Nuevos panoramas en teoría de la probabilidad y ciencia de datos.
(borrador).

You

kevincalderon@ciencias.unam.mx

Resumen

Se dará una breve exposición de la estadística algebraica y la teoría de la probabilidad homotopica , la primera es un área que redefine los métodos de la estadística usando geometría algebraica , combinatoria algebraica y álgebra conmutativa (permitiendo que la estadística pueda considerarse como matemáticas) , la segunda es una nueva redefinicion de la teoría de la probabilidad vía teoría de homotopias. Este texto supondrá que el lector ya tiene buen conocimiento de estadística y probabilidad.

Índice

Advertencia al lector :

Este texto es solo una mención (se debe de pensar a este articulo como una platica de café o de pasillo , tal y como si yo lo estuviese contando tal y como descubrí estos temas) a la estadística algebraica y la teoría de la probabilidad homotopica , debido a que son áreas muy recientes se carece del suficiente material para armar una exposición detallada (que sea breve a diferencia de las monografías que hay de los temas) , y sobre todo debido a que ocupan herramientas que se estudian cuando uno entra de lleno a las matemáticas puras , complican mucho hacer la exposición accesible. Otro punto es que este este texto es un borrador de un futuro articulo de la misma temática y contenidos pero de mejor calidad. Existen otras mas teorías interesantes que se basan de unir áreas de alto calibre de las matemáticas puras o física teórica con estadística y probabilidad pero decidí tratar con estadística algebraica ya que es mi principal área de interés , y además me gusta la teoría de homotopias. Por ultimo quiero recalcar lo siguiente , soy :

1. 100 % Matemático.

2. 1 % Físico.

3. 0 % Científico de la computación.

4. 0 % Científico de datos.

Por eso no se debe de esperar que el fuerte de este texto sea la parte estadística. Recomiendo a las Científicos de datos (y sus inferiores los estadísticos) , físicos y actuarios darle una oportunidad a estos tópicos ya que muchas de las quejas de la probabilidad y la estadística se esfuman usando estos métodos menos tediosos y mas rápidos. Si se tienen sugerencias o alguna pregunta/información importante son libres de escribirme.

Considero importante el proceso de abstraer , formalizar , generalizar ideas , al fin y al cabo esto nos ayuda a simplificar procesos , técnicas y métodos , áreas como el calculo nacieron bajo esa misma premisa.

Aunque no lo incluyo aquí recomiendo al lector entusiasta leer [1] , [2] y principalmente [3] .

1. Preliminares.

En esta sección se dará un rápido resumen a los conocimientos y resultados que se ocuparan durante la exposición (tal y como son) y solo los que se ven. Se puede leer esta sección a modo de guía para poder saber que temas estudiar a profundidad para un eventual estudio mas serio de los temas presentados en el articulo. Debido a que la demostraciones de muchas proposiciones en el texto quedan fuera del alcance de este se le dejara como ejercicio al lector dicha labor.

1.1. Álgebra .

1.1.1. Categorías.

Aunque no sea parte del álgebra , se mostrara en esta sección por que son resultados e ideas que se ocuparan en esta sección.

Una categoría $\mathcal{C}$ esta formada por una clase de objetos $\mathrm{Obj}(\mathcal{C}) = X, Y, Z, ...$ y una de morfismos $\mathrm{Mor} = f, g, h, ...$ tales que cada morfismo tiene un único dominio y codominio en los objetos , $f : X \to Y$ y $g : Y \to Z$ entonces existe la composición de morfismos $g \circ f : X \to Z$. Para cada objeto X existe un morfismo identidad. Los morfismos son asociativos.

Un functor $F : \mathcal{C} \to \mathcal{D}$ es un par de funciones $F : \mathrm{Obj}(\mathcal{C}) \to \mathrm{Obj}(\mathcal{D})$ y $F : \mathrm{Mor}(\mathcal{C}) \to \mathrm{Mor}(\mathcal{D})$. Los functores preservan la identidad y composición. Diremos que un functor es contravariante si $F : \mathrm{Hom}_{\mathcal{C}}(A, B) \to \mathrm{Hom}_{\mathcal{D}}(F(B), F(A))$ y es covariante si $F : \mathrm{Hom}_{\mathcal{C}}(A, B) \to \mathrm{Hom}_{\mathcal{D}}(F(A), F(B))$.

Dado un morfismo $f : X \to Y$ y otro $g_Z \to Y$, un levantamiento de f a Z es un morfismo $h : X \to Z$ tal que $f = g \circ h$.

1.1.2. Álgebra conmutativa.

Definición 0 : Un conjunto R junto con dos leyes de composición $+$ y $\cdot$ lo llamaremos anillo (con unidad) , si satisface las siguientes propiedades :

1. R es un grupo abeliano[1] con respecto a $+$. El correspondiente elemento cero es denotado por $0 \in R$.

2. $\cdot$ es asociativa i.e $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

3. Existe un elemento unidad en R , el cual significa un elemento $1 \in R$ tal que $1 \cdot a = a = a \cdot 1$ para todo $a \in R$.

4. La multiplicación es distributiva sobre $+$ i.e para $a, b, c \in R$ tenemos : $a \cdot (b + c) = a \cdot b + a \cdot c$

[1]Dado un conjunto G con una operación que es asociativa , conmutativa , distributiva y con elemento neutro e inverso , es un grupo abeliano.

, $(a+b)\cdot c = a\cdot c + b\cdot c$.

El anillo R lo llamaremos conmutativo si $\cdot$ es conmutativo.

Definición 1 : Sea $a \in R$ es llamado invertible o unidad si existe algún elemento $b \in R$ tal que $ab = 1 = ba$. Se sigue que el conjunto

$$R^* = \{a \in R : a \text{ es una unidad en } R\}.$$

sea un grupo con respecto a $\cdot$ dada en R.

Un elemento $a \in R$ es llamado divisor cero si existe un elemento $b \in R - \{0\}$ tal que $ab = 0$ o $ba = 0$. Además un anillo conmutativo $R \neq 0$[2] es llamado dominio entero si este no contiene divisores cero (no triviales) i.e si $ab = 0$ con $a, b \in R$ implica $a = 0$ o $b = 0$.

Definición 2 : Una función $\phi : R \to R'$ entre anillos es llamada homomorfismo de anillos[3] si para todo $a, b \in R$ las siguientes condiciones se satisfacen :

1. $\varphi(a+b) = \varphi(a) + \varphi(b)$.
2. $\varphi(a\cdot b) = \varphi(a)\cdot\varphi(b)$.
3. $\varphi(1) = 1$.

Los mono- , epi- , iso- , endo- y automorfismos de anillos los podemos definir de manera usual.

Sea $R' \subset R$, lo llamaremos un subanillo de R si $1 \in R'$ y si $a, b \in R' \Rightarrow a - b, a\cdot b \in R'$. Entonces R' es un anillo bajo $+$ y $\cdot$ heredadas de R. Cuando estemos dentro del álgebra conmutativa solo vamos a considerar anillos que sean conmutativos con unidad , por lo que usaremos el termino anillo para referirnos a este tipo de anillo (a menos que se mencione lo contrario).

Definición 3 : Sea R un anillo. Un subconjunto $\mathfrak{a} \subset R$ es llamando un ideal en R si :

1. $\mathfrak{a}$ es un subgrupo aditivo de R i.e $\mathfrak{a} \neq \varnothing$ y $a - b \in \mathfrak{a}$ para todo $a, b \in \mathfrak{a}$.
2. $ra \in \mathfrak{a}$ para todo $r \in R$ y $a \in \mathfrak{a}$.

Todo anillo R contiene ideales triviales , los cuales son el ideal cero 0 (consiste solo del elemento 0) y si mismo (también llamado ideal unitario). Para una familia $\{a_i\}_{i\in I}$ de elementos en R , podemos construir su ideal generado

$$\mathfrak{a} = \sum_{i\in I} Ra_i = \left\{\sum_{i\in I} r_i a_i : r_i \in R, r_i = 0 \text{ para casi todo } i \in I\right\}.$$

Este es el ideal mas pequeño en R que contiene a todos los elementos $a_i, i \in I$. Si consideramos el índice I finito podemos escribir :

$$\mathfrak{a} = \sum_{i=0}^{n} Ra_i = (a_0, ..., a_n).$$

Además , un ideal $\mathfrak{a} \subset R$ es llamado finitamente generado si como antes , puede ser generado apartir de una cantidad finita de elementos (a este le llamamos generador). Un ideal $\mathfrak{a} \subset R$ se llama principal si es generado por un solo elemento , entonces si hay un elemento $a \in \mathfrak{a}$ tal que $\mathfrak{a} = (a)$ diremos que es principal. Todo dominio entero R se llama dominio ideal principal si todo ideal de R es principal.

