

**Departamento de Matemáticas**

# **Memorias**

**de la XIX Semana Regional  
de Investigación y Docencia  
en Matemáticas**

Gabriel Alberto García Mireles  
Silvia Elena Ibarra Olmos (Eds.)

**Mosaicos Matemáticos No. 32  
Noviembre de 2009**

**Universidad de Sonora  
División de Ciencias Exactas y Naturales**

QA11.A1  
.S44  
2009

Semana Regional de Investigación y Docencia en Matemáticas (19 : 2009 : Hermosillo, Sonora)  
Memorias de la XIX Semana Regional de Investigación y Docencia en Matemáticas; Hermosillo, Sonora, noviembre de 2009 / Gabriel Alberto García Mireles, Silvia Elena Ibarra Olmos (Eds.).— Hermosillo, Sonora : Editorial Universidad de Sonora. Departamento de Matemáticas, 2009.  
184 p. : Il., Fotog.; 27 cm. -- (Mosaicos matemáticos; 32)  
ISBN 978-607-7782-16-2  
Incluye bibliografía.  
1. Matemáticas - Investigaciones - Congresos, conferencias, etc. 2. Matemáticas - Enseñanza - Congresos, conferencias, etc. I. García Mireles, Gabriel Alberto, ed. II. Ibarra Olmos, Silvia Elena, ed.

Área de Análisis Bibliográfico, Subdirección de Servicios de Apoyo Académico,  
DSU, Universidad de Sonora

Memorias de la XIX Semana Regional de Investigación y Docencia en Matemáticas  
editada por Gabriel Alberto García Mireles y Silvia Elena Ibarra Olmos

Derechos reservados para esta edición:

D. R. © 2009, Gabriel Alberto García Mireles y Silvia Elena Ibarra Olmos (Editores)  
D. R. © 2009, Universidad de Sonora

ISBN: 978-607-7782-16-2

Mosaicos Matemáticos No. 32  
Línea editorial del comité editorial del  
Departamento de Matemáticas  
Universidad de Sonora  
Rosales y Blvd. Encinas  
Hermosillo, Sonora  
C. P. 83000  
Teléfono (01-662) 259-2155  
Fax (01-662) 259-2219

Diseño de portada: Jorge Origel y David Urrutia

Impreso en México  
Printed in México

**Departamento de Matemáticas**

# **Memorias**

**de la XIX Semana Regional  
de Investigación y Docencia  
en Matemáticas**

Gabriel Alberto García Mireles  
Silvia Elena Ibarra Olmos (Eds.)

**Mosaicos Matemáticos No. 32  
Noviembre de 2009**



EL SABER DE MIS HIJOS  
HARÁ MI GRANDEZA

# Directorio

Dr. Heriberto Grijalva Monteverde  
Rector

Dr. Enrique Velázquez Contreras  
Secretario General Académico

M. E. Rosa Elena Trujillo Llanes  
Secretaria General Administrativa

Dra. Arminda Guadalupe García de León Peñúñuri  
Vicerrectora de la Unidad Regional Centro

M. C. Miguel Ángel Moreno Núñez  
Director de la División de Ciencias Exactas y Naturales

Dr. Jesús Adolfo Minjárez Sosa  
Jefe del Departamento de Matemáticas

M. D. O. Manuel Ignacio Guerra Robles  
Director de Vinculación y Difusión

Lic. Raúl Acevedo Savín  
Jefe del Área de Publicaciones

# Consejo editorial

Dr. Oscar Vega Amaya  
Coordinador general

Dr. Guillermo Dávila Rascón  
M. C. Francisco Cándido García Durán  
Dra. Martha Dolores Guzmán Partida  
Dr. Jesús Adolfo Minjárez Sosa  
M. C. Eduardo Tellechea Armenta

# **Mosaicos Matemáticos No. 32**

## **Noviembre de 2009**

### **XIX Semana Regional de Investigación y Docencia en Matemáticas**

#### **C**omité organizador

M. C. Eduardo Tellechea Armenta  
Presidente

Silvia Elena Ibarra Olmos  
Secretaria

Agustín Brau Rojas  
José Dolores Dávila Galindo  
Guillermo Dávila Rascón  
Lorena Armida Durazo Grijalva  
Gabriel Alberto García Mireles  
Rodrigo González González  
Lina Morales Peral

#### Comité Académico

Agustín Brau Rojas  
Silvia Elena Ibarra Olmos  
Rodrigo González González

#### Comité Editorial

Gabriel Alberto García Mireles  
Silvia Elena Ibarra Olmos

# Índice

	Pág.
<b><u>Ponencias generales</u></b>	<b>1</b>
Optimización del proceso de aprendizaje en línea personalizado usando redes bayesianas Francisco Javier Tapia Moreno y Claudio Alfredo López Miranda	3
Competencias matemáticas en educación primaria Gersón Hernández Martínez, Patricia Gómez Avilés, Ángel Muñoz Velázquez e Ignacia Elizabeth García Ramírez	17
Diseño, construcción y control a distancia de una grúa autoequilibrada Rogelio Francisco Antonio y Diego I. Chavero Núñez	25
Nuevas perspectivas en la didáctica de estadística en el área de las ciencias sociales María Elena Parra Ramos	33
La osculación como un medio para arribar a las series de Taylor y McLaurin apoyado con el uso de la tecnología Mario Silvino Ávila Sandoval y Carlos López Ruvalcaba	39
El uso de cubiertas en la demostración de teoremas del análisis real Evaristo Trujillo Luque y Guadalupe Ávila Godoy	45
La técnica del punto conocido aplicada al ajuste del modelo normal Gabriel Guillén Buendía y Ana María Islas Cortés	51
Notas para una interpretación de Herón de Alejandría Francisco Cándido García Durán	61
Conjuntos, grupos, música y <i>Mathematica</i> Inna Shingareva, Reyna Berenice Ibarra Huerta y Carlos Lizárraga Celaya	69
Evaluación de las habilidades de comprensión de código en C++ Gabriel Alberto García Mireles e Irene Rodríguez Castillo	77
Resolución de ecuaciones lineales a través de material manipulable Patricia Gómez Avilés, Gersón Hernández Martínez, Ángel Muñoz Velázquez e Ignacia Elizabeth García Ramírez	85
El análisis funcional: una perspectiva histórica Oswaldo González Gaxiola	91
Resultados de la implementación de tareas y exámenes en línea para los cursos de álgebra utilizando el software Maple T. A. Guadalupe del Castillo Bojórquez y Blanca Evelia Flores Soto	97
Formulaciones equivalentes de la propiedad Radon-Nikodým de un espacio de Banach Martha Guzmán Partida y Marysol Navarro Burruel	105

Estimación de un parámetro de interés en presencia de parámetros de estorbo Gudelia Figueroa Preciado	113
Verosimilitud para inferencia estadística sobre los parámetros de un proceso de Poisson María Elena Conde	121

## Taller de Topología 131

Dinámica en hiperespacios Gerardo Acosta	133
Existencia de funciones de Whitney Genaro Hernández Mada y Carlos Alberto Robles Corbalá	145
Dimensión en niveles de Whitney de una gráfica finita Roberto Torres Hernández	153
Funciones entre hiperespacios de continuos y relación de orden Gloria Andablo Reyes	163
La conexidad del hiperespacio $2^X$ Alejandra Fonseca Morales y Carlos Alberto Robles Corbalá	171
La propiedad de compacidad en el hiperespacio Carmen Giraldi Higuera Chan y Carlos Alberto Robles Corbalá	177

# **M**emorias

## **De la XIX Semana Regional de Investigación y Docencia en Matemáticas**

### **Presentación**

La Semana Regional de Investigación y Docencia en Matemáticas es organizada por el Departamento de Matemáticas como un foro de difusión y discusión de ideas relativas a la enseñanza y aprendizaje de las matemáticas así como un evento que invita a la exposición de los problemas, avances y resultados de la investigación de las diversas áreas de las matemáticas y de la computación. En este año, además de ser el evento décimo noveno de esta naturaleza, también fue el foro para celebrar los 45 años de la licenciatura en Matemáticas.

Como parte de dicha celebración, se organizaron diversos eventos académicos, en los cuales se discutieron los avances de investigación en áreas como teoría de gráficas, biomatemáticas, control, sistemas dinámicos, topología y estadística. Agradecemos la asistencia y participación de estudiantes, profesores e investigadores que nos acompañaron de distintos centros de investigación del país y del extranjero.

En esta memoria se recogen los trabajos presentados en el evento general de la Semana Regional de Investigación y Docencia en Matemáticas y algunas de las ponencias del Taller de Topología. Expresamos nuestro agradecimiento a todos aquellos que participaron, tanto a los autores de los trabajos que se presentan, como a los docentes que nos apoyaron en la revisión y edición de esta memoria.

Los editores,

Gabriel Alberto García Mireles  
Silvia Elena Ibarra Olmos

Hermosillo, Sonora  
Noviembre de 2009.

# **Ponencias generales**

# Optimización del proceso de aprendizaje en línea personalizado usando redes bayesianas

Francisco Javier Tapia Moreno y Claudio Alfredo López Miranda  
Departamento de Matemáticas  
Universidad de Sonora  
{ftapia, claudio}@gauss.mat.uson.mx

## Resumen

*En este artículo, proponemos un modelo basado en redes bayesianas para optimizar el proceso de aprendizaje en línea personalizado. Este modelo determina el tipo de personalización requerido por el alumno de acuerdo a sus necesidades reales. Las relaciones de causa y efecto se codificaron en una red bayesiana, para la cual consideramos cuatro fases que afectan el proceso de aprendizaje del alumno: fase de los conocimientos previos, fase del progreso del conocimiento, fase de los objetivos y metas del profesor/alumno y la fase de las experiencias y preferencias de navegación. Mostramos la eficacia de las redes bayesianas en la primera fase y deducimos el tipo de personalización requerido por el alumno a través de un conjunto de perfiles y sus probabilidades asociadas.*

## 1. Introducción

El proceso de aprendizaje en línea personalizado (PALP), es un sistema de enseñanza-aprendizaje en línea en el cual las fases del conocimiento se adaptan a las necesidades del alumno. En este sistema, se estudian las características principales del alumno en cada etapa del proceso de enseñanza-aprendizaje, para optimizar el PALP de acuerdo a sus necesidades reales, en conjunción con los objetivos y metas del profesor y del alumno, y en relación al programa de estudios. El proceso de aprendizaje en línea supone el uso de las tecnologías multimedia e hipermedia para desarrollar y mejorar nuevas estrategias de aprendizaje. Este proceso utiliza herramientas de la tecnología de la información tales como: CD-ROMs, Internet, intranet o dispositivos móviles para hacer el conocimiento accesible al mayor número de personas. Así, el conocimiento se obtiene por medio de cursos en línea, correos electrónicos, aprendizaje por computadora, libros electrónicos, CD-ROMs, simulación virtual y otros tipos de software tales como: wikis, fórum y otros espacios colaborativos.

Los trabajos previos sobre aprendizaje en línea basados en modelos bayesianos (MBs) se han implementado para identificar sólo una característica del alumno. Ejemplos de tales trabajos son los software: OLAE (*On-line assessment of expertise*), un sistema de cómputo para evaluar el conocimiento del alumno de física y mecánica Newtoniana, POLA (*Probabilistic on-line assessment*), ANDES, HYDRIVE, SIETTE (Sistema inteligente de evaluación mediante test), CAPIT, and POET (*The on line preference elicitation tool*). Los MBs usados en los software mencionados, son usados exitosamente para construir y actualizar el modelo del alumno, pero éstos sólo realizan diagnósticos del nivel de conocimientos del alumno, mientras que en otros MBs, únicamente sirven para predecir un sólo objetivo de personalización, por ejemplo, el

estilo de aprendizaje [3]. Consecuentemente, tales MBs no toman en cuenta las preferencias, necesidades, metas, intereses y otra información acerca del alumno, las cuales son muy importantes para determinar el perfil deseable de una manera realista.

En este artículo, presentamos un MB mejorado para optimizar el PALP. Este modelo probabilístico se desarrolló considerando los objetivos y objetos de adaptividad en las cuatro fases fundamentales: adaptividad para los conocimientos previos del alumno, adaptividad para el progreso de los conocimientos, adaptividad para los objetivos y las metas del profesor/alumno, y adaptividad para las experiencias y preferencias de navegación. Para optimizar el PALP de una manera individualizada, es necesario recopilar todas las posibles métricas de aprendizaje (MA). Las MAs son toda clase de información acerca de los procesos y actividades de aprendizaje, y todas las maneras de registrar el desarrollo del aprendizaje [11,12]. Las MAs consideradas fueron: nivel de conocimientos (alto, intermedio y bajo), estilo cognitivo (dependiente e independiente), estilo de comunicación (pasivo, asertivo y agresivo), estilo de aprendizaje (activo, reflexivo, teórico y pragmático), entre otros. El modelo se evalúa mediante la simulación de un curso en línea con 45 alumnos de diversas áreas del conocimiento, tales como bellas artes, ciencias exactas y naturales, ingeniería, biología, etc. Además, incluimos una lista de objetos y objetivos de personalización, para determinar las potencialidades, cualidades y preferencias personales del alumno.

El resto del documento está distribuido como sigue: en la Sección 2 revisamos algunas técnicas de modelado del alumno. En la Sección 3 presentamos las redes bayesianas (RBs). En la Sección 4 describimos el diseño del experimento. En la Sección 5 mostramos las variables que componen nuestro MB y definimos las fases del conocimiento, estas fases son fundamentales para la personalización del PALP. La Sección 6, contiene los resultados principales de esta investigación, los cuales son muy útiles para inferir las distribuciones de probabilidades condicionales y conjuntas, así como el tipo de perfil del alumno en cada una de las fases y las probabilidades previas y posteriores de los nodos padres y de los nodos hijos. Por último, en la Sección 7 presentamos las conclusiones de nuestra investigación.

## **2. Técnicas de modelado del alumno**

Para construir el modelo del alumno necesitamos inferir cierta característica de éste, tales como sus habilidades, creencias, motivos, preferencias individuales, necesidades personales, estilos de aprendizaje, conocimientos previos, acciones futuras y así sucesivamente. Esa característica, implica incertidumbre cuando se usa en un STI. Existen técnicas para tratar los casos de incertidumbre: 1) métodos deterministas: suponen que toda la información requerida puede cuantificarse a priori y hacerla disponible en caso necesario [1]; 2) métodos de extensión algorítmicos y deterministas: asumen que algunos algoritmos pueden abarcar todos los planes y sus acciones correspondientes [8]; 3) aprendizaje de máquina: los sistemas del modelado tradicional del usuario tienen desventajas, algunas de éstas pueden ser cubiertas con las técnicas del aprendizaje de máquina [4]; 4) lógica difusa: estas técnicas son usadas para representar y concluir con conceptos vagos, para imitar el estilo de razonamiento humano; 5) métodos probabilistas. Ejemplos de estos métodos son: las redes bayesianas de creencia, factores de certeza, Dempster-Shafer y otros más. Tales métodos se basan en la premisa de que asignar un cierto valor a la hipótesis del plan, reflejando la probabilidad de lo que está siendo

demandado por el usuario [7]. De acuerdo con Heckerman [5], una red bayesiana (RB) ofrece un gran número de ventajas para el análisis de datos, algunos de ellos son: el modelo puede manejar situaciones donde algunos datos de entrada están perdidos y el modelo también nos permite deducir relaciones causales entre variables o nodos.

### 3. Redes bayesianas

Conforme a Friedman y Goldsmidt [2], una RB es un modelo gráfico para representar eficientemente una distribución de probabilidad conjunta en un conjunto de variables aleatorias  $V$ . Una RB es denotada por  $(G, P)$  donde  $G$  es una gráfica acíclica dirigida (GAD) definida en  $V$  (esta gráfica codifica relaciones de independencia entre las variables contenidas en  $V$ ); y  $P$  denota un conjunto de distribuciones de probabilidad locales, una por cada variable condicionada por sus padres. Las variables están representadas por los *nodos* que indican conceptos y las *flechas* que denotan dependencias de causa-efecto entre los conceptos. Los nodos finales se visualizan como *efectos* (valores recabados de los ambientes de aprendizaje), mientras que el nodo de mayor nivel se toma como *causas*. Cada nodo tiene dos o más resultados posibles; cada resultado es denominado un *estado* de la variable. Así, la probabilidad asociada a cierto perfil del alumno se obtiene desde una GAD. Una vez que el perfil del estudiante es determinado, puede usarse ocasionalmente para construir el modelo de aprendizaje personalizado de este alumno.

Sea  $V = \{x_1, x_2, \dots, x_n\}$  el dominio, tal que su RB asociada represente la distribución de probabilidad conjunta  $P(x)$  sobre el conjunto de variables aleatorias  $x_i$ . La probabilidad conjunta se calcula mediante la fórmula:

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i / \Pi_i) \quad (1)$$

en donde  $\Pi_i$  es un conjunto de padres relativos a cada una de las variables  $x_i$ , tal que  $\Pi_i \subseteq \{x_1, x_2, \dots, x_{n-1}\}$  es un subconjunto de variables en donde  $x_i$  es condicionalmente dependiente. Usando la regla de la cadena para variables aleatorias [9] podemos reescribir la distribución de probabilidad conjunta como sigue:

$$P(x_1, x_2, \dots, x_n / e) = \prod_{i=1}^n P(x_i / x_1, x_2, \dots, x_{n-1}, e) \quad (2)$$

donde  $e$  representa la evidencia con respecto a la variable  $x_i$ .

Ahora, para cada  $x_i$  habrá alguno subconjunto  $\Pi_i \subseteq V$  tal que  $x_i$  y  $V$  sean condicionalmente independientes dado  $\Pi_i$ . Esto es,

$$P(x_i / x_1, x_2, \dots, x_{n-1}, e) = P(x_i / \Pi_i, e) \quad (3)$$

La estructura de la RB codifica las afirmaciones de independencia condicional como una gráfica acíclica dirigida tal que cada nodo corresponde a una sola variable y los padres del

nodo correspondiente a  $x_i$  son los nodos asociados a las variables en  $\Pi_i$ . El par formado por la estructura (la gráfica) y la colección de las distribuciones locales  $P(x_i|\Pi_i)$  de cada nodo en el dominio, constituye la RB para este dominio.

**Tabla 1. Variables del MB y sus estados.**

Variable	Estados o resultados posibles y su notación
<b>Objetivos de Personalización</b>	
1. Conocimientos previos	1) Bajo, 2) Intermedio, y 3) Alto
2. Estilo de aprendizaje	1) Activo, 2) Reflexivo, 3) Teórico, y 4) Pragmático
3. Estilo cognitivo	1) Dependiente (DEP), y 2) Independiente (IND)
4. Estilo de comunicación	1) Pasivo (PAS), 2) Asertivo (ASE), 3) Agresivo (AGR)
5. Estilo de enseñanza preferido	1) Autoritario, 2) Demostrador, 3) Facilitador, y 4) Delegador
6. Técnicas de aprendizaje	1) Para aprendizaje visual, 2) Para aprendizaje activo, y 3) Para aprendizaje colaborativo (COL)
7. Preferencias individuales	1) Visuales, 2) Auditivas, y 3) Kinestéticas (KIN)
8. Currículo	1) Ciencias Exactas y Naturales, 2) Ingeniería, 3) Biología y Ciencias de la Salud, 4) Ciencias Sociales, 5) Económicas y Administrativas, y 6) Humanidades y Bellas Artes
<b>Objetos de Personalización</b>	
9. Necesidades personales	1) Ambientales (AMB), 2) Emocionales (EMO), 3) Sociales (SOC), y 4) Fisiológicas (FIS)
10. Selección de los objetos de aprendizaje	1) En CD ROM, 2) En línea, y 3) Combinado
11. Presentación de los objetos de aprendizaje	1) Por necesidad del programa de enseñanza, y 2) Como una facilidad para tener acceso a un objeto de aprendizaje particular sugerido
12. Selección de métodos de entrada	1) Ratón, 2) Teclado, 3) Pulsador, 4) Sistema de reconocimiento de lenguaje (SRL)
13. Dispositivos de aprendizaje	1) Objetos inteligentes, (OI) 2) Infraestructuras de la información (II), y 3) Ambientes artificiales compartidos (AAC)
14. Usabilidad del sistema de software por parte del alumno	1) Buena, 2) Regular, 3) Deficiente
<b>Fase de personalización</b>	
15. Fase de los conocimientos previos	1) Adaptar, 2) No adaptar
16. Fase del progreso del conocimiento	1) Adaptar, 2) No adaptar
17. Fase de los objetivos y metas del profesor/alumno	1) Adaptar, 2) No adaptar
18. Fase de las preferencias y experiencia de navegación	1) Adaptar, 2) No adaptar
<b>Demanda</b>	
<b>Demanda fase 1</b>	
19. Sistema	1) Adecuación automática (AA), 2) Adecuación manual (AM)
20. Alumno	1) Capacitar (CAP), 2) No capacitar (NCAP)
<b>Demanda fase 2</b>	
21. Sistema	1) Adecuación automática (AA), 2) Adecuación manual (AM)
22. Alumno	1) Capacitar (CAP), 2) No capacitar (NCAP)
<b>Demanda fase 3</b>	
23. Sistema	1) Adecuación automática (AA), 2) Adecuación manual (AM)
24. Alumno	1) Capacitar (CAP), 2) No capacitar (NCAP)
<b>Demanda fase 4</b>	
25. Sistema	1) Adecuación automática (AA), 2) Adecuación manual (AM)
26. Alumno	1) Capacitar (CAP), 2) No capacitar (NCAP)

#### 4. Diseño del experimento

En nuestro experimento se asigna aleatoriamente un estilo de aprendizaje a cada uno de los 45 alumnos del curso mencionado en la Sección I, se asigna también un estilo cognitivo, un estilo de comunicación, un estilo de enseñanza preferido, técnicas de aprendizaje preferidas, un nivel de conocimientos previos, preferencias individuales, un curriculum o experiencia en el área (ingeniería, biología, etc.), y necesidades personales. Cada característica representa un objetivo de personalización del alumno en el MB. Además, asignamos aleatoriamente a cada uno de los estudiantes las siguientes particularidades: selección de objetos de aprendizaje (CD-ROM, en

línea, cualquier combinación de estas dos formas), selección de métodos de entrada (ratón, teclado, pulsador, sistema de reconocimiento de lenguaje), dispositivos de aprendizaje preferidos y nivel de usabilidad del sistema por parte del alumno. Cada característica representa un objeto de personalización del alumno en el MB. Tanto los objetivos como los objetos de personalización, son considerados como eventos independientes entre sí. Cada objeto u objetivo de personalización representa una causa que tiene un efecto directo en alguna de las cuatro fases del conocimiento antes mencionadas. Cada fase, a su vez, es considerada como una causa que tiene un efecto directo en el aprendizaje del alumno y en la adecuación del sistema de enseñanza/aprendizaje. También, el alumno y el sistema de cómputo son tomados como eventos mutuamente independientes. Así, es posible determinar el perfil deseable en cada fase. En la siguiente sección, vemos cómo se construyó el MB para la optimización del PALP.

## 5. Modelación de las fases del conocimiento con RBs

El análisis de datos se realizó considerando las probabilidades más altas de los resultados obtenidos en cada una de las fases de personalización. Estas probabilidades representan la credibilidad del STI usado para el proceso de enseñanza/aprendizaje en línea, acerca de las características del alumno que determinan su tipo de personalización. El resultado final se obtiene multiplicando las probabilidades calculadas en el nodo del alumno y en el nodo del STI.

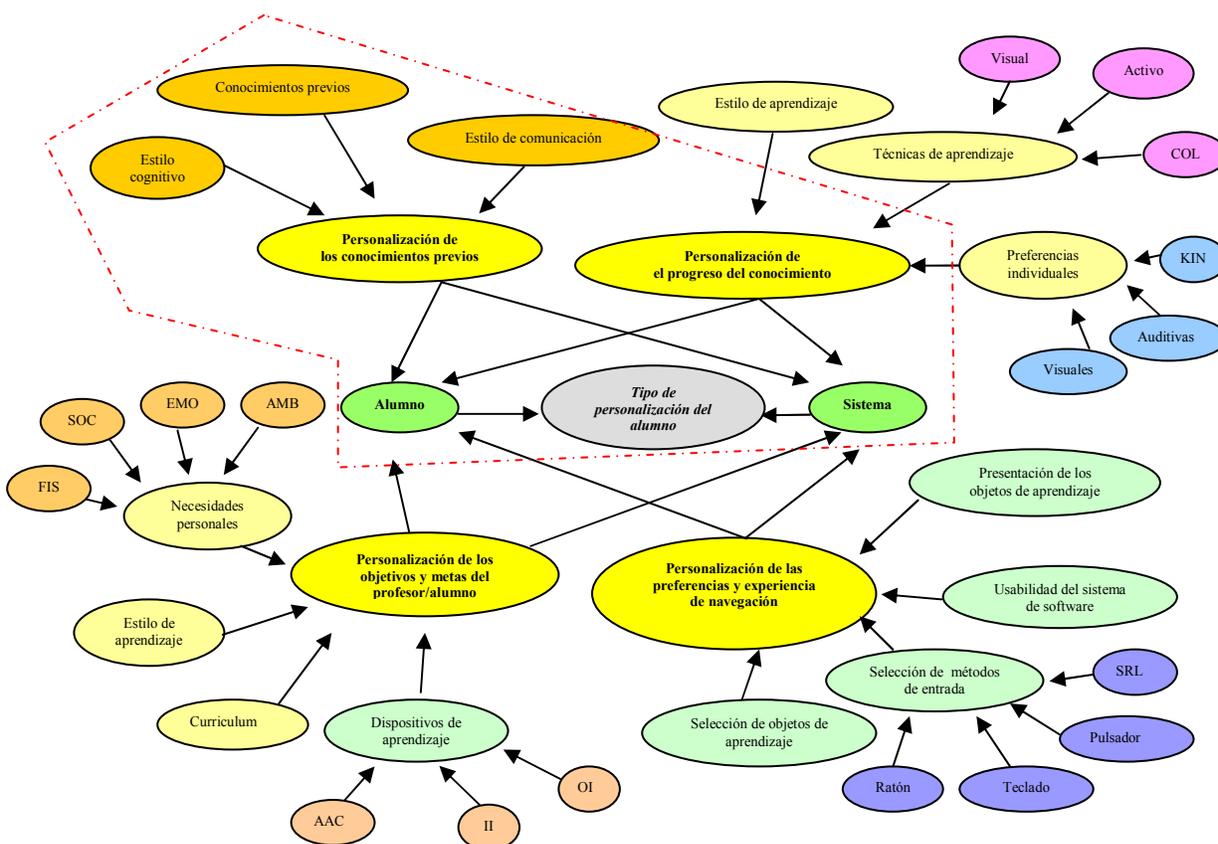


Figura 1. Estructura de la red bayesiana (ver nomenclatura en la Tabla 1).

Construir un MB para un dominio implica una diversidad de tareas [6,9]. La primera tarea consiste en identificar las variables significativas y sus posibles valores. En nuestro dominio aplicación, las variables representan objetivos y objetos de personalización, las fases de personalización, el alumno y el sistema (STI). La Tabla 1 muestra las variables y sus estados que se usan en este artículo.

La segunda tarea consiste en construir la parte cualitativa identificando las independencias entre las variables. Luego, debemos expresar esto en una GAD que codifique las afirmaciones de las independencias condicionales. Esta gráfica es denominada estructura de la RB y es mostrada en la Figura 1. En esta figura, el PALP está dividido en cuatro fases:

1. Fase de los conocimientos previos. En esta etapa, se detecta el nivel de conocimientos del alumno por medio de una evaluación individual, y mediante algún procedimiento, los objetos de aprendizaje son seleccionados y puestos al alcance del alumno dependiendo del nivel de conocimientos detectado. En nuestro MB esta fase es considerada como una causa de los siguientes objetivos de personalización: conocimientos previos del alumno, estilo cognitivo del alumno y estilo de comunicación. Teniendo estos objetivos en mente, es posible capacitar al alumno (si es necesario) para usar el STI óptimamente y obtener información particular del alumno. Así, el STI se podrá adaptar a las necesidades del alumno y estará listo para usarse en las fases siguientes.

2. Fase del progreso del conocimiento. En este paso, el proceso de aprendizaje del alumno se controla mediante trayectorias o itinerarios personales de aprendizaje, de acuerdo a algunas características específicas del alumno. Esta fase se considera como causa de los objetivos de personalización: estilo de aprendizaje, técnicas de aprendizaje, y el objeto de personalización de las preferencias individuales. Así, es posible capacitar al alumno y adecuar al STI de modo que el alumno obtenga el conocimiento deseado durante la etapa de aprendizaje, de acuerdo a los objetivos y objetos de personalización identificados en esta fase y en la primera.

3. Fase de los objetivos y metas del profesor/alumno. En esta etapa, se guía al alumno mediante trayectorias especiales de aprendizaje acordes a los objetivos y metas tanto del profesor como del alumno. Esta fase es considerada como una causa de los objetivos de personalización: curriculum o experiencia en el área, necesidades personales, estilo de enseñanza preferido y el objeto de personalización: dispositivos de aprendizaje preferidos. Con estos objetivos y objetos de personalización, es posible preparar al estudiante y al STI de acuerdo a los objetivos y metas tanto del profesor como del alumno, y seleccionar los contenidos y su presentación.

4. Fase de la experiencia y preferencias de navegación. En esta paso, se ofrecen al alumno varios apoyos a la navegación. Aquí, el estudiante puede tener la libertad total de navegación, o puede ser guiado a objetivos y metas específicos mediante itinerarios de aprendizaje previamente establecidos. En el MB, esta fase se considera como una causa de los objetos de personalización: Usabilidad del sistema por el alumno, elección de los métodos de entrada, selección de los objetos de aprendizaje y presentación de los objetos de aprendizaje. Conociendo estos objetos de personalización es posible preparar al alumno para la navegación y al STI de acuerdo a las preferencias y experiencia del alumno. El resultado

obtenido en cada una de las fases, es usado para determinar el modelo personalizado de aprendizaje del alumno.

## 6. Resultados del experimento

Ahora discutiremos las estadísticas obtenidas en nuestro experimento (curso en línea simulado), el cual describimos en la Sección 4. Las probabilidades mostradas en esta sección, son las frecuencias relativas obtenidas en la simulación.

### 6.1. Distribuciones de probabilidad condicional y probabilidad conjunta

Las independencias condicionales entre los objetos y los objetivos de personalización definen la estructura de la RB de la Figura 1. Esta estructura se usa para obtener probabilidades conjuntas de perfiles del alumno tales como

$$P(\text{Alto, IND, ASE, Adaptar, CAP, AA}) \quad (4)$$

en la Fase 1.

$$P(\text{Activo, COL, Visual, Adaptar, CAP, AA}) \quad (5)$$

en la Fase 2, y así sucesivamente.

En cada fase, los perfiles se obtienen mediante el producto de probabilidades tales como:

$$P(\text{Alto, IND, ASE, Adaptar, CAP, AA}) \times P(\text{Alto, IND, ASE, Adaptar, CAP, AA}) \quad (6)$$

donde el primer factor corresponde a la probabilidad del nodo del alumno y el segundo factor a la probabilidad del nodo del STI. Se selecciona el perfil con mayor probabilidad y se establece como el tipo de personalización del alumno (TPA) en la fase correspondiente. Este perfil representa la credibilidad del STI en relación a las características del estudiante. Para construir la RB completa, debemos calcular el valor de un conjunto probabilidades condicionales correspondientes a las distribuciones locales  $P(x_i|\Pi_i)$ . El modelo queda completo estableciendo el valor de las probabilidades asociadas a cada nodo de la gráfica. Esto es, en cada fase estableceremos una función de distribución de probabilidad (*FDP*),  $\Pi_i$ , para cada estado de los nodos. Las funciones de distribución de probabilidad asociadas con los nodos independientes tienen una distribución multinomial [9].

### 6.2. Perfiles del alumno en la primera fase

La sección punteada de la Figura 1 engloba la estructura de la RB de la personalización de los conocimientos previos en la Fase 1. Las Tablas 2, 3 y 4 contienen las estadísticas de esta fase, las cuales representan las distribuciones de probabilidad de los nodos padres: nivel de conocimientos previos, estilo cognitivo y estilo de comunicación, respectivamente.

**Tabla 2. FDP del nivel de conocimientos previos.**

1. Nivel de conocimientos previos	Prob.
1. Alto	0.2667
2. Intermedio	0.4222
3. Bajo	0.3111

**Tabla 3. FDP del estilo de comunicación.**

4. Estilo de comunicación	Prob.
1. Pasivo	0.4444
2. Asertivo	0.2223
3. Agresivo	0.3333

**Tabla 4. FDP del estilo de cognitivo.**

3. Estilo cognitivo	Prob.
1. Dependiente	0.5556
2. Independiente	0.4444

En la Tabla 2 note que el 42.22% de las veces, el nivel de conocimientos previos del alumno ha sido intermedio. Estos valores se actualizan al mismo tiempo que el administrador del sistema de cómputo, recopile información acerca del nivel de conocimiento de sus usuarios. Las probabilidades de las Tablas 3 y 4 indican algo similar.

Procediendo de una manera análoga, las Tablas 5, 6 y 7 muestran las probabilidades condicionales de los nodos hijos *Personalización de los conocimientos previos*, *Alumno y Sistema*, respectivamente. El cuarto renglón y la cuarta columna de la Tabla 6, indica la siguiente probabilidad condicional

$$P(\text{Adaptar/Bajo, Agresivo, Dependiente}) \quad (7)$$

**Tabla 5. Probabilidades condicionales para el nodo de la Fase 1.**

Nodos padres			Adecuación del sistema en la fase de los conocimientos previos	
Nivel de conocimientos previos	Estilo de comunicación	Estilo cognitivo	Adaptar	No adaptar
Bajo	Pasivo	Dependiente	0.5	0.5
		Independiente	1	0
	Asertivo	Dependiente	0	1
		Independiente	1	0
	Agresivo	Dependiente	0.6667	0.3333
		Independiente	0.3333	0.6667
Intermedio	Pasivo	Dependiente	0.25	0.75
		Independiente	0.5	0.5
	Asertivo	Dependiente	0.6	0.4
		Independiente	0	1
	Agresivo	Dependiente	1	0
		Independiente	0.6667	0.3333
Alto	Pasivo	Dependiente	0.5	0.5
		Independiente	1	0
	Asertivo	Dependiente	0.6667	0.3333
		Independiente	1	0
	Agresivo	Dependiente	0	1
		Independiente	0.6667	0.3333

Las probabilidades de las Tablas 6 y 7 indican algo similar.

**Tabla 6. Probabilidades condicionales para el nodo *alumno* de la Fase 1.**

Nodo	Alumno	
	Capacitar	No capacitar
Personalización de los conocimientos previos		
Adaptar	0.5769	0.4230
No adaptar	0.6316	0.3684

**Tabla 7. Probabilidades condicionales para el nodo *sistema* de la Fase 1.**

Nodo	Sistema	
	Adecuación automática	Adecuación manual
Personalización de los conocimientos previos		
Adaptar	0.6484	0.3516
No adaptar	0.5544	0.4456

Por cuestiones de espacio se omitieron los resultados de las Fases 2, 3 y 4, mismas que se pueden consultar en [12].

### 6.3. Probabilidades previas y posteriores de los perfiles

El propósito de esta sección es calcular las probabilidades de todos los perfiles posibles que se generan en cada fase de personalización. Por ejemplo, en la fase de los conocimientos previos obtenemos probabilidades previas tales como:

$$P(\text{Alto, Independiente, Asertivo, Adaptar, capacitar, Adecuación. Automática}) \quad (8)$$

Las probabilidades se calculan en dos partes. La primera parte proporciona la probabilidad del nodo del alumno, y la segunda la probabilidad del nodo del sistema. Por consecuencia, la probabilidad total se calcula mediante la multiplicación directa por ser nodos independientes.

#### 6.3.1. Probabilidades previas

Las probabilidades previas tales como (8), se calculan mediante el teorema de probabilidad total. Para información más detallada consultar [12]

#### 6.3.2. Probabilidades posteriores

Cuando se termina una actividad o módulo en el curso en línea, podemos hacer inferencias sobre la actividad o módulo siguiente, usando las probabilidades previas y calculando las probabilidades posteriores. Estas probabilidades pueden usarse para deducir las características y necesidades del alumno, adecuaciones del sistema y otros requerimientos del curso en línea. También, estas probabilidades pueden usarse para inferir perfiles parciales del alumno tales como  $P(\text{alto, pasivo, dependiente})$  o  $P(\text{capacitar/alto, pasivo, dependiente})$ .

Para calcular las probabilidades posteriores de todos los nodos de la Fase 1 (ver Figura 1 sección punteada en rojo), usamos las probabilidades previas mostradas en las Tablas II-V y el software MSBNX [10]. Los resultados se muestran en la Figura 2. En esta Figura, el primer valor de la columna State 0 de la Figura 2 (0.6009), representa la probabilidad de que un alumno dado necesite capacitarse. Por lo tanto, el STI tomará la decisión el 60.09 por ciento de las veces de sugerir alguna forma de capacitación al alumno. Similarmente, el tercer valor de la

columna State 0 de la Figura 2 (55.56), significa que existe un 55.56% de probabilidad de que un alumno particular tenga un estilo cognitivo dependiente. Así, el STI tomará esta probabilidad para asignar un estilo cognitivo al alumno con estilo cognitivo desconocido y así sucesivamente.

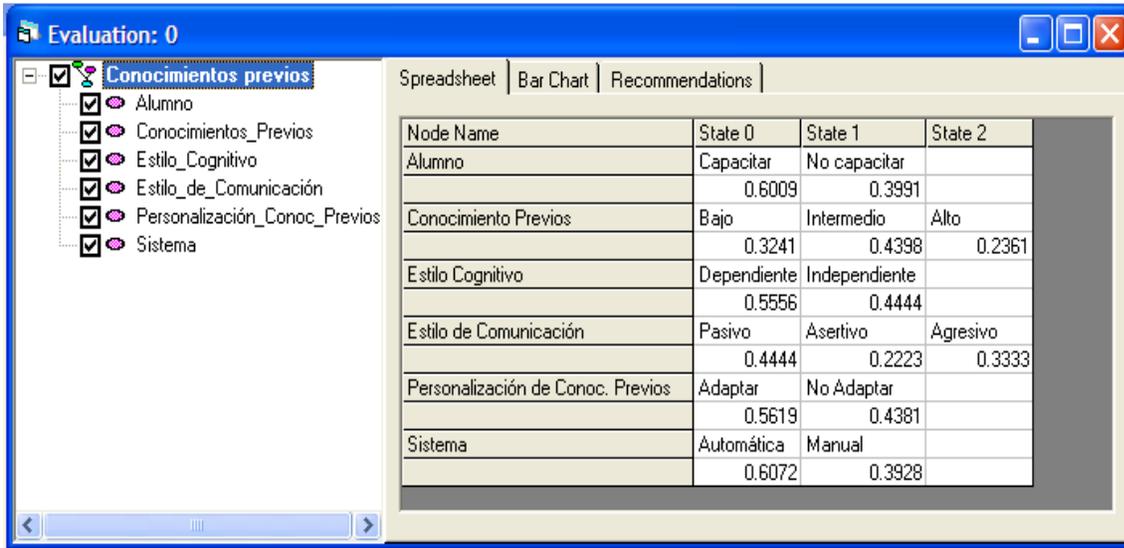


Figura 2. Probabilidades posteriores de la Fase 1.

Por otro lado, usando los valores mostrados en la Figura 2, podemos aplicar el Teorema de Bayes para calcular las probabilidades posteriores y hacer inferencias acerca del perfil parcial de personalización del alumno considerando su tipo de personalización. Este cálculo puede hacerse como sigue:

$$P(\text{Adaptar/Alto, Pasivo, Dependiente}) = \frac{P(\text{Alto, Pasivo, Dependiente/Adaptar})}{P(\text{Alto, Pasivo, Dependiente})}$$

$$0.5 = \frac{P(\text{Alto, Pasivo, Dependiente/Adaptar}) \times (0.5619)}{(0.2361) \times (0.4444) \times (0.5556)}$$

Así,

$$P(\text{Alto, Pasivo, Dependiente/Adaptar}) = 0.052$$

Esto indica que, que cuando sabemos que en alguna actividad o módulo del curso en línea se necesitó adaptar el sistema, hay un 5.2% de posibilidad de que un alumno tendrá un perfil parcial (Alto, Pasivo, Dependiente).

De una manera similar, podemos hacer deducciones acerca de un perfil de personalización parcial en un curso en línea calculando las probabilidades previas, usando el valor de la Tabla 5 (primera columna y treceavo renglón) y las probabilidades de las Tablas 2-4.

Tabla 8. Probabilidades condicionales del nodo *sistema* de la fase i.

Perfil	Conocimientos previos	Estilo cognitivo	Estilo de comunicación	Personalización de los conoc. previos	Estado del alumno	Adecuación del sistema	Prob.	
1	Bajo	Dependiente	Pasivo	Adaptar	Capacitar	Manual	0,0164	
2						Automático	0,0106	
3					No Capac.	Manual	0,0109	
4						Automático	0,0070	
5				No adaptar	Capacitar	Manual	0,0128	
6						Automático	0,0083	
7					No Capac.	Manual	0,0085	
8						Automático	0,0055	
.	<b>Truncado para Asertivo</b>							
.	<b>Truncado para Agresivo</b>							
.	<b>Truncado para Independiente</b>							
49	Intermedio	Dependiente	Pasivo	Adaptar	Capacitar	Manual	0,0223	
50						Automático	0,0144	
51					No Capac.	Manual	0,0148	
52						Automático	0,0096	
53				No adaptar	Capacitar	Manual	0,0174	
54						Automático	0,0112	
55					No Capac.	Manual	0,0115	
56						Automático	0,0075	
.		<b>Truncado para Asertivo</b>						
65		Independiente	Pasivo	Adaptar	Capacitar	Manual	0,0223	
66						Automático	0,0144	
67					No Capac.	Manual	0,0148	
68						Automático	0,0096	
69				No adaptar	Capacitar	Manual	0,0174	
70						Automático	0,0112	
71					No Capac.	Manual	0,0115	
72	Automático					0,0075		
73	Agresivo	Adaptar	Capacitar	Manual	0,0223			
74				Automático	0,0144			
75			No Capac.	Manual	0,0148			
76				Automático	0,0096			
77		No adaptar	Capacitar	Manual	0,0174			
78				Automático	0,0112			
79			No Capac.	Manual	0,0115			
80				Automático	0,0075			
81	<b>Truncado para Asertivo</b>							
.	<b>Truncado para Agresivo</b>							
.	<b>Truncado para Dependiente</b>							
.	<b>Truncado para Pasivo</b>							
.	<b>Truncado para Asertivo</b>							
137	Alto	Independiente	Agresivo	Adaptar	Capacitar	Manual	0,0120	
138						Automático	0,0077	
139					No Capac.	Manual	0,0079	
140						Automático	0,0051	
141				No adaptar	Capacitar	Manual	0,0093	
142						Automático	0,0060	
143					No Capac.	Manual	0,0062	
144						Automático	0,0040	

### 6.3.3. Perfiles de personalización del alumno

Usando los resultados de la Figura 2, obtenemos la Tabla 8 mediante multiplicaciones directas dado que los eventos son independientes. Esta tabla contiene todos los perfiles de personalización posibles en la fase 1 y sus probabilidades.

De acuerdo a nuestro MB, el perfil con la probabilidad más alta será la credibilidad del sistema experto (STI) acerca del TPA. En este caso, hay tres perfiles posibles, ellos son los perfiles 49, 65 y 73 de la Tabla 8. El SIT escogerá aleatoriamente uno de estos perfiles. En estos tres perfiles, hay tres cosas en común: 1) el sistema requiere adaptación, 2) el alumno necesita capacitación, 3) la adecuación del sistema debe ser manual.

## 7. Conclusiones

Hemos diseñado un modelo matemático útil para deducir el TPA, usando los objetivos y objetos de personalización. Nuestro modelo puede optimizar el proceso global del aprendizaje en línea del alumno siempre y cuando los contenidos, el apoyo, la infraestructura y la orientación adecuada sean dados al alumno. Por lo tanto, es necesaria una labor multidisciplinaria entre profesionales de la educación, psicología y ciencias de la computación, todas ellos apoyados en la ingeniería del conocimiento cuya aplicación puede responder, en general, a los requerimientos y problemas específicos de los alumnos y/o profesores. Las características principales del modelo propuesto aquí son: las relaciones causales entre los objetivos y objetos de aprendizaje, las fases del conocimiento, alumno y sistema electrónico usado para gestionar el proceso de enseñanza-aprendizaje del alumno. Obtuvimos un conjunto de perfiles de personalización considerando las principales características del alumno. Estos perfiles pueden usarse para proponer un modelo de enseñanza-aprendizaje al alumno, con el cual pueda optimizar su PALP de acuerdo a sus necesidades reales. Dicho MB puede servir, total o parcialmente, durante el proceso de enseñanza-aprendizaje, para realizar diagnósticos acerca de TPA, en caso de incertidumbre o de datos perdidos relativos a las características individuales del alumno. Es importante mencionar que el modelo por sí mismo no garantiza el aprendizaje del alumno, ya que el conocimiento de éste depende (en gran medida), de la actitud, el esfuerzo, el desempeño e interés por obtener el conocimiento.

