
ÍNDICE

Editorial	2
-----------------	---

MATEMÁTIC@ DEL NÚMERO

La matemática imperfecta	3
--------------------------------	---

AXIOMAS, TEOREMAS Y ALGO MÁS

Caracterización del Número Básico Reproductivo en Modelos Poblacionales	6
El problema del matrimonio estable	13

ATERORIZANDO IDEAS

El Método de Conjunto Activo para Problemas de Optimización Cuadráticos	20
Un problema novohispano	25
El Problema de Monty Hall	30
Una introducción práctica a las cópulas	34

ZONA OLÍMPICA

Lista de problemas	50
Pregunta de Erdős	50

EN EL HORIZONTE

Apología del matemático	51
Matemáticas prehispánicas: entrevista con Blair Madore	55
La flexibilidad de las matemáticas: entrevista con Rafael Prieto	62



Consejo Académico

Claudia Gómez Wulschner
César Luis García

Consejo Editorial

Directora

María José Domenzain Acevedo

Tesorero

Alejandro Méndez Lemus

Secretario

Daniel Salnikov

Edición

Dan Jinich Fainsod
Efrén Zagal Rodríguez
Guillaume Domenge Treuille
Ílan Jinich Fainsod
José Ángel Rodríguez
Karla Daniela Hernández
Ricardo Zúñiga Tello

Redes sociales

María Beatriz Montaña Murrieta

Diseño web

Alberto Albarrán Acuña

LABERINTOS INFINITOS, Año 2019, No. 49, enero-mayo 2019, es una publicación cuatrimestral editada por el Instituto Tecnológico Autónomo de México, a través de las Direcciones de Actuaría y Matemáticas del ITAM. Calle Río Hondo No. 1, Col. Progreso Tizapán, Delegación Álvaro Obregón, C.P. 01080, Tel. 56284000 ext 1732. www.itam.mx, raulranirezriba@gmail.com.

Editor responsable: María José Domenzain Acevedo. Reservas de Derechos al Uso Exclusivo No. 04-2016-112313125200-102, ISSN: en trámite, ambos otorgados por el Instituto Nacional del Derecho de Autor. Licitud de Título y contenido en trámite, otorgado por la Comisión Calificadora de Publicaciones y Revistas Ilustradas de la Secretaría de Gobernación, Permiso SEPOMEX en trámite.

Impresa por Antonio Bettech Shabot, ubicada en Eje Central Lázaro Cárdenas No. 214, Col. Guerrero, México CDMX, C.P. 06300. Este número se terminó de imprimir el 28 de noviembre del 2018 con un tiraje de 1500 ejemplares. Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor de la publicación.

Queda estrictamente prohibida la reproducción total o parcial de los contenidos e imágenes de la publicación sin previa autorización del Instituto Nacional del Derecho de Autor.

Editorial

Las matemáticas se encuentran en donde quiera que uno mire. Más allá de los salones de clases o de los libros académicos, las matemáticas ocupan nuestro día a día: en el café que tomamos por la mañana, en los días nublados, en las rutas que seguimos... En donde quiera que busquemos, podemos encontrar matemáticas; el reto está en, como un artista, mostrar la belleza en ellas a quienes no la encuentran.

Agradecimientos

A la División Académica de Actuaría, Estadística y Matemáticas del ITAM. En especial a Beatriz Rumbos, Claudia Gómez, César Luis García. A la Dirección Escolar del ITAM, específicamente a Magdalena Barba. Gracias a Raphson y Mathphile, representaciones de los alumnos de Actuaría y Matemáticas Aplicadas, respectivamente, por el apoyo brindado. Agradecemos también al Fondo de Organizaciones Estudiantiles y al Consejo Universitario de Honor y Excelencia.

$\sqrt{69}$

<http://laberintos.itam.mx>
laberintos@itam.mx



Diseño de portada:
Milenka Rufrancos Flores

Imagen de portada: Ricardo Zúñiga Tello

La matemática imperfecta

Efrén Zagal Rodríguez

Estudiante de Matemáticas Aplicadas del ITAM

"All in all, I have found great delight and pleasure in the pursuit of mathematics. I have been saved from boredom, dourness and self-absorption. One cannot ask for more."

Karen Uhlenbeck

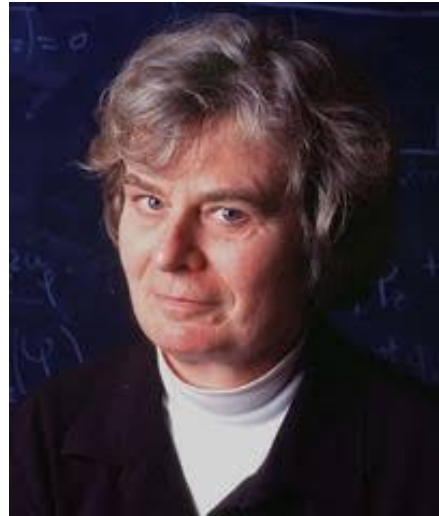
Este 21 de noviembre se entregará por primera vez el premio Abel (conocido como "El nobel de las Matemáticas") a una mujer. La ganadora de este prestigioso reconocimiento es Karen Uhlenbeck, una matemática estadounidense reconocida ampliamente por sus aportaciones en el análisis geométrico, entre otras áreas.

Aunque seguramente la Historia recordará a esta brillante matemática como la primera mujer en conseguir este premio, es reduccionista sólo hablar de este hito, pues Uhlenbeck tuvo muchos retos, triunfos y tragos amargos. En sus propias palabras: "Es difícil ser un modelo a seguir, porque lo que realmente tienes que hacer es demostrarle a los estudiantes qué tan imperfecta puede ser la gente y aún así tener éxito".

Karen (Keskulla) Uhlenbeck nació el 24 de Agosto de 1942 en Cleveland, Ohio. Hija de un ingeniero y una artista, Uhlenbeck fue una ávida lectora desde niña, como ella lo cuenta, leía toda la noche en su biblioteca local hasta el punto de terminar todos los libros disponibles.

Si bien ella misma reconoce que no entendió qué eran realmente las matemáticas hasta sus últimos años de la universidad, sí recuerda haber estado fascinada con su primer acercamiento al concepto abstracto de los distintos tipos de infinitos.

Aunque Uhlenbeck mostró un gran potencial como estudiante, no realizó su carrera en ninguno de los programas de Matemáticas más importantes del país, sino en la Universidad de Michigan, que no figuraba en aquel entonces. Sin embargo, tuvo la oportunidad de entrar en un prestigioso programa para alumnos con honores en esta institución. Uhlenbeck recuerda que en esta etapa de sus estudios tuvo clases con otras mujeres sumamente talentosas; ella atribuye esto, que era algo extraño en esa época, a que las familias americanas no estaban dispuestas a invertir mucho dinero en hijas con vocaciones académicas -aunque sí lo harían con hijos varones- de lo que se seguía que muchas mujeres brillantes terminarían estudiando en programas académicos secundarios en el panorama nacional.



Si bien Uhlenbeck siempre tuvo facilidad con las matemáticas gracias a su excelente pensamiento abstracto, ella empezó estudiando Física. Aunque esto fue sumamente relevante para las aplicaciones de su teoría en años posteriores, ella misma revela que nunca disfrutó mucho la parte experimental.

Otra razón por la que la carrera en Matemáticas fue su predilecta radicó en su personalidad introvertida, pues, de acuerdo con ella, al ser la menor de cuatro hermanos le costaba trabajo relacionarse con otras personas, y de esta forma concentraba su naturaleza competitiva en intentar superarse a sí misma día a día. Lo que Uhlenbeck ignoraba es que la naturaleza de su trabajo la llevaría a trabajar con múltiples profesores y alumnos. Aunque admite que le costó trabajo, esto no fue impedimento para que fuera una académica muy destacada durante muchos años.

Al terminar sus estudios universitarios, Uhlenbeck realizó un año de estudios en el Instituto Courant de Nueva York, para después casarse con un bioquímico que estaba a punto de realizar un posgrado en Harvard. De este modo, se fue a vivir con él y realizó su maestría en 1964 y posteriormente su doctorado en 1968 en la Universidad de Brandeis bajo la supervisión de Richard Palais.

Al momento de buscar un cargo permanente en la misma institución que su esposo, Uhlenbeck se enfrentó con lo que era una verdad incómoda de aquella época: los académicos de esa época no contrataban mujeres. Mientras que la excusa de las grandes universidades era de un carácter de nepotismo, no era difícil ver que la concepción que tenían era que las mujeres no pertenecían en academia. Así, Uhlenbeck terminó encontrando una posición en la Universidad de Illinois, Champaign Urbanam, en una época en la que ella describe que se sentía fuera de lugar en el ámbito matemático y social, pues el lugar era feo, burgués y plano.

Al abordar el tema de las mujeres en las matemáticas, Uhlenbeck reconoce haber cambiado su postura múltiples veces. Cuando era joven y relativamente exitosa en la academia, menciona que no tuvo mucha empatía con la tendencia feminista, pues a ella le iba bien en sus estudios sin la necesidad de un movimiento social. En ese momento, incluso pensaba que podía ser contraproducente, ya que no quería una compensación excesiva. En sus palabras: "Estaría igual de ofendida si me contrataran por ser mujer, yo quiero que me valoren por mi trabajo como matemática, no por ser parte de algún grupo en específico".

Cuando más adelante en su vida se enfrentó a un machismo sistemático por parte de las grandes universidades, se dio cuenta de la importancia de esta lucha en el área de las matemáticas y hasta la fecha se ve a sí misma como una figura a seguir para las mujeres matemáticas jóvenes que aún se enfrentan a estas actitudes. Prueba de esto es el programa de tutoría que fundó en el Instituto de Estudios Avanzados en Princeton para mujeres en 1994.

En cuanto a su investigación, esta exitosa matemática se desempeñó en el ámbito de ecuaciones diferenciales parciales. Esta pericia se juntó con conocimiento sobre geometría del espacio-tiempo, además de topología de variedades y teoría de campos. En particular, Uhlenbeck trabajó en herramientas analíticas para usar instantones como una herramienta geométrica. Estos avances les fueron útiles, entre otras personas, a Simon Donaldson y Edward

Witten, ambos ganadores de medallas Fields.

Aunque no es posible hacerle justicia a los avances científicos de Uhlenbeck en esta biografía, sus logros académicos hablan por sí mismos. Prueba de esto es su participación como MacArthur Fellow en 1983, su elección a la Academia Americana de Artes y Ciencias en 1985 y la Academia Nacional de Ciencias en 1985, así como la obtención de la Medalla Nacional de Ciencias en el año 2000 por sus sobresalientes aportaciones al conocimiento. Además de esto, ocupó el puesto como vicepresidenta de la *American Mathematical Society* y como ya se sabe, fue la primer mujer en ganar el premio Abel en el 2019.

A pesar de todos sus reconocimientos, Uhlenbeck mencionó en su discurso de aceptación del *Steele Prize* que sus mayores logros no son teoremas matemáticos, sino lo que ha podido hacer en el ámbito de la educación. Todos sus programas de mentoría y colaboración con estudiantes han cambiado la vida de muchos jóvenes y eso es lo que más le llena de orgullo en su carrera.

Consciente de su posición como figura a seguir, ella concluye que "Cualquiera sabe que si una persona es inteligente, divertida, bonita y bien vestida, tendrá éxito, pero también es posible triunfar con todas tus imperfecciones. Me tomó mucho tiempo darme cuenta de esto. Al respecto, estar en la posición de modelo a seguir es una posición con poco *glamour*, pues muestras a toda la gente tu lado malo. Puedo ser una gran matemática y ser famosa por ello, pero también soy muy humana".

Referencias

- [1] O'Connor, J.J. & Robertson, E.F.. Karen Keskulla Uhlenbeck. Consultado de: http://www-history.mcs.st-andrews.ac.uk/Biographies/Uhlenbeck_Karen.html
- [2] Chang, Kenneth. Karen Uhlenbeck Is First Woman to Win Abel Prize for Mathematics. The New York Times. Marzo 2019.
- [3] Riddle, Larry. Karen Uhlenbeck. Biographies of Women Mathematicians. Consultado de: <https://www.agnesscott.edu/lriddle/women/uhlenbk.htm>
- [4] S. Ambrose et al. "Journeys of Women in Science and and Engineering, No Universal Constants", Temple University Press. A Personal Profile of Karen K. Uhlenbeck. Consultado de: <https://web.ma.utexas.edu/users/uhlen/vita/pers.html>
- [5] Lee. Entrevista a Karen Uhlenbeck. Consultado de: <https://web.ma.utexas.edu/users/uhlen/vita/interview05.pdf>
- [6] Uhlenbeck, K. Response to the Award of the Steele Prize for Seminal Contribution to Research. Consultado de: <https://web.ma.utexas.edu/users/uhlen/vita/steele.pdf>
- [7] Noether, Emmy. 1988 Lecturer: Karen Uhlenbeck. Association for Women in Mathematics (AWM). Consultado de: <https://awm-math.org/awards/noether-lectures/noether-lectures-1988/>

Caracterización del Número Básico Reproductivo en Modelos Poblacionales

Alejandro Méndez Lemus
Estudiante de Matemáticas Aplicadas del ITAM

Mayra Núñez López
Departamento de Matemáticas, ITAM

Introducción

Consideremos una enfermedad viral que una vez contraída y recuperado el paciente confiere inmunidad total, se clasifica desde la perspectiva compartamental, como una enfermedad tipo *SIR* que describe a un individuo que pasa por tres etapas en el transcurso de la infección: Susceptible, Infeccioso e Inmune o Recuperado. Las ecuaciones básicas son las siguientes

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI, \\ \frac{dI}{dt} &= \beta SI - \gamma I, \\ \frac{dR}{dt} &= \gamma I,\end{aligned}$$

donde β se denomina la tasa de infección y γ la tasa de recuperación. De la segunda ecuación se puede ver que dI/dt es positiva o negativa dependiendo de que

$$R_0 = \frac{\beta S}{\gamma}$$

sea mayor o menor que 1. En modelos epidemiológicos R_0 es un umbral de suma importancia, denominado **número básico reproductivo**, que se define como el número de infecciones secundarias que un individuo infeccioso produce en una población formada de individuos susceptibles únicamente. Si $R_0 > 1$ los individuos infecciosos al inicio de la enfermedad incrementarán el número de nuevas infecciones y producirán un brote epidémico, si $R_0 < 1$ los individuos enfermos serán transmisores poco eficientes y la enfermedad tenderá a desaparecer, finalmente si $R_0 = 1$ cada individuo infeccioso se reemplazará a sí mismo y no se producirá un brote epidémico.

El objetivo de este escrito es presentar una metodología del cálculo e interpretación de R_0 en sistemas de ecuaciones discretos y sistemas continuos que describen la propagación de una epidemia, así como el papel que determina en la estabilidad del mismo.

Modelos poblacionales y matriciales

La primera contribución significativa a la teoría de poblaciones, desde una perspectiva matemática, fue la de Thomas Malthus, quien en 1798 publicó su **Ensayo sobre el Principio de**

la Población. Malthus introdujo el concepto de que la demanda de recursos naturales requeridos por una población que aumenta, en determinado momento, tiene que exceder el suministro. El modelo que deriva de la idea anterior propone que la población crece de manera exponencial, mientras que los recursos que consume crecen de manera aritmética. La expresión matemática para el modelo malthusiano es $P(t) = P_0 \exp(kt)$, que a su vez puede representarse como una ecuación diferencial de la forma:

$$\frac{dP}{dt} = kP$$

donde $P(t)$ denota el tamaño de la población al tiempo t , por lo que si $t = 0$ es el primer momento considerado entonces P_0 es la población inicial y k es la tasa de incremento de la población.

De esta manera, la ecuación presentada nos dice que si se tiene una población inicial de P_0 individuos, entonces, después de transcurrido cierto tiempo tendremos $\exp(kt)$ veces lo que se tenía al principio. Dado que la función exponencial es siempre creciente, a medida que transcurre el tiempo la población crecerá indefinidamente. Los modelos de crecimiento ilimitado tienen validez en la práctica para periodos cortos de tiempo en los cuales se pueden despreciar los distintos factores que alteran la tasa de crecimiento. Sin embargo, como los recursos son limitados, a medida que la población crece, los recursos disponibles por individuo decrecerán, así los individuos se verán obligados a competir por ellos y se producirá un efecto inhibitor sobre el crecimiento de la población. Modelos que tomaron en cuenta las consideraciones anteriores fueron desarrollados años después de la publicación del modelo malthusiano.

En 1837 el matemático holandés Pierre Verhulst desarrolló el modelo de crecimiento logístico, en el que sugiere que la tasa de incremento de la población depende de la densidad de ésta; a bajas densidades, la tasa de crecimiento es muy cercana a k . A medida que la población aumenta, la tasa de crecimiento va disminuyendo y tiende a 0 cuando $P = N$. El parámetro k se interpreta como la tasa de crecimiento de la población en ausencia de competencia intra-específica, y el parámetro N es el límite superior del crecimiento de la población o capacidad de carga, que comunmente se interpreta como el número máximo de organismos que pueden ser mantenidos con los recursos que hay en el medio. Es destacable que a diferencia del modelo propuesto por Malthus donde la población crecía sin cota conforme el tiempo transcurría, en este caso si el tamaño de la población excede N , entonces la tasa de crecimiento de la población será negativa y el tamaño de la población decrecerá. La dinámica de la población está descrita por la siguiente ecuación diferencial:

$$\frac{dP}{dt} = k \left(1 - \frac{P}{N} \right) P,$$

que se diferencia de la ecuación malthusiana por el término de crecimiento acotado $-\frac{k}{N}P^2$, el cual nos indica el efecto inhibitor que tiene la competencia intra-específica sobre el crecimiento de la población, ya que el promedio estadístico del número de encuentros entre dos individuos es proporcional a P^2 . De esta manera el modelo logístico combina

dos procesos ecológicos: reproducción y competencia. Los modelos anteriores son la base de sistemas más complejos utilizados para el análisis de enfermedades epidemiológicas, en las siguientes secciones algunos resultados teóricos y ejemplos.

Modelos discretos poblacionales

En esta subsección estudiaremos poblaciones estructuradas por edades descritas con modelos en tiempo discreto. Las herramientas matemáticas presentadas en esta sección provienen del Álgebra Lineal, específicamente de la teoría espectral de matrices no negativas. Los teoremas presentados a continuación requieren del entendimiento de cuatro aspectos matemáticos: las definiciones de *valor/vector propio*, la matriz no negativa (definida a continuación).

Definición 1. Una matriz cuadrada A con entradas reales, $A \in \mathbb{R}^{n \times n}$, es **no negativa** cuando sus entradas son no negativas, es decir, para cualesquiera $i, j \in \{1, 2, \dots, n\}$

$$a_{ij} \geq 0.$$

Una matriz es **irreducible** cuando no hay manera de llevarla a una forma triangular superior. Finalmente, la definición de radio espectral es la siguiente:

Definición 2. El radio espectral de una matriz A es el valor propio de mayor magnitud, es decir:

$$\rho(A) = \text{máx}\{|\lambda| \in \mathbb{R} : \lambda \text{ es valor propio de } A\}.$$

Con el objetivo de presentar una lectura más fluida, se omitirán las demostraciones respectivas de los resultados teóricos extraídos de [1] que se presentarán a continuación. Debido a que las matrices que tomaremos en cuenta son no negativas e irreducibles, con el fin de facilitar el análisis, es pertinente dar una caracterización de las mismas para relacionarlas con la teoría matemática aplicada a modelos epidemiológicos. En este sentido, el siguiente teorema establece las características del radio espectral de las matrices antes mencionadas.

Teorema 1. Teorema de Perron-Frobenius

Sea P una matriz irreducible no-negativa. Entonces:

- El radio espectral $\rho(P)$ de P es un valor propio simple y positivo de P con eigenvectores izquierdo y derecho únicos y positivos salvo reescalamientos.
- El radio espectral es el único eigenvalor con un eigenvector no-negativo.
- El radio espectral de la matriz P se incrementa/reduce estrictamente si cualquiera de sus entradas se incrementan/reducen.

Se entiende que una matriz A no negativa es **primitiva** si y sólo si existe un único valor propio que alcanza el valor del radio espectral $\rho(A)$.

Antes de enunciar el Teorema Fundamental de la Demografía, supongamos que el tiempo se mide en unidades de generación y se denomina a x_k al tamaño de la población en la

generación k con $k = 1, 2, 3, \dots$, entonces el primer teorema relevante aplicado a dinámica de poblaciones y su comportamiento se enuncia a continuación.

Teorema 2. Teorema Fundamental de la Demografía.