Veamos otras formas de construir ideales mediante una familia de ideales $\{\mathfrak{a}_i\}_{i\in I}$:

- $\sum_{i\in I} \mathfrak{a}_i = \left\{\sum_{i\in I} a_i : a_i \in \mathfrak{a}_i, a_i = 0 \text{ para casi todo } i \in I\right\}$ es un ideal en R , y es llamado la suma de los ideales $\mathfrak{a}_i$.

[2] mejor conocido como anillo trivial

[3] o morfismo de anillos.

- $\bigcap_{i\in I} \mathfrak{a}_i$ es un ideal en R.
- Si tenemos una coleccion finita de ideales $\mathfrak{a}_0, ..., \mathfrak{a}_n \subset R$ podemos construir su producto $\prod_{i=0}^{n} \mathfrak{a}_i$ el cual nuevamente es un ideal en R.

Proposición 0 : Sea $\varphi : R \to R'$ un morfismo de anillos. Entonces

$$\mathrm{Ker}(\varphi) = \{r \in R : \varphi(r) = 0\}$$

es un ideal en R y $\mathrm{im}(\varphi) = \varphi(R)$ es un subanillo de R'.

Dado un ideal $\mathfrak{a}$ en un anillo R , si queremos construir un homomorfismo suryectivo de anillos $\pi : R \to R'$ tal que $\mathrm{Ker}(\pi) = \mathfrak{a}$. Para hacer esto , consideremos el conjunto $R/\mathfrak{a} = \{r + \mathfrak{a} : r \in R\}$ [4] donde $r + \mathfrak{a} = \{r + a : a \in \mathfrak{a}\}$ para todos las clases (aditivas) de $\mathfrak{a}$ en R y definimos las dos leyes de composición en ellas por : $(r+\mathfrak{a})+(r'+\mathfrak{a}) := (r+r')+\mathfrak{a}$ y $(r+\mathfrak{a})\cdot(r'+\mathfrak{a}) = (r\cdot r')(\mathfrak{a})$. Veamos que $\pi_R \to R/\mathfrak{a}$, $r \mapsto r + \mathfrak{a}$, es un homomorfismo suryectivo de anillos que satisface $\mathrm{Ker}(\varphi) = \mathfrak{a}$ como buscabamos. Podemos caracterizar al anillo $R/\mathfrak{a}$ con el siguiente teorema :

Teorema fundamental de los homomorfismo : Sea $\varphi : R \to R'$ un morfismo de anillos y $\mathfrak{a} \subset R$ un ideal que satisface $\mathfrak{a} \subset \mathrm{Ker}(\varphi)$. Entonces existe un único homorfismo de anillos $\bar{\varphi} : R/\mathfrak{a} \to R'$ tal que el diagrama

$$\begin{tikzcd} R \arrow[r] \arrow[d, "\pi"'] & R' \\ R/\mathfrak{a} \arrow[ur, dotted, "\overline{\varphi}"'] & \end{tikzcd}$$

es conmutativo. Además :

1. $\bar{\varphi}$ inyectivo $\Leftrightarrow \mathfrak{a} = \mathrm{Ker}(\varphi)$.
2. $\bar{\varphi}$ suryectivo $\Leftrightarrow \varphi$ suryectivo.

Definición : Sea R un anillo.

1. Un ideal $\mathfrak{p} \subset R$ es llamado primo si $ab \in \mathfrak{p}$ implica $a \in \mathfrak{p}$ o $b \in \mathfrak{p}$, para cualquier $a, b \in R$.
2. Un ideal $\mathfrak{m} \subset R$ es llamado maximal si $\mathfrak{m} \subset \mathfrak{a}$ implica $\mathfrak{m} = \mathfrak{a}$, para cualquier ideal propio $\mathfrak{a} \subset R$.

Proposición : Cualquier anillo $R \neq 0$ contiene un ideal maximal.

Corolario : Sea R un anillo y $\mathfrak{a} \subset R$ un ideal propio en R. Entonces existe un ideal maximal $\mathfrak{m} \subset R$ tal que $\mathfrak{a} \subset \mathfrak{m}$.

Corolario : Sea R un anillo y $a \in R$ un anillo no unitario. Entonces existe un ideal $\mathfrak{m} \subset R$ tal que $a \in \mathfrak{a}$.

Para un anillo R el conjunto

$$\mathrm{Spec}(R) = \{\mathfrak{p} \subset R : \mathfrak{p} \text{ es un ideal primo en } R\}.$$

lo llamaremos el espectro primo de R. De igual manera el subconjunto

$$\mathrm{Spm}(R) = \{\mathfrak{m} \subset R : \mathfrak{m} \text{ es un ideal maximal en } R\}.$$

es llamado el espectro maximal de R.

[4]anillo cociente

Definición : Sea F un anillo en el cual todo elemento no cero tiene inverso bajo $\cdot$ lo llamaremos campo.

Definición : Un monomio de variables $x_1, ..., x_n$ es el producto $x^a = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ para un vector $a = (a_1, ..., a_n) \in \mathbb{N}^n$. Y un polinomio variables $x_1, .., x_n$ sobre un campo F es una combinación lineal finita de varios monomios , con coeficientes en F. O sea $f = \sum_{a \in A} c_a p^a$ donde A es un subconjunto finito de $\mathbb{N}^r$y cada coeficiente $c_a \in F$.

Sea $\mathbb{K}[x_1, .., x_n]$ el conjunto de todos los polinomios de variables $x_1, ..., x_n$ con coeficientes en $\mathbb{K}$. Este conjunto forma un anillo (considerando la suma y producto de polinomios usual). Cada polinomio puede ser pensado como una función multivaluada $\mathbb{K}^r \to \mathbb{K}$ por evaluación.

Recomiendo al lector proceder a leer la sección de Geometría Algebraica.[5]

Como espacio vectorial sobre $\mathbb{K}$, el anillo de polinomios $\mathbb{K}[x_1, ..., x_n]$ es una suma directa

$$\mathbb{K}[x_1, .., x_n] = \bigoplus_{a \in \mathbb{N}^n} \mathbb{K}_a$$

donde $\mathbb{K}_a = \mathbb{K}(x^a)$ es el subespacio vectorial de $\mathbb{K}[x_1, ..., x_n]$ generado por el monomio x^a.

Definición : Sea $S \subset \mathbb{K}[x_1, .., x_n]$ un conjunto de polinomios. La variedad definida por S es el conjunto

$$V(S) := \{a \in \mathbb{K}^r : f(a) = 0 \text{ para todo } f \in S\}$$

La variedad $V(S)$ es llamado el conjunto de ceros de S. Hay que notar que el conjunto de ceros depende de en cual campo $\mathbb{K}$ se este trabajando , la elección del campo depende del contexto del problema que se desea trabajar. Si no trabajamos en un campo algebraicamente cerrado se corre el riesgo de tener conjuntos de ceros vacíos. Vamos a mostrar como al tomar un objeto algebraico (un conjunto de polinomios) produce un objeto geométrico (el conjunto de ceros comunes) , y como esta construcción se puede invertir : a partir de un objeto geométrico (un subconjunto de $\mathbb{K}^r$) podemos construir el conjunto de polinomios que se anulan en el. Justamente en la interacción de ideales y variedades esta el nucleo de la geometría algebraica.

Definición : Sea $W \subset \mathbb{K}^r$. El conjunto de polinomios $I(W) = \{f \in \mathbb{K}[x_1, .., x_n] : f(a) = 0 \text{ para todo } \ a \in W\}$ es el ideal anulador de W.

Teorema de la base de Hilbert : Para todo ideal I en un anillo de polinomios $\mathbb{K}[x_1, .., x_n]$ debe de existir un conjunto de polinomios $\mathcal{F} \subset I$ tal que $I = (\mathcal{F})$ i.e todo ideal en $\mathbb{K}[x_1, ..., x_n]$ es finitamente generado.

De este teorema y de otras observaciones nos permiten concluir que toda variedad puede ser descrita como el conjunto común de ceros de un conjunto finito de polinomios.

Definición : Un ideal I es llamado radical si $f^k \in I$ para algún polinomio f y entero k implica que $f \in I$. El radical de I , denotado $\sqrt{I}$, es el ideal radical mas pequeño que contiene a I y consiste de los polinomios :

$$\sqrt{I} = \left\{f \in \mathbb{K}[x_1, .., x_n] : f^k \in I \text{ para algun } k \in \mathbb{N}\right\}$$

De la siguiente proposición se entiende la importancia de los ideales radicales en geometría algebraica.

Proposición : Para cualquier campo y cualquier conjunto $W \subset \mathbb{K}^r$, el ideal $I(W)$ es un radical ideal.

El siguiente teorema es el que nos podrá dar la conexión álgebra-geometría :

[5]Lo siguiente aunque es un poco de geometría algebraica clásica , esta mas enfocado a bases de Gröbner.

