. Definiremos el shift de H determinado por la palabra $\omega \in [n]^*$ como la composición:

$$(4.1.2) \quad \mathbb{S}_\omega = \mathbb{S}_{\omega_1} \circ \dots \circ \mathbb{S}_{\omega_r} : 2^{\mathcal{Y}^n} \longrightarrow 2^{\mathcal{Y}^n}$$

Para cada $H \subseteq \mathcal{Y}^n$, tendremos:

$$(4.1.3) \quad \mathbb{S}_\omega(H) := \mathbb{S}_{\omega_1}(\mathbb{S}_{\omega_2}(\dots(\mathbb{S}_{\omega_k}(H)))) .$$

OBSERVACIÓN 4.1.3 (LOS SHIFTS DETERMINAN BIYECCIONES QUE PRESERVAN \preceq). De nuevo debemos ser cuidadosos cuando tratemos elemento a elemento. Así, un shift \mathbb{S}_ω determinado por una palabra $\omega = \omega_1 \cdots \omega_k \in [n]^*$ determina una biyección entre los conjuntos $H \subseteq \mathcal{Y}^n$ y $\mathbb{S}_\omega(H)$, pero no es literalmente una composición de funciones, sino que toma la forma :

$$(4.1.4) \quad \widetilde{\mathbb{S}}_{H,\omega} := \widetilde{\mathbb{S}}_{H_1,\omega_1} \circ \widetilde{\mathbb{S}}_{H_2,\omega_2} \circ \cdots \circ \widetilde{\mathbb{S}}_{H_k,\omega_k},$$

donde $H_1 = H$ y para cada j , $2 \leq j \leq k$, se tiene $H_j := \mathbb{S}_i(H_{j-1})$. Por tanto, podemos concluir:

- i) Las aplicaciones $\widetilde{\mathbb{S}}_{H,\omega}$ son biyecciones entre H y $\mathbb{S}_\omega(H)$ que conservan \preceq (porque lo preserva cada una de las biyecciones que interviene en su definición).
- ii) $\sharp(H) = \sharp(\mathbb{S}_\omega(H))$.

OBSERVACIÓN 4.1.4. En [Hau, 1995] se demuestra (por lo demás evidente aplicando inducción y las observaciones precedentes) que, tras una secuencia de shifts para cada $H \subseteq \mathcal{Y}^n$ finito existe una palabra $\omega \in [n]^*$ que define una secuencia de shifts \mathbb{S}_ω , tal que la clase $H_\omega^* := \mathbb{S}_\omega(H)$, es cerrada hacia abajo con respecto al pre-orden de Dickson. Nótese que cada H tendrá asociada una (o varias) palabras $\omega := \omega(H)$ y que no es el mismo ω para todos los $H \subseteq \mathcal{Y}^n$. Sin embargo, el conjunto cerrado hacia abajo biyectable a H no es único y los invariantes pueden cambiar según la palabra elegida.

A modo de ejemplo retomemos la clase de conceptos H descrita en el Ejemplo 4.1.1 y observamos que tomando la palabra $\omega := 1 \in [n]^*$ ya tenemos que $H_{\omega^*} := \mathbb{S}_\omega(H)$ es cerrado hacia abajo para \preceq . Del mismo modo podemos tomar la palabra $\omega' := 2 \in [n]^*$ y considerar: $H_{\omega'}^* := \mathbb{S}_{\omega'}(H) = \mathbb{S}_2(H)$, que vuelve a ser cerrado hacia abajo pero ambos son distintos y tienen distintos invariantes. Trataremos de revisar eso más adelante en la Sección 4.2.1.

4.2. Shifting, cerrados hacia abajo y bases monomiales

En la conclusión de la Sección 1.1 concluíamos que las escaleras finitas se pueden obtener como complementarios de conjuntos de exponentes asociados a ciertos ideales, independientemente del orden monomial elegido. Ahora seguimos el camino inverso.

Fijamos un orden monomial (típicamente órdenes lexicográficos) y buscamos ideales que asocien las escaleras finitas con ideales especiales que resultan ser los ideales de polinomios que se anulan en los puntos de la propia escalera. Estas son las ideas de [Mes, 20] y su antecesor [FRR, 06]. La “prueba” de [Mes, 20] depende fuertemente de un resultado de [FRR, 06], quienes, a su vez, están fuertemente basados en elementos de Teoría de Juegos (el *lex game*). Como nos parecía que ya era hacer intervenir elementos muy diversos, nos hemos construido nuestra propia demostración del enunciado de [Mes, 20] sin usar nada más que elementos propios de la División de Weierstarss-Hironaka y las Bases de Gröbner. Comencemos con el ideal asociado a un subconjunto:

DEFINICIÓN 23. Sea K un cuerpo, $K \subseteq L$ una extensión de cuerpos y $V \subseteq \mathbb{A}^n(L)$ un conjunto finito. Llamaremos ideal asociado a V sobre K al ideal:

$$I_K(V) := \{f \in K[X_1, \dots, X_n] : f|_V \equiv 0\}.$$

Denotaremos mediante $K[V]$ al anillo de funciones polinomiales K -definibles en V , esto es,

$$K[V] := K[X_1, \dots, X_n]/I_K(V).$$

Cuando el contexto nos fije el cuerpo K escribiremos simplemente $I(V)$.

COROLARIO 4.2.1. Con las notaciones precedentes, supongamos $K = L$ y $V \subseteq \mathbb{A}^n(K)$ un conjunto finito. Entonces, para cualquier orden monomial \leq sobre el monoide $(\mathbb{N}^n, +)$ se tiene que $\mathbb{N}^n \setminus \exp(I_K(V))$ es un conjunto finito y, además, se tiene:

$$\sharp(V) = \dim_K(K[V]) = \sharp(\mathbb{N}^n \setminus \exp(I_K(V))).$$

Además, como en el Teorema 1.2.7, el conjunto

$$\mathcal{B}/I_K(V) := \{X_1^{\mu_1} \cdots X_n^{\mu_n} + I_K(V) : \mu := (\mu_1, \dots, \mu_n) \in \mathbb{N}^n \setminus \exp_{\leq}(I_K(V))\},$$

es una base de $K[V]$ como K -espacio vectorial, denominada base monomial asociada al orden monomial \leq y están en biyección con $\mathbb{N}^n \setminus \exp_{\leq}(I_K(V))$.

La prueba de este clásico corolario se sigue del teorema Chino de los Restos y se puede ver en [Par, 24a]. Como vimos en la Proposición 1.2.6, si $V \subseteq \mathbb{A}^n(K)$ es un conjunto finito, entonces $K[V]$ es una K -álgebra de Artin y el siguiente conjunto es cerrado hacia abajo:

$$\mathbb{N}^n \setminus \exp_{\leq} (I_K(V)),$$

para cualquier orden monomial \leq . A partir de este momento nos vamos a ocupar de estudiar una cierta forma de recíproco para lo cual usaremos solamente el orden lexicográfico estándar \leq_{lex} definido en el Ejemplo 1.2.2 de tal modo que:

$$X_n >_{\text{lex}} X_{n-1} >_{\text{lex}} \dots >_{\text{lex}} X_1.$$

Para simplificar un poco las notaciones, usaremos las siguientes notaciones para cualquier ideal $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$:

$$\exp_{\text{lex}}(\mathfrak{a}) := \exp_{\leq_{\text{lex}}}(\mathfrak{a}), \quad \text{Nexp}_{\text{lex}}(\mathfrak{a}) := \mathbb{N}^n \setminus \exp_{\leq_{\text{lex}}}(\mathfrak{a}).$$

Supongamos ahora $\mathcal{Y} := \{0, \dots, p-1\}$ como en secciones y capítulos anteriores. Supongamos que el cuerpo K satisface $\mathcal{Y} \subseteq K$ (con cualquier inmersión razonable) y sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Para cada $i \in [n]$, a diferencia de las proyecciones π_i usadas en secciones y capítulos precedentes, consideremos $\tilde{\pi}_i : \mathcal{Y}^n \rightarrow \mathcal{Y}^{n-i}$ la proyección que “olvida” las primeras i coordenadas de los elementos de \mathcal{Y}^n . Es decir,

$$\tilde{\pi}_i(x_1, \dots, x_n) := (x_{i+1}, \dots, x_n), \quad \forall (x_1, \dots, x_n) \in \mathcal{Y}^n.$$

Dado $q := (q_1, \dots, q_{n-i}) \in \mathcal{Y}^{n-i}$, denotaremos por $H(q)$ a la parte i -inicial de la fibra (en H) sobre q a partir de $\tilde{\pi}_i$, es decir,

$$(4.2.1) \quad H(q) := \{(\theta_1, \dots, \theta_i) \in \mathcal{Y}^i : (\theta_1, \dots, \theta_i, q_1, \dots, q_{n-i}) \in \tilde{\pi}_i^{-1}(\{q\}) \cap H\}.$$

LEMA 4.2.2 (DE ELIMINACIÓN CUANTIFICADA). *Con las notaciones precedentes, sea $\nu := (\nu_1, \dots, \nu_n) \in \mathbb{N}^n$ un exponente monomial. Supongamos que se verifica*

$$\#\{q \in \mathcal{Y} : (\nu_1, \dots, \nu_{n-1}) \in \text{Nexp}_{\text{lex}}(I(H(q)))\} \geq \nu_n + 1.$$

Entonces $\nu \in \text{Nexp}_{\text{lex}}(I(H))$.

DEMOSTRACIÓN. El caso $n = 1$ es obvio y no merece más discusión.

Supongamos $n \geq 2$ y definamos $B_1 := \{q \in \mathcal{Y} : (\nu_1, \dots, \nu_{n-1}) \in \text{Nexp}_{\text{lex}}(I(H(q)))\}$. Supongamos que $\#(B_1) \geq \nu_n + 1$ y razonemos por reducción al absurdo. Supongamos que $\nu \in \exp_{\text{lex}}(I(H))$. Entonces, existe un polinomio $f \in I(H)$ tal que:

$$(4.2.2) \quad f := X_1^{\nu_1} \cdots X_n^{\nu_n} + g(X_1, \dots, X_n),$$

donde los términos no nulos de $g(X_1, \dots, X_n)$ tienen exponentes estrictamente menores que ν para \leq_{lex} . Es decir, supongamos que:

$$(4.2.3) \quad g := \sum_{\mu <_{\text{lex}} \nu} a_{\mu} X_1^{\mu_1} \cdots X_n^{\mu_n}.$$

Ahora observamos que hay dos tipos de términos en g :

- *Caso 1:* Términos cuyo exponente monomial es de la forma $(\nu_1, \dots, \nu_{n-1}, k)$, en cuyo caso $k \leq \nu_n - 1$.
- *Caso 2:* $\mu := (\mu_1, \dots, \mu_{n-1}, r) \in \mathbb{N}^n$, con r cualquiera, pero $(\mu_1, \dots, \mu_{n-1}) <_{\text{lex}} (\nu_1, \dots, \nu_{n-1})$.

Agrupemos los exponentes monomiales que responden al Caso 1 y definamos:

$$h_n(X_n) := X_n^{\nu_n} + \sum_{k=0}^{\nu_n-1} a_{(\nu_1, \dots, \nu_{n-1}, k)} X_n^k.$$

Entonces reescribimos el polinomio f descrito en la Igualdad (4.2.2) del modo siguiente:

$$f := X_1^{\nu_1} \cdots X_{n-1}^{\nu_{n-1}} h_n(X_n) + h(X_1, \dots, X_n),$$

donde todos los términos no nulos de h pertenecen al Caso 2. Es decir, h tiene la forma siguiente:

$$h := \sum_{\theta \in J} a_{\theta} X_1^{\theta_1} \cdots X_n^{\theta_n},$$

donde $J \subseteq \mathbb{N}^n$ es un conjunto finito y $\theta := (\theta_1, \dots, \theta_n) \in J$ si y solamente si:

$$(\theta_1, \dots, \theta_{n-1}) <_{\text{lex}} (\nu_1, \dots, \nu_{n-1}).$$

Consideremos ahora $b \in \mathcal{Y}$ y el polinomio $f(X_1, \dots, X_{n-1}, b) := X_1^{\nu_1} \cdots X_{n-1}^{\nu_{n-1}} h_n(b) + h(X_1, \dots, X_{n-1}, b)$.

Por la forma en la que hemos elegido f se tiene que:

- $f(X_1, \dots, X_{n-1}, b)$ se anula en $H(b)$.
- Si $h_n(b) \neq 0$, $\ell t_{\leq_{\text{lex}}}(f(X_1, \dots, X_{n-1}, b)) = h_n(b) X_1^{\nu_1} \cdots X_{n-1}^{\nu_{n-1}}$.

Por tanto, si $h_n(b) \neq 0$, entonces $(\nu_1, \dots, \nu_{n-1}) \in \exp_{1\text{ex}}(I(H(b)))$ y, por tanto, $b \notin B_1$. En conclusión,

$$B_1 \subseteq \{b \in \mathcal{Y} : h_n(b) = 0\},$$

pero como h_n es un polinomio univariado de grado ν_n , se tendrá que el número de raíces posibles está acotado por ν_n , es decir $\sharp(B_1) \leq \nu_n$ contradiciendo nuestra hipótesis. Y el Lema queda probado. \square

LEMA 4.2.3 (EXISTENCIA DE FACTORES TRIANGULARES). *Con las mismas notaciones precedentes, sea $\nu := (\nu_1, \dots, \nu_n) \in \mathbb{N}^n$. Supongamos que $n \geq 2$ y que se verifica:*

$$\sharp(\{q \in \mathcal{Y} : (\nu_1, \dots, \nu_{n-1}) \in \text{Nexp}_{1\text{ex}}(I(H(q)))\}) \leq \nu_n.$$

Entonces, existirán $g_1, \dots, g_n \in [X_1, \dots, X_n]$ tales que verifican:

- i) $g_i \in K[X_i, \dots, X_n]$ es un polinomio mónico con respecto a la variable X_i , verificando $\deg_{X_i}(g_i) \leq \nu_i$, para cada i , $1 \leq i \leq n$.
- ii) El polinomio $h := \prod_{i=1}^n g_i \in I(H)$.

En particular, $(\nu_1, \dots, \nu_n) = \exp_{\leq 1\text{ex}}(h) \in \exp_{1\text{ex}}(I(H))$.

DEMOSTRACIÓN. Comencemos con el caso $n = 2$. Definamos $B_1 \subseteq \mathcal{Y}$ como el conjunto:

$$B_1 := \{b \in \mathcal{Y} : \nu_1 \in \text{Nexp}_{1\text{ex}}(I(H(b)))\},$$

y supongamos $\sharp(B_1) \leq \nu_1$. Entonces, existe un polinomio $g_2 \in K[X_2]$ de grado acotado por ν_2 y tal que g_2 se anula en B_1 . Nótese que, en particular, g_2 se anula en:

$$\bigcup_{b \in B_1} H(b) \times \{b\}.$$

Consideremos ahora $B_2 := \{a \in \mathcal{Y} : \nu_1 \in \exp_{1\text{ex}}(I(H(a)))\}$. Como $\nu_1 \in \exp_{1\text{ex}}(I(H(a)))$, para cada $a \in B_2$ existirá un polinomio univariado mónico $g_1^{(a)} \in K[X_1]$ tal que se verifica: $\deg(g_1^{(a)}) \leq \nu_1$, y tal que $g_1^{(a)}$ se anula en $H(a)$. Es decir, para cada $a \in B_2$, podemos escribir, sin pérdida de la generalidad, que se tiene:

$$g_1^{(a)} := X_1^{\nu_1} + \sum_{k=0}^{\nu_1-1} a_k^{(a)} X_1^k.$$

Nótese que si $d_1 := \deg(g_1^{(a)}) < \nu_1$ bastaría con multiplicar por $X_1^{\nu_1-d_1}$ a $g_1^{(a)}$ y tendríamos la misma expresión. Aplicando interpolación de Lagrange, consideremos los polinomios:

$$g_{1,k}(X_2) := \sum_{a \in B_2} a_k^{(a)} \frac{\prod_{a' \in B_1 \setminus \{a\}} (X_2 - a')}{\prod_{a' \in B_1 \setminus \{a\}} (a - a')}.$$

Obsérvese que $g_{1,k}(a) = a_k^{(a)}$ para cada $a \in B_2$ y consideremos, finalmente, el polinomio:

$$g_1(X_1, X_2) := X_1^{\nu_1} + \sum_{k=0}^{\nu_1-1} g_{1,k}(X_2) X_1^k.$$

Observamos que $g_1(X_1, a) = g_1^{(a)}(X_1)$ para cada $a \in B_2$ y, por tanto, g_1 se anula sobre la unión siguiente:

$$\bigcup_{a \in B_2} H(a) \times \{a\}.$$

En particular, el polinomio $g = g_1(X_1, X_2)g_2(X_2)$ se anula en todo H porque:

$$H := \left(\bigcup_{b \in B_1} H(b) \times \{b\} \right) \cup \left(\bigcup_{a \in B_2} H(a) \times \{a\} \right).$$

Además es claro que, usando nuestro orden lexicográfico, $\ell t_{\leq 1\text{ex}}(h) = X_1^{\nu_1} X_2^{\nu_2}$.

Para el paso inductivo con $n > 2$ basta con repetir un proceso similar, que reproducimos por ser tan completos como sea posible. Definamos, de modo análogo:

$$B_1 := \{b \in \mathcal{Y} : (\nu_1, \dots, \nu_{n-1}) \in \text{Nexp}_{1\text{ex}}(I(H(b)))\},$$

y supongamos que $\sharp(B_1) \leq \nu_n$.