Sea $P \in \mathbb{R}^{n \times n}$ una matriz de proyección para un modelo de poblaciones dinámicas x_k dado por:

$$x_k = Px_{k-1}, k = 1, 2, \dots$$

donde x_0, x_1, x_2, \dots es una secuencia de vectores no negativos de longitud fija. Suponga que P es una matriz primitiva con radio espectral $\rho(P) = r$ y tiene eigenvectores izquierdo, v^T , y derecho u , normalizados tales que $v^T u = 1$. Entonces:

$$\lim_{k \rightarrow \infty} \frac{x_k}{r^k} = (v^T x_0)u$$

De manera que, si $|w|$ denota la suma de las entradas de un vector w , entonces $|x_k|$ denotará la población total al tiempo k en el modelo poblacional, entonces:

$$\lim_{k \rightarrow \infty} |x_k| = \begin{cases} 0 & \text{si } r < 1 \\ |(v^T x_0)u| & \text{si } r = 1 \\ \infty & \text{si } r > 1. \end{cases}$$

Para establecer la abundancia de la población en la generación $k - 1$ basta conocer la sobrevivencia y fertilidad en $k - 1$ (dicha información está contenida en P), de tal forma que se tiene una proyección de la población en k . La importancia del teorema anterior radica en que el comportamiento de una población básicamente dependerá del valor del radio espectral de la matriz asociada P . De tal forma que obtenemos un umbral que nos permite determinar si una población crecerá sin cota, si se mantendrá en un valor determinado o se extinguirá conforme el tiempo avanza, es decir, en un modelo poblacional, el radio espectral $\rho(P) = r$ denota la tasa de crecimiento. De acuerdo a la teoría de Perrón-Frobenius, se estudia el radio espectral en un contexto demográfico como una medida de éxito reproductivo individual en ecología de poblaciones, que traducido en un modelo epidemiológico basado en un modelo matricial discreto, es el número de infecciones secundarias generadas por un individuo infeccioso conocido como $\rho(P) = R_0$.

Criterios para el brote de una enfermedad

Con el fin de extender los resultados anteriores al caso continuo, basta con definir al vector $x(t)$ que depende continuamente de t y sólo contiene compartimientos infecciosos, y un sistema de ecuaciones diferenciales de la forma:

$$\dot{x} = H(x).$$

Además, si $H(x)$ puede expresarse como $H(x) = F(x) - Mx$, con M una matriz no singular, donde F está conformada por las tasas a las cuales las infecciones secundarias se

incrementan en el i -ésimo compartimento infeccioso, y M por las tasas a la cual la progresión de la enfermedad decrece así como la muerte en el i -ésimo compartimento. El sistema anterior puede reescribirse de la forma:

$$\dot{x} = F(x) - Mx.$$

Para determinar el comportamiento del sistema es importante calcular en primera instancia los puntos de equilibrio del mismo, esto es, encontrar los vectores x para los cuales:

$$\dot{x} = F(x) - Mx = 0, \tag{1}$$

o lo que es lo mismo, obtener los vectores tales que:

$$x = M^{-1}F(x), \tag{2}$$

de manera que obtener las raíces del sistema (1) es equivalente a encontrar los puntos fijos del sistema (2). Con respecto al análisis de estabilidad de los puntos de equilibrio calculados, es posible probar que los valores propios del jacobiano del sistema (1), dado por $DF(x) - M$, contienen información equivalente respecto a la estabilidad de los puntos de equilibrio:

$$\Phi(x) = M^{-1}F(x).$$

Una vez evaluados los puntos de equilibrio x^* en el jacobiano de (1), empleando la definición de cota espectral de una matriz A como el número real

$$s(A) = \text{máx}\{Re(\lambda) : \lambda \text{ es valor propio de } A\}$$

y utilizando el teorema 2 de [2]:

Teorema 3. Sean T y V matrices cuadradas con entradas reales tales que T es no negativa, V es una matriz diagonal con elementos positivos y las entradas $(T - V)_{ij} \geq 0$ para cualesquiera $\{i, j\} \in \{1, \dots, n\}$ con $i \neq j$. Entonces: $s(T - V) < 0$ si y sólo si $\rho(V^{-1}T) < 1$.

Si $T = DF(x^*)$ y $V = M$ luego tenemos que $s(DF(x^*) - M) < 0$ si y sólo si el radio espectral $\rho(M^{-1}DF(x^*)) < 1$. Por lo tanto el punto de equilibrio x^* es asintóticamente estable si $\rho(M^{-1}DF(x^*)) < 1$ e inestable si $\rho(M^{-1}DF(x^*)) > 1$.

En el caso de una enfermedad, si x^{PLE} es un punto de equilibrio libre de la enfermedad (si todos los miembros de la población son susceptibles). De acuerdo a [3], el número básico reproductivo se define como $R_0 = \rho(M^{-1}DF(x^{PLE}))$ y aplicando el criterio obtenido anteriormente podemos conocer cómo se comportará la enfermedad en una vecindad cercana al punto x^{PLE} tal que:

- Si $R_0 < 1$ entonces no existirá un brote epidémico.
- Si $R_0 > 1$ entonces existirá un brote epidémico.

Es decir, el cálculo de R_0 está basado en la linealización del sistema de ecuaciones alrededor del equilibrio libre de enfermedad.

Ejemplos de modelos epidemiológicos.

Un modelo epidemiológico de compartimentos puede definirse a partir de las clases de individuos en que se puede dividir una población afectada por una enfermedad. Las tasas de transición son estimadas del conocimiento cualitativo y evidente en la historia natural de la enfermedad. Cuando el modelo ha sido establecido y las clases epidemiológicas y tasas de transición están definidas, la dinámica de la infección en la población puede ser expresada matemáticamente por un sistema de ecuaciones diferenciales.

Ejemplo 1. En muchas enfermedades los individuos afectados no pasan directamente de la clase susceptible a la infecciosa. Puede existir una etapa latente o de exposición a la enfermedad y por ello el período de incubación se define como el intervalo de tiempo entre el momento de adquisición de la enfermedad y el momento en que aparecen los primeros síntomas, estas características corresponden al modelo *SEIR* (Susceptibles-Expuestos-Infectados-Recuperados).

$$\begin{aligned}\frac{dS}{dt} &= \Lambda - \beta S \frac{I}{N} - \mu S, \\ \frac{dE}{dt} &= \beta S \frac{I}{N} - (\mu + k)E, \\ \frac{dI}{dt} &= kE - (\gamma + \mu)I, \\ \frac{dR}{dt} &= \gamma I - \mu R.\end{aligned}$$

donde Λ es la tasa de nacimiento, μ la tasa de muerte natural, β la tasa de infección, γ la tasa de recuperación y k la tasa de transición del estado latente a infeccioso.

El punto de equilibrio libre de infección es $E_0 = (\frac{\Lambda}{\mu}, 0, 0, 0)$. En este caso $X = (E, I)^T$ y así el sistema a analizar se reduce:

$$\dot{X} = FX - NX = \begin{pmatrix} 0 & \beta \\ k & 0 \end{pmatrix} \begin{pmatrix} E \\ I \end{pmatrix} - \begin{pmatrix} \mu + k & 0 \\ 0 & \mu + \gamma \end{pmatrix} \begin{pmatrix} E \\ I \end{pmatrix}.$$

Ahora calculamos el eigenvalor dominante de MN^{-1} evaluado en el punto de equilibrio libre de la infección donde

$$N^{-1} = \begin{pmatrix} \frac{1}{\mu+k} & 0 \\ 0 & \frac{1}{\mu+\gamma} \end{pmatrix}$$

los eigenvalores son $\lambda_1 = \sqrt{\frac{k\beta}{(\mu+k)(\mu+\gamma)}}$ y $\lambda_2 = -\sqrt{\frac{k\beta}{(\mu+k)(\mu+\gamma)}}$. Por lo tanto:

$$R_0 = \sqrt{\frac{k\beta}{(\mu+k)(\mu+\gamma)}}.$$

la expresión obtenida para R_0 representa los dos estados de infección, los que están en estado de latencia y los infecciosos.

Ejemplo 2. En enfermedades infecciosas como la influenza o el paludismo, se presentan diferentes cepas o subtipos de virus que compiten por el mismo recurso con el fin de colonizar al hospedero. En este sentido se presenta el siguiente modelo con dos tipos de infectados,

$$\begin{aligned}\frac{dS}{dt} &= \Lambda - \left(\beta_1 \frac{I_1}{N} + \beta_2 \frac{I_2}{N} \right) S - \mu S, \\ \frac{dI_1}{dt} &= \beta_1 S \frac{I_1}{N} - (\mu + k_1) I_1 - \beta_2 I_1 \frac{I_2}{N}, \\ \frac{dI_2}{dt} &= \beta_2 S \frac{I_2}{N} - (\mu + k_2) I_2 + \beta_2 I_1 \frac{I_2}{N}, \\ \frac{dR}{dt} &= k_1 I_1 + k_2 I_2 - \mu R.\end{aligned}$$

La población total la denotamos como $N = S + I_1 + I_2 + R$, así el punto de equilibrio libre de la enfermedad es $E_0 = (\frac{\Lambda}{\mu}, 0, 0, 0)$, ya que al no presentarse una infección $S \approx \frac{\Lambda}{\mu}$. El vector X está conformado por los compartimientos que generan la infección, $X = (I_1, I_2)^T$, así \dot{X} está descrito como

$$\dot{X} = FX - NX = \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} - \begin{pmatrix} \mu + k_1 & 0 \\ 0 & \mu + k_2 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}$$

luego el eigenvalor dominante esta dado como $R_0 = \max\{\frac{\beta_1}{\mu+k_1}, \frac{\beta_2}{\mu+k_2}\}$, donde β_1 y β_2 son las tasas de infección respectivas de cada cepa, y en función de la que sea competitivamente superior es la que generará el mayor número de casos secundarios de infección.

Referencias

- [1] Li C.K., Schneider H., "Applications of Perron-Frobenius theory to population dynamics", *J. Mathematical Biology* 44 pp. 450-462, 2002
- [2] Velasco-Hernández, J.X. "Sobre enfermedades infecciosas", *Miscelánea Matemática* 29 pp. 51-72, 1999.
- [3] Driessche P., Watmough J., "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission", *Mathematical Biosciences* 180 pp. 29-48, 2002

El problema del matrimonio estable¹

Ilan Jinich Fainsod

Ex-alumno de Matemáticas Aplicadas y Actuaría del ITAM

“Matchmaker, matchmaker, make me a match! Find me a find, catch me a catch! Night after night in the dark I’m alone, so find me a match of my own!”

En el pequeño Shtetl de Anatevka, había muchos hombres y muchas mujeres jóvenes con el deseo de casarse. Yenta, la casamentera del pueblo, tenía el trabajo y la obligación de emparejar a todos estos jóvenes. Esta tarea suena fácil pero no lo es, cada joven tenía distintas preferencias y cada uno de ellos necesitaba un trato especial, “el matrimonio es para toda la vida y no es algo que hay que tomar como si fuera cualquier cosa” decía Yenta. Retomando los conceptos de la última sección, se desea que el emparejamiento de estos jóvenes sea estable y óptimo si es que es posible.

Para definir las preferencias de una forma un poco más formal, hacemos uso de la matriz de preferencias.

Definición 1. Definimos la **matriz de preferencias** para un problema con n hombres y con m mujeres como una matriz con n filas y m columnas donde la primera entrada de la celda (i, j) representa el orden de preferencia que le asigna el hombre i a la mujer j y análogamente la segunda entrada representa el orden de preferencia que le da la mujer j al hombre i . La matriz podría tener lugares vacíos en caso de listas no exhaustivas.

Para que dejar más clara la definición vale la pena mostrar un ejemplo.

Ejemplo 1. Supongamos que contamos con tres hombres α , β y γ y con tres mujeres A , B y C y la matriz de preferencias

$$\begin{pmatrix} & A & B & C \\ \alpha & 1,3 & 2,2 & 3,1 \\ \beta & 3,1 & 1,3 & 2,2 \\ \gamma & 2,2 & 3,1 & 1,3 \end{pmatrix},$$

aquí el orden preferencias de α es (A, B, C) y el orden de preferencias de C es (γ, β, α) .

Para continuar, introducimos un criterio de estabilidad y otro de optimalidad, y para ello enunciamos dos definiciones.

Definición 2. [2] Decimos que un emparejamiento es **inestable** si existen dos hombres, α y β casados con mujeres A y B respectivamente, con la propiedad de que α prefiere estar con B que con A y B prefiere estar con α que con β . Es decir, existe una mujer y un hombre que se prefieren entre ellos a sus respectivas parejas.

Alternativamente un emparejamiento es **estable** si no es inestable.

¹Nota: Este problema hace el supuesto de que las relaciones solo pueden ser entre un hombre y una mujer y que esto es reflejado por sus preferencias. Esto no manifiesta la realidad o las opiniones del autor y se hace únicamente con fines matemáticos.

Definición 3. [2] Un emparejamiento es considerado **óptimo** si cada hombre esta mejor o igual respecto a sus preferencias que en cualquier otra asignación estable.

A continuación mostramos un par de ejemplos de cómo funcionan las definiciones para dejar todo más claro.

Ejemplo 2. Retomando el ejemplo 1.

El conjunto de parejas (α, A) , (β, B) y (γ, C) es estable porque a pesar de las preferencias de las mujeres cada hombre esta con su primera opción, es decir, los hombres prefieren a su mujer más que a cualquier otra. Ésta además es óptima porque cada hombre está casado con su primera opción.

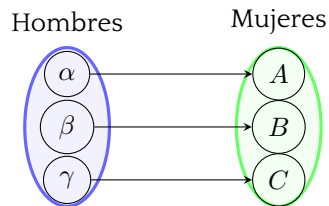


Figura 1: Emparejamiento estable.

Ejemplo 3. Retomando el ejemplo 1.

El conjunto de parejas (α, A) , (γ, B) y (β, C) es inestable porque γ prefiere a A que a su pareja actual (B) y A prefiere a γ que a su esposo actual (α) .

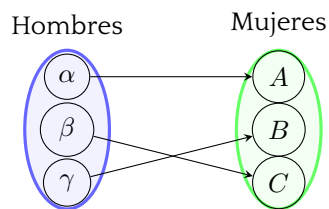


Figura 2: Emparejamiento inestable.

Un supuesto adicional que haremos es que el número de hombres es exactamente igual al número de mujeres. Se verá después que es relativamente sencillo generalizar al caso en el que el número de hombres y de mujeres no es igual.

Algoritmo de Gale Shapley

En 1962 David Gale y Lloyd Stowell Shapley demostraron que para el problema del matrimonio estable siempre se puede encontrar un emparejamiento estable óptimo para los

hombres, esto se hizo enunciando un algoritmo y mostrando su convergencia. El algoritmo queda representado por el siguiente código y es conocido como el algoritmo de Gale Shapley.

Algoritmo 1: Gale Shapley

input : Una matriz de preferencias para n hombres y n mujeres

output: Un emparejamiento.

- 1 Cada hombre le propone a la primera mujer de su lista; Cada mujer que recibe más de una propuesta acepta la que este más arriba en su lista y rechaza al resto ;
- repeat**
- 2 Los hombres no emparejados le proponen a la siguiente mujer en su lista; Cada mujer que recibe una propuesta escoge la que está más arriba en su lista entre las propuestas que recibió y su pareja actual; Las mujeres rechazan las propuestas que no aceptaron;
- 3 **until** hasta que todos los hombres tengan pareja;

Ejemplo 4. [2] Supongamos que para 4 hombres y 4 mujeres tenemos la matriz de preferencias

$$\begin{pmatrix} & A & B & C & D \\ \alpha & 1,3 & 2,2 & 3,1 & 4,3 \\ \beta & 1,4 & 2,3 & 3,2 & 4,4 \\ \gamma & 3,1 & 1,4 & 2,3 & 4,2 \\ \delta & 2,2 & 3,1 & 1,4 & 4,1 \end{pmatrix}.$$

Después de aplicar diez iteraciones (13 propuestas) del algoritmo llegamos al siguiente emparejamiento:

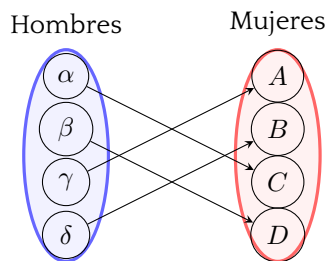


Figura 3: Resultado del algoritmo.

Lo importante de este algoritmo no es sólo que produce un emparejamiento, adicional a esto, como lo muestra el siguiente teorema, el emparejamiento resultante es estable.

Teorema 4 (Teorema de Gale Shaley). [2]

El algoritmo de Gale Shapley termina en un emparejamiento estable.

Demostración. Supongamos que el emparejamiento producido por el algoritmo no es estable. Esto es, un hombre α y una mujer A se prefieren entre ellos que a sus respectivas

parejas.

Como α prefiere a A más que a su esposa entonces, α le propuso matrimonio primero a A que a su propia esposa. Además, como A prefiere a α que a su esposo entonces, A hubiera rechazado a su esposo y se hubiera quedado casada con α lo cual es una contradicción. Por lo tanto, el algoritmo de Gale Shapley termina siempre en un emparejamiento estable. \square

Un resultado inmediato de esto es que siempre, sin importar como sean las preferencias, existe un emparejamiento estable.

Corolario 5. *Dada una matriz de preferencias arbitraria existe un emparejamiento estable.*

Demostración. Si aplicamos el algoritmo de Gale Shapley, sabemos por el teorema 4 que éste siempre acaba en un emparejamiento estable. Por lo tanto, siempre existe un emparejamiento estable. \square

Una vez que sabemos que el algoritmo siempre converge, nos interesa conocer qué tan rápido lo hace y si tiene algunas ventajas el emparejamiento resultante. El lema 6 nos ayuda a llegar a estos resultados y los corolarios 7 y 8 nos dan una idea de qué tan bueno es el algoritmo.

Lema 6. [5]

Bajo el emparejamiento obtenido por Gale Shapley, sólo un hombre puede terminar con la última mujer de su lista como pareja.

Demostración. Supongamos que en el emparejamiento de Gale Shapley m ($m \geq 2$) hombres terminan con la última mujer de su lista como pareja, eso significa que cada uno de esos m hombres invitó a salir a todas las mujeres. Entonces cada mujer fue invitada a salir por lo menos m veces, lo cual es una contradicción porque el algoritmo acaba cuando invitan a salir a la última mujer y a ésta sólo la invitan a salir una vez. Por lo tanto, sólo un hombre puede terminar con la última mujer de su lista como pareja. \square

Dos consecuencias casi inmediatas de esto son que bajo el algoritmo la gran mayoría de los hombres no terminan con su última opción y que el algoritmo converge relativamente rápido.

Corolario 7. *Si en un emparejamiento estable por lo menos dos hombres están emparejados con la última mujer de sus respectivas listas, entonces existen dos o más emparejamientos estables en el problema.*

Demostración. Llamemos m al emparejamiento estable donde por lo menos dos hombres están emparejados con la última mujer de sus respectivas listas y llamemos m' al emparejamiento de Gale Shapley. Por el teorema 4 sabemos que el algoritmo de Gale Shapley siempre acaba en un emparejamiento estable y además en ese emparejamiento estable sólo un hombre puede terminar con la última mujer de su lista como pareja por el lema 6. Por lo tanto m y m' son diferentes y como ambos son emparejamientos estables entonces el número de emparejamientos estables en el problema es mayor o igual a dos. \square

Corolario 8. *El número máximo de propuestas en el algoritmo es $n^2 - n + 1$*

Demostración. Por el lema 6 sabemos que a lo más un hombre acaba emparejado con la última mujer de su lista, por lo tanto, el peor emparejamiento posible para el algoritmo es uno donde $n - 1$ hombres terminan con la penúltima mujer de sus respectivas listas y un hombre termina con la última. Para llegar a esta situación, de acuerdo con el algoritmo, los $n - 1$ hombres deben de realizar $n - 1$ propuestas cada uno y el otro hombre debe de realizar n propuestas. Esto es $(n - 1)(n - 1) + n$ propuestas, que es igual a $n^2 - n + 1$ propuestas. \square

Observación 9. La complejidad del algoritmo de Gale Shapley es del orden de n^2 .

A partir de esto ya podemos resolver el problema en el que la cantidad de hombres y de mujeres es distinta. El siguiente resultado muestra que sin importar el número de hombres o de mujeres, el algoritmo igual converge.

Teorema 10. [2]

Dada una matriz de preferencias arbitraria con n hombres y m mujeres, el algoritmo de Gale Shapley converge a un emparejamiento estable

Demostración. Supongamos que la cantidad de hombres es más chica que la cantidad de mujeres ($n < m$), en este caso el algoritmo acaba cuando n de las m mujeres reciben una propuesta. Si suponemos que la cantidad de hombres es mayor a la cantidad de mujeres ($n > m$), en este caso el algoritmo acaba después de que $n - m$ hombres son rechazados por todas las mujeres y las propuestas de m de los hombres son aceptadas.