Nullstellensatz : Sea $\mathbb{K}$ un campo algebraicamente cerrado. Entonces el ideal anulador de la variedad de un ideal es el radical del ideal : $I(V(I)) = \sqrt{I}$.

Proposición : El conjunto $V(I(W))$ es llamado clausura de Zariski de W. Es la variedad algebraica mas pequeña que contiene a W.

Correspondencia ideal-variedad : Sea $\mathbb{K}$ un campo algebraicamente cerrado. Entonces las funciones :

$$V : \left\{ I \subset \mathbb{K}[x_1, .., x_n] : I = \sqrt{I} \right\} \to \{W \subset \mathbb{K}^r : W = V(J) \text{ para algun } J \subset \mathbb{K}[x_1, ..., x_n]\}$$

$$I : \{W \subset \mathbb{K}^r : W = V(J) \text{ para algun } J \subset \mathbb{K}[x_1, ..., x_n]\} \to \left\{ I \subset \mathbb{K}[x_1, .., x_n] : I = \sqrt{I} \right\}$$

son biyecciones que invierten la inclusión entre el conjunto de ideales radicales y el conjunto de variedades.

Notese que forzosamente $\mathbb{K}$ debe ser cerrado algebraicamente. El anterior teorema nos sirve como un traductor entre el álgebra y la geometría :

Álgebra	Geometría.
Ideales maximales de $\mathbb{K}[x_1, ..., x_n]$	Puntos de $\mathbb{K}^n$
Ideales primos de $\mathbb{K}[x_1, ..., x_n]$	Subconjuntos irreducibles de $\mathbb{K}^n$

Proposición : Sean I y J ideales de $\mathbb{K}[x_1, ..., x_n]$. Entonces

$$V(I + J) = V(I) \cap V(J).$$

Si $\mathbb{K}$ es algebraicamente cerrado y V, W subconjuntos arbitarios de $\mathbb{K}^r$,entonces

$$I(V \cap W) = \sqrt{I(V) + I(W)}.$$

Proposición : Sean I y J ideales de $\mathbb{K}[x_1, ..., x_n]$. Entonces

$$V(I \cap J) = V(I) \cup V(J).$$

Si $\mathbb{K}$ es algebraicamente cerrado y V, W subconjuntos arbitarios de $\mathbb{K}^r$,entonces

$$I(V \cup W) = I(V) \cap I(W).$$

Las bases de Gröbner es una herramienta que nos permite estudiar ideales de manera computacional , es una pieza clave de la maquinaria de la estadística algebraica. El resto de la sección estara dedicada a dicho topico.

Definición : Un orden de termino $\prec$ es un orden total en el conjunto de monomios en el anillo de polinomios $\mathbb{K}[x_1, ..., x_n]$ el cual satisface dos condiciones :

1. $1 = p^0 \preceq p^u$ para todo $u \in \mathbb{K}^r$.
2. $p^u \prec p^v$ implica que $p^w \cdot p^u \prec p^w \cdot p^v$ para todo $w \in \mathbb{N}^r$.

Definición : Dado un orden de termino $\prec$ y un polinomio $f \in \mathbb{K}[x_1, ..., x_n]$, el monomio inicial , o termino inicial de f con respecto a $\prec$ es el monomio p^u mas grande de f con respecto al orden $\prec$ entre todos los monomios que aparecen en f con coeficiente no cero. Al monomio inicial lo denotaremos con $\text{in}_\prec(f)$.

Definición : El ideal inicial $in_\prec(f)$ de un ideal $I \subseteq \mathbb{K}[x_1, ..., x_n]$ con respecto al orden de termino $\prec$ es el ideal generado por los monomios iniciales de todos los polinomios en I : $in_\prec(I) = (in_\prec(f) : f \in I)$.

Definición : Una base de Gröbner de el ideal I con respecto al orden de termino $\prec$ es una colección finita de polinomios $\mathcal{G} \subset I$ tal que el monomio inicial de los elementos de $\mathcal{G}$ genere el ideal inicial :

$$(\text{in}_\prec(g) : g \in \mathcal{G}) = \text{in}_\prec(I).$$

Algoritmo :

- **Entrada :** Sea $\mathcal{G} = \{g_1, .., g_k\}$ un conjunto finito de polinomios , f otro polinomio , y $\prec$ un orden de termino.
- **Salida :** Una representación $f = \sum_{i=1}^{k} h_i g_i + r$ tal que ningún termino de r es divisible por un termino inicial de cualquiera de los polinomios en $\mathcal{G}$

  1. Establezca $h_i = 0$ para todo i y $r = f$.
  2. Mientras r tenga un termino $c_a p^a$ divisible por un termino inicial de algún g_i , remplace h_i por $h_i + \frac{c_a p^a}{in_\prec(g_i))}$ y r por $r - \frac{c_a p^a}{in_\prec(g_i))} g_i$.

El residuo r al dividir f por $\mathcal{G}$ se denomina forma normal de f y se denota por $NF_{\mathcal{G}}(f)$. Por supuesto , se debe tener cuidado con esta definición porque la forma normal no esta determinada únicamente por $\mathcal{G}$ y f.

Teorema : Sea I un ideal , $\prec$ un orden de termino , y $\mathcal{G}$ un subconjunto finito de I. Entonces $NF_{\mathcal{G}} = 0$ para todo $f \in I$ si y solo si $\mathcal{G}$ es una base de Gröbner para I con respecto a $\prec$.

Corolario : Sea I un ideal , $\prec$ un orden de termino , y $\mathcal{G}$ una base de Gröbner para I con respecto de $\prec$. Entonces $\mathcal{G}$ es un generador de I.

Definimos el $S-$polinomio de f y g como

$$S(f,g) = \frac{p^v}{\text{gdc}(p^u, p^v)} f - \frac{p^u}{\text{gcd}(p^u, p^v)} g$$

El siguiente es un teorema que nos permite caracterizar a las bases de Gröbner :

Criterio de Buchberger : Sea I un ideal , $\prec$ un orden de termino y $\mathcal{G}$ un conjunto finito generador de I. Entonces $\mathcal{G}$ es una base de Gröbner de I con respecto de $\prec$ si y solo si $NF_{\mathcal{G}}(S(f,g)) = 0$ para todo $f, g \in \mathcal{G}$.

El criterio de Buchberger conduce a un sencillo algoritmo para calcular las bases ideales de Gröbner a partir de sus conjuntos generadores.

Algoritmo :

- **Entrada :** Un conjunto finitamente generado $\mathcal{F}$ de I y orden de termino $\prec$.
- **Salida :** Una base de Gröbner $\mathcal{G}$ para I con respecto a $\prec$.

  1. Fijamos $\mathcal{G} := \mathcal{F}$.
  2. Mientras $\mathcal{G}$ no cumpla el criterio de Buchberger , hacer :
     a) Existe un par $f, g \in \mathcal{G}$ tal que $NF_{\mathcal{G}}(S(f,g)) = h \neq 0$.
     b) Fijar $\mathcal{G} := \mathcal{G} \cup \{h\}$.
  3. **Salida :** Una base de Gröbner $\mathcal{G}$ para I con respecto a $\prec$.

Recomiendo al lector interesado , leer acerca de Ideales de Stanley-Resiner y combinatoria algebraica. Se puede estudiar mas de bases de Gröbner en [4]

Definición : Sea R un anillo , un R-modulo consiste de un conjunto M junto con una ley de composición interna $M \times M \to M$, $(a, b) \mapsto a + b$ llamada suma y una ley de composición externa $R \times M \to M$, $(\alpha, a) \mapsto a \cdot a$ llamada multiplicación por un escalar , tal que :

1. Mes un grupo abeliano respecto a la suma.
2. $\cdot$ es distributiva respecto de $+$.
3. La multiplicación por escalar es asociativa.
4. Existe un elemento unitario en $1 \in R$ tal que $1 \cdot a = a$ para $a \in M$.

Un modelo es la generalización de un espacio vectorial , si $R = K$ donde K es un campo recuperamos un espacio vectorial. Definimos los morfismos de R-módulos de manera usual.

Definición : El producto tensorial $M \otimes_R N$ es un R-modulo equipado con una función bilineal $M \times N \to^{\oplus} M \times_R N$ tal que cada función bilineal $M \times N \to^{B} P$ existe una única transformacional lineal $M \otimes_R N \to^{L} P$ haciendo que el siguiente diagrama conmute :

$$\begin{tikzcd} & M \otimes_R N \arrow[d, dotted, "L"] \\ M \times N \arrow[ur, "\otimes"] \arrow[r, "B"'] & P \end{tikzcd}$$

Una secuencia de R-módulos es una cadena de morfismos de R-módulos

$$\cdots \to^{f_{n-2}} M_{n-1} \to^{f_{n-1}} M_n \to^{f_n} M_{n+1} \to^{f_{n+1}} \cdots .$$

diremos que es un complejo de cadenas si $f_n \circ f_{n-1} = 0$ o equivalentemente $\mathrm{im} f_{n-1} \subset \ker f_n$, además diremos que la secuencia es exacta en M_n si $\mathrm{im} f_{n-1} = \ker f_n$, así mismo diremos que la secuencia es exacta si lo es en todas partes. Podemos remplazar los módulos por cualquier otro objeto (mientras siga teniendo sentido).