## 8. Referencias

- [1] Ahmad, A. R., Basir, O. A., Hassanein, K., "Decision Preferences, Constraints, and Evaluation Objectives in Layout Design: A Review of Modeling Techniques", Proc. of the 5th Int'l Conf. on Op. and Quan. Manag (ICOQM-V), Oct. 2004, Seoul, Korea, to appear.
- [2] Friedman, N. and Goldszmidt, M.: "Learning Bayesian networks from data". Aaai-98 tutorial. 1998. <http://ai.stanford.edu/~nir/tutorial/index.html> (Última visita en noviembre, 2006).
- [3] García P., Amandi A. Schiaffino S. Campo M. "Using Bayesian Networks to Detect Student's Learning Styles in Web-based education system" 7o. Simposio Argentino de Inteligencia Artificial -ISAI2005. Rosario, 29-30 de Agosto de 2005.

- [4] Goecks, J., and Shavlik, J., "Learning Users Interests by Unobtrusively Observing Their Normal Behavior", IUI'00, New Orleans, LA. 2000.
- [5] Heckerman, D. "A Tutorial on Learning Bayesian Networks". Technical Report MSR-TR-95 06, Microsoft Research, 1995.
- [6] Henrion, M., "Some Practical Issues in Constructing Belief Networks". In: Kanl, L. N., Levitt, T. S., Lemmer, J. F. (eds), *Uncertainty in Artificial Intelligence 3*. Elsevier Science Publishers, 1989. <http://wundt.unigraz.at/publicdocs/publications/file1084438217.doc> (Visto Noviembre 2006).
- [7] Jameson, A. "Modelling both the context & user", *Personal Technologies*, 5(1), 1-4. 2001.
- [8] Kautz, H. A., "A formal theory of plan recognition", PhD thesis, U. of Rochester, NY. 1987.
- [9] Neapolitan R.E. "Learning Bayesian Networks" Prentice Hall, 2004.
- [10] Microsoft Research. "Bayesian Network Editor and Tool Kit." <http://research.microsoft.com/adapt/msbnx/> . 2001 (Última visita Febrero 2009).
- [11] Tapia, F. Galán M., Ocón A., Rubio E., "Using Bayesian Networks In The Global Adaptive e-Learning Process". EUNIS, 2005. [http://www.mc.manchester.ac.uk/eunis2005/medialibrary/papers/paper\\_130.pdf](http://www.mc.manchester.ac.uk/eunis2005/medialibrary/papers/paper_130.pdf). Última visita, Junio 2007.
- [12] Tapia, F. Lopez Claudio., Galan Manuel., Rubio E., "Bayesian Model for Optimization Adaptive e-Learning Process". *International Journal Emerging Technologies in Learning (I-JET)*. Vol. 3, No. 2. 2008. <http://online-journals.org/i-jet/> Última visita, Enero 2009.



# Competencias matemáticas en educación primaria

Gersón Hernández Martínez, Patricia Gómez Avilés  
Ángel Muñoz Velázquez e Ignacia Elizabeth García Ramírez  
Dirección de Investigación Educativa y Fortalecimiento Institucional  
Secretaría de Educación Pública Hidalgo  
gerson\_hm@hotmail.com

## Resumen

*La presente ponencia es un reporte de investigación realizada en el Estado de Hidalgo, con la finalidad de dar a conocer y analizar cómo se promueve el desarrollo de competencias matemáticas en educación primaria e identificar las problemáticas que se presentan durante el proceso enseñanza aprendizaje; con el propósito de implementar una propuesta de innovación a través de actividades con apoyo de materiales manipulables que permitan el desarrollo de competencias matemáticas (planteamiento y resolución de problemas, comunicación, argumentación y utilización de técnicas).*

## 1. Introducción

La principal intención y la gran apuesta de los cambios de enfoque propuestos en los Programas de Estudio del Acuerdo Nacional para la Modernización Educativa de 1992, fueron la reorientación de la práctica educativa, de manera que el desarrollo de capacidades y competencias cobrase primacía por sobre la visión predominantemente memorística e informativa del aprendizaje que la educación básica conserva desde su origen.

Esta reforma curricular en educación primaria en específico, enfatiza los procesos de aprendizaje con una nueva concepción de las competencias de lectura, escritura y matemáticas, de llevar a las aulas actividades de estudio que despierten el interés de los alumnos y los inviten a reflexionar. Sin embargo, en la escuela se generan y reproducen prácticas que provocan esta persistencia en la enseñanza y la evaluación que impiden mejorar los resultados educativos, pese a los esfuerzos de la política educativa (producción y distribución de materiales, actualización de profesores y programas de estímulos al desempeño profesional).

Considerando lo anterior es necesario retomar el trabajo y la experiencia de quienes laboran en las instituciones educativas. La educación primaria debe proporcionar a todos sus educandos herramientas suficientes que les permitan avanzar a los grados superiores, pero principalmente brindar una educación de calidad. En este sentido la Dirección de Investigación Educativa y Fortalecimiento Institucional del Estado de Hidalgo (DIEFI), se ha preocupado por darle prioridad al establecimiento de acciones que permitan contribuir al mejoramiento de una educación de calidad dentro del estado, mismas que están orientadas a promover de manera prioritaria la investigación educativa, cuyo rigor metodológico garantice incidir favorablemente en las principales problemáticas educativas en los distintos niveles y servicios que integran la Secretaría de Educación Pública de Hidalgo; por lo que se dio a la tarea de realizar la investigación “Competencias matemáticas en educación primaria”.

## 2. Antecedentes

El hecho de que se enseñen matemáticas en la escuela primaria responde a dos necesidades, la individual y la social; cada uno de nosotros debe saber un poco de matemáticas para poder resolver o por lo menos reconocer los problemas a los que nos enfrentamos mientras se convive con los demás. La presencia de las matemáticas en la escuela es una consecuencia de su presencia en la sociedad, por lo que cuando ésta se basa sólo en la enseñanza formal se reduce su valor social. Con relación a ello se han realizado recientemente evaluaciones en nuestro país que dan cuenta del logro que han tenido nuestros alumnos en esta materia, así mismo el identificar cuáles son los retos que tiene México en materia educativa, entre éstos destacan: los resultados de las pruebas nacionales aplicadas en 2003, indican que la mitad de los estudiantes de educación básica muestran un desempeño desfavorable en el área de razonamiento matemático, de acuerdo con el Instituto Nacional para la Evaluación de la Educación. La media nacional de respuestas correctas en los exámenes de ingreso a la educación media superior aplicados por el Centro Nacional de Evaluación para la Educación Superior en el 2003, fue de 46.7%, lo que significa que los casi 536 mil egresados de secundaria pudieron contestar correctamente en promedio sólo 60 de las 128 preguntas de una prueba diseñada para que la mayoría de los participantes acierte por lo menos 64. Se observó que el rendimiento de los jóvenes a quienes se les aplicó la prueba fue más desfavorable en las áreas de matemáticas (43.7%) y física (43.3%).

Durante el ciclo escolar 2004-2005, en el Estado de Hidalgo el Centro de investigación educativa y fortalecimiento institucional, a través del equipo técnico de la propuesta teórico-práctica de “Ludoteca interactiva de matemáticas”, aplicó una evaluación diagnóstica con contenidos matemáticos a grupos de primer grado en diferentes escuelas del nivel de secundaria (4 generales, 6 técnicas y 5 telesecundarias) en las regiones de Apan, Actopan, Huichapan, Zacualtipán, Tula, Tepehua, Pachuca, Tulancingo y Huejutla, con el propósito de conocer el grado de conocimientos con que cuentan los alumnos al ingresar a este nivel. Los resultados obtenidos en esta evaluación reflejan un bajo porcentaje de aprovechamiento general (35.10%). Las principales deficiencias detectadas fueron: el análisis y resolución de problemas que incluían dos o más operaciones, el uso de fórmulas para cálculo de perímetros, áreas y volúmenes, imaginación espacial, razonamiento matemático y operaciones básicas con números fraccionarios y decimales; así mismo, también se pudo identificar que prevalecen las prácticas de enseñanza rutinarias.

Considerando estos datos estadísticos se realiza un estudio retrospectivo en 1993, en donde la Secretaría de Educación Pública inicia un proceso de cambio al incorporar acciones como apoyo a la labor profesional que se lleva en las escuelas primarias, con la intención de facilitar el trabajo dentro de las aulas, así como proporcionar material de estudio para los maestros que deseen profundizar en el enfoque didáctico para la enseñanza, el estudio y el aprendizaje de las matemáticas, proporcionando recomendaciones generales para cada uno de los ejes temáticos que se trabajan en la educación primaria. Se sabe que durante el proceso enseñanza-aprendizaje se encuentran infinidad de factores que de alguna manera influyen en la baja calidad educativa, pero no basta con darse cuenta que existen problemas, sino se necesita tomar la decisión de hacer algo al respecto, por lo que es necesario propiciar una transformación decidida por parte de los actores que intervienen en el sector educativo, con el propósito de realizar transformaciones en lo que se hace, pero principalmente en el cómo se hace.

De esta reflexión nace el interés por realizar un proyecto de investigación que otorgue al docente una serie de información, que le permita tomar sus propias decisiones que sean de utilidad a los alumnos en su vida cotidiana, pero sobre todo que esos conocimientos sean duraderos; es importante reforzar las diferentes estrategias y recursos didácticos que permiten a docentes y alumnos favorecer el dominio conceptual de esta asignatura, teniendo siempre presentes los propósitos centrales de plan y programas oficiales y las sugerencias de los materiales de apoyo editados por la Secretaría de Educación Pública.

### **3. Justificación**

El Programa Nacional de Educación 2001-2006, en el capítulo correspondiente a educación básica hace referencia a promover el desarrollo de la investigación educativa y aprovechar sus resultados como insumo para el diseño, la evaluación y la reorientación oportuna de las políticas, programas y materiales educativos, así como para la transformación de las prácticas educativas en el aula y en la escuela; asimismo, fomentar la innovación educativa para el mejoramiento de la calidad y la equidad de los servicios educativos, así como las necesidades y problemas locales derivados de la diversidad regional y social del país (p.148). En el Plan Nacional de Desarrollo 2007–2012, se persigue la misma intención considerando en el Eje 3: igualdad de oportunidades, los objetivos de elevar la calidad educativa, reducir las desigualdades en las oportunidades educativas, impulsar el desarrollo y utilización de nuevas tecnologías en el sistema educativo para apoyar la inserción de los estudiantes en la sociedad del conocimiento y ampliar sus capacidades para la vida, así como el de promover la educación integral de las personas en todo el sistema educativo, planteando en cada uno de estos objetivos.

En este sentido la Dirección de Investigación Educativa y Fortalecimiento Institucional, se ha preocupado por darle prioridad al establecimiento de acciones que permitan contribuir al mejoramiento de una educación de calidad dentro del estado; para ello se ha considerado la importancia de construir diagnósticos y propuestas educativas que permitan abordar temáticas que a partir de las reformas se consideran relevantes, teniendo como eje central las áreas que han presentado problemáticas dentro de la educación como es el caso de la enseñanza y el aprendizaje de las matemáticas y el español. Para contribuir en la mejora de la calidad educativa se analizaron los resultados de distintas evaluaciones nacionales aplicadas a alumnos (ENLACE, EXCALE, PISA 2003 y 2006) para detectar e intervenir en los campos problemáticos. En cuanto a los proyectos de investigación que realiza la DIEFI se han considerado los niveles de educación secundaria y primaria; con respecto a este último durante el ciclo escolar 2006-2007 se estructuró el proyecto “Competencias matemáticas en educación primaria”.

La esencia del proyecto está planteada con base en el constructivismo, considerando los aspectos que destaca César Coll, donde se concibe al alumno como responsable y constructor de su propio aprendizaje, y al profesor, como un coordinador y guía del aprendizaje de los niños. Otro de los planteamientos que realiza es el referente a que los contenidos escolares no deben ser arbitrarios, considerando la concepción activa de los estudiantes, sin dejar de lado la acción pedagógica especialmente del profesor que es considerada en un término asociado con el de construcción, el del andamiaje o ajuste en la ayuda pedagógica, el cual va modificándose a lo largo del proceso de enseñanza-aprendizaje.

Para el diseño del proyecto de “Competencias matemáticas en educación primaria” se recuperaron elementos que refieren a las matemáticas como herramientas funcionales y flexibles que permiten resolver situaciones problemáticas, y a través de la acción alternada entre proponer y resolver problemas, se retoma la idea de Brousseau en relación a la utilización de elementos de la teoría de juegos y de la teoría de la información en la enseñanza de las matemáticas, además de considerar que “las matemáticas tienen un valor formativo que ayuda a estructurar todo el pensamiento y a agilizar el razonamiento deductivo”, las cuales constituyen las nociones centrales de la concepción de las matemáticas, así como de su enseñanza, además de que se deben considerar “diferentes perspectivas de la asignatura que significa conocer y enseñar matemáticas”.

#### **4. Metodología**

La metodología considerada para llevar a cabo este diagnóstico se plantea desde la perspectiva de investigación cualitativa y cuantitativa, centrando su interés principal en la búsqueda de información relacionada con los docentes y alumnos, lo que permitirá dar cuenta de lo que se hace, piensa y sabe con relación a las matemáticas en las escuelas de educación primaria en tres zonas escolares. Para realizar la investigación se seleccionó primeramente la muestra y se diseñaron instrumentos de recolección de información (guía de observación, guías de entrevista para docentes y alumnos, encuestas y cuestionarios para docentes y una encuesta para alumnos de tercero a sexto grado).

Las escuelas objeto de estudio para construir el diagnóstico, fueron seleccionadas bajo una muestra intencional, con el propósito de tener las facilidades necesarias y poder llevar a cabo la investigación, cubriendo las siguientes características: que pertenecieran a distintos contextos socioculturales, ubicadas en localidades de zonas urbanas y rurales, que fueran de organización completa y por último, contar con el interés y la aceptación voluntaria de directivos y docentes para participar en el proyecto de investigación.

Con la finalidad de observar y detectar de forma específica la situación que guarda la enseñanza y el aprendizaje de las matemáticas al interior de las aulas, el trabajo de campo se realizó en los seis grados de las diferentes escuelas, dando prioridad a los alumnos y docentes del segundo y tercer ciclo de educación primaria. Para dar inicio a este estudio se realizaron diferentes acciones que permitieron llevarlo a cabo, la primera de ellas fue la negociación del acceso, es decir el acercamiento con los supervisores y directivos de las tres zonas escolares, con el propósito de darles a conocer a grandes rasgos el proyecto de investigación; una vez aceptada la realización de la investigación, se procedió a oficializar ante las autoridades educativas estatales (director general de educación básica, director de educación primaria, supervisores escolares y directores de escuelas primarias). Posteriormente se realizaron reuniones colegiadas y personalizadas con los docentes involucrados, a quienes se les planteó de forma específica el trabajo a realizar, así como el objetivo central del trabajo de investigación. Por último, se les dio a conocer la estrategia de operatividad en donde se les proporcionó un cronograma de actividades que les permitiría identificar tiempos y fechas de visitas. Durante el trabajo de campo, se realizaron entrevistas a sujetos clave, éstos se seleccionaron considerando su antigüedad en la institución, así como sus referencias profesionales dentro de cada zona escolar.

Se realizaron visitas de exploración y reconocimiento a las 24 escuelas propuestas por los supervisores y directivos de las tres zonas escolares, donde se llevaron a cabo observaciones durante la clase de matemáticas, entrevistas a los docentes y alumnos de los seis grados de educación primaria, además de encuestas a alumnos de tercero, cuarto, quinto y sexto grado, y a los docentes de todos los grados.

La intención en todo momento fue, la recuperación de evidencias empíricas que mostraran las prácticas cotidianas de los docentes en el proceso constructivo de las matemáticas, el uso de estrategias y recursos didácticos, su planeación didáctica y la evaluación. Con lo que respecta a la aplicación de entrevistas y encuestas, el propósito central fue la búsqueda de información sobre su percepción, importancia y utilidad que tienen las matemáticas en el contexto cotidiano, las dificultades que enfrentan docentes y alumnos en el desarrollo de competencias matemáticas (habilidades, conocimientos y actitudes), las necesidades de capacitación y actualización prioritarias (docentes), así como obtener información específica que permitiera complementar las otras acciones de recolección utilizadas.

## 5. Resultados

En el quehacer docente, las profesoras y profesores de los diferentes niveles y modalidades de la educación básica ponen en juego estrategias diversas para lograr que los estudiantes aprendan de modo efectivo, sin embargo son muy frecuente los obstáculos con los que se enfrentan de forma cotidiana, es en este sentido importante identificar aquellos que representan mayor problema para el logro del desarrollo de competencias matemáticas.

Principales problemáticas que los docentes consideran que influyen en su práctica pedagógica:

- a) La puesta en práctica de las sugerencias metodológicas en la enseñanza y el aprendizaje de las matemáticas
- b) Pocos materiales y escasas actividades de manipulación
- c) Habilidades difíciles de desarrollar en los alumnos
- d) Falta de motivación de los alumnos por las matemáticas
- e) Contenidos que son difíciles de abordar con los alumnos.
- f) Grupos numerosos.
- g) Falta de tiempo para llevar a cabo las sesiones de matemáticas
- h) Falta de capacitación y actualización referente a estrategias didácticas
- i) Apoyo de los padres de familia

## 6. Propuesta

La propuesta está basada en el uso de materiales concretos que propician el desarrollo de competencias matemáticas (habilidades, conocimientos y actitudes) utilizando el enfoque del planteamiento y resolución de problemas en donde docentes y alumnos interactúan con diversos materiales manipulativos. Además del diseño y aplicación de planeaciones didácticas que permite a los docentes facilitar el proceso de enseñanza aprendizaje. En estas fichas de trabajo se establecen actividades que permiten identificar el eje temático a trabajar, temas y subtemas, así como de las consignas a trabajar con relación al planteamiento y resolución de problemas con el

apoyo de materiales manipulativos los cuales tienen una estructura que permite abordar el contenido de una forma fácil y posteriormente comprender el algoritmo formal que se plantea.

Se argumenta que el material concreto, juega un papel fundamental en la enseñanza-aprendizaje de las matemáticas, porque su correcta utilización constituye una importante base en la adquisición de conceptos, relaciones y métodos geométricos (por ejemplo) ya que posibilita una enseñanza activa de acuerdo a la evolución intelectual del alumno. También asumen la estructura de la propuesta como una filosofía que enfatiza el “aprender haciendo” y rompe frontalmente con los métodos de enseñanza formales: “Es un sistema basado en el aprendizaje activo y localizado en el proceso de aprendizaje, más que un proceso de enseñanza.”

## 7. Conclusiones

Es preciso terminar con la visión negativa que se tiene actualmente sobre las matemáticas, que origina rechazo a la misma, por ello es necesaria la implementación de estrategias para la enseñanza y el aprendizaje, en donde el alumno pueda aplicar dichos conocimientos en problemas afines a su realidad. Es momento de innovar, estableciendo y reconociendo la importancia que tiene que el alumno tome las riendas de su propio aprendizaje. La enseñanza de las matemáticas debe aproximarse a la cotidianidad del educando; para ello hay que reflexionar sobre cuáles son los centros de mayor interés que tiene la población de alumnos a los que se dirigen los esfuerzos, los cuales son en muchas ocasiones distintos a los del profesor que imparte la asignatura, por lo tanto, se debe ampliar el campo de visión para considerar aspectos que tradicionalmente no son abordados, en donde tenga cabida lo lúdico, lo imaginario, lo mítico.

Es importante reflexionar que aunque los propósitos planteados en el “Plan y programas de estudio de educación primaria” privilegian la adquisición y el desarrollo de competencias matemáticas, no existen elementos que puedan brindar información sobre los logros alcanzados. En este sentido se plantea como una necesidad una propuesta que apoye el proceso de enseñanza y aprendizaje de las matemáticas en educación primaria, que permita fortalecer el desarrollo de competencias y en donde el alumno sea protagonista de su propio aprendizaje y que el uso de recursos didácticos a través de materiales manipulables sea el eje central en este proceso.

## 8. Referencias

- [1] Alicia Ávila. Informes de Investigación. Temas prioritarios. La Reforma realizada. *La resolución de problemas como vía del aprendizaje en nuestras escuelas*. SEP.2004.
- [2] Block, David y Martha Dávila (1995), "*La matemática expulsada de la escuela*", en SEP, *La enseñanza de las matemáticas en la escuela primaria*. Lecturas, México.
- [3] Casanova, Maria. Antonia. (1989) "*La evaluación educativa*". Escuela básica. Biblioteca para la actualización del maestro. México. SEP.
- [4] CIEFI (2007). Base de datos: entrevistas, observaciones en el aula, encuestas, planes estratégicos.

- [5] *Fichero de actividades didácticas. Matemáticas. Educación primaria, 1° 3° 5° y 6° grado México.*
- [6] G. Polya. *Como plantear y resolver problemas.* Serie Matemáticas. Editorial Trillas 1996.
- [7] *Libro para el maestro. Matemáticas. Primaria 1° 3° 5° y 6° grado,* México.
- [8] SEP (1993), *Plan y programas de estudio. Educación básica. Primaria,* México.



# Diseño, construcción y control a distancia de una grúa autoequilibrada

Rogelio Francisco Antonio y Diego I. Chavero Nuñez  
División de Ingeniería Mecánica Electrónica  
Universidad Politécnica del Valle de México  
rfrancisco2002@yahoo.com.mx

## Resumen

*En este trabajo se presenta el diseño, construcción y el control en tiempo real (control remoto) de una grúa equilibra, donde se muestra las diferentes aplicaciones del control a distancia. Además se presenta las etapas del diseño mecánico, diseño electrónico que constituye la construcción del prototipo, así como la implementación en tiempo real, cuyo control empleado es por radiofrecuencia. El prototipo es de tipo didáctico, cuyo objetivo es mostrar las diferentes estrategias de control que puedan implementarse.*

## 1. Introducción

En la actualidad las industrias han estado cambiando debido a la transformación que sufre la tecnología y las exigencias de la sociedad actual. Una de las áreas más desarrollada es el control de procesos cuya finalidad es obtener un mejor rendimiento y optimización del proceso. El área de control puede afirmarse que tiene muchas aplicaciones en la industria actual, debido a la facilidad de adaptación e implementación. Considerando esta situación podremos destacar que para el transporte de material y equipo de gran dimensión y peso el control es relevante ya que se necesita desde máquinas simples como poleas, polipastos donde se aplican las propiedades fundamentales de la física y hasta el uso de máquinas sofisticadas como la grúa. Una grúa es una máquina de elevación, traslación y distribución de cargas. En general están constituidas por poleas acanaladas y contrapesos para poder mover pesos vertical y horizontalmente. Además, algunos utilizan máquinas simples (que requieren el impulso de una sola fuerza para trabajar) para crear ventaja mecánica y lograr mover grandes cargas. Haciendo algo de historia, las primeras grúas fueron inventadas por los griegos, a las cuales eran accionadas por hombres o animales. Estas grúas eran utilizadas para la construcción de estructuras altas. Posteriormente, fueron desarrolladas grúas más grandes utilizando poleas para permitir la elevación de mayores pesos. Las primeras grúas se construyeron de madera, pero desde la llegada de la revolución industrial los materiales más utilizados son el hierro fundido y el acero. Las grúas existen en una enorme variedad de formas, cada una adaptada a un uso específico entre ellas se encuentran fijas, apoyadas, empotradas móviles, con traslación, trepadora, telescópica, por su pluma (grúa de pluma horizontal, grúa de pluma abatible).

El prototipo que se presenta no sólo es de carácter educativo donde se pueda ilustrar las diferentes técnicas de diseño y leyes de control existentes en la actualidad, además este prototipo cumple con las especificaciones técnicas de diseño, las cuales son útiles para futuras implementaciones a nivel industrial, destacando que ahora sólo se presentan los primeros resultados de interés particular. En la Figura 1 se muestra una grúa a nivel industrial.

Actualmente el control de las grúas está pasando de operación manual, que depende de la habilidad de un operador, a un control automatizado, sobre todo si son muy grandes las cargas y se deben mover con altas velocidades, en consecuencia nuevos métodos de automatización se están desarrollando. Por ejemplo combinación de leyes de control clásico con leyes de control moderno, también algunas investigaciones utilizan controladores de lógica difusa para el control automatizado de las grúas.

El objetivo principal de este trabajo es: diseñar y construir una grúa para levantar una carga y desplazarla de un punto a otro en su espacio de trabajo, basando su funcionamiento en el auto equilibrio.

## 2. Descripción del sistema y etapas

La grúa que se propone en este proyecto es un sistema mecatrónico experimental que basa su funcionamiento en el equilibrio de momentos como se muestra en la Figura 1. En esta grúa se reduce en gran medida el lastre y elimina el anclaje, dado que va a ser 'auto-equilibrada'. El concepto de auto-equilibrio, implica prescindir de un sistema de anclaje, substituyendo por un contrapeso desplazable como se muestra en la figura indicando también las dimensiones de longitud, altura y carga. Las etapas que constituyen el trabajo son:

- Diseño del sistema mecánico
- Construir todos los elementos principales que componen una grúa auto-equilibrada
- Construir un sistema de manipulación a distancia por radio frecuencia
- Ensamblar todas las partes mecánicas y de control
- Obtener un buen funcionamiento del diseño satisfaciendo las necesidades de trabajo

## 3. Modelado matemático del sistema

Es muy importante mencionar que en esta primera propuesta de grúa auto equilibrada, se considera únicamente la oscilación de la carga en el plano  $xy$ , sabiendo que la carga (masa  $m_l$ ) puede tener una oscilación tanto en el plano  $xy$  como en el plano  $xz$ . En algunas grúas reales, el análisis dinámico es en dos dimensiones por el tipo de grúa y la forma de sujeción de la carga, como el caso de desplazar contenedores en los puertos y barcos cargueros.

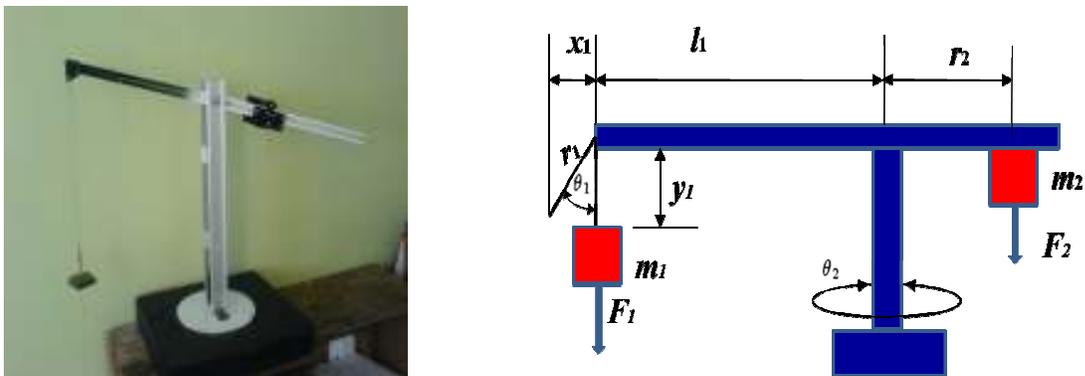


Figura 1. Equilibrio de momentos en la grúa.

Para compensar la oscilación lateral de la carga (en el plano  $xz$ ) se requiere de un mecanismo similar de compensación para lograr el equilibrio a la altura del mástil y de forma perpendicular al plano  $xy$  como se muestra en la Figura 1, el cual por el momento no se implementa en este sistema, debido a que antes de levantar la carga, la grúa se va a girar hasta obtener un ángulo nulo en el  $xz$ . Lo anterior se medirá con el sensor que se utiliza en la detección del ángulo  $\theta_1$ , el cual cuenta con dos potenciómetros, colocados uno perpendicularmente respecto del otro.

En las grúas anteriormente descritas el control está dado por la habilidad de un operador, quedando en él la responsabilidad de su adecuado funcionamiento. En la que se desarrolla, se incluye una ley de control para lograr el auto-equilibrio y funcionamiento adecuado, además de dispositivos actuadores y sensores, para el posicionamiento de los contrapesos.

El ángulo máximo de oscilación de la carga es un factor primordial en el diseño de la grúa auto-equilibrada. Para compensar y anular la oscilación de la masa  $m_1$ , el carro no debe desplazarse más del 10% de la longitud de la pluma, por los sobretiros en la respuesta de la ley de control para desplazar el carro; considerando que la pluma tiene aproximadamente 100 *cm*, el desplazamiento máximo  $x_1$  es de 10 *cm*. Con base en la Figura 1 se tiene que:

$$x_1 = r_1 \text{sen } \theta_1$$

Si  $r_1 \text{ max} = 130$  *cm*, y realizado un cálculo sencillo se obtiene  $\theta_1 \text{ max} = 4.413^\circ$ , valor que proporciona una estimación del ángulo de oscilación máximo en la carga. Sin embargo siendo conservadores y para la simulación se propone un  $\theta_1 = 3^\circ$ , dá como resultado un desplazamiento del carro de 6.8 *cm*. Con base en estos cálculos, para una grúa real que puede levantar una masa 30 *m*, con oscilación de  $\theta_1 = 3^\circ$  el desplazamiento del carro es de 1.57 *m* a lo largo de la pluma (eje *x*). Aparentemente el valor de dicho ángulo es muy pequeño, observando con estos datos que tiene un efecto considerable en la longitud de desplazamiento del carro para compensar la oscilación de la carga.

En esta parte se considera el levantamiento de la carga (masa  $m_1$ ) y el desplazamiento horizontal del contrapeso (masa  $m_2$ ) para mantener el equilibrio. Con base en la Figura 1, se hacen las siguientes consideraciones:

El valor del  $\text{sen}(\theta_1)$  se representará como  $S_1$ , o sea que:  $\text{sen}(\theta_1) = S_1$ ,

El valor del  $\text{cos}(\theta_1)$  se representará como  $C_1$ , o sea que:  $\text{cos}(\theta_1) = C_1$ ,

El método de Euler-Lagrange plantea que:

$$L = K_T - V_T$$

Donde

$L$  = Lagrangiano del sistema;  $K_T$  = Energía cinética total del sistema, y  $V_T$  = Energía potencial total del sistema.

Ahora, considerando los movimientos lineales, entonces de la Figura 3 se obtiene:

$$x_1 = -r_1 \text{sen}\theta_1 = -r_1 S_1$$

$$y_1 = r_1 \text{cos}\theta_1 = r_1 C_1$$

$$x_2 = r_2$$

$$y_2 = h_2 = 0$$

Por lo tanto

$$V_T = V_1 = m_1 g r_1 C_1$$

De aquí el Lagrangiano del sistema es:

$$L = \frac{1}{2} J_1 \dot{\theta}_1^2 + \frac{1}{2} m_1 r_1^2 \dot{\theta}_1^2 + \frac{1}{2} m_1 \dot{r}_1^2 + \frac{1}{2} m_2 \dot{r}_2^2 + m_1 g \dot{r}_1 C_1$$

Aplicando la ecuación de movimiento de Euler Lagrange a  $r_1$ :

$$\frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{r}_1} \right) - \left( \frac{\partial L}{\partial r_1} \right) = \tau_{r_1} \quad (1)$$

Donde 
$$\frac{\partial L}{\partial \dot{r}_1} = m_1 \dot{r}_1$$

$$\frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{r}_1} \right) = m_1 \ddot{r}_1 \quad (2)$$

$$\frac{\partial L}{\partial r_1} = m_1 \dot{\theta}_1^2 r_1 + m_1 g C_1 \quad (3)$$

Sustituyendo (3) y (2) en (1)

$$m_1 \ddot{r}_1 - m_1 \dot{\theta}_1^2 r_1 - m_1 g C_1 = \tau_{r_1} \quad (4)$$

Aplicando Euler Lagrange a  $\theta_1$ :

$$\frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{\theta}_1} \right) - \left( \frac{\partial L}{\partial \theta_1} \right) = \tau_{\theta_1}$$

$$\left( J_1 + m_1 r_1^2 \right) \ddot{\theta}_1 + m_1 g r_1 S_1 = 0 \quad (5)$$

$$m_1 = \frac{r_2}{l_1 C_1} m_2 \quad (6)$$

Sustituyendo (6) en (4) y (5) tenemos las ecuaciones que definen nuestro sistema en función de  $m_2$ :

$$\frac{r_2}{l_1 C_1} m_2 \ddot{r}_1 - \frac{r_2 r_1}{l_1 C_1} m_2 \dot{\theta}_1^2 - \frac{r_2}{l_1} m_2 g = \tau_{r_1}$$

$$\left( J_1 + \frac{r_2}{l_1 C_1} m_2 r_1^2 \right) \ddot{\theta}_1 + \frac{r_2}{l_1 C_1} m_2 g r_1 S_1 = 0$$

$$m_2 \dot{r}_2 = \tau_{r_2}$$

### 3.1 Modelado dinámico del segundo movimiento

Obteniendo el modelo dinámico cuando se tiene giro de la grúa ( $\theta_2$ ), como se muestra en la Figura 1. Para el análisis de este movimiento, se asume que la carga está en reposo, y por lo tanto  $\theta_1 = 0$ .

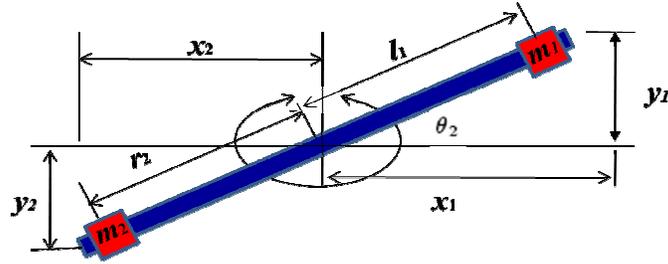


Figura 2. Movimiento angular de la grúa.

$$\begin{aligned}x_1 &= l_1 C_2 \\y_1 &= l_1 S_2 \\x_2 &= r_2 \cos \theta_2 = r_2 C_2\end{aligned}\quad (7)$$

$$y_2 = r_2 \text{ sen } \theta_2 = r_2 S_2 \quad (8)$$

$$|v_2|^2 = \dot{x}_2^2 + \dot{y}_2^2 \quad (9)$$

Derivando (7), (8) y sustituyendo en (9) tenemos:

$$\begin{aligned}|v_2|^2 &= \left(-r_2 \dot{\theta}_2 S_2\right)^2 + \left(r_2 \dot{\theta}_2 C_2\right)^2 \\|v_2|^2 &= r_2^2 \dot{\theta}_2^2\end{aligned}$$

Sustituyendo en  $K_T$

$$K_T = \frac{1}{2} J_2 \dot{\theta}_2^2 + \frac{1}{2} m_1 l_1^2 \dot{\theta}_2^2 + \frac{1}{2} m_2 r_2^2 \dot{\theta}_2^2$$

Como sabemos  $L = K - V$  y  $V=0$  dado que  $r_1$  es muy pequeña entonces  $L = K_T$ .

$$L = \frac{1}{2} J_2 \dot{\theta}_2^2 + \frac{1}{2} m_1 l_1^2 \dot{\theta}_2^2 + \frac{1}{2} m_2 r_2^2 \dot{\theta}_2^2$$

Aplicando el lagrangiano

$$\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{\theta}_2} \right] - \left[ \frac{\partial L}{\partial \theta_2} \right] = \tau_{\theta_1} \quad (10)$$

$$\frac{\partial L}{\partial \dot{\theta}_2} = J_2 \dot{\theta}_2 + m_1 l_1^2 \dot{\theta}_2$$

$$\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{\theta}_2} \right] = (J_2 + m_1 l_1^2 + m_2 r_2^2) \ddot{\theta}_2 \quad (11)$$

$$\frac{\partial L}{\partial \theta_2} = 0 \quad (12)$$

Sustituyendo (11) y (12) en (10) tenemos

$$(J_2 + m_1 l_1^2 + m_2 r_2^2) \ddot{\theta}_2 = \tau_{\theta_1}$$

Finalmente el modelo expresado en forma matricial es

$$\begin{bmatrix} \left(\frac{r_2}{l_1 \cos \theta_1}\right) m_2 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & J_1 + \left(\frac{r_2}{l_1 \cos \theta_1}\right) m_2 r_1^2 \end{bmatrix} \begin{bmatrix} \ddot{r}_1 \\ \ddot{r}_2 \\ \ddot{\theta}_1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & -\left(\frac{r_2 r_1}{l_1 \cos \theta_1}\right) m_2 \dot{\theta}_1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \dot{\theta}_1 \end{bmatrix} + \begin{bmatrix} -\left(\frac{r_2}{l_1}\right) m_2 g \\ 0 \\ \left(\frac{r_2 r_1}{l_1}\right) m_2 g \tan \theta_1 \end{bmatrix} = \begin{bmatrix} \tau_{r_1} \\ \tau_{r_2} \\ \tau_{\theta_1} \end{bmatrix} \quad (13)$$

Cuyo modelo general en forma matricial es:

$$M(r, \theta) [\ddot{r}, \ddot{\theta}] + C(r, \dot{\theta}) [\dot{\theta}] + G(r, \theta) = \tau \quad (14)$$

y si  $\theta_1 = 0$  con  $r_2 = \text{constante}$ , se tiene la siguiente ecuación independiente

$$(J_2 + m_1 l_1^2 + m_2 r_2^2) \ddot{\theta}_2 = \tau_{\theta_1} \quad (15)$$

#### 4. Desarrollo del prototipo

En esta etapa se muestran el diseño de la estructura de la grúa y las etapas de construcción e implementación física.

##### 4.1 Diseño electrónico

El sistema electrónico en la grúa auto-equilibrada se encarga de acondicionar y procesar las señales proporcionadas por los sensores, de posición de las diferentes articulaciones, fin de carrera, para que puedan ser utilizadas por el sistema de procesamiento. En la Figura 3 se presenta el diagrama del circuito electrónico para la manipulación de la grúa.

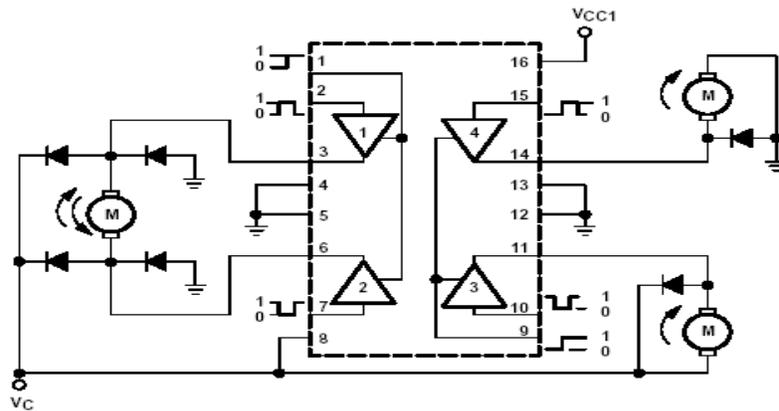


Figura 3. Diagrama del circuito de control.

##### 4.2 Diseño mecánico

En la Figura 4 se presenta el diseño y la simulación realizada en el software Solid Works



**Figura 4. Prototipo final.**

## 5. Conclusión

La grúa propuesta en este trabajo es una grúa auto equilibrada, que reduce en gran medida el lastre y elimina el anclaje, esta grúa es experimental, debido a que desarrollar un sistema mecánico de dimensiones reales es muy costoso y se corre el riesgo de destruirlo en las pruebas experimentales. El concepto de auto equilibrio, implica prescindir de un sistema de anclaje, substituyéndose por un contrapeso desplazable, como se muestra en la Figura 1. Además, este tipo de prototipos permite al estudiante poder implementar diferentes estrategias de control para ampliar su conocimiento.

## 6. Referencias

- [1] Ogata Katsuhiko, *Ingeniería de Control Moderna*, 4ª Edición, Pearson – Prentice, Hall Hispanoamericana, México, 2003
- [2] Craig John J., *Robótica*, 3ª Edición, Ed. Pearson Educación, S.A., México, 2006.
- [3] Bolton William, *Mecatrónica, Sistemas de control electrónico en ingeniería, mecánica y eléctrica*, 2ª Edición, Ed. Alfaomega, México, 2001.
- [4] Kelly, R. y Santibáñez V., *Control de Movimiento de Robots Manipuladores*, Pearson Educación, S.A., Madrid, 2003.
- [5] Barrientos A. , Peñín L. F. , *Fundamentos de Robótica*, 2ª. Edición, Mc. Graw Hill, 2007



# Nuevas perspectivas en la didáctica de estadística en el área de ciencias sociales

María Elena Parra Ramos  
Departamento de Matemáticas  
Universidad de Sonora  
meparra@gauss.mat.uson.mx

## Resumen

*En el presente trabajo, se pretende abordar uno de los problemas tradicionales en la enseñanza de la estadística en el área de ciencias sociales, consistente en el desfase entre los conceptos teóricos y la interpretación de éstos. Con el objetivo de promover un cambio en el paradigma del maestro como fuente única del conocimiento y desarrollar en el estudiante una formación activa, crítica y eficiente en el análisis de información. Lo anterior, mediante el uso tanto de paquetes computacionales como el SPSS (Statistical Package for the Social Sciences), uno de los más utilizados en investigación aplicada en esta disciplina, así como applets Java. Con ambos recursos se puede realizar actividades de tipo colaborativo en el aula que contribuyan a nuestro objetivo.*

## 1. Introducción

En la Universidad de Sonora, al igual que en otras instituciones de educación superior, se han realizado modificaciones en los planes de estudio de las diferentes disciplinas, en aras de una adecuación a los nuevos modelos educativos. Entre estas modificaciones se incluye la impartición de los cursos presenciales con la inclusión de nuevas tecnologías y la utilización de actividades en las que se involucre a los estudiantes en un aprendizaje más activo.

Acorde a tales propósitos institucionales, en el plan de estudios de estudios de la materia de estadística del área de ciencias sociales, se indica la necesidad de desarrollar en el estudiante una formación crítica y activa en el procesamiento de información concerniente a una situación concreta de la problemática social. Así mismo, se plantea el uso de software computacional haciendo especial énfasis en la utilización del SPSS. Como en la actualidad la mayoría de las investigaciones en ese ámbito involucra una gran cantidad de información a procesar, el gran volumen de información en cada estudio dificulta su procesamiento de forma manual. Lo anterior nos permite señalar que los paquetes tecnológicos para el tratamiento de datos son de gran ayuda en la realización de las tareas de organización, reducción y análisis haciendo que éstas sean más efectivas, precisas y exhaustivas, puesto que permiten el manejo de un gran volumen de información en menor tiempo, lo que permitiría profundizar en el análisis de datos.

## 2. Uso del paquete computacional SPSS en la enseñanza de la estadística

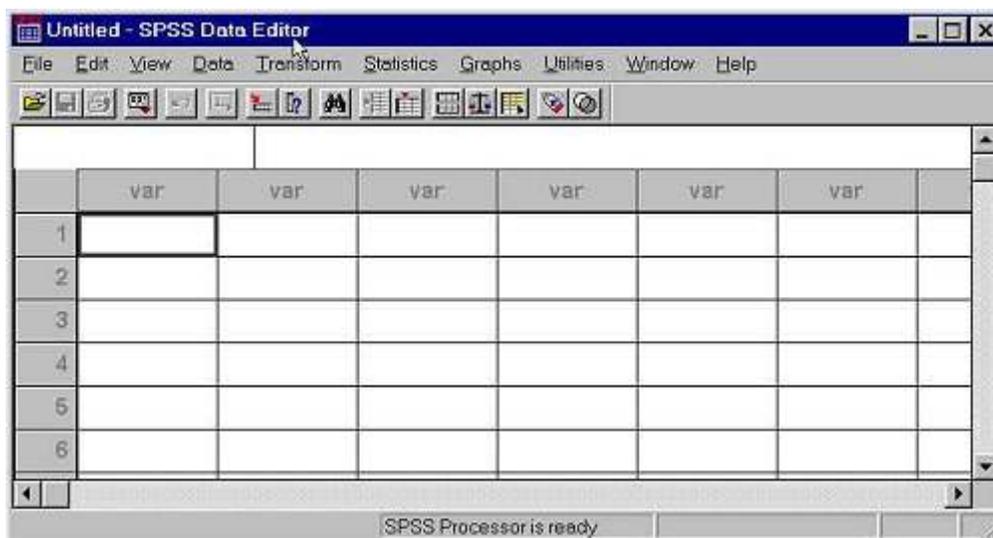
El tratamiento automatizado de la información ha sido una de las tecnologías que más impacto ha tenido en el uso de la estadística. Actualmente el uso de paquetes estadísticos permite la

aplicación de métodos y técnicas modernas de análisis de datos; entre éstos, se encuentra el SPSS que es uno de los más utilizados en investigación aplicada a las ciencias sociales.