Además de forma análoga al teorema 4 se puede ver que el emparejamiento producido es estable. \square

El problema de admisión a universidades

Supongamos que en una ciudad hay n personas que desean entrar a m universidades, los solicitantes tienen una lista de preferencias en donde reflejan a qué universidades prefieren entrar y de forma análoga las universidades tienen una lista de preferencias con la información de a quién prefieren admitir. Adicional a esto, para cada universidad existe un número máximo de alumnos que pueden admitir, esta restricción es natural porque la cantidad de personal y de espacio en las universidades es limitado. Es claro que esto es una generalización del problema del matrimonio estable. Podemos retomar las definiciones la matriz de preferencias, un emparejamiento estable y de un emparejamiento óptimo como análogas a las definiciones 1, 2 y 3.

En el siguiente código exhibimos un análogo al algoritmo de Gale-Shapley para este problema, el cual encuentra un emparejamiento estable y que además lo hace relativamente rápido. Para el caso general este algoritmo es conocido como el de "aceptación diferida".

Algoritmo 2: Gale Shapley para admisión a universidades

input : Una matriz de preferencias para n solicitantes y m universidades, un vector M en donde la entrada j representa la cota superior de la universidad j .

output: Un emparejamiento.

- 1 Cada solicitante aplica a la primera universidad de su lista; Cada universidad que recibe más de solicitudes a su cota superior acepta las que solicitudes que están más arriba en su lista y rechaza al resto. ; **repeat**
- 2

Los solicitantes no emparejados solicitan entrar a la siguiente universidad en su lista; Cada universidad que recibe alguna solicitud acepta hasta M_j solicitantes de los primeros de su lista entre los que aplicaron a ella y sus admitidos actuales; Las universidades rechazan a los alumnos no aceptados;

- 3 **until** hasta que cada uno de los solicitantes sea admitido por alguna universidad o rechazado por todas;

Este algoritmo, al igual que su análogo en el matrimonio estable, encuentra siempre un emparejamiento estable. La demostración se deja para el lector y es casi análoga al caso del matrimonio estable.

Además de producir un emparejamiento estable, en este algoritmo cada solicitante está mejor o igual en el emparejamiento resultante que en cualquier otro emparejamiento estable. Para hacer la demostración primero introducimos una definición y un lema.

Definición 11. [2] Decimos que una universidad es **posible** para un aplicante si existe una asignación estable en la que esta persona asiste a esa universidad.

Lema 12. [2]

Supongamos que en un paso arbitrario del algoritmo ningún estudiante ha sido rechazado por una universidad posible para él, además supongamos que una universidad A después de llenarse recibiendo a los estudiantes $\beta_1, \beta_2, \dots, \beta_q$ rechaza a α , entonces A no es posible para α .

Demostración. Sabemos por hipótesis que para toda $i = 1, \dots, q$, β_i prefiere a A que a todas las universidades que no lo han rechazado y además que cualquier universidad que lo rechazó previamente no es posible para él. Supongamos que existe un emparejamiento estable en el que α asiste a A , entonces alguna β_i no asiste a A porque α toma su lugar. Este emparejamiento es inestable porque β_i prefiere a A que a su asignación actual porque A es su mejor asignación posible y A prefiere tener a β_i que a α , lo cual es claramente una contradicción y por lo tanto A no es posible para α . \square

Teorema 13. [2]

Dada una matriz de preferencias arbitraria y un vector de cotas superiores arbitrario, el emparejamiento producido por el algoritmo es óptimo para los solicitantes.

Demostración. La prueba es por inducción. Primero que nada, sabemos que si en el primer paso del algoritmo una universidad rechaza a un alumno es porque esta universidad no es

posible para el aplicante, si suponemos que esta universidad es posible para él llegamos rápidamente a una contradicción porque esto querría decir que la universidad rechazó a un mejor estudiante que la tenía como primera opción para meterlo a él y por lo tanto el emparejamiento no sería estable. Luego por el lema 12 sabemos que en los siguientes pasos del algoritmo ningún estudiante es rechazado por una universidad posible para él y por lo tanto el emparejamiento obtenido es óptimo. \square

Conclusión

Lo más padre de las matemáticas es que podemos empezar con un problema sencillo que parece fácil de resolver y complicarlo de maneras espectaculares. En particular, si al lector le interesa más sobre el tema, existen muchas variantes de este problema significativamente más complicadas, muchas de ellas NP-Complejas (a diferencia de este problema). A partir de esto, el autor de este artículo recomienda al lector tomar un problema de un musical, una obra, una novela o una película y ver cuáles son las matemáticas detrás de éste y ver cómo hacerle para complicarlo y volverlo divertido.

Referencias

- [1] Biró, P., Fleiner, T., Irving, R. W., y Manlove, D. F. (2010). The college admissions problem with lower and common quotas. *Theoretical Computer Science*, 411 (34), 3136 - 3153.
- [2] Gale, D., y Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69 (1), 9-15.
- [3] Gale, D., y Sotomayor, M. (1985). Some remarks on the stable matching problem. *Discrete Applied Mathematics*, 11 (3), 223 - 232.
- [4] Gusfield, D., y Irving, R. W. (1989). *The stable marriage problem: Structure and algorithms*. Cambridge, MA, USA: MIT Press.
- [5] Knuth, D. E. (1997). *Stable marriage and it's relation to other combinatorial problems: An introduction to the mathematical analysis of algorithms* (Vol. 10). American Mathematical Society.
- [6] Roth, A. E. (2014). *Who gets what and why*. Eamon Dolan/Houghton Mifflin Harcourt.
- [7] Roth, A. E., y Sotomayor, M. A. O. (1990). *Two-sided matching: A study in game-theoretic modeling and analysis*. Cambridge University Press.

El Método de Conjunto Activo para Problemas de Optimización Cuadráticos

Adrian Tame Jacobo

Estudiante de Matemáticas Aplicadas del ITAM

Introducción

El mundo de la optimización empieza mucho antes que el nacimiento de las computadoras, con su concepto moderno introducido por Alan Turing en su ya famoso artículo [1]. Fermat y Lagrange son generalmente acreditados por presentar los primeros resultados importantes, utilizando resultados de Cálculo para poder identificar puntos óptimos de un problema, mientras que Newton y Gauss crearon métodos iterativos para poder encontrar una solución a problema de optimización. El método de Newton, una base central de muchos algoritmos de optimización, es presentado a detalle en [2].

Hoy en día, los problemas de optimización que se intentan resolver son mucho más complejos que hace tiempo, y se necesitan nuevas herramientas para poder resolverlos. Una de las formas interesantes que aparecen en problemas de optimización es el problema cuadrático, definido formalmente como

$$\begin{aligned} \underset{x}{\text{minimizar}} \quad & q(x) = \frac{1}{2}x^T Qx + x^T c \\ \text{sujeto a} \quad & a_i^T x = b_i, \quad i \in \mathcal{E} \\ & a_i^T x \geq b_i, \quad i \in \mathcal{I} \end{aligned} \tag{1}$$

con Q una matriz simétrica de $n \times n$, \mathcal{I}, \mathcal{E} conjuntos finitos de índices, y $x, c, \{a_i\}, i \in \mathcal{I} \cup \mathcal{E}$ vectores en \mathbb{R}^n [3]. Es de notar que las restricciones de este problema son puramente lineales. En la siguiente sección, se explora este problema y sus varias aplicaciones.

Programación Cuadrática ¹

El problema (1) es uno muy conocido y estudiado en el área de investigación de operaciones. Es también un subproblema importante en los algoritmos de programación cuadrática secuencial, métodos de Lagrange aumentados, y métodos de punto interior. En general, muchas de las aplicaciones de este problema caen en las áreas de finanzas, agricultura, producción y operaciones, economía y marketing.

Probablemente una de las razones por las cuales este problema es tan conocido es porque es una de las formas más simples de problemas no lineales. Los problemas de esta forma siempre pueden ser resueltos en una cantidad finita de iteraciones de un método, o en el caso contrario se puede demostrar que no tienen soluciones factibles. Se nota que el método de conjunto activo que se expone aquí no garantiza convergencia al problema general, pero sí a ciertos subcasos del problema, como por ejemplo el caso en el que solamente se

¹La mayoría de esta información viene de [3] y [4].

tienen restricciones de desigualdad.

Por completitud, y porque es un tema bastante interesante, continúan algunos casos detallados donde el problema se modela como un problema de la forma de (1).

- Markowitz [5] desarrolló aplicaciones importantes en finanzas basadas en modelos cuadráticos. Estos son para modelar portafolios de inversión sujetos a algún riesgo.
- Hay algunas aplicaciones interesantes a optimización de formas geométricas para optimizar el espacio utilizado al cortar diamantes, y gastar la menor cantidad posible de materia prima al cortar.
- El modelo de cambio de marca basado en procesos de Markov utiliza estos métodos. Theil y Ray desarrollaron un modelo para el cambio de un cliente a otro y se estima con un modelo cuadrático.
- Louwes, Boot y Wage desarrollaron un modelo para el uso óptimo de la leche en Países Bajos. Este modelo es usado en muchas aplicaciones de agricultura.
- Otra aplicación interesante es la de separación lineal. Se puede modelar un problema de separación lineal como la distancia en norma de un vector a un hiper-plano lineal, y minimizar las distancias para generar un hiper-plano que separe mejor los datos. La siguiente imagen es una implementación de esto al problema de Iris [6] con solamente 3 variables consideradas, que son la longitud del sépalos, ancho del sépalos y ancho del pétalo.

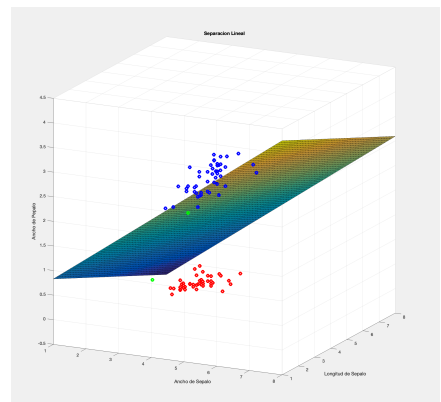
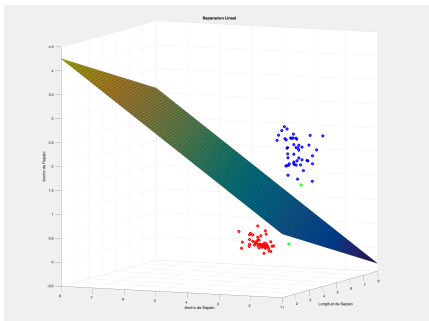


Figura 1: Separación lineal del conjunto de datos Iris.










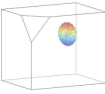
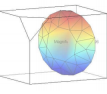
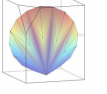



BFS	LP Solution	NLP Solution
<p>Vertical Orientation</p> 	 <p>Magnification = 3.78</p>	 <p>Magnification = 4.15</p>
<p>Oriented along diagonal between corner points</p> 	 <p>Magnification = 3.53</p>	 <p>Magnification = 4.15</p>
	 <p>Magnification = 3.53</p>	 <p>Magnification = 4.15</p>
<p>Oriented along diagonal between edge mid-points</p> 	 <p>Magnification = 4.00</p>	 <p>Magnification = 4.15</p>
	 <p>Magnification = 4.05</p>	 <p>Magnification = 4.15</p>

Figura 2: Corte de diamantes óptimo, tomado de [7]

El Método de Conjunto Activo

En principio, lo que hace este método es tomar un punto x , y checar qué condiciones de desigualdad se cumplen en ese punto de manera estricta. Esto es algo interesante, ya que el algoritmo eventualmente (en el caso de convergencia) convergerá a un punto, el cual nos puede decir mucho de la naturaleza del problema a resolver. Por ejemplo, saber cuáles restricciones están activas nos permite saber cuáles de ellas son las de mayor importancia o juegan un papel más importante en la minimización del problema.

El método resuelve un subproblema (este se puede ver de manera clara en el pseudocódigo, dado por (2)). En general, se trabaja en un conjunto convexo ya que las restricciones están dadas por igualdades y desigualdades. También, una condición importante para poder resolver el problema es que Q sea una matriz positiva definida y simétrica, pero en algunos casos con matrices indefinidas también se puede resolver, sin embargo, esto está fuera del alcance de este artículo.

Pseudocódigo

A continuación se presenta el pseudocódigo para resolver el problema (1) utilizando el método de conjunto activo. Este algoritmo está fuertemente basado en el algoritmo de conjunto activo presentado en [3]. Además, esta implementación considera $\mathcal{E} = \emptyset$.

Result: x un vector solución de tamaño n y μ un vector con los multiplicadores de Lagrange asociados.

Inicialización: Sea Q una matriz de tamaño $n \times n$ simétrica, positiva definida;

Sea c un vector de tamaño n ;

Sea μ un vector de tamaño n el vector de multiplicadores de lagrange completo ;

Sea λ un vector de tamaño cero inicialmente (\emptyset);

Sea F una matriz de coeficientes de restricciones de tamaño $m \times n$ con rango(F) = $m \leq n$;

Sea d un vector de valores de restricciones de tamaño m ;

while iteraciones no excedan un limite arbitrario **do**

Sea $g_k = Qx_k + c$;

Sea \mathcal{W}_k el conjunto de condiciones activas en el punto x_k ;

resolver el subproblema dado por:

$$\begin{aligned} & \underset{p}{\text{minimizar}} && \frac{1}{2} p_k^T Q p_k + g_k^T p_k \\ & \text{sujeto a} && a_i^T p_k = 0, \quad i \in \mathcal{W}_k \end{aligned} \tag{2}$$

if $\|p_k\|_2 = 0$ **then**

if $\lambda \leq 0$ **then**

END;

else

$j = \underset{j}{\text{argmin}} \{j \in I \cap \mathcal{W}_k, \hat{\mu}_j\}$;

 Actualizar x_k a x_{k+1} ;

$\mathcal{W}_{k+1} = \mathcal{W}_k - \{j\}$

end

else

 Activos = Todas las condiciones que cumplan que $Fx_k = d$;

$\mathcal{W}_c = \text{Activos} - \mathcal{W}_k$;

if Hay una condición que cumpla que $F(\mathcal{W}_k)p = 0$ **then**

$\mathcal{V} = \mathcal{W}_c [F(\mathcal{W}_c)] \cdot p_k = 0$;

$j = \min \{F(\mathcal{W}_c) \cdot p_k = 0\}$;

$\mathcal{W}_k = \mathcal{W}_k \cup \mathcal{V}_j$;

else

 Hacer búsqueda de línea con condiciones de Wolfe para encontrar k ;

$x_{k+1} = x_k + \alpha_k p_k$;

end

end

$\mu(\text{Activas}) = -\lambda$;

$\mu(\mathcal{W}_c) = 0$;

end

Referencias

- [1] A. M. Turing; On Computable Numbers, with an Application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society*, Volumen s2-42, Número 1, 1° de Enero de 1937, Páginas 230–265, <https://doi.org/10.1112/plms/s2-42.1.230>
- [2] Polyak, Boris. (2007). Newton's method and its use in optimization. *European Journal of Operational Research*. 181. 1086-1096. <https://doi.org/10.1016/j.ejor.2005.06.076>.
- [3] Nocedal, J. and Wright, S.J. (1999) Numerical Optimization. Springer, New York. <https://doi.org/10.1007/b98874>
- [4] Omprakash K. Gupta (1995) Applications of Quadratic Programming, *Journal of Information and Optimization Sciences*, 16:1, 177-194, <https://doi.org/10.1080/02522667.1995.10699213>
- [5] H. M. Markowitz (1952), Portfolio selection, *Journal of Finance*, Vol. 7, pp. 77-91. <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- [6] <http://archive.ics.uci.edu/ml/datasets/iris>
- [7] Viswambharan, Ambili. (1998). *Optimisation in Diamond Cutting*.

Un problema novohispano

Saúl Adrián Álvarez Tapia
Estudiante de Matemáticas Aplicadas del ITAM

Introducción

Una de las características que diferencian a las matemáticas de cualquier otra disciplina en el ámbito humano es su atemporalidad. En las humanidades, lo aceptado cambia de acuerdo a cómo cambia la sociedad. En las ciencias, el consenso cambia cuando se encuentran evidencias nuevas (por ejemplo, cuando se dio el cambio del modelo geocéntrico al heliocéntrico). Pero las matemáticas poseen verdades eternas: verdades que una vez son demostradas, no pueden cambiar ni aunque viviésemos en un universo distinto.

Debido a esto, resulta interesante buscar problemas antiguos y aplicar métodos nuevos, de forma que encontremos procedimientos más eficientes; pero también para poder valorar el hito que supone haber resuelto el problema sin las herramientas existentes hoy en día.

El problema que analizaremos en este artículo surgió en el siglo XVI, en el centro de la Nueva España. El problema lo presentó la doctora María del Carmen Jorge y Jorge el 19 de octubre del 2016, en una plática llamada “Dos códigos texcocanos del siglo XVI”, con motivo de la vigesimosexta semana de matemáticas.

El problema consiste en que en dos códigos texcocanos (Vergara y Santa María Asunción) hay información sobre cientos de terrenos. Esta información viene en tripletas: primero una tupla ordenada l_1, l_2, \dots, l_n de números reales, que corresponde a la longitud de los lados del terreno (en ese orden); una área A , y el tipo de suelo del terreno (en los códigos también están dibujos con la forma aproximada del terreno). En los códigos, n podía llegar hasta 23, las longitudes estaban medidas en varas, y el área en varas cuadradas. El problema de los antiguos texcocanos correspondía en asociar a cada terreno descrito en los códigos un terreno físico.

En esa época, sólo se contaba con elementos muy simples para calcular áreas; tanto en herramientas matemáticas como en herramientas para medir físicamente los terrenos. Por lo tanto, era común subestimar o sobreestimar las áreas calculadas a partir de sólo las longitudes de los lados y los ángulos entre ellos. Por lo tanto, era posible obtener medidas de área que no podían existir para un terreno dado.

Dejemos la historia para otro momento y definamos el problema que trataremos de resolver en este artículo. Nosotros no contamos con dibujos de los terrenos, ni podemos corroborar si los terrenos en efecto existen. Más aún, como los códigos son antiguos, no tenemos acceso a ellos. Pero lo que sí podemos hacer es determinar un método que determine si dados l_1, \dots, l_n y A , existe un polígono en \mathbb{R}^2 que tenga estas medidas como sus lados y su área, respectivamente.

Para este problema, permitiremos que los polígonos sean degenerados en el sentido de que dos lados sean el mismo en un segmento, o de que tres vértices consecutivos sean colineales, pero no permitiremos que los lados se crucen.

Desarrollo

Antes de continuar, definamos un poco de notación. \mathcal{P} denotará el polígono cuyos lados miden l_1, \dots, l_n . \mathcal{P} tendrá vértices $v_1, \dots, v_n \in \mathbb{R}^2$, de forma que $\|v_i - v_{i+1}\| = l_i$, para $i = 1 \dots n - 1$, y $\|v_n - v_1\| = l_n$. También, $v_i = (x_i, y_i)$. Por otra parte, $p_{\mathcal{P}} = \sum l_i$ denotará el perímetro de \mathcal{P} , y $s_{\mathcal{P}} = p_{\mathcal{P}}/2$ su semiperímetro. También, a_i representa la arista correspondiente a l_i . A un polígono cuyos lados tengan longitud l_1, \dots, l_n lo llamaremos un *polígono bueno*.

Primero, obtendremos una condición necesaria y suficiente para que un polígono con lados l_1, \dots, l_n pueda existir, sin importar A .

Teorema 1. *Un polígono \mathcal{P} cuyos lados tienen longitud l_1, \dots, l_n , en ese orden, existe si y sólo si para toda k , $2l_k \leq \sum l_i$.*

Esta condición es una desigualdad del triángulo generalizada.

Demostración. Realicemos lo siguiente:

\implies

Considérese que los lados son tales que \mathcal{P} existe. Fijémonos en $v_1 v_2 v_n$, y sea $u = \|v_2 - v_n\|$. Por una parte, se debe satisfacer que el triángulo $v_1 v_2 v_n$ es un triángulo válido, así que debe satisfacer la desigualdad del triángulo, así que $2l_n = 2\|v_1 - v_n\| \leq \|v_1 - v_2\| + \|v_2 - v_n\| + \|v_n - v_1\| = l_1 + u + l_n$. Por otra parte, el polígono \mathcal{P}_1 definido por $v_2 \dots v_n$ existe, de forma que aplicando inducción, podemos ver que $2u \leq l_2 + \dots + l_{n-1}$, de forma que

$$\begin{aligned} 2l_n &\leq l_1 + u + l_n \\ &\leq l_1 + (l_2 + \dots + l_{n-1}) + l_n \\ &= \sum l_i. \end{aligned}$$

Por supuesto, nos falta justificar el caso base de esta inducción, pero este caso sucede con $n = 3$, que es simplemente la desigualdad del triángulo sobre el triángulo $v_1 v_2 v_3$.

Aplicando este razonamiento a todos los lados, tenemos la condición del teorema.

\impliedby

Considérese una tupla l_1, \dots, l_n que cumpla con la condición del Teorema 1. Entonces, construyamos \mathcal{P} cuyos lados tengan estas longitudes.

Si $n = 3$, es fácil de construir el triángulo adecuado, usando el teorema del coseno para obtener los ángulos del triángulo.