Vale la pena señalar que independientemente del tema de este escrito es muy recomendable estudiar álgebra conmutativa , es una herramienta muy poderosa. Recomiendo revisar [5] , [6] y [7]

1.1.3. Operads.

Solo serán resultados muy selectos y autcontenidos , para hacer un resumen elemental de la teoría de operads tendría que suponer como mínimo que el lector ya sabe topología algebraica.

Definición : Sea $\mathcal{A}$ un $\mathbb{K}-$ espacio vectorial equipado con un producto bilineal ,diremos que $\mathcal{A}$ es una $\mathbb{C}$-álgebra.

Definición : Un álgebra asociativa sobre un campo $\mathbb{K}$ es un espacio vectorial equipado con la operación $\mu : A \otimes A \to A$ la cual es asociativa , o sea que se cumple el siguiente diagrama conmutativo :

$$\begin{tikzcd} A \otimes A \otimes A \arrow[r, "\mathrm{id} \otimes \mu"] \arrow[d, "\mu \otimes \mathrm{id}"'] & A \otimes A \arrow[d, "\mu"] \\ A \otimes A \arrow[r, "\mu"'] & A \end{tikzcd}$$

y es unital si existe una función $u : \mathbb{K} \to A$ que cumpla el diagrama conmutativo

$$\begin{tikzcd} \mathbb{K} \otimes A \arrow[r, "u \otimes \mathrm{id}"] \arrow[dr, "\cong"'] & A \otimes A \arrow[d, "\mu"] & A \otimes \mathbb{K} \arrow[l, "\mathrm{id} \otimes u"'] \arrow[dl, "\cong"] \\ & A & \end{tikzcd}$$

Definición : Un álgebra conmutativa es un espacio vectorial A sobre un campo $\mathbb{K}$ equipado con una operación binaria $\mu : A \otimes A \to A$, $\mu(a, b) = ab$ la cual es asociativa y conmutativa.

Definición : Un álgebra de Lie[6] es un espacio vectorial $\mathfrak{g}$ sobre un campo $\mathbb{K}$ equipado con una operación $c : \mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}$, $c(x,y) := [x,y]$ que satisface :

1. $[x,y] = -[y,x]$.
2. $[[x,y],z] = [x,[y,z]] + [[x,z],y]$ (identidad de Leibniz).
3. $[[x,y],z] + [[y,z],x] + [[z,x],y] = 0$ (identidad de Jacobi).

Definición : Un espacio vectorial graduado V es una familia de espacios vectoriales $\{V_n\}_{n\in\mathbb{Z}}$ tal que $V = \bigoplus_i V_i$.

Definición : Una coalgebra sobre $\mathbb{K}$ es un espacio vectorial C equipado con una co-operacion binaria asociativa $\Delta : C \to C \otimes C$.

Definición : El coproducto iterado $\Delta^n : C \to C^{\otimes n+1}$ esta definido como $\Delta^0 = \mathrm{id}$, $\Delta^1 = \Delta$ y $\Delta := (\Delta \otimes \mathrm{id} \otimes \cdots \otimes \mathrm{id}) \circ \Delta^{n-1}$.

Definición : Una coalgebra es coagulada si se da un morfismo de coalgebras $u : \mathbb{K} \to C$. Si C es coagulada entonces C es canonicamente isomorfo a $\ker_\epsilon \oplus \mathbb{K}1$ y el kernel lo representamos como $\bar{C}$.

Definición : El coproducto reducido $\bar{\Delta} : \bar{C} \to \bar{C} \times \bar{C}$esta dado por $\bar{\Delta}(x) =: \Delta(x) - x \otimes 1 - 1 \otimes x$.

Definición : Sea $C = \mathbb{K}1 \oplus \bar{C}$ un coalgebra coagulada , la filtración coradical en C es

1. $F_0C : \mathbb{K}1$.
2. $F_rC := \mathbb{K} \oplus \left\{x \in \bar{C} : \bar{\Delta}^n(x) = 0 \text{ para cualquier } n \geq r\right\}$ para un dado $r \geq 1$.

Definición : Una coalgebra C se dice que es conilpotente si su coagulación y su filtración cumple $C = \bigcup_r F_rC$

Definición : Sea $C = (C, \Delta)$una algebra conilpotente , una coderivacion es una transformación lineal $d : C \to C$ tal que $d(1) = 0$ y $\Delta \circ d = (d \otimes \mathrm{id}) \circ \Delta + (\mathrm{id} \otimes d) \circ \Delta$.

1.2. Geometría Algebraica.

1.2.1. Geometría algebraica clásica.

Vamos a denotar al anillo de polinomios $\mathbb{K}[x_1, ..., x_n]$ con P. Definimos el espacio afín $\mathbb{A}^n := \mathbb{K}^n$ dotado de cierta estructura topológica.

Definición : Un conjunto algebraico cerrado es un subconjunto de $\mathbb{A}^n$ de la forma $Z(S) = \{x \in \mathbb{A}^n : f(x) = 0 \text{ para todo } f \in S\}$ con $S \subseteq P$.

No todos los polinomios en S se van anular en $Z(S)$, esto depende del campo $\mathbb{K}$ en el que estemos trabajando.

Definición : Para $X \subseteq \mathbb{A}^n$definimos el ideal $I(x)$ que consiste de todos los polinomios en P que se anulan en X , esto es : $I(X) = \{f \in P : f(x) = 0 \text{para todo } x \in X\}$.

Esto nos da en automático una función inversa inclusiva del conjunto de subconjuntos del espacio afín al conjunto de ideales de P.

Nullstellensatz de Hilbert : Supongamos que $\mathbb{K}$ es algebraicamente cerrado , y $\mathfrak{a}$ un ideal en P , entonces tenemos que $I(Z(\mathfrak{a})) = \sqrt{\mathfrak{a}}$.

Este teorema nos va a permitir trabajar a placer pero solo en campos algebraicamente cerrados , en campos que no sean así la labor se complica demasiado.

[6] Como un ejemplo rapido , el producto cruz usual y $\mathbb{R}^n$ forman un algebra de Lie

Definición : Tomemos $X \subseteq \mathbb{A}^n$ un conjunto algebraico cerrado. El anillo cociente $A(x) = P/I(x)$ lo llamaremos anillo afín de coordenadas.

Podemos interpretar a los elementos en $A(x)$ como funciones polinomiales k-valuadas.

Los conjuntos algebraicos cerrados en $\mathbb{A}^n$ forman una topología la cual llamaremos topología de Zariski.

Proposición : Un conjunto algebraico $X \subseteq \mathbb{A}^n$ es irreducible si y solo si $I(x)$ es primo.

1.2.2. Geometría algebraica moderna.

Aunque los esquemas aun no han tenido un uso fuerte y grande en estadística algebraica , por mero placer y por si es en un futuro una referencia dejo un muy muy muy breve resumen de la geometría algebraica (moderna)

Definición : Sea X un conjunto. Una topología X es una familia $\mathcal{A}$ de subconjuntos en X llamados abiertos , tales que cumplen las siguientes condiciones :

1. Si $\{A_i\}_{i\in I}$, con I arbitrario , entonces $\bigcup_{i\in I} A_i \in \mathcal{A}$.
2. Si $\{A_i\}_{i\in I}$, con I finito , entonces $\cap A_i \in \mathcal{A}$.

A la pareja $(X, \mathcal{A})$ le llamaremos espacio topológico.

Definición : Sea X un espacio topológico , una pregavilla de grupos abelianos (o de cualquier otro objeto de otra categoria) $\mathcal{F}$ en X consiste en los dos siguientes grupos de información :

1. Para cada abierto $\mathcal{U} \subseteq X$, un grupo abeliano (un objeto de una categoría $\mathcal{C}$) $\mathcal{F}(\mathcal{U})$.
2. Para cada par de abiertos anidados $V \subseteq \mathcal{U}$ un homomorfismo de grupos (morfismo de elementos de la categoría del objeto que tomemos) $\rho_{\mathcal{U}\mathcal{V}} : \mathcal{F}(\mathcal{U}) \to \mathcal{F}(\mathcal{V})$.