El paquete estadístico SPSS permite analizar con detenimiento las variables implicadas en la investigación con el propósito de construir un modelo que sea capaz de explicar lo que aconteció, tanto antes como después del análisis estadístico. Asimismo, puede ser utilizado como hoja de cálculo, realizar funciones aritméticas, algebraicas y trigonométricas sobre un fichero de datos, gestionar de modo dinámico la información de un fichero de datos, pues se pueden actualizar los cambios operados (como ordenar, filtrar, etc.) o realizar informes personalizados de acuerdo con distintos criterios, etc. Así mismo, posibilita la preparación de atractivos informes de una investigación realizada al permitir incorporar en un mismo archivo el texto del reporte, las tablas y resultados estadísticos que el reporte necesite presentar e, incluso, los gráficos que se pudiesen generar. Como analizador de datos, se tiene la capacidad de extraer de un fichero de datos toda la información recogida, permitiendo realizar procedimientos estadísticos descriptivos, inferenciales y multivariantes y como ejecutor de minerías de datos.

Una de las ventajas de este paquete es la sencillez de manejo ya que mediante el despliegue de un menú interactivo permite realizar todo tipo de operaciones sobre los datos como así también aplicar distintas técnicas estadísticas.

Los pasos básicos en el análisis de datos utilizando este paquete son: a) creación de la base de datos desde SPSS o bien la captura de los datos desde otro archivo; b) selección de los procedimientos de análisis a emplear; c) selección de las variables para los distintos análisis; d) análisis de los resultados.



**Figura 1. Pantalla básica del SPSS.**

En la Figura 1 se exhibe una de las pantallas básicas con las que cuenta el programa. En ella se puede observar acciones que se pueden realizar con el mismo, tales como abrir, cerrar o exportar archivos, introducir variables, realizar cálculos estadístico y graficar entre otros.

Durante el transcurso de una sesión en SPSS se activan básicamente dos pantallas: la que corresponde al editor de datos y el visor de resultados. La primera de ellas se activa automáticamente cuando se inicia la sesión, mientras que la otra se habilita al efectuar algún procedimiento. Esta última se encuentra subdividida en dos partes: a la izquierda de la pantalla aparece una estructura de árbol que muestra los procedimientos realizados y en qué orden se efectuaron. A la derecha aparecen los resultados propiamente dichos, tales como tablas de frecuencias, gráficos, los resultados del cálculo de medidas descriptivas. Esto lo podemos observar en la Figura 2.



Figura 2. Visor de resultados.

En la Figura 3 se muestra la pantalla de datos. Este sistema permite realizar las acciones de una forma amigable al crear y analizar una base de datos muy extensa a partir de tamaños de muestra mayores y un número más grande de variables. Así mismo, se obtiene un importante ahorro de tiempo y esfuerzo, al ejecutar en segundos un trabajo que requeriría horas e incluso días.

En el aula, se puede realizar actividades de tipo colaborativo haciendo uso del SPSS tal como la que se enuncia a continuación:

1. Primeramente se forman equipos, a los cuales se les pide designar en cada equipo los estudiantes que representarán los diferentes roles de un trabajo colaborativo (secretario, abogado del diablo, entre otros ).
2. A continuación, con los datos obtenidos de una encuesta realizada a un grupo de estudiantes (que incluye preguntas sobre sus antecedentes académicos, condiciones de estudio y nivel económico), se les pide formar una base de datos con el paquete de cómputo SPSS.
3. Como un siguiente paso, se les pide obtener las medidas representativas de los datos y analizarlas.
4. Con el fin de obtener una visión preliminar de la distribución de los datos, se les solicita obtener los gráficos apropiados para algunas variables y realizar un análisis de los mismos.

- Para concluir esta actividad, se pide a las personas con el papel de secretarios dar a conocer sus resultados al resto del grupo, lo que permite el análisis de la información con la totalidad de los alumnos bajo la supervisión del maestro.

BASE COMPLETA SOLO 2003 - SPSS Data Editor

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1 : educ 3

	educ	estciv	edadcasa	Género	región	edad	pensami	niveduc
1	3	1	45	1	4	53	8	
2	0	1	14	1	3	63	4	
3	6	1	18	1	3	40	4	
4	12	1	25	2	3	52	8	
5	9	1	17	1	4	61	7	
6	9	1	16	1	4	27	4	
7	12	1	20	1	4	35	4	
8	7	1	16	1	4	61	2	
9	14	1	36	1	4	79	4	
10	10	1	18	1	1	54	4	

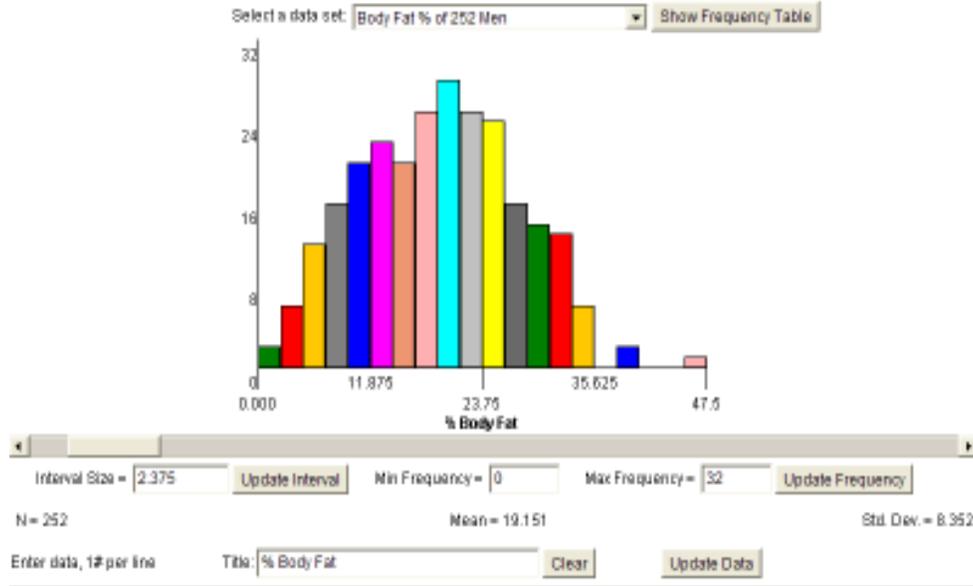
Vista de datos Vista de variables

Figura 3. Pantalla de datos.

### 3. Uso de otros recursos de aprendizaje

Otra de las acciones a realizar es el uso de *applets* Java. La propuesta de uno de ellos se muestra a continuación: El objetivo de la presente actividad es permitir a los estudiantes crear sus propios histogramas con diferentes conjuntos de datos e intervalos de clase; lo anterior, permite al estudiante desarrollar habilidades en la lectura de éstos. Así mismo, con este programa se puede propiciar el reconocimiento y comprensión de las representaciones de datos.

Para el logro de lo anteriormente mencionado, se hace uso de un *applet* de Java, el cual permite ver y crear histogramas a partir de conjuntos de datos pres existentes en el programa. También admite la construcción de datos por el usuario o la exportación de los mismos de otro archivo. Al realizar la gráfica, se puede modificar los intervalos de clase así como las frecuencias mínimas y máximas. Con lo anterior, se puede analizar su efecto en la imagen obtenida. Adicionalmente al gráfico, se presentan medidas representativas de los datos tal como se muestra en la Figura 4.

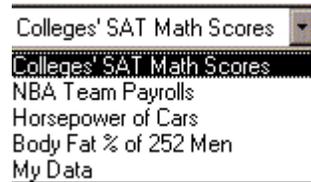


**Figura 4. Información que presenta el *applet*.**

Instrucciones para los estudiantes:

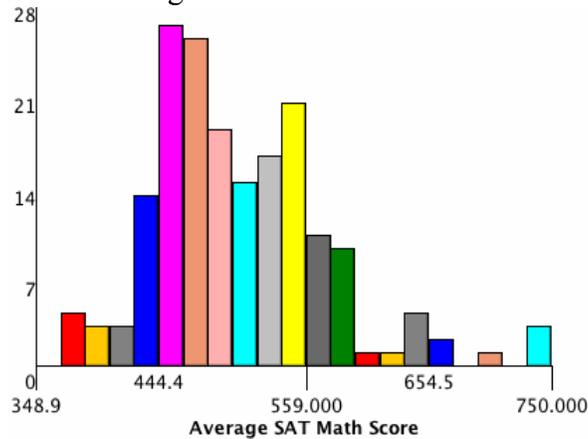
En la página web <http://www.shodor.org/interactivate/activities/Histogram/> hay un *applet* que emplearás para la siguiente actividad:

a) Elige un conjunto de datos de la lista de datos existentes en el menú desplegable del programa tal como se muestra en la Figura 5.



**Figura 5. Conjunto de datos del programa.**

b) Ya seleccionado el conjunto de datos, observa el histograma que se despliega y analiza la información que se presenta en la imagen obtenida.



**Figura 6. Histograma.**

c) En la parte inferior de la gráfica se presenta información sobre el tamaño de intervalo, así como los valores máximo y mínimo tal como se presenta en la Figura 7. Selecciona la opción *Update Frequency* con la que obtienes la tabla de frecuencias del conjunto de datos seleccionado.



**Figura 7. Selección del intervalo y valores de frecuencia.**

d) A continuación, modifica el tamaño del intervalo (introduce otro valor en la sección correspondiente a *Interval Size*); posteriormente, selecciona la opción *Update Interval* para que se aplique el cambio. ¿Qué observas? Así mismo, puedes modificar los valores máximo y mínimo de las frecuencias sin olvidar la selección de la opción *Update Frequency* para que aplique el cambio. ¿Qué diferencia notas?

e) Ahora selecciona la opción *My data* e introduce en la parte inferior del *applet* tus propios datos, los cuales debes introducir de uno en uno en forma de columna. O bien, puedes utilizar un conjunto de datos ya existente los cuales colocas con los comandos usuales en computadora (*copy, paste*) sin olvidar que estos datos se deben ordenar en columnas de un elemento. Selecciona la opción *Update Data* para que aplique el cambio. Realiza las acciones de los incisos c, d y e. Anota tus observaciones.

f) En la información que se presenta, además del histograma, se incluyen medidas de representación de los datos. Identificalas y anota tus observaciones de la imagen original y las obtenidas con los cambios.

## 4. Conclusiones

Al resolver un problema, se dependía del cálculo correcto de operaciones aritméticas, un tanto tediosas para los estudiantes especialmente los inscritos en las ciencias sociales. Actualmente, las herramientas tecnológicas entre ellas el uso de *applets* interactivos y paquetes computacionales como el SPSS son de gran utilidad en la enseñanza de estadística; lo que permite propiciar un entorno en el que se pueda llevar a cabo un aprendizaje significativo al optimizar el proceso educativo desarrollando y potenciando aptitudes intelectuales.

## 5. Referencias

- [1] Batanero, C. (2000). “¿Hacia dónde va la educación estadística?” Recuperado el 10 de octubre del 2008: <http://www.ugr.es/~batanero/ARTICULOS/BLAIX.html>.
- [2] Pallant, J. (2003). “*SPSS Survival Manual. A Step by Step Guide to Data Analysis Using SPSS*”. Ed. Open University Press.
- [3] Visauta, B. (1997). “*Análisis estadístico con SPSS para Windows*”. Volumen I y 11. Estadística Básica. Madrid: Mc Graw – Hill.

# La osculación como un medio para arribar a las series de Taylor y McLaurin apoyado con el uso de la tecnología

Mario Silvino Ávila Sandoval y Carlos López Ruvalcaba  
Departamento de Ciencias Básicas Exactas  
Universidad Autónoma de Ciudad Juárez  
m\_s\_avila@yahoo.com.mx, clopez@uacj.mx

## Resumen

*Una revisión de los textos de cálculo nos muestra que la forma en que se tratan las series de Taylor y McLaurin es desde una perspectiva estrictamente algebraica, generando en el estudiante que aborda estos temas una marcada escasez de sentido. El presente trabajo muestra una forma alternativa de arribar a las series de Taylor y McLaurin apoyada por el uso de la tecnología y la osculación de funciones. Los métodos implicados en este proceso nos dan pie a la construcción de series diferentes, no necesariamente polinomiales, que funcionan de manera análoga a las series conocidas, además de poder resolver problemas de movimiento entre otras aplicaciones.*

## 1. Introducción

El punto de partida es la posibilidad de encontrar la ecuación de una recta que pasa por los puntos del plano  $(x_1, y_1)$  y  $(x_2, y_2)$  por medio de la Expresión 1 [1].

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0 \quad (1)$$

Por otro lado, es posible obtener la ecuación de la parábola con eje focal paralelo al eje  $y$  que pasa por los puntos  $(x_1, y_1)$ ,  $(x_2, y_2)$  y  $(x_3, y_3)$  evaluando la Expresión 2 [2]:

$$\begin{vmatrix} x^2 & x & y & 1 \\ x_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0 \quad (2)$$

Se puede plantear la posibilidad de expandir el determinante para obtener ecuaciones de polinomios de mayor grado, dados los puntos necesarios. Por otro lado, dichos puntos pueden pertenecer a la curva de una función  $f(x)$ . En el presente trabajo, se explorará la idea de encontrar las ecuaciones límite en el proceso de aproximar dos o más puntos, situados sobre la gráfica de una función, con el fin de obtener polinomios tangentes u osculadores, haciendo uso del software *Mathematica*.

## 2. La ecuación de la recta tangente

Utilizando las ideas anteriores, calcularemos la ecuación de la recta tangente a la gráfica de la función  $f(x)$  en  $x = a$  por medio de evaluar la Expresión 3.

$$\lim_{n \rightarrow \infty} n \cdot \begin{vmatrix} x & y & 1 \\ a & f(a) & 1 \\ a + \frac{1}{n} & f(a + \frac{1}{n}) & 1 \end{vmatrix} = 0 \quad (3)$$

Notar que la anterior, calcula la ecuación límite de una recta cuando la separación entre dos puntos sobre la curva de  $f(x)$  tiende a cero.

Análogamente, y dado que una parábola necesita tres puntos para definirse, pueden colocarse dos de estos puntos sobre la gráfica de una función para hacerlos converger manteniendo el tercero fijo en cualquier punto del plano, obteniendo la ecuación de una parábola tangente. Otra posibilidad es colocar los tres puntos sobre la curva de una función  $f(x)$  (uno en  $(a, f(a))$ , otro en  $(a + \frac{1}{n}, f(a + \frac{1}{n}))$  y el tercero en  $(a - \frac{1}{n}, f(a - \frac{1}{n}))$ ) y calcular el límite de la ecuación cuando la separación entre ellos tiende a cero, es decir,  $n \rightarrow \infty$ . Esto nos lleva a la Expresión 4.

$$\lim_{n \rightarrow \infty} n^3 \cdot \begin{vmatrix} x^2 & x & y & 1 \\ a^2 & a & f(a) & 1 \\ (a + \frac{1}{n})^2 & a + \frac{1}{n} & f(a + \frac{1}{n}) & 1 \\ (a - \frac{1}{n})^2 & a - \frac{1}{n} & f(a - \frac{1}{n}) & 1 \end{vmatrix} = 0 \quad (4)$$

La ecuación obtenida a partir de la Expresión 4, no solamente corresponde a una parábola tangente a la gráfica de  $f(x)$ , sino también comparte con ésta la primera y la segunda derivada en  $x = a$ , a la que podemos llamar *parábola osculatriz*. A continuación se muestra un ejemplo que ilustra esta situación.

**Ejemplo 1.** Consideremos la función  $f(x) = e^x$  y  $a = 0$  para el último caso, usemos *Mathematica* para evaluar la expresión y obtener la gráfica y la ecuación de la parábola resultante mostrada en la Figura 1.

Extendiendo la idea, para obtener un polinomio de grado  $k$  será necesario usar  $k + 1$  puntos sobre una función, mismos que se pueden hacer converger a uno solo, tanto por la derecha como por la izquierda o por ambos lados, esto se puede generalizar por medio de la Expresión 5. Ahora, el polinomio obtenido, compartirá con la función, hasta la  $k$ -ésima derivada en  $x = a$ .

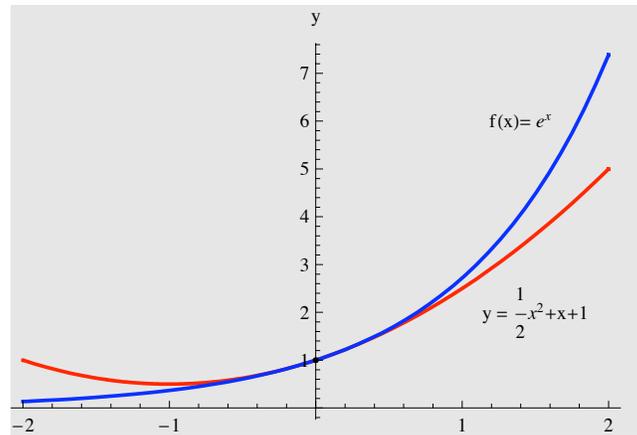


Figura 1. Parábola obtenida por la convergencia de tres puntos.

$$\lim_{n \rightarrow \infty} n^{\frac{k(k+1)}{2}} \cdot \begin{vmatrix} x^k & x^{k-1} & \dots & x^2 & x & y & 1 \\ a^k & a^{k-1} & \dots & a^2 & a & f(a) & 1 \\ (a + \frac{1}{n})^k & (a + \frac{1}{n})^{k-1} & \dots & (a + \frac{1}{n})^2 & a + \frac{1}{n} & f(a + \frac{1}{n}) & 1 \\ (a + \frac{2}{n})^k & (a + \frac{2}{n})^{k-1} & \dots & (a + \frac{2}{n})^2 & a + \frac{2}{n} & f(a + \frac{2}{n}) & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots \\ (a + \frac{k}{n})^k & (a + \frac{k}{n})^{k-1} & \dots & (a + \frac{k}{n})^2 & a + \frac{k}{n} & f(a + \frac{k}{n}) & 1 \end{vmatrix} = 0 \quad (5)$$

**Ejemplo 2.** Usando la Expresión 5, construyamos un polinomio de grado 4, situando 5 puntos sobre la misma función, alrededor del mismo punto. La Figura 2 muestra este polinomio y su gráfica junto con la obtenida en el ejemplo anterior.

Se puede observar en la Figura 2, que entre mayor es el grado del polinomio obtenido, más se aproxima localmente a la gráfica de la función, lo que marca un comportamiento similar a los polinomios de McLaurin. En la figura 3 se muestra el cálculo en *Mathematica* de un polinomio de McLaurin de cuarto grado. Observe que es el mismo que el obtenido usando la expresión 5. Por lo tanto, la anterior es una forma diferente de definir los polinomios de Taylor de grado  $k$ .

Lo anterior muestra una conceptualización de los polinomios de McLaurin y de Taylor (simplemente tomando  $a \neq 0$ ) desde una perspectiva geométrica:  $k + 1$  puntos sobre la gráfica de la función, convergiendo a uno solo formando un polinomio osculador de grado  $k$ . Con este método también es posible abordar problemas de movimiento uniformemente acelerado cuya función de posición es cuadrática con respecto al tiempo.

**Ejemplo 3.** Supongamos que se requiere encontrar la ecuación de posición de cierto móvil que se

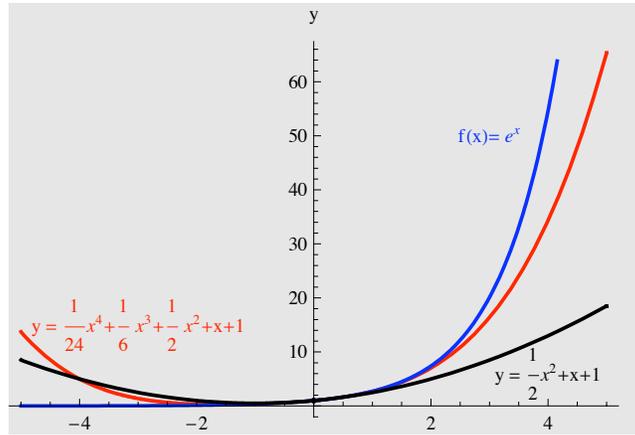


Figura 2. Polinomio de grado 4 junto con la parábola del ejemplo anterior.

```
In[7]:= f[x_] = e^x;
Series[f[x], {x, 0, 4}]
Out[8]= 1 + x + x^2/2 + x^3/6 + x^4/24 + O[x]^5
```

Figura 3. Polinomio de McLaurin de cuarto grado.

acelera de manera constante, dadas las siguientes condiciones de posición, velocidad y aceleración respectivamente:  $s(0) = 1$ ,  $v(0) = 2$ ,  $a = 3$ . El problema se reduce a definir los tres puntos necesarios para formar un polinomio de grado 2, basados en las condiciones del problema, por medio de la acción de predicción de donde se encontrará la partícula en  $t = \frac{1}{n}$  y en  $t = \frac{2}{n}$ . La ecuación de movimiento quedará definida por la Expresión 6.

$$\lim_{n \rightarrow \infty} n^3 \cdot \begin{vmatrix} t^2 & t & s & 1 \\ 0^2 & 0 & s(0) & 1 \\ (\frac{1}{n})^2 & \frac{1}{n} & s(0) + v(0) \cdot \frac{1}{n} & 1 \\ (\frac{2}{n})^2 & \frac{2}{n} & [s(0) + v(0) \cdot \frac{1}{n}] + [v(0) + a \cdot \frac{1}{n}] \cdot \frac{1}{n} & 1 \end{vmatrix} = 0 \quad (6)$$

La salida de la evaluación de ésta en *Mathematica* con los datos ya definidos, se muestra en la Figura 4.

```
In[10]:= 2 - 2 s + 4 t + 3 t^2 == 0
Out[10]= 2 - 2 s + 4 t + 3 t^2 == 0
```

Figura 4. Ecuación de posición.

De donde es fácil expresar la función de posición de manera explícita:  $s = \frac{3}{2}t^2 + 2t + 1$ .

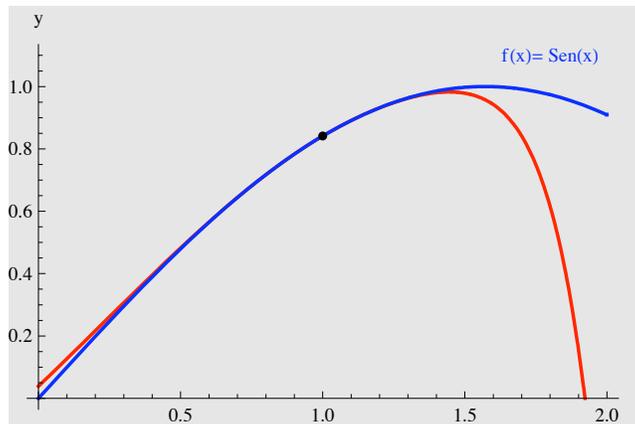
Como una extensión de las ideas plasmadas, se explora la posibilidad de aproximar localmente una función por medio de algunos elementos de series diferentes a las polinomiales, y que sin embargo, funcionan de manera similar a éstas, compartiendo de igual manera un cierto número de derivadas.

**Ejemplo 4.** Consideremos aproximar la función  $f(x) = \text{Sen}(x)$  en  $x = 1$  por medio de una expresión del tipo  $y = a_0e^x + a_1e^{2x} + a_2e^{3x} + a_3e^{4x}$ . Esto lo conseguimos por medio de la evaluación de la Expresión 7.

$$\lim_{n \rightarrow \infty} n^{10} \cdot \begin{vmatrix} e^{4x} & e^{3x} & e^{2x} & e^x & y & 1 \\ e^{4(1)} & e^{3(1)} & e^{2(1)} & e^{1(1)} & f(1) & 1 \\ e^{4(1+\frac{1}{n})} & e^{3(1+\frac{1}{n})} & e^{2(1+\frac{1}{n})} & e^{(1+\frac{1}{n})} & f(1+\frac{1}{n}) & 1 \\ e^{4(1-\frac{1}{n})} & e^{3(1-\frac{1}{n})} & e^{2(1-\frac{1}{n})} & e^{(1-\frac{1}{n})} & f(1-\frac{1}{n}) & 1 \\ e^{4(1+\frac{2}{n})} & e^{3(1+\frac{2}{n})} & e^{2(1+\frac{2}{n})} & e^{(1+\frac{2}{n})} & f(1+\frac{2}{n}) & 1 \\ e^{4(1-\frac{2}{n})} & e^{3(1-\frac{2}{n})} & e^{2(1-\frac{2}{n})} & e^{(1-\frac{2}{n})} & f(1-\frac{2}{n}) & 1 \end{vmatrix} = 0 \quad (7)$$

Al evaluarla, se obtuvo lo mostrado en la Figura 5.

$$-y + \frac{1}{12e^4} \left\{ -20e^4 \text{Cos}[1] + 30e^{3x} \text{Cos}[1] - 12e^{2+2x} \text{Cos}[1] + 2e^{1+3x} \text{Cos}[1] - 5e^4 \text{Sin}[1] - 5e^{4x} \text{Sin}[1] + 50e^{3+x} \text{Sin}[1] - 54e^{2+2x} \text{Sin}[1] + 26e^{1+3x} \text{Sin}[1] \right\} = 0$$



**Figura 5. Aproximación por suma de exponenciales.**

El lector puede comprobar que la expresión resultante comparte hasta la cuarta derivada con  $f(x)$  en  $x = 1$ , de la misma manera que lo haría un polinomio de Taylor de cuarto grado.

### 3. Conclusiones

Lo expuesto en este trabajo, ilustra que existen formas alternativas de abordar las series y polinomios de Taylor y McLaurin. En este trabajo se presentó una propuesta que se basa en la parte geométrica, ya que la manera tradicional de tratar estos temas, se centra en argumentos de corte algebraicos. Además, se muestra la posibilidad de resolver otro tipo de problemas, como los de movimiento uniformemente acelerado y el planteamiento de diferentes clases de series no polinomiales para aproximar funciones. Este artículo también brinda la posibilidad de percibir la naturaleza de los polinomios de Taylor: entre mayor es el grado de éste, mayor es el nivel de osculación y la aproximación local es cada vez mejor.

### 4. Referencias

- [1] Lehmann, C. (1996). *Geometría Analítica*. México. Limusa.
- [2] Weisstein, E. (2008). *Parábola*. Recuperado el 26 de enero de 2009 de MathWorld-A Wolfram Web Resource, <http://mathworld.wolfram.com/Parabola.html>.

# El uso de cubiertas en la demostración de teoremas del análisis real

Evaristo Trujillo Luque y Guadalupe Ávila Godoy  
Departamento de Matemáticas  
Universidad de Sonora  
evaristo.trujillol@alumno.uson.mx

## Resumen

*Este artículo es un reporte de tesis de licenciatura en el que se exponen dos conceptos especiales de cubiertas en el análisis matemático real y sus correspondientes lemas de cobertura. Para ilustrar la utilidad de estas ideas en el estudio de resultados básicos que son parte de los cursos introductorios de análisis real, se exponen en la tesis demostraciones no típicas de algunos teoremas fundamentales utilizando los lemas de cobertura mencionados, de las cuales incluimos cuatro en este reporte, tres relativas a propiedades de funciones continuas en un intervalo cerrado y la demostración del Teorema de Heine-Borel.*

## 1. Introducción

En este trabajo se presentan dos lemas de cobertura que pueden usarse para probar algunos resultados básicos del análisis. Para ejemplificar este uso, se aplican estos lemas en la prueba del Teorema de Heine-Borel y en la prueba de que la continuidad de una función en un intervalo cerrado implica que esta función es acotada, toma todos sus valores intermedios y es Riemman-integrable. PrimerO, en la Sección 2, se definen dos conceptos especiales de cubierta de un intervalo cerrado y acotado  $[a, b]$ , el de cubierta local y aditiva [1] y el de cubierta completa [2],[3]. Enseguida se prueban los lemas de cobertura mencionados. En el primero de ellos se muestra que el intervalo  $[a, b]$  es un elemento de cualquier cubierta local y aditiva de  $[a, b]$  y en el segundo se prueba que toda cubierta completa de  $[a, b]$  contiene una familia de subintervalos los cuales son una partición de  $[a, b]$ . Finalmente, en la Sección 3, se prueban los cuatro resultados mencionados aplicando cada uno de los dos lemas de cobertura.

## 2. Lemas de cobertura

**Definición 1.** Sea  $\mathcal{C}$  una familia de subconjuntos del intervalo  $[a, b]$ .

- Se dice que  $\mathcal{C}$  es local si cada punto  $x \in [a, b]$  tiene una vecindad, con respecto a la topología usual de  $\mathbb{R}$  relativa a  $[a, b]$ , la cual está en  $\mathcal{C}$ .
- Se dice que  $\mathcal{C}$  es aditiva si, siempre que  $C_1$  y  $C_2$  son miembros de  $\mathcal{C}$  tales que  $C_1 \cap C_2 \neq \emptyset$ , entonces  $C_1 \cup C_2 \in \mathcal{C}$ .

**Definición 2.** Una colección  $\mathcal{C}$  de subintervalos de  $[a, b]$  es una cubierta completa de  $[a, b]$  si a cada  $x \in [a, b]$  le corresponde un  $\delta(x) > 0$  tal que cualquier subintervalo cerrado de  $[a, b]$  que contiene a  $x$  y tiene longitud menor que  $\delta(x)$ , pertenece a  $\mathcal{C}$ .

**Lema 1.** Si  $\mathcal{C}$  es una familia local y aditiva de subintervalos cerrados de  $[a, b]$ , entonces  $[a, b] \in \mathcal{C}$ .

*Demostración.* Sea  $D = \{x : [a, x] \in \mathcal{C}\}$ . Para probar que  $[a, b] \in \mathcal{C}$  probaremos que  $b \in D$ . Como  $\mathcal{C}$  es local,  $a$  tiene una vecindad con respecto a la topología relativa de  $[a, b]$ , digamos  $[a, y]$  ( $y > a$ ), la cual está en  $\mathcal{C}$  y por lo tanto  $y \in D$ . Así, tenemos que  $D \neq \emptyset$ . Además,  $D$  está acotado ya que  $D \subseteq [a, b]$ . Entonces  $D$  tiene una máxima cota superior a la que llamaremos  $d$ . Ahora, para probar que  $b \in D$ , mostraremos: 1)  $d \in D$ , 2) si  $d \in D$  entonces  $d = b$ .

1) Como  $d \in [a, b]$ , existe  $[d', d''] \in \mathcal{C}$  con  $d' < d \leq d''$ . Ahora como  $d$  es el supremo de  $D$ , existe  $d_0 \in D$  tal que  $d' < d_0 < d \leq d''$ . Ya que  $\mathcal{C}$  es aditiva y  $[a, d_0] \cap [d', d''] \neq \emptyset$  se tiene que  $[a, d''] \in \mathcal{C}$  y por lo tanto  $d'' = d$ . Pero como  $[a, d''] \in \mathcal{C}$  se tiene que  $d'' \in D$ , esto es,  $d \in D$ .

2) Supongamos que  $d \in D$ . Si  $d \neq b$  entonces existe  $d' < d < d''$  tal que  $[d', d''] \in \mathcal{C}$ . Como  $[a, d] \in \mathcal{C}$  y  $[a, d] \cap [d', d''] \neq \emptyset$  se obtiene que  $[a, d''] \in \mathcal{C}$ , esto es,  $d'' \in D$  lo cual es una contradicción ya que  $d$  es el supremo de  $D$ .  $\square$

**Lema 2.** Si  $\mathcal{C}$  es una cubierta completa de  $[a, b]$ , entonces  $\mathcal{C}$  contiene una partición de  $[a, b]$ , i.e., existe  $x_0 = a < x_1 < x_2 < \dots < x_n = b$  tal que  $[x_{k-1}, x_k] \in \mathcal{C}$  para  $k = 1, \dots, n$ .

*Demostración.* Supongamos que  $\mathcal{C}$  no contiene partición alguna de  $[a, b]$ . Dividamos el intervalo  $[a, b]$  en dos intervalos de longitud  $(b-a)/2$ . Entonces se tiene que  $\mathcal{C}$  no contiene partición alguna de al menos uno de estos dos intervalos porque si  $\mathcal{C}$  tuviera una partición de cada uno de ellos entonces  $\mathcal{C}$  contendría una partición de  $[a, b]$ . Denotemos con  $J_1$  a este subintervalo de  $[a, b]$ . Dividamos ahora el intervalo  $J_1$  en dos intervalos de longitud  $(b-a)/2^2$  y denotemos con  $J_2$ , al subintervalo de  $J_1$  del cual  $\mathcal{C}$  no contiene partición alguna. La repetición de este procedimiento nos lleva a una sucesión  $\{J_n\}$  de subintervalos cerrados de  $[a, b]$  tales que  $J_n \supseteq J_{n+1}$  para cada  $n$ , además  $|J_n| \rightarrow 0$  y  $\mathcal{C}$  no contiene partición alguna de cada  $J_n$  ( $|J_n|$  denota la longitud del intervalo  $J_n$ ).

Por el Teorema de Intervalos Anidados se tiene que existe  $x$  en la intersección de las  $J_n$ 's. Como  $\mathcal{C}$  es una cubierta completa se tiene que existe  $\delta(x)$  tal que si  $I$  es un subintervalo cerrado de  $[a, b]$  con  $x \in I$  y  $|I| < \delta(x)$  entonces  $I \in \mathcal{C}$ . Así, ya que  $|J_n| \rightarrow 0$ , existe un natural  $N$  tal que  $|J_N| < \delta(x)$  y por lo tanto  $J_N \in \mathcal{C}$ , y de aquí se tiene que  $\mathcal{C}$  contiene una partición de  $J_N$ , lo cual es una contradicción.  $\square$

### 3. Uso de lemas de cobertura

En esta sección, para ejemplificar el uso de los lemas de cobertura, probaremos los Teoremas 3 y 4 utilizando el Lema 1 y los Teoremas 5 y 6 utilizando el Lema 2.

**Teorema 3.** Si  $f$  es continua en  $[a, b]$ , entonces  $f$  es acotada en  $[a, b]$ .

*Demostración.* Sea  $\mathcal{C} = \{I : I \text{ es subintervalo cerrado de } [a, b] \text{ y } f \text{ es acotada en } I\}$ .

Primero, probaremos que  $\mathcal{C}$  es local. Sea  $x \in [a, b]$ . Como  $f$  es continua en  $x$ , para  $\epsilon = 1$  existe  $[c, d]$  tal que  $x \in [c, d]$  y para toda  $y \in [c, d]$  se tiene que  $|f(x) - f(y)| < 1$ . De aquí que

$$|f(y)| < 1 + |f(x)| \text{ para toda } y \in [c, d].$$

Esto es,  $f$  está acotada en  $[c, d]$  y de aquí,  $[c, d] \in \mathcal{C}$ . Hemos encontrado una vecindad de  $x$  que pertenece a  $\mathcal{C}$  y por lo tanto  $\mathcal{C}$  es local. Ahora probaremos que  $\mathcal{C}$  es aditiva. Sean  $C_1, C_2 \in \mathcal{C}$ , es decir,  $f$  es acotada en  $C_1$  y en  $C_2$ . Entonces  $f$  es acotada en  $C_1 \cup C_2$ . Además, si  $C_1 \cap C_2 \neq \emptyset$ , entonces  $C_1 \cup C_2$  es un intervalo cerrado y por lo tanto  $C_1 \cup C_2 \in \mathcal{C}$ . Como  $\mathcal{C}$  es local y aditiva, aplicando el Lema 1 tenemos que  $[a, b] \in \mathcal{C}$ , esto es,  $f$  es acotada en  $[a, b]$ .  $\square$

**Teorema 4.** Sea  $f : [a, b] \rightarrow \mathbb{R}$  continua. Entonces  $f$  tiene la propiedad del valor intermedio en  $[a, b]$ . Esto es, si  $k$  es cualquier valor entre  $f(a)$  y  $f(b)$  (es decir,  $f(a) < k < f(b)$  o  $f(b) < k < f(a)$ ), entonces existe  $c \in (a, b)$  tal que  $f(c) = k$ .

*Demostración.* Es suficiente probar que si

$$f(a) < 0 < f(b) \tag{1}$$

entonces existe  $c \in (a, b)$  tal que  $f(c) = 0$ . La prueba se hará por contradicción. Supongamos que  $f$  no tiene ceros en  $[a, b]$ . Sea

$$\mathcal{C} = \{I : I \text{ es un subintervalo cerrado de } [a, b] \text{ y el signo de } f \text{ es constante sobre } I\}.$$

Mostraremos que  $[a, b] \in \mathcal{C}$  probando que  $\mathcal{C}$  es local y aditiva. Sea  $x \in [a, b]$ . Como  $f$  es continua en  $x$ , para  $\epsilon = |f(x)|$  existe una vecindad  $[c, d]$  de  $x$  tal que si  $y \in [c, d]$  se tiene  $|f(y) - f(x)| < |f(x)|$ . Esta desigualdad implica que  $f(x)$  y  $f(y)$  tienen el mismo signo, es decir, el signo de  $f$  es constante en  $[c, d]$ . De aquí que  $[c, d] \in \mathcal{C}$  y por tanto  $\mathcal{C}$  es local. Ahora, mostraremos que  $\mathcal{C}$  es aditiva. Sean  $C_1, C_2 \in \mathcal{C}$ , con  $C_1 \cap C_2 \neq \emptyset$ . Entonces  $f$  tiene signo constante en  $C_1$  y en  $C_2$ . Si  $x \in C_1 \cap C_2$ , entonces el signo de  $f$  sobre  $C_1$  y sobre  $C_2$  coincide con el de  $f(x)$ . De aquí que  $f$  tiene signo constante sobre el intervalo cerrado  $C_1 \cup C_2$ . Así,  $C_1 \cup C_2 \in \mathcal{C}$ . Tenemos por lo tanto que  $[a, b] \in \mathcal{C}$ , es decir,  $f$  tiene signo constante en  $[a, b]$ . Esto es,  $f(a)$  y  $f(b)$  son mayores que 0 ó  $f(a)$  y  $f(b)$  ambos son menores que 0 lo que contradice (1).  $\square$

**Teorema 5.** Si  $f$  es una función continua de variable real en  $[a, b]$ , entonces  $f$  es Riemann integrable en  $[a, b]$ .

*Demostración.* Sean  $\epsilon > 0$  y

$$\mathcal{C} = \left\{ I : f \text{ es acotada en } I \text{ y } \sup_I f - \inf_I f < \frac{\epsilon}{b-a} \right\}$$

Primero, probaremos que  $\mathcal{C}$  es una cubierta completa de  $[a, b]$ . Sea  $x$  cualquier punto en  $[a, b]$ . Dado que  $f$  es continua en  $x$ , entonces existe  $\delta(x) > 0$  tal que

$$|f(y) - f(x)| < \frac{\epsilon}{4(b-a)}, \quad \text{para toda } y \in (x - \delta(x), x + \delta(x)).$$

Sea  $I$  cualquier subintervalo cerrado de  $[a, b]$ , con  $x \in I$  y  $|I| < \delta(x)$ . Entonces claramente  $f$  es acotada en  $I$ . Sean  $u, v \in I$  tales que

$$\begin{aligned} f(u) &= \sup\{f(x) : x \in I\}, \\ f(v) &= \inf\{f(x) : x \in I\}. \end{aligned}$$

Entonces

$$\begin{aligned} |f(u) - f(v)| &\leq |f(u) - f(x)| + |f(x) - f(v)| \\ &\leq \frac{\epsilon}{4(b-a)} + \frac{\epsilon}{4(b-a)} \\ &= \frac{\epsilon}{2(b-a)} \\ &< \frac{\epsilon}{b-a}. \end{aligned}$$

Así que  $I \in \mathcal{C}$ . Por lo tanto  $\mathcal{C}$  es una cubierta completa de  $[a, b]$ . Ahora, por el Lema 2 se tiene que existe una partición de  $[a, b]$ ,  $P = \{a = x_0, x_1, \dots, x_n = b\}$ , tal que  $I_k = [x_{k-1}, x_k] \in \mathcal{C}$  para cada  $k$ . Como  $f$  es acotada en cada  $I_k$  se tiene que  $f$  es acotada en  $[a, b]$ . Tomando la diferencia entre la suma superior e inferior de  $f$  para la partición  $P$  y denotando con

$$M_k = \sup\{f(x) : x \in I_k\}, \quad m_k = \inf\{f(x) : x \in I_k\}$$

se obtiene

$$\begin{aligned} \overline{S}_p(f) - \underline{S}_p(f) &= \sum_{k=1}^n M_k |I_k| - \sum_{k=1}^n m_k |I_k| = \sum_{k=1}^n (M_k - m_k) |I_k| \\ &< \frac{\epsilon}{b-a} \sum_{k=1}^n |I_k| = \frac{\epsilon(b-a)}{b-a} = \epsilon. \end{aligned}$$

Esto es,  $f$  es Riemann integrable en  $[a, b]$ . □

**Teorema 6.** Toda cubierta abierta de  $[a, b]$  tiene una subcubierta finita.

*Demostración.* Sea  $\mathcal{G}$  una cubierta abierta de  $[a, b]$ . Sea

$$\mathcal{C} = \{I : I \text{ es un subintervalo cerrado de } [a, b], \quad I \subseteq G \text{ para algún } G \in \mathcal{G}\}.$$

Probaremos primero que  $\mathcal{C}$  es una cubierta completa. Sea  $x \in [a, b]$ . Entonces  $x \in G$  para algún  $G \in \mathcal{G}$ . Como  $G$  es abierto, existe  $\delta_x > 0$  tal que  $(x - \delta_x, x + \delta_x) \subseteq G$ . Sea  $I$  un subintervalo cerrado de  $[a, b]$  con  $x \in I$  y  $|I| < \delta_x$ . Entonces  $I \subseteq G$  y por lo tanto  $I \in \mathcal{C}$ , esto es,  $\mathcal{C}$  es completa. Ahora, por el Lema 2,  $\mathcal{C}$  contiene una partición  $I_1, I_2, \dots, I_n$  de  $[a, b]$ . Como cada  $I_k \in \mathcal{C}$ , se tiene que  $I_k \subseteq G_k$  para algún  $G_k \in \mathcal{G}$ . De aquí que  $\{G_1, G_2, \dots, G_n\}$  es una subcubierta finita de  $[a, b]$ . □

## 4. Referencias

- [1] Shanahan, Patrick (1972). *A Unified Proof of Several Basic Theorems of Real Analysis*. American Mathematical Monthly, 79, 8, 895–898.
- [2] Botsko, Michael W. (1987). *A Unified Treatment of Various Theorems in Elementary Analysis*. American Mathematical Monthly, 94, 5, 450–452.
- [3] Botsko, Michael W. (1989). *The Use of Full Covers in Real Analysis*. American Mathematical Monthly, 96, 5, 328–333.



# La técnica del punto conocido aplicada al ajuste del modelo normal

Gabriel Guillén Buendía<sup>1</sup> y Ana María Islas Cortes<sup>2</sup>

<sup>1</sup>IPN ESIME-UPA, Sta. Catarina, AZCAPOTZALCO, Cd. México

<sup>2</sup>IPN ESIT, Lindavista, GAM, Cd. México.

gguillen@ipn.mx

## Resumen

*En este trabajo se usaron histogramas de frecuencias correspondientes a las propiedades de alargamiento, fuerza, tenacidad y trabajo de rotura de tres hilados de poliéster de calibre 16.7 tex, procesados en una máquina texturizadora de doble torsión a tres niveles de temperatura 80, 90 y 100°C respectivamente y un nivel de torsión de 600 v/m. Los cuales se ajustaron al modelo normal usando la técnica del punto conocido y como métodos de comprobación el algoritmo de Guggenheim y el método Marquardt. La primera de las técnicas, permite la transformación lineal del modelo normal a través de un punto leído a voluntad sobre la curva envolvente; el grado de ajuste alcanzado depende de dicho punto. Sin embargo de todos los puntos escogidos se llega a valores convergentes a la optimización por regresión no lineal. Mientras que los resultados obtenidos usando el algoritmo de Guggenheim son similares al de la técnica que da pie al documento.*

## 1. Introducción

Los modelos estadísticos son de gran interés por su aplicación en diversas áreas del conocimiento, constantemente se buscan métodos numéricos que permitan un mayor ajuste de éstos a series de datos experimentales, que simplifiquen el método o sencillamente que muestren ventajas con respecto a los métodos ya existentes.