Para $n > 3$, constrúyase un triángulo con lados $l_1, l_2, p_{\mathcal{P}} - l_1 - l_2$. Luego, colóquese v_1 en la intersección de los lados con longitud l_1 y $p_{\mathcal{P}} - l_1 - l_2$; v_2 en la intersección de los lados con longitud l_1 y l_2 ; v_3 en la intersección del lado de longitud l_2 con el de longitud $p_{\mathcal{P}} - l_1 - l_2$; y por último, colóquese v_4, \dots, v_n sobre el lado de longitud $p_{\mathcal{P}} - l_1 - l_2$, de forma que se satisfaga el requisito de las distancias entre vértices.

De esta forma, obtenemos un polígono que, aunque degenerado, es válido según los criterios que establecimos. \square

Ahora, necesitamos ver que el problema que planteamos no es trivial; es decir, que para toda tupla de longitudes, existe un área que un polígono con esos lados no pueda tener. Es claro que existe una cota inferior para el área (0), pero nos gustaría ver si el ínfimo de ésta área es estrictamente positivo, y también nos gustaría ver que hay una cota superior para el área; y aún mejor, que esta cota se alcanza.

Definición 1 (Distancia entre polígonos de n lados). *Dados dos polígonos \mathcal{P} , \mathcal{Q} , con el mismo número de vértices, con vértices v_1, \dots, v_n y w_1, \dots, w_n respectivamente, se define la distancia entre ellos $d(\mathcal{P}, \mathcal{Q})$ como $\sum_{i=1}^n \|v_i - w_i\|^2$.*

Teorema 2. *El espacio de polígonos \mathcal{P} cuyos lados son l_1, \dots, l_n es completo bajo la distancia entre polígonos de n lados.*

Demostración. Considérese una sucesión de Cauchy $(\mathcal{P}_i)_{i \in \mathbb{N}}$ de polígonos con lados l_1, \dots, l_n . Sean $v_{i,1}, \dots, v_{i,n}$ los vértices del i -ésimo polígono de esta sucesión. Nótese que el hecho de que $(\mathcal{P}_i)_{i \in \mathbb{N}}$ sea una sucesión de Cauchy, implica que vértice a vértice, $(v_{i,j})_{i \in \mathbb{N}}$ es una sucesión de Cauchy. Como $v_{i,j}$ son puntos en \mathbb{R}^2 , por completitud de este espacio vectorial, existe v_j tal que $\lim_{i \rightarrow \infty} v_{i,j} = v_j$.

Usando esta propiedad para $j = 1, \dots, n$, podemos ver que para toda $\varepsilon > 0$, existe $M(\varepsilon) \in \mathbb{N}$ tal que si $i \geq N$, entonces $\|v_j - v_{i,j}\| < \varepsilon$, para toda j .

Dado esto, fijémonos en $\|v_j - v_{j+1}\|$. Utilizando la desigualdad del triángulo, podemos ver que:

$$\begin{aligned} \|v_j - v_{j+1}\| &\leq \|v_j - v_{i,j}\| + \|v_{i,j} - v_{i,j+1}\| + \|v_{i,j+1} - v_{j+1}\| \\ &\leq 2\varepsilon + \|v_{i,j} - v_{i,j+1}\| \text{ (tomando } i \geq M(\varepsilon)\text{)} \\ &= 2\varepsilon + l_j \text{ (recordando que } P_i \text{ es un polígono bueno).} \end{aligned}$$

Tomando el límite cuando ε tiende a 0, tenemos que $\|v_j - v_{j+1}\| \leq l_j$.

Por otra parte, usando la desigualdad del triángulo generalizada del teorema 1, vemos que

$$\|v_{i,j} - v_{i,j+1}\| \leq \|v_{i,j} - v_j\| + \|v_j - v_{j+1}\| + \|v_{j+1} - v_{i,j+1}\|.$$

Tomando $i \geq M(\varepsilon)$, y recordando que \mathcal{P}_i es un polígono bueno, tenemos que

$$\begin{aligned} \|v_j - v_{j+1}\| &\geq \|v_{i,j} - v_{i,j+1}\| - \|v_{i,j} - v_j\| - \|v_{j+1} - v_{i,j+1}\| \\ &\geq \|v_{i,j} - v_{i,j+1}\| - 2\varepsilon \\ &= l_j - 2\varepsilon. \end{aligned}$$

De nuevo tomando el límite cuando ε tiende a 0, tenemos que $\|v_j - v_{j+1}\| \geq l_j$. Combinando las dos desigualdades, $\|v_j - v_{j+1}\| = l_j$, y por lo tanto \mathcal{P} , el polígono formado por v_1, \dots, v_n es un polígono bueno. También, es claro que este polígono es el límite de $(\mathcal{P}_i)_{i \in \mathbb{N}}$, pues para

toda $i \geq M(\varepsilon)$,

$$\begin{aligned} d(\mathcal{P}, \mathcal{P}_i) &= \sum_{j=1}^n \|v_j - v_{i,j}\|^2 \\ &\leq \sum_{j=1}^n \varepsilon^2 \\ &= n\varepsilon^2 \end{aligned}$$

y cuando ε tiende a 0, también lo hace $n\varepsilon^2$, pues n es fijo. □

Ahora, consideremos la función área. El área de un polígono arbitrario puede ser considerada como la suma de áreas signadas de los triángulos $Ov_i v_{i+1}$, con $v_{n+1} = v_1$. Esta área es fácil de calcular con la fórmula $A(Ov_i v_{i+1}) = \frac{1}{2}|x_i y_{i+1} - x_{i+1} y_i|$. Entonces, tenemos lo siguiente.

Definición 2 (Área de un polígono). *Dado un polígono \mathcal{P} , definimos el área $A : \mathbb{R}^{2 \times n} \mapsto \mathbb{R}^+ \cup \{0\}$ como $A(\mathcal{P}) = \frac{1}{2} |\sum (x_i y_{i+1} - x_{i+1} y_i)|$, con $(x_{n+1}, y_{n+1}) = (x_1, y_1)$.*

Teorema 3. *El área es una función continua.*

Demostración. Es una consecuencia directa de que el área se exprese únicamente con sumas, restas y productos. □

Corolario. *La imagen del conjunto de polígonos buenos bajo A es compacta.*

Demostración. Primero, notemos que el área de los polígonos buenos es acotada. En efecto, es fácil ver que todo polígono bueno está contenido en un círculo de radio $\sum l_i$ (centrado en v_1), de forma que el área de un polígono bueno tiene que ser menor a $\pi(\sum l_i)^2$. Una vez que tenemos que el área es acotada, al ser la función A una función continua y ser el espacio de polígonos buenos completos, la imagen bajo A de este espacio debe cumplir que todas sus sucesiones de Cauchy convergen dentro de esta imagen. Entonces, esta imagen es cerrada. Más aún, como es acotada, debe ser compacta. □

Corolario. *Hay un área máxima y un área mínima para los polígonos buenos.*

Demostración. Este resultado es una consecuencia directa de la compacidad de la imagen de los polígonos buenos bajo el área. □

Con esto vemos que nuestro problema está bien definido, en el sentido de que hay áreas que no se pueden alcanzar.

Está fuera del alcance de este artículo, pero es intuitivo ver que el conjunto de polígonos buenos es un conjunto conexo, de forma que la imagen de este conjunto bajo el área es conexo, así que en el rango $[A_{min}, A_{max}]$, es posible alcanzar todos los valores de área. Entonces, encontrar los valores de A_{min} y A_{max} debe ser nuestra prioridad, pues una vez que lo logremos, resolveremos nuestro problema.

El problema de encontrar A_{min} , además de ser significativamente más difícil que encontrar A_{max} , es casi inútil, pues un polígono de área mínima es casi redundante. Entonces, concluyamos con la forma de encontrar A_{max} .

Voy a demostrar que el polígono que maximiza el área es el polígono cíclico; es decir, aquél que todos sus vértices yacen sobre la misma circunferencia. Lo realizaré por inducción.

El caso $n = 3$ es trivial, pues por criterio de congruencia **LLL**, hay un único triángulo con cierta sucesión de lados y todo triángulo es cíclico.

El caso $n = 4$ es un poco más complicado. La idea es utilizar la fórmula de Bretschneider para calcular el área de un cuadrilátero, que en nuestro caso es:

$$A = \sqrt{(s - l_1)(s - l_2)(s - l_3)(s - l_4) - l_1 l_2 l_3 l_4 \cos^2 \frac{\alpha + \beta}{2}},$$

donde α y β son respectivamente los ángulos en v_1 y v_3 . Nótese que el único término variable en la anterior fórmula es $\cos^2 \frac{\alpha + \beta}{2}$, así que minimizar este término nos llevará a maximizar el área. Al ser el cuadrado de algo, este término es mayor o igual a 0, así que para minimizarlo, necesitamos encontrar α, β tales que $\cos \frac{\alpha + \beta}{2} = 0$. Dado que $0 \leq (\alpha + \beta)/2 < \pi$, y al alcanzar la función coseno una única vez el 0 en este intervalo, tenemos que $(\alpha + \beta)/2 = \pi/2$, o $\alpha + \beta = \pi$. Pero sabemos que un cuadrilátero es cíclico si y sólo si los ángulos en vértices opuestos suman π , como en este caso. Entonces, nuestro cuadrilátero con mayor área es cíclico.

Finalmente, tomemos el caso $n > 4$. Considérese el polígono \mathcal{P} de mayor área con lados l_1, \dots, l_n . Sea $d_1 = \|v_2 - v_n\|$ y sea $d_n = \|v_1 - v_{n-1}\|$, nótese que podemos separar el área de \mathcal{P} como

$$A(v_n v_1 v_2) + A(v_2 v_3 \dots v_n).$$

Para maximizar esta suma, dado que conocemos todas las longitudes de los lados en $v_2 v_3 \dots v_n$ (pues son l_2, \dots, l_{n-1}, d_1), tenemos que $v_2 v_3 \dots v_n$ es cíclico, por hipótesis de inducción. Análogamente, separando $A(\mathcal{P})$ como $A(v_{n-1} v_n v_1) + A(v_1 v_2 \dots v_{n-1})$, tenemos que $v_1 \dots v_{n-1}$ es cíclico. Como $n > 4$, el conjunto $\{v_2, \dots, v_{n-1}\}$ tiene al menos tres elementos, y está en las dos circunferencias que definimos anteriormente. Pero tres puntos definen una única circunferencia. Entonces, éstas dos circunferencias son la misma, y por lo tanto, todos los vértices yacen sobre la misma circunferencia.

Conclusiones

De este procedimiento se vislumbra una forma de demostrar que el círculo es la figura en el plano con mayor área dado un perímetro fijo. En efecto, si partimos un lado de un polígono en dos, el nuevo vértice tenderá a pegarse al mismo círculo que el resto de los vértices, y el polígono se parecerá cada vez más al círculo.

Pero lo más importante, es la combinación de diversas áreas de las matemáticas, surgidas en distintos momentos de los últimos cinco siglos para resolver un problema antiguo, y mediante esta combinación de herramientas, obtener un resultado interesante.

El Problema de Monty Hall

Thomas Bladt

Alumno de Matemáticas Aplicadas del ITAM

Imagina que te encuentras en un concurso en donde se te presentan tres puertas cerradas. Detrás de una de ellas se encuentra un automóvil y detrás de las otras dos: cabras. Eliges una puerta y el presentador del concurso, que sabe qué hay detrás de cada una, abre una de las otras dos puertas, mostrando una cabra. En este momento te hace la pregunta: "¿Deseas quedarte con la puerta que elegiste, o prefieres cambiarla por la otra puerta que aún permanece cerrada?" Asumiendo que quieres ganar el coche ¿debes aceptar o cambiar?

Mientras que anteriormente ya existían problemas similares, es en 1975 cuando la revista estadounidense *American Statistician* publica una carta enviada por Steve Selvin, un matemático de la Universidad de California en Berkley.

En la carta, Selvin describe este problema (con un planteamiento similar) y propone que siempre conviene cambiar. Posteriormente, tuvo que escribir otra carta a la misma revista, titulada "On the Monty Hall Problem", defendiendo su postura. El curioso nombre proviene de un programa de juegos estadounidense de los setenta llamado *Let's Make a Deal*, cuyo presentador era conocido como Monty Hall.

Selvin planteó originalmente el ejercicio como una posible ocurrencia en este programa. Sin embargo, no fue sino hasta el 9 de septiembre de 1990 que el problema se volvió el centro de atención en la comunidad matemática, generando muchas disputas y polémica. En esa misma fecha apareció la pregunta, enviada por un lector, en la columna de preguntas y respuestas de Marilyn vos Savant, de la revista *Parade*. Marilyn era conocida por tener el récord del cociente intelectual más alto del mundo en ese entonces. En su respuesta, reafirmó lo que había concluido Steve Selvin tres lustros atrás: es mejor cambiar de puerta. Al hacerlo, decía, duplicas la probabilidad de ganar el coche. Esto desató una lluvia de críticas hacia ella por la comunidad matemática. Le llegaron miles de cartas, varias escritas por doctores en matemáticas, respondiéndole que estaba terriblemente equivocada; muchas de las cartas tenían un tono muy fuerte en contra de Vos Savant.

Por ejemplo, "May I suggest that you obtain and refer to a standard textbook on probability before you try to answer a question of this type again?" de la parte de Charles Reid, de la universidad de Florida.

E. Ray Bobo, de la universidad de Georgetown le escribió: "You are utterly incorrect about the game show question, and I hope this controversy will call some public attention to the serious national crisis in mathematical education. If you can admit your error, you will have contributed constructively towards the solution of a deplorable situation. How many irate mathematicians are needed to get you to change your mind?"

Por otro lado, Everett Harman, del US Army Research Institute apuntó: "You made a mistake, but look at the positive side. If all those Ph.D.'s were wrong, the country would be in some very serious trouble."

Inclusive recibió un comentario machista por parte de Don Edwards Sunriver: "Maybe women look at math problems differently than men."

Pero vos Savant tenía razón: cambiar de puerta duplica la probabilidad de ganar el coche.

Veamos por qué.

Vamos a partir del problema como fue descrito al principio y hagamos unos supuestos que deberían quitar cualquier ambigüedad del planteamiento.

Supondremos que Monty (el presentador) ya sabe de antemano qué hay detrás de cada puerta.

Asumiremos que la probabilidad de que el coche se encuentre detrás de una determinada puerta es de $\frac{1}{3}$, es decir, es igual de probable que esté detrás la puerta 1, 2 o 3.

Monty siempre abrirá una puerta que muestre una cabra y, en caso de que tenga dos opciones para esto, elegirá entre ellas aleatoriamente. Una prueba formal ocupa el Teorema de Bayes, pero antes de presentarla, vamos a convencernos de que la solución es correcta. Primero, una forma fácil de verlo es:

Intuición 1 (Simple visualización). Supongamos, sin pérdida de generalidad, que elegiste la puerta 1. A partir de esto, veamos todos los posibles escenarios que podrían ocurrir:

Escenario	Puerta 1	Puerta 2	Puerta 3	Quedándote	Cambiando
1	coche	cabra	cabra	ganas	pierdes
2	cabra	coche	cabra	pierdes	ganas
3	cabra	cabra	coche	pierdes	ganas

Esto quiere decir que en 2 de 3 casos sales ganando si cambias de puerta. Esto es porque es más probable elegir inicialmente una puerta que resguarda a una cabra, y luego Monty *siempre* va a abrir otra puerta que muestre una cabra. Por lo tanto, cambiar duplica tu probabilidad de ganar de $\frac{1}{3}$ a $\frac{2}{3}$.

Ahora, algo común en los matemáticos es llevar los casos al extremo. La siguiente intuición hace justamente esto.

Intuición 2 (Al extremo). Supongamos que, en lugar de 3, ahora en el juego hay 100 puertas. Igual que antes, sólo hay una puerta que esconde un coche y las 99 restantes tienen una cabra por detrás de ellas. Eliges una puerta; lo más probable es que no ganes el coche, pues la probabilidad de que hayas elegido correctamente es de $\frac{1}{100}$. Ahora Monty abre 98 puertas en donde se muestran puras cabras. ¿Cambiarías de puerta o te quedarías?

Aquí la decisión parece ser mucho más sencilla, pero es el mismo escenario, sólo que con un mayor número de puertas. Lo podemos ver como que la probabilidad de obtener el coche es de $\frac{1}{n}$ si te quedas con tu elección original, o de $\frac{n-1}{n}$ si decides cambiar de puerta, en donde n es el número de puertas que hay. Por ejemplo, en el caso de 3 puertas tenemos que las probabilidades de ganar son de $\frac{1}{3}$ o $\frac{2}{3}$ si decides quedarte o cambiarte, respectivamente; en el caso de las 100 puertas, las probabilidades de ganar son de $\frac{1}{100}$ o $\frac{99}{100}$ si decides quedarte o cambiarte, respectivamente. Esto, a su vez, implicaría que

$$\lim_{n \rightarrow \infty} \mathbb{P}(\{\text{ganar quedántote}\}) = \lim_{n \rightarrow \infty} \frac{1}{n} = 0,$$

donde n es el número de puertas. Por lo tanto, es mejor cambiarte.

Muchas personas no quedan convencidas de la solución con estos argumentos. Quizás el razonamiento parezca frágil o con errores. Una de estas personas fue el famoso matemático húngaro Paul Erdős, quien no creía en la respuesta proporcionada por Von Savant. Fue necesario que aparecieran las computadoras para que Erdős finalmente se convenciera.

Intuición 3 (Simulación). Para este apartado, hemos hecho un programa que simula el juego. En la siguiente tabla, n representa el número de veces que se realizó la simulación, GQ denota el número de veces que se ganó quedando, GC el número de veces que se ganó cambiando, mientras que $FR(GQ)$ y $FR(GC)$ denotan sus frecuencias relativas.

n	GQ	GC	$FR(GQ)$	$FR(GC)$
100	39	61	0.3900	0.6100
1,000	343	657	0.3430	0.6570
10,000	3,273	6,727	0.3273	0.6727
100,000	33,308	66,692	0.3331	0.6669
1,000,000	333,335	666,665	0.3333	0.6667

Bajo un punto de vista frecuentista, tenemos que conforme n crece la frecuencia relativa de un evento tiende a la probabilidad del mismo. Las simulaciones nos harían pensar que $FR(GQ)$ tiende a $0.\overline{66}$ y $FR(GC)$ tiende a $0.\overline{33}$, que es lo mismo que nos dicen las otras intuiciones. Veamos una gráfica de la frecuencia relativa usando 1000 simulaciones.

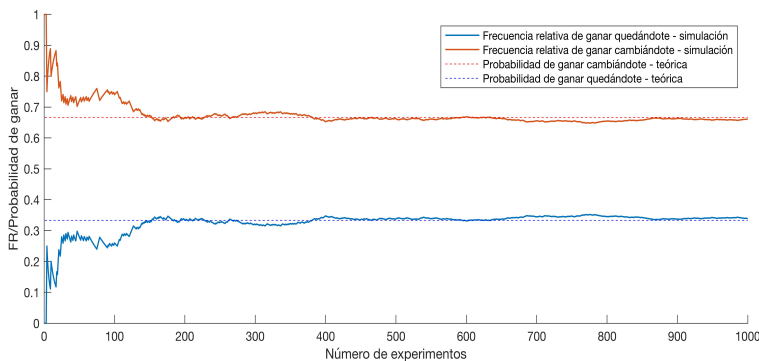


Figura 1: Frecuencia relativa de una simulación de $n = 1000$ juegos.

Lo que la Figura 1 nos muestra es que con pocas simulaciones las observaciones varían mucho, pero aproximadamente a las 200 se estabilizan.

Ahora que podemos estar convencidos de que lo que afirmaron tanto Selvin como Von Savant es cierto, pasemos a ver una prueba formal de ello.

Demostración. Supongamos, sin pérdida de generalidad, que elegimos la puerta 1. Sea el evento H definido por

$$H := \{ \text{Puerta 1 tiene el coche detrás de ella} \}$$

y sea el evento E definido por

$$E := \{ \text{Monty abre una puerta que muestra una cabra} \}.$$

Buscamos la probabilidad de que el coche esté en la puerta 1 dado que Monty nos acaba de abrir otra puerta que resguarda a una cabra. Esto es, buscamos $\mathbb{P}(H|E)$, por el teorema de Bayes tenemos que

$$\mathbb{P}(H|E) = \frac{\mathbb{P}(H \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|H)\mathbb{P}(H)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|H)\mathbb{P}(H)}{\mathbb{P}(E|H)\mathbb{P}(H) + \mathbb{P}(E|H^C)\mathbb{P}(H^C)}.$$

Luego,

- $\mathbb{P}(H)$ es la probabilidad de que la puerta tenga un coche detrás de ella. Por nuestros supuestos, esta probabilidad es de $\frac{1}{3}$.
- $\mathbb{P}(H^C)$ es la probabilidad de que la puerta tenga una cabra detrás de ella. Esta probabilidad es de $\frac{2}{3}$.
- $\mathbb{P}(E|H) = \mathbb{P}(E|H^C) = 1$, pues son la probabilidad de que Monty nos abra una puerta con una cabra en los dos escenarios. Como esto *siempre* va a pasar, la probabilidad es 1.