Como en el caso $n = 2$, existirá $g_n \in K[X_n]$ un polinomio univariado de grado ν_n , mónico con respecto a la (única) variable X_n y tal que g_n se anula en B_1 . En particular, se tiene:

$$\bigcup_{b \in B_1} H(b) \times \{b\} \subseteq V_{\mathbb{A}^n(K)}(g_n) := \{\zeta = (z_1, \dots, z_n) \in \mathcal{Y}^n : g_n(\zeta) = g_n(z_n) = 0\}.$$

Consideremos ahora el conjunto B_2 , disjunto con B_1 tal que:

$$B_2 := \{a \in \mathcal{Y} : (\nu_1, \dots, \nu_{n-1}) \in \exp_{\mathbf{lex}}(I(H(a)))\}.$$

Dado $a \in B_1$, aplicando el Lema 4.2.2 anterior a $H(a)$, como $(\nu_1, \dots, \nu_{n-1}) \in \exp_{\mathbf{lex}}(I(H(a)))$, entonces el siguiente conjunto no puede tener cardinal mayor o igual que $\nu_{n-1} + 1$:

$$B_1(a) := \{c \in \mathcal{Y} : (\nu_1, \dots, \nu_{n-2}) \in \text{Nexp}_{\mathbf{lex}}(I(H(a)(c)))\}.$$

Como $\sharp(B_1(a)) \leq \nu_{n-1}$, aplicado la hipótesis inductiva a $B_1(a)$ existirán polinomios $g_1^{(a)}, \dots, g_{n-1}^{(a)}$ en el anillo $K[X_1, \dots, X_{n-1}]$, tales que se verifica:

- i) $g_i^{(a)} \in K[X_i, \dots, X_{n-1}]$ es un polinomio mónico con respecto a la variable X_i , $\deg_{X_i}(g_i^{(a)}) \leq \nu_i$
- ii) El polinomio $h^{(a)}$ siguiente se anula en $H(a)$, es decir:

$$h^{(a)} := \prod_{i=1}^{n-1} g_i^{(a)} \in I(H(a)).$$

Como $B_2 \subseteq \mathcal{Y} \subseteq K$ y K es un cuerpo, podemos considerar los anillos $R_i := K[X_i, \dots, X_{n-1}]$ y podemos hacer interpolación de Lagrange sobre $R_i[X_n]$ siempre que el denominador sea una constante no nula en K . Dicho de otro modo, supongamos:

$$g_i^{(a)} := X_i^{\nu_i} + \sum_{k=0}^{\nu_i-1} g_{i,k}^{(a)} X_i^k,$$

donde $g_{i,k}^{(a)} \in R_{i+1} = K[X_{i+1}, \dots, X_{n-1}]$ son los coeficientes con respecto a las variables X_i . Aplicando Lagrange, consideremos los polinomios:

$$g_{i,k}(X_{i+1}, \dots, X_n) := \sum_{a \in B_2} g_{i,k}^{(a)}(X_{i+1}, \dots, X_{n-1}) \frac{\prod_{a' \in B_2 \setminus \{a\}} (X_n - a')}{\prod_{a' \in B_2 \setminus \{a\}} (a - a')}.$$

Nótese que podemos hacer esta interpolación porque el discriminante es un elemento no nulo en K . Además, se verifica que $g_{i,k} \in K[X_{i+1}, \dots, X_n]$ y se tiene que

$$g_{i,k}(X_{i+1}, \dots, X_{n-1}, a) = g_{i,k}^{(a)}(X_{i+1}, \dots, X_{n-1}),$$

para cualquier $a \in B_2$. Definamos, finalmente, el siguiente polinomio:

$$g_i := X_i^{\nu_i} + \sum_{k=0}^{\nu_i-1} g_{i,k}(X_{i+1}, \dots, X_n) X_i^k.$$

Observamos que $g_i(X_i, \dots, X_{n-1}, a) = g_i^{(a)}(X_i, \dots, X_{n-1})$. Definiendo

$$h_0 := \prod_{i=1}^{n-1} g_i,$$

concluimos que para cada $a \in B_2$ y para todo $(z_1, \dots, z_{n-1}) \in H(a)$ se tiene:

$$h_0(z_1, \dots, z_{n-1}, a) = \prod_{i=1}^{n-1} g_i(z_1, \dots, z_{n-1}, a) = \prod_{i=1}^{n-1} g_i^{(a)}(z_1, \dots, z_{n-1}) = h^{(a)}(z_1, \dots, z_{n-1}) = 0.$$

Por tanto, $h_0 \in H(a) \times \{a\}$ para cualquier $a \in B_2$. Añadiendo $g_n \in K[X_n]$, descrito anteriormente, que se anula sobre B_1 (y, por tanto, sobre $H(b) \times \{b\}$ para cualquier $b \in B_1$), tendremos la familia buscada con g_1, \dots, g_n y el polinomio $h := \prod_{i=1}^n g_i = h_0 \cdot g_n$ satisface que $h \in I(H)$ porque tenemos la descomposición de H dada mediante:

$$H = \left(\bigcup_{b \in B_1} H(b) \times \{b\} \right) \cup \left(\bigcup_{a \in B_2} H(a) \times \{a\} \right).$$

La razón es que g_n se anula en la primera unión, mientras que h_0 se anula en la segunda. Finalmente, por haber elegido el orden lexicográfico $\leq_{\mathbf{lex}}$ concluimos que:

$$\ell_{\leq_{\mathbf{lex}}}(h) = X_1^{\nu_1} \cdots X_n^{\nu_n},$$

y, por tanto, $(\nu_1, \dots, \nu_n) \in \exp_{\mathbf{lex}}(I(H))$, como se pretendía. \square

Estos dos lemas nos generan una prueba nueva, totalmente independiente y original, del siguiente teorema debido a [FRR, 06]:

TEOREMA 4.2.4 ([FRR, 06]). *Sea \mathcal{Y} un conjunto finito y K un cuerpo, supongamos $\mathcal{Y} \subseteq K$ (inmerso de alguna forma) y sea $H \subseteq \mathcal{Y}^n$. Son equivalentes para un exponente monomial $\nu := (\nu_1, \dots, \nu_n) \in \mathbb{N}^n$:*

- i*) $\nu \in \text{Nexp}_{1\text{ex}}(I(H))$.
- ii*) se verifica la siguiente desigualdad:

$$\#(\{b \in \mathcal{Y} : (\nu_1, \dots, \nu_{n-1}) \in \text{Nexp}_{1\text{ex}}(I(H(b)))\}) \geq \nu_n + 1.$$

DEMOSTRACIÓN. Cada implicación es uno de los lemas:

- *ii*) \implies *i*): Es el Lema 4.2.2
- *i*) \implies *ii*): Es el contrarrecíproco del Lema 4.2.3 tras recordar que $\text{Nexp}_{1\text{ex}}(I(H)) = \mathbb{N}^n \setminus \text{exp}_{1\text{ex}}(I(H))$. \square

4.2.1. Las ideas T. Mészáros en [Mes, 20]. Armado con el poderoso Teorema 4.2.4, cuya prueba alternativa (completamente original) acabamos de aportar, [Mes, 20] obtiene una sencilla caracterización que liga el shifting con las bases monomiales. Vamos a tratar de recordarlo en esta subsección.

Como en secciones y capítulos precedentes, tenemos un conjunto finito $\mathcal{Y} := \{0, \dots, p-1\}$ un cuerpo con una inmersión $\mathcal{Y} \subseteq K$ y una clase de conceptos finita $H \subseteq \mathcal{Y}^n$. Consideremos la estructura de hiper-grafo de una inclusión $(G(H) := (H, E(H)))$ discutida con anterioridad y los operadores de shifting para cada $i \in [n]$:

$$\mathbb{S}_i : 2^{\mathcal{Y}^n} \longrightarrow 2^{\mathcal{Y}^n},$$

definidas anteriormente. Por último, para cada $b \in \mathcal{Y}^n$ sea $H(b)$ el conjunto definido en (4.2.1):

$$H(b) := \{(q_1, \dots, q_{n-1}) \in \mathcal{Y}^{n-1} : (q_1, \dots, q_{n-1}, b) \in H\}.$$

Dado $H \subseteq \mathcal{Y}^n$ y dado $b \in \mathcal{Y}^n$ tenemos dos estructuras de hiper-grafo de una inclusión:

- $G(H) := (H, E(H))$,
- $G(H(b)) := (H(b), E(H(b)))$.

Sobre ambos hiper-grafos podemos definir las correspondientes operaciones de shifting y las palabras asociadas en $[n]^*$. En particular, nos interesan los dos conjuntos siguientes:

$$\mathbb{S}_{x'}(H) \quad , \quad \mathbb{S}_{x'}(H(b)),$$

donde $x' := n-1 \ n-2 \ \dots \ 2 \ 1 \in [n]^*$. La siguiente proposición nos permite relacionar el resultado de ambos operadores:

PROPOSICIÓN 4.2.5. *Con las notaciones precedentes se tiene:*

$$\mathbb{S}_{x'}(H(b)) \times \{b\} \subseteq \mathbb{S}_{x'}(H).$$

De hecho, $\mathbb{S}_{x'} \times \{b\}$ está contenido en la fibra:

$$\{(\mu_1, \dots, \mu_n) \in \mathbb{S}_{x'}(H) : \mu_n = b\} \subseteq \mathbb{S}_{x'}(H).$$

DEMOSTRACIÓN. El argumento es sencillo: los shifts involucrados en $\mathbb{S}_{x'}$ no tocan en ningún caso la última coordenada de los elementos de \mathcal{Y}^n , que permanece inalterada. \square

Con ella probamos el siguiente teorema:

TEOREMA 4.2.6 (BASES MONOMIALES (PARA $\leq_{1\text{ex}}$) DE $K[H]$ COINCIDEN CON $\mathbb{S}_{n \dots 1}(H)$). *Con las notaciones precedentes, se tiene:*

$$(4.2.4) \quad \text{Nexp}_{1\text{ex}}(I(H)) = (\mathbb{N}^n \setminus \text{exp}_{\leq_{1\text{ex}}}(I(H))) = \mathbb{S}_{n \dots 1}(H).$$

DEMOSTRACIÓN. Hagamos la demostración por inducción. El caso $n = 1$ es obvio, así que supondremos $n \geq 2$. Observemos la siguiente cadena de igualdades:

$$\#(\mathbb{S}_{n \dots 1}(H)) = \#(H) = \dim_K(K[H]) = \#(\text{Nexp}_{1\text{ex}}(I(H))).$$

Por tanto, bastará con probar una de las inclusiones de la Identidad (4.2.4) para tener la igualdad por ser ambos conjuntos del mismo cardinal. Así que nos limitaremos a probar la inclusión \subseteq . Consideremos $(\nu_1, \dots, \nu_n) \in \text{Nexp}_{1\text{ex}}(I(H))$ y la palabra $x' := n-1 \ n-2 \ \dots \ 2 \ 1 \in [n]^*$. Definamos el conjunto:

$$B_1 := \{b \in \mathcal{Y} : (\nu_1, \dots, \nu_{n-1}) \in \text{Nexp}_{1\text{ex}}(I(H(b)))\}.$$

Por el Teorema 4.2.4 precedente, se tiene que:

$$\#(B_1) = \#(\{b \in \mathcal{Y} : (\nu_1, \dots, \nu_{n-1}) \in \text{Nexp}_{1\text{ex}}(I(H(b)))\}) \geq \nu_n + 1.$$

Por otro lado, aplicando la hipótesis inductiva, para cada $b \in \mathcal{Y}$ tal que $H(b) \neq \emptyset$ se tiene:

$$\text{Nexp}_{1\text{ex}}(I(H(b))) = \mathbb{S}_{x'}(H(b)).$$

Como $x' = n - 1 \ n - 1 \ \cdots \ 2 \ 1 \in [n]^*$, la Proposición 4.2.5 precedente implica que para cada $b \in B_1$ se tiene:

$$(4.2.5) \quad (\nu_1, \dots, \nu_{n-1}, b) \in \mathbb{S}_{x'}(H(b)) \times \{b\} \subseteq \mathbb{S}_{x'}(H).$$

Tomando $f := (\nu_1, \dots, \nu_{n-1}) \in \mathcal{Y}^{n-1}$, podemos considerar la arista $e_{f,n}$ del grafo $G(\mathbb{S}_{x'}(H))$ dada mediante:

$$e_{f,n} = \mathbb{S}_{x'}(H) \cap \pi_n^{-1}(\{f\}),$$

donde $\pi_n : \mathcal{Y}^n \rightarrow \mathcal{Y}^{n-1}$ es, como siempre, la proyección que “olvida” la última coordenada. Pero por la inclusión descrita en la Ecuación (4.2.5), esto significa que $\{f\} \times \{b\} = (\nu_1, \dots, \nu_{n-1}, b) \in \mathbb{S}_{x'}(H)$ y, obviamente, π_n proyecta $(\nu_1, \dots, \nu_{n-1}, b)$ sobre f . Es decir, tenemos probado el contenido: $e_{f,n} \supseteq \{(\nu_1, \dots, \nu_{n-1})\} \times B_1$. Por tanto, podemos concluir la siguiente desigualdad de cardinales:

$$\#(e_{f,n}) \geq \#(B_1) \geq \nu_n + 1.$$

Entonces, $\mathbb{S}_n(e_{f,n}) \supseteq \{(\nu_1, \dots, \nu_{n-1}, t) : 0 \leq t \leq \nu_n\}$. Luego,

$$(\nu_1, \dots, \nu_n) \in \mathbb{S}_n(\mathbb{S}_{x'}(H)) = \mathbb{S}_{n \ n-1 \ \dots \ 1}(H).$$

Hemos probado, por tanto, que $\text{Nexp}_{1\text{ex}}(I(H)) \subseteq \mathbb{S}_{n \ n-1 \ \dots \ 1}(H)$ y el resultado se sigue porque ambos conjuntos tienen el mismo cardinal. \square

Ya vimos en el Corolario 1.2.10 una caracterización de los conjuntos finitos cerrados hacia abajo como complementarios de conjunto de exponentes monomiales de ideales cuyos anillos residuales son K -álgebras de Artin. Ahora podemos ser más finos en la forma siguiente:

COROLARIO 4.2.7 (H ES CERRADO HACIA ABAJO SII $H = \text{Nexp}_{1\text{ex}}(I(H))$). *Con las notaciones precedentes, un subconjunto $H \subseteq \mathcal{Y}^n$ es cerrado hacia abajo si y solamente si $H = \text{Nexp}_{1\text{ex}}(I(H))$, donde $I(H)$ es el conjunto de todos los polinomios $p \in K[X_1, \dots, X_n]$ que sea anulados en H .*

DEMOSTRACIÓN. Ya hemos visto que los conjuntos de la forma $\text{Nexp}_{1\text{ex}}(I(H))$ son cerrados hacia abajo en la Proposición 1.2.6. De hecho, lo vimos para cualquier ideal. En el Corolario 1.2.10 también vimos que todo conjunto finito cerrado hacia abajo es el complementario de $\text{exp}_{\leq}(\mathfrak{a})$ para cualquier orden monomial y para un ideal \mathfrak{a} construido ad hoc. Ahora tenemos una caracterización aún más fuerte. Obsérvese que si H es cerrado hacia abajo, entonces es invariante por la acción de los shifts, es decir, para cualquier $i \in [n]$ $\mathbb{S}_i(H) = H$. Por tanto, si H es cerrado hacia abajo, se tiene $\mathbb{S}_n \dots \mathbb{S}_1(H) = H$, lo cual se prueba de manera sencilla por inducción. Pero el Teorema 4.2.6 anterior nos precisa un orden y un ideal cuya base monomial coincide con H , es decir, $H = \text{Nexp}_{1\text{ex}}(I(H))$, lo que prueba la equivalencia buscada. \square

4.3. Shifting e invariantes de clases de conceptos finitas: un par de mejoras estrictas de [BCDMY, 22]

Hasta este momento hemos visto cuatro elementos esenciales en el estudio de la parte combinatoria del aprendizaje multi-clase:

- i) El hiper-grafo de una inclusión OIG (descrito en la Sección 2.1).
- ii) Los conjuntos cerrados hacia abajo (cuyo significado se ha observado en la Subsección 1.2.2 y en la Sección 4.2 precedentes).
- iii) Diversas “medidas” de las clases de conceptos (potencialmente ligadas al aprendizaje multi-clase). Algunas conocidas como $\text{avd}(H)$, $\text{dim}_{\text{DS}}(H)$, $\text{dim}_{\text{E}}(H)$, $\text{dim}_{\text{N}}(H)$ y otras introducidas por vez primera en la literatura científica en este texto: $\text{outdeg}_{\text{av}}(H)$ y $p \text{dim}_{\text{DS}}(H)$.
- iv) Un operador esencial para alcanzar el estadio de escalera en el caso de clases de conceptos finitas: el shifting $\mathbb{S}_x(H)$, con $x \in [n]^*$ una palabra sobre el alfabeto $[n]$.

Omitiremos la discusión de la dimensión de Natarajan por los resultados de [BCDMY, 22] que la descartan como invariante para caracterizar el aprendizaje multi-clase. Hemos estudiado relaciones entre las distintas nociones y distintas caracterizaciones de los objetos involucrados, pero nos queda UNA CUESTIÓN FUNDAMENTAL: *Si los invariantes cuantitativos pretenden medir parámetros en el aprendizaje multi-clase, ¿cómo se comportan tras realizar operaciones de shifting?*

Dedicaremos esta sección a estudiar alguna de ellas. Comencemos con la más simple, cuya prueba (por lo demás evidente, usando la biyección asociada a \mathbb{S}_i) puede encontrarse como Corollary 23 en [BCDMY, 22].

PROPOSICIÓN 4.3.1 (DIMENSIÓN EXPONENCIAL Y SHIFTING). *Con las notaciones precedentes, sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Para cada $i \in [n]$ y $S \subseteq [n]$, tenemos:*

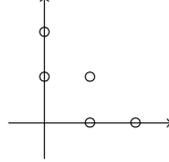
$$\#(\mathbb{S}_i(H)|_S) \leq \#(H|_S).$$

Por tanto, también tenemos $\text{dim}_{\text{E}}(\mathbb{S}_i(H)) \leq \text{dim}_{\text{E}}(H)$.

Si bien la dimensión exponencial es monótona con respecto al shifting (disminuye conforme se van haciendo operaciones del tipo $\mathbb{S}_x(H)$, con $x \in [n]^*$), la dimensión DS puede crecer tras un shift como se prueba en el Ejemplo 19 de [BCDMY, 22] que reproducimos por completar los detalles no explícitos en [BCDMY, 22]. Nótese que esto significa que aunque nos acerquemos a una escalera (i.e. cerrado hacia abajo) usando diversas iteraciones del shift, puede ser que estemos haciendo crecer la dimensión DS en el camino o al final:

EJEMPLO 4.3.2 (EJEMPLO 19 DE [BCDMY, 22] CON SUS DETALLES). La dimensión DS (de [DS, 14]) puede crecer después de realizar un shift. Sea $H \subseteq \{0, 1, 2\}^2$ dado por la siguiente igualdad:

$$H := \{(0, 1), (0, 2), (1, 0), (1, 1), (2, 0)\}.$$



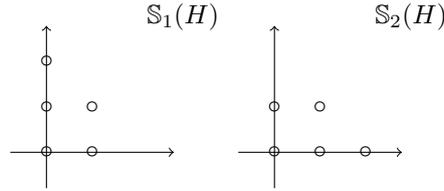
Aplicamos un shift (son transformaciones muy sencillas de realizar) en la dirección de la proyección π_1 y obtenemos:

$$\mathbb{S}_1(H) = \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1)\}.$$

Apliquemos también un shift a H en la dirección de π_2 y obtenemos:

$$\mathbb{S}_2(H) = \{(0, 0), (0, 1), (1, 0), (1, 1), (2, 0)\}.$$

El siguiente gráfico muestra las dos transformaciones:



En ambos casos obtenemos conjuntos cerrados hacia abajo (i.e. escaleras) pero ninguno de ellos es un pseudo-cubo, aunque ambos contienen un pseudo-cubo de dimensión 2. Haciendo los cálculos con los medios y nociones que ya hemos estudiado, observamos que la dimensión exponencial permanece inalterada, mientras que la dimensión DS crece. Es decir,

- $\dim_{\text{DS}}(\mathbb{S}_1(H)) = \dim_{\text{DS}}(\mathbb{S}_2(H)) = 2 > \dim_{\text{DS}}(H) = 1.$
- $\dim_{\text{E}}(\mathbb{S}_1(H)) = \dim_{\text{E}}(\mathbb{S}_2(H)) = \dim_{\text{E}}(H) = 2 > \dim_{\text{DS}}(H) = 1.$

En ambos casos el número de aristas decrece (esto es lo natural con el shifting, como demostraremos más adelante) y, por tanto, disminuye la densidad del grafo y aumenta el grado de salida promedio), pero también decrece el número de aristas unarias. Es decir, obtendremos

$$1 = 2 - 5/5 = \text{outdeg}_{\text{av}}(\mathbb{S}_2(H)) = \text{outdeg}_{\text{av}}(\mathbb{S}_1(H)) > \text{outdeg}_{\text{av}}(H) = 2 - 6/5 = 4/5.$$

$$9/5 = 2 - 1/5 = \text{avd}(\mathbb{S}_2(H)) = \text{avd}(\mathbb{S}_1(H)) > \text{avd}(H) = 2 - 2/5 = 8/5.$$

Como consecuencia de este ejemplo, podemos concluir que dos invariantes pueden tener comportamientos más “erráticos” (impredecibles) tras aplicar una o varias operaciones de shifting: avd y \dim_{DS} . En [PSZ, 24] se hacen algunos estudios más detallados de algunas de estas variaciones. Por ahora nos vamos a concentrar en los otros dos invariantes $\text{outdeg}_{\text{av}}(H)$ y $\dim_{\text{E}}(H)$.