Este documento manifiesta el interés de aplicar una técnica estadística para ajustar el modelo normal a la campana envolvente de histogramas unimodal de moda centrada. Es un procedimiento de aproximación preliminar de las constantes numéricas del modelo normal que son convergentes a la optimización de éstas constantes por un programa iterativo para regresión no lineal, reduciendo el tiempo de máquina a través de un número mínimo de iteraciones que efectúe el programa. La expresión del modelo normal [1] es:

$$f = A \exp \left( -\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2 \right) \quad (1)$$

Las constantes numéricas del modelo anterior se determinan a través de la técnica del punto conocido y del algoritmo de Guggenheim [2] y el método Marquardt [3] que sirve como método de comprobación a los resultados obtenidos por la primera de ellas.

### 1.1. La técnica del punto conocido

Esta técnica en esencia propone, aproximar por inspección una campana al histograma unimodal de moda centrada y sobre ésta leer a voluntad un punto que se denomina punto conocido

$P_k(x_k, y_k)$  que cumple la expresión siguiente:

$$f_k = A \exp^{-\frac{1}{2} \left( \frac{x_k - \mu}{\sigma} \right)^2} \quad (2)$$

Dividiendo miembro a miembro la Ecuación 1 entre la Ecuación 2 y aplicando logaritmos neperianos resulta:

$$\ln(f/f_k) = -\frac{1}{2} [(x - \mu)/\sigma]^2 + \frac{1}{2} [(x_k - \mu)/\sigma]^2 \quad (3)$$

Desarrollando algebraicamente la Expresión 3 se llega a la Ecuación 4 que es la transformación lineal del modelo normal:

$$\ln(f/f_k)/(x - x_k) = -\frac{1}{2\sigma^2} x + \frac{1}{2\sigma^2} (2\mu - x_k) \quad (4)$$

De ésta ecuación resulta que graficando la variable  $x$  contra la columna de transformación  $\ln(f/f_k)/(x - x_k)$  se obtiene una recta, que es más acentuada mientras más se ajusta el histograma unimodal al modelo normal. Entonces de esta Ecuación 4 se calcula la varianza  $\sigma^2$  a partir de la pendiente de la recta  $m$ , y la media aritmética  $\mu$  de la intersección al origen  $b$ , como se indica:

$$m = -\frac{1}{2\sigma^2}; \quad b = \frac{1}{2\sigma^2} (2\mu - x_k) \quad (5)$$

El parámetro  $A$  del modelo normal se despeja de la Ecuación 1 y se calcula para todos los puntos experimentales y se obtiene el promedio de dicho parámetro:

$$\bar{A} = f / \exp^{-\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2} \quad (6)$$

### 1.2. El algoritmo de Guggenheim

Indica el algoritmo de Guggenheim [2] que tomando valores sobre la curva igualmente espaciados en las abscisas, se puede establecer dos subconjuntos de  $n/2$  puntos cada uno  $(x, f)$  y  $(x', f')$ . Donde los  $(x, f)$  son los primeros  $n/2$  puntos de la curva y  $(x', f')$  son los siguientes  $n/2$  puntos de la misma. Evidentemente, por haber tomado igualmente espaciados la diferencia es constante  $\tau$  y se denomina constante de desplazamiento de Guggenheim. Es decir, las expresiones siguientes son la esencia del algoritmo de Guggenheim:

$$\begin{aligned} x &\rightarrow f; \\ x + \tau &\rightarrow f' \end{aligned} \quad (7)$$

Aplicando las Expresiones 7 en la Ecuación normal 1, tenemos:

$$f = A \exp \left[ -\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2 \right] \quad (8a)$$

$$f' = A \exp \left[ -\frac{1}{2} \left( \frac{x+\tau-\mu}{\sigma} \right)^2 \right] \quad (8b)$$

Dividiendo miembro a miembro la Ecuación 8b entre la 8a y simplificando, llegamos a la transformación lineal del modelo normal, que se expresa:

$$\ln(f'/f) = -\frac{\tau}{\sigma^2}x + \frac{\tau}{2\sigma^2}(2\mu - \tau) \quad (9)$$

De ésta ecuación resulta que graficando la expresión  $x$  contra  $\ln(f'/f)$  se obtiene una recta, entonces de la intersección al origen  $b$  y pendiente  $m$  de la misma llegamos a la media aritmética y la varianza, como se señala a continuación:

$$m = -\frac{\tau}{\sigma^2}; \quad b = \frac{\tau}{2\sigma^2}(2\mu - \tau), \quad (10)$$

En tanto que el parámetro  $A$  se determina de la Expresión 6 y se calcula la media de todos los puntos experimentales.

## 2. Parte experimental

Para el desarrollo de este trabajo se realizó el ensayo de tracción a tres hilados de poliéster filamento continuo, bajo norma técnica vigente [4]. Los citados hilados son de calibre 16.7 tex (se entiende por un tex a la masa en gramos del hilado por cada 1000 metros de longitud) que fueron procesados a temperaturas de 80, 90 y 100°C en una máquina de doble torsión [5] GC96FF de la marca RPR y todos ellos a un nivel de torsión de 600 v/m. Se utilizó para el ensayo un dinamómetro universal de la marca Statimat M de Textechno ilustrando en la Figura 1, el cual es de uso común en la industria textil.



Figura 1. Dinamómetro universal Statimat M Textechno.

Se obtuvieron 250 valores experimentales de las propiedades mecánica [6], como el alargamiento (%), fuerza (cN), tenacidad (cN/tex) y trabajo (cN.cm) de cada uno de los hilados citados. Se procedió a construir histogramas de frecuencias, en la Tabla 1 aparecen los datos para las propiedades mecánicas arriba citadas correspondiente al hilado número 1, el cual es un filamento continuo de poliéster calibre 16.7 tex procesado a 600 v/m y temperatura de 80°C en una máquina de doble torsión.

**Tabla 1. Datos experimentales de propiedades mecánicas (hilado 1).**

Máquina texturizadora de doble torsión GC96FF de RPR Calibre 16.7 tex, torsión 600 v/m y temperatura 80°C							
Alargamiento		Fuerza		Tenacidad		Trabajo	
$A$ (%)	Frecuencia (No.)	$F$ (cN)	frecuencia (No.)	$\sigma$ (cN/tex)	frecuencia (No.)	$W$ (cN.cm)	frecuencia (No.)
23.125	3	384.25	6	34.719	5	2746.875	6
24.375	15	388.75	15	35.156	13	2940.625	26
25.625	40	393.25	13	35.594	16	3134.375	35
26.875	74	397.75	27	36.031	27	3328.125	71
28.125	96	402.25	51	36.469	51	3521.875	73
29.375	14	406.75	68	36.906	68	3715.625	37
30.625	1	411.25	60	37.344	55	3909.375	1
31.875	7	415.75	10	37.781	15	4103.125	1

En la Tabla 2 aparecen los datos para las propiedades de alargamiento, fuerza, tenacidad y trabajo de rotura correspondiente al hilado número 2, el cual es un filamento continuo de poliéster calibre 16.7 tex, procesado a 600 v/m y temperatura de 90°C en una máquina de doble torsión.

**Tabla 2. Datos experimentales de propiedades mecánicas (hilado 2) (Marquardt).**

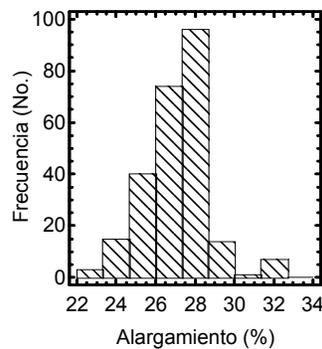
Máquina texturizadora de doble torsión GC96FF de RPR Calibre 16.7 tex, torsión 600 v/m y temperatura 90°C							
Alargamiento		Fuerza		Tenacidad		Trabajo	
$A$ (%)	frecuencia (No.)	$F$ (cN)	frecuencia (No.)	$\sigma$ (cN/tex)	frecuencia (No.)	$W$ (cN.cm)	frecuencia (No.)
25.513	1	383.625	8	34.688	3	2809.375	1
26.138	8	386.875	6	35.063	11	2928.125	7
26.763	17	390.125	13	35.438	34	3046.875	21
27.388	36	393.375	58	35.813	59	3165.625	24
28.013	68	396.625	85	36.188	96	3284.375	71
28.638	74	399.875	44	36.563	38	3403.125	85
29.263	36	403.125	28	36.938	8	3521.875	36
29.888	10	406.375	8	37.313	1	3640.625	5

La Tabla 3 muestra los datos de las propiedades de alargamiento, fuerza, tenacidad y trabajo de rotura del hilado número 3, que es un filamento continuo de poliéster calibre 16.7 tex, procesado a 600 v/m y temperatura de 100°C en una máquina de doble torsión.

**Tabla 3. Datos experimentales de propiedades mecánicas (hilado 3).**

Máquina texturizadora de doble torsión GC96FF de RPR Calibre 16.7 tex, torsión 600 v/m, y temperatura 100°C							
Alargamiento		Fuerza		Tenacidad		Trabajo	
$A$ (%)	frecuencia (No.)	$F$ (cN)	Frecuencia (No.)	$\sigma$ (cN/tex)	frecuencia (No.)	$W$ (cN.cm)	F (No.)
25.281	5	352.188	5	29.2	2	2546.875	10
25.844	18	356.563	26	29.6	10	2640.625	30
26.406	43	360.938	52	30	57	2734.375	25
26.969	49	365.313	65	30.4	51	2828.125	60
27.531	57	369.688	46	30.8	63	2921.875	63
28.094	50	374.063	37	31.2	46	3015.625	28
28.656	26	378.438	16	31.6	18	3109.375	26
29.219	2	382.813	3	32	3	3203.125	8

Para mostrar el ajuste del modelo normal sobre los histogramas de frecuencias de las Tablas anteriores, tomamos los datos del alargamiento a la rotura (%) del hilado número 1 según señala la Tabla 1, que conduce al histograma de frecuencias de la Figura 2.

**Figura 2. Histograma de frecuencias de datos de alargamiento a la rotura del hilado número 1.**

Se procedió a escoger un punto a voluntad sobre dicho histograma, al punto escogido le llamamos punto conocido, en este caso es:

$$P_k(25, 25)$$

De acuerdo a la expresión (4) al relacionar los datos de  $x$  contra  $\ln(f/f_k)/(x-x_k)$  se obtiene la transformación lineal del modelo normal; y de ésta se obtienen sus parámetros, es decir, la pendiente de la recta, la intersección al origen y el coeficiente de correlación, para el caso que estamos tratando éstos últimos son:

$$m = -0.18058488, \quad b = 5.31852423, \quad r = -0.94074111 \quad (11)$$

Sustituyendo los valores anteriores en (5) y (6), y despejando los valores numéricos de  $\mu$ ,  $\sigma$  y  $A$ , entonces se obtiene la expresión numérico-funcional de la normal:

$$f = 95.6444818 * \exp \left( \frac{1}{2} \left( \frac{x - 27.2258291}{1.66396546} \right)^2 \right) \quad (12)$$

Los parámetros anteriores son optimizados por el procedimiento Marquardt [3], llegando al siguiente resultado:

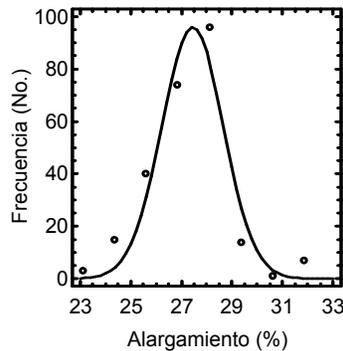
$$f = 95.8836 * \exp \left( \frac{1}{2} \left( \frac{x - 27.4397}{1.22774} \right)^2 \right) \quad (13)$$

El análisis de varianza correspondiente se muestra en la Tabla 4.

**Tabla 4. Análisis de varianza del modelo normal (Marquardt).**

Causa	Suma de Cuadrados	g. l.	Cuadrado Medio
Modelo	16002.8	2	5334.270
Residual	769.200	5	153.8400
Total	16772.0	8	
Total (corr.)	8959.50	7	
R <sup>2</sup>	91.4147		
R <sup>2</sup> (ajustado g.l)	87.9806		

El coeficiente de determinación es del 91.4147%, mientras que la varianza residual es de 153.84, ambos son significativos al 5% de confianza estadística, la bondad de ajuste se ilustra en la Figura 3.



**Figura 3. Bondad de ajuste logrado con el método Marquardt.**

En la Tabla siguiente se muestra los valores numéricos de los parámetros del modelo normal y su coeficiente de determinación (%), calculados a partir de los diversos puntos leídos a voluntad sobre el histograma ilustrado en la Figura 2.

**Tabla 5. Parámetros del modelo normal en relación al punto escogido.**

Técnica del punto conocido				
Punto escogido	$A$	$\mu$	$\sigma$	$R^2$ (%)
P <sub>1</sub> (24.5, 15.0)	57.7123458	27.2549568	1.94027416	11.560664
P <sub>2</sub> (25.0, 25.0)	95.6444818	27.2258291	1.66396546	79.755778
P <sub>3</sub> (26.0, 50.0)	110.107165	27.2415608	1.61277789	77.098054
P <sub>4</sub> (26.5, 65.0)	125.091810	27.2423028	1.57499449	71.937000
P <sub>5</sub> (27.0, 80.0)	145.188177	27.3455362	1.51619147	65.892480
P <sub>6</sub> (28.0, 95.0)	260.770298	27.2728250	1.41283393	38.537080
P <sub>7</sub> (29.0, 40.0)	1075.06543	26.8140212	1.34897459	8.1896430

La tabla indica, que la mayor bondad de ajuste del modelo normal logrado con la técnica del punto conocido, es cuando se trabajó con el punto P<sub>2</sub> (25.0, 25.0) alcanzando un coeficiente de determinación del 79.7557%, mientras que el menor ajuste se obtuvo con el punto P<sub>7</sub> (29.0, 40.00) con un coeficiente de determinación del 8.1896%. Sin embargo, todos los puntos escogidos son convergentes a la optimización por el método Marquardt. Usando ahora, el algoritmo de Guggenheim para evaluar las constantes del modelo normal al histograma ya indicado, se procedió a construir el arreglo rectangular de Guggenheim, donde se observa la constante de desplazamiento del mismo nombre, designada como  $\tau$ , y que su valor es de 5.0 en este caso.

**Tabla 6. Arreglo rectangular de Guggenheim.**

Alargamiento (%) $x$	Alargamiento (%) $x'$	Frecuencia $f$	Frecuencia $f'$	Constante $\tau$
23.125	28.125	3	96	5.00
24.375	29.375	15	14	5.00
25.625	30.625	40	1	5.00
26.875	31.875	74	7	5.00

Aplicando la Expresión 9 que es a la transformación lineal del modelo normal obtenido por el algoritmo de Guggenheim, y relacionando las columnas que indica con los datos de la Tabla 6, se llega a los valores de la intersección al origen y la pendiente de la recta:

$$m = -1.68732473, \quad b = 41.5205454, \quad r = -0.86937552 \quad (14)$$

Sustituyendo los valores anteriores en las Expresiones 10 y 6, nos conduce a la expresión numérico-funcional siguiente:

$$f = 88.6247761 * \exp \left[ \frac{1}{2} \left( \frac{x - 27.1073235}{1.72141533} \right)^2 \right] \quad (15)$$

En la tabla siguiente se comparan la bondad de ajuste del modelo normal ajustado al histograma de datos de alargamiento a la rotura de los hilados en estudio usando las diversa técnicas expuestas en el documento, en el caso de la técnica del punto conocido se seleccionó el máximo ajuste logrado con el punto P<sub>6</sub> (25.0, 25.0).

**Tabla 7. Valores numéricos del modelo normal usando diversas técnicas de ajuste numérico.**

Modelo normal				
Método	$A$	$\mu$	$\sigma$	$R^2$ (%)
Guggenheim	88.6247761	27.1073235	1.72141533	76.280785
Punto conocido*	95.6444818	27.2258291	1.66396546	79.755778
Marquardt	95.8836000	27.4397000	1.22774000	91.414700

A continuación en la Tabla 8 se muestran los resultados obtenidos al ajustar el modelo normal usando la técnica del punto conocido, el algoritmo de Guggenheim, y el método Marquardt a las propiedades de fuerza (cN), tenacidad (cN/tex) y trabajo (cN.cm) del hilado número 1.

**Tabla 8. Ajuste del modelo normal usando diversas técnicas (hilado 1).**

Máquina texturizadora de doble torsión GC96FF de RPR Calibre 16.7 tex, torsión 600 v/m y temperatura de 80°C					
Propiedad	Método de ajuste:	Parámetros de la normal			$R^2$ (%)
		$A$	$\mu$	$\sigma$	
Alargamiento (%) $P_k(25, 25)$	Guggenheim	88.6247761	27.1073235	1.72141533	76.280785
	Punto conocido	95.6444818	27.2258291	1.66396546	79.755778
	Marquardt	95.881500	27.439600	1.22779000	91.41460
Fuerza (cN) $P_k(404, 60)$	Guggenheim	53.159819	405.02114	9.2916838	48.37072
	Punto conocido	75.200006	406.11086	8.4241756	71.13616
	Marquardt	69.009000	406.23100	6.1877700	88.08200
Tenacidad (cN/tex) $P_k(37.2, 60)$	Guggenheim	53.073106	36.802619	0.9146561	68.45182
	Punto conocido	87.901690	36.729118	0.7278696	70.08422
	Marquardt	66.231700	36.844600	0.6408260	91.89520
Trabajo (%) $P_k(3600, 60)$	Guggenheim	83.052059	3325.5613	249.13403	82.58681
	Punto conocido	137.57671	3368.3876	221.22253	65.83243
	Marquardt	76.232400	3407.8600	251.0850	92.13270

En la Tabla 9 se resumen los resultados del ajuste del modelo normal a los histogramas de propiedades mecánicas del hilado número 2.

**Tabla 9. Ajuste del modelo normal usando diversas técnicas (hilado 2).**

Máquina texturizadora de doble torsión GC96FF de RPR Calibre 16.7 tex, torsión 600 v/m y temperatura de 90°C.					
Propiedad	Método de ajuste:	Parámetros de la normal			$R^2$ (%)
		$A$	$\mu$	$\sigma$	
Alargamiento (%) $P_k(27.2, 30)$	Guggenheim	67.755347	28.21452	0.932326	93.28628
	Punto conocido	58.895109	28.35823	1.030195	82.50857
	Marquardt	75.312400	28.31990	0.811228	98.10260
Fuerza (cN) $P_k(395, 70)$	Guggenheim	63.520690	396.7904	5.443609	71.25128
	Punto conocido	104.07948	396.8992	4.744728	71.59670
	Marquardt	79.245700	396.6650	3.858950	93.43080
Tenacidad (cN/tex) $P_k(36.8, 15)$	Guggenheim	79.046058	35.97017	0.471368	84.75456
	Punto conocido	93.279638	35.94751	0.450619	84.28212
	Marquardt	88.352800	36.07420	0.408130	94.34930
Trabajo (%) $P_k(3200, 35)$	Guggenheim	69.025701	3312.449	170.2886	72.78817
	Punto conocido	70.546015	3403.431	201.1363	61.04222
	Marquardt	86.685100	3358.510	127.4240	94.69210

En la Tabla 10 se resumen los resultados del ajuste del modelo normal a los histogramas de propiedades mecánicas del hilado número 3.

**Tabla 10. Ajuste del modelo normal usando diversas técnicas (hilado 3).**

Máquina texturizadora de doble torsión GC96FF de RPR Calibre 16.7 texc, torsión 600 v/m y temperatura de 100°C.					
Propiedad	Método de ajuste:	Parámetros de la normal			$R^2$ (%)
		$A$	$\mu$	$\sigma$	
Alargamiento (%) $P_k(26.8, 48)$	Guggenheim	71.35533	27.23118	0.828939	86.806434
	Punto conocido	56.61931	27.19311	0.934893	86.159633
	Marquardt	59.17360	27.35450	0.985521	94.928200
Fuerza (cN) $P_k(360, 45)$	Guggenheim	67.19573	366.4439	6.630232	94.220434
	Punto conocido	73.25519	366.2167	6.426366	92.332733
	Marquardt	62.15830	365.9130	7.185990	96.179900
Tenacidad (cN/tex) $P_k(30.3, 50)$	Guggenheim	74.14017	30.64599	0.545236	85.988586
	Punto conocido	59.60628	30.58253	0.603040	80.784811
	Marquardt	64.83650	30.60590	0.638729	89.590500
Trabajo (%) $P_k(2800, 50)$	Guggenheim	55.16861	2870.825	175.1637	73.740613
	Punto conocido	65.26364	2901.523	163.5170	74.286635
	Marquardt	57.27250	2874.950	163.9150	80.659100

### 3. Conclusiones

Del presente trabajo se puede concluir lo siguiente:

1. La técnica del punto conocido permite determinar las constantes numéricas del modelo normal, aunque la bondad de ajuste depende del punto escogido, sin embargo de todos ellos se calculan estimadores convergentes al procedimiento iterativo Marquardt.
2. Se usa también el algoritmo de Guggenheim para determinar los parámetros del modelo normal, y sus resultados son cercanos a la técnica del punto conocido. Cabe señalar que dicho fue propuesto originalmente en 1926, y aún continúa vigente.
3. Las técnicas de transformación lineal de modelo no lineales, ha sido tratado constantemente en estadística, sin embargo su utilidad docente se incrementa al introducir a jóvenes estudiantes a este campo del conocimiento, simplificando ésta tarea.
4. Los datos experimentales usados en este trabajo corresponden a la ingeniería textil, pero la aplicación de la técnica matemática expuesta se extiende a cualquier área del conocimiento.

### 4. Referencias

- [1] Montgomery, D. C., Peck, E. A. & Vining, G. G.; (2002), "Introducción al análisis de regresión lineal", Primera Edición, Ed. CECSA, México, 373-379.

- [2] Guillén, G., Islas, A. M., Guillén, E.; (2005), “*Ajuste a modelos no lineales sin el uso de computadoras*“, Revista ContactoS, México, No. 58, 11-20.
- [3] Marquardt, D. W.; (1963), “*An Algorithm for Least-Squares Estimation of Nonlinear Parameters*“, Journal for the Society of Industrial and Applied Mathematics, 11: 431-41.
- [4] ASTM Standard; (1992), D:2256-90, “*Annual Books of ASTM Standars*“, Section 7 (textiles), Volumen 07.01, 596.  
Información Técnica firma RPR; (1995).
- [5] Morton, W. E & Hearle, J. W. S.; (1997); “*Physical Properties of Textile Fibres*“, The Textile Institute, England. 266-275.

# Notas para una interpretación del método de Herón de Alejandría

Francisco C. García Durán  
Departamento de Matemáticas  
Universidad de Sonora  
fcgarcia@gauss.mat.uson.mx

## Resumen

*Históricamente, hasta la fecha, la primera descripción explícita de un método para calcular una aproximación de la raíz cuadrada de un natural que no es cuadrado perfecto es debida a Herón de Alejandría, en su obra “Metrica” del primer siglo de nuestra era. En este trabajo se da una interpretación de dicho método de aproximación. También se presentan las ideas filosóficas, y matemáticas de la civilización griega clásica que sustentan la plausibilidad de la interpretación dada, en particular las referidas a la teoría de medias pitagórica. Con lo anterior, este método de cálculo numérico se muestra como un inesperado ejemplo paradigmático de la extendida afirmación sobre la gran influencia filosófica en la concepción de las matemáticas de los antiguos griegos.*

## 1. Introducción

Es comúnmente aceptado por la inmensa mayoría de los historiadores de las Matemáticas que su carácter de ciencia deductiva es la herencia inobjetable de la civilización griega del período clásico del 600 a.C. al 300 a.C., aproximadamente. El epitome de estas matemáticas se encuentra en *Los Elementos* de Euclides (aproximadamente 325-265 a.C.) que nos muestran los avances alcanzados por esta civilización durante el antedicho período, y que constituyó por dos milenios el paradigma de trabajo matemático, con sus rasgos distintivos de basarse en la axiomatización y el uso de la demostración como criterio de validación de las afirmaciones matemáticas (o teoremas) establecidas. Estos logros matemáticos fueron posibles gracias a la creencia griega del poder de la razón y del razonamiento para la comprensión del mundo en todas sus facetas. Con la aplicación del raciocinio al estudio de las matemáticas los griegos dejaron atrás la concepción instrumental, utilitaria y empírica de las matemáticas típica de las predecesoras culturas egipcias y babilónicas, dando origen a una concepción abstracta, formal y teórica que sobrevive hasta nuestros días, y que llevada a sus últimas consecuencias le permite a Jacobi (1804-1851) afirmar a mediados del siglo XIX que “las matemáticas sólo sirven al honor del espíritu del hombre”. Ciertamente los griegos clásicos no llegaron a los extremos de Jacobi ya que, recordemos, ellos consideraban que el estudio de las matemáticas era indispensable para el desarrollo de las capacidades intelectuales del hombre, en particular, del hombre de estado (Véase [1]). Además de esta función educativa fundamental, para la cultura griega clásica, debido a la influencia fundadora de Pitágoras de Samos (569-475 a.C., aproximadamente) y su escuela, las matemáticas eran la ciencia idónea para la interpretación intelectual del mundo como un sistema racional, ordenado y armonioso, para nada caótico, caprichoso y arbitrario. Esta aplicación de las

matemáticas fue posible gracias al punto de vista filosófico pitagórico de que todo en la naturaleza era expresable en términos de números enteros positivos y de relaciones entre ellos: *todo es número*.

Como es de sobra conocido, gracias a sus mismos estudios esta concepción filosófica pitagórica recibió un duro golpe adverso con el descubrimiento de la inconmensurabilidad del lado y la diagonal del cuadrado: irracionalidad de  $\sqrt{2}$  en términos modernos. Una muestra tanto del genio griego, como del acierto de su confianza en la racionalidad la hallamos en la manera en que absorbieron el impacto: la cuestión de la inconmensurabilidad la sacaron del dominio de la teoría de números o aritmética griega y la llevaron a los terrenos de la geometría. En esta última rama enfrentaron la cuestión elaborando una teoría de proporciones, incluyendo *magnitudes* inconmensurables, que daba cabal cuenta matemática del tema; labor achacada por los historiadores a Eudoxo de Cnidos ( ca. 390 a.C.- ca. 337 a.C.) y plasmada por Euclides en el libro X de su famosa obra. La pertinencia del desarrollo de tal teoría se derivó del descubrimiento de que  $\sqrt{2}$  no era el único de su tipo, sino que también eran inconmensurables  $\sqrt{3}$ ,  $\sqrt{5}$ ,  $\sqrt{6}$ ,  $\sqrt{7}$ ,  $\sqrt{8}$ ,  $\sqrt{10}$ ,  $\sqrt{11}$ ,  $\sqrt{12}$ ,  $\sqrt{13}$ ,  $\sqrt{14}$ ,  $\sqrt{15}$ , y  $\sqrt{17}$ , como nos informa el contemporáneo de Eudoxo, Platón (428 a.C.-347 a.C.), cuando hace decir a Teetetes:

Teodoro nos enseñaba algún cálculo sobre las raíces de los números, demostrándonos que las de tres y de cinco no son conmensurables en longitud con la de uno, y en seguida continuó así hasta la de diez y siete, en la que se detuvo. Juzgando, pues, que las raíces eran infinitas en número, nos vino al pensamiento intentar el comprenderlas bajo un solo nombre, que conviene a todas [2].

Así, para cuando Euclides publica sus *Elementos*, en los albores del período alejandrino o helenístico (cerca del 300 a.C.-600 d.C.), la teoría sobre las magnitudes inconmensurables ya formaba parte del corpus matemático vigente y aceptado por la civilización griega. Recordemos que esta nueva época histórica inicia con las conquistas de Alejandro Magno, quien creó un vasto imperio con una recién fundada capital en Egipto, Alejandría, y quien deliberadamente se propuso fusionar la cultura griega con la del Cercano Oriente. Esto trajo como consecuencia que la concepción teórica de los griegos clásicos se mezclase con la visión práctica de los babilonios y egipcios para producir un nuevo pensamiento matemático que aunque de espíritu predominante geométrico (teórico) no dejaba de lado las aplicaciones prácticas de sus hallazgos teóricos. Ejemplos de esta nueva actitud en la matemática alejandrina los hallamos en la obra de los matemáticos griegos Arquímedes de Siracusa (ca. 287-212 a.C ) y Apolonio de Perga (ca. 262-ca.200 a.C). Misma actitud que se halla en aquéllos que pretendían, al estilo clásico, estudiar matemáticamente la naturaleza y que al hacerlo afrontaron la resolución de problemas que requerían el manejo operatorio de datos cuantitativos, como fueron los casos de los astrónomos Hiparco de Nicea ( ca. 180- ca. 125 a.C.) y Claudio Ptolomeo ( ca. 85- ca. 165 d.C.) Más aún, para la segunda mitad del período helenístico se deja atrás la tendencia clásica “hacia una ciencia matemática pura, estudiada por sí misma, en la que la búsqueda de la verdad es lo único que debe contar” [3], para orientarse a lo que puede considerarse matemáticas aplicadas, produciéndose trabajos como el de Diofanto de Alejandría (ca. 200- ca. 284 d.C.) Inevitablemente, al aplicar las matemáticas surgió el problema de calcular datos cuantitativos que no correspondían a números racionales, como fue el caso para  $\pi$ , por lo que en la práctica hubo que conformarse con *aproximaciones* calculadas con técnicas derivadas del conocimiento teórico y no del conocimiento empírico. Por supuesto que entre estos casos se dio la aproximación de raíces

cuadradas de números que no eran cuadrados perfectos, es decir, irracionales, y el ejemplo por excelencia lo constituye el método diseñado por Herón de Alejandría (ca. 10- ca. 75 d.C.), del cual nos ocupamos en este trabajo. Este método es el *primero* de su tipo *explícitamente declarado* del que a la fecha se tenga noticia, ya que si bien hay registro históricos de raíces cuadradas calculadas (Tablilla Yale YBC7289) no se conoce *cómo* se obtuvieron.

## 2. Método de Herón para el cálculo de raíz cuadrada

Lo poco que se conoce de Herón lo perfila como lo que hoy llamaríamos matemático aplicado ya que sus numerosos y variados trabajos constituyen toda una enciclopedia de las matemáticas y la física contemporáneas. Ella puede dividirse en dos partes: una de índole geométrica donde aborda principalmente problemas de medidas de longitudes, áreas y volúmenes en campos diversos como la óptica, la mecánica y la geodesia; la otra parte inserta sobre todo en la mecánica se centra en la descripción de decenas de aparatos mecánicos y juguetes ingeniosos, tales como, un sifón, un odómetro, un reloj de agua, un lanzallamas, un dispositivo de apertura de puertas automático por medio de fuego para un templo, un órgano hidráulico, y máquinas neumáticas automáticas para la guerra, levantar cargas, etc. Su tratado principal sobre la medición, *Métrica*, se compone de tres libros, y en todos ellos se da el respaldo teórico de las fórmulas empleadas por lo que no son simples enumeraciones de reglas a emplear; además, no emplea en sus ejemplos medidas sino números sin unidades de medición. En el libro I Herón presenta una fórmula, demostrada matemáticamente, para calcular el área de un triángulo a partir de sus tres lados, que hoy lleva su nombre aunque ya se sabe que es debida a Arquímedes. La llamada fórmula de Herón es  $A^2 = s(s-a)(s-b)(s-c)$ , donde  $s$  es el semiperímetro, es decir,  $s = (a+b+c)/2$ . Desde el punto de vista epistemológico esta fórmula es interesante pues requiere efectuar una multiplicación de *cuatro* longitudes, lo cual para los matemáticos griegos clásicos no tenía sentido geométrico, y por ende, no tenía sentido matemático. Tomando en cuenta la motivación de aplicabilidad práctica de los trabajos de Herón parece razonable suponer que tomaba las longitudes como números, quizá con un sentido logístico más que aritmético. Además, siendo una fórmula aplicable para *cualquier* triángulo el resultado de la multiplicación no tenía que ser siempre un cuadrado perfecto (número equilátero) sino que podía no serlo (número oblongo). Más aún, para obtener el área del triángulo,  $A$ , todavía hay que calcular la *raíz cuadrada* del resultado de la multiplicación, y en el caso de que éste fuese oblongo ¿cómo hacerlo? Ya que su libro estaba escrito para su uso en aplicaciones Herón *debía* proporcionar una técnica para aproximar la raíz cuadrada en esta última eventualidad, porque de otra forma no hubiera tenido utilidad práctica la fórmula dada.

La necesidad de proporcionar adicionalmente un método de aproximación debe haber sido más que evidente para Herón ya que, por un lado, para ilustrar el uso de la fórmula del área propone un triángulo de lados 7, 8, y 9, lo cual lo lleva al resultado oblongo 720, y a la necesidad de calcular  $\sqrt{720}$  para obtener el área  $A$ . Y por otro lado, a continuación inmediata presenta su método de aproximación para la raíz cuadrada ilustrándolo con la obtención de una aproximación precisamente para  $\sqrt{720}$ . En una de sus versiones inglesas este párrafo es:

Since 720 has not its side rational, we can obtain its side within a very small difference as follows. Since the next succeeding square number is 729, which has 27 for its side, divide 720 by 27. This gives  $26\frac{2}{3}$ . Add 27 to this, making  $53\frac{2}{3}$ , and take half this or

$26 \frac{5}{6}$ . The side of 720 will therefore be very nearly  $26 \frac{5}{6}$ . In fact, if we multiply  $26 \frac{5}{6}$  by itself, the product is  $720 \frac{1}{36}$ , so the difference in the square is  $\frac{1}{36}$ . If we desire to make the difference smaller still than  $\frac{1}{36}$ , we shall take  $720 \frac{1}{36}$  instead of 729, and by proceeding in the same way we shall find the resulting difference much less than  $\frac{1}{36}$  [4].

Con notación aritmética actual, tanto la ilustración de Herón, como su generalización para aproximar la raíz cuadrada de un entero positivo que no es cuadrado perfecto, se dan en la Tabla 1. Analizando el párrafo de Herón con la terminología y desde el punto de vista matemáticos comunes en los cursos universitarios de introducción al Análisis Numérico podemos observar algunos aspectos notables en la técnica ilustrada. Para empezar, ¿qué problema está intentando resolver Herón? Observemos que el sabio griego señala que 720 no tiene raíz cuadrada racional, pero que se puede obtener un racional con una muy pequeña diferencia con  $\sqrt{720}$ ; es decir, aunque  $\sqrt{720}$  es *irracional* se puede obtener una *aproximación* de él. Además, Herón inicia hablando de 720, es decir, del *cuadrado* de la raíz irracional; y una vez obtenido el racional  $26 \frac{5}{6}$  afirma que será una aproximación, pasando a apoyar su afirmación de cercanía dando el cuadrado  $720 \frac{1}{36}$  y señalando que la diferencia en *cuadrado* es  $\frac{1}{36}$ ; es decir, la bondad de la aproximación calculada se remite a una comparación en los *cuadrados*. En términos modernos, Herón está tratando con la ecuación  $x^2 = 720$ , o equivalentemente,  $x^2 - 720 = 0$ . Así, que la técnica de Herón es un método para la solución aproximada de ecuaciones de la forma  $x^2 - q = 0$ , para  $q$  entero positivo no cuadrado perfecto. Un problema elemental en los cursos de métodos numéricos donde se disponen de varios como Bisección, Falsa Posición, Secante, y Newton-Raphson, para resolverlo en forma aproximada.

Tabla 1. Interpretación del método de Herón.

Herón: $\sqrt{720}$	Resultados de Herón	Para $\sqrt{q}$ , $q > 0$ , no cuadrado perfecto	
$720 < 729 (= 27^2)$		Sea $p^2$ el primer cuadrado perfecto mayor que $q$ : $q < p^2$	
27		Tome $p$	Paso 1
$720/27$	$26 \frac{2}{3}$	Calcule $q/p$	Paso 2
$(720/27 + 27)$	$53 \frac{2}{3}$	Añada el resultado del paso anterior a $p$	Paso 3
$(1/2)(720/27 + 27)$	$26 \frac{5}{6}$	Calcule la mitad del resultado del paso anterior	Paso 4
$\sqrt{720}$ es muy cercano a $26 \frac{5}{6}$		Tome el resultado $r$ del paso anterior como aproximación de $\sqrt{q}$ : $\sqrt{q} \approx r$	Paso 5
En lugar de 729 tomar $720 \frac{1}{36} (= (26 \frac{5}{6})^2)$ y repetir método		Repetir los pasos 1 al 5 con $r$ en el lugar de $p$	Paso 6

### 3. Interpretación analítica numérica del método de Herón

Continuando con nuestro análisis, notemos que si denotamos  $A = 26 \frac{2}{3}$ ,  $B = 27$ , y  $x = 26 \frac{5}{6}$ , entonces la aproximación  $x$  es la media aritmética de  $A$  y  $B$ , ya que  $x = (A + B)/2$ . En términos geométricos, dado que  $A < B$ ,  $x$  es el punto medio del intervalo  $[A, B]$ , por lo que Herón efectuado una *bisección* de un segmento lineal. Además, aunque él no lo haya hecho, plantea la posibilidad de conseguir una mejor aproximación que la ya calculada *repitiendo* el procedimiento con  $x$  en el papel de  $B$ ; es decir, está proponiendo un método *iterativo* para aproximar la raíz

cuadrada. Al hacerse  $B = x$  tendremos  $A = 720/x$ , lo que nos da un nuevo intervalo  $[A, B]$ , mismo que se deberá *bisectar* para calcular la nueva aproximación mejorada (ya que será más cercana que la primera). Notemos que el nuevo intervalo es más pequeño que el inicial pues es un subintervalo de este último, y más aún, ¡el nuevo intervalo es más pequeño que la mitad del intervalo inicial! Si continuamos el proceso de Herón con este nuevo intervalo, tomando a la nueva aproximación como la siguiente  $B$  tendremos otro nuevo intervalo contenido en el intervalo inmediato anterior y de longitud menor que la mitad de este último, y así sucesivamente.

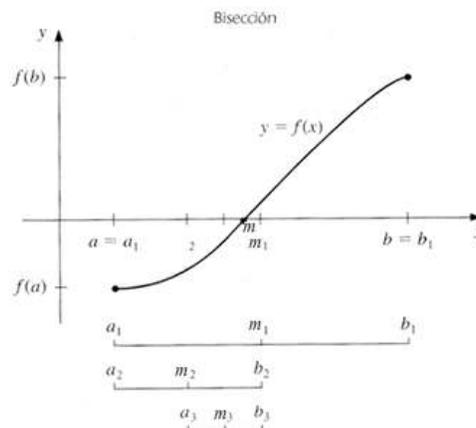
Aprovechando la concisión y claridad de la notación usual en Análisis Numérico podemos expresar de manera general el método de Herón como sigue,

Sea el intervalo inicial  $I_1 = [A_1, B_1]$  tal que  $A_1 = q/p$ ,  $B_1 = p$ , con  $q < p^2$ . A partir de  $I_1$  genérese la sucesión de aproximaciones  $\{x_n\}_{n=1}^{\infty}$ , cuyo término  $n$ -ésimo,  $n > 1$ , viene dado por

$$x_n = (A_n + B_n)/2$$

Con el intervalo  $n$ -ésimo  $I_n = [A_n, B_n]$  de extremos:  $A_n = q/x_{n-1}$  y  $B_n = x_{n-1}$ .

Donde este método sería específico para ecuaciones no lineales de la forma  $x^2 - q = 0$ , con  $q$  entero positivo no cuadrado perfecto. La anterior presentación hace evidente que Herón ha diseñado un método numérico que es una variante del tradicional método de Bisección (véase Figura 1) de los cursos universitarios introductorios en Análisis Numérico.



**Figura 1. Método de Bisección actual.**

Señalemos que la variación consiste en el *dinamismo* de los extremos del intervalo  $n$ -ésimo,  $I_n$ , en particular del extremo izquierdo  $A_n$ , ya que el método de Bisección tradicional aplicado al tipo específico de ecuaciones mencionado, y con el mismo intervalo inicial  $I_1$ , produce intervalos tales que  $A_n = A_{n-1}$  y  $B_n = x_{n-1}$ , ¡con el extremo izquierdo fijo! Unos interesantes ejercicios matemáticos serían demostrar, tanto la última afirmación sobre lo *estático* de  $A_n$  en Bisección tradicional, como el hecho de que  $A_n > A_{n-1}$ , para  $n > 1$ , en el método de Herón, y por tanto que  $\text{longitud } I_n < (1/2)\text{longitud } I_{n-1}$ , para  $n > 1$ , siempre y cuando se tenga  $A_n < B_n$ . Resolver los ejercicios no requiere más que el conocimiento algebraico del nivel bachillerato por lo que por razones de espacio no lo hacemos en este texto. Observemos que Herón, si bien demuestra

matemáticamente todas las fórmulas que propone en su *Métrica*, no se toma la molestia ni considera necesario demostrar y/o justificar matemáticamente *por qué* su técnica efectivamente trabaja produciendo la aproximación buscada. Tarea esta última indispensable desde el punto de vista del Análisis Numérico moderno, así que surge la interrogante, ¿porqué no sería necesario explicitarle al lector contemporáneo de la obra la demostración matemática, o al menos la justificación matemática de la efectividad del método propuesto?

#### 4. Algunos influyentes conocimientos matemáticos de los pitagóricos

Aunque Tales de Mileto (ca. 624-ca. 548 a.C) sea el primer hombre conocido a quien se le atribuye la demostración de resultados matemáticos, la asignación histórica del título de *fundador* de las matemáticas griegas corresponde a Pitágoras de Samos, gracias a los estudios matemáticos realizados por él y por los miembros de la secta que fundó. Los puntos de vista y los descubrimientos pitagóricos dieron el tono y marcaron la pauta para las investigaciones en matemáticas realizadas durante el esplendor de la civilización griega antigua. Entre los resultados de Pitágoras y los primeros pitagóricos se halla la elaboración de una teoría de las proporciones válida únicamente para magnitudes conmensurables, la cual una vez descubierta la existencia de magnitudes inconmensurables no pudo dar cuenta matemática de ellas por lo que posteriormente se desarrolló una nueva teoría que superaba esta limitante; labor atribuida a Eudoxo y aparecida en el *Libro V* de *Los Elementos* de Euclides. Asociada a la inicial teoría de las proporciones se hallaba una teoría aritmética de las *medias*, de las cuales los pitagóricos llegaron a distinguir diez, entre ellas: la media  $m$  aritmética, la media  $g$  geométrica, y la media  $h$  subcontraria, después llamada armónica. Estas tres últimas, para enteros positivos  $a$  y  $b$ ,  $a \leq b$ , se definían como las medias que cumplían, respectivamente, con:

$$\frac{m - a}{b - m} = \frac{a}{a} \qquad \frac{g - a}{b - g} = \frac{a}{g} \qquad \frac{h - a}{b - h} = \frac{a}{b}$$

Definiciones que se da por aceptado conoció Pitágoras durante sus viajes iniciáticos por el Medio Oriente, antes de establecer su secta en Crotona. Los pitagóricos también sabían que si  $a$  y  $b$  no eran iguales entonces  $a < h < g < m < b$ .

Es de sobra conocido que para el Pitagorismo la educación debía ser matemática (en el sentido general etimológico del término *mathema*), y que para ello era indispensable el estudio de cuatro ciencias: geometría, aritmética, armonía y astronomía. Principio didáctico basado en el misticismo numérico de la doctrina pitagórica, respaldado fundamentalmente este último en dos resultados matemáticos también aprendidos por Pitágoras en sus viajes iniciáticos. Estas verdades matemáticas eran, primero, el cumplimiento por la media  $g$  geométrica de la proporción,

$$\frac{a}{g} = \frac{g}{b}$$

Y más importante aún, el cumplimiento por las medias  $m$  y  $h$  de la proporción:

$$\frac{a}{h} = \frac{m}{b}$$

Proporción calificada como *dorada*, *más perfecta*, o *musical*, y que el mito dice que Pitágoras descubrió empíricamente que se cumplía en la música. Algunos autores dicen que lo hizo con ayuda de un monocordio, es decir, un instrumento donde una cuerda tensa vibra emitiendo un sonido agradable de un tono, y marcado con 12 subdivisiones iguales de arriba abajo. Así, al

oprimir en el 6 y pulsar la parte inferior de la cuerda se obtiene un *octavo* del tono original; al oprimir en el 8 y pulsar la cuerda inferior se obtiene un *quinto* del tono; y al oprimir en 9 y pulsar se obtiene una *cuarta* del tono original. Tomando los números 6, 8, 9, y 12, es fácil comprobar que 8 es la media armónica y 9 la media aritmética entre 6 y 12, y que ¡cumplen estos cuatro números con la proporción más perfecta! Además, las razones de los números 6, 8, 9, y 12, con 12, son las mismas que hay entre  $1/2$ ,  $2/3$ ,  $3/4$ , y 1, con 1, respectivamente, y también,  $2/3$  y  $3/4$  son medias armónica y aritmética, en ese orden, entre  $1/2$  y 1, cumpliendo asimismo estos números con la proporción dorada. Estos hechos aritméticos filosóficamente eran importantes pues las razones  $1/2$ ,  $2/3$ ,  $3/4$ , son las más sencillas que se pueden construir con la sagrada *Tetractys*: unidad y los números 2, 3, y 4, cuaterna básica para la geometría del Cosmos pitagórico.