Por lo tanto,

$$P(H|E) = \frac{1 \cdot (1/3)}{1 \cdot (1/3) + 1 \cdot (2/3)} = \frac{1}{3}.$$

Esto es, la probabilidad de que el coche esté en la misma puerta que elegimos, dado que Monty nos abre otra con una cabra, es de $\frac{1}{3}$. Por lo tanto, si nos quedamos, ganaremos con probabilidad de $\frac{1}{3}$ y si cambiamos ganamos con probabilidad de $\frac{2}{3}$. \square

La clave del problema yace en que Monty siempre abre una puerta con una cabra, mas no abre una de maneara aleatoria. Si esto último fuese el caso, el término $\mathbb{P}(E|H^C)$ sería igual a $1/2$ pues Monty debe elegir aleatoriamente entre abrir una puerta con una cabra y una puerta con un coche detrás. Sustituyendo este valor en la demostración pasada nos daría que $P(H|E) = 1/2$.

Referencias

- [1] Rosenhouse, J. *The monty hall problem: the remarkable story of maths most contentious brain teaser*. Oxford University Press, Estados Unidos, 2009.

Una introducción práctica a las cópulas

Jorge Francisco de la Vega Góngora
Profesor del Departamento de Estadística del ITAM

“Fall in love with some activity, and do it! Nobody ever figures out what life is all about, and it doesn’t matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough. Work as hard and as much as you want to on the things you like to do the best.”

Richard Phillips Feynman

El concepto de dependencia estocástica es complejo. En los cursos generales usuales de estadística y probabilidad a nivel universitario se da una idea general de lo que conforman un conjunto de eventos independientes y dependientes, para luego revisar distribuciones conjuntas, marginales y condicionales. Después, todo son básicamente muestras aleatorias: conjuntos de variables aleatorias independientes e idénticamente distribuidas. En muchas ocasiones los estudiantes se quedan con una visión parcial del concepto de dependencia, típicamente caracterizando a la **correlación de Pearson** como *la* medida de dependencia con la que pueden medir el ajuste de una regresión lineal. Sólo para aquellos valerosos estudiantes que se animan (o están obligados por sus programas) a tomar cursos en procesos estocásticos, amplían sus horizontes para entender otro tipo de dependencia estocástica: la dependencia markoviana, que simplifica el problema complejo de entender la distribución de un proceso estocástico y permite realizar aplicaciones concretas.

Pero la historia no termina ahí, ya que en muchas aplicaciones de la vida real se requiere entender la relación de dependencia entre conjuntos de variables aleatorias con distribuciones marginales dadas, y poder simular dichas variables con una estructura de dependencia dada y es ahí donde entran las cópulas. El propósito de esta nota es dar una introducción práctica al concepto de cópula y a algunas de sus aplicaciones prácticas sobre todo en simulación y en el contexto financiero.

Introducción

El estudio de las cópulas no es un tema nuevo, tiene más de 60 años a la fecha. Fue en 1959 que Abe Sklar¹ introdujo el concepto de cópulas como parte de una pregunta que Maurice Fréchet formuló sobre la relación que existe entre una función de distribución conjunta y sus funciones de distribución marginales.

Las cópulas son funciones que ligan distribuciones multivariadas a sus marginales univariadas. Su estudio entre 1959 y 1974, particularmente en el caso bivariado, permitió usar

¹Abe Sklar fue alumno de Tom M. Apostol y fue profesor en el Illinois Institute of Technology.

estas funciones para definir medidas no paramétricas de dependencia para pares de variables aleatorias.

Las aplicaciones de las cópulas se han extendido con singular fuerza e importancia en los últimos 20 años, y se les encuentra de manera común mencionadas en contextos estadísticos, financieros y particularmente en el de medición de riesgos, donde la medición de eventos extremos que se pueden dar de manera conjunta es de mucha importancia.

Se pueden encontrar una gran cantidad de artículos y referencias que introducen el contexto teórico básico de las cópulas, pero no se encuentran en la misma proporción documentos que expliquen o ejemplifiquen la estimación y aplicaciones prácticas de la cópulas, en parte porque sólo hasta fechas recientes se han desarrollado de manera más completa herramientas computacionales que permiten llevar a cabo los cálculos operacionales necesarios tanto para su estimación como uso práctico, particularmente en el ambiente de \mathbb{R} . En esta breve nota, se pone un énfasis al uso adecuado de las herramientas computacionales disponibles que complementa a las notas teóricas disponibles (por ejemplo, consídese Erdely [4]). En lo que sigue se dará un breve contexto conceptual y se mostrará el uso práctico de las herramientas y las consideraciones que se deben hacer para un uso adecuado.

Marco teórico

En esta sección daremos una definición operacional de cópulas y se comentarán las herramientas computacionales disponibles en \mathbb{R} para trabajar con ellas.

Conceptos básicos necesarios

En esta sección se darán las definiciones necesarias y algunos resultados importantes sin demostración. Las pruebas de los resultados y mayor detalle de toda la teoría se puede consultar en Joe [1] o en Nelsen [2].

Definición 1. Una *cópula* es una función de distribución conjunta $C[0, 1]^n \rightarrow [0, 1]$ cuyas distribuciones marginales son todas $\mathcal{U}(0, 1)$.

Dada la definición anterior, podemos construir una cópula de la siguiente manera. Primero considérense dos variables aleatorias (X_1, X_2) con distribución conjunta F y marginales $F_1(x_1)$ y $F_2(x_2)$. Definamos la función C de la siguiente manera:

$$C(u, v) = F(F_1^-(u), F_2^-(v)) \quad \forall u, v \in [0, 1],$$

donde F^- indica la inversa generalizada de F^2 . Entonces C es una cópula:

²Para una función creciente $T : \mathbb{R} \rightarrow \mathbb{R}$ con $T(-\infty) = \lim_{x \downarrow -\infty} T(x)$ y $T(\infty) = \lim_{x \uparrow \infty} T(x)$, la **inversa generalizada** $T^- : \mathbb{R} \rightarrow [-\infty, \infty]$ de T se define como: $T^-(y) = \inf\{x \in \mathbb{R} | T(x) \geq y\}$, $y \in \mathbb{R}$. También se le llama **función cuantil**.

$$\begin{aligned} C(u, v) &= F(F_1^-(u), F_2^-(v)) = P(X \leq F_1^-(u), Y \leq F_2^-(v)) \\ &= P(F_1(X) \leq u, F_2(Y) \leq v) = P(U \leq u, V \leq v), \end{aligned}$$

y claramente las marginales de $C(u, v)$ son uniformes.

Podemos reescribir la fórmula de una cópula de tal manera que resalte una de sus características más interesantes. Como $F_X(X) \sim \mathcal{U}(0, 1)$ para cualquier variable aleatoria X . Si tomamos $X = (X_1, X_2)$ entonces, por definición, la función $C(F_1(x_1), F_2(x_2))$ es una cópula. Como C es una función de distribución si reescribimos $u = F_1(x_1)$ y $v = F_2(x_2)$ obtenemos la siguiente relación:

$$\begin{aligned} C(F_1(x_1), F_2(x_2)) = C(u, v) &= P(U \leq u, V \leq v) = P(F_1(X_1) \leq u, F_2(X_2) \leq v) \\ &= P(X_1 \leq F_1^{-1}(u), X_2 \leq F_2^{-1}(v)) \\ &= F_X(F_1^{-1}(u), F_2^{-1}(v)) = F_X(x_1, x_2) \end{aligned}$$

Juntando los extremos, obtenemos: $F_X(x_1, x_2) = C(F_1(x_1), F_2(x_2))$. De este modo, F_X se desacopla en dos piezas: una cópula C que contiene sólo la información de las dependencias de $X = (X_1, X_2)$; la otra parte corresponde a las distribuciones marginales de X_1 y X_2 , que sólo contienen la información de las variables de manera marginal. Esta separación es justamente lo que prueba el teorema de Sklar (1959).

Teorema 1. Teorema de Sklar (1959)

Sea F una función de distribución conjunta de un vector aleatorio p -dimensional $X = (X_1, \dots, X_p)$ con distribuciones marginales F_1, \dots, F_p . Entonces existe una cópula $C : [0, 1]^p \rightarrow [0, 1]$ tal que $\forall x \in \mathbb{R}^p$,

$$F(x_1, \dots, x_p) = C(F_1(x_1), \dots, F_p(x_p))$$

Además, si las marginales son continuas, entonces C es única. Conversamente, si C es una cópula y F_1, \dots, F_p son funciones de distribución, entonces F como se definió arriba, es una distribución con marginales F_1, \dots, F_p .

En resumen, usamos las cópulas para especificar una distribución conjunta en un proceso de dos etapas: primero, especificamos el tipo de distribuciones marginales que se desea conjuntar y segundo, especificamos qué cópula usar, que corresponda a la dependencia conjunta de las variables aleatorias. Como las cópulas solo especifican la estructura de la dependencia, diferentes cópulas producen diferentes distribuciones conjuntas cuando se aplican a las mismas marginales. Es importante notar que para un par de marginales, se pueden usar infinitas cópulas para crear una distribución conjunta, así que es necesario "ajustar" la que más se apegue al comportamiento de dependencia entre las variables.

Aterrizando ideas, veremos algunos ejemplos de cópulas.

- **Cópula de independencia:** Definamos $C : [0, 1]^2 \rightarrow [0, 1]$ como $C(u_1, u_2) = u_1 u_2$. Entonces, si F_1 y F_2 son distribuciones,

$$C(F_1(x_1), F_2(x_2)) = F_1(x_1)F_2(x_2) = F(x_1, x_2)$$

Entonces X_1 es independiente de X_2 .

- **Cópula de co-monotonicidad:** Si $U \sim \mathcal{U}(0, 1)$ y $\mathbf{U} = (U, U)$ (dos copias de U , así que las variables son completamente dependientes)

$$C^M(u_1, u_2) = P(U \leq u_1, U \leq u_2) = P(U \leq \min\{u_1, u_2\}) = \min\{u_1, u_2\}$$

- **Cópula de contra-monotonicidad:** Si $\mathbf{U} = (U, 1 - U)$

$$C^{CM}(u_1, u_2) = P(U \leq u_1, 1 - U \leq u_2) = P(1 - u_2 \leq U \leq u_1) = \max\{u_1 + u_2 - 1, 0\}$$

- **Cópula Gaussiana:** Si $Z \sim \mathcal{N}_2(\mathbf{0}, \Sigma)$ con marginales $Z_1 \sim \mathcal{N}(0, \sigma^2)$ y $Z_2 \sim \mathcal{N}(0, \sigma^2)$ y $\rho(Z_1, Z_2) = \rho$, entonces $C(u_1, u_2) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \Phi^{-1}(u_2))$ es una cópula. La fórmula explícita de la cópula es:

$$C_{\rho}(u_1, u_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{s^2 + t^2 - 2\rho st}{2(1-\rho^2)}\right\} ds dt$$

con $x_i = \phi^{-1}(u_i)$.

La Figura 1 muestra cómo se ven los ejemplos considerados.

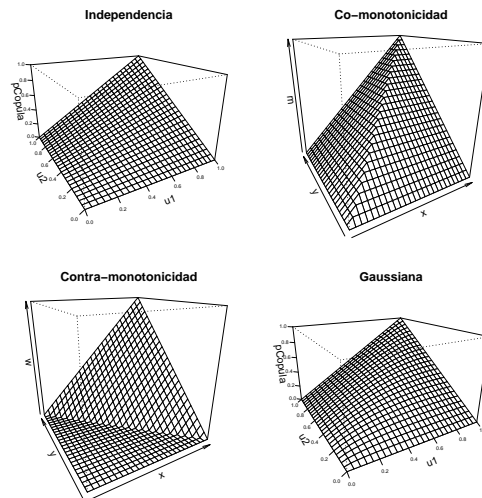


Figura 1: Ejemplos de cópulas.

Un resultado importante es que las cópulas de co-monotonicidad y contra-monotonicidad son los extremos que cualquier cópula puede tomar (es decir, son cotas mínima y máxima):

Teorema 2. Cotas inferior y superior de Fréchet-Hoeffding

Dada una cópula $C, \forall u_1, \dots, u_n \in [0, 1]$:

$$\max\{u_1 + \dots + u_n - n + 1, 0\} \leq C(u_1, \dots, u_n) \leq \min\{u_1, \dots, u_n\}$$

Entonces todas las cópulas viven entre estas fronteras, dentro del cubo unitario.

Familia Arquimediana de cópulas

Un grupo importante de cópulas para las aplicaciones son las cópulas que se obtienen a partir de ciertas funciones generadoras.

Definición 2. Una cópula Arquimediana con generador ϕ tiene la forma:

$$C(u_1, \dots, u_p) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_1))$$

donde ϕ satisface (1) $\phi : [0, 1] \rightarrow [0, \infty]$ es continua y estrictamente decreciente, (2) $\phi(0) = \infty$ y (3) $\phi(1) = 0$.

Diferentes generadores generan diferentes cópulas. Algunas de las más comunes son las que se definen a continuación.

Cópula de Frank

La cópula de Frank tiene función generadora:

$$\phi^{Fr}(u) = -\log\left\{\frac{e^{-\theta u} - 1}{e^{-\theta} - 1}\right\}, \theta \in \mathbb{R}$$

Verificando las propiedades de cópula arquimediana:

- $\phi^{Fr}(0) = -\log\left\{\frac{1-1}{e^{-\theta}-1}\right\} = -\log(0) = \infty$
- $\phi^{Fr}(1) = -\log\left\{\frac{e^{-\theta}-1}{e^{-\theta}-1}\right\} = -\log(1) = 0$
- Si $y = -\log\left\{\frac{e^{-\theta u}-1}{e^{-\theta}-1}\right\}$, entonces:

$$\begin{aligned} e^{-y} &= \frac{e^{-\theta u} - 1}{e^{-\theta} - 1} \\ (e^{-\theta} - 1)e^{-y} + 1 &= e^{-\theta u} \\ u &= -\frac{1}{\theta} \log\{(e^{-\theta} - 1)e^{-y} + 1\} \end{aligned}$$

- Por lo tanto: $\phi^{Fr^{-1}}(y) = -\frac{1}{\theta} \log \{ (e^{-\theta} - 1)e^{-y} + 1 \}$.

Con ambas funciones $\phi^{Fr^{-1}}$ y ϕ^{Fr} , hay que resolver la ecuación:

$$\phi^{Fr^{-1}}[\phi^{Fr}(u_1) + \phi^{Fr}(u_2)]$$

Finalmente, la cópula que se obtiene (en el caso bidimensional) es la siguiente:

$$C^{Fr}(u_1, u_2) = -\frac{1}{\theta} \log \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right).$$

Cópula de Clayton

La cópula de Clayton tiene función generadora: $\phi^C(u) = \frac{u^{-\theta} - 1}{\theta}$, $\theta > 0$ y de manera similar a como se hizo previamente, se obtiene la ecuación de la cópula:

$$C^C(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}$$

Cópula de Gumbel

La función generadora de la cópula de Gumbel tiene la forma: $\phi^G(u) = (-\log u)^\theta$, $\theta \geq 1$ y la ecuación de la cópula queda de la siguiente manera:

$$C^G(u_1, u_2) = \exp \{ -[(-\log u_1)^\theta + (-\log u_2)^\theta] \}$$

Cópula de valores extremos

Sea $A : [0, 1] \rightarrow [1/2, 1]$ una función convexa que satisface la siguiente condición para $w \in [0, 1]$: $\max\{w, 1 - w\} \leq A(x) \leq 1$. La familia de cópulas de valores extremos [5] se define a partir de la función A como:

$$C(u, v) = \exp \left[\log(uv) A \left(\frac{\log(v)}{\log(uv)} \right) \right]$$

Algunos ejemplos de casos particulares:

- $A(w) = 1$ es la cópula de independencia.
- $A(w) = \max\{w, 1 - w\}$ es la cópula co-monótona.
- $A(w) = [w^\theta + (1 - w)^\theta]^{1/\theta}$, $\theta \geq 1$ es la cópula de Gumbel.
- $A(w) = 1 - [w^{-\theta} + (1 - w)^{-\theta}]^{-1/\theta}$, $\theta \geq 0$ es la **cópula de Galambos**.

Hay muchísimas otras que se han desarrollado para aplicaciones en seguros y finanzas. Un resumen de las familias de cópulas más comunes se da en la siguiente tabla, y la figura 2, que se genera con el código en el siguiente recuadro, muestra gráficas de funciones generadoras para cada cópula y diferentes valores del parámetro θ .

Familia	Espacio parametral θ	$\psi(t)$	$\psi^{-1}(t)$	$C(u, v)$
Clayton	$\theta \geq 0$	$\frac{t^{-\theta}-1}{\theta}$	$(1 + \theta t)^{-1/\theta}$	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$
Frank	$\theta \geq 0$	$-\log \frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$-\theta^{-1} \log(1 + e^{-t}(e^{-\theta} - 1))$	$-\frac{1}{\theta} \log \left(1 + \frac{(e^{-\theta}u-1)e^{-\theta}v-1}{e^{-\theta}-1} \right)$
Gumbel	$\theta \geq 1$	$(-\log t)^\theta$	$e^{-t^{1/\theta}}$	$\exp \left[-\{(-\log u)^\theta + (-\log v)^\theta\}^{1/\theta} \right]$

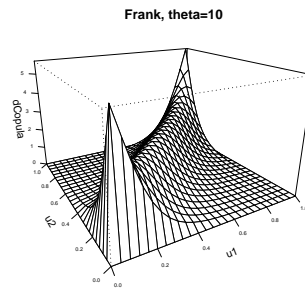
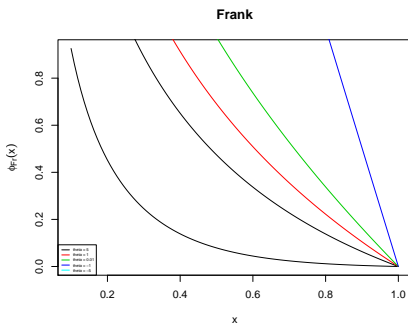
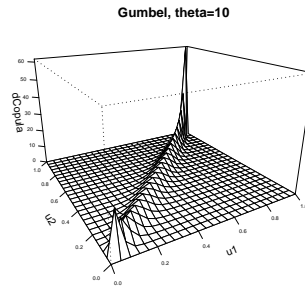
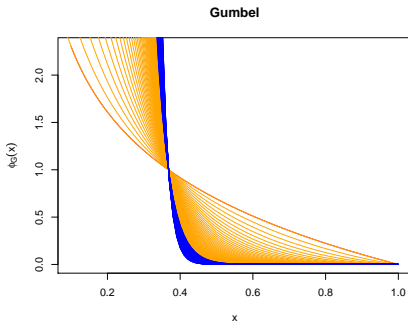
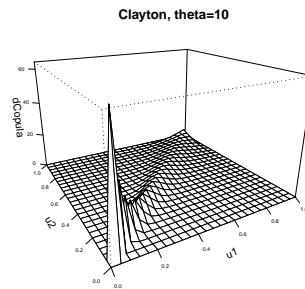
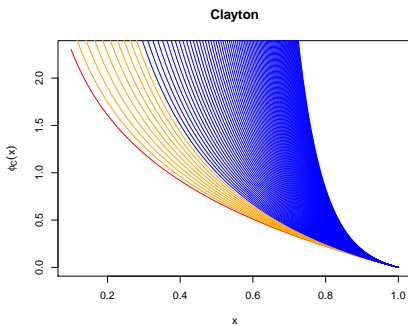


Figura 2: funciones generadoras y ejemplos de densidades de cópulas Arquimedianas

Medidas de dependencia

Como ya se mencionó previamente, la correlación de Pearson $\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$ es una medida de dependencia limitada, ya que por una parte, sólo mide el grado de asociación *lineal* entre dos variables, y por otra parte su cálculo involucra tanto a la distribución conjunta como a las marginales.

Otras medidas de dependencia son más relevantes en el contexto de cópulas, que miden tipos de dependencia más generales. En este contexto, las **medidas de correlación basada en rangos**, la tau de Kendall y la rho de Spearman, son medidas escalares que dependen sólo de la cópula de la distribución bivariada y no de las marginales.

τ (tau) de Kendall

Definición 3. Sea (X, Y) un vector aleatorio y (X^*, Y^*) un vector con la misma distribución e independiente de (X, Y) (es una copia de (X, Y)). Entonces (X, Y) y (X^*, Y^*) son pares **concordantes (discordantes)** si $(X - X^*)(Y - Y^*) > 0$ ($(X - X^*)(Y - Y^*) < 0$).