Dados $i, j \in [n]$, con $i \neq j$, y $H \subseteq \mathcal{Y}^n$, denotaremos $\pi_{i,j} : \mathcal{Y}^n \rightarrow \mathcal{Y}^n$ la proyección que olvida las coordenadas i y j :

$$(4.3.1) \quad \begin{aligned} \pi_{i,j} : \mathcal{Y}^n &\longrightarrow \mathcal{Y}^{n \setminus \{i,j\}} \\ h &\longmapsto h|_{[n] \setminus \{i,j\}}, \end{aligned}$$

donde $h|_{[n] \setminus \{i,j\}}$ quiere decir que hemos eliminado las coordenadas i y j del h original. El siguiente Lema tiene una prueba casi inmediata que omitiremos:

LEMA 4.3.3. *Con las notaciones precedentes, dados $i \neq j$ tenemos:*

$$\pi_{i,j}(H) = \pi_{i,j}(\mathbb{S}_i(H)).$$

Introduzcamos dos subconjuntos para $\ell \in \pi_{i,j}(H)$:

$$(4.3.2) \quad A_{\ell,i} := \bigcup_{\substack{f \in \pi_i(H) \\ f|_{[n] \setminus \{i,j\}} = \ell}} e_{f,i}.$$

$$(4.3.3) \quad B_{\ell,i} := \bigcup_{\substack{f \in \pi_i(H) \\ f|_{[n] \setminus \{i,j\}} = \ell}} \mathbb{S}_i(e_{f,i}).$$

Observemos que el lado derecho de las identidades (4.3.2) y (4.3.3) son uniones disjuntas (de ejes en la dirección determinada por π_i). Más aún, tenemos:

LEMA 4.3.4. *Con las notaciones precedentes, las siguientes dos igualdades exhiben particiones respectivamente de H y $\mathbb{S}_i(H)$:*

$$(4.3.4) \quad H := \bigsqcup_{\ell \in \pi_{i,j}(H)} A_{\ell,i}.$$

$$(4.3.5) \quad \mathbb{S}_i(H) := \bigsqcup_{\ell \in \pi_{i,j}(H)} B_{\ell,i}.$$

DEMOSTRACIÓN. Veamos primero que ambas son uniones disjuntas:

- Dados $\ell_1, \ell_2 \in \pi_{i,j}(H)$, si $A_{\ell_1,i} \cap A_{\ell_2,i} \neq \emptyset$, entonces existe $h \in A_{\ell_1,i} \cap A_{\ell_2,i}$. Por tanto, conforme a la Identidad (4.3.2) :

$$\ell_1 = \pi_i(h)|_{[n] \setminus \{i,j\}} = \pi_{i,j}(h) = \ell_2.$$

Y, en conclusión, el término de la derecha de la Identidad (4.3.4) describe una unión disjunta.

- De modo análogo, dados $\ell_1, \ell_2 \in \pi_{i,j}(H) = \pi_{i,j}(\mathbb{S}_i(H))$, si existe $h' \in B_{\ell_1,i} \cap B_{\ell_2,i} \neq \emptyset$, entonces tendríamos $\ell_1 = \pi_{i,j}(h') = \ell_2$, y la Identidad (4.3.5) describe una unión disjunta.

Obviamente, se tiene que $A_{\ell,i} \subseteq H$ y $B_{\ell,i} \subseteq \mathbb{S}_i(H)$. Entonces, las inclusiones en la dirección \supseteq se siguen en ambas identidades. Para la otra inclusión, probemos simplemente la segunda de las identidades, porque la primera se sigue de argumentos similares. Sea $h' \in \mathbb{S}_i(H)$. Consideremos la proyección $f := \pi_i(h')$ y tenemos que $h' \in \mathbb{S}_i(e_{f,i})$. Entonces, tomando $\ell := f|_{[n] \setminus \{i,j\}} = \pi_{i,j}(h)$ tenemos $h' \in \mathbb{S}_i(e_{f,i}) \subseteq B_{\ell,i}$, y la inclusión en el sentido \subseteq se sigue. \square

LEMA 4.3.5. *Con las notaciones precedentes, dados $i, j \in [n]$, con $i \neq j$, tenemos:*

$$(4.3.6) \quad \#(\pi_j(H)) = \sum_{\ell \in \pi_{i,j}(H)} \#(\pi_j(A_{\ell,i})).$$

$$(4.3.7) \quad \#(\pi_j(\mathbb{S}_i(H))) = \sum_{\ell \in \pi_{i,j}(H)} \#(\pi_j(B_{\ell,i})).$$

DEMOSTRACIÓN. Por los resultados y definiciones previos, todo lo que tenemos que hacer es probar que π_j respeta las uniones disjuntas descritas en las identidades (4.3.4) y (4.3.5). Es decir, todo lo que tenemos que probar es que las siguientes son uniones disjuntas :

$$(4.3.8) \quad \pi_j(H) = \bigsqcup_{\ell \in \pi_{i,j}(H)} \pi_j(A_{\ell,i}).$$

$$(4.3.9) \quad \pi_j(\mathbb{S}_i(H)) = \bigsqcup_{\ell \in \pi_{i,j}(H)} \pi_j(B_{\ell,i}).$$

Y esas uniones disjuntas se siguen de las afirmaciones siguientes:

- Si $\pi_j(A_{\ell_1,i}) \cap \pi_j(A_{\ell_2,i}) \neq \emptyset$, entonces $\ell_1 = \ell_2$.
- Si $\pi_j(B_{\ell_1,i}) \cap \pi_j(B_{\ell_2,i}) \neq \emptyset$, entonces $\ell_1 = \ell_2$.

Estas dos afirmaciones se siguen de argumentos similares. Es decir, si existe $g \in \pi_j(A_{\ell_1,i}) \cap \pi_j(A_{\ell_2,i}) \neq \emptyset$ (respectivamente, $g' \in \pi_j(B_{\ell_1,i}) \cap \pi_j(B_{\ell_2,i}) \neq \emptyset$) es porque existe $f \in \pi_i(H)$ (resp. $f' \in \pi_i(\mathbb{S}_i(H)) = \pi_i(H)$) tal que $h \in e_{f,i}$ (resp. $h' \in \mathbb{S}_i(e_{f,i})$) y $\ell_1 = f|_{[n] \setminus \{i,j\}} = \pi_{i,j}(h) = \ell_2$ (resp. $\ell_1 = f'|_{[n] \setminus \{i,j\}} = \pi_{i,j}(h') = \ell_2$).

Las igualdades entre los cardinales se siguen obviamente de disponer de uniones disjuntas y el lema queda demostrado. \square

Introduzcamos dos cantidades relacionadas con la proyección $\pi_{i,j}$. Dados $i, j \in [n]$, $i \neq j$, y $\ell \in \pi_{i,j}(H)$, denotemos mediante $H(\ell)$ la fibra de ℓ en H :

$$H(\ell) := \pi_{i,j}^{-1}(\ell) \cap H := \{h \in H : \pi_{i,j}(h) = \ell\}.$$

Observemos que:

$$\pi_i(H(\ell)) = \{\pi_i(h) : h \in H, h|_{[n] \setminus \{i,j\}} = \ell\}.$$

DEFINICIÓN 24. Con las notaciones precedentes, dados $i, j \in [n]$, $i \neq j$, y $\ell \in \pi_{i,j}(H)$, definimos los máximos de los cardinales de las aristas que se proyectan sobre ℓ como:

$$(4.3.10) \quad \mu_i(\ell) := \max\{\#\{e_{f,i} : f \in \pi_i(H(\ell))\}\}.$$

DEFINICIÓN 25. Con las mismas notaciones, dados $i, j \in [n]$, $i \neq j$, y $\ell \in \pi_{i,j}(H)$, elijamos $f_0 \in \pi_i(H) \cap H(\ell)$ tal que $\#\{e_{f_0,i}\} = \mu_i(\ell)$. Definimos el “segundo” cardinal máximo de las aristas que se proyectan sobre ℓ del modo siguiente:

- Si $\#\{\pi_i(H(\ell))\} \geq 2$, entonces definimos:

$$(4.3.11) \quad \mu_i^{(2)}(\ell) := \max\{\#\{e_{f,i} : f \in \pi_i(H(\ell)) \text{ } g \neq f_0\}\},$$

- en otro caso ($\#\{\pi_i(H(\ell))\} = 1$), definamos $\mu_i^{(2)}(\ell) := 0$.

Los siguientes dos lemas, parcialmente explican las variaciones del número de arista en el OIG tras un shift, usando las cantidades máximas anteriores:

LEMA 4.3.6. Con las mismas notaciones que en los Lemas previos, tenemos:

$$\#\{\pi_j(\mathbb{S}_i(H))\} = \sum_{\ell \in \pi_{i,j}(H)} \mu_i(\ell).$$

Por tanto,

$$\#\{E(\mathbb{S}_i(H))\} = \#\{E_i(H)\} + \sum_{j \neq i} \left(\sum_{\ell \in \pi_{i,j}(H)} \mu_i(\ell) \right).$$

DEMOSTRACIÓN. Como $\pi_{i,j}(H) = \pi_{i,j}(\mathbb{S}_i(H))$, conforme al Lema 4.3.5, tenemos:

$$\#\{\pi_j(\mathbb{S}_i(H))\} = \sum_{\ell \in \pi_{i,j}(H)} \#\{\pi_j(B_{\ell,i})\}.$$

Todo lo que nos queda por probar es la siguiente igualdad:

$$(4.3.12) \quad \#\{\pi_j(B_{\ell,i})\} = \mu_i(\ell).$$

Sin pérdida de la generalidad, supongamos $i = 1$ y $j = 2$. Entonces, observemos que dado $f \in \pi_1(H)$ tal que $f|_{[n] \setminus \{1,2\}} = \ell$ y dado $\mathbb{S}_1(e_{f,1})$ tenemos:

$$\pi_2(\mathbb{S}_1(e_{f,1})) = \{0, \dots, \#\{e_{f,1}\} - 1\} \times \{(\ell(3), \ell(4), \dots, \ell(n))\},$$

Entonces, concluimos:

$$\pi_2(B_{\ell,1}) = \bigcup_{\substack{f \in \pi_1(H) \\ f|_{[n] \setminus \{1,2\}} = \ell}} \{0, \dots, \#\{e_{f,1}\} - 1\} \times \{(\ell(3), \dots, \ell(n))\} = \{0, \dots, \mu_i(\ell) - 1\} \times \{(\ell(3), \dots, \ell(n))\}.$$

Y, por tanto, la Identidad (4.3.12) se sigue. Para la segunda, observemos simplemente que:

$$\#\{E(\mathbb{S}_i(H))\} = \sum_{k=1}^n \#\{E_k(\mathbb{S}_i(H))\}.$$

Como $\#\{E_i(\mathbb{S}_i(H))\} = \#\{E_i(H)\}$, la última igualdad se sigue de nuestro estudio precedente. \square

LEMA 4.3.7. Con las mismas notaciones anteriores, dados $i, j \in [n]$, $i \neq j$, $\ell \in \pi_{i,j}(H)$, se verifica la siguiente igualdad:

$$\#\{\pi_j(A_{\ell,i})\} \geq \mu_i(\ell).$$

Por tanto, también tenemos:

$$(4.3.13) \quad \#\{\pi_j(H)\} \geq \sum_{\ell \in \pi_{i,j}(H)} \mu_i(\ell).$$

DEMOSTRACIÓN. Desde la definición de $A_{\ell,i}$ (Identidad (4.3.2) anterior) tenemos:

$$\pi_j(A_{\ell,i}) = \bigcup_{\substack{f \in \pi_i(H) \\ f|_{[n] \setminus \{i,j\}} = \ell}} \pi_j(e_{f,i}).$$

Entonces, concluimos:

$$\#(\pi_j(A_{\ell,i})) \geq \max \#(\pi_j(e_{f,i})) : f \in \pi_i(H), f|_{[n] \setminus \{i,j\}} = \ell.$$

Todo lo que queda por ver es que π_j es inyectiva cuando se restringe a $e_{f,i}$. Pero esto es inmediato puesto que dados $h \in e_{f,i}$ y $f \in \pi_i(H)$ tales que $f|_{[n] \setminus \{i,j\}} = \ell$ ya sabemos que $\pi_j(h)_k = \ell_k$ para cada $k \neq i, j$. Pero $\pi_j(h)_i = f_i$ y, por tanto, $\pi_j(h)$ está unívocamente determinado por f para cada $h \in e_{f,i}$.

Como π_j restringido a $e_{f,i}$ es inyectiva, concluimos : $\#(\pi_j(e_{f,i})) = \#(e_{f,i})$ y, por tanto,

$$\#(\pi_j(A_{\ell,i})) \geq \max \{ \#(e_{f,i}) : f \in \pi_i(H), f|_{[n] \setminus \{i,j\}} = \ell \} = \mu_i(\ell),$$

como se pretendía. \square

Estos resultados técnicos nos llevan a nuestra primera estimación.

TEOREMA 4.3.8. *Con las notaciones anteriores, dado $H \subseteq \mathcal{Y}^n$ y dado $i \in [n]$, el número de aristas y la densidad del hiper-grafo no crecen tras realizar un shift, i.e.*

$$\begin{aligned} \#(E(\mathbb{S}_i(H))) &\leq \#(E(H)), \\ \text{gd}(\mathbb{S}_i(H)) &\leq \text{gd}(H). \end{aligned}$$

En particular, el grado de salida promedio no decrece tras el shifting, i.e.:

$$\text{outdeg}_{\text{av}}(\mathbb{S}_i(H)) \geq \text{outdeg}_{\text{av}}(H).$$

DEMOSTRACIÓN. Por la Identidad (2.1.3) tenemos:

$$\#(E(\mathbb{S}_i(H))) = \sum_{j=1}^n \#(E_j(\mathbb{S}_i(H))) = \sum_{j=1}^n \#(\pi_j(\mathbb{S}_i(H))).$$

Observemos que $\pi_i(\mathbb{S}_i(H)) = \pi_i(H)$ y tenemos:

$$\#(E(\mathbb{S}_i(H))) = \#(E_i(H)) + \sum_{j \neq i} \#(\pi_j(\mathbb{S}_i(H))).$$

Por el Lema 4.3.6 concluimos:

$$\#(E(\mathbb{S}_i(H))) = \#(E_i(H)) + \sum_{j \neq i} \left(\sum_{\ell \in \pi_{i,j}(H)} \mu_i(\ell) \right).$$

Conforme a la Desigualdad (4.3.13) del Lema 4.3.7, concluimos:

$$\#(E(\mathbb{S}_i(H))) \leq \#(E_i(H)) + \sum_{j \neq i} \#(\pi_j(H)).$$

De nuevo, por la Identidad (2.1.3), concluimos:

$$\#(E(\mathbb{S}_i(H))) \leq \#(E_i(H)) + \sum_{j \neq i} \#(E_j(H)) = \#(E(H)).$$

Esto demuestra la primera de las desigualdades de la proposición. Como $\#(\mathbb{S}_i(H)) = \#(H)$, la segunda, relativa a la densidad del grafo, se sigue de la primera. En cuanto a la desigualdad final, se sigue de la Afirmación *ii*) de Proposición 2.2.1. \square

OBSERVACIÓN 4.3.9. De acuerdo con nuestro análisis, la dimensión exponencial y el grado de salida en promedio evolucionan en direcciones “opuestas” conforme se aplican secuencias de shifts. Más aún, hemos observado que el número de aristas del resultado de aplicar un shift verifica la siguiente igualdad:

$$\#(E(\mathbb{S}_i(H))) = \#(E_i(H)) + \sum_{j \neq i} \sum_{\ell \in \pi_{i,j}(H)} \mu_j(\ell).$$

Pero no podemos “permutar” los sumatorios con índices $j \neq i$ y $\ell \in \pi_{i,j}(H)$, puesto que $\pi_{i,j}(H)$ varía con j .

En el caso binario (i.e. $p = 2$) también $\text{avd}(H)$ tiene un buen comportamiento tras realizar un shift:

COROLARIO 4.3.10. En el caso $p = 2$, dado $i \in [n]$ y $H \subseteq \{0, 1\}^n$, tenemos:

$$\text{avd}(\mathbb{S}_i(H)) \geq \text{avd}(H).$$

DEMOSTRACIÓN. Conforme a la Identidad (2.2.6) de la Observación 2.2.6, dado un shift \mathbb{S}_i , tenemos:

$$\text{avd}(\mathbb{S}_i(H)) - \text{outdeg}_{\text{av}}(\mathbb{S}_i(H)) = \frac{\#(\text{Haussl}(\mathbb{S}_i(H)))}{2\#(H)},$$

Por la Afirmación (2) de la demostración del Lema 2 de [Hau, 1995], sabemos que se tiene:

$$\#(\text{Haussl}(\mathbb{S}_i(H))) \geq \#(\text{Haussl}(H)).$$

Por tanto, tenemos:

$$\text{avd}(\mathbb{S}_i(H)) - \text{outdeg}_{\text{av}}(\mathbb{S}_i(H)) \geq \frac{\#(\text{Haussl}(H))}{2\#(H)} = \text{avd}(H) - \text{outdeg}_{\text{av}}(H).$$

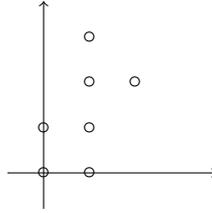
Por el Teorema 4.3.8, concluimos:

$$\text{avd}(\mathbb{S}_i(H)) - \text{avd}(H) \geq \text{outdeg}_{\text{av}}(\mathbb{S}_i(H)) - \text{outdeg}_{\text{av}}(H) \geq 0.$$

Y concluimos el resultado anunciado. \square

Sin embargo, en el caso $p \geq 3$, el grado promedio del hiper-grafo $G(H)$ tiene un comportamiento errático, como muestra el siguiente ejemplo. En [PSZ, 24] se trata de controlar ese comportamiento errático, pero los resultados son muy primitivos aún.

EJEMPLO 4.3.11 (EL GRADO PROMEDIO PUEDE CRECER O DECRECER TRAS UN SHIFT CUANDO $p \geq 3$). Consideremos $\mathcal{Y} := \{0, 1, 2, 3\}$, $n = 2$, y definamos $H := \{(0, 0), (0, 1), (1, 0), (1, 1), (1, 2), (1, 3), (2, 2)\}$. Una imagen de este conjunto es:



Tenemos que $\#(H) = 7$ y el hiper-grafo asociado $G(H)$ tiene 7 aristas y, por tanto,

$$\text{outdeg}_{\text{av}}(H) = 2 - 7/7 = 1,$$

mientras que contiene un cubo de dimensión 2 y, por tanto, $\dim_{\text{DS}}(H) = \dim_{\mathbb{E}}(H) = 2$.

Adicionalmente, posee 2 aristas unarias (i.e. $\#(E^{(1)}(H)) = 2$). Por tanto, $\text{avd}(H) = 2 - 2/7 = 12/7$.