El descubrimiento del cumplimiento de la proporción más perfecta en la música llevó a su extrapolación a la astronomía dando pie al mito pitagórico de *la música de las esferas*: los cuerpos que se movían en el espacio emitían sonidos musicales en función de su velocidad de rotación y distancia a la tierra. Más aún, las distancias entre los planetas estaban armónicamente determinadas ya que el Demiurgo al hacer el alma del Cosmos uso dos progresiones geométricas, una con razón 2 y la otra con 3, tal como describe Platón en el *Timeo*, donde además,

After this he filled up the double intervals and the triple [i.e. between 1, 3, 9, 27] cutting off yet other portions from the mixture and placing them in the intervals, so that in each interval there were two kinds of means, the one exceeding and exceeded by equal parts of its extremes, the other being that kind of mean which exceeds and is exceeded by an equal number [5].

Es decir, el intervalo entre dos términos consecutivos de la progresión es partido en tres partes con la media armónica y la media aritmética, y con ello logrando la armonía tan apreciada por los pitagóricos. En suma, la verdad matemática expresada por la proporción más perfecta era una ley de la música y de la naturaleza, y su cumplimiento dotaba de belleza y armonía tanto al Cosmos como al arte musical. Más aún, el hablar de belleza y armonía era el tema total de las matemáticas griegas como bien expresa Aristóteles cuando escribe:

Incurren en un error los que pretenden que las ciencias matemáticas no hablan ni de lo bello ni del bien. De lo bello es de lo que principalmente hablan, y lo bello es lo que demuestran. No hay razón para decir que no hablan de lo bello porque no lo nombren; mas indican sus efectos y sus relaciones. ¿No son las más imponentes formas de lo bello el orden, la simetría y la limitación? Pues esto es en lo que principalmente hacen resaltar las ciencias matemáticas [6].

Observemos, adicionalmente, que el concepto de media aritmética está en consonancia con la anterior concepción sobre las matemáticas ya que corresponde geoméricamente a la acción de *partir en dos partes iguales* un segmento lineal. Acción esta última la más elemental de efectuarse y que produce armonía; la bipartición es fácil, pero no siempre produce *armonía*, y la armonía más sencilla la proporciona la *simetría*. No en balde, la geometría griega abunda en resultados y procedimientos que involucran la acción de partir en dos partes iguales, lo que hoy llamamos *bisectar*; recordemos, por ejemplo, que uno de los primeros resultados geométricos *demostrados* (por Thales) fue el de la bisección de la circunferencia por su diámetro. Con la

acción de bisección se obtienen dos magnitudes que están entre sí en la razón 1: 1, y cada una de ellas está con el todo original en la razón 1:2. Razón esta última entre la unidad 1 y el primer número 2, la razón más elemental para considerar aritméticamente cuando se busca que el mundo sea armonioso y bello, tal como pretendían los pitagóricos.

## 5. El método de Herón a la luz del discurso matemático de los griegos clásicos

Con los elementos del discurso matemático griego establecidos en la sección anterior basta para encontrar motivos para la ausencia de demostración y justificación matemáticas en la presentación de Herón de su método de aproximación. Señalemos, para empezar, que Herón encierra inicialmente el número oblongo  $q$  en un intervalo tal que la media geométrica de sus extremos al cuadrado,  $g^2$ , es dicho número:  $g^2 = A_1 * B_1 = (q/p) * p = q$ ; misma situación que se mantiene en cada paso  $n$  con el intervalo  $[A_n, B_n]$ , por lo que podemos afirmar que Herón mantiene el cumplimiento siempre de la primera de las proporciones conocidas por Pitágoras. Además, un sencillo trabajo algebraico nos demuestra que al tomar, en cada paso  $n$ , como extremos a  $A_n = q / x_{n-1}$  y  $B_n = x_{n-1}$ , con  $x_{n-1} = (A_{n-1} + B_{n-1})/2$ , está tomando la media  $h$  armónica y la media  $m$  aritmética, respectivamente, del intervalo inmediato anterior  $[A_{n-1}, B_{n-1}]$ . Con esta selección ingeniosa de extremos Herón está dando cumplimiento a la proporción más perfecta e introduciendo armonía en su técnica y, como ya mencionamos antes, dotándola de un dinamismo del que carece en este caso el acostumbrado método de Bisección de los cursos actuales de métodos numéricos. En suma, Herón ha aprovechado resultados matemáticos verdaderos ya establecidos en la pitagórica teoría de las medias, en particular los de la proporción sobre la media geométrica y el de la proporción dorada, para diseñar un método efectivo de cálculo de aproximaciones de raíces cuadradas no racionales. Como es sensato suponer que 700 años después, en los tiempos de Herón, estos resultados pitagóricos eran ya parte del acervo cultural de cualquier persona medianamente instruida en matemáticas y por ende bastante conocidos, parece razonable considerar que por ello Herón no necesitaba justificar, ni mucho menos demostrar, matemáticamente que su técnica numérica lograba dar una aproximación aceptable en la práctica. Es esta inmersión total en el discurso matemático pitagórico lo que nos lleva a calificar al método de Herón como un elegante e inesperado ejemplo paradigmático de la extendida afirmación sobre la gran influencia filosófica en la concepción de las matemáticas de los antiguos griegos.

## 6. Referencias

- [1] Platón (1979). *Diálogos*, 18ª ed. México: Editorial Porrúa.
- [2] *Ibidem*.
- [3] Collette, J.-P. (2003). *Historia de las Matemáticas I*, 6ª ed. México: Siglo xxi editores.
- [4] Heath, T. (1981). *A History of Greek Mathematics, vol. II*. New York: Dover Publications, Inc.
- [5] Plato (s/f). *Timaeus*. Recuperado el 12, enero, 2009 de <http://classics.mit.edu/Plato/timaeus.html>
- [6] Aristóteles (1997). *Metafísica, libro décimotercero*. Recuperado el 12, enero, 2009 de [http://www.cervantesvirtual.com/servlet/SirveObras/12260620880115953087846/p0000004.htm#I\\_140\\_](http://www.cervantesvirtual.com/servlet/SirveObras/12260620880115953087846/p0000004.htm#I_140_)

# Conjuntos, grupos, música y *Mathematica*

Inna Shingareva, Reyna Berenice Ibarra Huerta y Carlos Lizárraga Celaya  
Departamento de Matemáticas  
Universidad de Sonora  
inna@fisica.uson.mx, reynaibarra@gmail.com, clizarraga@gmail.com

## Resumen

*En matemáticas simbólicas, la teoría de conjuntos y de grupos computacional es una rama de las matemáticas modernas actualmente muy activa en investigación, que desarrolla y analiza los métodos simbólicos para obtener información sobre grupos. Para muchos grupos interesantes es impráctico realizar cálculos sin ayuda de la computadora. En este trabajo se estudian conjuntos y grupos finitos. Se obtienen nuevos resultados y se muestran las soluciones simbólicas de algunos problemas con ayuda del sistema de álgebra simbólica Mathematica. Se consideran algunas aplicaciones de la teoría de conjuntos y de grupos a la teoría de música tonal.*

## 1. Conjuntos

Uno de los aspectos sorprendentes de la matemática moderna es un *método abstracto* que maneja con símbolos matemáticos y que permite considerar nuevos problemas, obtener nuevos resultados y abrir nuevas áreas de las matemáticas. El *álgebra* (o álgebra abstracta) es una de las áreas más importantes de la investigación matemática que estudia las *estructuras algebraicas*. Una estructura algebraica puede describirse como un *conjunto* de objetos, junto con una serie de operaciones definidas entre ellos.

Desde un punto de vista intuitivo, es conocido que un *conjunto* implica la idea de una colección de objetos bien definidos mediante algunas propiedades comunes. En este trabajo no intentaremos dar una definición formal de un conjunto de acuerdo con la teoría axiomática de conjuntos. Desde un punto de vista formal, podemos considerar un conjunto como un concepto primitivo que no se define. Estudiaremos los conjuntos y obtenemos nuevos resultados con ayuda del sistema de álgebra simbólica. Mientras que el concepto de conjunto es bien conocido para todos, la manera en la cual las computadoras los manipulan es quizás menos conocido [1]. En este trabajo consideraremos el sistema de álgebra simbólica *Mathematica*.

Consideramos el conjunto  $A$  que consiste de los elementos  $\{m, a, t, e, m, a, t, i, c, a\}$ . Podemos mostrar que el número de todos los subconjuntos de  $A$  es igual a 64 y seleccionar todos los subconjuntos que tengan por lo menos un elemento.

$A = \{m, a, t, e, i, c\}; \{B = \text{Subsets}[A], \text{Length}[B], \text{Subsets}[A, \{1\}]\}$

Algunos de los elementos están repetidas y, por lo tanto, el conjunto  $A = \{m, a, t, e, i, c\}$  tiene 6 elementos. Primero,  $A$  es un subconjunto de  $A$ . Existen 6 subconjuntos de  $A$  que tienen 5 elementos:  $\{m, a, e, i, c\}$ ,  $\{a, t, e, i, c\}$ ,  $\{m, a, t, e, c\}$ ,  $\{m, a, t, i, c\}$ ,  $\{m, a, t, e, i\}$ ,  $\{m, t, e, i, c\}$ . Hay

15 subconjuntos de  $A$  que tienen 4 elementos cada uno, 20 subconjuntos que tienen 3 elementos, 15 subconjuntos que tienen 2 elementos, 6 subconjuntos que tienen un solo elemento y un conjunto que consiste de ningún elemento, el *conjunto vacío*. Todos los subconjuntos que tengan por lo menos un elemento son:  $\{a\}, \{c\}, \{e\}, \{i\}, \{m\}, \{t\}$ .

Ahora sea  $A = \{a_1, a_2, \dots, a_n\}$  el conjunto que consiste de  $n$  elementos distintos. Mostramos que  $A$  tiene  $2^n$  subconjuntos distintos.

`1+Sum[Binomial[n,k], {k,1,n}]`

El número de subconjuntos de  $k$  ( $k \geq 1$ ) elementos distintos es el número de maneras de seleccionar  $k$  elementos a partir de  $n$  elementos y es el coeficiente binomial  $\binom{n}{k}$ . Por lo tanto, el número total de subconjuntos distintos (incluyendo conjunto vacío  $\emptyset$ ) es  $1 + \sum_{k=1}^n \binom{n}{k} = 2^n$ .

Consideramos algunas aplicaciones de la teoría de conjuntos a la música tonal [2]. Primero, escribimos la música tonal en el lenguaje matemático. Consideramos una melodía para un instrumento musical (piano), por ejemplo, la Sonata en C-major KV545 de W.A. Mozart (Viena 1788):



Sean  $A_T$ ,  $F_M$  y  $D_M$  conjuntos principales determinados por la asignación, respectivamente, de altura tonal, figuras musicales (o duración musical) y dinámica musical de una nota en una composición musical para un instrumento tonal (con su correspondiente timbre), donde  $A_T = \{a \mid 1 \leq a \leq n, a, n \in \mathbb{N}\}$ ,  $F_M = \{f \mid 1 \leq f \leq m, f, m \in \mathbb{N}\}$ , y  $D_M = \{d \mid 1 \leq d \leq k, d, k \in \mathbb{N}\}$ . Los números  $n, m, k$  que describen la longitud de cada conjunto son dados. Construye dos conjuntos importantes en el área de composición musical: el conjunto  $C$  con todas las posibles composiciones musicales y el conjunto  $M$  determinado por una melodía particular (de nuestro problema). Luego tenemos  $C = A_T \times F_M \times D_M$ ,  $M = \{(a, f, d) \mid a \in A_T \wedge f \in F_M \wedge d \in D_M\}$ ,  $M \subseteq C$ .

```
<<Combinatorica` ; {aT=Table[i, {i, 1, 89}], fM=Table[i, {i, 1, 22}],
                    dM=Table[i, {i, 1, 8}]}
{c=CartesianProduct[CartesianProduct[aT, fM], dM], Length[c]}
```

Para construir  $A_T$  consideramos las alturas que produce un piano. Podemos denotar todas las octavas como el conjunto  $\{-4, -3, -2, -1, 0, 1, 2, 3, 4\}$ . Las alturas de cada octava pertenecen al conjunto:

$$\{C, C\#/D\flat, D, D\#/E\flat, E, F, F\#/G\flat, G, G\#/A\flat, A, A\#/B\flat, B\},$$

donde, por ejemplo,  $C\#/D\flat$  significa la equivalencia de las alturas  $C\#$  y  $D\flat$ . Por lo tanto, podemos describir todas las alturas en la forma de la siguiente tabla (donde excluimos alturas equivalentes tales como  $C\#$  y  $D\flat$ ), por ejemplo,  $A[-4] = 1$ ,  $A\#[-4] = 2$ ,  $B[-4] = 3$ ,  $C[-3] = 4, \dots, C[4] = 88$ , Silencio = 89. Consideramos, por ejemplo, el número máximo de las figuras musicales que describen la duración musical de una nota es 40, por ejemplo:



Formamos el conjunto de las figuras musicales, donde cada figura musical tiene su número natural  $(1, 2, \dots, 22)$ , ver en la figura arriba. Sea el conjunto de la dinámica musical de una nota en una composición musical consistente de 8 elementos,  $\{ppp, pp, p, mp, mf, f, ff, fff\}$ . Luego formamos el conjunto

$$D_M = \{d \mid 1 \leq d \leq 8, d \in \mathbb{N}\},$$

donde  $ppp = 1$ ,  $pp = 2$ , etc. Como resultado podemos, por ejemplo, encontrar 15664 melodías a una sola voz. Con esto es posible formar un gran conjunto  $C$ . El conjunto  $M$  de nuestro problema tiene la forma:

$$M = \{(52, 2, 4), (56, 3, 4), (59, 3, 4), (51, 17, 5), (52, 5, 4), (54, 5, 4), (52, 3, 3), (89, 10, 1), (61, 2, 5), (59, 3, 4), (64, 3, 4), (59, 3, 4), (57, 20, 4), (56, 5, 3), (57, 5, 3), (56, 3, 3), (89, 10, 1)\}.$$

Sean el conjunto de todas las octavas del piano

$$X = \{-4, -3, -2, -1, 0, 1, 2, 3, 4\},$$

el conjunto de todas las alturas de cada octava:

$$f = \{CC, C1, DD, D1, EE, F, F1, G, G1, A, A1, B\},$$

el conjunto de los números naturales  $Y = \{n \mid 1 \leq n \leq 89\}$ . Sea  $f : X \rightarrow Y$  definida por  $A[-4] = 1$ ,  $A1[-4] = 2$ ,  $B[-4] = 3$ ,  $CC[-3] = 4$ ,  $\dots$ ,  $CC[4] = 88$ , Silencio = 89. Construya conjunto  $A_T$  de las alturas que produce el piano.

```
{f={CC,C1,DD,D1,EE,F,F1,G,G1,A,A1,B},
 x={-3,-2,-1,0,1,2,3}, y=Range[1,89],
 fn=Length[f], xn=Length[x], yn=Length[y],
 ln={A[-4],A1[-4],B[-4]}, rn={CC[4],Silence},
 l={}, lf={}}
Pr[j_]:=Module[{i},Do[l=Append[l,f[[i]][x[[j]]]],
 {i,l,fn}];l];
Do[Print[Pr[i]],{i,l,xn-1}]; mn=Pr[xn]
lt=Append[Append[ln,mn],rn]//Flatten
Do[lf=Append[lf,lt[[i1]]==i1],{i1,l,yn}];
at=lf
```

Con esta solución obtenemos que el conjunto de todas las alturas del piano tiene la siguiente forma:

$AT = \{A[-4] = 1, A1[-4] = 2, B[-4] = 3, CC[-3] = 4, C1[-3] = 5, DD[-3] = 6,$   
 $D1[-3] = 7, EE[-3] = 8, F[-3] = 9, F1[-3] = 10, G[-3] = 11, G1[-3] = 12,$   
 $A[-3] = 13, A1[-3] = 14, B[-3] = 15, CC[-2] = 16, C1[-2] = 17, DD[-2] = 18,$   
 $D1[-2] = 19, EE[-2] = 20, F[-2] = 21, F1[-2] = 22, G[-2] = 23, G1[-2] = 24,$   
 $A[-2] = 25, A1[-2] = 26, B[-2] = 27, CC[-1] = 28, C1[-1] = 29, DD[-1] = 30,$   
 $D1[-1] = 31, EE[-1] = 32, F[-1] = 33, F1[-1] = 34, G[-1] = 35, G1[-1] = 36,$   
 $A[-1] = 37, A1[-1] = 38, B[-1] = 39, CC[0] = 40, C1[0] = 41, DD[0] = 42,$   
 $D1[0] = 43, EE[0] = 44, F[0] = 45, F1[0] = 46, G[0] = 47, G1[0] = 48, A[0] = 49,$   
 $A1[0] = 50, B[0] = 51, CC[1] = 52, C1[1] = 53, DD[1] = 54, D1[1] = 55, EE[1] = 56,$   
 $F[1] = 57, F1[1] = 58, G[1] = 59, G1[1] = 60, A[1] = 61, A1[1] = 62, B[1] = 63,$   
 $CC[2] = 64, C1[2] = 65, DD[2] = 66, D1[2] = 67, EE[2] = 68, F[2] = 69, F1[2] = 70,$   
 $G[2] = 71, G1[2] = 72, A[2] = 73, A1[2] = 74, B[2] = 75, CC[3] = 76, C1[3] = 77,$   
 $DD[3] = 78, D1[3] = 79, EE[3] = 80, F[3] = 81, F1[3] = 82, G[3] = 83, G1[3] = 84,$   
 $A[3] = 85, A1[3] = 86, B[3] = 87, CC[4] = 88, \text{Silencio} = 89\}$

## 2. Grupos

La teoría de grupos nos brinda un método general y formal que nos permite analizar sistemas abstractos y sistemas del mundo real complicados (por ejemplo, sistemas en mecánica cuántica) en los cuales se presenta la *simetría*. El álgebra abstracta estudia varias estructuras algebraicas (magma, cuasigrupo, loop, semigrupo, monoide, grupo, anillo, campo). La teoría de grupos estudia varios tipos de grupos, por ejemplo, grupos de permutaciones, grupos de simetría, grupos abelianos de torsión, grupos finitos (o infinitos), grupos libres, grupos de Galois, grupos de Lie,  $p$ -grupos, subgrupos de Sylow, etc. En matemáticas simbólicas, la teoría de grupos computacional es una rama de las matemáticas modernas actualmente muy activa en investigación, que desarrolla y analiza los métodos simbólicos y estructuras de datos para obtener la información y características sobre grupos. Para muchos grupos interesantes es impráctico realizar cálculos sin apoyo de una computadora.

Una aplicación uno a uno de un conjunto (no vacío)  $X$  sobre sí mismo recibe el nombre de permutación de  $X$ . El conjunto  $S(X)$  de todas las permutaciones de  $X$  es un grupo que se llama *grupo simétrico* sobre  $X$ . Cualquier subgrupo de  $S(X)$  recibe el nombre de *grupo de permutaciones*. Sea  $S_n$  el grupo simétrico de todas las permutaciones sobre los símbolos  $1, 2, \dots, n$  ( $n$  está dado). Determine: a) todas las permutaciones  $s_i \in S_4$ , b) el número de las permutaciones ( $i = 1, \dots, 24$ ), c) todos los ciclos ajenos, d) para cualquier dos permutaciones  $p_1, p_2$ , calcule el producto de permutaciones, e) todos elementos del grupo  $S_4$  generados por los ciclos ajenos  $(1, 2)$  y  $(1, 2, 3, 4)$  y despliega todos elementos en la notación standard matemática para permutaciones.

```

<<Combinatorica`; {s4=Permutations[Range[4]], ls4=Length[s4],
p1=ToCycles[s4[[3]]],

```

```

p2=ToCycles[s4[[4]]], p11=FromCycles[p1], p21=FromCycles[p2],
m1=Permute[p11,p21],m2=Permute[p21,p11], Map[ToCycles,{m1,m2}]
{l4={}, l4=Append[l4, Table[ToCycles[s4[[i]]],{i,1,ls4}]],
HoldForm[l4]==l4, p12=FromCycles[{{1,2},{3},{4}}],
p22=FromCycles[{{1,2,3,4}}], g4=SymmetricGroup[4],
PermutationGroupQ[g4], lg4=Length[g4]
Map[MatrixForm,Table[{{1,2,3,4},g4[[i]]},{i,1,lg4}]]

```

Sean  $p_1$  y  $p_2$  dos permutaciones  $\in S_7$ . Encontramos  $p_1^{-1}$  y verificamos que  $p_1^{-1}p_2p_1 \neq p_2$ .

```

<<Combinatorica`; {p1={4,6,7,3,5,1,2}, p2={3,6,7,1,5,4,2},
p1Inv=InversePermutation[p1], p2p1=Permute[p2,p1],
p1Invp2p1=Permute[p1Inv,p2p1],p1Invp2p1===p2}

```

Sea  $S_3$  el grupo simétrico de todas las permutaciones. Determinemos todos los elementos de  $S_3$ , el orden de cada elemento del grupo y el orden del grupo  $S_3$ .

```

<<Combinatorica`; {n=3,PermutationGroupQ[SymmetricGroup[n]],
eg3=SymmetricGroup[n], grouporder=n!, leg3=Length[eg3]}
Map[MatrixForm,Table[{{1,2,3},eg3[[i]]},{i,1,leg3}]]
{s31=Permutations[Range[3]],ls31=Length[s31]}
Table[ToCycles[s31[[i]]],{i,1,ls31}]
{Permute[eg3[[2]],eg3[[2]]],Permute[eg3[[3]],eg3[[3]]],
Permute[eg3[[6]],eg3[[6]]],
Permute[Permute[eg3[[4]], eg3[[4]]], eg3[[4]]],
Permute[Permute[eg3[[5]], eg3[[5]]], eg3[[5]]]}

```

La primera permutación en la lista EG3 es la permutación identidad  $(1, 2, 3)$  y se denota en *Mathematica* por  $[\ ]$ . Las permutaciones  $(2, 1, 3)$ ,  $(1, 3, 2)$ , y  $(3, 2, 1)$  tienen el orden 2, las dos permutaciones restantes  $(2, 3, 1)$  y  $(3, 1, 2)$  tienen el orden 3.

Un método importante para estudiar las características de un grupo finito es la construcción de la tabla abstracta del grupo. Sea  $G$  el grupo multiplicativo de seis elementos  $\{e, a, b, c, d, f\}$ . El teorema de Cayley indica que cada grupo finito  $G$  es isomorfo a un subgrupo  $SG$  del grupo simétrico  $S_n$  sobre  $G$ . Sea  $f : G \rightarrow S_3$  un isomorfismo, donde  $f(e) = (1, 2, 3)$ ,  $f(a) = (2, 3, 1)$ ,  $f(b) = (3, 1, 2)$ ,  $f(c) = (2, 1, 3)$ ,  $f(d) = (3, 2, 1)$ ,  $f(f) = (1, 3, 2)$ . Construya la tabla de multiplicación (la *tabla de Cayley*) de elementos de  $G$ .

```

<<Combinatorica`; eg3={{1,2,3},{2,3,1},{3,1,2},{2,1,3},{3,2,1},
{1,3,2}} {n=3,g={e,a,b,c,d,f},leg3=Length[eg3]}
subs=Table[eg3[[i]]->g[[i]],{i,1,leg3}]
multiplication[p1_,p2_] :=Permute[p1,p2];
{mTable1=MultiplicationTable[eg3,multiplication]//MatrixForm,
mTable2=MultiplicationTable[eg3,
multiplication]/.{m_Integer:>eg3[[m]]}//MatrixForm,mTable2/.subs}

```

El isomorfismo se describe en la lista EG3 de acuerdo a la definición del problema. Pero se puede generar un grupo de permutaciones (con generadores o relaciones) y construir la tabla de Cayley. En este caso cambiamos la solución:

```
<<Combinatorica`; {n=3,g={e,a,b,c,d,f},
  tH={"e","a","b","c","d","f"},eg3=SymmetricGroup[n],leg3=Length[eg3]}
subs=Table[eg3[[i]]->g[[i]],{i,1,leg3}]
mult[p1_,p2_] :=Permute[p1,p2];
{mTable1=MultiplicationTable[eg3,mult]//MatrixForm,
  mTable2=MultiplicationTable[eg3,mult]/.{m_Integer:>eg3[[m]]},
  mTable2//MatrixForm, mTable2/.subs}
TableForm[mTable2/.subs,TableHeadings->{tH,tH}]
```

Generamos el grupo cuatro de Klein  $G$ , con los generadores  $\{x, y, z\}$  y las relaciones  $xz = y$ ,  $yz = x$ ,  $xy = z$ ,  $yz = z$ . Determinemos el orden del grupo, todos los elementos del grupo  $G$ , y construimos la tabla de Cayley.

```
<<Combinatorica` m[a1___,a2___]:=NonCommutativeMultiply[a1,a2];
rules={m[a1___,e,a2___]:>m[a1,a2],m[a1___,e,a2___]:>m[a1,a2],
m[a1___,x,x,a2___]:>m[a1,e,a2],m[a1___,y,y,a2___]:>m[a1,e,a2],
m[a1___,z,z,a2___]:>m[a1,e,a2],m[a1___,x,z,a2___]:>m[a1,y,a2],
m[a1___,x,y,a2___]:>m[a1,z,a2],m[a1___,y,z,a2___]:>m[a1,x,a2],
m[a1___,y,x,a2___]:>m[a1,z,a2],m[a1___,z,x,a2___]:>m[a1,y,a2],
m[a1___,z,y,a2___]:>m[a1,x,a2],m[a1_]:>a1};
{gr4Klein={ {e**e,e**x,e**y,e**z}, {x**e,x**x,x**y,x**z},
{y**e,y**x,y**y,y**z}, {z**e,z**x,z**y,z**z} }//.rules,
  Length[gr4Klein],tH={"e","x","y","z"} }
mTable=TableForm[gr4Klein,TableHeadings->{tH,tH}]
```

Ahora consideramos una aplicación de la teoría de grupos y de anillos al sistema tonal de la Música Occidental. Sabemos que este sistema tiene la escala cromática de 12 notas y es un sistema del temperamento igual. Supongamos que un piano tiene 12 octavas y no hay diferencias entre los tonos tales como  $C\#/Db$ , etc. Frecuentemente, para investigaciones matemáticas y computacionales en la teoría de música, la diferencia entre las octavas se puede ignorar. Construimos una enumeración para todas las notas de un piano dado, definimos una función que describe la transposición musical, definimos una función que describe la inversión musical.

```
<<Combinatorica`; {n=12,n1=9, z12=Table[i,{i,0,n-1}],
  eqC1=Table[Table[m+12*k,{k,0,n-1}],{m,0,n-1}];
  eqC1[[1]], eqC1[[2]]}
mT[x_List,k_]:=Map[Mod[#+k,12]&,x];
{mT[eqC1[[1]],n1],p1=mT[z12,n1], l1={},
  Do[l1=Append[l1,mT[z12,i]],{i,1,n}]; l2=l1/.{0->12},
  p11=p1/.{0->12}, ToCycles[p11],Map[ToCycles,12]}
mI[x_List,k_]:=Map[Mod[2*k-#,12]&,x];
{l1[[1]],mI[z12,0],l1[[2]],mI[l1[[1]],1],l1[[3]],mI[l1[[2]],2]}
```

Primero, seleccionamos un tono principal arbitrario  $a_1 = 0$ , luego podemos enumerar fácilmente los tonos siguientes construyendo una biyección entre el sistema de todos los tonos y el sistema de todos los números enteros. Puesto que excluimos la diferencia entre las octavas, podemos clasificar todos los tonos en las clases de equivalencia  $\mathbb{Z}_{12} = \{m + 12k, k \in \mathbb{Z}, m \in [0, 11]\}$ . Por lo tanto, la enumeración del sistema tonal de la Música Occidental, descrita como un anillo conmutativo  $\mathbb{Z}_{12}$  con las dos operaciones  $(+, *)$  sobre los números enteros modulo 12 y la identidad  $1 \neq 0$ , se puede escribir de la siguiente forma:



Recordemos que la transposición musical por  $N$  tonos más alto o más bajo es equivalente a sustituir cada tono por el tono correspondiente (esto es,  $N$  tonos más altos o más bajos). Por lo tanto, la función de transposición se puede escribir en la forma:  $MT(x, N) = x + N \pmod{12}$ , donde  $x = \{m + 12k, k \in \mathbb{Z}, m \in [0, 11]\}$  y  $N \in [0, 11]$ . Esta función describe todas las permutaciones posibles de  $\mathbb{Z}_n$ . Notemos que para trabajar con permutaciones, sustituimos el elemento 0 de  $\mathbb{Z}_n$  por 12. La función de inversión se puede definir en la forma  $MI(x, N) = 2N - x \pmod{12}$ , donde  $x = \{m + 12k, k \in \mathbb{Z}, m \in [0, 11]\}$  y  $N \in [0, 11]$ .

### 3. Conclusiones

En el presente trabajo se obtienen nuevas soluciones simbólicas de varios problemas computacionales en teoría de conjuntos y de grupos finitos con el sistema de álgebra simbólica *Mathematica*. Se consideren algunas aplicaciones de la teoría de conjuntos y de grupos a la teoría de música tonal. Las soluciones simbólicas de problemas considerados demuestran que el proceso de la solución de problemas y de la enseñanza en teoría de grupos finitos llega a ser más fácil y más interactivo debido a la representación visual de los sistemas de álgebra simbólica.

### 4. Referencias

- [1] Shingareva I., Lizárraga-Celaya C. (2009). *Maple and Mathematica. A Problem Solving Approach for Mathematics*. Wien, New York Springer Verlag, Second Edition.
- [2] Nierhaus G. (2009) *Algorithmic Composition. Paradigms of Automated Music Generation*. Wien, New York Springer Verlag.



# Evaluación de las habilidades de comprensión de código en C++

Gabriel Alberto García Mireles e Irene Rodríguez Castillo  
Departamento de Matemáticas  
Universidad de Sonora  
{mireles, irene}@gauss.mat.uson.mx

## Resumen

*En este documento se presentan los avances logrados en la construcción de un instrumento de evaluación, elaborado con el propósito de ser una herramienta de apoyo para la certificación de conocimientos y habilidades que un estudiante de licenciatura adquiere en el curso de programación de computadoras con el lenguaje C++. El instrumento fue elaborado y validado por un grupo de docentes con experiencia en la enseñanza del lenguaje en cuestión. Consiste en 12 reactivos, los cuales abordan seis temas principales sobre la resolución de problemas mediante programas de computadora. Se realizó la prueba piloto sobre dos grupos de la licenciatura en ciencias de la computación de la Universidad de Sonora. Aunque los resultados obtenidos en dicha prueba indican que el instrumento propuesto, en términos de complejidad y discriminación, es aceptable, todavía necesita ajustes.*

## 1. Introducción

Los cursos de programación de computadoras generalmente se imparten en el primer año de estudios en una licenciatura del área de computación. En éstos, el estudiante elabora programas en un lenguaje, como C, C++ o Java entre otros con la finalidad de resolver problemas de procesamiento de información. Sin embargo, lograr el dominio de los contenidos de estos cursos se convierte en una meta compleja para muchos estudiantes.

La complejidad del aprendizaje de la programación se refleja en las altas tasas de deserción y reprobación en los programas del área de computación [6]. Por su parte Woszczyński, Guthire y Shade [19] indican que los porcentajes de reprobación y de deserción de las clases de programación exceden el 50%. De acuerdo a la experiencia que se tiene en la licenciatura en Ciencias de la Computación de la Universidad de Sonora, el curso tiene una tasa de aprobación cercana al 40% [7]. De hecho, en el estudio de egresados se considera que los estudiantes tienen trayectoria baja y que las materias que tienen el índice de promoción más bajo son las que siguen a la asignatura de programación de computadoras [13].

Una forma de conocer la amplitud y profundidad del aprendizaje del estudiante es a través del uso de un instrumento de evaluación. Aunque existen distintos instrumentos que permiten certificar las habilidades desarrolladas y los conocimientos adquiridos en el área de computación [4], este instrumento se utiliza cuando el alumno egresó de la licenciatura. Por otro lado, se ha

dado atención a la evaluación de los aspectos de cobertura de las políticas públicas en educación y a la acreditación y certificación, dando muy baja prioridad a la calidad del aprendizaje que logra el estudiante en cada uno de los cursos [3]. Por tanto, el objetivo de este trabajo es presentar los avances logrados en el desarrollo de un instrumento de evaluación que permita conocer las habilidades de comprensión adquiridas por los estudiantes en un curso de programación en el lenguaje C++.

Se presentará primeramente un panorama general sobre el significado del aprendizaje de la programación de computadoras y cómo éste repercute en la forma en que se evalúa tal aprendizaje. En seguida se describirán algunos instrumentos de evaluación que han sido elaborados para determinar el nivel de aprendizaje de los estudiantes en el curso de programación de computadoras, pero que por su naturaleza o contexto, no satisfacen las necesidades del ambiente que nos interesa. Se mostrará la metodología utilizada en la elaboración del instrumento de evaluación propuesto, así como los resultados de la prueba piloto aplicada. Por último, se indicarán las conclusiones y el trabajo a futuro para el mejoramiento de dicho instrumento de evaluación.

## **2. Aprendizaje de la programación**

Aprender a programar una computadora significa que el individuo sea capaz de analizar una problemática de procesamiento electrónico de información, diseñe soluciones y escoja la mejor, de acuerdo a las restricciones establecidas; instrumente el plan elegido en un lenguaje de programación y genere un programa de computadora, o software, que responda a las necesidades planteadas. Además, este programa deberá ejecutarse bajo cierto sistema operativo; y en el caso de defectos en el programa, o solicitudes de mejora, encontrar, seguir y depurar el código generado. Independientemente del estilo y paradigma utilizado (orientado a objetos o imperativo), los estudiantes del primer curso de programación deben aprender las estructuras de control básicas y los conceptos y procedimientos para manejar las variables, paso de parámetros entre funciones y recursión [9].

Actualmente, existe bibliografía básica para el curso de programación de computadoras en C++ cuyo contenido temático toma como referencia la sintaxis y estructura del lenguaje de programación. Pero también, abordan algunos conceptos elementales de ingeniería de software como diseño de software, herramientas y entorno de software, requisitos y especificaciones de software [8][2][5].

El estudio de la problemática asociada al aprendizaje de la programación se ha realizado desde los ochenta [11][14][17]. Se reconoce que un programador novato tiene mayor facilidad para aprender la sintaxis y semántica de las instrucciones individuales del lenguaje de programación, pero tiene dificultades para combinarlas en programas válidos. Bayman y Mayer [1] desarrollaron un modelo en el cual se clasifican los distintos tipos de conocimientos adquiridos al aprender programación: sintáctico, conceptual y estratégico. El conocimiento sintáctico se refiere a los hechos específicos del lenguaje y las reglas para su uso, mientras que el conocimiento conceptual está relacionado con las estructuras de los programas y principios que las sustentan. Finalmente, para ellos el conocimiento estratégico está relacionado con la elaboración de versiones específicas de programas para resolver problemas.

Sin embargo, se reconoce que no se ha encontrado una estrategia para elaborar el currículo y enseñar la programación de computadoras de tal manera que sea aplicable a todos los contextos específicos de programas educativos [15]. Además, se recomienda a las instituciones educativas y profesores continuar con la experimentación en esta área.

### **3. Métodos de evaluación para los cursos de programación**

La intervención educativa puede mejorar los resultados de aprendizaje si se incorporan procedimientos de evaluación que permitan identificar riesgos y prescribir mejoras en mecanismos que fomenten el aprendizaje eficiente de los estudiantes [3]. En la literatura se reportan distintos instrumentos aplicados sobre poblaciones de estudiantes que están por terminar o terminaron el primer curso de programación. Los objetivos y contenidos evaluados difieren en gran medida y es difícil tomarlos como referencia para aplicarlos a un contexto local. A continuación se presentan los principales instrumentos de evaluación utilizados.

McCracken et al. [12] desarrollaron un instrumento de evaluación para conocer la habilidad del estudiante en la elaboración de un programa en el cual se requería una pila para procesar las entradas en notación infija y posfija. La elección del lenguaje de programación se le dejaba libre al estudiante. Los resultados de la prueba piloto mostraron que los estudiantes tuvieron un desempeño pobre. De acuerdo a Whalley et al. [16], la discusión de los resultados indica que el estudiante, aunque estaba familiarizado con las estructuras fundamentales de la programación, carecía de las habilidades para resolver problemas.

Por su parte, Lister [10] desarrolló un examen de opción múltiple para el curso introductorio del lenguaje de programación Java. Parte del supuesto que una evaluación basada en criterio no debería probar la habilidad del estudiante para escribir código, como ocurrió con el instrumento de McCracken, sino que debería enfocarse en la habilidad del estudiante para leer código. El instrumento se elaboró considerando que el estudiante comprende: los conceptos básicos de la programación orientada a objetos; las estructuras básicas de control (secuencia, selección, iteración); y, el código que implementa los algoritmos básicos de ordenación y búsqueda en arreglos. El instrumento consta de 26 preguntas, las cuales en conjunto, abordan los tres objetivos anteriores. El instrumento se aplicó a 336 estudiantes australianos y el número de aciertos obtenidos estuvo en el rango de 4 a 26, en donde alrededor del 45 % tuvo menos de 10 aciertos.

Por último, Whalley et al. [16] desarrollaron un instrumento para evaluar las habilidades de lectura y comprensión de los programadores novatos. El análisis y elaboración de los reactivos de este instrumento se basó en la versión revisada de la taxonomía de Bloom [16] para los primeros nueve reactivos de opción múltiple, cuyo propósito es medir la comprensión lograda por el estudiante. Las preguntas se pueden clasificar como de código fijo y de esqueleto; en las primeras, el estudiante realiza una corrida de escritorio para determinar el valor de salida, mientras que en la otra categoría, el estudiante tiene que completar la línea de código faltante en el problema presentado. El último reactivo consistió en un segmento de código que se le da al estudiante y se le pide que describa el propósito de éste en el lenguaje materno. El instrumento se aplicó a una muestra de 117 estudiantes, y los resultados sugieren que los estudiantes que no pueden leer ni describir un segmento de código corto, tampoco han desarrollado las habilidades

para escribir código por ellos mismos. Finalmente, indican que la evaluación de la habilidad de programación que tiene el estudiante es una tarea compleja y desafiante, para la cual los educadores del área de programación carecen de marcos de trabajo claros y de herramientas apropiadas.

#### 4. Método

El grupo de profesores que participó en la elaboración y revisión del instrumento está conformado por cinco docentes. La formación académica de cuatro de ellos es en la licenciatura en Matemáticas; y el otro en Ciencias Computacionales. Además, cuatro de ellos tienen grado de maestría en Ciencias de la Computación. En cuanto al curso introductorio de programación de computadoras, al menos ellos lo han impartido durante cuatro años, en diferentes momentos de su actividad profesional. Los docentes tienen experiencia en enseñar los lenguajes de programación C y C++, que básicamente forman parte de los lenguajes que se utilizan actualmente para aprender a programar.

Las etapas para construir el instrumento y probarlo fueron: planeación, elaboración de los reactivos, validación de reactivos y prueba piloto. En la planeación se analizaron los objetivos y contenidos del programa vigente del curso de programación de computadoras, que se imparte en la licenciatura en ciencias de la computación de la Universidad de Sonora.

En la etapa de elaboración de reactivos, en reunión con el grupo de docentes participantes, se decidió por consenso que los temas básicos que deberían ser evaluados a nivel aplicación y resolución de problemas serían: identificación de variables y tipos de datos asociados a partir del enunciado del problema, problemas en donde se involucre el uso de la estructura de control secuencial y las operaciones aritméticas con énfasis a la aritmética de enteros, problemas asociados al uso de las estructuras de control secuencial y repetitiva en las modalidades simples y anidadas, manejo de arreglos, y finalmente, uso de funciones. Se consideró desarrollar un instrumento que abordara los temas elegidos en el cual los reactivos fueran de opción múltiple, con cinco opciones, y que para cada tema principal se desarrollaran al menos dos reactivos. En total se desarrollarían doce reactivos. Se tomó como referencia el trabajo de Wahley et al. [16] para construir los reactivos del examen. En la tabla 1 se muestran los temas abordados por cada reactivo.

El equipo de docentes participantes validó los reactivos. Cada uno de ellos tuvo al menos una semana con el material para revisarlo individualmente antes de la sesión grupal. Los aspectos por evaluar consistían en determinar si el reactivo era adecuado al tema, identificar la opción correcta, verificar que las opciones incorrectas fueran creíbles para alumnos con pocos conocimientos, verificar que cada reactivo tuviera una opción correcta, verificar que la redacción del reactivo sea clara y comprensible por los estudiantes, y finalmente, contestar el examen para estimar el tiempo que lleva resolverlo.

En las sesiones se hicieron observaciones sobre el estilo de redacción de cuatro reactivos. Se solicitó que el estilo de presentación de código fuera consistente, y para uno de ellos, verificar las opciones de respuesta. Las observaciones se atendieron y se presentó la versión actualizada a los docentes nuevamente. En esta reunión, se decidió por consenso que el examen podría ser un buen

indicador de la habilidad de los estudiantes para resolver problemas de programación relacionados con los temas básicos de la asignatura.

La prueba piloto se aplicó en el 2006 a los dos grupos de programación que se imparte en el programa de licenciatura en ciencias de la computación. Ambos grupos estaban a cargo de la misma maestra, quien participó en la revisión de los reactivos. En total fueron 18 estudiantes, en el grupo matutino fueron ocho hombres y tres mujeres, mientras que en el grupo vespertino sólo fueron siete hombres. La moda es de 19 años para la edad de los participantes. A cada estudiante se le entregó una hoja de respuestas y el examen en papel.

**Tabla 1. Descripción de las doce preguntas del instrumento de evaluación.**

<b>Núm.</b>	<b>Tema</b>	<b>Descripción</b>
1	Variables	Identificar las variables que se requieren para calcular el descuento en una transacción
2	Variables	Identificar las variables que se requieren en el programa para completar los espacios de la plantilla del informe
3	Operaciones aritméticas	A partir del total de una factura, deducir las operaciones para determinar el importe y el impuesto
4	Aritmética de enteros	Predecir la salida de tres variables, las cuales tienen asignadas operaciones aritméticas con enteros
5	Estructuras selectivas simples	Determinar la opción de código que correctamente respete las condiciones especificadas en el planteamiento del problema
6	Estructuras selectivas anidadas	Se presenta un segmento de código incompleto, y el estudiante escoge la opción que satisface las condiciones del problema, en donde intervienen tres variables
7	Estructuras de control repetitivas	El estudiante escoge la opción que especifica la cantidad de elementos, el valor inicial y el factor de incremento que corresponden a la secuencia que genera el código planteado
8	Estructuras de control repetitivas	Determinar la sección en el código, en donde debe incluirse la inicialización del ciclo interno del programa
9	Vectores	El estudiante realiza una corrida de escritorio para determinar el valor que está asignado en una celda específica del arreglo de enteros
10	Matrices	Multiplicar un vector por una matriz para determinar el valor que solicita el reactivo
11	Funciones	Escoger el segmento de código que realiza la llamada de función con los argumentos adecuados al planteamiento del problema
12	Funciones y paso de parámetros	El estudiante escoge la opción que satisface la llamada de una función con paso de parámetros por valor y de arreglos

## 5. Resultados

Los estudiantes estaban por terminar el ciclo escolar y la calificación obtenida varió entre 4 y 12 aciertos. Al examen se le aplicaron las siguientes pruebas: índice de dificultad de los reactivos, nivel de discriminación y consistencia interna (alfa de Cronbach). Se utilizan los procedimientos especificados en Woolfok [18].

Para calcular el parámetro de consistencia interna se utilizó el SPSS 15.0 para Windows. La tabla 2 muestra los resultados del cálculo del índice de dificultad y nivel de discriminación de los reactivos. Se señala que los reactivos 1 y 5, resultaron muy fáciles para los estudiantes; mientras que el reactivo 2, fue realmente complejo. Respecto al nivel de discriminación, el reactivo 2 no discrimina. Para los estudiantes que realizaron el examen, encontraron que este instrumento tiene un índice de dificultad de 0.41 y un nivel de discriminación de 0.43. Estos valores son consistentes con un buen nivel de examen [18:564]. Los resultados que brinda el software SPSS respecto al alfa de Cronbach con los 18 casos es de 0.625.