La τ de Kendall es la diferencia de probabilidades de par concordante y de par discordante:

$$\begin{aligned} \rho_\tau(X, Y) &= P((X - X^*)(Y - Y^*) > 0) - P((X - X^*)(Y - Y^*) < 0) \\ &= E(\text{sgn}\{(X - X^*)(Y - Y^*)\}) \end{aligned}$$

La τ de Kendall muestral está dada por:

$$\hat{\rho}_\tau = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \text{sgn}\{(X_i - X_j^*)(Y_i - Y_j^*)\} = \frac{C - D}{C + D} = \frac{C - D}{\binom{n}{2}}$$

donde C son los pares concordantes y D son los pares discordantes.

ρ_S (rho) de Spearman

Definición 4. La **rho o correlación de Spearman** es la correlación de Pearson sobre los valores evaluados en los rangos inducidos por las distribuciones marginales de los datos:

$$\rho_S(X, Y) = \text{cor}(F_1(\text{rango}(X)), F_2(\text{rango}(Y)))$$

Podemos ver que la correlación de Spearman es la correlación de la cópula de (X, Y) . La ρ_S muestral está dada por:

$$\hat{\rho}_S(X, Y) = \frac{12}{n(n^2 - 1)} \sum_{i=1}^n \left(\text{rango}(X_i) - \frac{n+1}{2} \right) \left(\text{rango}(Y_i) - \frac{n+1}{2} \right)$$

Las medidas de correlación aún toman valores entre -1 y 1 . $\rho_\tau = \rho_S = 0$ para variables independientes (la converso no se cumple) $\rho_\tau = \rho_S = 1$ cuando X y Y son co-monótonas.

$\rho_\tau = \rho_S = -1$ cuando X y Y son contra-monótonas. En términos de cópulas, se puede ver que:

$$\rho_\tau = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1 \quad \text{y} \quad \rho_S = 12 \int_0^1 \int_0^1 (C(u, v) - uv) dudv$$

Aunque en la práctica estas fórmulas no son muy útiles, salvo en el caso de variables aleatorias normales. En el caso de la normal bivariada, hay una relación biyectiva entre las medidas no paramétricas y el coeficiente de correlación, que es útil en el contexto de simulación de cópulas gaussianas:

$$\rho_\tau = \frac{2}{\pi} \arcsin(\rho) \quad \text{y} \quad \rho_S = \frac{6}{\pi} \arcsin(\rho/2)$$

La Figura 3 muestra la relación entre los coeficientes en el caso de la distribución normal.

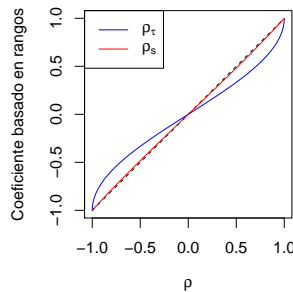


Figura 3: Medidas de dependencia en el caso normal

Ejemplos y aplicaciones

Simulación de una distribución conjunta con marginales dadas

En algunas aplicaciones, se requiere simular variables dependientes con marginales conocidas. Se puede dar un método general para construir variables dependientes con distribuciones arbitrarias: $X_1 \sim F_1, X_2 \sim F_2$, utilizando una cópula dada. Por ejemplo, para el caso de la cópula gaussiana:

1. Generar $\mathbf{Z} = (Z_1, Z_2) \sim \mathcal{N}_2(\mathbf{0}, \Sigma(\rho))$, donde $cor(Z_1, Z_2) = \rho$.
2. Obtener $\mathbf{u} = (u_1, u_2) \sim (\Phi(Z_1), \Phi(Z_2))$.³

³Notar que la cópula normal se usa en este punto, ya que $C(u_1, u_2) = \Phi(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) = \Phi(\Phi^{-1}(\Phi(Z_1)), \Phi^{-1}(\Phi(Z_2))) = \Phi(Z_1, Z_2) = \mathcal{N}_2(\mathbf{0}, \Sigma(\rho))$.

3. Obtener $\mathbf{X} = (X_1, X_2) \sim (F^{-1}(u_1), G^{-1}(u_2))$

X_1 y X_2 son dependientes. Sin embargo la correlación no es necesariamente la deseada: $cor(X_1, X_2) \neq \rho$, ya que se aplicaron transformaciones no lineales. Entonces, es necesario utilizar alguna de las medidas de dependencia basadas en rangos ya que son invariantes ante ese tipo de transformaciones.

Consideremos el siguiente ejemplo: queremos obtener una muestra aleatoria de un vector aleatorio $W = (X, Y, Z)$ donde $X \sim N(4, 25)$, $Y \sim t_4$ y $Z \sim Binom(25, 0.4)$, considerando las siguientes restricciones: $\tau(X, Y) = 0.7$, $\tau(X, Z) = 0.3$ y $\tau(Y, Z) = 0.4$.

La primera alternativa es usar una cópula Gaussiana para construir nuestra muestra, incorporando la restricción de la dependencia. El procedimiento que se utilizará es el siguiente:

1. Transformar las τ de Kendall a la correlación para construir la matriz de correlaciones necesaria Σ . Para transformar los valores usamos la fórmula conocida para la distribución normal: $\rho = \sin(\frac{\pi * \tau}{2})$. Entonces la matriz Σ tiene la forma

$$\Sigma = \begin{pmatrix} 1 & 0.8910065 & 0.4539905 \\ 0.8910065 & 1 & 0.5877853 \\ 0.4539905 & 0.5877853 & 1 \end{pmatrix}$$

2. Obtener un vector $Z \sim N(0, \Sigma)$
3. Obtener un vector $U \sim (\Phi(Z_1), \Phi(Z_2), \Phi(Z_3))$, donde $\Phi(x)$ es la función de distribución normal estándar
4. Obtener un vector $W \sim (F_1^{-1}(U_1), F_2^{-1}(U_2), F_3^{-1}(U_3))$, donde las F_i son las distribuciones deseadas. Entonces W tiene la composición deseada.

Justo los pasos 2 y 3 son los que corresponden a obtener una muestra aleatoria de una cópula, en este caso una cópula gaussiana. Entonces podemos escribir un código relativamente sencillo:

```
copula_normal_3 <- normalCopula(c(sin(0.7*pi/2), sin(0.4*pi/2), sin(0.3*pi/2)), dim = 3, dispstr = "un")
set.seed(100) #fija una semilla
U <- rCopula(1000, copula_normal_3) #Genera la muestra aleatoria
```

La Figura 4 muestra los valores U en pares. Estos valores se conocen como **pseudo-observaciones**.

```
pairs(U, pch=16, cex=0.5)
```

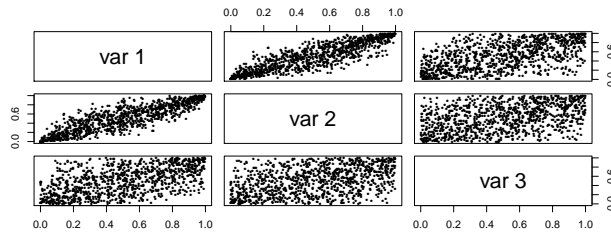


Figura 4: pseudo-observaciones de la cópula normal

Con el código previo, se realizan los primeros tres pasos de la simulación. Antes de continuar, podemos ver cómo se ven las muestras generadas por pares, y podemos ver si tenemos la condición establecida como restricción sobre las covarianzas. La matriz de tau de Kendall está dada por:

```
round(cor(U, method = "kendall"), 2)

      [,1] [,2] [,3]
[1,] 1.00 0.69 0.39
[2,] 0.69 1.00 0.28
[3,] 0.39 0.28 1.00
```

Ahora procedemos al paso 4, para generar nuestro vector y hacemos los histogramas para ver si tienen el comportamiento deseado, y se muestra un ejemplo de los valores generados. La Figura 5 muestra valores simulados de las distribuciones bivariadas de la cópula normal.

```
W <- cbind(qnorm(U[,1], mean = 4, sd = 5), qt(U[,2], 4), qbinom(p = U[,3], size = 25, prob = 0.4))
head(W, 3) #ejemplo de la muestra generada

      [,1]      [,2] [,3]
[1,] 2.180664 -0.1682176  9
[2,] 8.355016  0.6595940 11
[3,] 2.275934  0.1751841  8

pairs(W, pch = 16, cex = 0.5)
```

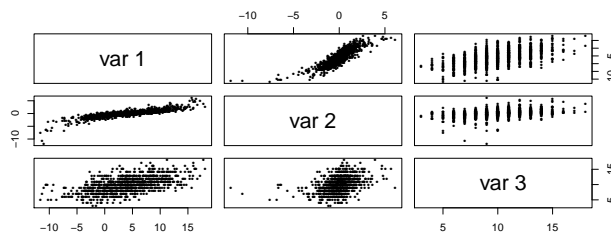


Figura 5: valores simulados de las distribuciones bivariadas de la cópula normal

Finalmente, las marginales de la cópula se muestran en la Figura 6.

```
#Grafica los histogramas y agrega densidades con las distribuciones deseadas para ver
par(mfrow = c(1,3))
hist(W[,1], prob = T,breaks=50); points(sort(W[,1]),dnorm(sort(W[,1]),4,5),type="l",col="red")
hist(W[,2], prob = T,breaks=50); points(sort(W[,2]),dt(sort(W[,2]),4),type="l",col="red")
hist(W[,3], prob = T);
points(sort(W[,3]),dbinom(sort(W[,3]),size=25,prob=0.4),type="l",col="red")
```

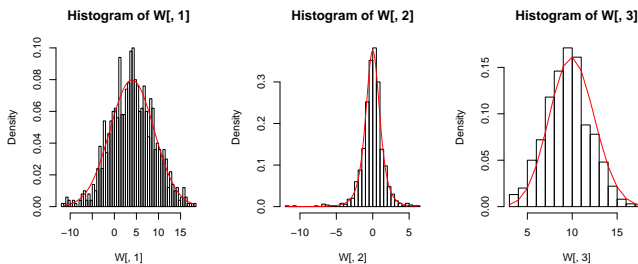


Figura 6: Marginales de la cópula normal simulada

Podemos corroborar que alcanzamos la medida de dependencia requerida

```
cor(W, method = "kendall")
      [,1] [,2] [,3]
[1,] 1.0000000 0.6884324 0.4120571
[2,] 0.6884324 1.0000000 0.3006960
[3,] 0.4120571 0.3006960 1.0000000
```

Estimación de cópula a partir de datos observados

Finalmente, consideraremos el caso de estimación de una cópula a partir de observaciones. El problema tiene dificultades técnicas y trampas con las que hay que ser muy cuidadoso ya que la estimación de cópulas implica usualmente que cada marginal debe ser evaluada y conectada a una distribución multivariada estimada.

A continuación veremos un ejemplo de estimación. Usaremos varios paquetes de R para realizar el ejercicio:

- paquetes `Ecdat`, es un conjunto de datos econométricos. Los datos `CRSPday` contienen una base de datos de rendimientos diarios de acciones del Center for Research in Security Prices (CRSP), del 3 de enero de 1969 al 31 de diciembre de 1998.
- `copula`, que es el paquete principal de donde se obtienen la mayoría de las características que se aplicarán.
- `fGarch` para el uso de la densidad t estandarizada
- `MASS` para el uso de las funciones `fitdistr`, `kde2d`

- `fCopulae` para funciones adicionales de copulas: `pempiricalCopula`, `ellipticalCopula-fit`.

Nos concentraremos en las dos variables `ibm`, que es el rendimiento de IBM y `crsp` que es un índice ponderado de rendimientos construido por el CRSP. La Figura 7 muestra una gráfica de la estructura conjunta de los rendimientos de ambos instrumentos. En la gráfica se puede apreciar que los datos muestran una alta concentración en el origen, y hay algunas observaciones que tienden a ocurrir simultáneamente en los extremos de ambas distribuciones marginales.

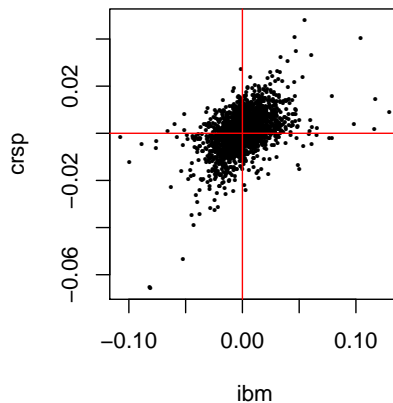


Figura 7: Rendimientos de IBM y CRSP para los datos dados

A continuación se ajustará una distribución t a cada una de las variables marginales. Los valores que se guardan corresponden a los valores estimados de las distribuciones t marginales. Cada distribución marginal puede ajustar diferentes grados de libertad.

La función `fitdistr` del paquete MASS estima las características de una función de distribución usando máxima verosimilitud (en el caso de la t , su media, escala, y grados de libertad).

```
est.ibm <- as.numeric(fitdistr(ibm,"t")$estimate) #parámetros t: media, escala, gl
est.crsp <- as.numeric(fitdistr(crsp,"t")$estimate)
#Convierte los parámetros de escala a desviaciones estándar en el caso de la t
est.ibm[2] <- est.ibm[2]*sqrt(est.ibm[3]/(est.ibm[3]-2))
est.crsp[2] <- est.crsp[2]*sqrt(est.crsp[3]/(est.crsp[3]-2))
#Grados de libertad para cada caso
est.ibm[3]

[1] 4.276156

est.crsp[3]

[1] 3.473982
```

Como un ejercicio inicial, supongamos que se quiere ajustar una cópula específica, por ejemplo, una cópula t . Para estimar una t -cópula por máxima verosimilitud, se requiere una estimación de la correlación y un valor inicial adecuado. Se usarán las densidades t estimadas como valores iniciales. Se define la cópula t con 2 grados de libertad

```
tau <- cor(ibm,crsp,method = "kendall")
omega <- 2/pi*asin(tau)
c(tau,omega)

[1] 0.3308049 0.2146404

copula2 <- tCopula(omega,dim=2,dispstr = "un",df = 2)
```

Ahora hay que ajustar la copula a los datos uniformes transformados:

```
# La función pstd es la distribución estándar t
# por el método de máxima verosimilitud
d1 <- cbind(pstd(ibm, mean = est.ibm[1], sd = est.ibm[2], nu = est.ibm[3]),
           pstd(crsp, mean = est.crsp[1], sd = est.crsp[2], nu = est.crsp[3]))

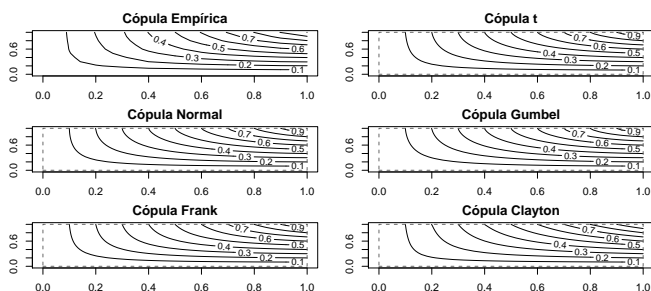
fit1 <- fitCopula(copula2, method = "ml", optim.method = "L-BFGS-B", data = d1,
                 start = c(omega,5), lower = c(0,2.5), upper = c(0.5,15))
fit1

Call: fitCopula(copula, data = data, method = "ml", start = .3, lower = .4,
               upper = .5, optim.method = "L-BFGS-B")
Fit based on "maximum likelihood" and 2528 2-dimensional observations.
Copula: tCopula
rho.1    df
0.4937  9.8537
The maximized loglikelihood is 362
Optimization converged
```

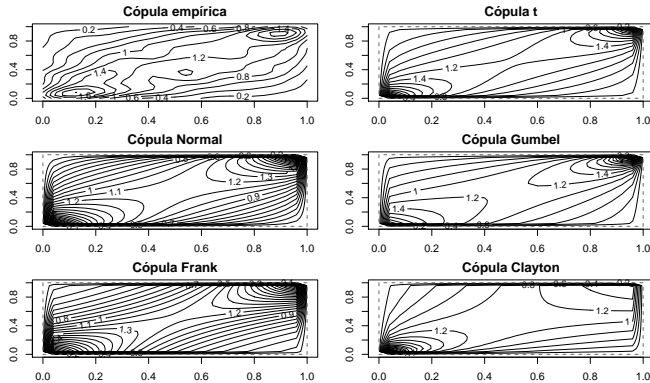
Para efectos de comparación, consideremos ahora el ajuste de otras cópulas a los datos:

```
#ajusta copula normal
fnorm <- fitCopula(data=d1,copula=normalCopula(-0.3,dim=2),method="ml",optim.method="BFGS",start=0.5)
#ajusta Gumbel
fgumbel <- fitCopula(data=d1,copula=gumbelCopula(3,dim=2),method="ml",optim.method="BFGS",start=2)
#ajusta Frank
ffrank <- fitCopula(data=d1,copula=frankCopula(3,dim=2), method="ml",optim.method="BFGS",start=2)
#ajusta Clayton
fclayton <- fitCopula(data=d1,copula=claytonCopula(3,dim=2), method="ml",optim.method="BFGS",start=2)
```

Las copulas estimadas se compararán con la cópula empírica y se verá si hay alguna estimación que quede cerca a la cópula que se obtiene de los datos. Podemos hacer una comparación gráfica:



En el siguiente conjunto de gráficas se comparará la estimación de la densidad bivariada entre las diferentes estimaciones paramétricas:



Por último, podemos comparar los AIC. El criterio de información de Akaike se utiliza para comparar modelos. Aquel con el mayor valor absoluto nos dice cuál es el mejor modelo. En este caso corresponde a la cópula t .

```
#AIC Normal
2*length(fnorm@estimate)-2*fnorm@loglik

[1] -692.3688

#AIC Gumbel
2*length(fgumbel@estimate)-2*fgumbel@loglik

[1] -624.4514

#AIC frank
2*length(ffrank@estimate)-2*ffrank@loglik

[1] -648.5734

#AIC Clayton
2*length(fclayton@estimate)-2*fclayton@loglik

[1] -584.2204

#AIC t
2*length(fit1@estimate)-2*fit1@loglik

[1] -719.9693
```


Conclusiones

En este artículo, se hizo una introducción muy general sobre el concepto de cópulas, su uso para modelar dependencia de variables aleatorias, se mostró un ejemplo de su aplicación en simulación y finalmente se mostraron los pasos a seguir mediante un ejemplo de cómo se puede estimar una cópula a partir de observaciones dadas. Dado el poco espacio, no se discutieron algunos ejemplos de cópulas importantes como las cópulas para modelar valores extremos, o tipos de dependencias como la dependencia en colas de las distribuciones [11]. El lector interesado puede consultar el libro muy reciente de Hofert, et.al [3].

Referencias

- [1] Joe, Harry. *Dependence Modelling with Copulas*. CRC Press, 2015.
- [2] Nelsen, Roger B. *An Introduction to Copulas, 2nd. Ed.*. Springer, 2006.
- [3] Hofert, Marius, et.al. *Elements of Copula Modeling with R*. Springer, 2018.
- [4] Erdelyi, Arturo. "Cópulas y dependencia de variables aleatorias: Una introducción", *Miscelánea Matemática* 48 (2009): 7-28.
- [5] Pickands, J. "Multivariate extreme value distributions", *Proceedings of the 43rd Session of the International Statistical Institute. Bull. Inst. Internat. Statistic* 49 (1981): 859-878, 894-902.
- [6] Yan, Yun. "Enjoy the Joy of Copulas: With the Package copula", *journal of Statistical Software* 21-4 (2007). <https://www.jstatsoft.org/index.php/jss/article/view/v021i04/v21i04.pdf>
- [7] Marius Hofert, Ivan Kojadinovic, Martin Maechler & Jun Yan. "copula: Multivariate Dependence with Copulas", *2018,R package version 0.999-19* <https://CRAN.R-project.org/package=copula>
- [8] "Abe Sklar" *Mathematics Genealogy Project*: . Consultado 15 de enero de 2019. <https://www.genealogy.math.ndsu.nodak.edu/id.php?id=1313>
- [9] Wuertz, Diethelm & Tobias Setz & Yohan Chalabi. "fCopulae: Rmetrics - Bivariate Dependence Structures with Copulae", *2017,R package version 3042.82* <https://CRAN.R-project.org/package=fCopulae>
- [10] "R metrics" *R metrics*. Consultado 15 de enero de 2019. <https://www.rmetrics.org/software>
- [11] Carmona, René. *Statistical Analysis of Financial Data in R, 2nd Ed*. Springer, 2014.

Zona Olímpica

1. Sea ABC un triángulo agudo y sea M el punto medio de AC . Un círculo pasa por B y M y toca los lados AB y BC en P y Q respectivamente. Sea T el punto tal que el cuadrilátero $BPTQ$ es un paralelogramo. Suponga que T se encuentra en el circuncírculo del triángulo ABC . Determine todos los valores posibles de BT/BM .

2. Encuentra todas las ternas (x, y, z) de enteros positivos tal que $x \leq y \leq zy$

$$x^3 * (y^3 + z^3) = 2019(xyz + 2).$$

3. Encuentra todas las parejas ordenadas (a, b) tales que:

$$1/a + 1/b = 3/2019.$$

4. Sean x, y, z números reales distintos de 0 que cumplen $xyz = 1$. Pruebe la siguiente desigualdad:

$$x^2/(x-1)^2 + y^2/(y-1)^2 + z^2/(z-1)^2 \geq 1.$$

5. Sea f una función definida de los números reales en si misma que satisface

$$f(x+y) \leq yf(x) + f(f(x))$$

para todos los números reales x, y . Pruebe que $f(x) = 0$ para toda $x \leq 0$.