Hagamos un primer shifting en la dirección de la proyección π_2 (i.e. la proyección que “olvida” la segunda coordenada):

$$H' := \mathbb{S}_2(H) = \{(0, 0), (1, 0), (2, 0), (0, 1), (1, 1), (1, 2), (1, 3)\}.$$

Las dimensiones exponencial y DS se preservan. También se preserva el número total de aristas y, por tanto, se preservan la densidad del grafo y el grado de salida promedio (i.e. $\text{outdeg}_{\text{av}}(\mathbb{S}_2(H)) = \text{outdeg}_{\text{av}}(H) = 1$). Sin embargo, el número de aristas unarias decrece, lo que, por nuestros resultados previos hace aumentar el grado promedio, i.e.

$$3 = \#(E^{(1)}(\mathbb{S}_2(H))) > \#(E^{(1)}(H)) = 2, \quad 11/7 = \text{avd}(\mathbb{S}_2(H)) < \text{avd}(H) = 12/7.$$

Por otro lado, si hacemos un shift en la dirección de la proyección π_1 obtenemos (sencillo cálculo a mano):

$$\mathbb{S}_1(H) := \{(0, 0), (1, 0), (0, 1), (1, 1), (0, 2), (1, 2), (0, 3)\}.$$

El conjunto $\mathbb{S}_1(H)$ es cerrado hacia abajo, pero no es un pseudo-cubo. Las dimensiones exponencial y DS se preservan. Pero el número de aristas decrece estrictamente y, por tanto, el grado de salida promedio crece estrictamente. Es decir,

$$\#(E(\mathbb{S}_1(H))) = 6 < \#(E(H)), \quad 8/7 = \text{outdeg}_{\text{av}}(\mathbb{S}_1(H)) > \text{outdeg}_{\text{av}}(H) = 1.$$

Esta vez, el número de aristas unarias también decrece (i.e. $\#(E^{(1)}(\mathbb{S}_1(H))) = 1 < \#(E^{(1)}(H)) = 2$) y, por tanto, el grado promedio crece estrictamente:

$$2 - 1/7 = 13/7 = \text{avd}(\mathbb{S}_1(H)) > \text{avd}(H) = 12/7.$$

Estamos ya en condiciones de probar nuestro primer resultado importante en esta sección:

TEOREMA 4.3.12 (EL GRADO DE SALIDA EN PROMEDIO ES ESTRICTAMENTE MENOR QUE LA DIMENSIÓN EXPONENCIAL). *Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Entonces, tenemos:*

$$\text{outdeg}_{\text{av}}(H) < \dim_{\mathbb{E}}(H).$$

DEMOSTRACIÓN. Como en la Observación 4.1.4, seguimos la idea de usar secuencias de shifts. Para cada $H \subseteq \mathcal{Y}^n$ existe una palabra $x := x_1 \cdots x_r \in [n]^*$ que nos permite concluir que \mathbb{S}_x satisface que la clase $H_* := \mathbb{S}_x(H)$ es cerrada hacia abajo con respecto al pre-orden de Dickson. Por la Proposición 4.3.8 concluimos que:

$$\text{outdeg}_{\text{av}}(H) \leq \text{outdeg}_{\text{av}}(H_*).$$

Por la Proposición 4.3.1, también concluimos que:

$$\dim_{\mathbb{E}}(H_*) \leq \dim_{\mathbb{E}}(H).$$

Finalmente, por la Proposición 3.3.6, como H_* es cerrado hacia abajo, tenemos también que:

$$\text{outdeg}_{\text{av}}(H_*) < \text{avd}(H_*) \leq \dim_{\mathbb{E}}(H_*).$$

Y esto concluye la prueba de nuestro resultado. \square

Este teorema implica las siguientes consecuencias, que suponen una mejora estricta de la Proposición 27 de [BCDMY, 22]. La mejora estricta consiste en transformar $\text{avd}(H) \leq 4 \dim_{\mathbb{E}}(H)$ en la desigualdad que aparece en la Afirmación *ii*) ($\text{avd}(H) < 2 \dim_{\mathbb{E}}(H)$) del corolario siguiente. El uso de $\text{outdeg}_{\text{av}}(H)$ es esencial en esta pequeña mejora:

COROLARIO 4.3.13. *Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Entonces, tenemos:*

i) Las dimensiones satisfacen:

$$(4.3.14) \quad \dim_{\text{DS}}(H) < 2 \dim_{\mathbb{E}}(H).$$

ii) El grado promedio satisface:

$$(4.3.15) \quad \text{avd}(H) < 2 \dim_{\mathbb{E}}(H).$$

DEMOSTRACIÓN. Sea $S \subseteq [n]$ maximal con la propiedad de que $H|_S$ contiene un pseudo-cubo de dimensión $r = \sharp(S) = \dim_{\text{DS}}(H)$. Sea $\mathcal{C} \subseteq H|_S$ tal pseudo-cubo. Por el Teorema 3.1.4 tenemos:

$$\dim_{\text{DS}}(H) = r = \sharp(S) = \text{avd}(\mathcal{C}).$$

Por la Afirmación *iii*) del Teorema 2.2.1 también tenemos:

$$\text{avd}(\mathcal{C}) \leq 2 \text{outdeg}_{\text{av}}(\mathcal{C}).$$

Por el Teorema 4.3.12 previo concluimos:

$$\dim_{\text{DS}}(H) = r < 2 \dim_{\mathbb{E}}(\mathcal{C}).$$

Por la Observación 3.2.2, se concluye inmediatamente que se verifica:

$$\dim_{\text{DS}}(H) = r < 2 \dim_{\mathbb{E}}(\mathcal{C}) \leq 2 \dim_{\mathbb{E}}(H|_S) \leq 2 \dim_{\mathbb{E}}(H),$$

y hemos probado la Desigualdad (4.3.14).

Para la segunda desigualdad, por el Teorema 2.2.1 tenemos:

$$\text{avd}(H) \leq 2 \text{outdeg}_{\text{av}}(H).$$

Y el Teorema 4.3.12 anterior implica:

$$\text{avd}(H) < 2 \dim_{\mathbb{E}}(H),$$

Lo que demuestra la Desigualdad (4.3.15). \square

Esto nos permite demostrar una mejora estricta sobre la existencia de orientaciones con grado de salida maximal controlado que se expresa en el Corolario 28 de [BCDMY, 22]. En ese corolario se prueba la existencia una orientación $\sigma : E(H) \rightarrow H$ que satisface:

$$\text{outdeg}_{\text{max}}(\sigma) \leq 4 \dim_{\mathbb{E}}(H).$$

Lo hacemos mediante nuestro Teorema 4.3.14, especialmente en la Afirmación *i*) y la Desigualdad (4.3.22). Introduzcamos la noción de grado mínimo del hiper-grafo $G(H)$ definido por una clase de conceptos finita $H \subseteq \mathcal{Y}^n$ como:

$$\text{deg}_{\text{min}}(H) := \min\{\text{deg}(h) : h \in H\}.$$

En [DS, 14], los autores consideran también el grado promedio “maximal”, cuya definición reproducimos aquí:

$$(4.3.16) \quad \max - \text{avd}(H) := \max\{\text{avd}(F) : F \subseteq H\},$$

De modo similar a [DS, 14], podemos introducir el grado de salida promedio “maximal” mediante:

$$(4.3.17) \quad \max -\text{outdeg}_{\text{av}}(H) := \max\{\text{outdeg}_{\text{av}}(F) : F \subseteq H\}.$$

Como se observa de nuestros resultados, se tiene:

$$(4.3.18) \quad \max -\text{avd}(H) \leq 2 \max -\text{outdeg}_{\text{av}}(H).$$

TEOREMA 4.3.14 (MEJORA ESTRICTA DEL COROLARIO 28 DE [BCDMY, 22] SOBRE EL GRADO DE SALIDA MÁXIMO DE ORIENTACIONES). *Con las mismas notaciones precedentes, sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Sea $\varphi : 2^H \rightarrow \mathbb{R}$ una función monótona con respecto a la inclusión en 2^H . Supongamos que la función φ también satisface la siguiente hipótesis:*

$$(4.3.19) \quad \forall F \in 2^H, \text{deg}_{\text{min}}(F) \leq \varphi(F).$$

Entonces, existe una orientación $\sigma : E(H) \rightarrow H$ tal que se verifica:

$$(4.3.20) \quad \text{outdeg}_{\text{max}}(\sigma) \leq \varphi(H).$$

Adicionalmente, tenemos:

i) La Hipótesis (4.3.19) se verifica siempre que la siguiente condición sea cierta:

$$(4.3.21) \quad \text{avd}(F) \leq \varphi(F), \forall F \subseteq H.$$

En particular, la función monótona $\varphi(F) := 2 \dim_{\mathbb{E}}(F)$ satisface la Desigualdad (4.3.21) y, por tanto, existe una orientación $\sigma : E(H) \rightarrow H$ tal que:

$$(4.3.22) \quad \text{outdeg}_{\text{max}}(\sigma) \leq 2 \dim_{\mathbb{E}}(H).$$

ii) El grado promedio maximal también satisface la Hipótesis (4.3.21) y, por tanto, existe una orientación $\sigma : G(H) \rightarrow H$ tal que:

$$\text{outdeg}_{\text{max}}(\sigma) \leq \max -\text{avd}(H) \leq 2 \max -\text{outdeg}_{\text{av}}(H).$$

iii) El $\text{outdeg}_{\text{av}}(H)$ controla la complejidad del algoritmo OIG dado que:

$$(4.3.23) \quad \text{outdeg}_{\text{av}}(H) \leq \min -\text{outdeg}_{\text{max}}(H) \leq 2 \max -\text{outdeg}_{\text{av}}(H),$$

DEMOSTRACIÓN. Probaremos el resultado general por inducción en $\sharp(H)$ (formalizando con mayor cuidado y precisión algunos de los argumentos de [BCDMY, 22]). El caso $\sharp(H) = 1$ es obvio.

Por la Identidad (4.3.19), sea $h \in H$ tal que $\text{deg}(h) \leq \varphi(H)$. Consideremos la clase $H' := H \setminus \{h\}$ y su hiper-grafo de una inclusión asociado $G(H')$. Clasifiquemos las aristas de $E(H)$ en las siguiente tres clases disjuntas:

- La clase de aristas que no contienen a h :

$$E_h^{(0)}(H) := \{e \in E(H) : h \notin e\}.$$

- La clase de aristas que no contiene a h pero no son unarias:

$$E_h^{(2)}(H) := \{e \in E(H) : h \in e, \sharp(e) \geq 2\}.$$

- Las aristas de $E(H)$ cuyo único elemento es h :

$$E_h^{(1)}(H) := \{e \in E^{(1)}(H) : h \in e\}.$$

Tenemos una partición de $E(H)$:

$$E(H) = E_h^{(0)}(H) \sqcup E_h^{(1)}(H) \sqcup E_h^{(2)}(H).$$

Aplicamos la hipótesis inductiva y supongamos que existe una orientación $\sigma' : E(H') \rightarrow H'$ tal que:

$$\text{outdeg}_{\text{max}}(\sigma') \leq \varphi(H').$$

Entonces, extendemos σ' a una orientación $\sigma : E(H) \rightarrow H$ de la forma siguiente:

- Si $e \in E_h^{(0)}(H) \sqcup E_h^{(2)}(H)$, entonces definimos:

$$\sigma(e) := \sigma'(e \setminus \{h\}) \in H' \subseteq H.$$

- En otro caso, si $e \in E_h^{(1)}(H)$, definamos $\sigma(e) := h$.

Observemos que para cada $g \in H'$ (i.e. $g \neq h$), tenemos $\text{outdeg}(\sigma, g) = \text{outdeg}(\sigma', g)$. Por otro lado, observemos que se tiene:

$$\#\{e \in E(H) : h \in e, \sigma(e) \neq h\} = \#\left(E_h^{(2)}(H)\right) = \text{deg}(h).$$

Por tanto, $\text{outdeg}(\sigma, h) = \text{deg}(h)$ y, por tanto, tenemos $\text{outdeg}_{\max}(\sigma) := \max\{\text{outdeg}_{\max}(\sigma'), \text{deg}(h)\}$. Observemos también que se tiene:

- Por la hipótesis inductiva y el hecho de que φ es monótona con respecto a la inclusión, tenemos:

$$\text{outdeg}_{\max}(\sigma') \leq \varphi(H') \leq \varphi(H).$$

- Por nuestra Hipótesis (4.3.19), h ha sido elegido de tal modo que se satisfaga:

$$\text{deg}(h) \leq \varphi(H).$$

Por tanto, concluimos que $\text{outdeg}_{\max}(\sigma) \leq \varphi(H)$.

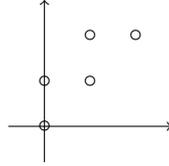
Para la Afirmación *i*), observemos que si $\text{avd}(F) \leq \varphi(F)$ se verifica para todo $F \subseteq H$, entonces la Hipótesis (4.3.19) también se verifica. Finalmente, como la función $\varphi(H) := 2 \dim_{\mathbb{E}}(H)$ es monótona con respecto a la inclusión, por el Corolario 4.3.13, esta función satisface la Hipótesis (4.3.19), lo que implica la Afirmación *i*).

La Afirmación *ii*) se verifica porque $\max - \text{avd}$ es monótona con respecto a la inclusión y, obviamente, satisface la Hipótesis (4.3.21). La segunda desigualdad de la Afirmación *ii*) es inmediata por la Desigualdad (4.3.18). \square

OBSERVACIÓN 4.3.15 (CONSIDERACIONES FINALES). Observemos que la Afirmación *ii*) es el Lema 3 de [DS, 14], añadiendo el doble del grado promedio maximal, mientras que la Afirmación *i*) mejora estrictamente el Corolario 28 de [BCDMY, 22] por un factor estrictamente mayor que 2. Sin embargo, no hemos sido capaces de probar resultados similares para la función (naturalmente monótona) \dim_{DS} . El Ejemplo 4.3.16 siguiente prueba que la Hipótesis (4.3.21) no es satisfecha por la dimensión DS. Más aún, este ejemplo prueba también que $\max - \text{avd}(H)$ puede ser mayor que $\dim_{\text{DS}}(H)$. Por tanto, el Lema 3 de [DS, 14] no se aplica para producir orientaciones cuyo grado de salida máximo esté acotado por $\dim_{\text{DS}}(H)$. En consecuencia, parece que $2 \max - \text{outdeg}_{\text{av}}(H)$ es la cantidad más fina que podemos encontrar tal que sea estable por shifting y que permita controlar el grado de salida máximo de alguna orientación en un OIG sobre H . Pero sigue siendo un objeto de estudio en [PSZ, 24].

EJEMPLO 4.3.16 (EL GRADO PROMEDIO NO SIEMPRE ES MENOR QUE LA DIMENSION DS Y OTROS). Consideremos $n = 2$, $\mathcal{Y} := \{0, 1, 2\}$ y $H \subseteq \mathcal{Y}^{[2]}$ la clase dada por la siguiente igualdad:

$$H := \{(0, 0), (0, 1), (1, 1), (1, 2), (2, 2)\}.$$



Unos pocos cálculos permiten observar que $\#\{E(H)\} = 6$ y $\#\{E^{(1)}(H)\} = 2$. Por tanto, $\text{avd}(H) = 8/5 > 1$, $\text{outdeg}_{\text{av}}(H) = 4/5 < 1$. Puesto que $\#\{H\} = 5 \geq 2^2$, obtenemos inmediatamente que $\dim_{\mathbb{E}}(H) = 2$. Sin embargo, H no es un pseudo-cubo. Una discusión de diversos casos, prueba que si eliminamos un sólo punto de H , los cinco puntos restantes no son un pseudo-cubo. Si eliminamos 2 puntos en H , también obtenemos un H' que no es un pseudo-cubo. Esto nos llevará a concluir que H sólo contiene pseudo-cubos de dimensión 1. Por tanto, $\dim_{\text{DS}}(H) = 1$ y este ejemplo verifica que se tiene:

$$\text{outdeg}_{\text{av}}(H) = 4/5 < \dim_{\text{DS}}(H) = 1 < \text{avd}(H) = 8/5 < \dim_{\mathbb{E}}(H) = 2.$$

Por tanto, la Hipótesis (4.3.21) no se sostiene para $\dim_{\text{DS}}(H)$. Ni el Lema 3 de [DS, 14] ayuda con este ejemplo, porque $\dim_{\text{DS}}(H) < \max - \text{avd}(H)$. Sin embargo, la Hipótesis (4.3.19) sí es cierta en este ejemplo y la tesis del Lema 13 de [BCDMY, 22] se salva. Nos preguntamos si el Lema 13 de [BCDMY, 22] seguirá siendo cierto en un contexto más general.

Pre-orden de Dickson, órdenes monomiales y Bases de Gröbner, con algo más de detalle en las pruebas

Índice

A.1. Pre-orden de Dickson y clases de conceptos cerradas hacia abajo: escalones y escaleras	51
A.2. Órdenes monomiales y el lenguaje de las Bases de Gröbner	53
A.2.1. Unas palabras sobre la existencia de algoritmos para el cálculo de las Bases de Gröner	58
A.2.2. El caso de los anillos de Artin y las escaleras	59

A.1. Pre-orden de Dickson y clases de conceptos cerradas hacia abajo: escalones y escaleras

Comenzaremos con un sencillo resultado combinatorio sobre el monoide conmutativo $(\mathbb{N}^n, +)$ conocido como LEMA DE DICKSON (cf. [Dix, 1913]), aunque parece que también era conocido por P. Gourdan.

DEFINICIÓN 26 (PRE-ORDEN DE DICKSON). *Consideremos el monoide $(\mathbb{N}^n, +)$ y consideremos la relación (de orden parcial) siguiente:*

Dados $\alpha = (\alpha_1, \dots, \alpha_n), \beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$, diremos que α es menor que β con respecto al pre-orden de Dickson si se satisface $\exists \gamma \in \mathbb{N}^n, \alpha + \gamma = \beta$.

Escribiremos $\alpha \preceq \beta$ para indicar esa relación. Se denomina pre-orden de Dickson

Algunas sencillas propiedades que satisface este pre-orden se resumen en el siguiente enunciado, de fácil demostración:

PROPOSICIÓN A.1.1. *Con las notaciones anteriores, se verifica:*

- i) En el caso $n = 1$ el orden \preceq es el orden usual en \mathbb{N} .*
- ii) La relación \preceq es una relación de orden parcial sobre \mathbb{N}^n .*
- iii) Para cada $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, denotemos por:*

$$\underline{X}^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n} \in K[X_1, \dots, X_n],$$

al monomio cuyos exponentes vienen dados por las coordenadas de α , siendo K un cuerpo cualquiera. Entonces, para cada $\alpha, \beta \in \mathbb{N}^n$ se tiene que

$$\alpha \preceq \beta \iff \underline{X}^\alpha \mid \underline{X}^\beta,$$

donde \mid significa “divide” en $K[X_1, \dots, X_n]$.

El resultado fundamental de las ideas de Dickson se resume en el siguiente resultado (conocido como *Lema de Dickson*) cuya prueba puede seguirse fácilmente de los pasos descritos en [Par, 24a].