**Tabla 2. Índice de dificultad y grado de discriminación para los reactivos de la prueba.**

<b>Índice de dificultad</b>												
<b>Criterio de aceptación .20 &gt; IDF &lt; .80</b>												
R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	Prom
0	5	1	2	0	1	1	0	0	0	3	1	
1	5	3	3	2	2	5	4	5	4	6	5	
12	12	12	12	12	12	12	12	12	12	12	12	
0.08	0.83	0.33	0.42	0.17	0.25	0.50	0.33	0.42	0.33	0.75	0.50	0.41

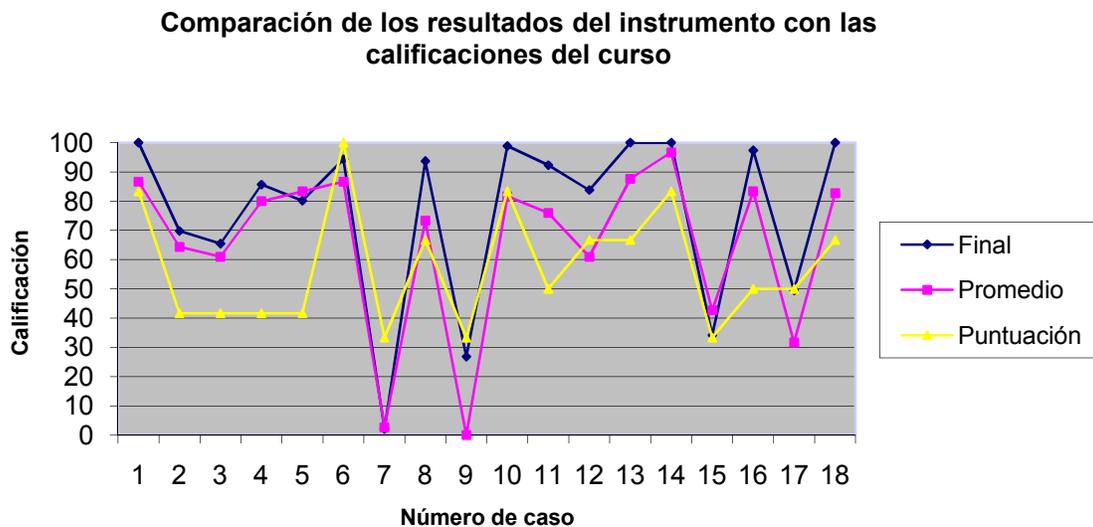
<b>Nivel de discriminación</b>												
<b>Criterio de aceptación IDC &gt; 0</b>												
R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	Prom
6	1	5	4	6	5	5	6	6	6	3	5	
5	1	3	3	4	4	1	2	1	2	0	1	
6	6	6	6	6	6	6	6	6	6	6	6	
0.17	0	0.3	0.17	0.33	0.17	0.67	0.67	0.8	0.7	0.5	0.67	0.43

En cuanto a la comparación de la puntuación obtenida por los estudiantes con este instrumento, se solicitó al docente responsable del curso que proporcionara una copia de las evaluaciones parciales y finales para validar los resultados obtenidos en el examen. Como se observa en la figura 1, el patrón de calificaciones es semejante entre ambos enfoques de evaluación. La etiqueta promedio se refiere a la calificación lograda en los exámenes parciales del curso, mientras que la etiqueta final, agrega otros elementos de evaluación como prácticas y ejercicios realizados. La etiqueta puntuación muestra las calificaciones logradas en la prueba piloto. De acuerdo al análisis de correlación entre el promedio obtenido y la puntuación lograda en el instrumento de evaluación se tiene que la correlación de 0.650 es significativa entre ambas variables, con un nivel de significancia de 0.01.

## 6. Conclusiones

Los docentes del área de programación de computadoras necesitan herramientas confiables que permitan evaluar el nivel de aprendizaje logrado por los estudiantes. En este trabajo se presenta una propuesta para conocer el nivel de comprensión que logra un estudiante de programación de los temas básicos que se imparten en estos cursos. Como los resultados indican, es un instrumento que tiene elementos de validez de contenido, pero que necesita ajuste para lograr los niveles de confiabilidad adecuados.

En un futuro se espera contar con un modelo de aprendizaje de la programación, que guíe el desarrollo del instrumento de evaluación para determinar el nivel de dominio que logra el estudiante en los cursos de programación de computadoras.



**Figura 1. Representación gráfica de la calificación final obtenida en el curso, el promedio de los exámenes y la puntuación lograda en el instrumento piloteado.**

## 7. Referencias

- [1] Bayman, P. & Mayer, R. E. (1988). *Using Conceptual Models to Teach BASIC Computer Programming*. Journal of Educational Psychology, v80, n3, p291-98.
- [2] Bronson, G. J. (2000). *C++ para ingeniería y ciencias*. International Thomson Editores. México.
- [3] Castañeda F., S. (2004). *Educación, aprendizaje y cognición. Teoría en la práctica*. México: El Manual Moderno.
- [4] Centro Nacional para la Evaluación de la Educación Superior, A. C. (2005). Varianzas. El EGEL Informática - Computación. Boletín CENEVAL núm. 19 Archivo adjunto Boletin19.pdf. Consultado el 8 de mayo de 2006 en: (palabra clave: computación) <http://www.ceneval.edu.mx/portalceneval/>

- [5] Deitel, H. M. y Deitel, P. J. (2003). *Cómo programar en C++*. Cuarta edición. Pearson Educación, México.
- [6] Ellerbracke, S. y Lomelí, E. (s/f). *Análisis de la evolución (1971-2001) y tendencias (2002-2005) de los programas de Informática y Computación en México*. Consultado el 16 de mayo de 2006 en: [http://www.software.net.mx/desarrolladores/prosoft/Estudios/evol\\_progs.htm](http://www.software.net.mx/desarrolladores/prosoft/Estudios/evol_progs.htm)
- [7] García Mireles, G. A. y Rodríguez Castillo, I. (2005). *Exploración y Utilización del Proceso de Desarrollo de Software en un Curso Introductorio de Programación*. En memorias del XVIII Congreso Nacional y IV Congreso Internacional de Informática y Computación. Del 26 al 28 de octubre de 2005. Torreón, Coahuila.
- [8] Joyanes Aguilar, L. (2003). *Fundamentos de programación. Algoritmos, estructuras de datos y objetos*. Tercera edición. Madrid: McGrawHill.
- [9] Lahtinen, E; Ala-Mutka, K. y Hannu-Matti J. (2005). *A Study of the difficulties of Novice Programmers*. En Memorias del ITiCSE '05 Junio 27-29, 2005, Monte de Caparica, Portugal.
- [10] Lister, R. (2005). *One Small Step Toward a Culture of Peer Review and Multi-Institutional Sharing of Educational Resources: A Multiple Choice Exam for First Semester Programming Students*. Australasian Computing Education Conference 2005, Newcastle, Australia. P155-164.
- [11] Mayer, R. E. (1981). *The Psychology of How Novices Learn Computer Programming*. Computing Surveys 13(1) 121 – 141.
- [12] McCracken, M., V. Almstrum, D. Diaz, M. Guzdial, D. Hagen, Y. Kolikant, C. Laxer, L. Thomas, I. Utting, T. Wilusz, (2001): *A Multi-National, Multi-Institutional Study of Assessment of Programming Skills of First-year CS Students*, SIGCSE Bulletin, 33(4):125-140.
- [13] Pacheco, M. y Burgos, B. (2005). *Estudio de Trayectoria Escolar*. Colección Documentos de Investigación educativa: Universidad de Sonora.
- [14] Soloway, E. (1986). *Learning to program = learning to construct mechanisms and explanations*. Communications of the ACM. 29(9) 850 – 858.
- [15] *The Joint Task Force on Computing Curricula* (2005). Computing Curricula 2005. The Overview Report. Computing Curricula Series. Consultado el 9 de mayo de 2006 en: [http://www.acm.org/education/curric\\_vols/CC2005-March06Final.pdf](http://www.acm.org/education/curric_vols/CC2005-March06Final.pdf)
- [16] Whalley, J. L.; Raymond Lister, Errol Thompson, Tony Clear y Phil Robbins, P. K. Ajith Kumar, Christine Prasad. (2006). *An Australasian Study of reading and comprehension Skills in Novice Programmers, using Bloom and SOLO Taxonomies*. Proceedings of the 8th Australasian Computing Education Conference, pp 243-252.
- [17] Winslow, L. E. (1996). *Programming Pedagogy – A Psychological Overview*. SIGSE Bulletin 28(3). 17-25.
- [18] Woolfok, A. E. (1999). *Psicología Educativa*. Prentice-Hall, 7ma. Edición. México.
- [19] Woszczyński, A.; T. Guthrie; S. Shade. (2005). *Personality and Programming*. Journal of Information Systems Education. 16(3) 293 – 299.

# Resolución de ecuaciones lineales a través de material manipulable

Patricia Gómez Avilés, Gersón Hernández Martínez,  
Angel Muñoz Velázquez e Ignacia Elizabeth García Ramírez  
Dirección de Investigación Educativa y Fortalecimiento Institucional  
Secretaría de Educación Pública Hidalgo  
patygomeza@hotmail.com

## Resumen

*El presente trabajo es una propuesta producto de la operatividad del Proyecto “Ludoteca Interactiva de Matemáticas” en el estado de Hidalgo; con la finalidad de abatir una de las problemáticas que enfrentan los docentes durante el proceso enseñanza-aprendizaje con relación al álgebra; misma que propone el uso de material manipulativo que propicia en los estudiantes el desarrollo de competencias matemáticas (planteamiento y resolución de problemas; la comunicación; argumentación y la aplicación de técnicas).*

## 1. Introducción

Durante el ciclo escolar 2005-2006 se dio inicio con un estudio complementario en el nivel de educación secundaria “Resolución de Ecuaciones Lineales a través de un Material Manipulable” éste tuvo la finalidad de poder detectar a aquellos docentes que tienen el interés de poner en práctica el uso de nuevas metodologías y estrategias didácticas en apoyo a la enseñanza y el aprendizaje de las matemáticas; se da bajo la intención de conocer las necesidades y problemáticas que tienen los docentes de educación secundaria al abordar temas relacionados con el álgebra.

Surge como resultado del seguimiento y evaluación de la operatividad de la propuesta de intervención *Ludoteca Interactiva de Matemáticas* la cual coordina la Dirección de Investigación Educativa y Fortalecimiento Institucional, propone el uso de materiales manipulativos los cuales propician en el estudiante la reflexión, el análisis, el razonamiento, la socialización, la argumentación y la investigación, al mismo tiempo se crea un ambiente propio para hacer matemáticas dentro del aula. Con relación a lo anterior se presenta la estructura del estudio, los resultados encontrados y un apartado de conclusiones.

## 2. Antecedentes

*Ludoteca Interactiva de Matemáticas* es una propuesta de intervención se da de una investigación en 1999, de la cual se detectaron cuatro problemáticas en la enseñanza y el aprendizaje de las matemáticas: 1) bajo aprovechamiento de la asignatura; 2) falta de materiales didácticos en apoyo a la asignatura; 3) poca vinculación con otras asignaturas; y 4) escasa relación maestro-alumno.

Ante estas problemáticas se diseña la propuesta de intervención que enfatiza su interés en propiciar diferentes estrategias metodológicas dirigidas a los docentes sobre el uso y aplicación de materiales manipulativos que favorezcan en el estudiante el desarrollo de competencias matemáticas, despertando su interés por la investigación y el desarrollo de habilidades a través de un enfoque lúdico; reconociendo el papel fundamental que tienen las matemáticas en la ciencia, la tecnología y en la vida cotidiana considerando que el enseñar no solo refiere a tener los conocimientos necesarios sino que se requiere considerar didácticas y estrategias que favorezcan la construcción de conocimientos significativos en los estudiantes.

La fase de pilotaje inicia en 11 escuelas secundarias en sus tres modalidades (Generales, Técnicas y Telesecundarias) en 7 regiones del estado, mismas a las que se les proporciono materiales manipulativos y escritos con sugerencias didácticas para su uso y aplicación de estos. Durante esta etapa y los ciclos escolares subsecuentes surgen solicitudes de varias escuelas para integrarse a la propuesta, por lo que se amplía la cobertura y actualmente 204 escuelas de 11 regiones del estado la implementan.

### **3. Propósito**

Identificar a los docentes que fundamentan su práctica en la utilización de recursos didácticos apoyados con materiales manipulables como apoyo al proceso enseñanza y aprendizaje de temas relacionados al Álgebra en educación secundaria, con la intención de poner en práctica dichos recursos con los docentes de las 204 escuelas que integran el proyecto.

### **4. Justificación**

En este sentido se dio inicio con un estudio que permitiera identificar si los docentes que imparten la asignatura de matemáticas en educación secundaria y que se encuentran trabajando actualmente la propuesta de *Ludoteca Interactiva de Matemáticas* ponen en práctica las estrategias didácticas basadas en la utilización de material manipulable y cuáles han sido las principales problemáticas al trabajar ecuaciones lineales.

Para su desarrollo la Dirección de Investigación Educativa y Fortalecimiento Institucional contó con la participación activa de los directivos, docentes y alumnos de escuelas secundarias de diferentes municipios del estado de Hidalgo, sin embargo los resultados que se presentan solo considera la zona escolar Número 18 de secundarias generales ubicada en Zimapán, Hgo.

Considerando lo anterior, el estudio permitirá tomar acciones que logren la optimización y operatividad de la propuesta en las escuelas involucradas en el estudio y fortalecerá aquellos aspectos que no se hayan considerado de forma inicial.

### **5. Metodología**

Se diseñaron instrumentos de recolección de información como: encuestas para docentes y alumnos con preguntas abiertas y cerradas, con el propósito de que permitieran al docente especificar de forma más amplia sobre los cuestionamientos realizados; entrevistas a profundidad

y observaciones de clase a los grupos, con la finalidad de obtener información sistemática y útil que diera a conocer las problemáticas.

El estudio da inicio durante los meses de septiembre y octubre de 2005, con un muestreo de las escuelas secundarias de la Zona Escolar Número 18 de Secundarias Generales en donde opera la propuesta, el acercamiento a estas instituciones fue a través de la realización de cursos taller los cuales fueron el medio para poder conjuntar a los docentes de lugares retirados y de difícil acceso, al finalizar los cursos, se les aplicó encuestas y se realizaron entrevistas a los docentes, lo cual permitió tener una muestra mayor a la programada; además se programaron las visitas a las escuelas para realizar el trabajo con los alumnos.

Los ejes conductores de este estudio estuvieron orientados por los siguientes objetivos:

- Conocer si los docentes ponen en práctica estrategias didácticas en la enseñanza y el aprendizaje apoyadas de materiales manipulables propuestos por ludoteca.
- Identificar que materiales manipulables son utilizados durante el proceso enseñanza-aprendizaje de contenidos relacionados al álgebra.
- Obtener información sobre los docentes que fundamentan su práctica educativa en el enfoque actual de la enseñanza y el aprendizaje de las matemáticas.

## 6. Resultados

Se presentan los resultados obtenidos en el estudio los cuales muestran un panorama de aquellos docentes que imparten la asignatura de matemáticas que fundamentan su práctica en la utilización de estrategias didácticas y el apoyo con materiales manipulables, la información se presenta de forma gráfica y en comentarios generales sobre cuestionamientos abiertos que se hicieron directamente a los docentes: Es importante reconocer que en el quehacer docente y el uso de materiales manipulables favorece el libre acceso al conocimiento y al desarrollo de habilidades que plantea los planes y programas de estudio en educación básica; centra a la enseñanza de las matemáticas como una herramienta útil y como un objeto de conocimiento sujeto a cuestionamiento, análisis y experimentación.

El estudio parte de este supuesto en donde el uso de estrategias didácticas y de materiales manipulables son elementos esenciales en el logro de los propósitos de la asignatura, ya que se plantean recomendaciones para que las matemáticas puedan disfrutarse, fomentar el gusto por ellas y propiciar el logro de aprendizajes significativos. La información obtenida de las encuestas refleja como una problemática, el que los docentes utilizan los materiales manipulables durante el proceso enseñanza aprendizaje, pero no cuentan con suficientes materiales que les permitan abordar temas relacionados con el álgebra. Sin embargo dentro de las entrevistas realizadas se pudo identificar que no cuentan con la capacitación requerida sobre la aplicación de los materiales manipulables, los cursos que se han realizado a nivel zona y que han retomado el uso de estos materiales se han dado de forma parcial por lo que el proceso metodológico de los materiales manipulables en su totalidad no se conoce. Con relación al proceso de enseñanza y aprendizaje de las matemáticas y la puesta en práctica del enfoque didáctico se obtuvo la siguiente información: en términos de aprendizaje se identificó que es poco frecuente la relación de lo transmitido por los docentes con la vida cotidiana o su utilidad, así como el planteamiento y

resolución de problemas, por lo que el fortalecer estos aspectos permitirá a los alumnos identificar la utilidad que se obtiene de ellos dentro del ambiente en el que se desarrollan.

En este sentido una de las preguntas que se realizó a docentes y alumnos fue la relacionada a la enseñanza y el aprendizaje del álgebra, la cual para los docentes la consideran de difícil enseñanza, además de carecer del apoyo de materiales manipulables que les permita facilitar el proceso enseñanza y aprendizaje de tales temas y para los alumnos se les complica entenderlos, así como les resulta difícil explicar su uso en la vida cotidiana.

Respecto a la utilización de estrategias didácticas y de materiales manipulables estos se plantearon como un punto central en las entrevistas y encuestas aplicadas a los docentes con el siguiente cuestionamiento: ¿La metodología basada en el uso de estrategias didácticas favorece un aprendizaje significativo en los estudiantes de educación básica?

Con relación a este cuestionamiento, se recuperó que en su mayoría los docentes consideran que la utilización de estrategias didácticas apoyada de materiales manipulables favorece que los conocimientos se adquieran de forma objetiva; además de que la utilización del juego didáctico es una forma de apoyar el proceso de enseñanza y aprendizaje de matemáticas, permitiendo el desarrollo de habilidades y la socialización de conocimientos de los educandos. Otro de los cuestionamientos realizados a los docentes fue: ¿Qué tan importante es la utilización de materiales manipulativos en la introducción, reforzamiento y evaluación de conocimientos?

Este cuestionamiento permitió que los docentes aportaran las siguientes respuestas: Estimula sus conocimientos, actitudes y habilidades. Al manipular los materiales juegan, se divierte y les interesan más las matemáticas. Facilita la comprensión de temas y el aprendizaje significativo. Propicia la reflexión y el razonamiento haciéndolos prácticos. Porque se apega a nuestras necesidades de enseñanza. No contesto.

Con base en estas respuestas, y como se hizo mención anteriormente que es un estudio complementario ya que su realización buscó identificar si los docentes utilizan estrategias didácticas para la enseñanza de las matemáticas, específicamente en contenidos de Álgebra y en este sentido apoyarlos con estrategias didácticas que propone Ludoteca.

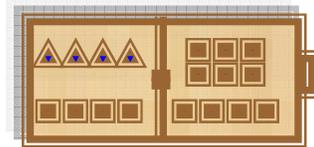
## 7. Propuesta

Una vez identificadas las problemáticas en la aplicación de estrategias didácticas para la enseñanza de las matemáticas por parte de los docentes y con la finalidad de apoyarlos con algunas de ellas que permiten la introducción de la enseñanza de ecuaciones lineales en educación secundaria, se da a conocer una que forma parte del acervo de materiales con los que cuenta ludoteca. El material referido se llama *tablero de fichas* el cual fue diseñado por docentes que imparten la asignatura de Matemáticas en la Escuela Secundaria Técnica No. 18, ubicada en Xuchitlán, municipio de San Salvador, Hidalgo.

La finalidad de este material manipulativo es introducir el contenido de ecuaciones lineales de forma significativa para el alumno, en donde él pueda identificar el proceso de cómo resolver una ecuación lineal y como llegar al procedimiento convencional a través del planteamiento y

resolución de problemas: tiene un proceso metodológico en el cual se debe organizar al grupo en equipos, entregarles el material y darles a conocer el valor de cada una de las fichas que se utilizan.

Y a través del planteamiento y resolución de problemas el alumno utiliza el “Tablero de Fichas” para resolverlo y posteriormente a su representación se llega a la forma convencional de resolver una ecuación lineal y al mismo tiempo el alumno identifica la utilidad de las ecuaciones lineales al plantear y resolver problemas. Se presenta el Modelo *Tablero de Fichas* y su procedimiento metodológico.



#### *Valores de las fichas*

Las fichas cuadradas representan los valores positivos, y las rojas los negativos; las fichas triangulares representan las incógnitas.

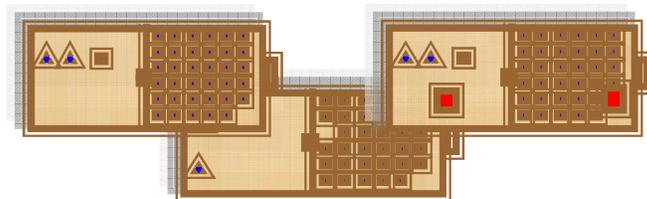


En el lado izquierdo del tablero se coloca el primer miembro de la ecuación y en lado derecho el segundo miembro. Para resolver la ecuación se procede a despejar la incógnita, aplicando las propiedades fundamentales de la igualdad a ambos miembros de la ecuación para que la igualdad permanezca y la incógnita quede sola. Para eliminar una incógnita o una variable se procede a colocar el mismo número de fichas de un mismo color en ambos lados del tablero, es decir si se colocan 3 fichas cuadradas rojas de un lado se tendrán que colocar otros 3 fichas cuadradas rojas del otro lado del signo igual y al término de este deben quedar triángulos de un solo lado y cuadrados en el otro. Para obtener la solución de la ecuación, se dividen las literales entre las incógnitas. Utilización de modelos manipulativos donde el juego didáctico presenta una forma diferente y agradable de dar solución a un determinado problema presentado a los alumnos:

#### *Ejemplo de problema:*

¡ADIÓS MIS 35 PALOMAS!

- Dijo un gavián a una parvada de palomas.
  - No somos 35 -Contestó una de ellas, y agregó:
  - Somos estas, más otras tantas como estas más usted, señor gavián, las 35 serán.
- ¿Cuántas palomas había en la parvada?



### *Procedimiento Convencional*

$$\begin{array}{rcl}
 x + x + 1 & = & 35 \\
 2x + 1 & = & 35 \\
 2x + 1 - 1 & = & 35 - 1 \\
 2x & = & 34 \\
 x & = & 34/2 \\
 x & = & 17
 \end{array}$$

## 8. Conclusiones

Es importante hacer mención que en el quehacer docente el uso de estrategias didácticas y material manipulable favorece el libre acceso al conocimiento y al desarrollo de habilidades que plantea los planes y programas de estudio en educación básica; centrando a la enseñanza y el aprendizaje de las matemáticas como una herramienta útil y como un objeto de conocimiento sujeto a cuestionamiento, análisis, experimentación, construcción e innovación.

En este sentido la información obtenida permitió detectar cuales han sido las principales problemáticas a las que se enfrentan los docentes al abordar contenidos relacionados al Álgebra con el apoyo de estrategias didácticas y materiales manipulativos; por lo que se han puesto en práctica algunas acciones en el ciclo escolar 2006-2007 apoyando a los docentes con material escrito y manipulativo que les permita fortalecer el proceso enseñanza-aprendizaje, así mismo se han planteado estrategias de reproducción de materiales manipulativos de bajo costo que permita al docente elaborarlos con los estudiantes. Se espera que ésta información permita reorganizar y reforzar el trabajo en las escuelas con la operatividad adecuada de la propuesta *Ludoteca Interactiva de Matemáticas* logrando así aprendizajes significativos que permita elevar la calidad de la educación.

## 9. Referencias

- [1] NUNES Terezinha y BRYANT Peter.- (1998).- *Las matemáticas y su aplicación: La perspectiva del niño*. Editorial siglo XXI editores.
- [2] RESNICK Lauren B.- (1991).- *La enseñanza de las matemáticas y sus fundamentos Psicológicos. Temas de educación* Paidós. Ministerio de Educación y Ciencia.
- [3] POLYA G.- (1996).- *Como plantear y resolver problemas. Serie Matemáticas*. Editorial Trillas.
- [4] DÍAZ, Vega José Luís.- (2002).- *El juego y el juguete en el desarrollo del niño*.-Trillas, México.
- [5] SEP.- (2001).- *Libro para el Maestro de Matemáticas. Educación Secundaria*, Comisión Nacional de Libros de Texto Gratuitos, Talleres de Editorial Offset, S. A. de C. V.- México.
- [6] SEP.- (1994).- *Plan y Programas de estudio 1993. Educación Básica. Secundarias*.- Editorial Offset, S. A. de C. V.- México.

# El análisis funcional: una perspectiva histórica

Oswaldo González Gaxiola  
Universidad Autónoma Metropolitana-Cuajimalpa

## Resumen

*El análisis funcional ha sido definido por J. Dieudonné como “El estudio de los espacios vectoriales topológicos y de las aplicaciones definidas entre subconjuntos de los mismos, sujetas a distintas condiciones algebraicas y topológicas.” Esta definición, aunque formalmente correcta, tiene la desventaja de estar hecha desde el punto de vista actual, cuando la teoría está muy desarrollada y metodológicamente organizada, pero no dice nada sobre los contenidos concretos y el tipo de problemas que dieron origen al análisis funcional. En el presente trabajo se dará un breve panorama histórico del surgimiento del análisis funcional además se hablará sobre la clase de problemas que originaron esta rama de las matemáticas así como de los personajes involucrados en el establecimiento de dicha teoría.*

## 1. Introducción y antecedentes

El *análisis funcional* ha sido definido por J. Dieudonné como “El estudio de los espacios vectoriales topológicos y de las aplicaciones definidas entre subconjuntos de los mismos, sujetas a distintas condiciones algebraicas y topológicas”. Esta definición, aunque formalmente correcta, tiene la desventaja de estar hecha desde el punto de vista actual, cuando la teoría está muy desarrollada y metodológicamente organizada, pero no dice nada sobre los contenidos concretos y el tipo de problemas que dieron origen al análisis funcional.

Dentro de su ambigüedad, la definición que da Dieudonné pone de manifiesto algunas de las características más importantes del análisis funcional: la tendencia hacia la algebrización del análisis, el énfasis en resultados de carácter estructural y la fuerte influencia de la topología. De hecho, como el propio Dieudonné señala, es prácticamente imposible dissociar los comienzos de la topología general y del análisis funcional. En cualquier caso, como toda otra teoría matemática, el análisis funcional surge de la necesidad de encontrar nuevas técnicas para abordar una gran cantidad de problemas que los métodos tradicionales no podían resolver.

Ya desde el comienzo del cálculo diferencial fue poniéndose de manifiesto la conveniencia de considerar conjuntos cuyos elementos, a diferencia de lo que sucede en el análisis clásico, no son puntos del espacio euclídeo ordinario, sino funciones. Y éste es el origen mismo del nombre de análisis funcional; el estudio de los espacios funcionales, es decir, conjuntos formados por funciones, dotados de determinadas estructuras que permiten realizar en ellos gran parte de las operaciones habituales del análisis. Sin embargo, no hay duda de que se pueden hallar antecedentes claros del análisis funcional desde el mismo comienzo del cálculo diferencial, pues el estudio de las ecuaciones diferenciales lleva inmediatamente a la necesidad de considerar el conjunto de las soluciones y, eventualmente, al estudio de sus propiedades.

## 2. Pasar de lo finito a lo infinito

Como ya hemos dicho, uno de los rasgos distintivos del análisis funcional, es la algebraización del análisis. Los métodos algebraicos se han desarrollado casi siempre antes que los analíticos y, al considerar esencialmente conjuntos finitos, suelen ser más fáciles de usar. Por ello, una idea reiteradamente utilizada por los analistas ha sido la de considerar las ecuaciones funcionales como casos límites de ecuaciones algebraicas, cuya solución es más sencilla. Así, por ejemplo, la deducción que hace D. Bernoulli en 1750 de la solución general del problema de la cuerda vibrante, se basa en sustituir la cuerda por  $n$  masas puntuales, calcular la posición general del sistema a lo largo del tiempo y hacer tender formalmente  $n$  a infinito.

El descubrimiento por D' Alembert de la ecuación diferencial que rige el movimiento y el desarrollo de las técnicas analíticas relegó el método de Bernoulli a un segundo plano. Sin embargo, la idea persistió y tuvo una influencia decisiva en los trabajos sobre física de Lagrange y, sobre todo, de Fourier, para la obtención de las ecuaciones diferenciales que describen la transmisión del calor. Al mismo tiempo esta idea, del paso de lo finito a lo infinito, fue sistemáticamente utilizada por Fourier para la obtención de soluciones correctas. Pero mejor será ilustrar un poco el método con un ejemplo, de los muchos que aparecen en la *La théorie analytique de la chaleur* (1822): consideremos el problema de encontrar una solución de la ecuación diferencial

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, & x > 0, \quad -\frac{\pi}{2} < y < \frac{\pi}{2} \\ u(0, y) = 0, \\ u(x, \pm \frac{\pi}{2}) = 0, & x \rightarrow \infty. \end{cases}$$

Se trata de un modelo matemático de la temperatura estacionaria en el interior de una placa infinita de forma rectangular, cuyos bordes se mantienen a la temperatura prefijada.

La solución que Fourier obtiene es

$$u(x, y) = \sum_{n=1}^{\infty} a_n u_n(x, y),$$

donde  $a_n = \frac{4}{\pi} \frac{(-1)^{n-1}}{2n-1}$  y  $u_n(x, y) = e^{-(2n-1)x} \cos(2n-1)y$ ,  $n \in \mathbb{N}$ .

El mismo Fourier no está muy convencido de su forma de proceder para hallar la solución y dice: “Como estos resultados parecen desviarse de las consecuencias ordinarias del Cálculo, es necesario examinarlos con cuidado e interpretarlos en su verdadero sentido”. Y prueba directamente que la suma de la serie obtenida para  $x = 0$  es constante e igual a 1 en el intervalo señalado y finalmente afirma que  $u$  es solución de su problema.

## 3. El problema de Sturm-Liouville o el comienzo de la teoría espectral

Los trabajos de Fourier influyeron decisivamente en el tratamiento posterior de las ecuaciones diferenciales. El método de separación de variables, aplicado a otras ecuaciones diferenciales (en general, no homogéneas), conduce al estudio de la ecuación diferencial ordinaria de segundo orden

$$y'' - q(x)y + \lambda y = 0,$$

donde  $\lambda$  es un parámetro complejo,  $q(x)$  es real y la función incógnita  $y$  es por lo menos dos veces diferenciable en algún  $(a, b)$  y satisface las condiciones de frontera:

$$\alpha_1 y(a) + \beta_1 y'(a) = 0, \quad \alpha_2 y(b) + \beta_2 y'(b) = 0.$$

Ch. Sturm (1836) y J. Liouville (1837) desarrollaron una teoría general para abordar este tipo de problemas. Los resultados obtenidos tuvieron una gran influencia en el desarrollo posterior. La contribución principal de Sturm fue la demostración de que el problema planteado sólo tiene solución para una sucesión estrictamente creciente  $\{\lambda_n\}$ , valores reales del parámetro  $\lambda$  (los valores propios del problema), con lo que se sientan las bases de la teoría espectral. Las propiedades de ortogonalidad de las correspondientes funciones propias  $\{u_n\}$ , llevaron a Liouville a tratar de generalizar el desarrollo en serie de Fourier, y expresar cualquier función continua  $u$  como una serie  $\sum a_n u_n$ , donde

$$a_n = \frac{\int u \cdot u_n}{\int u_n^2}.$$

Liouville logra demostrar la convergencia de la serie, siempre que la serie de Fourier de  $u$  sea convergente.

#### 4. El cálculo de variaciones y el problema de Dirichlet

Probablemente los antecedentes más claros del análisis funcional se pueden encontrar en el cálculo de variaciones. Con este nombre se conoce a una serie de problemas en los que se trata de maximizar o minimizar no ya una función real definida sobre un subconjunto de  $\mathbb{R}^n$ , sino una expresión del tipo

$$J(\varphi) = \int_a^b F(\varphi(x), \varphi'(x), \dots) dx,$$

siendo  $F$  una función regular, y las variables  $\varphi$  un adecuado conjunto de curvas regulares parametrizadas en  $[a, b]$ . Es éste el contexto donde aparece primero la idea de *campo funcional*, como conjunto de funciones admisibles.

En el caso general, para estudiar problemas de extremos de funcionales de la forma dada en la ecuación anterior, se solía razonar por analogía al caso de funciones reales. Así por ejemplo, se solía admitir como evidente que si  $F$  estaba acotada,  $J$  alcanzaba su máximo o su mínimo en alguna función admisible. Sin embargo, la formalización del análisis iniciada por Weierstrass, puso pronto de manifiesto la debilidad de este argumento; uno de los primeros problemas fue estudiar condiciones bajo las cuales el límite (puntual) de una sucesión de funciones conserva las buenas propiedades que pudieran tener las funciones de la sucesión. Los primeros intentos en esta dirección, consistieron en poner condiciones más restrictivas sobre la forma de converger de la sucesión. Así surgió la noción de *convergencia uniforme* (Weierstrass, 1841; Stokes, 1847; Von Seidel, 1848; Cauchy, 1853). La postura de los italianos Dini, Ascoli y Arzelá, fue totalmente diferente; en lugar de modificar la noción de convergencia empleada, dieron una condición general sobre el conjunto formado por la sucesión de funciones (la *equicontinuidad*, Ascoli, 1883), de tal modo que el límite puntual es necesariamente continuo.

## 5. Las ecuaciones integrales y el desarrollo del análisis funcional

Las ecuaciones integrales, son probablemente el ejemplo más representativo de la necesidad de establecer al análisis funcional como un área de las matemáticas. A lo largo del siglo XIX se habían planteado algunas ecuaciones integrales involucradas en problemas de la física. Así por ejemplo, Abel había resuelto en 1823 la ecuación relacionada con la tautócrona:

$$f(x) = \int_0^x \frac{\phi(y)}{\sqrt{x-y}} dy.$$

Este es un ejemplo de *ecuación integral de primera especie* en notación de Hilbert, ya que la función incógnita  $\phi(x)$  aparece sólo bajo el signo de la integral. Otra importante clase de ecuaciones integrales aparece en relación con el método de Beer-Neumann para la solución del problema de Dirichlet. Fue en 1888 cuando P. du Bois-Reymond sugirió el nombre de ecuaciones integrales para designar a este tipo de problemas y propuso desarrollar una teoría general de tales ecuaciones como método alternativo para resolver ecuaciones diferenciales. Los primeros resultados en esta dirección, fueron obtenidos por J.M. Le Roux (1894) y V. Volterra (1896); ambos establecieron teoremas de existencia y unicidad mediante hipótesis adecuadas sobre el núcleo  $k$ , para ecuaciones del tipo:

$$f(x) + \int_a^x k(x,t)f(t)dt = g(x). \quad (1)$$

I. Fredholm, estudiante de Mittag-Leffler y más tarde su colega en Estocolmo, realizó una visita a París en 1899, entrando en contacto con Poincaré y otros famosos matemáticos de la época y en 1900 publicó una nota, titulada *Sur une nouvelle méthode pour la resolution du problème de Dirichlet*, completada más tarde por un artículo en Acta Mathematica que iban a provocar impacto en la comunidad matemática; en su trabajo Fredholm considera el método de Beer-Neumann y trata de resolver la ecuación integral (1).

## 6. La contribución de Hilbert, F. Riesz y M. Fréchet

Los sensacionales resultados de Fredholm se extendieron rápidamente y en 1901 E. Holgrem expone estos resultados en Göttingen, en el seminario de Hilbert, quien se interesó mucho por el tema y, entre 1904 y 1910, publicó seis artículos sobre ecuaciones integrales en el Göttingen Nachrichten, que fueron posteriormente reunidos en un libro (1912) que contiene varias aplicaciones tanto matemáticas como físicas. En dicho libro hay nociones y directrices novedosas que después, en manos de matemáticos como E. Schmidt y F. Riesz, van a convertirse en los fundamentos del análisis funcional. Hilbert también consideró el caso de formas cuadráticas no necesariamente completamente continuas, sino simplemente continuas (o equivalentemente, acotadas sobre la bola unitaria) e incluso formas no acotadas, estableciendo así el germen de la teoría espectral moderna. De hecho en el trabajo de Hilbert aparecen ya algunas de las clases más importantes de operadores (de Hilbert-Schmidt, nucleares, etc.) aunque realmente la formulación moderna se debe a F. Riesz. Es evidente, en una perspectiva actual, que estos resultados están prefigurando la teoría de los espacios de Hilbert.

Por otro lado, en 1906 aparece también la famosa tesis doctoral de M. Fréchet “*Sur quelques points du calcul fonctionnel*”, en ella Fréchet introduce la noción abstracta de distancia en un conjunto, lo que permite extender las nociones habituales de entornos, límites, continuidad, compacidad, completitud y separabilidad en conjuntos abstractos y las estudió en espacios funcionales teniendo esto una tremenda influencia para el desarrollo del análisis funcional moderno. Por esta época los jóvenes matemáticos E. Fischer y F. Riesz, asumieron la visión geométrica y topológica de los espacios de Hilbert, lo que les llevó a descubrir en 1907 (independientemente) el llamado *teorema de Fischer-Riesz*, que a su vez establece una inesperada relación de estos temas con otro gran descubrimiento de la época: la teoría de integración de Lebesgue y en el mismo año Fréchet y Riesz (independientemente), obtienen la representación de cualquier forma lineal continua  $T$  sobre el espacio  $L^2$  en la forma

$$T(f) = (f, g) = \int f(x)g(x)dx, \quad \text{para alguna } g \text{ del mismo espacio.}$$

Dos años más tarde, en 1909, Riesz resuelve completamente el problema abordado por Hadamard y Fréchet en 1903 y 1904, probando que cualquier funcional lineal y continuo  $T$  sobre el espacio  $C([a, b])$ , puede escribirse en forma de integral de Stieljes:

$$T(f) = \int_a^b f(x)d\alpha(x), \quad \text{donde } \alpha \text{ es una función de variación acotada.}$$

En 1910, Riesz introduce los espacios  $L^p$ , para  $1 < p < \infty$ , como generalización natural de  $L^2$ . Se plantea Riesz la resolución de un sistema de infinitas ecuaciones del tipo

$$\int_a^b f_i(x)g(x)dx = c_i, \quad (i \in I),$$

donde las  $f_i$  y los escalares  $c_i$  son los datos, y se trata de encontrar una solución  $g$ .

## 7. Conclusiones

Las contribuciones fundamentales que hemos venido refiriendo, preparan el camino para el desarrollo de una teoría general de espacios normados, funcionales y operadores lineales entre ellos y esto aconteció en 1920 cuando S. Banach, en su tesis doctoral, presentó la definición axiomática de los espacios que hoy llevan su nombre (dos años después publicada en *Fundamenta Mathematicae*). Desde la introducción de su trabajo Banach declara su intención de demostrar una serie de resultados válidos en distintos campos funcionales, para lo cual establece un conjunto de teoremas en un marco muy general que dan lugar a los distintos resultados buscados. El marco general en cuestión es precisamente lo que hoy conocemos como espacios normados completos (espacios de Banach); da la definición axiomática de espacio vectorial real, normado y completo y la tesis contiene, entre otros, el *principio de acotación uniforme* y la forma general del *principio de contracción* en espacios métricos completos. Finalmente, en 1922, P. Levy, alumno de Hadamard publica el libro *Lecons d'Analyse Fonctionnelle*, en donde por primera vez aparece el nombre de análisis funcional y con este breve recorrido por la historia contemporánea podemos considerar, que termina el proceso fundacional de esta rama de las matemáticas.

## 8. Referencias

- [1] Birkhoff G. ; Kreyszig E. *The Establishment of Functional Analysis*; Historia Math; 11, (1984), 258-321.
- [2] Bourbaki N. *Elementos de Historia de la Matemática*; Alianza Ed. Madrid, 1972.
- [3] Dieudonné J. *History of Functional Analysis*; North Holland, Amsterdam; 1981.
- [4] Temple G. *100 Years of Mathematics*; Duckworth, Londres, 1981.

# Resultados de la implementación de tareas y exámenes en línea para los cursos de Álgebra, utilizando el software Maple T. A.

Ana Gpe. Del Castillo Bojórquez y Blanca Evelia Flores Soto  
Departamento de Matemáticas  
Universidad de Sonora  
acastillo@gauss.mat.uson.mx, bflores@gauss.mat.uson.mx.

## Resumen

*El propósito de este trabajo es mostrar algunos resultados relacionados con el diseño e implementación de tareas y exámenes en línea con el software Maple T.A., para los cursos de Álgebra en Ciencias e Ingeniería de la Universidad de Sonora, durante el semestre 2008-2, como una alternativa de apoyo al trabajo extra-clase de los estudiantes, al trabajo docente y, particularmente, a la evaluación y auto-evaluación. Este trabajo emerge de las actividades del proyecto “Seguimiento de la Impartición de los Cursos de Álgebra bajo el Esquema del Nuevo Modelo Curricular de los Programas de la División de Ingeniería de la Universidad de Sonora” que se enmarca dentro del “Programa Permanente de Mejoramiento de los Servicios que Ofrece el Departamento de Matemáticas”.*

## 1. Introducción

Dentro de las actividades del proyecto “Seguimiento de la Impartición de los Cursos de Álgebra bajo el Esquema del Nuevo Modelo Curricular de los Programas de la División de Ingeniería de la Universidad de Sonora”, se iniciaron reflexiones sobre la evaluación de la calidad de la enseñanza y aprendizaje de los contenidos del curso de álgebra, motivadas por la solicitud del Departamento de Ingeniería Industrial de incorporar, a su esquema de exámenes departamentales, las materias que ofrece el Departamento de Matemáticas. Aunado a esto, se tuvo acceso a un sistema en línea de evaluación y entrenamiento en matemáticas Maple T.A. adquirido dentro del proyecto “Homogenización y Certificación de los Programas de Matemáticas del Sistema de Educación Pública en Sonora”.

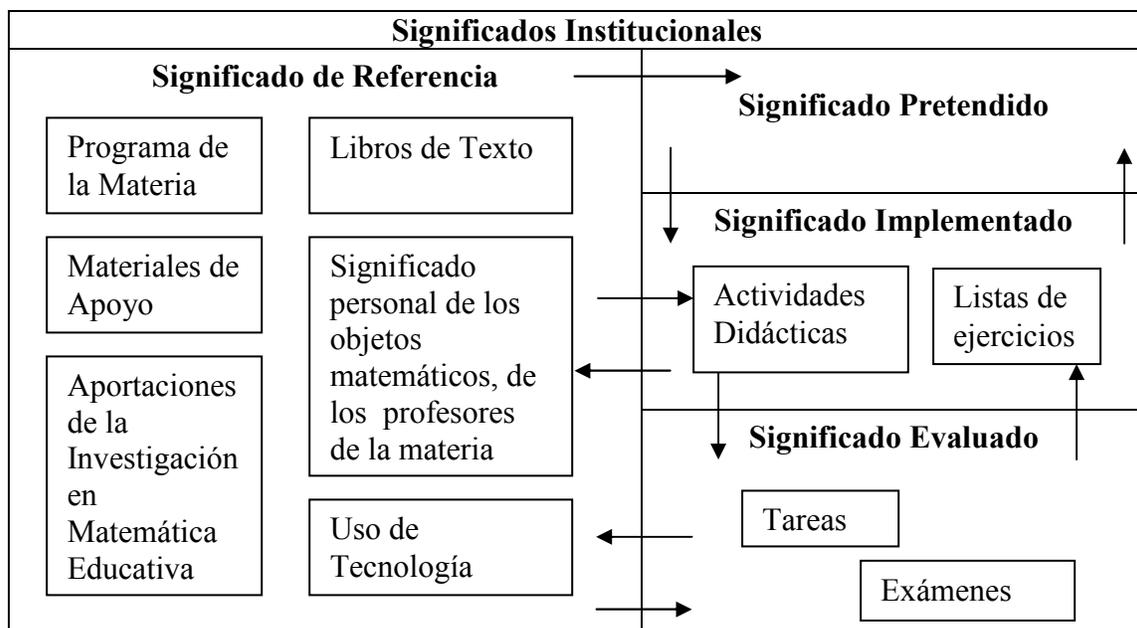
El objetivo de este trabajo es mostrar algunos resultados relacionados con el diseño e implementación de tareas y exámenes departamentales en línea para los cursos de Álgebra del área de Ciencias e Ingeniería, con el uso de Maple T.A., como una alternativa de apoyo al trabajo extra-clase de los estudiantes, al trabajo docente y, particularmente, a la evaluación y auto-evaluación.