Esta función satisface que para cualquier abierto $\mathcal{U} \subseteq X$ $\rho_{\mathcal{U}\mathcal{U}} = \mathrm{id}_{\mathcal{F}(\mathcal{U})}$ y para cualesquiera abiertos anidados $\mathcal{W} \subseteq \mathcal{V} \subseteq \mathcal{U}$ se tiene $\rho_{\mathcal{U}\mathcal{V}} = \rho_{\mathcal{V}\mathcal{W}} \circ \rho_{\mathcal{U}\mathcal{V}}$.

Definición : Una pregavilla $\mathcal{F}$ es una gavilla si cumple :

1. Axioma de localidad : Dado un abierto $\mathcal{U} \subseteq X$ con una cubierta abierta $\mathcal{U}$ y una sección $s \in \mathcal{F}(\mathcal{U})$, si $s|_{\mathcal{U}_i} = 0$ para todo i , entonces $s = 0 \in \mathcal{F}(\mathcal{U})$.
2. Axioma de pegado : Si $\mathcal{U}$ y $\mathcal{U}'$ estan dados como en el axioma de la localidad , y si $s_i \in \mathcal{F}(\mathcal{U}_i)$ son secciones que cumplen $s_i|_{\mathcal{U}_i\cap\mathcal{U}_j} = s_j|_{\mathcal{U}_i\cap\mathcal{U}_j}$ para todo $i, j \in I$, entonces existe una sección $s \in \mathcal{F}(U)$ tal que $s|_{U_i} = s_i$ para todo i.

Definición : Sea R un anillo , definimos el n-espacio afín $\mathbb{A}^n$ como $\mathbb{A}^n_R = \text{Spec } R[x_1, ..., x_n]$.

Definición : Un espacio anillado $(X, \mathcal{O}_X)$ donde X es un espacio topológico y $\mathcal{O}_X$ una gavilla de anillos en X.

Definición : Un espacio localmente anillado es un espacio anillado $(X, \mathcal{O}_X)$ con la condición de que para cada punto $x \in X$ el tallo $\mathcal{O}_{X_x}$ es un anillo local.

Definición : Un esquema afín es un espacio localmente anillado $(X, \mathcal{O}_X)$ el cual es isomorfo a $(\text{Spec}(R), \mathcal{O}_{\text{Spec R}})$para algún anillo R.

Definición : Un esquema es un espacio anillado localmente $(X, \mathcal{O}_X)$ el cual es localmente isomorfo a un esquema afín.[7]

[7]esta la "generalización"de las variedades mostradas.

1.3. Combinatoria.

Definición : Dados $n+1$ puntos independientes $a_0, ..., a_n$, un n simplejo es la envolvente convexa de los puntos $a_0, ...a_n$-

2. Estadística Algebraica.

La estadística algebraica se ocupa de desarrollar técnicas en geometría algebraica , álgebra conmutativa y combinatoria , para abordar problemas en estadística y sus aplicaciones. Gracias a esto la estadística adquiere una estructura mas cerca a la de la una teoría matemática.

La razón de existir de la estadística algebraica es que muchos conceptos de la estadística tienen un análogos bastante orgánicos en la geometría algebraica , de tal manera que se forma un puente solido entre la matemática y la ciencia de datos.

Estadística	Geometría Algebraica.
Independencia	Variedad de Segre.
Modelo log-lineal	Variedad Torica.
Familia de Curvas exponenciales	Variedades
Modelo de mezcla	Unión de variedades.
Estimación MAP	Tropicalización (geometría tropical).
Modelo de selección	Geometría de singularidades.
...	...

Cuadro 1: Diccionario Estadística-Geometría Algebraica.

Gran parte de la investigación actual en estadística algebraica es en seguir llenado mas entradas al diccionario anterior y seguir ampliando la conexión. Esta unión estadística-geometría algebraica no es una simple curiosidad que solo disfrutemos los matemáticos , si no que tiene una enorme utilidad aun no explotada , y es que el anterior diccionario nos permite "traducir"problemas o preguntas de estadística en problemas o preguntas de geometría algebraica , recordemos la monumental teoría que se ha construido durante el ultimo siglo en geometría algebraica así como la enorme cantidad de investigadores en el área (siendo el área a nivel mundial con mayor cantidad de gente trabajando en ello) y sus enormes conexiones con toda la matemática y gran parte de la física , hacen que la estadística se vea muy beneficiada , ya que sus poderosas herramientas que constantemente se van perfeccionando y ampliando pueden permitir el ataque y resolución de cuestiones estadísticas ; y de igual forma , las preguntas estadísticas pueden motivar el desarrollo de mas técnicas en geometría algebraica , por lo que esta nueva área da un perfecto intercambio de ideas entre un área que se enfoca en preguntas aplicadas y en otra que se enfoca en preguntas puramente en el terreno matemático.

Lamentablemente no es posible en este texto abarcar lo mas posible la estadística algebraica (ya que básicamente usa la mayoría de áreas de la geometría algebraica y bastante de la combinatoria algebraica) , solo nos enfocaremos en la geometría algebraica clásica. El que guste un una introduccion totalmente aplicada le recomiendo la primera lectura de [8]

2.1. El articulo de Sturmfels

Esta área es relativamente nueva , ya que podemos considerar su inicio con la publicación del trabajo de Diaconis (un famoso estadístico) y Sturmfels (una de las mas prominentes figuras en geometría algebraica de la actualidad) a finales de los años 90"s [9] en el cual construyen un algoritmo en términos de bases de Gröbner y álgebra conmutativa , de cadenas de Markov para el muestro de familias exponenciales discretas condicionadas a una estadística[8] suficiente. El algoritmo desarrollado por Diaconis y Sturmfels permite evitar las compilaciones de la teoría asintótica.

La genialidad de la ideal de Diaconis-Sturmfels fue la de poder dar equivalencias algebraicas de

[8] i.e una función medible.

ciertos enunciados estadísticos (que están en términos del lenguaje del análisis) , por ejemplo al inicio definen a una base de Markov como la colección de funciones $f_i : \mathfrak{X} \to \mathbb{Z}$ $(1 \le i \le L)$ tal que

$$\sum_x f_i(x)T(x) = 0$$

y que para cualquier t y $f, f' \in \mathcal{F}_t$ (una estadística) existen $(\epsilon_1, f_{i_1}, ..., \epsilon_A, f_{i_A})$ con $\epsilon = \pm 1$, $f' = f + \sum_{j=1}^{A} \epsilon_i f_{i_j}$ y $f + \sum_{j=1}^{a} \epsilon_j f_{i_j} \ge 0$ para $1 \le a \le A$. Y posteriormente prueban que la siguiente definición es equivalente a la anterior :

Una colección de funciones $f_1, ..., f_L$es una base de Markov si y solo si el conjunto $\mathcal{X}^{f_i^+} - \mathcal{X}^{f_i^-}$ $(1 \le i \le L)$ genera al ideal $\mathcal{J}_T$.

Donde $\mathcal{X}$ es un conjunto finito , $\mathcal{X}^g$ un monomio donde $g : \mathcal{X} \to \mathbb{N}$, $f^+ = \max(f(x), 0)$ y definido de manera análoga , y $\mathcal{J}_T = \mathrm{Ker}\varphi_T$ (de echo una parte del estudio de [9] es estudiar este objeto) , donde $\phi_T : \mathbb{K}[\mathcal{X}] \to \mathcal{K}[t_1, ..., t_d]$ con $x \mapsto t_1^{T(x)_1} t_2^{T_2(x)} \cdots t_d^{T(x)_d}$, donde $T(x)_i$ es la i-esima coordenada de $T(x) \in \mathbb{N}^d$ donde $T : \mathcal{X} \to \mathbb{N}$.

2.2. Viendo la estadística con nuevos lentes.

2.2.1. Variables aleatorias binomiales.

Variables aleatorias binomiales : Consideremos la función polinomial $\phi : \mathbb{C} \to \mathbb{C}^{r+1}$ cuya i-esima función de coordenadas es :

$$\phi_i(t) = \binom{r}{i} t^i (1-t)^{r-i}, \ i = 0, 1, ..., r.$$

Para un parámetro real $\theta \in [0, 1]$, el valor $\phi_i(\theta)$ es la probabilidad $P(X = i)$, donde X es una variable aleatoria binomial con probabilidad de éxito θ. Por ejemplo , si θ es la probabilidad de obtener sol en un lanzamiento de una moneda sesgada , $\phi_i(\theta)$ es la probabilidad de ver i soles en r lanzamientos independiente s de la misma moneda sesgada , por lo tanto el vector $\phi(\theta)$ denota la distribución de una variable aleatoria binomial.

La imagen del intervalo real $[0, 1]$ es una curva dentro del simplejo de probabilidad

$$\Delta_r = \left\{ p \in \mathbb{R}^{r+1} : \sum_{i=0}^{r} p_i = 1, p_i \ge 0 \text{ para todo } i \right\}$$

que consta de todas las posibles distribuciones de probabilidad de una variable aleatoria binomial.