TEOREMA A.1.2 (LEMA DE DICKSON). *Con las notaciones precedentes, todo subconjunto no vacío $S \subseteq \mathbb{N}^n$ posee un número finito de elementos minimales para el orden \preceq .*

DEFINICIÓN 27 (CONJUNTOS CERRADOS HACIA ABAJO). *Con las notaciones precedentes, diremos que una clase de conceptos $H \subseteq \mathcal{Y}^n$ es cerrada hacia abajo (con respecto al pre-orden de Dickson \preceq) si satisface la siguiente propiedad:*

$$\forall h \in H, \forall g \in \mathcal{Y}^n, g \preceq h \implies g \in H.$$

EJEMPLO A.1.3 (ESCALÓN). Un sencillo ejemplo de clase de conceptos cerrada hacia abajo es el ejemplo del *escalón* determinado por un exponente monomial. Así, dado $h \in \mathbb{N}^n$ definimos el escalón determinado por h como el conjunto siguiente:

$$(A.1.1) \quad \mathcal{E}_h := \{g \in \mathbb{N}^n : g \preceq h\}.$$

Obviamente, los escalones son conjuntos cerrados hacia abajo y para cada $h \in H$, si H es cerrado hacia abajo, se tiene $\mathcal{E}_h \subseteq H$.

La siguiente proposición explica el término de escalón y, en el caso finito, la “escalera” que determinan.

PROPOSICIÓN A.1.4 (ESCALERA (FINITA)). *Una clase de conceptos finita $H \subseteq \mathcal{Y}^n$ es cerrada hacia abajo si y solamente si H admite una descomposición del tipo siguiente:*

Existen $h_1, \dots, h_s \in \mathbb{N}^n$ tales que las siguientes dos propiedades se verifican:

- *Cada elemento $h_i \in H$ no está en los escalones determinados por los demás elementos, i.e.:*

$$h_i \notin \bigcup_{j \neq i} \mathcal{E}_{h_j}.$$

- *La clase H es una unión finita de tales escalones:*

$$(A.1.2) \quad H := \bigcup_{i=1}^s \mathcal{E}_{h_i}.$$

Diremos que H es una escalera cuyos escalones son determinados por $\mathcal{E}_{h_1}, \dots, \mathcal{E}_{h_s}$.

Más aún, si $H \subseteq \mathbb{N}^n$ es una escalera, los escalones $\mathcal{E}_{h_1}, \dots, \mathcal{E}_{h_s}$ están determinados de manera única por H y los elementos h_1, \dots, h_s son los elementos maximales de H con respecto al pre-orden de Dickson \preceq .

Los escalones (únicamente determinados) $\mathcal{E}_{h_1}, \dots, \mathcal{E}_{h_s}$ serán denominados las componentes (de la escalera) del conjunto finito cerrado hacia abajo H .

DEMOSTRACIÓN. Ciertamente, si H es finito sólo puede tener un número finito de elementos maximales con respecto a \preceq . Como \mathcal{E}_h es siempre cerrado hacia abajo, tenemos que si $h \in H$, el escalón que determina satisface $\mathcal{E}_h \subseteq H$. Esto prueba que si H es cerrado hacia abajo, entonces es la unión finita de los escalones de sus elementos maximales con respecto a \preceq , lo que demuestra la Identidad (A.1.2). Más aún, elementos maximales de H con respecto a \preceq no serán comparables entre sí. Por tanto, para cada $i \in [s]$, $h_i \notin \bigcup_{j \neq i} \mathcal{E}_{h_j}$, lo que prueba la primera propiedad.

Obviamente cada unión finita de escalones es un conjunto cerrado hacia abajo, lo que prueba el recíproco. \square

OBSERVACIÓN A.1.5. Dado un subconjunto $H \subseteq \mathbb{N}^n$ como una unión finita de escalones (y, por tanto, es también un conjunto finito):

$$H := \bigcup_{k=1}^t \mathcal{E}_{h_k},$$

concluimos que H es cerrado hacia abajo. Pero la descomposición puede refinarse hasta obtener una descomposición que satisfaga las dos propiedades de la proposición precedente. Basta con elegir los elementos maximales $S \subseteq H$ para \preceq de entre la clase $\{h_1, \dots, h_t\}$.

Por otro lado, podemos considerar el “anti-escalón” asociado a un elemento de \mathbb{N}^n . Es decir, podemos introducir la notación siguiente:

Dado $h \in \mathbb{N}^n$, definamos el siguiente subconjunto de \mathbb{N}^n :

$$h + \mathbb{N}^n := \{g \in \mathbb{N}^n : \exists m \in \mathbb{N}^n, h + m = g\} = \{g \in \mathbb{N}^n : h \preceq g\},$$

donde $h + m$ denota la suma estándar en el monoide $(\mathbb{N}^n, +)$. Se tiene la siguiente caracterización alternativa de los conjuntos cerrados hacia abajo:

PROPOSICIÓN A.1.6. *Con las notaciones precedentes, para cada $H \subseteq \mathcal{Y}^n$, no necesariamente finito, H es cerrado hacia abajo si y solamente si existen $g_1, \dots, g_t \in \mathbb{N}^n$, no comparables entre ellos con respecto a \preceq tales que:*

$$H := \bigcap_{i=1}^t (\mathbb{N}^n - (g_i + \mathbb{N}^n)).$$

Más aún, los elementos g_1, \dots, g_t están determinados de manera única por H por ser los elementos minimales para el pre-orden de Dickson \preceq del conjunto $\mathbb{N}^n \setminus H$.

DEMOSTRACIÓN. Se trata simplemente de considerar el conjunto $\mathbb{N}^n \setminus H$ y usar el Lema de Dickson (Teorema A.1.2 anterior) para concluir que hay solamente un número finito de elementos minimales de $\mathbb{N}^n \setminus H$: $\{g_1, \dots, g_t\}$. Esos son los elementos a los que hace referencia el enunciado y, obviamente, las propiedades se verifican de manera elemental. \square

Obsérvese que la condición de ser cerrada hacia abajo es hereditaria mediante restricciones.

LEMA A.1.7. *Sea $H \subseteq \mathcal{Y}^n$ una clase de conceptos finita. Entonces, H es cerrada hacia abajo si y solamente si lo son todas sus restricciones $H|_S$ para cualquier $S \subseteq [n]$.*

DEMOSTRACIÓN. Supongamos que H es cerrada hacia abajo y sea $S \subseteq [n]$ un subconjunto. Dada $g \in H$ una función en la clase de conceptos y consideramos $g|_S \in H|_S$. Entonces, para cualquier $f : S \rightarrow \mathcal{Y}$ tal que $f \preceq g|_S$, definamos un “lifting” $\tilde{f} : [n] \rightarrow \mathcal{Y}$ de f a $[n]$, mediante:

$$\tilde{f}(i) := \begin{cases} f(i), & \text{if } i \in S, \\ g(i), & \text{if } i \notin S. \end{cases}$$

Obviamente, $\tilde{f}|_S = f$ y $\tilde{f} \preceq g$. Entonces, como H es cerrada hacia abajo, $\tilde{f} \in H$ y, por tanto, $f = \tilde{f}|_S \in H|_S$, lo que prueba que $H|_S$ es cerrada hacia abajo.

El recíproco es obvio simplemente tomando $S = [n]$. □

A.2. Órdenes monomiales y el lenguaje de las Bases de Gröbner

Uno de los elementos esenciales en el trabajo con polinomios univariados (especialmente en la División euclídea y en el desarrollo de la Teoría de la Eliminación basada en resultantes), es la existencia de un orden natural en el monoide $(\mathbb{N}, +)$. Para el trabajo con polinomios multivariados la extensión natural de esta relación de orden son los *órdenes monomiales* sobre el monoide $(\mathbb{N}^n, +)$. Queremos introducir un modelo de división que se “parezca” a la División euclídea y que se denominará *División de Weierstarss-Hironaka*.

OBSERVACIÓN A.2.1 (ORDENAR EL MONOIDE $(\mathbb{N}^n, +)$ MEDIANTE BIYECCIONES CON $(\mathbb{N}, +)$). Podemos identificar \mathbb{N}^n con \mathbb{N} de diversas formas. Cada biyección $\varphi : \mathbb{N}^n \rightarrow \mathbb{N}$ induce una relación de orden sobre \mathbb{N}^n del modo siguiente:

Dados $\mu, \theta \in \mathbb{N}^n$, $\mu \leq_\varphi \theta$ si y solamente si $\varphi(\mu) \leq \varphi(\theta)$. Nos interesa no solamente una biyección sino que, además, se comporte bien sobre la estructura de monoide. En este sentido, en lugar de tratar con biyecciones, interesan, sobre todo, los *órdenes monomiales*.

DEFINICIÓN 28 (ORDEN MONOMIAL). *Un orden monomial sobre $(\mathbb{N}^n, +)$ es una relación de orden \leq sobre \mathbb{N}^n que verifica las siguientes propiedades:*

- i) *Es un buen orden (i.e. todo subconjunto no vacío de \mathbb{N}^n posee mínimo para \leq).*
- ii) *$\forall \mu, \theta, \tau \in \mathbb{N}^n$, si $\mu \leq \tau$, entonces $\mu + \theta \leq \tau + \theta$.*
- iii) *El elemento $0 = (0, \dots, 0) \in \mathbb{N}^n$ es el mínimo de (\mathbb{N}^n, \leq) .*

EJEMPLO A.2.2 (ORDEN LEXICOGRÁFICO CANÓNICO). El orden lexicográfico \leq_{lex} es el orden dado por la regla siguiente:

Dados $\mu := (\mu_1, \dots, \mu_n) \in \mathbb{N}^n$ y $\nu := (\nu_1, \dots, \nu_n) \in \mathbb{N}^n$, diremos que $\mu \leq_{\text{lex}} \nu$ si $\mu = \nu$ ó $\mu <_{\text{lex}} \nu$, es decir, existe $k \in \{0, \dots, n-1\}$ tal que se verifican las dos propiedades siguientes:

- $\mu_i = \nu_i$, para $1 \leq i \leq k$. Obsérvese que si $k = 0$, ésta es la condición vacía.
- $\mu_{k+1} < \nu_{k+1}$.

Es claramente un orden monomial en \mathbb{N}^n , pero no permite biyectar \mathbb{N}^n con \mathbb{N} para $n \geq 2$. Si el orden \leq_{lex} viniera de una biyección de \mathbb{N}^n con \mathbb{N} para $n \geq 2$, dado $\mu \in \mathbb{N}^n$, el conjunto de los elementos $\{\theta \in \mathbb{N}^n : \theta \leq_{\text{lex}} \mu\}$ sería un conjunto finito porque así sucede en \mathbb{N} . Pero eso no es cierto para $n \geq 2$. Si elegimos $\mu = (1, 0, \dots, 0) \in \mathbb{N}^n$ observamos que este conjunto es infinito:

$$\{\theta \in \mathbb{N}^n : \theta \leq_{\text{lex}} (1, 0, \dots, 0)\} = (\{0\} \times \mathbb{N}^{n-1}) \cup \{(1, 0, \dots, 0)\}.$$

EJEMPLO A.2.3 (ORDEN LEXICOGRÁFICO SALVO PERMUTACIÓN). Consideremos el grupo simétrico (\mathcal{S}_n, \circ) de orden $n!$ y sea $\sigma \in \mathcal{S}_n$ una permutación. Se define el orden lexicográfico en el monoide $(\mathbb{N}^n, +)$ salvo la permutación σ , que denotamos mediante \leq_{lex}^σ , como el orden dado por el orden lexicográfico tras permutar las coordenadas. Es decir, dado $\mu := (\mu_1, \dots, \mu_n) \in \mathbb{N}^n$, definamos la permutación de sus coordenadas como:

$$\sigma(\mu) := (\mu_{\sigma(1)}, \dots, \mu_{\sigma(n)}) \in \mathbb{N}^n.$$

Entonces, dados $\mu, \theta \in \mathbb{N}^n$ y $\sigma \in \mathcal{S}_n$, definimos:

$$\mu \leq_{\text{lex}}^\sigma \theta := \sigma(\mu) \leq_{\text{lex}} \sigma(\theta).$$

Es un orden monomial en el que cambiamos el orden de eliminación de las variables que resultan ordenadas mediante:

$$X_{\sigma(n)} > \dots > X_{\sigma(1)}.$$

EJEMPLO A.2.4 (ORDEN GRADO+LEXICOGRÁFICO). Otro ejemplo de orden monomial es el \leq_{deglex} , también denominado “grado + lexicográfico” (\leq_{deglex}) y en ocasiones denotado mediante grlex , que se define del modo

siguiente:

Dado un multi-índice $\mu := (\mu_1, \dots, \mu_n) \in \mathbb{N}^n$ se define su *grado total*, mediante la identidad siguiente:

$$|\mu| := \mu_1 + \dots + \mu_n.$$

Así, sea \leq_{lex} el orden lexicográfico en \mathbb{N}^n . Ya hemos indicado que \leq_{lex} no es un orden que permita biyectar \mathbb{N}^n con \mathbb{N} . Podemos definir, entre otras, alternativas como la siguiente: dados $\mu, \theta \in \mathbb{N}^n$, diremos que $\mu \leq_{deglex} \theta$ si se verifica:

$$[|\mu| < |\theta|] \vee [(|\mu| = |\theta|) \wedge (\mu \leq_{lex} \theta)].$$

En este caso sí se puede establecer un isomorfismo de monoides entre $(\mathbb{N}^n, +)$ y $(\mathbb{N}, +)$ que, además, respeta las respectivas relaciones de orden.

Los órdenes monomiales juegan un papel esencial en la ordenación de los monomios en varias variables. Así, consideremos un conjunto de variables $\{X_1, \dots, X_n\}$, algebraicamente independientes sobre un cuerpo K . Y consideremos la base monomial del K -espacio vectorial definida por esos monomios:

$$\mathcal{M}(X_1, \dots, X_n) := \{X_1^{\mu_1} \cdots X_n^{\mu_n} : (\mu_1, \dots, \mu_n) \in \mathbb{N}^n\}.$$

Tenemos claramente un isomorfismo de monoides, donde en $\mathcal{M}(X_1, \dots, X_n)$ se considera la restricción de la operación producto natural del anillo de polinomios $K[X_1, \dots, X_n]$ dada del modo siguiente:

$$(A.2.1) \quad \begin{array}{ccc} \exp : (\mathcal{M}(X_1, \dots, X_n), \cdot) & \longrightarrow & (\mathbb{N}^n, +) \\ X_1^{\mu_1} \cdots X_n^{\mu_n} & \longmapsto & (\mu_1, \dots, \mu_n) \end{array}$$

Dado un término $T = aX_1^{\alpha_1} \cdots X_n^{\alpha_n} \in K[X_1, \dots, X_n]$, con $a \neq 0$, llamaremos exponente monomial del término al exponente de su monomio, i.e.

$$\exp(T) := \exp(X_1^{\alpha_1} \cdots X_n^{\alpha_n}) := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n.$$

Los órdenes monomiales también tienen una relación particular con el pre-orden de Dickson, aunque este último no sea un orden monomial.

PROPOSICIÓN A.2.5. *Con las notaciones precedentes, se verifica:*

- i) La relación \preceq es una relación de orden parcial sobre $(\mathbb{N}^n, +)$, que no es un orden monomial.
- ii) Si \leq es un orden monomial cualquiera en $(\mathbb{N}^n, +)$, dados $\alpha, \beta \in \mathbb{N}^n$, si $\alpha \preceq \beta$, entonces $\alpha \leq \beta$. El recíproco no es cierto, como lo prueban ejemplos muy elementales.
- iii) Dados $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ y $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$, son equivalentes:

$$\alpha \preceq \beta \iff X_1^{\alpha_1} \cdots X_n^{\alpha_n} \mid X_1^{\beta_1} \cdots X_n^{\beta_n},$$

donde \mid significa “divide” en $K[X_1, \dots, X_n]$.

- iv) Dado un orden monomial cualquiera \leq sobre el monoide $(\mathbb{N}^n, +)$ y cualquier subconjunto no vacío $S \subseteq \mathbb{N}^n$, existe un mínimo de S para \leq . Más aún, ese mínimo de S se encuentra entre los elementos minimales (en número finito) de S con respecto al pre-orden de Dickson \preceq .

DEMOSTRACIÓN. Todas las afirmaciones son inmediatas. Por exponer alguna, probemos la Afirmación iv). Supongamos que m es el mínimo de S , que existe porque el orden monomial \leq es un buen orden. Denotemos por $m(S)$ el conjunto (finito) de los elementos minimales de S . Observamos anteriormente que:

$$S \subseteq \bigcup_{m' \in m(S)} m' + \mathbb{N}^n.$$

Todo elemento de S debe estar “por encima” de algún elemento minimal (i.e. para cada $m'' \in S$ debe existir $m' \in m(S)$ tal que $m' \preceq m''$). Pero eso significa que existe $m' \in m(S)$ tal que $m \in m' + \mathbb{N}^n$ o, equivalentemente, que $m' \preceq m$. Pero, por la Afirmación ii), si $m' \preceq m$, entonces $m' \leq m$ y, como m es minimal de S para \leq , se concluye que $m' = m$. \square

Fijado un cuerpo K (o un anillo R) los polinomios en n variables con coeficientes en K son simplemente funciones $f \in K^{\mathbb{N}^n}$ con *soporte* finito, donde el soporte de f es el conjunto dado por la siguiente identidad:

$$(A.2.2) \quad \text{supp}(f) := \{\mu \in \mathbb{N}^n : f(\mu) \neq 0\}.$$

Dicho de otro modo, el anillo de polinomios $K[X_1, \dots, X_n]$ es simplemente el subanillo del anillo de las series de potencias formales $K[[X_1, \dots, X_n]]$ dado por la siguiente igualdad:

$$K[X_1, \dots, X_n] := \{f \in K[[X_1, \dots, X_n]] : \#(\text{supp}(f)) < \infty\}.$$

Así, un monomio $X_1^{\mu_1} \cdots X_n^{\mu_n} \in K[X_1, \dots, X_n]$, con $\mu := (\mu_1, \dots, \mu_n) \in \mathbb{N}^n$ es simplemente la función:

$$(A.2.3) \quad \begin{aligned} X_1^{\mu_1} \cdots X_n^{\mu_n} : \mathbb{N}^n &\longrightarrow K \\ \theta &\longmapsto \begin{cases} 1, & \text{si } \theta = \mu \\ 0, & \text{en otro caso.} \end{cases} \end{aligned}$$

Esto conduce a la representación usual de los polinomios como combinaciones lineales de monomios. A diferencia de la “representación densa”, nos concentraremos en la representación “dispersa” (conocida en inglés como “sparse”, “fewnomial” o, en francés, como “creuse”) en la que sólo destacaremos los monomios con coeficiente no nulo. Esto es, presentamos los polinomios como el siguiente tipo de combinación lineal:

$$(A.2.4) \quad f := \sum_{\mu \in \text{supp}(f)} a_\mu X_1^{\mu_1} \cdots X_n^{\mu_n} \in K[X_1, \dots, X_n],$$

donde, por definición, $a_\mu := f(\mu) \in K$. Obviamente, esta presentación es única, lo que recuerda que $\mathcal{M}(X_1, \dots, X_n)$ es una base de $K[X_1, \dots, X_n]$ como K -espacio vectorial. A los polinomios no nulos de la forma $T := a_\mu X_1^{\mu_1} \cdots X_n^{\mu_n} \in K[X_1, \dots, X_n] \setminus \{0\}$ se les denomina *términos* y al elemento $a_\mu \in K \setminus \{0\}$ se le denomina coeficiente del término no nulo. Al elemento $\mu = (\mu_1, \dots, \mu_n)$ se le denomina exponente del término $\text{exp}(T) := \mu$, mientras que se denomina grado del término al valor $|\mu| = \mu_1 + \cdots + \mu_n \in \mathbb{N}$.