Es importante enfatizar que el grupo de profesores participantes, tenía dos años de trabajo previo, en el análisis del programa de la materia, los materiales disponibles, uso de tecnología, y compartiendo experiencias y actividades didácticas implementadas en el aula.

## 2. Aspectos teóricos

En este trabajo se han tomado en consideración algunos elementos teóricos del Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (EOS), desarrollado por Godino (2003). Entre los elementos considerados, podemos mencionar los objetos personales e institucionales, sus significados sistémicos, los elementos básicos del significado y las relaciones que se establecen entre ellos (funciones semióticas). En la siguiente tabla se muestran los distintos tipos de significados institucionales considerados, así como las posibles interacciones entre ellos:

Tabla 1. Significados Institucionales.



Para la evaluación del aprendizaje a través de un sistema en línea, se considera necesario un sistema de categorías para clasificar los diferentes niveles de preguntas y respuestas esperadas. Tomamos como base la taxonomía SOLO (*Structure of the Observed Learning Outcome*), (Collis y Biggs, 1982) pero reinterpretemos los niveles básicos de respuesta, en términos de algunos elementos del EOS.

Esta taxonomía incluye cinco niveles básicos de respuesta que, en orden de complejidad creciente, son:

- **Pre-estructural:** el estudiante tiene parte de la información, ninguna organización y ni sentido. (Vacíos de significación, disparidad de interpretaciones).
- **Uni-estructural:** el estudiante puede establecer relaciones directas simples entre elementos básicos del significado, de manera aislada, sin conexión unas con otras.

- **Multi-estructural:** el estudiante puede establecer varias relaciones entre elementos básicos del significado, considerándolas en una secuencia.
- **Relacional:** el estudiante puede establecer una red de relaciones entre elementos básicos del significado e integrarlas en un todo.
- **Abstracción Extendida:** los estudiantes pueden establecer relaciones más allá del tema inmediato; generalizan y transfieren los principios de lo específico a lo abstracto.

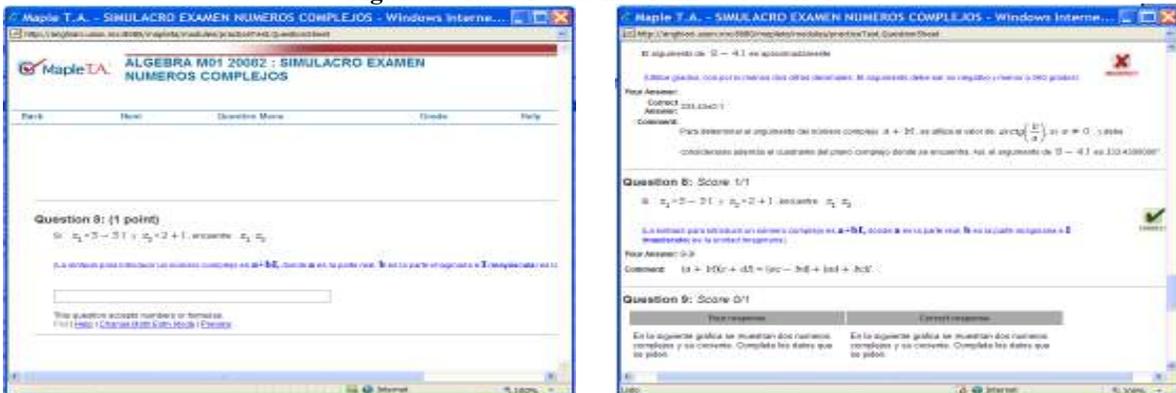
### 3. Diseño de reactivos en Maple T.A.

Con base en la taxonomía presentada y analizando problemas, listas de ejercicios y exámenes propuestos por los profesores participantes en el proyecto, se dio inicio al diseño de reactivos a implementar con el sistema Maple T.A. Se propuso elaborar reactivos cuya respuesta esperada correspondiera a los niveles uni-estructural, multi-estructural o relacional.

*Ejemplo de reactivo de nivel uni-estructural:*

En este reactivo, se pide calcular el producto de dos números complejos dados. La ventaja de utilizar Maple T.A. es que este reactivo se diseña de forma algorítmica, de modo que se programa el producto de los números  $(a + bi)(c + di)$ , donde las variables  $a$ ,  $b$ ,  $c$  y  $d$  toman valores en un rango especificado. Así, con un solo reactivo en Maple T.A. se tienen muchas versiones del mismo ejercicio.

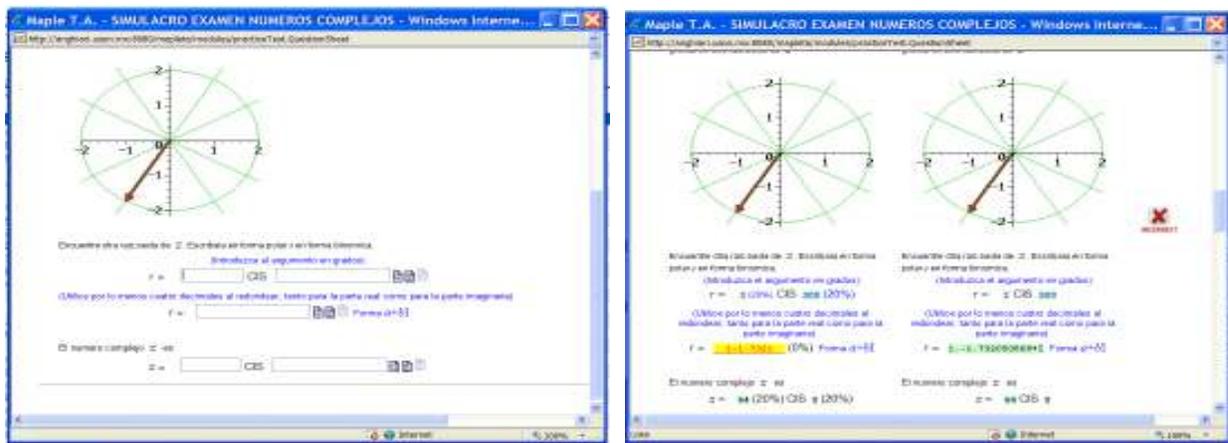
Figura 1. Reactivo de nivel uni-estructural.



*Ejemplo de reactivo de nivel relacional:*

En este reactivo, se presenta gráficamente una raíz  $n$ 'ésima (cuadrada, cúbica, cuarta, quinta o sexta) de un número complejo desconocido. Se pide al estudiante encontrar otra de tales raíces, representarla en forma polar y cartesiana, y encontrar el número complejo del cual es raíz.

Figura 2. Reactivo de nivel uni-estructural.



#### 4. Diseño de tareas y exámenes en línea

En el sistema Maple T.A. es posible diseñar diferentes tipos de tareas y exámenes. Los reactivos pueden ser seleccionados fija o aleatoriamente, de uno o varios grupos de ellos, a partir de los bancos de reactivos. De este modo, a partir de un mismo diseño de examen, es posible generar un número importante de versiones del mismo.

En este trabajo se utilizaron tres tipos distintos:

- Prácticas anónimas: Son tareas tipo examen que el estudiante puede realizar libremente, sin límite de tiempo, el número de veces que desee, y tienen la característica de que, aunque el sistema evalúa su trabajo, éste no queda registrado en el sistema. El estudiante tiene acceso a este tipo de tareas desde cualquier computadora con Internet.
- Tareas registradas: Son tareas tipo examen que, a diferencia de las anteriores, quedan registradas en el sistema, tanto el trabajo del estudiante como el puntaje obtenido. El estudiante tiene acceso a este tipo de tareas desde cualquier computadora con Internet. Se ha tomado el acuerdo de que pueden realizarla hasta tres veces, tomándose en consideración la puntuación más alta, como parte de la calificación del parcial correspondiente.
- Exámenes departamentales supervisados: Este tipo de examen requiere la autorización del profesor para que cada estudiante en particular tenga acceso a él, por lo que se lleva a cabo de manera presencial en un laboratorio de computadoras.

#### 5. Prueba piloto semestre 2008-2

Aunque se llevaron a cabo pruebas piloto durante los semestres 2007-2 y 2008-1, fue durante el semestre 2008-2, donde se realizaron de una manera más sistemática. Al inicio del semestre se

ofreció un taller de capacitación a los profesores para el manejo del sistema. Se creó un curso raíz en el que se cargaron los bancos de reactivos, tareas y exámenes y, a partir de éste, se generaron los cursos de los demás profesores, de modo que heredaran los contenidos del curso raíz.

Se elaboraron seis bancos de reactivos: uno de ellos para el diseño de un examen de diagnóstico al inicio del semestre, y cinco relacionados con los contenidos del curso de álgebra.

**Figura 3. Bancos de reactivos en Maple T.A.**

Question Bank	Questions	In Use	Last Saved
1 HSB_ECUACIONES ARITMETICAS Y ALGEBRAICAS	15	delete	yes 2/10/08
2 NUMEROS COMPLEJOS	43	delete	yes 4/10/08
3 ECUACIONES DE SEGUNDO Y TERCER GRADO	36	delete	yes 8/10/08
4 POLINOMIOS	32	delete	yes 19/11/08
5 SISTEMAS ECUACIONES LINEALES	42	delete	yes 8/12/08
6 MATRICES	29	delete	yes 24/11/08

Estos bancos deberán ampliarse y reestructurarse, con base en los resultados obtenidos. También está pendiente la elaboración de otros dos bancos de reactivos: uno relacionado con Transformaciones Lineales, y otro con Valores y Vectores Propios.

A partir de los bancos de reactivos, se diseñaron 15 prácticas anónimas, 13 tareas registradas y cuatro exámenes departamentales supervisados (tres ordinarios y uno extraordinario).

Para el semestre en cuestión, se formaron 31 grupos de Álgebra en Ciencias e Ingeniería en la Universidad de Sonora, Unidad Regional Centro, los cuales estuvieron a cargo de 15 profesores. Trece profesores se involucraron en el proyecto, pero sólo nueve aplicaron las tareas y exámenes, de manera sistemática, a un total de 17 grupos.

## 6. Aspectos importantes a destacar

Durante la experiencia de aplicar exámenes en línea durante el semestre 2008-2, se puede destacar que:

Concebimos el sistema Maple T.A. como un fuerte apoyo a la actividad del estudiante fuera del salón de clases, con la ventaja de brindar retroalimentación inmediata a las acciones llevadas a cabo por él mismo, para la solución de los problemas y ejercicios que se les presentan. Sin embargo, consideramos necesario llevar a los estudiantes a un centro de cómputo, al inicio del semestre resolviendo problemas de acceso y para que se familiaricen con el ambiente general de Maple T.A. En este punto, se tuvo la desventaja de contar con laboratorios de a lo más 30 computadoras, mientras que los grupos contaban generalmente con 40 estudiantes. Por esta razón, la aplicación de los exámenes departamentales supervisados, se llevó a cabo en dos sesiones por grupo.

Cada vez que un estudiante selecciona una tarea, se le presenta una versión distinta, con reactivos seleccionados aleatoriamente, y éstos, a su vez, son algorítmicos, es decir, incluyen parámetros que cambian cada vez que aparecen en una tarea. Así, podemos decir que las tareas son individualizadas. Esto redujo enormemente la posibilidad de copiar las respuestas de los compañeros y promovió, en cambio, el intercambio de estrategias y métodos de solución.

El uso de prácticas anónimas, promovió que los estudiantes formaran equipos y respondieran entre todos una misma tarea, ya que ésta no era registrada para nadie. También, les dio oportunidad de resolver varias versiones de una misma tarea, antes de realizar la que quedaría registrada en el sistema.

Algunos estudiantes, al realizar varias veces la “misma tarea” y recibir retroalimentación inmediata del sistema, identificando sus aciertos y errores, se concentraron sobre sus deficiencias y fueron capaces de plantear preguntas más específicas durante las sesiones de clase.

Pero también hubo alumnos que expresaron su descontento por este tipo de exámenes con los que no se sentían del todo cómodos, no realizaron las prácticas anónimas ni tareas, y sólo se presentaron al examen departamental, a veces sin recordar su nombre de usuario y contraseña. Ellos aludieron su bajo rendimiento a la utilización del sistema. Entre los comentarios adversos de los estudiantes, podemos mencionar:

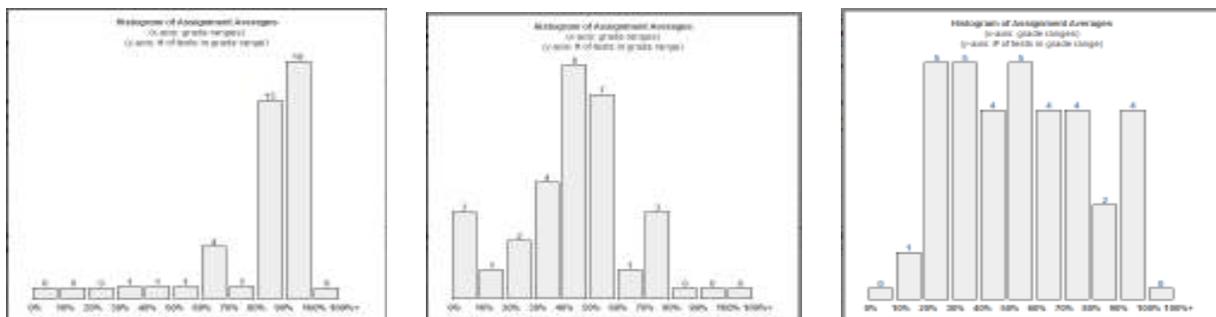
- a. Los menús para la navegación en el sistema están en inglés y se sentían en desventaja.
- b. Tienen problemas con la sintaxis utilizada en las preguntas abiertas.
- c. El sistema no siempre califica bien respuestas equivalentes.
- d. Su práctica se interrumpió por problemas de conectividad en Internet.
- e. El sistema presentó inhabilitaciones y saturación teniendo problemas de acceso.

Creemos que una mejor comunicación entre alumnos y profesores puede ayudar a llevar a cabo las acciones pertinentes para que se tenga una mayor aceptación por parte de estos estudiantes.

Los estudiantes que realizaron las prácticas anónimas y tareas registradas, obtuvieron mejores resultados en los exámenes departamentales. A nivel de grupo, consideramos muy importante el papel del profesor para promover de manera sistemática la realización de estas tareas.

Se muestran los resultados obtenidos por tres grupos en el primer examen parcial, en una de las formas mostradas por el sistema Maple T.A.

**Figura 4. Histogramas de resultados de tres grupos.**



## 7. Referencias

- [1] Godino, J. D. (2003). *Teoría de las funciones semióticas. Un enfoque ontológico-semiótico de la cognición e instrucción matemática*. Departamento de Didáctica de la Matemática. Universidad de Granada. Recuperado de: [http://www.ugr.es/local/jgodino/indice\\_tfs.htm](http://www.ugr.es/local/jgodino/indice_tfs.htm).
- [2] Godino, J., Batanero, C. (1994) Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, Vol. 14, No. 3, pp. 325-355.
- [3] Biggs, J.B.; Collis, K.F.; (1982). *Evaluating the quality of learning. The SOLO taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press.
- [4] Huerta, P.; (1999). *Los niveles de Van Hiele y la taxonomía SOLO: "Un análisis comparado, una integración necesaria*. Departamento de Matemática de la Universidad de Valencia. *Enseñanza de las Ciencias*, (17(2), 291-309).
- [5] Atherton, J. S.; (2005) *Learning and Teaching: SOLO taxonomy*. Recuperado el 20 de Enero de 2009 de la dirección electrónica <http://www.learningandteaching.info/learning/solo.htm>.
- [6] MapleSoft (2009), *Maple T.A.: Content Center*. Recuperado el 20 de Enero de 2009 de la dirección electrónica <http://www.maplesoft.com/tacontent/index.aspx>.



# Formulaciones equivalentes de la propiedad de Radon-Nikodým de un espacio de Banach

Martha Guzmán Partida y Marysol Navarro Burruel  
Departamento de Matemáticas  
Universidad de Sonora  
manavarro@correoa.uson.mx

## Resumen

*En este trabajo presentaremos dos formulaciones equivalentes al teorema de Radon-Nikodým en espacios de Banach. La primera, es la equivalencia con el teorema de representación de Riesz para operadores lineales y continuos. La segunda, con el teorema de Fatou para funciones armónicas definidas en el disco unitario con valores en un espacio de Banach  $X$ , que asegura la existencia de límites radiales ctp. de dichas funciones. Finalmente, exhibimos un ejemplo de una función vectorial armónica y acotada que no tiene límites radiales ctp., es decir, el espacio donde toma valores no tiene la propiedad de Fatou.*

## 1. Introducción

J.A. Clarkson en 1936 y N. Dunford y A.P. Morse en el mismo año, crearon las nociones de convexidad uniforme y de base acotadamente completa de un espacio de Banach  $X$ , respectivamente, con el objeto de demostrar que toda función absolutamente continua definida en el espacio euclideo y con valores en  $X$  es la integral de su derivada. Estos resultados fueron reconocidos como auténticos teoremas de Radon-Nikodým para la integral de Bochner. De hecho, fueron los primeros teoremas de tipo Radon-Nikodým para medidas vectoriales en espacios de medida abstractos.

En la primera sección del presente trabajo introducimos la noción de medida vectorial y algunos conceptos relacionados con ésta, así como la integral de Bochner la cual es una generalización directa de la integral de Lebesgue en el contexto de espacios de Banach. También veremos que, sin embargo, para la integral de Bochner es imposible establecer de manera general el teorema de Radon-Nikodým. La clase de espacios de Banach donde este teorema se verifica se conoce como la familia de espacios de Banach con la propiedad de Radon-Nikodým. En la segunda sección establecemos sin demostración que este teorema es equivalente al Teorema de Representación de Riesz. En la tercera sección presentamos la equivalencia con el teorema de Fatou para funciones armónicas vectoriales definidas en el disco unitario. Por último veremos que el espacio  $c_0$  no tiene la propiedad de Fatou; en realidad lo que ocurre es que este espacio no tiene la propiedad de Radon-Nikodým.

## 2. Medida vectorial

Empezaremos introduciendo la definición de medida vectorial, así como algunos conceptos relacionados. En todo lo que sigue,  $\Omega$  será un conjunto no vacío,  $\mathcal{F}$  será un álgebra de subconjuntos de  $\Omega$ ,  $\Sigma$  una  $\sigma$ -álgebra de subconjuntos de  $X$  y  $X$  un espacio de Banach.

## 2.1. Conceptos Preliminares

**Definición 1.** Diremos que una función  $F : \mathcal{F} \rightarrow X$  es una medida vectorial finitamente aditiva, o simplemente una medida vectorial, si para cada par  $E_1, E_2$  de elementos ajenos de  $\mathcal{F}$  se tiene que

$$F(E_1 \cup E_2) = F(E_1) + F(E_2).$$

Si además, cada vez que  $(E_n)_{n=1}^{\infty} \subset \mathcal{F}$  con  $E_i \cap E_j = \emptyset$  si  $i \neq j$  y  $\bigcup_{n=1}^{\infty} E_n \in \mathcal{F}$ , se tiene que

$$F\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} F(E_n),$$

diremos que  $F$  es una medida vectorial aditiva numerable o numerablemente aditiva.

**Definición 2.** Sea  $F : \mathcal{F} \rightarrow X$  una medida vectorial. La **variación de  $F$**  es la función  $|F| : \mathcal{F} \rightarrow \bar{\mathbb{R}}$  definida por

$$|F|(E) = \sup_{\pi} \sum_{A \in \pi} \|F(A)\|_X, \quad (1)$$

donde el supremo se toma sobre todas las particiones  $\pi$  de  $E$  en un número finito de elementos de  $\mathcal{F}$  ajenos por pares.

Evidentemente tenemos que  $|F| \geq 0$ . Cuando  $|F|(\Omega) < \infty$  diremos que  $F$  es una **medida de variación acotada**.

**Definición 3.** Sea  $\mathcal{F}$  un álgebra de subconjuntos de  $\Omega$ ,  $F : \mathcal{F} \rightarrow X$  una medida vectorial y  $\mu$  una medida a valores reales y no negativa (finita) definida en  $\mathcal{F}$ . Diremos que  $F$  es  $\mu$ -continua, lo cual denotamos  $F \ll \mu$ , si  $\lim_{\mu(E) \rightarrow 0} F(E) = 0$ .

## 2.2. Integración

A continuación daremos la definición y algunas propiedades de la integral de Bochner, también conocida como la “integral de Dunford y Schwartz” y la cual es una generalización directa de la clásica integral de Lebesgue.

**Definición 4.** Una función  $f : \Omega \rightarrow X$  se llama  $\mu$ -medible o **fuertemente medible** si existe una sucesión de funciones simples  $(f_n)_{n=1}^{\infty}$  tal que

$$\lim_{n \rightarrow \infty} \|f_n(w) - f(w)\| = 0$$

para  $\mu$ -casi toda  $w \in \Omega$ .

**Definición 5.** Sea  $f : \Omega \rightarrow X$  una función fuertemente medible. Diremos que es Bochner integrable si existe una sucesión  $\{\varphi_n\}_{n=1}^{\infty}$  de funciones  $X$ -escalonadas tal que la función real medible  $\|f - \varphi_n\|$  es Lebesgue integrable para cada  $n \in \mathbb{N}$  y

$$\lim_{n \rightarrow \infty} \int \|f - \varphi_n\| d\mu = 0.$$

En este caso, para cada  $E \in \Sigma$  la integral de Bochner de  $f$  sobre  $E$  se define por

$$\int_E f d\mu = \lim_{n \rightarrow \infty} \int_E \varphi_n d\mu, \quad (2)$$

donde el límite se toma en la topología de la norma en  $X$ .

Sea  $1 \leq p < \infty$ .  $L^p(\Omega, \Sigma, \mu, X)$  o simplemente  $L^p(\mu, X)$  denotará todas las clases de equivalencia de las funciones  $f : \Omega \rightarrow X$  que son  $\mu$ -Bochner integrables y tales que

$$\|f\|_p = \left[ \int_{\Omega} \|f\|_X^p d\mu \right]^{1/p} < \infty.$$

Igual que en el caso escalar, puede demostrarse que  $(L^p(\mu, X), \|\cdot\|_p)$  es un espacio de Banach.

$L^\infty(\Omega, \Sigma, \mu, X)$  o simplemente  $L^\infty(\mu, X)$  denotará todas las clases de equivalencia de funciones esencialmente acotadas  $\mu$ -Bochner integrables  $f : \Omega \rightarrow X$ . Para  $f \in L^\infty(\mu, X)$  denotamos por

$$\|f\|_\infty = \text{ess sup } \|f\|_X,$$

y también como en el caso escalar puede verse que  $(L^\infty(\mu, X), \|\cdot\|_\infty)$  es un espacio de Banach.

El símbolo  $L^p(\mu)$ ,  $1 \leq p \leq \infty$ , siempre denotará  $L^p(\mu, X)$  cuando  $X$  es el campo de escalares.

Uno de los aspectos más interesantes de la teoría de la integral de Bochner se centra en la siguiente pregunta: ¿Cuándo una medida vectorial surge como una integral de Bochner indefinida?

Sea  $(\Omega, \Sigma, \mu)$  un espacio de medida finita y  $F : \Sigma \rightarrow X$  una medida vectorial de la forma

$$F(E) = \int_E f d\mu$$

para alguna función Bochner integrable  $f$ .

Podemos ver que  $F$  es numerablemente aditiva,  $\mu$ -continua y de variación acotada. Recíprocamente, si  $F : \Sigma \rightarrow X$  es cualquier medida aditiva numerable,  $\mu$ -continua y de variación acotada con un rango finito dimensional entonces el clásico Teorema de Radon-Nikodým produce una función Bochner-integrable  $f$  tal que

$$F(E) = \int_E f d\mu.$$

Para la integral de Bochner general, esto ya no necesariamente se verifica.

**Ejemplo 1.** Una medida vectorial aditiva numerable con valores en  $c_0$ , de variación acotada que no tiene derivada de Radon-Nikodým.

Construcción.

Sea  $\mu$  la medida de Lebesgue en  $[0, 1]$ . Si  $E \subset [0, 1]$  es un conjunto medible, sean

$$\lambda_n(E) = \int_E \text{sen}(2^n \pi t) dt,$$

y  $F(E) = (\lambda_1(E), \lambda_2(E), \dots, \lambda_n(E), \dots)$ .

$F$  es una medida vectorial finitamente aditiva con valores en  $c_0$ . Se puede ver que  $F$  es  $\mu$ -continua, aditiva numerable y de variación acotada pero no tiene derivada de Radon Nikodým.

### 3. El teorema de Radon-Nikodým y operadores Riesz representables en $L^1(\mu)$

**Definición 6.** Diremos que un espacio de Banach  $X$  tiene la **propiedad de Radon-Nikodým con respecto a**  $(\Omega, \Sigma, \mu)$  si para cada medida vectorial numerablemente aditiva  $G : \Sigma \rightarrow X$ ,  $\mu$ -continua y de variación acotada, existe  $g \in L^1(\mu, X)$  tal que

$$G(E) = \int_E g d\mu \quad \text{para toda } E \in \Sigma.$$

**Definición 7.** Diremos que un espacio de Banach  $X$  tiene la **propiedad de Radon-Nikodým** si  $X$  tiene la propiedad de Radon-Nikodým con respecto a todo espacio de medida finita.

**Definición 8.** Diremos que un operador lineal acotado  $T : L^1(\mu) \rightarrow X$  es Riesz Representable (o simplemente, representable) si existe  $g \in L^\infty(\mu, X)$  tal que

$$Tf = \int_\Omega f g d\mu \quad \text{para toda } f \in L^1(\mu).$$

**Teorema 1.** Sea  $X$  un espacio de Banach y  $(\Omega, \Sigma, \mu)$  un espacio de medida finita. Entonces  $X$  tiene la propiedad de Radon-Nikodým con respecto a  $(\Omega, \Sigma, \mu)$  si y sólo si cada operador lineal continuo  $T : L^1(\mu) \rightarrow X$  es representable.

### 4. La propiedad de Radon-Nikodým y el teorema de Fatou

En esta parte demostraremos la equivalencia entre el teorema de Radon-Nikodým y el teorema de Fatou sobre la existencia de límites radiales en casi todo punto de funciones armónicas y acotadas definidas en el disco unitario y con valores en un espacio de Banach  $X$ .

**Definición 9.** Sea  $u : D \rightarrow X$ , donde  $D$  es el disco unitario y  $X$  es un espacio de Banach. Diremos que  $u$  es **armónica** si  $u \in C^2(D, X)$ , es decir,  $u$  es continua con segundas derivadas continuas y satisface la ecuación de Laplace  $\Delta u = 0$ .

**Definición 10.** Un espacio de Banach  $X$  se dice que tiene la propiedad de Fatou si para toda función armónica  $u$  con valores en  $X$  y acotada, es decir  $\sup_{z \in D} \|u(z)\|_X < \infty$ , existe  $f \in L^\infty([0, 1], X)$  tal que

$$\lim_{r \rightarrow 1^-} u(re^{2\pi it}) = f(t)$$

para casi todo  $t \in [0, 1]$ .

**Definición 11.** El núcleo de Poisson para el disco unitario  $D$  es la función

$$P_r(t) = \sum_{k=-\infty}^{\infty} r^{|k|} e^{ikt} \quad t \in [-\pi, \pi]. \quad (3)$$

Además, podemos observar que

$$P_r(t) = \sum_{k=-\infty}^{\infty} r^{|k|} e^{ikt} = \frac{1 - r^2}{1 - 2r \cos t + r^2} \quad t \in [-\pi, \pi].$$

**Lema 2.** Sea  $f \in L^1([0, 1], X)$ . Entonces la integral de Poisson de  $f$ , definida por

$$P(f)(re^{2\pi it}) = \int_0^1 f(s)P_r(t - s)ds$$

es una función armónica con valores en  $X$  que verifica

$$\lim_{r \rightarrow 1^-} P(f)(re^{it}) = f(t)$$

para casi todo  $t \in [0, 1]$ .

**Teorema 3.**  $X$  tiene la propiedad de Radon-Nikodým si y sólo si  $X$  tiene la propiedad de Fatou.

*Demostración.* Supongamos que  $X$  tiene la propiedad de Radon-Nikodým. Sea  $u : D \rightarrow X$  una función armónica y acotada.

Sean  $r_n = 1 - \frac{1}{2^n}$  y  $f_n(t) = u(r_n e^{2\pi it})$ .  $(f_n)_{n=1}^{\infty}$  es una sucesión de funciones continuas con valores en  $X$  tal que

$$\sup_{n \in \mathbb{N}} \|f_n\|_{\infty} \leq \sup_{z \in D} \|u(z)\| < \infty.$$

Como  $X \hookrightarrow X^{**}$  tenemos las inclusiones isométricas

$$C([0, 1], X) \hookrightarrow \mathcal{L}(L^1([0, 1]), X) \hookrightarrow \mathcal{L}(L^1([0, 1]), X^{**}) \quad (4)$$

dadas por  $f \mapsto T_f$  tal que

$$T_f(\phi) = \int_0^1 \phi(t)f(t)dm(t)$$

y  $T \mapsto jT$  donde  $j : X \rightarrow X^{**}$  es la inclusión canónica.

Dado que

$$(L^1([0, 1]) \widehat{\otimes}_{\pi} X^*)^* = (L^1([0, 1], X^*))^* = \mathcal{L}(L^1([0, 1]), X^{**}),$$

y puesto que la sucesión  $(f_n)_{n=1}^{\infty}$  es una sucesión acotada en el espacio  $L^1([0, 1], X^*)^*$ , podemos aplicar el Teorema de Banach-Alaoglu y encontrar  $T \in \mathcal{L}(L^1([0, 1]), X^{**})$  y una subsucesión  $(n_k)_{k=1}^{\infty}$  de números naturales tales que  $(f_{n_k})_{k=1}^{\infty}$  converge a  $T$  en la topología débil-\* de  $L^1([0, 1], X^*)^*$ .

Se tiene que

$$T_{f_{n_k}}(\phi) \longrightarrow T(\phi) \quad \text{para toda } \phi \in L^1([0, 1]), \quad (5)$$

y que  $T(\phi) \in X$  y  $P(T) = u$ .

Definamos ahora la siguiente medida vectorial

$$\mu(E) = T(\chi_E) = \lim_{k \rightarrow \infty} \int_E u(r_{n_k} e^{2\pi i t}) dt.$$

Podemos ver que  $\mu$  es numerablemente aditiva,  $\mu \ll m$  y de variación acotada.

Puesto que  $X$  tiene la propiedad de Radon-Nikodým, existe  $f \in L^1([0, 1], X)$  tal que

$$\mu(E) = \int_E f d\mu \quad \text{para todo boreliano } E.$$

Además  $f \in L^\infty([0, 1], X)$  y puesto que

$$P(f)(re^{2\pi i t}) = u(re^{2\pi i t}),$$

se sigue del Lema 2 la existencia de

$$\lim_{r \rightarrow 1} P(f)(re^{2\pi i t}) = f(t)$$

para casi toda  $t \in [0, 1]$ . Por tanto,  $X$  tiene la propiedad de Fatou.

Recíprocamente, supongamos que  $X$  tiene la propiedad de Fatou y sea  $\mu$  una medida vectorial con valores en  $X$  tal que  $\mu$  es numerablemente aditiva,  $\mu \ll m$  y  $|\mu|([0, 1]) < \infty$ . Estudiemos primero el siguiente caso para  $\mu$  y  $m$ :

Existe  $c > 0$  tal que  $|\mu|(E) \leq cm(E)$  para todo boreliano  $E$ .

Definamos el operador  $T_\mu : L^1([0, 1]) \rightarrow X$  del modo siguiente: si

$$\varphi = \sum_{k=1}^m \alpha_k \chi_{E_k},$$

con  $\alpha_k$  escalar,  $E_k$  boreliano para  $1 \leq k \leq m$ ,  $E_k \cap E_l = \emptyset$  si  $k \neq l$ , entonces

$$T_\mu(\varphi) = \sum_{k=1}^m \alpha_k \mu(E_k).$$

Podemos extender continuamente  $T_\mu$  a todo  $L^1([0, 1])$ , y esta extensión posee la misma norma.

Ahora definamos la integral de Poisson del operador  $T_\mu$ , esto es,  $u = P(T_\mu)$ . Tenemos que  $u$  es una función armónica con valores en  $X$  y además

$$\|u(re^{2\pi i t})\|_X \leq C, \quad \text{para toda } z \in D.$$

Como  $X$  tiene la propiedad de Fatou existe  $f \in L^\infty([0, 1], X)$  tal que

$$\lim_{r \rightarrow 1} u(re^{2\pi i t}) = f(t) \quad \text{para casi todo } t \in [0, 1].$$

Podemos ver que

$$\mu(E) = T_\mu(\chi_E) = \int_E f d\mu.$$

Para resolver el caso general, dada  $\mu \ll m$ , basta considerar la función maximal

$$\mu^*(\theta) = \left\{ \frac{|\mu|(I)}{m(I)} : \theta \in I, I \text{ intervalo} \right\}.$$

(el supremo se toma sobre todos los intervalos  $I$  que contienen a  $\theta$ ).

□

Veamos ahora un ejemplo de un espacio que no tiene la Propiedad de Fatou.

**Ejemplo 2.** Una función vectorial armónica y acotada que no tiene límites radiales ctp.

Construcción:

Sea  $X = c_0$  y  $T : L^1[0, 1] \rightarrow c_0$  tal que

$$T(f) = \left( \int_{[0,1]} f(t) \text{sen}(2^n \pi t) dt \right)_{n=0}^{\infty}$$

Consideramos la integral de Poisson del operador  $T$ , esto es,  $\varphi : D \rightarrow c_0$  tal que

$$\varphi(re^{2\pi it}) = P(T)(re^{2\pi it}) = T(P_r(t - \cdot))$$

Sabemos que  $\varphi$  es una función armónica y acotada.

Mostraremos que existe un subconjunto Lebesgue medible  $E \subset [0, 1]$  con  $m(E) > 0$  tal que no existe  $\lim_{r \rightarrow 1} \varphi(re^{2\pi it})$  en  $(c_0, \|\cdot\|_{\infty})$  para toda  $t \in E$ .

Observemos que si  $t \in [0, 1]$  y  $0 < r < 1$

$$\varphi(re^{2\pi it}) = \left( \int_0^1 P_r(t - \theta) \text{sen}(2^n \pi \theta) d\theta \right)_{n=0}^{\infty}$$

y se sabe que para cada  $n \in \mathbb{N}$

$$\lim_{r \rightarrow 1} \int_0^1 P_r(t - \theta) \text{sen}(2^n \pi \theta) d\theta = \text{sen}(2^n \pi t) \text{ para casi toda } t \in [0, 1].$$

Fijemos uno de tales  $t$ 's. Si existiera  $f(t) \in c_0$ , digamos  $f(t) = (f_n(t))_{n=1}^{\infty}$  tal que

$$\lim_{r \rightarrow 1} \varphi(re^{2\pi it}) = f(t) \text{ en } \|\cdot\|_{\infty},$$

entonces debido a que convergencia en  $\|\cdot\|_{\infty}$  implica convergencia coordenada a coordenada tendríamos que

$$f_n(t) = \text{sen}(2^n \pi t) \text{ para toda } n \in \mathbb{N},$$

pero la sucesión  $(\text{sen}(2^n \pi t))_{n=1}^{\infty} \notin c_0$  a menos que  $t = k/2^m$ , para algún  $k \in \mathbb{Z}$  y algún  $m \in \mathbb{N}$ , y puesto que el conjunto  $\{t = k/2^m : k \in \mathbb{Z}, m \in \mathbb{N}\}$  es numerable y por tanto, tiene medida de Lebesgue cero, entonces hay un subconjunto  $E$  de  $[0, 1]$  Lebesgue medible con medida de Lebesgue positiva tal que para toda  $t \in E$   $(f_n(t))_{n=1}^{\infty} \notin c_0$ .

Por tanto  $\varphi : D \rightarrow c_0$  no tiene límites radiales ctp., es decir,  $c_0$  no tiene la propiedad de Fatou. En realidad, lo que ocurre es que el espacio  $c_0$  no tiene la Propiedad de Radon-Nikodým, la cual es equivalente a la propiedad de Fatou.

## 5. Referencias

- [1] R. G. Bartle, *The Elements of Integration and Lebesgue Measure*, John Wiley, 1995.
- [2] O. Blasco, *Radon-Nikodým versus Fatou*, Aportaciones Matemáticas, Serie Comunicaciones 20 (1997), 1-5.
- [3] O. Blasco, *Boundary Values of Functions in Vector-Valued Hardy Spaces and Geometry on Banach Spaces*, J. Funct. Anal. 78 (1988), 346-364.
- [4] J. Diestel, J.J. Uhl, *Vector Measures*, Mathematical Surveys 15, AMS, 1977.
- [5] J. García-Cuerva, J.L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North-Holland, 1985.
- [6] W. Hensgen, *Some remarks on boundary values of vector-valued harmonic and analytic functions*, Arch. Math. 57 (1991), 88-96.
- [7] E.M. Stein, *Harmonic Analysis. Real-variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, 1993.
- [8] F. Trèves, *Topological Vector Spaces, Distributions and Kernels*, Academic Press, 1967.

# Estimación de un parámetro de interés en presencia de parámetros de estorbo

Gudelia Figueroa Preciado  
Departamento de Matemáticas  
Universidad de Sonora  
gfiguero@gauss.mat.uson.mx

## Resumen

*Para estudiar la mayoría de los problemas reales generalmente se requieren modelos multi-paramétricos y la función de verosimilitud es una herramienta que nos permite obtener información sobre los parámetros de interés de un modelo. Esta función puede ser compleja cuando se trabaja con varios parámetros, y una alternativa es describir un parámetro a la vez, considerando los restantes como parámetros de estorbo. Para efectuar este tipo de análisis se pueden utilizar diferentes métodos, entre los cuales se examinan: la verosimilitud perfil, la verosimilitud condicional y la verosimilitud integrada. Se describen las características de cada una de ellas y se desarrolla un ejemplo utilizando estos tres métodos.*

## 1. Introducción

El problema de especificar un modelo estadístico para observaciones recolectadas en una muestra puede llegar a ser una tarea difícil, ya que el modelo seleccionado puede tener varios parámetros o características de interés. En ciertos casos podemos conocer algunos de ellos y enfocarnos solamente en estimar el parámetro que nos interesa, pero en otros debemos encontrar la manera de estimar el parámetro de interés en presencia de los otros parámetros que denominaremos de estorbo. Pawitan [3] afirma que los parámetros de estorbo aparecen en escena como una consecuencia natural de nuestro esfuerzo al usar mejores modelos y de mayor dimensión. Este problema central en inferencia estadística ha sido abordado de diversas maneras, pero solo consideraremos tres de estos métodos: la verosimilitud perfil, la verosimilitud condicional y la verosimilitud integrada. Antes de describirlos debemos definir algunos de los conceptos y notación a utilizar.

Primeramente consideremos uno de los conceptos más importantes en estadística que es el de verosimilitud, el cual nos permite inferir sobre la población de la cual se selecciona la muestra. Fisher [1] introdujo el concepto de verosimilitud para cualquier valor particular de  $\theta$ , como proporcional a la probabilidad de observar la muestra. Así, si  $y$  es una variable aleatoria discreta cuya función de probabilidad es  $f(y; \theta) = P(y; \theta)$ , la función de verosimilitud de  $\theta$  la denotamos por  $L(\theta; y)$  y se calcula como:

$$L(\theta; y_o) = C(y_o)f(y_o; \theta) \propto P(y = y_o; \theta)$$

donde, como señala Sprott [4],  $y_o$  denota  $y$  observada y  $C(y_o)$  es una función de  $y_o$  arbitraria, positiva y acotada, que no depende de  $\theta$ .

Solo es razonable interpretar el cociente de verosimilitudes, por ello para comparar  $\theta_2$  y  $\theta_1$  se calcula  $L(\theta_2; y)/L(\theta_1; y)$ ; si este cociente es igual a  $k$  decimos que  $\theta_2$  es  $k$  veces más plausible que  $\theta_1$ .

Es conveniente normalizar la función de verosimilitud para que su máximo sea la unidad y obtener con ello una representación única. Para esto dividimos la función por su máximo, resultando lo que se conoce como función de verosimilitud relativa que denotamos por  $R(\theta; y)$  y calculamos como:

$$R(\theta; y) = \frac{L(\theta; y)}{\sup_{\theta} L(\theta; y)} = \frac{L(\theta; y)}{L(\hat{\theta}; y)}$$

El máximo de la función de verosimilitud es de gran utilidad, pero debemos evitar la tendencia común de enfocarnos solamente en éste y desechar la función en sí.

A continuación se presentan tres métodos de verosimilitud que nos permiten trabajar modelos que involucran parámetros de estorbo.

## 2. Verosimilitud perfil

En la verosimilitud perfil, para eliminar un parámetro de estorbo se le sustituye por su estimador máximo verosímil en cada valor fijo del parámetro de interés. Así podemos escribir la verosimilitud perfil como:

$$L_{max}(\delta; y) \propto f[y; \delta, \hat{\xi}(\delta)]$$

donde  $\hat{\xi}(\delta)$  es el estimador de máxima verosimilitud restringido de  $\xi$  para un valor específico de  $\delta$ . Al estandarizar con respecto al máximo sobre  $\delta$  obtenemos la verosimilitud relativa perfil:

$$R_{max}(\delta; y) = f[y; \delta, \hat{\xi}(\delta)] / f(y; \hat{\delta}, \hat{\xi})$$

Cuando la dimensión de  $\xi$  es grande, puede no ser muy conveniente utilizar la verosimilitud relativa perfil ya que el método supone que para cada  $\delta$  fijo,  $\xi$  es conocido e igual a su estimador máximo verosímil.

## 3. Verosimilitud condicional

La verosimilitud condicional es otro método que nos permite estimar parámetros de interés en presencia de parámetros de estorbo. En este método se analiza cuidadosamente la situación y se trata de identificar una distribución condicional que contenga toda la información concerniente al parámetro de interés. La función de verosimilitud derivada de esta distribución se le llama función de verosimilitud condicional.

Si tenemos un vector de parámetros  $\theta = (\delta, \xi)$  con la siguiente estructura en su función de verosimilitud:

$$\begin{aligned} L(\delta, \xi; y) &\propto f(y; \delta, \xi) = f(t; \delta, \xi) f(y; \delta/t) \\ &\propto L_{res}(\delta, \xi; t) L_c(\delta; y) \end{aligned}$$

La cantidad  $L_c(\delta; y)$  es la distribución condicional de  $\delta$  pues está basada en la distribución condicional de la muestra  $y$  dado  $t$ . El estadístico  $t$  es minimal suficiente para  $\xi$  para cualquier valor específico de  $\delta$ , ya que el segundo factor no depende de  $\xi$ , Spratt [4].

Si bien este método es muy utilizado, en algunas ocasiones el trabajar con la verosimilitud condicional puede llegar a ser más complicado que utilizar la verosimilitud original.

## 4. Verosimilitud integrada

En este método, para eliminar el parámetro de estorbo se utiliza un proceso de integración. Para ello se requiere conocer una densidad inicial o a priori para el parámetro de estorbo, que a su vez dependa del parámetro de interés. Así, multiplicando la verosimilitud por esta densidad e integrando con respecto al parámetro de estorbo, se elimina este último.

Si tenemos el vector de parámetros  $\theta = (\lambda, \delta)$  donde  $\lambda$  es un parámetro de estorbo para el cual conocemos la distribución a priori  $\Pi(\lambda; \delta)$ , entonces la función de verosimilitud integrada de  $\delta$  es:

$$L_I(\delta; y) \propto \int_{\Lambda} L(\delta, \lambda; y) \Pi(\lambda; \delta) d\lambda$$

La verosimilitud relativa se calcula como se ha expuesto anteriormente. La dificultad con este método de eliminación de parámetros es que depende de un conocimiento previo del cual no siempre se dispone.

## 5. Una aplicación

A continuación se mostrará, por medio de un ejemplo, el uso de estos tres métodos de eliminación de parámetros. Los datos se toman de Fisher [2] y muestran el número de garrapatas encontradas en cada una de 60 ovejas.

No. de garrapatas	0	1	2	3	4	5	6	7	8	9	10
No. de ovejas ( $f_i$ )	7	9	8	13	8	5	4	3	0	1	2

Considerando a  $X$  como la variable aleatoria que cuenta el número de garrapatas por oveja, Fisher [2] supone para  $X$  una distribución binomial negativa, esto es

$$P[X = x, \theta, p] = \binom{\theta + x - 1}{x} p^\theta (1 - p)^x$$

En este ejemplo  $\theta$  es el parámetro de interés y  $p$  será considerado como parámetro de estorbo. Enseguida se muestra el análisis de estos datos utilizando los métodos antes expuestos.