6. Supón que hay 1000 personas paradas en un círculo. Pruebe que existe un entero z con $100 \leq z \leq 300$ tal que hay un grupo de $2k$ personas tal que la primera mitad del grupo tiene el mismo número de niñas que la segunda mitad.

7. Determine todo los enteros $z \geq 2$ tal que toda n con $m/3 \leq n \leq m/2$ divide el coeficiente binomial $\binom{n}{m-2n}$.

Pregunta de Erdős

Sean x y y enteros positivos. Si $x^2n - 1$ es divisible por $2^n * y + 1$ para todo entero positivo n . Pruebe que $x = 1$.

Apología del matemático, o un bosquejo de una demostración de nuestro amor por nuestra disciplina

Héctor Garrido

Ex-alumno de Economía y Matemáticas Aplicadas, ITAM

“Cualquiera puede cocinar”

Auguste Gusteau

Nuestro lenguaje moldea nuestra realidad en la medida en la que es la estructura que se nos fue “impuesta” para pensar desde pequeños. Impuesta porque, una vez adoptada, nos es difícil pensar sin recurrir a las palabras. Entre comillas porque, si bien toda estructura es una restricción por definición estricta, es evidente que dicha estructura mejora nuestras capacidades de razonamiento y le dan un sentido más allá de uno mismo para poder comunicarnos con otros, ya que el lenguaje, al estructurar nuestro pensamiento, nos permite realizar construcciones mentales más sofisticadas (se le da forma a lo que pensamos), nos permite hacer razonamientos más allá de los límites de nuestra memoria por medio de la escritura y nos abre a la posibilidad de adquirir los conocimientos de otros al proveernos de un suelo común con el que nos podemos comunicar con ellos. Al abrirnos a sus conocimientos, también nos abrimos a sus experiencias que han forjado sus maneras de pensar.

Agregado a esto, el mismo mensaje no será interpretado de la misma manera por todas las personas, si no que al contrario, cada quien terminará con un mensaje distinto, pues el contenido neto de la información que recibimos de afuera (ya sea una conversación con alguien, ya sea una experiencia que no involucre a nadie más) se “contamina” de todo aquello que tenemos ex-ante en nuestra mente, proveyéndole de nuestro color personal. Por todo lo anterior, el lenguaje abona a nuestra identidad y nos da voz y memoria, tanto personal como colectiva.

A todo esto, el lector podría replicar:

¿Qué no se supone que este era un texto de matemáticas? de ser así, ¿por qué me estás hablando de lenguaje?

Le responderé diciendo que lo ya escrito podría verse como una motivación para el corazón de este ensayo, el cual se encuentra condensado en el siguiente enunciado que buscaré defender:

Las matemáticas, contrario a lo que la mayoría de la gente piensa, en esencia no son un asunto de aritmética, sino de gramática, o de lógica para ser más exactos.

Y es que al escuchar sobre ellas, la gran mayoría de la gente automáticamente las relaciona solo con sumas, restas y hacer cuentas en general, pero en realidad estudian las formas de pensamiento y todo lo que eso implica. Un camino que hace valer mi punto se puede encontrar al examinar nuestros primeros acercamientos a las matemáticas con algo de

cuidado.

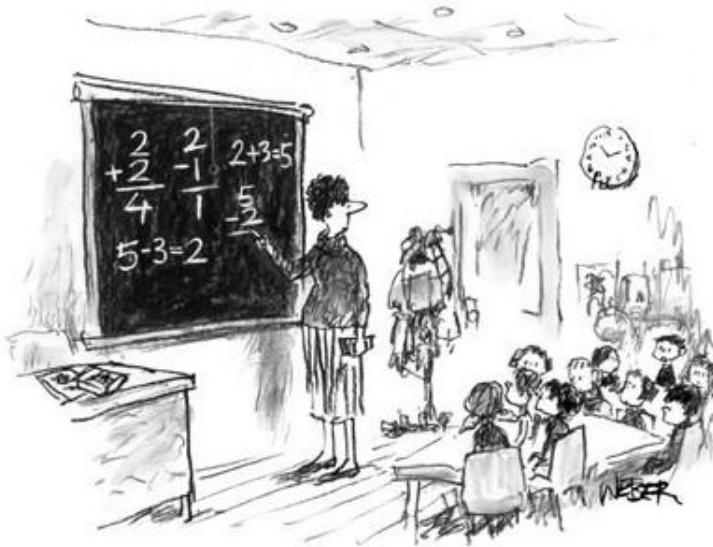


Figura 1: Autor: Weber

Figura 2: *Tenemos 132 manzanas. Si tenemos 12 personas, ¿cuántas manzanas le tocará a cada uno?*

Recordando nuestros años de primaria en los que los maestros nos enfrentaban a problemas en el pizarrón como el de la imagen, esas primeras experiencias eran un ejercicio de abstracción mayor al de las otras disciplinas. Si con el lenguaje (o la materia de español en el contexto descrito) comenzamos a jugar con el mundo real y sus abstracciones al asignarles nombres a las cosas que observábamos, con las matemáticas dimos un paso más al manipular dichos objetos en nuestra mente de una manera tal que pudimos llegar a conclusiones que no necesariamente habiéramos podido inferir con facilidad en el mundo real sin hacer uso de dichas herramientas.

Recordemos como, en alguna ocasión, todos llegamos a hacer cuentas con los dedos (muchas veces inútilmente) durante los exámenes ante la frustración de no saber emplear las técnicas que nos habían enseñado para resolver los problemas, tanto en aquel entonces como en nuestros parciales de Cálculo y Economía de ahora.

Si bien es evidente que todos aquellos que tuvieron un acercamiento inicial a las matemáticas exitoso y le agarraron gusto a las mismas no pensaron en todo esto en aquel entonces (la primaria), sí tenían una intuición de ello, del gusto por hacer ejercicios mentales de abstracción, no meramente "hacer cuentas".

En la medida en la que uno se va acostumbrando a la jerga de la disciplina y va familiarizándose con ella a fuerza de práctica y ejercicios, poco a poco, sin darnos cuenta, la intuición

que mencionamos arriba va madurando en nuestra cabeza para convertirse en la idea de que la matemática es un metalenguaje que, a la par que se nos enseña, nos impulsa a cuestionar construcciones mentales que el resto asume de manera implícita en su quehacer diario. Y esto último se debe a que (regresando a la motivación inicial), sin importar nuestro origen, nuestro lenguaje materno se nos fue enseñado nombrando las cosas que observábamos, de afuera hacia adentro. Al limitar nuestras primeras enseñanzas del lenguaje a lo observable, el cual a su vez condiciona nuestros pensamientos, restringimos de manera silenciosa nuestra manera de pensar y de concebir el mundo. Si Wittgenstein dijo que los límites de nuestro lenguaje eran los límites de nuestro mundo¹, yo tomo su enunciado en sentido inverso: los límites de nuestro mundo han marcado los límites de nuestro lenguaje, y con ello los límites de nuestro razonamiento y de aquello que podemos aportar al mundo.

Ahora bien, con las matemáticas siento que opera a la inversa, de adentro hacia afuera. En ellas, primero empezamos por aprender sus “reglas gramaticales” y a partir de estas, por medio de tareas y ejercicios, fuimos empujados a construir estructuras mentales que en más de una ocasión, debido a su carácter, superaban la barrera de nuestro mundo inmediato observable y con ello de nuestras experiencias sensoriales (objetos de muchas dimensiones, diferentes tipos de infinitos, expresiones geométricas irreverentes pero congruentes en sí mismas, etc.), proveyéndonos de una ventaja especial a la hora de pensar e imaginar problemas en comparación de aquellos que sólo poseen la herramienta de pensamiento convencional derivada del lenguaje común.

De ahí que los matemáticos digan que su disciplina involucre mucha creatividad, e incluso algunos se aventuren a decir que es una forma de poesía², cosa que ha hecho a más de uno fuera del gremio levantar una ceja. Volteando a Wittgenstein de vuelta a su forma original, considero que aquí más que en ningún otro lado tiene sentido la frase “imaginar un lenguaje significa imaginar una forma de vida”. [2]

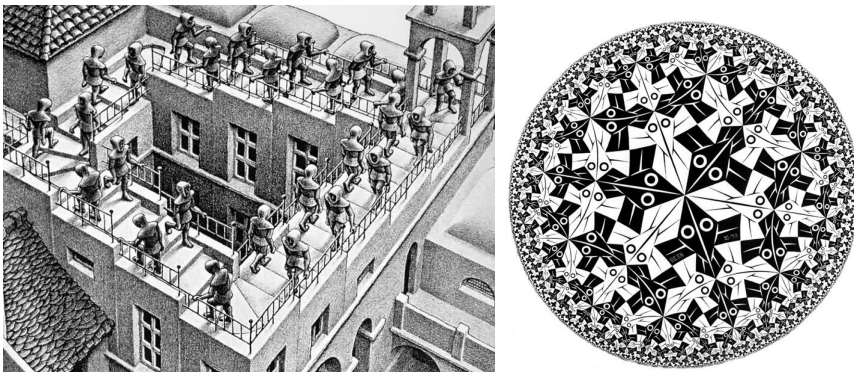


Figura 3: Ascending and Descending / Circle Limit I, M. C. Escher

¹“Los límites de mi lenguaje significan los límites de mi mundo.” [1]

²En mi opinión se acercaría más a una forma de pintura

Y creo que esto último captura bastante bien la razón por la cual los matemáticos le tenemos tanto cariño a nuestra disciplina. Es difícil verla sólo como nuestra herramienta de sustento. Esta permea más allá en nosotros, en nuestras maneras de pensar y de comunicarnos con el mundo y con los demás. Más que una profesión, lo sentimos como parte de lo que nos define como personas y al mismo tiempo de lo que nos hace humanos, y es por esto último que repetimos hasta el cansancio que las matemáticas no son sólo para unos cuantos, si no que son para todos y están dentro de todos nosotros, así como el lenguaje y las artes. De ahí el sentido de la frase citada al inicio de este artículo.³

Sin embargo, admito que muchos sufren al acercarse a ellas. Como todo nuevo idioma que uno comienza a aprender, el inicio es complicado para (casi) todos. Pero también, como todo nuevo idioma, a base de práctica y constancia lo hacemos nuestro. Por ello, quisiera cerrar con la siguiente reflexión del maestro Ramón Espinoza [3], que demuestra que más de uno de nosotros cree en la universalidad las matemáticas y su valor:

Las matemáticas no son fáciles (ni siquiera para los matemáticos profesionales), y uno no se debe desilusionar si siente que no puede avanzar tan rápido como quisiera. El trabajo constante y sistemático tarde o temprano comienza a rendir frutos. Con el tiempo se aprende a disfrutar de las matemáticas, de la misma manera que se puede disfrutar de la música o la literatura.

Referencias

- [1] Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. 1921.
- [2] Wittgenstein, Ludwig. *Investigaciones Filosóficas*. 1953.
- [3] Espinoza Armenta, Ramón. *Matemáticas Discretas*. Alfaomega Grupo Editor, 1^o Edición, 2010.

³Una disculpa, pues su interpretación, al igual que con Wittgenstein, fue tomada de manera distinta a la de su sentido original.

Matemáticas prehispánicas: Entrevista con Blair Madore

María José Domenzain
Guillaume Domenge
Paula Langle

Blair Madore es Doctor en Matemáticas, y profesor del departamento de matemáticas de la universidad estatal de Nueva York, SUNY Potsdam. Algunos de sus intereses incluyen la teoría Ergódica, la historia de las matemáticas, principalmente aquellas desarrolladas en Latinoamérica durante la época prehispánica, la teoría de grupos, y la teoría de operadores.

Laberintos e Infinitos (L&I) - ¿Por qué decidiste estudiar matemáticas?

Blair Madore (BM) - Saben, cuando estaba en preparatoria, siempre pensé que estudiaría Computación, Física, o algo por el estilo. Cuando entré al último año de preparatoria yo era muy, muy bueno con las matemáticas, más que en cualquier otra materia. Incluso había una competencia de matemáticas con las otras preparatorias de mi ciudad, y yo estaba en el equipo de mi escuela; a pesar de que no ganamos, a mi equipo le fue muy bien. Así es que yo me sentía bastante seguro cuando se trataba de matemáticas, y eso era lo que quería hacer.

Uno de mis profesores pensaba que debía ser ingeniero, y pensé que sonaba interesante, y más tarde me di cuenta de que la razón por la que no estudié ingeniería es porque no sabía exactamente lo que era, ¿saben? Vengo de un ambiente muy modesto: mi papá era carpintero y mi mamá ama de casa, mis tíos eran mecánicos, pintores, cosas por el estilo. Personalmente, no conocía a nadie que fuera un ingeniero, así que no tenía ni idea de lo que se trataba.

(L&I) - Si tuvieras que escoger de nuevo, ¿volverías a escoger Matemáticas?

(BM) - ¡Por supuesto! Tal vez Matemáticas no es el campo mejor pagado, pero probablemente es el más estimulante e interesante, y al compararla con otras áreas, una de las cosas lindas de las matemáticas es que está separada de la política, ¿no?

En cuanto a comunicación, es complicado evadir grandes diferencias de opiniones, pero en Matemáticas, las matemáticas hablan por sí mismas: las matemáticas están bien, o están mal. Lo único que se puede argumentar es qué tan importantes son, y eso normalmente no es tan problemático hasta que de conviertes en una súper estrella de clase mundial, lo cual yo no soy... Así que no tengo que preocuparme por eso.



(L&I) - ¿Tienes algún consejo para los estudiantes de Matemáticas? De ser así, ¿cuál sería?

(BM) - Primero que nada, que tienen que trabajar duro. Mucha gente piensa que les será fácil porque les iba bien en matemáticas. No importa quién seas, puedes ser el mejor estudiante, y aún así tendrás que trabajarle, y sacarás provecho de ello.

Visiten a sus profesores y háganles muchas preguntas, no se queden en casa y digan "Tal vez entendí esto, tal vez no", sus profesores son las mejores personas para ayudarles a realmente entender las cosas. Muchos estudiantes tienen miedo de ir a hablar con ellos. Ahora, a los profesores les gusta hablar con los estudiantes, por eso son profesores. Si no les gustara platicar con ellos, serían matemáticos aplicados trabajando en la industria, o en el gobierno, o algo por el estilo; a los profesores en realidad les gusta platicar con los estudiantes. Nuestras vidas serían muy pobres sin ellos, los estudiantes llenan nuestras vidas de riqueza. Así que sí, si cien estudiantes vinieran a verme en un día, tal vez sería mucho, pero si ninguno viene, entonces es algo decepcionante.

¿Otro consejo? Siempre es bueno estudiar campos complementarios a Matemáticas. Pueden estudiar matemáticas por sí mismas, pero en realidad las matemáticas son más poderosas cuando las usan en servicio de otras áreas, por ejemplo Computación, Biología, Economía, Finanzas... Quiero decir, pueden escoger casi cualquier ámbito, y aplicarle matemáticas, y eso es increíblemente poderoso; usualmente se darán cuenta que si deciden enfocarse en Biología, por ejemplo, y tienen cientos matemáticos, entonces podrán investigar en áreas que el resto de sus compañeros no podrán porque no tienen las bases matemáticas que ustedes tienen. Ustedes podrán hacer todo lo que ellos hagan, pero ellos no podrán hacer todo lo que ustedes puedan hacer. Las Matemáticas te dan un poder que no tendrías de otra forma.

Yo estudié en la Universidad de Waterloo en Ontario, y ahí tienen una Facultad de Matemáticas con al rededor de 5 mil estudiantes. En la Facultad estudian Computación, Matemáticas Aplicadas, Matemáticas Puras, Estadística, Actuaría, Matemáticas en la Industria y Comercio, entre otras. Yo tuve la oportunidad de hacer un doble título en Matemáticas Puras y en Combinatoria y Optimización. Optimización es la extensión natural de lo que se aprende en Cálculo: el cálculo de mínimos y máximos en un nivel más amplio, y con mayor número de condiciones. Combinatoria es una materia que no se estudia mucho a nivel licenciatura: es todo aquello que tiene que ver con contar. Mi doble título encaja con esta noción de "No sólo estudies una cosa, haz también otras cosas", ¿verdad?

(L&I) - ¿Tienes un teorema favorito? Si la respuesta es sí, ¿cuál es?

(BM) - Creo que mi teorema favorito es el Teorema Ergódico. Es el teorema fundamental que establece las bases del área de investigación en la que trabajo y en la que hice mi doctorado. El teorema surgió alrededor de 1890, de personas como Boltzmann que estaban estudiando los gases. Hay ciertas propiedades de los gases, así como de otros objetos, que varían dependiendo del estado físico en el que se encuentran. Piensen por ejemplo: si tomaran la temperatura de una piedra, ésta seguirá siendo la misma si se esperan un par de minutos. O su masa, ésta seguirá siendo la misma después de un tiempo. Pero con los gases hay ciertas cosas, como la presión que tienen, que varían con el tiempo y que es interesante

medir de la misma forma que se mide la gravedad, por decir algo. Así pues, la idea de Boltzmann era "Si tomo una medida cada segundo, aunque las magnitudes cambien, puedo tomar el promedio de éstas durante mil segundos y ver qué pasa". Pero entonces alguien le dijo, "Espera, ¿por qué mil segundos y no cien? ¿Por qué no diez mil?", así que él dijo "Bueno, si hago esto el tiempo suficiente -y yo creo que mil segundos es tiempo suficiente-, entonces finalmente las diferentes medidas estarán arbitrariamente cerca de cada medida posible". Así, él pensó que naturalmente todos los gases pasarían por ciertos estados que estarían cerca de todos sus estados posibles, en todas las configuraciones de moléculas posibles, y a la velocidad correspondiente; lo cual es una conjetura bastante radical. No tenía ninguna prueba de esto, ni ninguna justificación, pero hizo sus medidas de esta forma y obtuvo resultados sorprendentemente precisos, que le permitieron hacer aplicaciones en otras ciencias y le funcionaron.

Durante treinta años la gente sospechaba que lo que él había hecho era correcto, pero no había prueba científica o matemática que lo fundamentara. Entonces, curiosamente en 1932 dos matemáticos, cada uno por un camino diferente, probaron que Boltzmann estaba en lo correcto. Uno de ellos lo probó utilizando funciones de tipo L^2 , en espacios de Hilbert, y otro en espacios de L^p ; sus pruebas eran completamente diferentes, y aparecieron con 6 meses de diferencia entre sí, lo cual es sorprendente.

La mejor manera de describir el Teorema Ergódico, en términos que tengan sentido para la mayoría de las personas es: en Calculo aprendes a hacer integrales de Riemann, divides el intervalo en partes, mides cada parte, las sumas y obtienes el área, tienes n intervalos con un valor de $\frac{1}{n}$. Lo que el Teorema Ergódico dice es que, en vez de dividir el intervalo en $\frac{1}{n}$, usa una transformación Ergódica que saltará de un punto a otro, de manera aparentemente aleatoria, y que finalmente visitará los puntos suficientes del espacio que el promedio de ellos será de la misma magnitud que el intervalo.

Cuando la gente aprende acerca de transformaciones Ergódicas, empiezan a ver propiedades como las mezclas. La idea es, en términos simples, como cuando tienes una bebida con alcohol y jugo, y mezclas buscando que el alcohol esté presente en toda la bebida. ¿Llegará a pasar esto? ¿Cuánto tiempo tendrás que mezclar para que suceda? Bueno, esta propiedad lo que dice es que si tomas dos conjuntos, y uno de ellos lo mueves, la proporción de veces que tocará al otro conjunto es exactamente igual al tamaño de la intersección de ambos conjuntos. Este tipo de propiedades se usan para probar cosas relacionadas con las tormentas, por ejemplo.

(L&I) - Escribiste un libro para estudiantes de preparatoria. ¿Es difícil tomar cosas en un nivel elevado, y aterrizárlas de forma que sean comprensibles para estudiantes de ese nivel?

(BM) - El proceso de escribir ese libro fue muy interesante porque en realidad uno no escribe los libros de textos, y mucho menos para preparatoria, de la misma forma que escribes libros para universidad. Usualmente, cuando se escribe un libro para la universidad hay uno, o dos profesores que trabajan en él juntos; cuando se escribe un libro para preparatoria, hay alrededor de veinte personas trabajando en él. En esencia lo que se hace es que se le asigna a cada persona un capítulo, con una lista de lo que éste debe incluir, y luego una

persona lee todos los capítulos, y ajusta el estilo para que la lectura fluya como un todo. Afortunadamente ese era mi trabajo, yo era el editor que tenía que leer todos los capítulos y cambiar la forma en que se expresaban ciertas cosas para que pareciera que el libro había sido escrito por una sola persona.