Obviamente, dado $f \in K[X_1, \dots, X_n]$, su soporte $\text{supp}(f) \subseteq \mathbb{N}^n$ define una clase de conceptos finita que puede ser ordenada mediante dos estrategias:

- *Mediante el pre-orden de Dickson:* En este caso, dado $f \in K[X_1, \dots, X_n] \setminus \{0\}$, $\text{supp}(f)$ tendrá un número finito de elementos minimales y maximales. En el caso en que $\text{supp}(f)$ sea cerrado hacia abajo diremos que el polinomio es (genéricamente) denso. El conjunto $\text{supp}(f)$ tendrá un sólo mínimo y constituirá una escalera (finita) determinada por sus elementos maximales.
- *Mediante un orden monomial \leq :* Por otro lado, dado $f \in K[X_1, \dots, X_n] \setminus \{0\}$, el orden inducido sobre el conjunto finito $\text{supp}(f)$ por \leq será un orden total y un buen orden, con lo que tendrá un único máximo y un único mínimo, coincidentes si y solamente si f es un término nulo. Así, fijado el orden monomial \leq sobre $(\mathbb{N}^n, +)$, y dado $f \in K[X_1, \dots, X_n]$, supongamos que $\mu := \max_{\leq}(\text{supp}(f))$ es el elemento maximal del soporte de f para \leq , llamaremos *término director de f* al término no nulo cuyo exponente sea, precisamente, ese máximo μ con respecto a \leq . Lo denotaremos mediante $\ell t_{\leq}(f)$ y se define formalmente como sigue:

$$\ell t_{\leq}(f) := a_\mu X_1^{\mu_1} \cdots X_n^{\mu_n} \in K[X_1, \dots, X_n],$$

donde $\mu := (\mu_1, \dots, \mu_n) = \max_{\leq}(\text{supp}(f))$. Como en el caso de monomios y términos, llamaremos a ese máximo y lo denotaremos mediante:

$$\text{exp}_{\leq}(f) := \max_{\leq}(\text{supp}(f)).$$

Obviamente, exponentes y términos directores de cada polinomio dependen del orden monomial elegido.

Otra medida distinta del “tamaño” de la representación dispersa de un polinomio es su grado que resulta más sencillo de definir por el único orden en el monoide $(\mathbb{N}, +)$:

DEFINICIÓN 29 (GRADO Y GRADOS PARCIALES). *Sea $f \in K[X_1, \dots, X_n] \setminus \{0\}$ un polinomio no nulo.*

i) Llamaremos grado (total) de f a la cantidad dada por la siguiente igualdad:

$$\text{deg}(f) := \max\{|\mu| : \mu \in \text{supp}(f)\}.$$

ii) Fijado $i \in [n]$, consideremos el cuerpo de funciones racionales:

$$K_i := K(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n),$$

que no contiene a la variable X_i . Como $f \in K_i[X_i]$, llamaremos grado de f con respecto a la variable X_i y lo denotaremos mediante $\text{deg}_{X_i}(f)$ al grado de f como elemento del anillo de polinomios $K_i[X_i]$.

Como la Matemática tiene esa tendencia a llamar con nombres distintos a las mismas cosas, dependiendo del contexto, vamos a tratar de ver que lo descrito en esta sección está fuertemente ligado a lo descrito en la Sección A.1 precedente en una sencilla proposición que no requiere prueba por lo expuesto hasta ahora.

PROPOSICIÓN A.2.6. *Con las notaciones precedentes, sea $\mathcal{Y} = \{0, \dots, p-1\}$. Entonces, se tiene:*

i) Dado cualquier polinomio no nulo $f \in K[X_1, \dots, X_n]$, tal que $\text{deg}_{X_i}(f) \leq p-1$, para cada $i \in [n]$, su soporte $\text{supp}(f) \subseteq \mathcal{Y}^n$ es una clase de conceptos finita.

ii) Recíprocamente, dada una clase de conceptos $H \subseteq \mathcal{Y}^n \subseteq \mathbb{N}^n$, entonces, existe un espacio vectorial $W(H) \subseteq K[X_1, \dots, X_n]$, de dimensión igual al cardinal $\sharp(H)$, formado por polinomios tales que $\deg_{X_i}(f) \leq p - 1$, para cada $i \in [n]$, de tal que modo que para cada polinomio no nulo $f \in W(H)$ se tiene que:

$$\text{supp}(f) \subseteq H.$$

iii) El conjunto $W(H)^* := W(H) \cap (K^\times)^{\mathbb{N}^n}$ (tórico, i.e. con todos sus términos con coeficientes no nulos y, por tanto, un abierto Zariski en $W(H)$) satisface:

$$\text{supp}(f) = H \iff f \in W(H)^*.$$

Más aún, la clase de conceptos $H \subseteq \mathcal{Y}^n$ es cerrada hacia abajo si y solamente si todos los polinomios $f \in W(H)^*$ son (genéricamente) densos.

La existencia de un orden es un elemento esencial en la División euclídea y, por tanto, en el desarrollo de algoritmos que trabajan con polinomios univariados. Del mismo modo, la existencia y uso de órdenes monomiales es esencial para concebir una algorítmica de división con polinomios multi-variados: es la *División de Weierstarss-Hironaka* o, simplemente División de Hironaka (cf. sus trabajos [Hir, 1964a] y [Hir, 1964b], que le valieron la medalla Fields del año 1970), que discutimos a continuación:

TEOREMA A.2.7 (DIVISIÓN DE WEIERSTARSS-HIRONAKA). Sea \leq un orden monomial sobre el monoide $(\mathbb{N}^n, +)$. Dados $f_1, \dots, f_r \in K[X_1, \dots, X_n]$ y $g \in K[X_1, \dots, X_n]$. Supongamos fijados los exponentes (con respecto a \leq) de cada f_i mediante $\mu_i := \text{exp}_{\leq}(f_i)$, $i \in [r]$.

Definamos los siguientes conjuntos de exponentes monomiales en el monoide $(\mathbb{N}^n, +)$:

$$\begin{aligned} \Delta_1 &:= \mu_1 + \mathbb{N}^n, \\ \Delta_2 &:= (\mu_2 + \mathbb{N}^n) \setminus \Delta_1, \\ \Delta_3 &:= (\mu_3 + \mathbb{N}^n) \setminus (\Delta_1 \cup \Delta_2), \\ &\vdots \\ \Delta_r &:= (\mu_r + \mathbb{N}^n) \setminus (\cup_{i=1}^{r-1} \Delta_i), \\ \bar{\Delta} &:= \mathbb{N}^n \setminus (\cup_{i=1}^r \Delta_i). \end{aligned}$$

Entonces, existen polinomios $g_1, \dots, g_r, h \in K[X_1, \dots, X_n]$ tales que se satisfacen las siguientes propiedades :

- i) Para cada $i \in [r]$, los soportes satisfacen $\text{supp}(g_i X^{\mu_i}) \subseteq \Delta_i$.
- ii) El soporte del polinomio h satisface: $\text{supp}(h) \subseteq \bar{\Delta}$.
- iii) Se verifica la siguiente igualdad :

$$g := g_1 f_1 + \dots + g_r f_r + h.$$

- iv) Los polinomios g_1, \dots, g_r verificando las anteriores propiedades son únicos y se denominan cocientes de la División de Hironaka.
- v) El polinomio h es único y se denomina resto de la División de Hironaka.

Además, si el orden monomial \leq es algoritmizable, existe un algoritmo que calcula los polinomios g_1, \dots, g_r, h en su forma dispersa (i.e. como en la Identidad (A.2.4)).

Ejemplos de los algoritmo que se citan en el enunciado se pueden ver en [Par, 24a] o en [CLO, 07].

Del mismo modo que hemos introducido exponentes monomiales para monomios, términos y polinomios podemos introducir los exponentes de un ideal:

DEFINICIÓN 30 (EXPONENTE DE UN IDEAL). Sea K un cuerpo y $K[X_1, \dots, X_n]$ el anillo de polinomios en n variables con coeficientes en el cuerpo K . Sea $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$ un ideal de ese anillo, y \leq un orden monomial en el monoide $(\mathbb{N}^n, +)$. Llamaremos exponente del ideal \mathfrak{a} con respecto a \leq al conjunto formado por los exponentes de todos los elementos del ideal, i.e.

$$(A.2.5) \quad \text{exp}_{\leq}(\mathfrak{a}) := \{\text{exp}_{\leq}(f) : f \in \mathfrak{a}\} \subseteq \mathbb{N}^n.$$

PROPOSICIÓN A.2.8. Sea \leq un orden monomial sobre el monoide $(\mathbb{N}^n, +)$ y sean K un cuerpo y $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$ un ideal de ese anillos de polinomios. Entonces, se tiene

- i) El conjunto $\text{exp}_{\leq}(\mathfrak{a})$ es cerrado hacia arriba para el pre-orden de Dickson \preceq .
- ii) Su complementario $\mathbb{N}^n \setminus \text{exp}_{\leq}(\mathfrak{a})$ es cerrado hacia abajo para el pre-orden de Dickson \preceq .

DEMOSTRACIÓN. Es casi inmediato. Probemos i): si $\alpha \preceq \beta$ y $\alpha \in \text{exp}_{\leq}(\mathfrak{a})$ es porque existen:

- $\gamma := (\gamma_1, \dots, \gamma_n) \in \mathbb{N}^n$ tal que $\beta = \alpha + \gamma$ en el monoide $(\mathbb{N}^n, +)$.

- Existe $f \in \mathfrak{a}$ tal que $lt_{\leq}(f) = \alpha$. Consideremos el polinomio:

$$g := X_1^{\gamma_1} \cdots X_n^{\gamma_n} f \in \mathfrak{a}.$$

Como los órdenes monomiales respetan la estructura del monoide $(\mathbb{N}^n, +)$ tendremos que:

$$lt(g) = \gamma + lt(f) = \gamma + \alpha = \beta \in \exp_{\leq}(\mathfrak{a}).$$

Esto prueba que $\exp_{\leq}(\mathfrak{a})$ es cerrado hacia arriba para \preceq . Obviamente, su complementario será cerrado hacia abajo para \preceq . \square

TEOREMA A.2.9 (EXISTENCIA DE BASES DE GRÖBNER). *Sea K un cuerpo, y sea $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$ un ideal en el anillo de polinomios $K[X_1, \dots, X_n]$. Sea \leq un orden monomial en el monoide $(\mathbb{N}^n, +)$. Entonces, existe una familia finita $G := \{f_1, \dots, f_s\} \subseteq \mathfrak{a}$ de elementos, dos a dos distintos, en el ideal \mathfrak{a} de tal modo que se verifica:*

$$(A.2.6) \quad \exp_{\leq}(\mathfrak{a}) = \bigcup_{i=1}^s (\exp_{\leq}(f_i) + \mathbb{N}^n).$$

En particular, el ideal \mathfrak{a} es generado por el conjunto finito $\mathcal{G} := \{f_1, \dots, f_s\}$ (lo que implica el Basissatz de Hilbert para el anillo de polinomios $K[X_1, \dots, X_n]$).

A cualquier conjunto finito de elementos \mathcal{G} que satisface la Identidad (A.2.6) se le denomina base de Gröbner del ideal \mathfrak{a} para el orden monomial \leq .

Más aún, una base como K -espacio vectorial del anillo cociente $K[X_1, \dots, X_n]/\mathfrak{a}$ es dada por el conjunto:

$$(A.2.7) \quad \mathcal{B}/\mathfrak{a} := \{X_1^{\mu_1} \cdots X_n^{\mu_n} + \mathfrak{a} : \mu := (\mu_1, \dots, \mu_n) \in \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})\}.$$

Al conjunto \mathcal{B}/\mathfrak{a} se le denomina base monomial del anillo cociente $K[X_1, \dots, X_n]/\mathfrak{a}$.

DEMOSTRACIÓN. El enunciado es consecuencia casi inmediata de combinar el Lema de Dickson (cf. Teorema A.1.2), con la División de Weierstarss-Hironaka (cf. Teorema A.2.7). Comencemos probando la Igualdad (A.2.6): Dada la inclusión, $\exp_{\leq}(\mathfrak{a}) \subseteq \mathbb{N}^n$, el Lema de Dickson garantiza que posee un conjunto finito de elementos minimales para \preceq . Supongamos que ese conjunto de elementos minimales es el conjunto $\mathcal{M} := \{\mu_1, \dots, \mu_s\}$ y, sin pérdida de la generalidad, podemos suponer que son dos a dos distintos, esto es, $\mu_i \neq \mu_j$, para cada $i, j \in [s]$, con $i \neq j$. Como el orden monomial es un buen orden podemos también suponer:

$$\mu_s < \mu_{s-1} < \cdots < \mu_1,$$

donde $< := \preceq$ y \leq es el orden monomial fijado en el enunciado.

En consecuencia, para cada $j \in [s]$, como cada $\mu_j \in \exp_{\leq}(\mathfrak{a})$, existirá al menos un polinomio $f_j \in \mathfrak{a}$ tal que $\mu_j := \exp_{\leq}(f_j)$. En particular, los términos directores de esta familia $\mathcal{G} := \{f_1, \dots, f_s\} \subseteq \mathfrak{a}$ tienen la forma siguiente:

$$lt_{\leq}(f_j) = a_{\mu_j} X_1^{\mu_{j,1}} \cdots X_n^{\mu_{j,n}},$$

con $a_{\mu_j} \in K \setminus \{0\}$. Por tanto, al tener coeficientes directores distintos, podemos también garantizar que dados $i, j \in [s]$, con $i \neq j$, se satisface $f_i \neq f_j$ y, por supuesto, $lt_{\leq}(f_i) \neq lt_{\leq}(f_j)$ para el mismo par de índices distintos. Seguidamente, observemos que para $\theta := (\theta_1, \dots, \theta_n) \in \mathbb{N}^n$, se tiene que:

- $X_1^{\theta_1} \cdots X_n^{\theta_n} f_j \in \mathfrak{a}$, para cada $j \in [s]$, simplemente porque \mathfrak{a} es un ideal.
- Como el orden \leq es un orden monomial, se tiene para cada $j \in [s]$ que:

$$lt_{\leq}(X_1^{\theta_1} \cdots X_n^{\theta_n} f_j) = lt_{\leq}((X_1^{\theta_1} \cdots X_n^{\theta_n}) \cdot lt_{\leq}(f_j)) = \theta + \mu_j \in \exp_{\leq}(f_j) + \mathbb{N}^n.$$

Por tanto concluimos una de las inclusiones descritas en la Identidad (A.2.6):

$$\exp_{\leq}(\mathfrak{a}) \supseteq \bigcup_{i=1}^s (\exp_{\leq}(f_i) + \mathbb{N}^n).$$

Para la otra incusión, observemos que para cada $\theta \in \exp_{\leq}(\mathfrak{a})$ deberá existir un elementos $j \in [s]$ tal que $\mu_j \preceq \theta$ para el pre-orden de Dickson. Por la Afirmación *ii*) de la Proposición 1.2.3 concluimos también que $\mu_j \leq \theta$ para el orden monomial fijado. Más aún, existirá $\gamma := (\gamma_1, \dots, \gamma_n) \in \mathbb{N}^n$ tal que $\theta = \mu_j + \gamma \in \mu_j + \mathbb{N}^n$. Con ello concluimos el otro contenido:

$$\exp_{\leq}(\mathfrak{a}) \subseteq \bigcup_{i=1}^s (\exp_{\leq}(f_i) + \mathbb{N}^n),$$

y la Igualdad (A.2.6).

Por otro lado, ya hemos visto que $f_i \in \mathfrak{a}$ para cada $i \in [s]$, con lo que tenemos la inclusión de ideales $(f_1, \dots, f_s) \subseteq \mathfrak{a}$.

Para la otra inclusión de ideales, consideremos $g \in \mathfrak{a}$ y la familia de subconjuntos $\Delta_1, \dots, \Delta_s, \overline{\Delta} \subseteq \mathbb{N}^n$ definidos

(a partir de la familia $\{f_1, \dots, f_s\}$) como en el Teorema A.2.7. Observamos que, por construcción, se tiene la siguiente cadena de igualdades:

$$(A.2.8) \quad \bigcup_{i=1}^s \Delta_i = \bigcup_{i=1}^s (\exp_{\leq}(f_i) + \mathbb{N}^n) = \exp_{\leq}(\mathfrak{a}).$$

También observamos que se tiene:

$$(A.2.9) \quad \overline{\Delta} = \mathbb{N}^n \setminus \left(\bigcup_{i=1}^s \Delta_i \right) = \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a}).$$

Por el Teorema A.2.7, existirán polinomios $g_1, \dots, g_s, h \in K[X_1, \dots, X_n]$ tales que se verifican las afirmaciones del citado teorema. En particular, observamos que

$$g = g_1 f_1 + \dots + g_s f_s + h,$$

con lo que, dada la inclusión de ideales $(f_1, \dots, f_s) \subseteq \mathfrak{a}$, podemos concluir que $h \in \mathfrak{a}$. Pero si $h \neq 0$, entonces $\exp_{\leq}(h) \in \exp_{\leq}(\mathfrak{a})$. Pero por la Afirmación *ii*) del Teorema A.2.7, tenemos que $\ell t_{\leq}(h) \in \text{supp}(h) \subseteq \overline{\Delta} = \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})$. Esto nos lleva a contradicción y, por tanto, $h = 0$, lo que implica:

$$g = g_1 f_1 + \dots + g_s f_s \in (f_1, \dots, f_s),$$

y tendremos la igualdad de ideales afirmada en el enunciado:

$$(A.2.10) \quad \mathfrak{a} = (f_1, \dots, f_s).$$

Por otro lado, de nuevo por la División de Hironaka los elementos del conjunto \mathcal{B}/\mathfrak{a} son los que reproducimos a continuación:

$$\mathcal{B}/\mathfrak{a} := \{X_1^{\mu_1} \dots X_n^{\mu_n} + \mathfrak{a} : \mu := (\mu_1, \dots, \mu_n) \in \overline{\Delta}\},$$

donde $\overline{\Delta} = \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})$ en el lenguaje del Teorema A.2.7. Consideremos ahora una combinación lineal (con coeficientes en K) de los elementos de \mathcal{B}/\mathfrak{a} igualada a cero en $K[X_1, \dots, X_n]/\mathfrak{a}$:

$$0 + \mathfrak{a} := \sum_{\mu \in \overline{\Delta}} a_{\mu} (X_1^{\mu_1} \dots X_n^{\mu_n} + \mathfrak{a}) = h + \mathfrak{a},$$

donde

$$h := \sum_{\mu \in \overline{\Delta}} a_{\mu} X_1^{\mu_1} \dots X_n^{\mu_n} \in K[X_1, \dots, X_n].$$

Por la Identidad (A.2.10), combinada con la División de Hironaka, concluimos que $h \in \mathfrak{a}$ si y solamente si $h = 0$ en $K[X_1, \dots, X_n]$. Pero eso sólo es posible si todos sus coeficientes son nulos, esto es, si $a_{\mu} = 0, \forall \mu \in \overline{\Delta}$. Esto demuestra que la familia \mathcal{B}/\mathfrak{a} es una familia libre en el anillo cociente $K[X_1, \dots, X_n]/\mathfrak{a}$ visto como K -espacio vectorial. De otro lado, a partir de la Igualdad (A.2.10) de ideales, combinada, de nuevo, con el Teorema A.2.7, aplicado a la familia $\{f_1, \dots, f_s\}$ que genera el ideal \mathfrak{a} , concluimos que para cada $g \in K[X_1, \dots, X_n]$ existe un elemento de la forma $h + \mathfrak{a}$ en el K -subespacio vectorial generado por \mathcal{B}/\mathfrak{a} de tal modo que:

$$g + \mathfrak{a} = g + (f_1, \dots, f_s) = h + \mathfrak{a}.$$

Por tanto \mathcal{B}/\mathfrak{a} es una base de $K[X_1, \dots, X_n]/\mathfrak{a}$ como K -espacio vectorial y habremos concluido tanto la existencia de Bases de Gröbner como la existencia de bases monomiales de anillos residuales $K[X_1, \dots, X_n]/\mathfrak{a}$ para órdenes monomiales prefijados. \square

Obsérvese que las Bases de Gröbner \mathcal{G} no son, en realidad, bases del ideal como K -espacio vectorial, sino sistemas finitos de generadores. En cambios, las clases \mathcal{B}/\mathfrak{a} , definidas como las clases módulo \mathfrak{a} del complementario de $\exp_{\leq}(\mathfrak{a})$, sí constituyen una base del anillo residual $K[X_1, \dots, X_n]/\mathfrak{a}$ como K -espacio vectorial. Es como si la terminología estuviera cambiada.