La imagen de la parametrización compleja $\phi(\mathbb{C})$ es una curva en $\mathbb{C}^{r+1}$ que también es una variedad algebraica. En particular es el conjunto de puntos que satisfacen la ecuación trivial $\sum_{i=0}^{p} p_i = 1$, junto con la anulación de todos los subdeterminantes de $2 \times r$ de la matriz :

$$\begin{pmatrix} p_0/\binom{r}{0} & p_1/\binom{r}{1} & p_2/\binom{r}{2} & \cdots & p_{r-1}/\binom{r}{r-1} \\ p_1/\binom{r}{1} & p_2/\binom{r}{2} & p_3/\binom{r}{3} & \cdots & p_r/\binom{r}{r} \end{pmatrix}$$

Algo que podemos notar es que la imagen del intervalo $[0, 1]$ en Δ_r no es un conjunto cerrado de Zariski. La clausura de Zariski $\phi([0, 1])$ es igual a $\phi(\mathbb{C})$, esto es por que $\phi(\mathbb{C})$ no es cierre de Zariski , y la clausura de Zariski de $[0, 1] \subseteq \mathbb{C}$ es todo $\mathbb{C}$. Si solo nos interesan los subconjuntos de Δ_r , entonces $\phi([0, 1])$ es igual a la clausura de Zariski intersectado con el simplejo de probabilidad , en este caso la variedad da una buena aproximación al modelo estadístico de las variables aleatorias binomiales.

Sea $\mathcal{P} \subseteq \Delta_{r-1}$una familia de distribuciones de probabilidad[9] y sea $s > 0$ un entero. El $s-$esimo modelo de mezcla $\mathrm{Mixt}^s(\mathcal{P})$ consiste en todas las distribuciones de probabilidad que surgen como combinaciones convexas de s distribuciones en $\mathcal{P}$. Osea

$$\mathrm{Mixt}^s(\mathcal{P}) = \left\{ \sum_{j=1}^{s} \pi_j : \pi \in \Delta_{s-1} \text{ y } p^{(j)} \in \mathcal{P} \text{ para todo } j \right\}$$

[9]un modelo estadístico.

Los modelos de mezcla nos permiten construir un modelo complejo a partir de un modelo simple $\mathcal{P}$ (al final comentaremos mas acerca de esto).

El caso especial de las mezclas de variables aleatorias binomiales surge en la estadística Bayesiana vía intercambiabilidad mediante el teorema de De Finetti. Considerando el caso $r = 2$ y $s = 2$, podemos notar que $\mathrm{Mixt}^2(\mathcal{P})$ es precisisamente la región bajo la curva en la figura []. Dado que $\mathrm{Mixt}^2(\mathcal{P})$ es un subconjunto 2-dimensional de Δ_2 así que la clausura de Zariski de $\mathrm{Mixt}^s(\mathcal{P})$ es el plano entero $V(p_0 + p_1 + p_2 - 1)$. Entonces la intersección de la clausura de Zariski y el simplejo de probabilidad Δ_2 no es el modelo de mezcla.

En este caso como en algunos otros es necesario de apoyarse de cotas , de esta forma podemos obtener el modelo de mezlcla :

$$\Delta_2 \cap \left\{ p : \det \begin{pmatrix} 2p_0 & p_1 \\ p_1 & 2p_2 \end{pmatrix} \geq 0 \right\}$$

.

Esto es un problema patologico de la geometría algebraica real (subárea de la geometría algebraica donde nos encontramos debido a que consideramos polinomios de coeficientes reales) debido a que $\mathbb{R}$ no es algebraicamente cerrado.

Uno de los problemas clave en estadística algebraica es el de implicitación. Antes de mencionarlo hay que ver otro importante objeto geométrico que es la parametrización (de conjuntos algebraicos) , supongamos que tenemos los siguientes polinomios $\phi_1, .., \phi_r \in \mathbb{K}[x_1, ..., x_n]$, y ahora vamos a tomar la función $\phi : \mathbb{K}^d \to \mathbb{K}^r$, $(\theta_1, ..., \theta_d) \mapsto (\phi(\theta), ..., \phi_r(\theta))$. Si $\Theta \subseteq \mathbb{K}^d$, entonces$\phi(\Theta) = \{\phi(\Theta) : \theta \in \Theta\}$ es la imagen de la parametrización ϕ. Aunque para mayor generalidad podemos considerar el caso para función racional. Es común en estadística algebraica no solo considerar funciones de coordenadas polinomiales , si no también funciones racionales[10] tales funciones son de la forma $\frac{f}{g}$ donde f, g son polinomios. Por lo tanto una función racional , es la función $\phi : \mathbb{C}^d \to \mathbb{C}^r$ donde cada función de coordenadas es de la forma $\frac{f_i}{g_i}$ donde $f_i, g_i \in \mathbb{C}[x_1, ..., x_n]$ son polinomios . Ahora si estamos listos para hablar del problema de implicitación , consideremos una parametrizacion bajo una función racional $\phi : \Theta \to \mathbb{K}^r$, ahora el problema nos pide tomar los conjuntos Θ y la función racional ϕ y calcular el ideal de anulación $I(\phi(\Theta))$, o lo que es lo mismo , si yo tengo $V \subseteq \mathbb{K}^r$ parametrizados

$$x_1 = g_1(t_1, ..., t_m), ..., x_r = g_k(t_1, ..., t_m)$$

Si los g_i son polinomios o funciones racionales , entonces V debe de ser una variedad algebraica (o parte de ella).

Pero usualmente la imagen de la parametrización no es necesariamente siempre una variedad , en el caso de que $\mathbb{K}$ sea algebraicamente cerrado , se aproxima a ser una variedad como lo indica el siguiente teorema.

Teorema : Sea $\mathbb{K}$ algebraicamente cerrado , sea $V \subseteq \mathbb{K}^{r_2}$ una variedad , y sea ϕ una función racional $\phi : V \to \mathbb{K}^{r_1}$. Entonces hay una sucesión finita de variedades $W_1 \supset W_2 \supset \cdots W_k$ en $\mathbb{K}^{r_1}$ tal que

$$\phi(V) = W_1 - (W_2 - (\cdots(-W_k)\cdots))$$

Los conjuntos que surgen del teorema anterior se llaman conjuntos constructibles , por que se construyen mediante una diferencia iterada de variedades. Las imágenes de variedades algebraicas reales (las que usualmente aparecen en estadística algebraica) bajo funciones racionales tienen una estructura especial que nos el teorema de Tarski-Seidenberg , lamentablemente eso excedería el margen planteado para exponer.

El álgebra lineal nos puede ayudar a resolver el problema de implicitacion. Ya que después de un cambio de coordenadas (que es precisamente parametrizar) podemos suponer que dicha transformación lineal es una proyección de coordenadas . El siguiente teorema nos ayudara a mostrar que el ideal de anulación de una proyección de coordenadas puede calcularse usando bases de Gröbner.

[10]Recomiendo leer mas al respecto en un libro de Geometría Algebraica

Teorema : Sea $\pi : \mathbb{K}^{r_1+r_2} \to \mathbb{K}^{r_1}$ una proyección de coordendas $(a_1, .., a_{r_1}, b_1, ..., b_{r_2}) \mapsto (a_1, ..., a_r)$. Sea $V \subseteq \mathbb{K}^{r_1+r_2}$ una variedad y sea $I = I(V) \subseteq \mathbb{K}[p_1, .., p_{r_1}, q_1, ..., q_{r_2}]$ su ideal de anulación. Entonces el ideal de anulación de la imagen $\pi(V)$ es el ideal

$$I(\pi(V)) = I \cap \mathbb{K}[p_1, ..., p_{r_1}]$$

En álgebra computacional , si tenemos un ideal $I \subseteq \mathbb{K}[p_1, .., p_{r_1}, q_1, ..., q_{r_2}]$, el ideal $I \cap \mathbb{K}[p_1, ..., p_{r_1}]$ es llamado el ideal de eliminación por que las variables $q_1, ..., q_{r_2}$ fueron eliminadas. Los generadores para los ideales de eliminación pueden ser determinados por el calculando una base de Gröbner con un orden apropiado de las variables.

Teorema : Sea $I \subseteq \mathbb{K}[p_1, .., p_{r_1}, q_1, ..., q_{r_2}]$ y sea $\prec$ un orden de termino lexicográfico tal que

$$q_1 \prec \cdots \prec q_{r_2} \prec p_q \prec \cdots \prec p_{r_2}$$

Si $\mathcal{G}$ es una base de Gröbner de I con respecto de $\prec$, entonces

$$I \cap \mathbb{K}[p_1, \cdots, p_{r_|}] = (\mathcal{G} \cap \mathbb{K}[p_1, ..., p_{r_2}])$$

y además $\mathcal{G} \cap \mathbb{K}[p_1, ..., p_{r_2}]$ es una base de Gröbner para el ideal de eliminación $I \cap \mathbb{K}[p_1, .., p_{r_1}]$.