Las matemáticas son difíciles de enseñar, no importa el nivel. Sin embargo, a veces el nivel de preparatoria es más difícil que el nivel de la universidad. Tal vez no quieran escuchar esto pero cuando un profesor les enseña algo, ya sea sobre Matemáticas o cualquier otra ciencia, no les dicen toda la verdad; únicamente les dicen la porción de la verdad que ustedes están listos para entender, y a menudo hay cosas que se relacionan con eso, y que ustedes desconocen y no están listos para entender. La primera vez que aprendieron a factorizar polinomios, todos los polinomios se factorizaban de forma sencilla, normalmente habían dos raíces con fracciones bonitas o números enteros. Bueno, eso no siempre es verdad, hay miles y miles de polinomios que no se pueden factorizar, y no siempre los factores están en el campo de los números reales: entonces hay que enseñar sobre los números complejos. Lo mismo para con otras materias como Cálculo, cuando les enseñan sobre integración y les dejan con la impresión de que cualquier función con la que se encuentren es integrable, y eso no es cierto.

En cierta forma, los profesores escondemos la totalidad de la verdad, y conforme van aprendiendo, van estando más preparados para aprender todavía más cosas. Es por esto que creo que es difícil enseñar en preparatoria, porque tienes que tomar todas las cosas que los alumnos no están listos para aprender, hacerlas a un lado, y tomar las cosas que sí están listos para entender y hacerlas ver interesantes, ¿saben? Hay mucha gente que enseña matemáticas a nivel preparatoria, o niveles inferiores, y lo hacen ver sumamente aburrido porque todo lo que enseñan es "Ah, suben este número, multiplican por esto, ponen esto abajo..." Si, eso no es interesante, y así a mí tampoco me gustaría estudiar. Como profesor tienes que hablar de los *por qué*s, por qué esto tiene sentido, para qué te va a servir esto. Si no hablas de eso, las Matemáticas se convierten en un puñado de procedimientos aburridos de los que nadie quiere saber nada.

Pero en la universidad ustedes están listos para escuchar casi todo. Están en el punto en el que si les enseñan algo y les dicen "Por cierto, esto no es todo, vamos a profundizar más", ustedes se sienten cómodos con ello, de hecho hasta se sienten emocionados de saber que hay más que aprender. En preparatoria no están emocionados, tienen miedo. En preparatoria se trata de despertar el interés de los estudiantes sin intimidarlos o asustarlos.

(L&I) - ¿Podrías hablarnos un poco sobre tu investigación de las Matemáticas en América Latina?

(BM) - Cuando recién me gradué de la universidad, y estaba haciendo investigación en varios campos, hice algo de investigación con el Doctor Martínez Avendaño. Alrededor de ese tiempo comencé a venir a México. La primera vez que estuve en México tuve unas excelentes experiencias conociendo gente, lugares interesantes, y quería que mis estudiantes en Estados Unidos tuvieran un poco de las mismas experiencias. Traté durante dos o tres años de organizar alguna forma de traer un grupo de estudiantes durante una semana y tener alguna conferencia de Matemáticas con estudiantes mexicanos y estadounidenses;

pero cada vez que intentaba organizar esto no había suficiente dinero para llevarlo a cabo. Entonces alguien me dijo "Si quieres llevar estudiantes estadounidenses fuera, tienes que organizar un curso y hacer que paguen por él", y eso funciona porque de alguna forma, ellos están mucho más dispuestos a pagar por algo de lo que se van a beneficiar. Obtienen experiencia universitaria, y obtienen tres créditos por el curso. Así que sí, cuesta más porque tienen que pagar por los gastos del viaje, pero la estructura del curso, y la forma en que los estudiantes lo costean funciona. Por ejemplo, si es un curso, y los alumnos tienen préstamos estudiantiles, entonces el préstamo les permite pagarlo porque se trata de un curso formal; pero si no fuera así, aunque se tratara de una experiencia académica, los préstamos no pagarían por ello.

Así pues, empecé a traer a mis estudiantes a la Universidad de Hidalgo hace seis años. La primera vez que venimos fue en 2013, y en cuatro ocasiones diferentes trajimos un grupo de estudiantes a México, y pasamos dos semanas aprendiendo sobre la historia de las Matemáticas alrededor del mundo. ¿Por qué hago esto? Porque es el único curso que puedo dar y que me permitía traer estudiantes a otro lugar del mundo. Usualmente, los cursos que permiten viajar son clases de Arte, donde visitan museos como el Louvre. ¿Un curso de Matemáticas que te permite viajar?, nadie había escuchado de algo parecido. Este era el único curso que podía dar y que permitía viajar. Así, con una de mis colegas creamos el curso de Historia de las Matemáticas que exploraba las matemáticas de Egipto, Mesopotamia, China, de algunas culturas árabes, India... Pero no sólo eso, también las matemáticas de las culturas maya, azteca, olmeca, tolteca, otomí, mixteca, y zapoteca; de forma que mis estudiantes obtienen una amplia gama de las culturas prehispánicas.

Cuando estoy en Estados Unidos, y platico con profesores que se especializan en la Historia de las Matemáticas, muy pocos de ellos se enfocan en las Matemáticas de las culturas de Mesoamérica. Hay mucha área de oportunidad ahí. He sido algo flojo y no he escrito mucho sobre ello, pero siento que aún estoy aprendiendo. En cuanto a Mesoamérica, no se conoce mucho, y a la par tienes que aprender acerca de Atropología, así que en realidad tienes que adentrarte en lo que se conoce de las culturas y conectar los puntos sobre qué es lo que está pasando.

Para que se den una idea de lo complicado que esto puede ser, una de las cosas que es fascinante de las culturas mesoamericanas es que tenían muchos calendarios diferentes, ¿saben? Nosotros usualmente utilizamos el calendario conocido como Calendario Gregoriano, que surgió en los tiempos del Papa Gregorio, de quien recibe el nombre. En las culturas mesoamericanas, sé de doce calendarios que se usaban de forma regular. Algunos de los principales calendarios que utilizaban eran los mayas, había uno que contaba los días desde el origen de los tiempos; había algo llamado el Calendario Sagrado, que cuenta un ciclo de 260 días, basado supuestamente en el periodo de gestación del feto humano; alguna gente que se dedica a la Astrología lo llaman el Calendario Astrológico Maya, pero eso no es justo porque no es realmente cierto. Existe, también, el Calendario Solar, el cual es similar al calendario anual que utilizamos. Solían poner el Calendario Solar y el Sagrado juntos para conseguir un calendario que representaba de forma única cada día durante un periodo de 52 años.

Esto se puede complicar más, y tratar de entender cómo usaba los calendarios cada cultura, y tratar de asegurarse de estar en lo cierto ha llevado a la gente especializada en es-

tudios mesoamericanos a encontrarse una y otra vez con información que no encaja. Hay muchas razones para creer que una cierta fecha en el calendario Gregoriano corresponde a una fecha maya determinada, pero los expertos pueden mostrar diez contraejemplos que contradicen estas correspondencias, a veces hay errores de un día, a veces el error es mayor. Lo mismo pasa con cada uno de los calendarios; para poder compararlos hay que saber cuál corresponde a cuál; hay quienes piensan que los errores vienen de la forma en que llevaba cada comunidad el calendario. Hay investigadores que piensan, y esto me parece realmente fascinante, que el cambio de día en los calendarios ocurría a diferente hora del día. Piensan que uno cambiaba al amanecer, mientras que otro lo hacía a medianoche. ¡Este tipo de cosas son realmente fascinantes!

Si quieren ver más Matemáticas en esto, de alguna forma esta gente, hace 1200 años o más, era capaz de usar sus calendarios y calcular de forma más precisa que la información que tenemos de hace cien años cuánto tiempo le toma a Venus dar la vuelta al Sol. ¿Cómo le hacían? Aún no sabemos. La parte triste es que cuando los españoles llegaron, y después de que conquistaron, hicieron un gran esfuerzo en destruir la religión indígena, pero más que eso, ¡estaban destruyendo la cultura indígena! Punto. Habían miles y miles de libros que había sido escritos en piel de cabra que fueron quemados, en particular hubo un incendio masivo donde quemaron miles de estos libros a la vez. Al día de hoy, cuando hablas de escritos mayas, sólo hay cinco que sabemos que existen, y esto por suerte. Escritos mayas hay más, no se exactamente cuántos pero por lo menos veinte o treinta; así que la evidencia que tenemos es mínima. Hay que añadir que tanto los mayas como los aztecas esculpían muchas cosas en piedra, y de esto aún hay varias evidencias.

(L&I) - En tu investigación sobre las Matemáticas mesoamericanas, ¿alguna vez te encontraste con algo que te hiciera pensar "¿Cómo puede ser que no haya más gente que sea sobre esto?¿

(BM) - Esta es una muy buena pregunta. Tengo dos ejemplos de esto, y ambos vienen del curso que doy a mis alumnos. Cuando vamos al Templo Mayor hay una gran estatua de Tlaltecuhltli, que es la diosa de la Tierra... Bueno, dios/diosa. Tiene una identidad dual, y la mayoría de las veces no ves una imagen de esta deidad porque solía ser esculpida en la parte inferior de las estatuas de forma que quedaba hacia el piso; se supone que no debía haber imágenes de esta deidad expuestas al aire porque era la diosa de la Tierra y tenía que estar en contacto con ésta todo el tiempo. Es fascinante. En fin, en 2003 encontraron esta gran escultura, y cuando la encontraron, el arqueólogo que estaba a la cabeza en ese momento, e incluso el arqueólogo principal del INAH, afirmaron que se trataba de la marca de la entrada a la tumba de un emperador azteca; y la razón por la que pensaban que este era el caso, es porque en la garra de la estatua había una fecha azteca esculpida. Tal vez no lo sepan, pero los aztecas solían dar un nombre a cada año, basado en la fecha del Calendario Sagrado; lo que estaba esculpido en la garra, era la imagen de un conejo, a un lado del conejo habían diez círculos, y al otro lado dos círculos. Los representantes del INAH dijeron que esto representaba el año 1506 o 103, que es alrededor del tiempo antes de que Moctezuma muriera. Así pues, estaban seguros de que esta iba a ser la entrada a la tumba de Moctezuma.

Entonces, todo ese tiempo yo había estado llevando a mis estudiantes y les había estado diciendo: "Yo no sé ustedes, pero a mi parecer esto dice 12 conejo, no 10 conejo. ¿Podrían calcular para mí qué año año sería 10 conejo, y qué año 12 conejo?".^{Esto es parte de lo que les enseñamos, cómo convertir estas cosas a nuestro calendario actual. De hecho, porque ese calendario es de un ciclo de 52 años, una vez que encuentras un año, bien podría ser 52 años en el futuro, o 52 años atrás. Entonces les dije a mis alumnos que investigaran a ver si encontraban algún evento importante en el Imperio Azteca que sirviera de motivación para escribir ese año. No había nada que fuera de interés particular.}

En 2017, o 2018, el director del INAH admitió que estaban equivocados. Me sentí reivindicado, ¡estaba tan emocionado! ¡Era 12 conejo! Bueno, no dijeron eso exactamente, aun no saben lo que significa, pero admitieron que estaban equivocados sobre 10 conejo y sobre el año que representaba. Fue genial para mis estudiantes porque estaban tomando un curso de primer año, e inmediatamente habían sido envueltos en un problema que los llevó en medio de un tema sumamente importante en el ámbito de la investigación moderna de antropología, y sin el entendimiento matemático suficiente, no se puede estudiar ni investigar sobre eso. De echo, estoy convencido de que la razón por la cuál los antropólogos cometen este tipo de errores es porque no saben matemáticas, deberían saber más sobre matemáticas.

Ese es un ejemplo, el otro ejemplo es similar a ese. Supongo que han ido al Museo Nacional de Antropología en Chapultepec. Nosotros siempre vamos ahí, y a mí me gusta que mis alumnos traduzcan algo que está en maya. Hay un sinfín de artefactos con fechas mayas en ellos, pero hay uno que es el más visible de todos, se trata de una gran estela maya que se encuentra en medio del jardín. La primera vez que les pedí que tradujeran esto, yo ya lo había hecho de antemano y me había quedado sorprendido con mi resultado, pero lo chequé, y lo revisé con un colega, y yo definitivamente estaba en lo correcto. La fecha en la estela maya era 11 o 12 de septiembre del 1945, lo cual no tiene nada de sentido. Traté de entender por qué, pedí ayuda en el museo, y nadie podía ayudarme a entender. Entonces me rendí, y les asigné a mis alumnos la misma tarea, y ¡ellos llegaron a la misma respuesta! Todos pensamos "¿Qué? Algo debe estar mal alguien debe haberse equivocado". La próxima vez que visité el museo, descubrí que de acuerdo con los encargados del museo, la fecha en la piedra debía de ser septiembre de 1967, que es la fecha en la que abrió el museo, lo cual tendría mucho más sentido, ¿no? Desde entonces he tratado de varias formas entender cómo pudieron haber hecho un error de cálculo tal que terminara dando como resultado 1945 en lugar de 1967. Creo que los mexicanos merecen saber que la fecha en la estela maya está mal.

En fin, eso es todo.

(L&I) - ¿Algún último comentario que quisieras agregar?

(BM) - Es un verdadero placer para mí el venir y platicar con estudiantes. Mi universidad no es una universidad de investigación per se, nuestro principal enfoque es trabajar con los alumnos, así que donde quiera que vaya, siempre trato de hablar con algunos estudiantes para darme una idea de cómo es la escuela y cómo son los estudiantes. Hasta ahora ha resultado ser muy fascinante.

La flexibilidad de las matemáticas: Entrevista con Rafael Prieto

Consejo Editorial

Rafael Prieto Curiel estudió la licenciatura en Matemáticas Aplicadas en el ITAM, donde fue director de Laberintos & Infinitos entre 2008 y 2009. En su reciente visita a México, nos entrevistamos con Rafael y tuvimos la oportunidad de que nos platicara de su experiencia después de este periodo, y nos hablara sobre sus opiniones acerca de las Matemáticas y sus aplicaciones, así como su relación con temas sociales.

Rafael Prieto (RP) - Creo que el tema crítico de las matemáticas es fijarnos siempre en la equidad de género y asegurarnos de que estemos haciendo siempre lo correcto para abrir espacios y no cerrarlos.

Me da mucho orgullo ver que en el ITAM hay una preocupación latente por tener equidad de género; en particular en el Departamento de Matemáticas las mujeres siempre han sido la fuerza, la virtud, las profesoras con más empuje, y siempre son las que van llevando las cosas adelante. No por minimizar a los profesores, sino por enfatizar que aquí la condición de "soy mujer" no te pone en una posición de menos privilegio, lo cual es algo fascinante del Departamento.

Después de Laberintos & Infinitos, Rafael comenzó a trabajar en la publicidad durante cinco años. Trabajó en el Centro de Atención a Emergencias y Protección Ciudadana de la Ciudad de México (C4 en aquella época o C5 actualmente) como Director de Análisis Estratégico, pronosticando delitos para hacer los recursos de seguridad de la Ciudad más eficientes. Más adelante estudió una maestría en Estadística y un doctorado en Matemáticas, Seguridad y Crimen en Londres en University College London, UCL.

RP - Recuerdo que el primer año de la maestría me dio el "síndrome del impostor". De todas las personas que me rodeaban, yo sentía que era el menos capaz y el menos preparado de ellas. Sentía que no tenía nada que hacer ahí. Conforme pasó el tiempo, entraron nuevas personas, y el sentimiento se fue reduciendo. De digamos mil personas, yo ya me sentía al rededor de la posición doscientos. Además, conforme platicaba con mis colegas, me di cuenta que no era el único que se sentía así. Casi todos pasamos por lo mismo, y casi todos compartíamos esa sensación de no ser lo suficientemente buenos para estar en donde estábamos.

Otra anécdota es cuando estaba trabajando en la investigación para mi doctorado. Trabajé con una infinidad de datos sobre el crimen en México, tenía super computadoras que me



prestó la Universidad para procesar toda esa cantidad de información corriendo día y noche durante semanas. Además, la investigación tenía límite de tiempo, por lo que tenía que trabajar en mi investigación los siete días de la semana. Hubo momentos en los que pensé que no acabaría a tiempo. Recuerdo una frase de mi asesor cuando fui con él a decirle todo lo que me estaba costando procesar la información con la que estaba trabajando: "If that's what it takes, that's what it takes. You wanted a PhD, right?".

Durante su tiempo en UCL, y utilizando los conocimientos que *Laberintos & Infinitos* le dio durante su tiempo como director editorial, Rafael fundó la revista *Chalkdust*.

RP - Me acuerdo de una Navidad que vine en mi primer semestre de doctorado. Pasé por el ITAM a saludar, y por suerte me llevé conmigo algunas revistas de *Laberintos & Infinitos* a Londres, y con mis amigos del Doctorado les dije: "Tenemos que hacer esto en Inglés". Ellos vieron la revista, y les encantó la idea; nos costó un poco de trabajo, pero a inicios del 2015 iniciamos una revista que se llama *Chalkdust* que, siguiendo la filosofía de *Laberintos & Infinitos* permite que los estudiantes escriban sus ideas interesantes, chistes, cómics, rompecabezas y todo.

Las matemáticas, aunque no son para todos, no tienen por qué ser difíciles, oscuras, aburridas o tediosas. Te puedes encontrar problemas de todo tipo: cáncer, calentamiento global, tráfico, medio ambiente, venas en los ojos, ¡de todo hay problemas! Y además son bellísimos, cada vez que me encuentro con alguno de estos problemas digo ¡sí, por esto soy matemático! Porque te da una satisfacción máxima ver que con matemáticas se puede resolver muchísimas cosas. Por ejemplo en mi caso, ver que se puede luchar contra el delito, y aplicarlo al crimen para poder sentarme con el jefe de la policía y poder decirle "esto no", porque tengo matemáticas atrás que lo muestran y lo respaldan.

De mi experiencia de llevarme de *Laberintos & Infinitos* a *Chalkdust* es lo que pude exportar de México al Reino Unido. *Chalkdust* es una revista tanto digital como impresa, de cada edición imprimimos 4 mil copias, sin embargo este impacto es muy fijo y limitado. Por otro lado, el impacto digital le llega a 70 mil personas alrededor del mundo, en Estados Unidos, Reino Unido, Nueva Zelanda, entre otros.

De *Chalkdust* tuve una experiencia padrísima, la revista sigue porque la idea es, al igual que en *Laberintos & Infinitos*, tener un equipo que pertenece a una generación y que se va reciclando. Yo ya tengo más de un año de no pertenecer a *Chalkdust*, pero ellos siguen, y seguirán nuevas generaciones y seguirán publicando.

A pesar de que los miembros actuales del comité editorial son muy buenos amigos míos, estoy consciente de que pasará lo mismo que me pasó con *Laberintos & Infinitos*, fui director de 2008 a 2009 y cuando yo dejé la revista poco a poco fui dejando de conocer las personas que conformaban cada una de las generaciones encargadas de la revista.

Laberintos & Infinitos ha pasado de generación en generación de manera increíble, llevar 17 años haciendo esto es impresionante. Imagínense que en el equipo hemos estado 200 o más personas, y eso es increíble porque somos generaciones que nos han tocado de todo. Inclusive, mi primer trabajo fue con los fundadores de *Laberintos & Infinitos* en el año 2009 en una consultoría que ellos pusieron después de fundar *Laberintos & Infinitos*. Y aunque ellos hoy en día están casados, tienen hijos y familia, los veo con gusto porque fueron una

semilla y cada generación va dejando algo.

También algo padre es que cada generación ha hecho algo diferente, entonces cada una de ellas le da su toque personal, ves las portadas de las revistas y dices "esto es una generación nueva, este es su estilo" eso está increíble.

Además fíjense, bajita la mano han publicado 800 artículos, eso hace de ella una revista enorme. Compíte a nivel nacional con *Miscelánea Matemática* en cuanto a cantidad de artículos. En realidad yo pienso que, en México *Laberintos & Infinitos* es una de las revistas de Matemáticas más grandes que hay. Con 800 artículos publicados, tienen un acervo de artículos súper interesantes, entrevistas con personas increíbles. Por ejemplo, vino el matemático Cédric Villani y ustedes lo entrevistaron, ¡eso está increíble! ¿Cuántas personas tienen la oportunidad de sentarse y platicar con él? Y el equipo de *Laberintos & Infinitos* fue uno de ellos.

Actualmente, Rafael es investigador postdoctoral en la Universidad de Oxford en el Mathematical Institute en un proyecto sobre ciudades, llamado Peak Urban. Durante el 2019 estará en Colombia realizando investigación sobre matemáticas aplicadas a temas como migración y crimen en dicho país.

Además, Rafael trabaja para la Organización para la Cooperación y Desarrollo Económicos (OCDE), donde realiza análisis sobre políticas de crecimiento urbano y demográfico en el Oeste de África.

Para saber más pueden consultar su blog: <https://rafaelprietoc.wordpress.com/>