A.2.1. Unas palabras sobre la existencia de algoritmos para el cálculo de las Bases de Gröbner.

Aunque este Trabajo Fin de Grado no trata de algoritmos sobre Bases de Gröbner, citaremos un poco de la historia y de las primeras contribuciones debidas a Bruno Buchberger. La historia, en los tiempos recientes, comienza con los trabajos de H. Hironaka sobre *resolución de singularidades en característica cero* (cf. [Hir, 1964a] y [Hir, 1964b]), aunque hay trazos de la misma idea en la obra de K. Weierstrass (cf. [Wei, 1854, Wei, 1876]). En estos trabajos ya aparece la idea de la División de Weierstrass-Hironaka y la idea de *bases estándar* que preluirá la obra de B. Buchberger.

En [Buc, 1965] y en [Buc, 1985], B. Buchberger trata de interpretar la noción de *bases estándar* usada por Hironaka, pero adaptándola al anillo de polinomios $K[X_1, \dots, X_n]$. Las denomina *Bases de Gröbner* en honor

de su director de tesis W. Gröbner quien, a su vez, hizo investigación bajo la dirección de E. Noether en Göttingen entre 1932 y la marcha de E. Noether a Suiza.

Lo que hace Buchberger es desarrollar un algoritmo elemental (de ahí su popularidad) basado en *técnicas de reescritura en monoides*. No es un algoritmo eficiente ni teóricamente ni experimentalmente. Los algoritmos experimentalmente más eficientes hasta la fecha son los debidos a J.C. Faugère quien, sin embargo, abandonó el ámbito académico en la cumbre de su éxito (siendo ya, como poco, *Directeur de Recherche du CNRS*), para ganar más dinero en el mundo de la industria de software. Ha habido muchas variantes del algoritmo de Buchberger, muchos resultados que muestran que las técnicas de reescritura no son del todo las más naturales y existe toda una bibliografía extensísima con los pros y los contras de este tipo de técnicas. Nos limitamos con dar al lector unas pocas referencias para que inicie su búsqueda en caso de tener interés: [CLO, 07], [Mor, 16], [Stur, 1993], [HHPS, 20], [HHPS, 21] y un extensísimo etcétera en el que no hemos incluido, precisamente, las alternativas mucho más eficientes para los propósitos de la Teoría de la Eliminación.

A.2.2. El caso de los anillos de Artin y las escaleras. El matemático austriaco Emil Artin se incorpora a la Universidad de Göttingen en 1921. Allí coincidirá con D. Hilbert y colaborará con E. Noether y H. Hasse, hasta su marcha, en 1923, a la Universidad de Hamburgo y su exilio en Princeton, por causa del nazismo, en 1937. Mientras se encuentra en Göttingen, colabora en el seminario de E. Noether. No fue E. Artin, sino E. Noether quien introdujo en su seminario la condición de cadena numerable descendente para ideales en anillos. Era una terminología enmarcada en sus conferencias sobre Teoría de Grupos, como noción dual de sus propias ideas sobre la condición de cadena ascendente numerable. El testimonio proviene del texto histórico [vdW, 1985], cuyo autor fue testigo privilegiado de aquel tiempo.

En 1927, ya en Hamburgo, Artin generaliza el teorema de Wedderburn de 1908 sobre números hiper-complejos (es decir, una teoría de álgebras sobre cuerpos arbitrarios). La generalización de Artin en [Artin, 1927] reposa sobre esa condición de cadena descendente numerable de Noether (de nuevo, según testimonio de [vdW, 1985]). En nuestro contexto, nos interesan especialmente las K -álgebras de Artin (terminología más actualizada de los sistemas hiper-complejos) y, por extensión, los *anillos de Artin*, que será el objeto de este capítulo. Todo el material que sigue puede verificarse en [Par, 24a].

DEFINICIÓN 31 (ANILLO DE ARTIN). *Un anillo R se dice anillo de Artin (o anillo artiniano) si satisface la siguiente propiedad, conocida como la Condición de Cadena Numerable descendente de ideales : Toda cadena descendente numerable de ideales se estabiliza. Es decir, dada una cadena descendente numerable de ideales de R :*

$$\mathfrak{a}_0 \supseteq \mathfrak{a}_1 \supseteq \cdots \supseteq \mathfrak{a}_n \supseteq \cdots ,$$

existe $m \in \mathbb{N}$ tal que $\mathfrak{a}_n = \mathfrak{a}_m$, $\forall n \geq m$.

DEFINICIÓN 32 (ANILLOS DE DIMENSIÓN CERO). *Un anillo R se dice anillo de dimensión (de Krull) cero (o anillo cero-dimensional) si satisface: $\text{Spec}(R) = \text{MaxSpec}(R)$, donde $\text{Spec}(R)$ es el espectro primo del anillo y $\text{MaxSpec}(R)$ es el espectro maximal de R .*

TEOREMA A.2.10 (TEOREMA DE AKIZUKI). *Un anillo R es artiniano si y solamente si se verifican las dos propiedades siguientes:*

- i) R es noetheriano, es decir, todo ideal de R es finitamente generado.*
- ii) Se verifica $\text{Spec}(R) = \text{MaxSpec}(R)$ (i.e. todo ideal \mathfrak{p} primo en R es un ideal maximal en R).*

En otras palabras, un anillo de Artin es un anillo cero-dimensional en el que todo ideal es finitamente generado.

Debe recordarse que ser de dimensión cero no garantiza inmediatamente la condición de ser anillo de Artin (i.e. la condición noetheriana es esencial). Un ejemplo puede verse en [Par, 24a], Capítulo 10:

EJEMPLO A.2.11 (DIMENSIÓN CERO NO GARANTIZA SER ANILLO DE ARTIN). Sea K un cuerpo y, por tanto, un anillo artiniano. Consideramos una familia numerable de variables $\{X_n : n \in \mathbb{N}\}$ y el anillo de polinomios en ese conjunto numerable de variables:

$$R := K[X_n : n \in \mathbb{N}].$$

En ese anillo, que es K -álgebra, podemos considerar el ideal \mathfrak{a} generado por los cuadrados de esas variables. Esto es, consideremos:

$$\mathfrak{a} := (\{X_n^2 : n \in \mathbb{N}\}).$$

Consideremos el la K -álgebra cociente de R por ese ideal:

$$A := R/\mathfrak{a}.$$

La K -álgebra A será un anillo cero-dimensional: sólo posee como maximal el ideal:

$$\mathfrak{m}/\mathfrak{a},$$

donde

$$\mathfrak{m} := (X_n : n \in \mathbb{N}),$$

pero no es un anillo de Artin porque ni siquiera es noetheriano: el ideal $\mathfrak{m}/\mathfrak{a}$ no es finitamente generado.

En [Par, 24a] se prueba una larga lista de equivalencias sobre la condición de ser K -álgebra de Artin. Nos conformaremos con mostrar algunas de ellas. Comencemos con una notación para referirnos a las variedades algebraicas:

DEFINICIÓN 33 (VARIETADES ALGEBRAICAS AFINES K -DEFINIBLES). Sea K un cuerpo, \mathbb{K} una extensión algebraicamente cerrada y $\mathbb{A}^n(\mathbb{K})$ el espacio afín sobre \mathbb{K} de dimensión n . Consideremos el anillo de polinomios en n variables $K[X_1, \dots, X_n]$ y su extensión $\mathbb{K}[X_1, \dots, X_n]$ con coeficientes en \mathbb{K} .

i) Sea $f \in K[X_1, \dots, X_n]$, llamaremos hiper-superficie afín definida por f en $\mathbb{A}^n(\mathbb{K})$, que denotaremos por $V_{\mathbb{A}^n(\mathbb{K})}(f)$, al conjunto:

$$V_{\mathbb{A}^n(\mathbb{K})}(f) := \{x \in \mathbb{A}^n(\mathbb{K}) : f(x) = 0\} = f^{-1}(\{0\}).$$

ii) Llamaremos intersección finita de hiper-superficies algebraicas K -definibles a la intersección de un número finito de tales hiper-superficies definidas por una familia finita $\{f_1, \dots, f_s\} \subseteq K[X_1, \dots, X_n]$ y lo denotaremos mediante:

$$V_{\mathbb{A}^n(\mathbb{K})}(f_1, \dots, f_s) := \bigcap_{i=1}^s V_{\mathbb{A}^n(\mathbb{K})}(f_i).$$

iii) Dado un ideal $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$, llamaremos variedad algebraica K -definible asociada al ideal \mathfrak{a} al conjunto

$$V_{\mathbb{A}^n(\mathbb{K})}(\mathfrak{a}) := \bigcup_{f \in \mathfrak{a}} V_{\mathbb{A}^n(\mathbb{K})}(f) = \{\zeta \in \mathbb{A}^n(\mathbb{K}) : f(\zeta) = 0, \forall f \in \mathfrak{a}\}.$$

En el caso de que no haya confusión entre las dimensiones (n), los cuerpos K y \mathbb{K} involucrados (especialmente cuando $K = \mathbb{K}$), nos conformaremos con usar las notaciones $V_{\mathbb{A}}(f)$, $V_{\mathbb{A}}(f_1, \dots, f_s)$ y $V_{\mathbb{A}}(\mathfrak{a})$ para los mismos objetos.

Por supuesto el famoso Basissatz de Hilbert implica que las variedades algebraicas K -definibles son exactamente las intersecciones finitas de hiper-superficies K -definibles. Y que eso no depende del sistema generador finito del ideal elegido para esa intersección finita de hiper-superficies.

En ocasiones no nos interesan solamente las soluciones de un sistema de ecuaciones polinomiales K - en una extensión algebraica del cuerpo K sino en el propio cuerpo de coeficientes (o, incluso, en un subcuerpo, aunque es un caso menos habitual).

DEFINICIÓN 34 (VARIETADES Y PUNTOS ALGEBRAICOS L -RACIONALES). Con notaciones similares a las de la anterior definición, sean $L \subseteq K$ una extensión de cuerpos, $\mathbb{A}^n(L)$ el espacio afín de dimensión n sobre L . Sea $\mathfrak{a} \subseteq K[X_1, \dots, X_n]$ un ideal. Llamaremos variedad de puntos L -racionales asociada a \mathfrak{a} al conjunto siguiente:

$$V_{\mathbb{A}^n(L)}(\mathfrak{a}) := V_{\mathbb{A}^n(\mathbb{K})}(\mathfrak{a}) \cap \mathbb{A}^n(L).$$

A los puntos de $V_{\mathbb{A}^n(L)}(\mathfrak{a})$ se les denomina denomina puntos L -racionales de la variedad algebraica $V_{\mathbb{A}^n(\mathbb{K})}(\mathfrak{a})$. De nuevo, cuando no haya confusión con los cuerpos L, K y \mathbb{K} o las dimensiones escribiremos simplemente $V_{\mathbb{A}(L)}(\mathfrak{a})$.

En el caso de sistemas hiper-complejos (i.e. K -álgebras finitamente generadas que son anillos de Artin), se tiene la siguiente caracterización que también puede verse en [Par, 24a]. Se trata del siguiente enunciado:

TEOREMA A.2.12. Sea K un cuerpo y \mathbb{K} su clausura algebraica. Sea $\{f_1, \dots, f_s\}$ un conjunto finito de polinomios en $K[X_1, \dots, X_n]$. Sea $\mathfrak{a} = (f_1, \dots, f_s)$ el ideal que genera ese conjunto en $K[X_1, \dots, X_n]$ y sea \mathfrak{a}^e la extensión de \mathfrak{a} al anillo en $\mathbb{K}[X_1, \dots, X_n]$. Sea $V_{\mathbb{A}}(\mathfrak{a}) = V_{\mathbb{A}}(f_1, \dots, f_s) = V_{\mathbb{A}}(\mathfrak{a}^e) \subseteq \mathbb{A}^n(\mathbb{K})$ la variedad algebraica K -definible de sus ceros comunes. Son equivalentes:

i) El sistema de ecuaciones siguiente:

$$(A.2.11) \quad \begin{cases} f_1(X_1, \dots, X_n) = 0, \\ f_2(X_1, \dots, X_n) = 0, \\ \vdots \\ f_s(X_1, \dots, X_n) = 0, \end{cases}$$

posee solamente un número finito de soluciones en $\mathbb{A}^n(\mathbb{K})$.

ii) La variedad $V_{\mathbb{A}}(\mathfrak{a}) = V_{\mathbb{A}}(\mathfrak{a}^e)$ es un conjunto finito de puntos en $\mathbb{A}^n(\mathbb{K})$.

- iii) El anillo $K[X_1, \dots, X_n]/\mathfrak{a}$ es un anillo de Artin.
- iv) El anillo $\mathbb{K}[X_1, \dots, X_n]/\mathfrak{a}^e$ es un anillo de Artin.
- v) El anillo $K[X_1, \dots, X_n]/\mathfrak{a}$ es un K -espacio vectorial de dimensión finita.
- vi) El anillo $\mathbb{K}[X_1, \dots, X_n]/\mathfrak{a}^e$ es un \mathbb{K} -espacio vectorial de dimensión finita.

Además, en cualquiera de estos casos, se tienen las siguientes igualdades de dimensiones:

$$\dim_K(K[X_1, \dots, X_n]/\mathfrak{a}) = \dim_{\mathbb{K}}(\mathbb{K}[X_1, \dots, X_n]/\mathfrak{a}^e).$$

De toda esta discusión abstracta, extraemos simplemente la equivalencia entre la condición de ser una K -álgebra de Artin y las clases de conceptos finitas introducidas en la Sección 1.1 precedente. En realidad, el resultado no debe ser muy desconocido, pero lo escribimos y lo probamos por falta de una referencia más apropiada.

COROLARIO A.2.13 (ESCALERAS FINITAS EQUIVALEN A BASES MONOMIALES DE K -ÁLGEBRAS DE ARTIN). Sea K un cuerpo de característica distinta de 2. Sea $p \in \mathbb{N}$ un entero positivo con $p \geq 2$. Sea $\mathcal{Y} := \{0, \dots, p-1\}^n$ y sea $H \subseteq \mathcal{Y}^n$ un subconjunto finito. Entonces, son equivalentes:

- i) H es una escalera finita (i.e. un conjunto finito cerrado hacia abajo con respecto a \preceq).
- ii) Para cualquier orden monomial \preceq sobre el monoide $(\mathbb{N}^n, +)$, existe un ideal \mathfrak{a} de tal modo que:
 - $H := \mathbb{N}^n \setminus \exp_{\preceq}(\mathfrak{a})$.
 - La K -álgebra $R := K[X_1, \dots, X_n]/\mathfrak{a}$ es una K -álgebra de Artin y la siguiente es una base monomial de R como K -espacio vectorial:

$$\mathcal{B}(H)/\mathfrak{a} := \{X_1^{\mu_1} \cdots X_n^{\mu_n} + \mathfrak{a} : \mu := (\mu_1, \dots, \mu_n) \in H\},$$

con lo que, además, se tiene la siguiente igualdad de dimensiones como K -espacios vectoriales:

$$\dim_K(K[X_1, \dots, X_n]/\mathfrak{a})_{\#}(H).$$

Además, en ii) se puede elegir que el ideal \mathfrak{a} sea un ideal monomial de $K[X_1, \dots, X_n]$ aunque no es imprescindible.

DEMOSTRACIÓN. Hagamos cada implicación por su lado.