Para proyecciones lineales solo es un simple paso para calcular las imágenes. Supongamos que $\phi : \mathbb{K}^{r_2} \to \mathbb{K}^{r_1}$ es una función polinomial (cada componente de la función es un polinomio) , la gráfica de la función es el conjunto de puntos :

$$\Gamma_\phi = \{(\phi(\theta), \theta) : \theta \in \mathbb{K}^{r_2}\} \subset \mathbb{K}^{r_1+r_2}$$

Si $\mathbb{K}$ es un campo finito el ideal de anulación de la grafica es :

$$I(\Gamma_\phi) = (p_1 - \phi(q), ..., p_{r_1} - \phi_{r_1}(q)).$$

Con el siguiente teorema damos por resuelto el problema de implicitación.

Teorema : Sea $I \subseteq \mathbb{K}[q_1, ..., q_{r_2}]$el ideal de anulación de alguna variedad $V \subseteq \mathbb{K}^{r_2}$. Sea ϕ un mapeo polinomial $\phi : \mathbb{K}^{r_2} \to \mathbb{K}^{r_1}$. Entonces el ideal de anulación de la clausura de Zariski es el ideal de eliminación

$$I(\phi(V)) = I(I + I(\Gamma_\theta)) \cap \mathbb{K}[p_1, ..., p_{r_2}].$$

El ideal $I + I(\Gamma_\theta)$ es generado por el conjunto de todos los polinomios generadores de I mas los r_1 polinomios $p_1 - \phi(q), ..., p_{r_1} - \phi_{r_1}(q)$. Entonces el problema de implicitación puede ser resuelto calculando una base de Gröbner.

Vamos a estudiar el problema de implicitación para las variables aleatorias binomiales. Vamos a considerar la parametrización que dimos inicialmente para dos ensayos. Para calcular la descripción implícita de este modelo , vamos a comenzar con el ideal de anulación I del simplejo de probabilidad Δ_1 , el cual es el ideal

$$I = (s + t - 1) \subseteq \mathbb{R}[s, t]$$

La parametrización tiene la forma

$$\phi(s, t) = (s^2, 2st, t^2)$$

Podemos formar el ideal de la gráfica $I + \Gamma_\theta = (s+t-1, p_0 - s^2, p_1 - 2st, p_2 - t^2) \subseteq \mathbb{R}[s, t, p_0, p_1, p_2]$.

Calculando la base de Gröbner de este ideal con $s \prec t \prec p_0 \prec p_1 \prec p_2$ produce el ideal de anulación de este modelo : $(p_0 + p_1 + p_2 - 1, 4p_0p_2 - p_1^2)$.

Debido a que estas herramientas se emplean en estadística el uso de Software es vital , afortunadamente existen dos muy buenos , que son `Macaulay2` y `Singular` con los cuales se pueden efectuar de manera computacional todo estos cálculos que en la practica obtenerlos a mano suele ser muy laborioso y tardado , esto ya es terreno de la geometría algebraica numérica y computacional.

Por ultimo , con técnicas mucho mas refinadas de la geometría algebraica (geometría tropical y geometría torcía) es posible encontrar un modelo para variables aleatorias binomiales Binom$(2, \theta)$

$$\mathcal{M} = \left\{((1-\theta)^2, 2\theta(1-\theta), \theta^2) : \theta \in (0,1)\right\}$$

Que nos permite estudiar con soltura este tipo de variables.

2.2.2. Algunos modelos estadísticos algebraicos elementales.

Vamos a llamar espacio de estado al conjunto : $[m] = \{1, ..., m\}$, Una distribución de probabilidad en el conjunto $[m]$ es un punto en el simplejo de probabilidad

$$\Delta_{m-1} = \left\{(p_1, ..., p_m) \in [0,1]^m : \sum_i p_i = 1\right\}$$

Un modelo estadístico algebraico esta definido por un mapeo polinomial $f : \mathbb{R}^d \to \mathbb{R}^m$dado por $(\theta_1, ..., \theta_d) \mapsto (f_1(\theta), ...f_m(\theta))$ donde $\theta_1, ..., \theta_d$ son parámetros del modelo y $f_1, ..., f_m \in \mathbb{R}[x_1, ..., x_n]$ polinomios o funciones racionales. El vector de parámetros $(\theta_1, ..., \theta_d)$ oscila sobre un subconjunto abierto no vació adecuado Θ de $\mathbb{R}^d$, el espacio de parámetros de f. Vamos a suponer que el espacio de parámetros satisface

$$\Theta \subseteq \left\{\theta \in \mathbb{R}^d : f_i(\theta) > 0, 1 \leq i \leq m\right\}$$

Entonces tenemos que $f(\Theta) \subseteq \Delta_{m-1} \Leftrightarrow \sum_{i=0}^{m} f_i(\theta_i) = 1$. Si se cumple la anterior condición tenemos que $f(\Theta)$ es el modelo. No siempre la información modelada satisface la condición , y en dicho caso los vectores en $f(\Theta)$ pueden ser escalados para obtener una familia de distribuciones de probabilidad en $[m]$: $\frac{1}{\sum_i f_i(\theta_i)}(f_1(\theta), ..., f_m(\theta))$ con $\theta \in \Theta$.

Los datos de una muestra suelen estar dados por una secuencia de valores del espacio de estado $i_1, i_2, .., i_N$ donde N es el tamaño de la muestra. Si suponemos que los valores son independientes e idénticamente distribuidos , entonces los datos los podemos resumir por el vector de frecuencia : $u = (u_1, u_2, ..., u_m)$ donde $u_i$el es número de ocurrencias de $i \in [m]$ en la información $(1 \leq i \leq m)$, por lo que $u_1 + ... + u_m = N$ y la distribución empírica de los datos esta dada por el vector $\frac{1}{N}(u_1, ..., u_m)$ el cual pertenece al simplejo de probabilidad Δ_{m-1} . Y las coordenadas $\frac{u_i}{N}$ son las frecuencias relativas observadas de los resultados.

Vamos a considerar un modelo estadístico algebraico $f : \mathbb{R}^d \to \mathbb{R}^m$ para la información. La probabilidad de observar los datos $i_1, .., i_N$ viene dada por $L(\theta) = f_{i_1}(\theta) f_{i_2}(\theta) \cdots f_{i_N} = f_1(\theta)^{u_1} \cdots f_{i_m}(\theta)^{u_m}$. Si el vector de frecuencia u se mantiene fijo , la función de verosimilitud L es una función del espacio de parámetros θ a $\mathbb{R}^+$. La probabilidad de observar el vector de frecuencia u viene dada por

$$\binom{N}{u_1, ..., u_m} L(\theta)$$

El vector de frecuencia es un estadístico suficiente para el modelo f dado que la función de verosimilitud $L(\theta)$ depende de los datos a través de u (y no a través de los datos en si). Para evitar los problemas numéricos convencionales usamos la transformación logarítmica y representación la función de verosimilitud por una función de verosimilitud logarítmica : $\mathcal{L}(\theta) = u_1 \log L(\theta) + \cdots + u_m \log f_m(\theta)$.

La función log-verosimilitud $\mathcal{L}(\theta)$ es una función del espacio de parámetros Θ a $\mathbb{R}^-$. El problema de la estimación de máxima verosimilitud es maximizar la función de verosimilitud $L(\theta)$ de manera equivalente , o de manera equivalente , la función escalada log-verosimilitud $\mathcal{L}(\Theta)$ sobre el espacio de parámetros :

$$\max \mathcal{L}(\theta) \quad \text{tal que} \quad \theta \in \Theta$$

Una solución a este problema de optimización es llamado máximo estimador de verosimilitud de θ con respecto del modelo f y la información u. Como paréntesis , este tipo de problemas caen en el área de terreno de la geometría tropical (que al final se convierten en problemas combinatorios).

Dependiendo del tipo de modelo estadístico algebraico que se tome puede variar la facilidad con el que resuelve el problema , veremos el mas simple de ellos que es el modelo lineal.

Un modelo estadistico algebraico $f : \mathbb{R}^d \to \mathbb{R}^m$ es llamado modelo lineal si las funciones coordendas $f_i(\theta)$ son funciones lineales. O sea , para los números reales $a_{i1}, ..., a_{id}$ y b_i , $1 \leq i \leq d$, tal que

$$f_i(\theta) = \sum_{j=1}^{d} a_{ij}\theta_j + b_j \ 1 \leq i \leq d.$$

Teorema : Para cualquier modelo lineal $f : \mathbb{R}^d \to \mathbb{R}^m$ y una estadística suficiente $u \in \mathbb{N}_0^m$, la función log-verosimilitud

$$\mathcal{L}(\theta) = \sum_{i=1}^{m} u_i \log f_i(\theta)$$

es cóncava. Si $f$es inyectiva y la información u_i $(1 \leq i \leq m)$ es positiva , entonces la función log-verosimilitud $\mathcal{L}(\theta)$ es estrictamente convexa.

El máximo estimador de verosimilitud de un modelo lineal viene dado por los puntos críticos de la función log-verosimilitud.
Corolario : Si el modelo lineal $f : \mathbb{R}^d \to \mathbb{R}^n$es inyectivo y la información u_i $(1 \leq i \leq m)$ es positiva , entonces cada punto critico de la función log-verosimilitud θ es un máximo local.

Para mas información recomiendo al lector [10] , [11] y [12]

3. Teoría de la probabilidad homotópica.

Debido a que la teoría de la probabilidad homotopica es un área que aun esta en pañales (ni siquiera existe un libro del tema) y que es muy joven (se gesto en la década pasada) , además del enorme cumulo de conocimientos previos solo expondré los puntos mas clave de la base para adentrarse en la teoría .

Al igual que en la estadística algebraica , esta área de estudio surge de encontrar vinculo entre dos áreas diferentes , en este caso el álgebra homotopica y la teoría de la probabilidad . Y es que existe una coincidencia entre los morfismos de homotopia en álgebra homotopica y la funciones cumulantes en la teoría de la probabilidad.

En esta linea de investigación se da la noción de espacio de probabilidad homotopico , el cual enriquece de la noción del espacio de probabilidad algebraico (que a su vez es una mejora de los espacios de probabilidad clásicos). Esta mejora emplea la caracterización de las leyes de variables aleatorias en un espacio de probabilidad en términos de la simetría de la expectación. Ahora dichas leyes las vemos como invariantes de los tipos de homotopia de los morfismos entre ciertas álgebras de homotopia.

Vamos a ver como se construye esta teoría , consideremos un espacio vectorial graduado V , y a $TV = \bigoplus_{n=1}^{\infty} V^{\oplus n}$[11] que es también una coalgebra.

Definición : Una A_∞ álgebra es el par (V, D) donde D es la coderivacion $D : TV \to TV$ que satisface $D^2 = 0$. Y un morfismo entre A_∞ álgebras es un una función de (V, D) a (V', D') es una función entre coalgebras $F : TV \to TV$ que cumple $FD = D'F$.

Definición : Sea $C = (V, d)$ un complejo de cadenas , una A_∞ estructura en C es una A_∞ álgebra (V, D) con $d_1 = d$.[12]. Con $\mathcal{M}_C$ denotamos el conjunto de A_∞ estructuras en C.

Vamos a emplear de base una definición mas moderna de espacio de probabilidad , pues lo vamos a definir la terna (V, e, a) , donde V es un espacio vectorial complejo (a sus elementos los llamaremos variables aleatorias) , $e : V \to \mathbb{C}$ una transformación lineal donde al número $e(X)$ lo llamaremos expectación de la variable aleatoria $X \in V$, y $a : V \times V \to V$es un producto bilineal

[11] $V^{\oplus n}$ denota al n-esimo producto vectorial de V
[12] d_1 es la primera componente de D

asociativo.

Definición : Sea (V, a, e) un espacio de probabilidad. El n-esimo cumulante de (V, e, a) es la transformación lineal $\kappa_n : V^{\otimes n} \to \mathbb{C}$ definido recursivamente como

$$e \circ \alpha_n = \sum_{k=1}^{n} \alpha'_k \circ (\sum_{P} \bigotimes_{j}^{k} \kappa_{n_j})$$

La siguiente observación motiva la definición de espacio de probabilidad homotopico : Sea (V, e, a) un espacio de probabilidad. $(V, 0), (\mathbb{C}, 0)$ son A_∞ álgebras y la función $e : V \to \mathbb{C}$ y la función $e : V \to \mathbb{C}$ define un morfismo de A_∞ álgebras. $E : TV \to T\mathbb{C}$ denota el levantamiento de e como una función de coalgebras. Además sea $a = \exp(A) := \mathrm{id} + A + \frac{A^2}{2!} + \ldots.$ donde $A : TV \to TV$ es el levantamiento de $a : V \otimes V \to V$ como coderivación y de igual manera $a' = \exp(a')$ donde $A' : T\mathbb{C} \to T\mathbb{C}$ es el levantamiento de $a' : \mathbb{C} \otimes \mathbb{C} \to \mathbb{C}$ a una coderivacion. Entonces $E^{a,a'} : TV \to T\mathbb{C}$ es un A_∞ morfismo entre $(V, 0^a) = (V, 0)$ y $(\mathbb{C}, 0^{a'}) = (\mathbb{C}, 0)$. Con $e_n : V^{\otimes n} \to \mathbb{C}$ denota las componentes de un A_∞morfismo $E^{a,a'}$. Lo asombroso y clave es que es cierta la siguiente proposición :

$$\kappa_n = e_n$$

esto nos motiva a desarrollar una teoría de la probabilidad homotopica. Para construir una definición de espacio de probabilidad homotopico hay que remplazar el espacio V de variables aleatorias y remplazarlo por un complejo de cadenas de variables aleatorias.

Definición de un espacio de probabilidad homotopico: Un espacio de probabilidad homotopico consiste de un complejo de cadenas $C = (V, d)$, una función de cadena $e : C \to \mathbb{C}$ y un producto asociativo $a : V^{\otimes 2} \to V$.

Esta coincidencia entre los cumulantes de un espacio de probabilidad y un A_∞ morfismo en la expectación nos dan un vistazo de como trabajar con nueva óptica. Ahora podemos remplazar la antigua definición de cumulante y pensarlo ahora como A_∞ morfismos $E^{a,a'}$ en la función de cadena $e : (V, d) \to (\mathbb{C}, 0)$ asociado al producto $a : V^{\otimes 2} \to V$ y $a' : \mathbb{C}^{\otimes 2} \to \mathbb{C}$ de números complejos. Entonces los comulantes son un morfismo de A_∞ entre la estructura A_∞ (V, D^a) y la A_∞ estructura $(\mathbb{C}, D^{a'})$.

Esta nueva área se relaciona con otras áreas de investigación de alto calibre como las categorías superiores , la teoría de ∞-categorías y la geometría algebraica derivada. Se puede consultar mas en [13] y en [14]

4. Referencias.

En caso de que el lector quiera comenzar a profundizar mas en las áreas puede guiarse la siguiente bibliografia. Además como autopublicidad , tengo unas notas de curso de los temas expuestos en : https://kevinciencias.github.io/aeml.html

Referencias

[1] John Terilla Gabriel C. Drummond-Cole. Homotopy probability theory on a riemannian manifold and the euler equation. *New York Journal of Mathematics*, 23:1065–1085, 2017.

[2] Bertrand Michel Frédéric Chazal. An introduction to topological data analysis : fundamental and practical aspect for data scientists. *Pre-print*, 2021.

[3] Persistent Homology For Metric Mesure Spaces, Robust Statistics For Hypotesis Testing, and Confidence Intervals. Andrew j. blumberg , itamar gak , michael a. mandell , matthew pancia. *Pre-print*, 2012.

[4] Antonio Montes. *The Gröbner Cover*, volume 27 of *Algortithms and Computation in Mathematics*. Springer, first edition, 2018.

[5] Steven Kleiman Allen Altman. *A term of commutative algebra*. Worldwide Center of Mathematics, 2017.

[6] I.G. MacDonald M.F.Atiyah. *Introduction to Commutative Algebra*. Addison-Wesley , Oxford University, 1969.

[7] Bruno Vallette Jean-Luois Loday. *Algebraic Operads*, volume 346 of *Grundlehren der mathematischen Wissenschaften*. Springer, first edition, 2012.

[8] Bernd Sturmfels Lior Patcher. *Algebraic Statistics for Computational Biology*. Cambridge University Press, 2005.

[9] Bernd Sturmfels Persi Diaconis. Algebraic algorithms for sampling from conditional distributions. *The Annals of Statistics*, 26(1):363–397, 1998.

[10] Seth Sullivant Mathias Drton, Bernd Sturmfels. *Lectures on Algebraic Statistics*, volume 39 of *Oberwolfach*. Birkhäuser, first edition, 2009.

[11] Seth Sullivan. *Algebraic Statistics*, volume 194 of *Graduate Studies in Mathematics*. American Mathematical Society, first edition, 2018.

[12] Henry P. Wynn Giovanni Pistone, Eva Riccomagno. *Algebraic Statistics computational Commutative Algebra in Statistics*. Chapman & Hall, 2001.

[13] John Terilla Gabriel C.Drummond-Cole, Jae-Suk Park. Homotopy probability i. *Homotopy Related Structures*, 10:425–435, 2013.

[14] Jae-Suk Park. *Homotopy Theory of Probability Spaces I*. ArXiv, 2015.