- i) \implies ii): Comencemos considerando el ideal monomial $\mathfrak{b} := (X_1^p, \dots, X_n^p)$ y, claramente, el anillo residual $A := K[X_1, \dots, X_n]/\mathfrak{b}$ es un anillo de Artin y un K -espacio vectorial de dimensión p^n . Seguidamente consideremos el conjunto $\mathcal{Y}^n \setminus H$ y consideremos el conjunto finito de monomios asociado:

$$\mathcal{M}(\mathcal{Y}^n \setminus H) := \{X_1^{\mu_1} \cdots X_n^{\mu_n} : \mu := (\mu_1, \dots, \mu_n) \in \mathcal{Y}^n \setminus H\}.$$

Definamos los exponentes $\{p \cdot e_1, \dots, p \cdot e_n\} \subseteq \mathbb{N}^n$, donde $\{e_1, \dots, e_n\} \subseteq \{0, 1\}^n$ es la base “canónica” usual de \mathbb{R}^n y consideremos el conjunto complementario de H en \mathbb{N}^n dado mediante:

$$H^c := \mathbb{N}^n \setminus H := (\mathcal{Y}^n \setminus H) \cup (\cup_{i=1}^n (p \cdot e_i + \mathbb{N}^n)).$$

Consideremos el conjunto de elementos minimales de H^c con respecto al pre-orden de Dickson y lo denotamos mediante $\mathfrak{m}(H^c)$. Tenemos que, por la propia definición de los elementos minimales de este conjunto y por ser ese conjunto cerrado hacia arriba, se tiene que:

$$(A.2.12) \quad H^c = \bigcup_{\mu \in \mathfrak{m}} (\mu + \mathbb{N}^n).$$

Escribamos:

$$\mathfrak{m}(H^c) := \{\mu_1, \dots, \mu_t\},$$

con la condición de que sean distintos dos a dos. Obviamente, se tiene el contenido siguiente: $\mathfrak{m}(H^c) \subseteq H^c$. Finalmente, definamos la familia de monomios siguiente:

$$\mathcal{M}_0(H^c) := \{X_1^{\mu_{i,1}} \cdots X_n^{\mu_{i,n}} : \mu_i := (\mu_{i,1}, \dots, \mu_{i,n}) \in \mathfrak{m}(H^c)\}.$$

Y consideremos el ideal monomial \mathfrak{a} generado por $\mathcal{M}_0(H^c)$, es decir:

$$\mathfrak{a} := (\mathcal{M}_0(H^c)) = (\mathcal{M}_0(H^c)) + \mathfrak{b},$$

porque los exponentes $p \cdot e_i$ de los monomios de la forma X_i^p tienen que estar en algún $\mu_i + \mathbb{N}^n$, para algún $i \in [t]$ minimal en H^c para \preceq . Por tanto $\mathfrak{b} \subseteq \mathfrak{a}$ y podemos considerar el ideal $\bar{\mathfrak{a}} := \mathfrak{a}/\mathfrak{b}$ de las clases de restos de elementos del ideal \mathfrak{a} módulo el ideal \mathfrak{b} . Además, tenemos un epimorfismo natural de K -álgebras que, combinado con el Segundo teorema de Isomorfía, nos da la siguiente cadena:

$$\pi : A \longrightarrow A/\bar{\mathfrak{a}} \cong K[X_1, \dots, X_n]/\mathfrak{a}.$$

En particular, π es un epimorfismo de K -espacios vectoriales. Por el Teorema A.2.12, concluimos que $K[X_1, \dots, X_n]/\mathfrak{a}$ es un K -espacio vectorial de dimensión finita y una K -álgebra de Artin. Además, observamos que el conjunto siguiente no solamente es un sistema generador del ideal monomial \mathfrak{a} de

$K[X_1, \dots, X_n]$ sino que es lo que en [CLO, 07] (Problem 8, Section 4, Chapter 2) se denomina una base minimal del ideal monomial \mathfrak{a} . La razón es que no puede haber relación de divisibilidad entre ellos porque sus exponentes son los elementos minimales de H^c para \preceq y son distintos dos a dos. Finalmente, observamos que para cualquier orden monomial \leq se ha de tener:

$$(A.2.13) \quad \exp_{\leq}(\mathfrak{a}) = \bigcup_{i=1}^t (\mu_i + \mathbb{N}^n).$$

Un contenido es obvio dado que \mathfrak{a} está generado por los monomios $\mathfrak{m}(H^c) := \{\mu_1, \dots, \mu_t\}$. El otro contenido se sigue del hecho de que un orden monomial \leq induce una graduación sobre el anillo de polinomios $K[X_1, \dots, X_n]$ del tipo siguiente:

$$K[X_1, \dots, X_n] := \bigoplus_{\mu \in \mathbb{N}^n} K \langle X_1^{\mu_1} \cdots X_n^{\mu_n} \rangle,$$

donde $K \langle X_1^{\mu_1} \cdots X_n^{\mu_n} \rangle$ es el K -espacio vectorial de dimensión 1 generado por el monomio $X_1^{\mu_1} \cdots X_n^{\mu_n}$. Los ideales monomiales son, justamente, ideales homogéneos con respecto a esa graduación y, por tanto, si $f \in \mathfrak{a}$, entonces todos sus términos están en \mathfrak{a} . En particular, $\ell_{t_{\leq}}(f) \in \mathfrak{a}$ y, por tanto, debe ser divisible por alguno de los monomios en $\mathcal{M}_o(H^c)$ (que es lo que viene a decir el Lemma 2, Section 4 of Chapter 2 de [CLO, 07]). Por tanto, para cada $f \in \mathfrak{a}$, se tiene:

$$\exp_{\leq}(f) \in \bigcup_{i=1}^t (\mu_i + \mathbb{N}^n) = \bigcup_{\mu \in \mathfrak{m}(H^c)} (\mu + \mathbb{N}^n).$$

Esto prueba la Identidad (A.2.13) que podemos reescribir mediante la cadena de igualdades siguiente usando la Igualdad (A.2.12)

$$\exp_{\leq}(\mathfrak{a}) = \bigcup_{\mu \in \mathfrak{m}(H^c)} (\mu + \mathbb{N}^n) = H^c.$$

Finalmente, el teorema sobre existencia de Bases de Gröbner (cf. Teorema A.2.9) nos permite concluir que $\mathcal{B}(H)/\mathfrak{a}$ es una base de $K[X_1, \dots, X_n]/\mathfrak{a}$ como K -espacio vectorial, de donde se siguen el resto de afirmaciones.

- $ii) \implies i)$: Para probar esta implicación basta con demostrar que $\exp_{\leq}(\mathfrak{a})^c := \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})$ es siempre cerrado hacia abajo y finito, bajo las hipótesis de $ii)$. Comencemos considerando $\mathfrak{b} = (X_1^p, \dots, X_n^p)$ y observando que el anillo residual siguiente es ya un anillo de Artin:

$$A := K[X_1, \dots, X_n]/\mathfrak{b},$$

porque es un espacio vectorial de dimensión p^n cuya base es sencilla de expresar como las clases residuales módulo \mathfrak{b} de los elementos de \mathcal{Y}^n y usando el Teorema 1.2.9 precedente. Como en la implicación precedente, definiendo $\bar{\mathfrak{a}} := \mathfrak{a}/\mathfrak{b}$ como las clases de restos de los elementos de $\mathfrak{a} \supseteq \mathfrak{b}$ módulo el ideal \mathfrak{b} , tenemos claramente un epimorfismo de K -álgebras (dado por el famoso Segundo teorema de Isomorfía) del tipo siguiente:

$$\pi : A \longrightarrow A/\bar{\mathfrak{a}} \cong K[X_1, \dots, X_n]/\mathfrak{a}$$

Por tanto, usando el Teorema 1.2.9 concluimos que $\mathcal{B}/\bar{\mathfrak{a}}$ es base del anillo residual $K[X_1, \dots, X_n]/\mathfrak{a}$ como anillo residual. En consecuencia, $\exp_{\leq}(\mathfrak{a})^c$ es un conjunto finito (biyectable a $\mathcal{B}(H)/\bar{\mathfrak{a}}$ para cualquier orden monomial \leq). Queda por ver que $\exp_{\leq}(\mathfrak{a})^c$ es cerrado hacia abajo. Pero eso es consecuencia casi inmediata de combinar la División de Weierstarss-Hironaka (Teorema A.2.7) con el Teorema 1.2.7 sobre existencia de Bases de Gröbner. Por existir una base de Gröbner del ideal \mathfrak{a} , entonces existen $\{\mu_1, \dots, \mu_s\}$, todos ellos distintos dos a dos y tales que

$$\exp_{\leq}(\mathfrak{a}) = \bigcup_{i=1}^s (\mu_i + \mathbb{N}^n).$$

Consideremos ahora el conjunto complementario

$$\exp_{\leq}(\mathfrak{a})^c = \bigcap_{i=1}^s (\mu_i + \mathbb{N}^n)^c.$$

Veamos que ese conjunto así descrito es cerrado hacia abajo, lo cual es del todo evidente por su forma. Para ello, sean γ, θ tales que $\theta \in \exp_{\leq}(\mathfrak{a})^c$ y $\gamma \preceq \theta$. Entonces, si $\gamma \in \exp_{\leq}(\mathfrak{a})$, como $\gamma \preceq \theta$ concluiríamos que $\mu_i \preceq \gamma \preceq \theta \in \mu_i + \mathbb{N}^n$ para algún $i \in [s]$. Y esto último es imposible por construcción. Por tanto, $\gamma \notin \exp_{\leq}(\mathfrak{a})$ y tendremos que $\exp_{\leq}(\mathfrak{a})^c$ es cerrado hacia abajo y finito (i.e. una escalera finita). \square

En conclusión, dado cualquier conjunto finito cerrado hacia abajo $H \subseteq \mathcal{Y}^n$ (escalera finita) existe un ideal tal que para cualquier orden monomial verifica $H := \mathbb{N}^n \setminus \exp_{\leq}(\mathfrak{a})$ y ese ideal se puede elegir que sea un ideal monomial.

Bibliografía

- [Aln, 1983] N. Alon, *On the density of sets of vectors*. Discrete Math. **46** (1983), 199-202.
- [Artin, 1927] E. ARTIN, *Zur Theorie der hyperkomplexen Zahlen*. Abhandlungen math. seminar Hamburg **5** (1927), 251-260.
- [Buc, 1965] B. Buchberger, *“Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal”*. Tesis Doctoral, Universidad Leopold-Franzens, Innsbruck, 1965.
- [Buc, 1985] B. Buchberger, *Gröbner Basis : An Algorithmic Method in Polynomial Ideal theory*. In “Multi-dimensional Systems Theory” , Mathematics and its Applications **6**, Reidel, 1985, 184-232.
- [BCDMY, 22] N. Brukhim, D. Carmon, I. Dinur, S. Moran, A. Yehudayoff, *A Characterization of Multiclass Learnability*. Electronic Colloquium on Computational Complexity, TR22-035, 2022.
- [CLO, 07] D.A. Cox, J. Little, D. O’Shea, *“Ideals, Varieties, and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra”*. Springer, 2nd Edition, 2007.
- [DS, 14] A. Daniely, S. Shalev-Shwartz, *Optimal Learners for Multiclass Problems*. In Proc. COLT’2014, J. Machine Learning Research, **35** (2014), 1-30.
- [Dix, 1913] L.E. Dickson, *Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors*. American Journal of Mathematics **35** (1913), 413-422.
- [FRR, 06] B. Felszeghy, B. Ráth, L. Rónyai, *The lex game and some applications*. J. of Symb. Comput. **41** (1006), 663-681.
- [Fra, 1983] P. Frankl, *On the Trace of Finite Sets*. J. of Combinatorial Theory, Series A, **34** (1983), 41-45.
- [Hau, 1995] D. Haussler, *Sphere packing numbers for subsets of the Boolean n -cube with bounded Vapnik-Chervonenkis dimension*. J. Comb. Theory, Ser. A **69** (1995), 217-232.
- [HLW, 1994] D. Haussler, N. Littlestone, M.K. Warmuth, *Predicting $\{0, 1\}$ functions on randomly drawn points*. Information and Computation, **115** (1994), 284-293.
- [HHPS, 20] A. Hashemi, J. Heintz, L. M. Pardo, P. Solernó, *Intrinsic complexity for constructing zero-dimensional Gröbner bases*. In “Computer Algebra and Scientific Computing” , Lecture Notes in Computer Science **12291**, Springer, 2020, 245-265.
- [HHPS, 21] A. Hashemi, J. Heintz, L.M. Pardo, P. Solernó, *On Bézout inequalities for non-homogeneous polynomial ideals*. J. of Symb- Comput. **106** (2021), 1-22.
- [Hir, 1964a] H. Hironaka, *Resolution of Singularities of an Algebraic Variety over a Field of Characteristic Zero I*. Annals of Math. **79** (1964), 109-203.
- [Hir, 1964b] ———, *Resolution of Singularities of an Algebraic Variety over a Field of Characteristic Zero II*. Annals of Math. **79** (1964), 205-326.
- [Mes, 20] T. Mészáros, *Standard monomials and extremal point sets*. Discrete Mathematics **343** (2020), 111785.
- [Mor, 16] T. Mora, *“Solving Polynomial Equation Systems IV: Volume 4, Buchberger Theory and Beyond”*. Encyclopedia of Mathematics and its Applications **158**, Cambridge University Press, 2016.
- [Nat, 1988] B. K. Natarajan, *Some results on learning*. Unpublished manuscript, 1988.
- [Nat, 1989] B. K. Natarajan, *On learning sets and functions*. Machine Learning **4** (1989), 67-97.
- [Par, 24a] L.M. Pardo, *“Draconis Venationes Enhciridion, Pras Prima”*. Manuscrito en permanente preparación, 2024.
- [Par, 24b] L.M. Pardo, *Exploring implications of Trace (Inversion) formula and Artin algebras in extremal combinatorics*. Applic. Algebra in Eng., Comm. and Comput., to appear (2024).
- [PSZ, 24] L.M.Pardo, D. Sebastián, J. Zhang, *Some combinatorial aspects of the one-inclusion graph if multi-class learning: average out-degree, pseudodimension DS and its relations with other known invariants*. Manuscrito en preparación, 2024.
- [RBR, 06] B.I. Rubinstein, P.L. Bartlett, J. Hyam Rubinstein. *Shifting, one-inclusion mistake bounds and tight multiclass expected risk bounds*. In “Advances in Neural Information Processing Systems **19**, Proc. of the 2006 Conference”, MIT Press 2006, 1193-1200.
- [Stur, 1993] B. STURMFELS, *Sparse elimination Theory*. In D. Eisenbud, L. Robbiano, eds., “Proc. of Computational Algebraic Geometry and Commutative Algebra, Cortona, 1991”, Cambridge University Press, 1993, 264-298.
- [Val, 1984] L.G. Valiant, *A theory of the learnable*. Communications of the ACM **27** (1984), 1134-1142
- [VC, 1971] V.N. Vapnik, A.Ya. Chervonenkis, *On the uniform Convergence of relative Frequencies of rvents to their probabilities*. Theory Probab. Appl. **16** (1971), 264-280.
- [vdW, 1985] B.L. van der Waerden, *“The History of Algebra”*. Springer, 1985.
- [Wei, 1854] K. Weierstrass, *“Zur Theorie der Abelschen Functionen”*. J. für Reine und Angew. Math. (1854).
- [Wei, 1876] K. Weierstrass, *“Zur Theorie der eindeutigen analytischen Functionen”*. Berl. Abh. (1876), 11-60.

Glosario de Términos

- i -vecino, 25
- DS-desmenuzado, 27
- arista unaria, 12
- aristas del OIG, 11
- aristas no unarias, conjunto de, 15
- Artin, anillo de, 7, 59
- base monomial, 4, 54
- base monomial de cociente $K[X_1, \dots, X_n]/\mathfrak{a}$, 7, 57
- cero-dimensional, anillo, 7, 59
- cerrado hacia abajo, conjunto, 2
- coeficiente, 4, 55
- condición de cadena numerable descendente de ideales, 7, 59
- cubo, 25
- denso, polinomio, 5, 55
- desmenuzado “à la Natarajan”, 28
- Dickson, pre-orden, 1, 51
- dimensión DS, 27
- dimensión de Natarajan, 28
- dimensión exponencial, 27
- distancia de Hamming, 12
- distancia de Haussler, 12
- escalera, 2, 52
- escalón, 2
- exponente de un ideal, 6, 56
- exponente director de un polinomio, 5, 55
- exponente monomial del término, 4, 54
- Gröbner, base de un ideal, 6, 57
- grado de h en $G(H)$, 13
- grado de salida (de una orientación), 13
- grado de salida promedio (de una orientación), 13
- grado de salida promedio “maximal”, ix, 49
- grado de un monomio, 54
- grado promedio de un OIG, 13
- grado, de un polinomio, 5, 55
- grafo de una inclusión, 12
- hiper-superficie afín, 8, 60
- orden monomial, 3, 53
- orientación en $G(H)$, 13
- poli-cubo, 25
- pseudo-cubo, 25
- pseudo-cubo de dimensión n , 25
- pseudo-dimensión DS, 29
- puntos L -racionales, 8, 60
- shift dado por una palabra $\omega \in [n]^*$, 36
- shift de H en la dirección de π_i , 35
- shift de una arista, 35
- soporte, 4
- término, 4, 55
- término director, 5, 55
- variedad algebraica K -definible, 8, 60
- variedad de puntos L -racionales, 8, 60
- $V_{\mathbb{A}^n(\mathbb{K})}(f_1, \dots, f_s)$, 8, 60

Glosario de teoremas y Resultados

- acotación del $\text{outdeg}_{\max}(\sigma)$, teorema de, 49
- Akizuki, teorema de, 8, 59
- Bases monomiales (para $\leq_{1\text{ex}}$) de $K[H]$ es igual a $\mathbb{S}_{n \dots 1}(H)$, teorema, 41
- Basissatz, para $K[X_1, \dots, X_n]$, 6, 57
- caracterización de $\text{outdeg}_{\text{av}}(\sigma) = \text{outdeg}_{\text{av}}(H)$, teorema de, 14
- caracterización de pseudo-cubos cerrados hacia abajo, corolario de, 31
- caracterización de pseudo-cubos, teorema de, 26
- Dickson, Lema, 2, 51
- $\dim_{\text{DS}}(H) < 2 \dim_{\mathbb{E}}(H)$, corolario destacable, 48
- $\dim_{\text{DS}}(H) \leq 2 \text{outdeg}_{\max}(\sigma)$ (Lema 12 de [BCDMY, 22]), corolario, 28
- $\dim_{\mathbb{N}}(H) \leq \dim_{\mathbb{E}}(H)$, corolario, 29
- dimensión DS es un máximo de la pseudo-dimensión DS, corolario, 30
- dimensión exponencial no crece tras un shift, proposición, 42
- eliminación cuantificada, Lema de, 38
- escalera finita H es cerrado hacia abajo sii $H = \text{Nexp}_{1\text{ex}}(I(H))$, corolario destacable, 42
- escaleras finitas= bases monomiales de K -álgebras de Artin, corolario destacable, 9, 61
- existencia de Bases de Gröbner, teorema de, 6, 57
- existencia de factores triangulares, Lema de, 39
- Felszeghy-Ráth-Rónyai (nueva demostración), teorema de, 40
- grado de salida en promedio es estrictamente menor que la dimensión exponencial, teorema, 48
- grado de salida promedio no decrece tras varios shifts, teorema, 46
- grado promedio no decrece en el caso $p = 2$, corolario, 47
- mejora de un resultado de [BCDMY, 22], proposición, 33
- “promedios” no son funciones con respecto a la inclusión conjuntista, proposición, 19
- propiedades básicas de la $\dim_{\text{DS}}(H)$, Lema de, 21
- propiedades de \mathcal{E}_h (conjuntamente con Proposición 2.3.2), proposición, 30
- pseudo-dimensión DS es máximo de ψ_H , proposición, 29
- Weierstrass-Hironaka, división, 6, 56

Glosario de Símbolos y Abreviaturas

<p> $\text{avd}(H)$, 13 CCD, 7, 59 $\text{deg}_H(h)$, 13 $\text{deg}_{\min}(H)$, 48 $\text{dim}_{\text{DS}}(H)$, 27 $\text{dim}_{\mathbb{E}}(H)$, 27 $\text{dim}_{\mathbb{N}}(H)$, 28 $p \text{dim}_{\text{DS}}(H)$, 29 $E(H)_h$, 13 $E^{(k)}(H)$, 13 $E(H)$, 11 $E^{(1)}(H)$, 12 $e_{f,i}$, 11 \mathcal{E}_h, 2, 51 $\text{exp}_{1\text{ex}}(\mathbf{a})$, 38 $G(H)$, 12 $\max\text{-outdeg}_{\text{av}}(H)$, 49 $I(V)$, 37 $I_K(V)$, 37 $K[V]$, 37 $\ell t_{\leq}(f)$, 5, 55 $\max\text{-avd}(H)$, 48 $\max\text{-outdeg}_{\text{av}}(H)$, 49 $\mathcal{M}(X_1, \dots, X_n)$, 4, 54 </p>	<p> $[n]$, i $\text{Nexp}_{1\text{ex}}(\mathbf{a})$, 38 $E_{>}^{(1)}(H)$, 15 $\leq_{1\text{ex}}$, 3, 53 $\leq_{\text{deg}1\text{ex}}$, 53 $\text{outdeg}_H(\sigma)$, 13 π_i, 11 \preceq, 1, 51 $\rho(f, g)$, 12 $\#(X)$, i $\widetilde{\mathbb{S}}_{f,i}$, 35 $\widetilde{\mathbb{S}}_{H,\omega}$, 37 $\widetilde{\mathbb{S}}_{H,i}$, 36 \mathbb{S}_{ω}, 36 $\mathbb{S}_i(H)$, 35 $\text{supp}(f)$, 4, 54 $V_{\mathbb{A}^n(\mathbb{K})}(f)$, 8, 60 $V_{\mathbb{A}^n(L)}(\mathbf{a})$, 8, 60 $V_{\mathbb{A}^n(\mathbb{K})}(\mathbf{a})$, 8, 60 $\mathcal{Y}^{[n]}$, i $\mathcal{Y}^{\mathcal{X}}$, i </p>
--	---