- Verosimilitud perfil.

Calculamos primero la función de verosimilitud para este ejemplo, que está dada por:

$$\begin{aligned} L(\theta; p) &\propto \prod_{x=0}^{10} P[X = x]^{f_x} = \prod_{x=0}^{10} \left[ \binom{\theta + x - 1}{x} p^\theta (1 - p)^x \right]^{f_x} \\ &= p^{\theta \sum_{x=0}^{10} f_x} (1 - p)^{\sum_{x=0}^{10} x f_x} \prod_{x=0}^{10} \left[ \binom{\theta + x - 1}{x} \right]^{f_x} \end{aligned}$$

Tomando logaritmos y derivando con respecto a  $p$  llegamos a que el estimador  $\hat{p}(\theta)$  es:

$$\hat{p}(\theta) = \frac{N\theta}{N\theta + \sum x f_x}$$

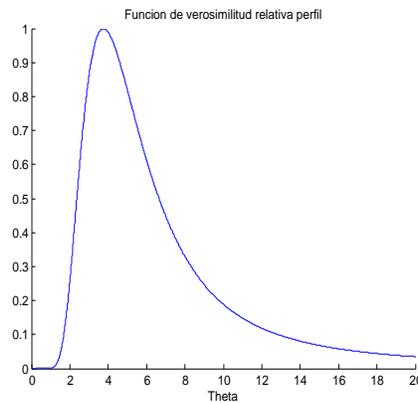
Podemos ahora calcular la verosimilitud relativa de  $\theta$  como:

$$R(\theta) = \frac{L_p(\theta)}{\max_{\theta \in \Theta} L_p(\theta)}$$

donde

$$L_p(\theta) \propto L[\theta, p = \hat{p}(\theta)]$$

A continuación se presenta la gráfica obtenida en Matlab al programar esta verosimilitud relativa perfil



**Figura 1.**

- Verosimilitud condicional.

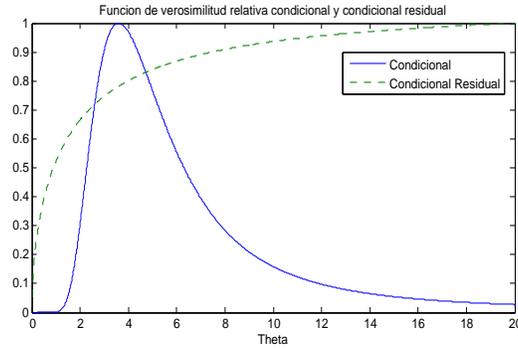
Para el ejemplo citado la verosimilitud condicional se calcula como:

$$\begin{aligned} L_c(\theta) &\propto P[\vec{X}|T = t] = P[X_1 = x_1, \dots, X_n = x_n, T = t] \\ &= \frac{p^{N\theta}(1-p)^t \prod_{i=1}^n \binom{\theta+x_i-1}{x_i}}{p^{N\theta}(1-p)^t \binom{N\theta+t-1}{t}} \\ &= \frac{\prod_{i=1}^n \binom{\theta+x_i-1}{x_i}}{\binom{N\theta+t-1}{t}} \end{aligned}$$

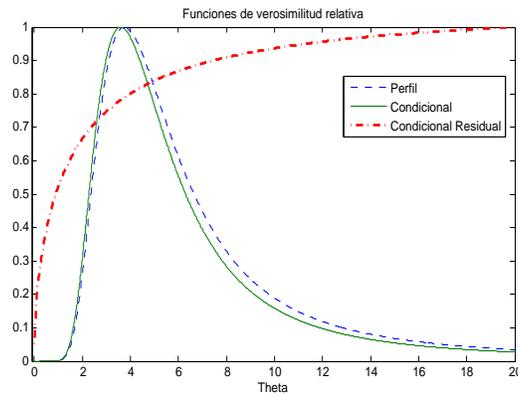
de donde obtenemos la condicional residual

$$P[T = t; \theta, p] = \binom{N\theta + t - 1}{t} p^{N\theta} (1-p)^t$$

En la Figura 2. podemos observar las verosimilitudes condicional y condicional residual. Esta última es plana, por lo cual se deduce que no posee información acerca de  $\theta$ . La Figura 3. hace una comparación de las verosimilitudes perfil y condicional.



**Figura 2.**



**Figura 3.**

■ Verosimilitud integrada.

Teníamos que la función de verosimilitud para el caso de una binomial negativa es:

$$L(\theta, p) = \prod_{x=0}^n \left[ \binom{\theta + x - 1}{x} p^\theta (1 - p)^x \right]^{f_x}$$

donde  $\theta$  es el parámetro de interés y  $p$  es el parámetro de estorbo.

La verosimilitud integrada la calculamos como:

$$L_I(\theta) \propto \int L(\theta, p, x) \Pi(p) dp$$

Suponiendo para  $p$  una distribución Beta, esto es:

$$\Pi(p) = \frac{(1 - p)^{\beta-1} p^{\alpha-1}}{B(\alpha, \beta)}$$

Entonces la verosimilitud integrada es:

$$L_I(\theta) \propto \prod_{x=0}^{10} \left[ \binom{\theta + x - 1}{x} \right]^{f_x} B(\theta \sum f_x + \alpha, \sum x f_x + \beta)$$

Y nuevamente la verosimilitud relativa integrada calculada como:

$$R_I(\theta) = \frac{L_I(\theta)}{\max_{\theta \in \Theta} L_I(\theta)}$$

se programó en Matlab, para diferentes parámetros de la distribución Beta que se utilizó como distribución a priori. En las siguientes gráficas se muestran las verosimilitudes integradas obtenidas.

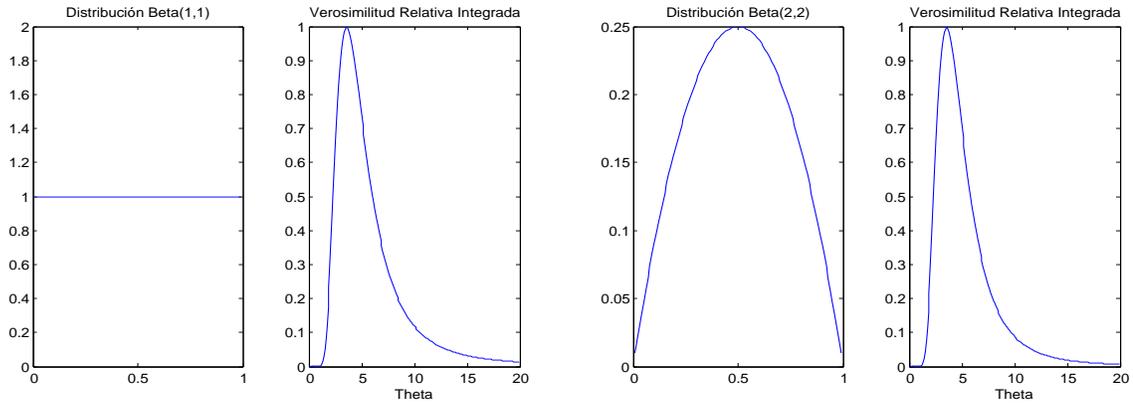


Figura 4.

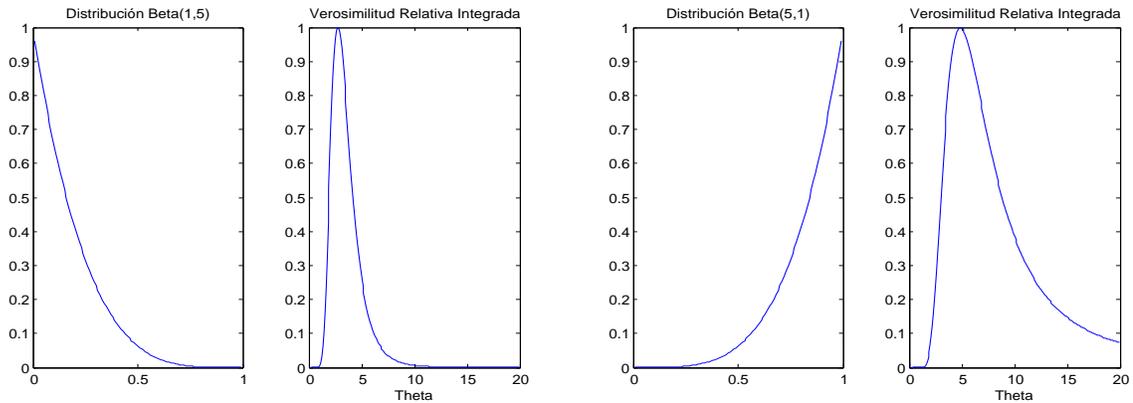


Figura 5.

Se puede observar que para utilizar este método es necesario un conocimiento previo acerca del problema en estudio, pues solamente así se puede seleccionar el valor adecuado para los parámetros de la distribución a priori utilizada.

## 6. Conclusiones

De lo anteriormente expuesto se deduce que para el análisis de modelos multiparamétricos, problema central en inferencia estadística, no existen normas establecidas sobre el método más conveniente a utilizar. Sobre los métodos expuestos se han vertido ya algunos comentarios, y se puede agregar que la verosimilitud perfil está considerada como el método más sencillo a utilizar

y está completamente determinado por la probabilidad de los datos observados. La verosimilitud condicional puede ser afectada por el esquema muestral o arreglo ideado para el espacio muestral, y en el caso de la verosimilitud integrada puede ser muy difícil construir distribuciones a priori cuando se tienen varios parámetros de estorbo, aunque algunos consideren más seguro efectuar un proceso de integración que uno de maximización.

## 7. Referencias

- [1] Fisher, R.A. (1921). On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron*, 1, 3 – 32.
- [2] Fisher, R.A. (1941). The Negative Binomial Distribution. *Annals of Eugenics*, 11, 182 – 187.
- [3] Pawitan, Y. (2001). *In All Likelihood*. New York. Oxford.
- [4] Sprott, D.A. (2000) *Statistical Inference in Science* New York. Springer Verlag.



# Verosimilitud para inferencia estadística sobre los parámetros de un proceso de Poisson

María Elena Conde Hernández  
Departamento de Matemáticas  
Universidad de Sonora  
male\_conde@hotmail.com

## Resumen

*El objetivo principal de este artículo, es el de realizar inferencia estadística sobre los parámetros involucrados en la función de intensidad de un proceso de Poisson no homogéneo, mediante el enfoque de verosimilitud; ésta última se basa solamente en los datos observados y el conjunto de medidas de probabilidad para el sistema que es objeto de estudio. Mediante un ejemplo con datos reales en meteorología y apoyados de la herramienta computacional, mostraremos el vínculo establecido entre las disciplinas de probabilidad y estadística.*

## 1. Introducción

El proceso de Poisson es un modelo fundamentado en teoría de probabilidad y utilizado en situaciones de conteo, donde se deducen propiedades con base en axiomas y suposiciones técnicas, que consideraremos para la simulación del mismo. En nuestro caso a tratar consideraremos a través de la variable de estudio *rachas de calor*, datos reales de un proceso de Poisson no homogéneo en un intervalo de tiempo fijo y con cierta función de intensidad gama, que incluye parámetros  $\alpha$  y  $\beta$ , es decir, observaremos la realización aleatoria de los tiempos de ocurrencia de una muestra de una población que cumple con ciertas características; la incertidumbre estocástica se hace presente en el momento de desconocer qué tiempos y cuántos eventos resultarán en la muestra. Una vez que han sido registradas las observaciones, éstas son fijas, esto es, ya no existe más la incertidumbre estocástica; estamos ahora interesados en inferir sobre los parámetros del modelo generado por esos datos observados, es decir, se hace presente la incertidumbre estadística.

Formalmente, el objetivo es hacer inferencia sobre los números  $\alpha$  y  $\beta$  a través de intervalos de verosimilitud y una aproximación de éstos con los intervalos de confianza, basándonos en una muestra aleatoria  $\underline{S} = (s_1, s_2, \dots, s_n)$  del espacio de estados  $E = \{1, 2, \dots\}$ , donde  $\{S_n\}_{n \geq 1}$  es un proceso puntual en orden creciente que representa los tiempos de ocurrencia de las rachas de calor; la sucesión  $\{T_n = S_n - S_{n-1}\}_{n \geq 1}$  son los tiempos entre rachas consecutivas, los cuales se distribuyen de manera exponencial, cumpliendo con las condiciones para que el proceso sea de Poisson con cierta función de intensidad  $\lambda(t)$ , o tasa de ocurrencia, y parámetros  $\alpha, \beta$  desconocidos. Ahora bien  $\alpha$  y  $\beta$  son valores que satisfacen la condición de que la densidad de probabilidad  $f_n(\underline{S}/\alpha, \beta)$  es grande, esto es, para cualquier vector  $\underline{S}$  este razonamiento nos conduce a considerar valores de  $\alpha$  y  $\beta$  cuya función de verosimilitud  $f_n(\underline{S}/\alpha, \beta)$  es un máximo y utilizar estos valores como una estimación de  $\alpha$  y  $\beta$ .

Por un lado tenemos, un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$  donde  $P$  es conocida; el objetivo de la probabilidad es deducir propiedades sobre  $P$  en base a axiomas que desarrollaremos en la Sección

1 tomados de Billingsley[2] principalmente. Por otro lado, en estadística donde  $P$  es desconocida, el objetivo es inferir propiedades de  $P$  con base en observaciones realizadas del fenómeno aleatorio. Para tal fin, es preciso desarrollar la teoría estadística necesaria que permita realizar inferencia acerca de los parámetros considerados en el modelo. Para ello se retoma la teoría de verosimilitud de Kalbfleisch[3] y Sprott[9]. Esta teoría, mediante la función e intervalos de verosimilitud, proporciona una estimación razonable de los parámetros, a través del estimador máximo verosímil.

## 2. Procesos de Poisson

Estudiaremos, primeramente, los requisitos que un proceso debe cumplir para que sea de Poisson y éstos se retoman de Billingsley[2] principalmente. De manera intuitiva, un proceso de Poisson con cierta tasa de ocurrencia  $\lambda$  en la recta real, es un proceso que contabiliza el número de eventos ocurridos hasta un tiempo  $t$ , el proceso es *homogéneo* si el número de eventos que ocurren dependen solamente de la longitud del intervalo de tiempo.

Consideremos  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad,  $E = \{1, 2, \dots\}$  el espacio de estados. Sea  $\{S_n\}_{n \geq 1}$ ,  $S_n : \Omega \rightarrow E$  un proceso puntual en orden creciente; en nuestro caso  $S_n$  representa el *tiempo en el que ocurre el enésimo evento*. Dado el proceso puntual  $\{S_n\}_{n \geq 1}$ , definimos el proceso  $\{T_n\}_{n \geq 1}$  que llamaremos *tiempo entre eventos consecutivos*, de la siguiente forma:  $T_n := S_n - S_{n-1}$ ;  $n \geq 1$ ; las variables aleatorias  $\{T_n\}_{n \geq 1}$  son independientes e idénticamente distribuidas exponencial con parámetro  $\lambda$ , característica fundamental de un proceso de Poisson de este tipo. Una generalización del proceso de Poisson es permitir que el parámetro  $\lambda$  sea una función de  $t$ . Esto es, en muchas aplicaciones se consideran parámetros que varían con el tiempo en forma no homogénea, tales procesos son denominados procesos de Poisson no homogéneos o no estacionarios.

Denotemos por  $\xi$  la  $\sigma$ -álgebra generada por los intervalos de la forma  $(a, b]$  de  $E$ ,  $\mu$  una medida en  $(E, \xi)$ , tal que es finita para conjuntos en  $\xi$ ,  $\mu(\cdot) : \xi \rightarrow \mathbb{R}$ , y absolutamente continua con densidad  $\lambda(t)$ ; esto es  $\mu(t) = \int_0^t \lambda(t)dt$ , donde  $\lambda(t)$  es la función de intensidad del proceso.

**Definición 2.1.** *Un proceso de conteo  $\{N(t)\}_{t \geq 0}$  es un proceso de Poisson no homogéneo con función de intensidad  $\lambda(t)$ , si*

(i)  $\Pr[N(0) = 0] = 1.$

(ii) Para  $0 \leq s_0 < s_1 < s_2 < \dots < s_k < \dots < s_n$  tenemos que  $N((s_i, s_{i+1} ])$ ,  $i = 1, 2, \dots, n$  son variables aleatorias independientes.

(iii) Para cada  $t \geq 0$

$$\Pr[N((0, t]) = n] = e^{-\{\mu(t)\}} \frac{(\mu(t))^n}{n!}; \quad \text{para } n \geq 0;$$

donde  $\mu(t) = \int_0^t \lambda(u)du.$

### 3. Verosimilitud para inferencia estadística

Un concepto básico en la teoría de verosimilitud es precisamente la función de verosimilitud. Es en esta donde se basa principalmente el desarrollo de esta sección, que se apoya en Kalbfleisch[3] y Sprott[9] básicamente. Consideremos que  $\chi$  es una muestra de variables aleatorias discretas independientes idénticamente distribuidas con función de probabilidad  $P(x; \theta)$ , que depende de un número finito de parámetros reales desconocidos  $\theta = (\theta_1, \theta_2, \dots, \theta_n) \in \Theta \subseteq \mathbb{R}^k$ .

#### 3.1. Función de verosimilitud

La estructura inferencial de un modelo es completamente especificado por los siguientes tres elementos: el espacio muestral, conjunto de todas las muestras aleatorias posibles  $\chi = \{X_1, X_2, \dots, X_n\}$ ; el espacio paramétrico  $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$  y la densidad de probabilidad  $f(x; \theta)$ . Este es el modelo estadístico usual de un experimento; el conjunto de funciones de densidades con parámetros definidos en un espacio paramétrico,  $\mathcal{F} = \{f(\cdot; \theta) : \theta \in \Theta \subseteq \mathbb{R}^k\}$ .

**Definición 3.1.** *La función de verosimilitud de una muestra o conjunto de variables aleatorias  $\underline{x} = (x_1, x_2, \dots, x_n)$ , se define como proporcional a la probabilidad de observar  $X = \underline{x}$ . La función  $f : \Theta \rightarrow \mathbb{R}^+ \cup \{0\}$  escrita como  $L(\theta; \underline{x})$  está dada por:*

$$L(\theta; \underline{x}) \propto P(x_1, x_2, \dots, x_n; \theta);$$

donde  $\theta$  puede ser un vector de parámetros.

**Observación 3.1.** *La función de verosimilitud tiene toda la información del parámetro contenida en la muestra y aunque el valor de  $L(\theta; \underline{x})$  es determinado por una distribución de probabilidad, la función de verosimilitud no es una función de probabilidad.*

Un *estimador de máxima verosimilitud (EMV)* es un elemento  $\hat{\theta} \in \Theta$  en el cual  $L(\theta; \underline{x})$  alcanza el valor máximo en  $\Theta$ , esto es,

$$L(\hat{\theta}; \underline{x}) = \max_{\theta \in \Theta} L(\theta; \underline{x}).$$

El *EMV*  $\hat{\theta}$ , es el valor más plausible de  $\theta$  en el sentido de proporcionar a la muestra observada la más alta probabilidad de ocurrencia. Como generalmente la función de verosimilitud es compleja y difícil de evaluar, y dado que existe una relación biunívoca entre una función y su logaritmo, entonces se prefiere trabajar con el logaritmo de la función de verosimilitud,

$$\ell(\theta; \underline{x}) = \log L(\theta; \underline{x});$$

y a esta función se le conoce como *función log-verosimilitud*. Siempre podemos trabajar con el logaritmo de  $L(\theta; \underline{x})$  ya que la función de verosimilitud es no negativa, con la convención de que  $\ell(\theta; \underline{x}) = -\infty$ , si  $L(\theta; \underline{x}) = 0$ . Dada una muestra aleatoria con resultados  $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ , la función de verosimilitud es proporcional a la probabilidad de observar  $X = \underline{x}$ , entonces, el objetivo que se pretende con este método de estimación es encontrar aquellos valores de los parámetros que maximicen la probabilidad de obtener los valores que se dieron en la muestra. Por lo tanto, para encontrar estos estimadores se debe derivar la función de verosimilitud con respecto a cada uno de los parámetros a estimar, igualar a cero y despejar el respectivo valor.

La función de verosimilitud por si sola no nos dice nada, lo interesante es trabajar con los cocientes de las funciones de verosimilitud, esto tiene sentido y se interpreta como una medida de plausibilidad entre dos valores del parámetro basada en la muestra observada  $\underline{x}$ . Para ello es preciso introducir la siguiente definición.

*Definición 3.2 La función de verosimilitud relativa  $R(\theta; \underline{x})$  de  $\theta$  está definida como la razón de la función de verosimilitud  $L(\theta; \underline{x})$  y su máximo  $L(\hat{\theta}; \underline{x})$  :*

$$R(\theta; \underline{x}) = \frac{L(\theta; \underline{x})}{\sup_{\theta} L(\theta; \underline{x})} = \frac{L(\theta; \underline{x})}{L(\hat{\theta}; \underline{x})}.$$

*Observación 3.2. Note que  $L(\theta; \underline{x}) \leq L(\hat{\theta}; \underline{x})$  para todos los valores posibles de  $\theta$ , se sigue por tanto que  $0 \leq R(\theta; \underline{x}) \leq 1$ .*

Para el caso de un solo parámetro, el gráfico de la función de verosimilitud relativa nos muestra la precisión o plausibilidad de un valor específico de  $\theta$  relativo al *EMV*  $\hat{\theta}$ ; ya que el valor de  $\hat{\theta}$  es una medida de la posición de la función de verosimilitud con respecto al eje  $\theta$ . Nos indica también los correspondientes valores de  $\theta$  que llegan a ser implausibles. En pocas palabras  $R(\theta; \underline{x})$  es un resumen de toda la información sobre  $\theta$  contenida en la muestra.

Supongamos que el modelo de probabilidad para un experimento involucra dos parámetros desconocidos,  $\alpha$  y  $\beta$ . El propósito no se desvía de la intención principal de este trabajo, producir inferencia estadística sobre los parámetros desconocidos del modelo. Supongamos que sólo nos interesa hacer inferencia sobre  $\beta$ , con  $\alpha$  considerado como un parámetro de estorbo; la verosimilitud perfil es una alternativa que se ha desarrollado para eliminar tal parámetro de estorbo, de tal forma que la verosimilitud se puede escribir solamente como una función del parámetro de interés.

*Definición 3.3. Definimos la función de verosimilitud perfil de  $\beta$ ,  $L_{\text{máx}}(\beta; \underline{x})$ , como el máximo de  $L(\alpha, \beta; \underline{x})$  sobre  $\alpha$  con  $\beta$  fijo, en otras palabras*

$$L_{\text{máx}}(\beta; \underline{x}) = \max_{\alpha} L(\alpha, \beta; \underline{x}) = L(\hat{\alpha}(\beta), \beta; \underline{x}),$$

donde  $\hat{\alpha}(\beta)$  es el *EMV* restringido de  $\alpha$  por un valor específico de  $\beta$ .

### 3.2. Intervalos de verosimilitud

Hasta ahora se ha visto cómo determinar el estimador de máxima verosimilitud. Sin embargo, una estimación puntual no proporciona ninguna información sobre la precisión involucrada en la estimación. Por ello es necesario determinar un intervalo de valores plausibles del parámetro de la función de verosimilitud. A ese intervalo se le conoce como intervalo de verosimilitud.

*Definición 3.4. El conjunto de valores de  $\theta$  para los cuales  $R(\theta; \underline{x}) \geq c$ , donde  $0 \leq c \leq 1$  es conocido como un  $100c\%$  región de verosimilitud para  $\theta$ .*

Si  $\theta$  es un valor real, la  $100 c \%$  región de verosimilitud consiste de un intervalo que llamaremos *100 c % intervalo de verosimilitud* y denotaremos  $IV(c)$  para  $\theta$ . Cada valor específico de  $\theta$  dentro de la región tiene una verosimilitud relativa  $R(\theta; \underline{x}) \geq c$ , y cada valor específico de  $\theta$  fuera de la región tiene una verosimilitud relativa  $R(\theta; \underline{x}) < c$ .

Los intervalos o regiones de verosimilitud son de gran utilidad para comparar la plausibilidad de los valores del parámetro  $\theta$  con respecto a  $\hat{\theta}$  especificando rangos de los valores más plausibles del parámetro. Gráficamente los intervalos de verosimilitud son resultado de trazar una línea horizontal a través de la gráfica de  $R(\theta; \underline{x})$  a una distancia  $c$  del eje de  $\theta$ . Variando  $c$  de 0 a 1 se produce un conjunto completo de intervalos de verosimilitud anidados que converge al estimador de máxima verosimilitud  $\hat{\theta}$  cuando  $c \rightarrow 1$ . De esta manera  $\hat{\theta}$  es un valor común en todos los intervalos, y eso sirve para especificar su localización. El conjunto completo de estos intervalos es equivalente a la función de verosimilitud relativa  $R(\theta; \underline{x})$ , y reproduce su gráfica. La desviación de  $\hat{\theta}$  del centro geométrico del intervalo nos da una idea del sesgo de la función de verosimilitud y por ende de la vecindad de plausibilidad dentro del intervalo.

Usualmente consideraremos 25 %, 15 % y 3.6 % intervalos o regiones de verosimilitud, ya que los  $IV(c)$  con  $c = 0.258, 0.147$  y  $0.036$  tienen una probabilidad de cobertura aproximada del 90 %, 95 % y 99 % respectivamente. No desarrollaremos este resultado en el artículo, ya que es un tema amplio, para mayor referencia ver Kalbfleisch[3] y Sprott[9].

## 4. Simulación

En muchas ocasiones determinar los estimadores de máxima verosimilitud no es analíticamente calculable o sencillo realizarlo, por lo que se recurre a métodos numéricos de aproximación por simulación, los cuales son material de la presente sección. Como primer paso, generamos, con la ayuda del software MATLAB, un simulador del experimento para observar procesos de Poisson no homogéneos y estimar el o los parámetros involucrados. Paso segundo y de principal importancia, mostrar propiedades frecuentistas de la verosimilitud para realizar inferencia sobre el o los parámetros desconocidos mediante la estimación de los intervalos de verosimilitud y la probabilidad de cobertura real de los mismos que se aproximan a los intervalos de confianza, seleccionados con fines evaluativos y comparativos.

## 5. Proceso de Poisson no homogéneo en meteorología

El siguiente caso es de mayor importancia para lograr los objetivos de este artículo. En el aislamos dos propósitos, el primero es reflejar los argumentos que soportan, de alguna forma, las condiciones que se deben cumplir para que la corrida de tiempos puedan ser modelados como un proceso de Poisson no homogéneo con función de intensidad gama; el segundo, para constrastrar las diferentes estimaciones de los intervalos de verosimilitud.

Los eventos observados para este caso de estudio, están integrados en una base de datos en Excel; éstos refieren las mediciones de las temperaturas durante los meses de diciembre del año 2005 a

diciembre del año 2006, registradas por la estación meteorológica ubicada en el Centro de Investigación de Matemáticas (CIMAT) en el municipio de Guanajuato, Gto.

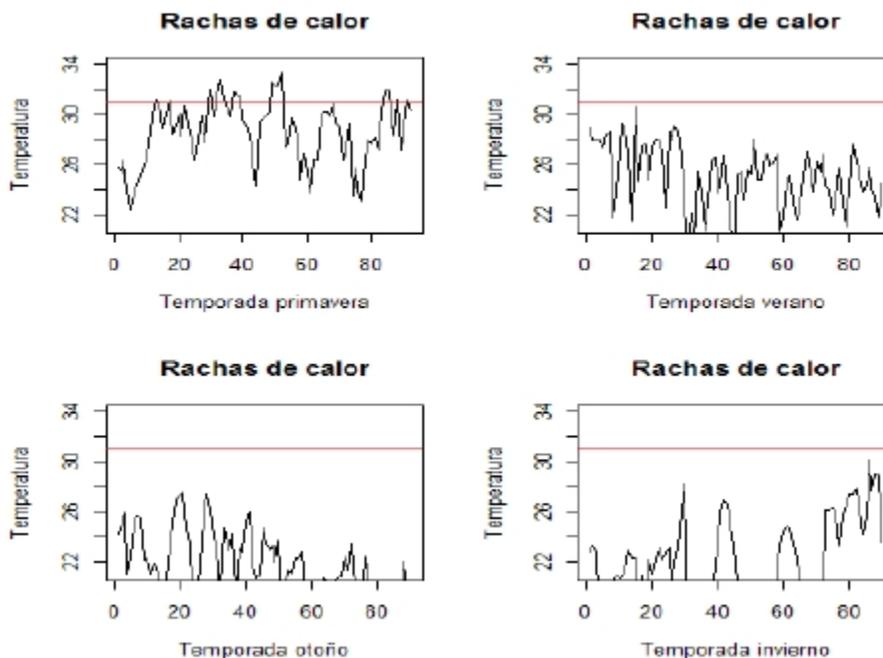
La base de datos muestra las mediciones registradas cada media hora aproximadamente de: temperaturas tanto interior como exterior en grados Celcius, dirección, velocidad y rachas del viento en millas por hora, porcentaje de humedad interior como exterior, presión barométrica y lluvia en pulgadas, así como presión barométrica al nivel del mar, y otros. Nuestro interés en este caso, es el estudio de la posible presencia de un proceso de Poisson en los registros de temperatura exterior a través del tiempo, por lo que enfocaremos nuestro estudio a esta parte de la información de la base de datos.

Iniciaremos el planteamiento de este caso definiendo la variable de estudio que nos ocupa, rachas de calor. Una *racha de calor* son las temperaturas máximas diarias que exceden los  $31^{\circ}$  C. Los objetivos a tratar en este caso de estudio son:

**Objetivo 1.** Estudiar la intensidades de las rachas de calor durante todo el año.

**Objetivo 2.** Determinar la tendencia de las rachas de calor, esto es, determinar si la intensidad de las rachas de calor es constante o no; si no es constante, verificar si presenta alguna tendencia con respecto al tiempo.

La Figura 1 nos muestra el comportamiento de las rachas de calor por estadios. Es muy notoria la presencia de las mismas en el estadio de primavera principalmente, mientras que en el resto del año es nula.



**Figura 1.** Rachas de calor en los diferentes estadios del año.

Los datos observados fueron registrados de manera continua en un intervalo de tiempo fijo; en el conjunto de observaciones, denotaremos por  $N(t)$  al proceso que cuenta el número de rachas de calor presentadas durante el intervalo de tiempo fijo  $[0, 365]$ , la sucesión  $S_n$ ,  $n = 1, 2, \dots, k$  representa los tiempos en los que ocurre una racha de calor, los cuales estarán representados por el vector de posiciones:

$$\underline{S} = (13, 17, 30, 32, 33, 34, 37, 38, 39, 49, 50, 51, 52, 84, 85, 88, 91);$$

donde cada entrada del vector representa el día en el que se presentó una racha de calor durante el intervalo de tiempo  $[0, 365]$ . Así mismo los tiempos entre rachas consecutivas  $T_n = S_n - S_{n-1}$ ,  $n = 1, 2, \dots, k$ , estarán expresados por el vector:

$$\underline{T} = (4, 13, 2, 3, 10, 32, 3, 3);$$

donde a partir del primer tiempo de ocurrencia de una racha de calor, primer entrada con valor 13 del vector  $\underline{S}$ , el valor de cada entrada del vector  $\underline{T}$  representa el tiempo que hay entre una racha de calor y la subsecuente racha.

Analizando el vector de tiempos entre rachas consecutivas  $\underline{T}$ , observamos que éstos cumplen con una de las condiciones restrictivas del proceso de Poisson, esta es, son independientes. Ya que, el tiempo de duración de una racha de calor no depende del tiempo de duración de otra.

Realizamos el histograma de la distribución que presentan estos tiempos y una prueba de ajuste de distribución mediante un gráfico cuantil-cuantil simulando una exponencial, y observamos que la distribución de los datos reales y la de los datos simulados son muy similares, éstas razones sustentan que una distribución exponencial es razonable para los datos observados, esto es, cumplen con la característica principal que sustenta que el proceso  $N$  es de Poisson.

Por el análisis exploratorio de los datos en la Figura 1, notamos que las rachas de calor se presentan con mayor frecuencia en la primera mitad del período de primavera alcanzando un máximo, para luego descender de manera paulatina en la segunda mitad del mismo período, y siendo nula su presencia en el resto del año. Razón por la que el proceso de Poisson es no homogéneo, por tanto propondremos como hipótesis de la función de intensidad del modelo estadístico la expresión:  $\lambda(t) = \frac{1}{\Gamma(\alpha)\beta^\alpha} t^{\alpha-1} \exp(-t/\beta)$ ,  $\alpha, \beta > 0$ ; donde  $\lambda(s)$  es la función de intensidad de las rachas de calor y,  $\alpha$  y  $\beta$  son parámetros desconocidos. La función de verosimilitud nos queda:

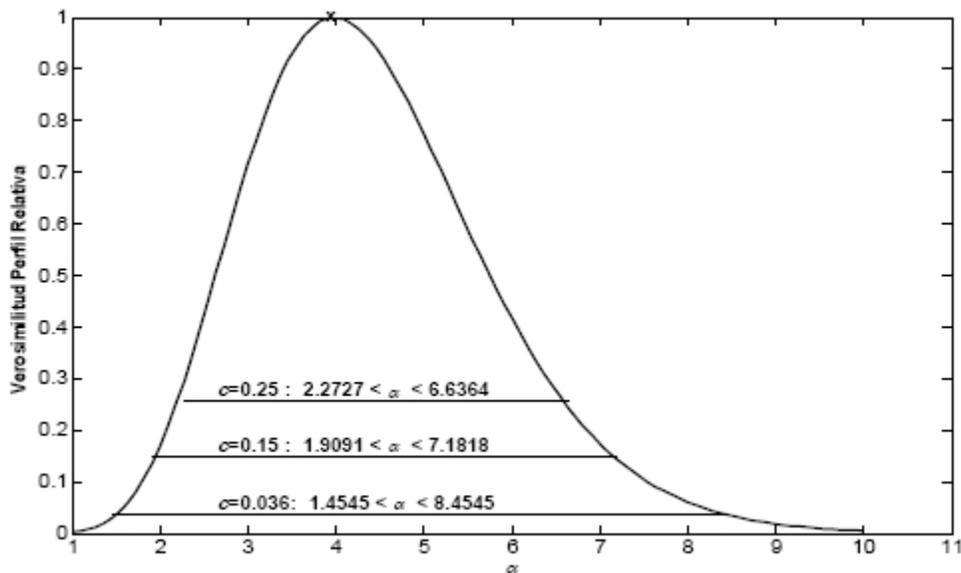
$$\begin{aligned} L(\alpha, \beta; \underline{S}) &\propto \prod_{k=1}^N \left( \lambda(s_k; \alpha, \beta) e^{-\int_{s_{k-1}}^{s_k} \lambda(t; \alpha, \beta) dt} \right); \\ &= \prod_{k=1}^N \left( \frac{1}{\Gamma(\alpha)\beta^\alpha} s_k^{\alpha-1} \exp(-s_k/\beta) \right) e^{-\frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{T_0} t^{\alpha-1} \exp(-t/\beta) dt}. \end{aligned}$$

Considerando ahora el logaritmo, tenemos que la función log-verosimilitud es

$$\begin{aligned}\ell(\alpha, \beta; \underline{S}) &= \log L(\alpha, \beta; \underline{S}); \\ &= \sum_{k=1}^N \log \left( \frac{1}{\Gamma(\alpha)\beta^\alpha} s_k^{\alpha-1} \exp(-s_k/\beta) \right) - \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{T_0} t^{\alpha-1} \exp(-t/\beta) dt,\end{aligned}$$

ya que la expresión de la función de verosimilitud resulta ser complicada, procederemos a obtener la estimación de verosimilitud para los parámetros  $\alpha$  y  $\beta$  de manera numérica. Los resultados obtenidos fueron  $\hat{\alpha} = 3.98$  y  $\hat{\beta} = 12.22$ . Por lo que la función de intensidad de las rachas de calor está expresada por  $\lambda(t; \alpha, \beta) = \frac{1}{124110.6} t^{2.98} \exp(-t/12.22)$ .

Ya que nuestro modelo contempla dos parámetros, determinaremos ahora la función de verosimilitud perfil para aproximar los intervalos de verosimilitud para cada uno de ellos. Considerando primeramente a  $\alpha$  como parámetro de interés. La Figura 2 nos muestra la función de verosimilitud perfil de  $\alpha$ , con sus respectivos intervalos de verosimilitud para los valores de  $c = 0.25, 0.15$  y  $0.036$  que como vimos en la Sección 3 suelen estar asociados a niveles del 90, 95 y 99 % de confianza:



**Figura 2.** Función de verosimilitud perfil relativa para  $\alpha$ .

Como lo muestra la Figura 2, la función de verosimilitud perfil para  $\alpha$  se presenta de manera asimétrica con respecto al valor de  $\hat{\alpha} = 3.98$ , para valores de  $\alpha$  menores que 1 y mayores que 10 son implausibles a la luz de los datos observados. Ahora considerando a  $\beta$  como parámetro de interés, se observa una clara asimetría con cola pesada hacia la derecha con respecto a  $\hat{\beta} = 12.22$ .

Tratar de dar una interpretación física a los valores de los *EMV* no tiene sentido aquí. En cambio lo que verdaderamente tiene sentido es hablar acerca de un rango de la forma que la función de

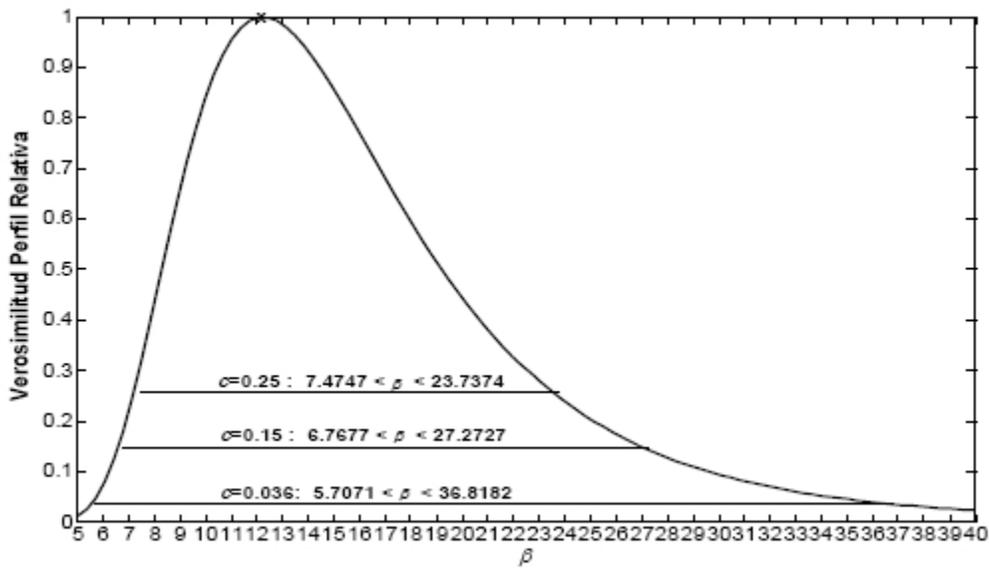


Figura 3. Función de verosimilitud perfil relativa para  $\beta$ .

intensidad toma. Esto con el objetivo de predecir que futuras mediciones de rachas de calor tendrán un comportamiento dentro de una franja establecida, que es precisamente lo que la Figura 4 nos muestra. Marca una función de intensidad mínima y una máxima a partir de la función de intensidad media obtenida en este caso de estudio.

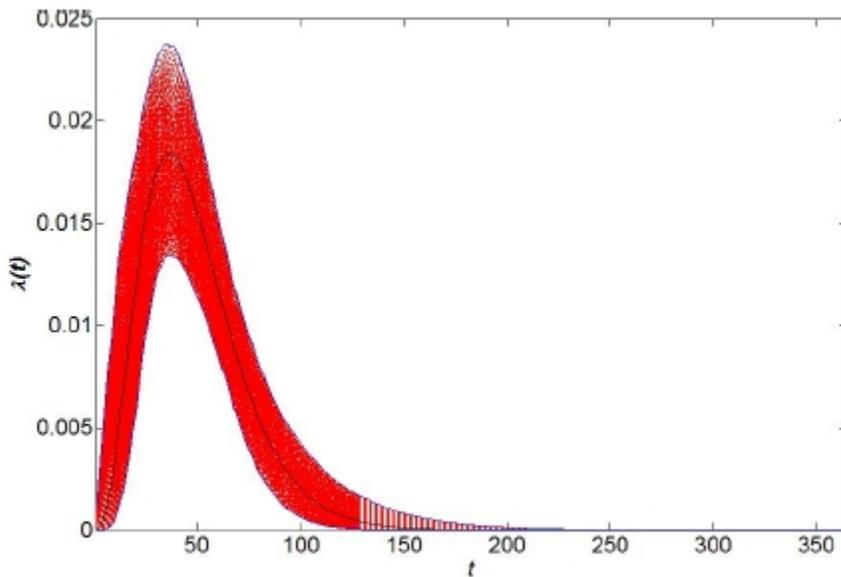


Figura 4. Rangos de la función de intensidad gama  $(\alpha, \beta)$  del PPNH.

## 6. Conclusiones

Aunque el objetivo principal del presente artículo fue el desarrollo de la teoría de verosimilitud para producir inferencia estadística sobre los parámetros de procesos de Poisson, no tendría mayor relevancia si no encontráramos en nuestro entorno la aplicación de esta formulación a un problema científico, ya que es la culminación práctica de una teoría desarrollada de manera abstracta. Como puede observarse, para alcanzar dichos objetivos se ha desarrollado un algoritmo para generar un proceso de Poisson no homogéneo con función de intensidad gama, basado en las propiedades y condiciones probabilísticas que éste debe cumplir; el cual pudimos aplicar a una situación real. Se estimó puntualmente cada uno de los parámetros involucrados en el proceso y finalmente se realizó inferencia sobre éstos mediante la teoría de verosimilitud. Nótese que las condiciones establecidas son muy generales por lo que el algoritmo puede considerar una amplia variedad de funciones de intensidad  $\lambda(t)$ .

La implementación computacional permitió realizar teoría en inferencia estadística en procesos de Poisson tanto homogéneos como no homogéneos, incluyendo los diferentes tipos de inspección en las cuales se pueden llevar a cabo estos procesos, esto es, en intervalos de tiempo fijo, como aleatorio o inspección continua con un plan de muestreo y realizaciones discretas para mayor referencia consultar Basawa[1]. La programación implementada fue llevada a cabo en MATLAB y soporta cualquier variación a las condiciones probabilísticas del proceso de Poisson, como lo son: los parámetros involucrados y la función de intensidad. De igual manera, estadísticamente hablando soporta cualquier variación a los niveles de  $c$ , al momento de determinar los intervalos de verosimilitud.

## 7. Referencias

- [1] Basawa, I. & Prakasa, R., (1980), *Statistical Inference for Stochastic Process*, Academic Press Inc.
- [2] Billingsley, P., (1979), *Probability and measure*, John Wiley & Sons, Inc.
- [3] Kalbfleisch, J., (1985), *Probability and Statistical Inference*, Springer-Verlag.
- [4] Karlin, S., (1975), *A first course in sthochastic processes*, Academic Press Inc.
- [5] Karlin, S. & Howard, T., (1998), *An introduction to Sthochastic Modeling*, Academic Press Inc.
- [6] Kulkarni, V., (1995), *Modeling and analisis of stochastic systems*, Chapman & Hall.
- [7] Pranab K. & Singer, J., (1993), *Large sample methods in statistics and introduction with applications*, Chapman & Hall.
- [8] Resnick, S., (1992), *Adventures in Stochastic Processes*, d. Birkhäuser Boston.
- [9] Sprott. D., (2000), *Statistical Inference in Science*, Springer.

# **T**aller de **Topología**



# Dinámica en hiperespacios

Gerardo Acosta  
Instituto de Matemáticas  
Universidad Nacional Autónoma de México  
gacosta@matem.unam.mx

## Resumen

*Para un continuo  $X$ , consideramos los hiperespacios  $2^X$  y  $C(X)$ , de los subconjuntos cerrados de  $X$  y de los subcontinuos de  $X$ , respectivamente, ambos con la topología inducida por la métrica de Hausdorff. En el presente artículo mostramos que si  $X$  es un continuo que contiene un arco libre y  $f: X \rightarrow X$  es una función continua, entonces la función inducida  $C(f): C(X) \rightarrow C(X)$  no es transitiva. Esto generaliza los resultados probados en [2, p. 683], [10, Corolario 29] y [25, Corolario 3.2]. También mostramos un continuo  $X$ , de dimensión infinita, y un homeomorfismo  $f: X \rightarrow X$  de forma que la función inducida  $C(f)$  es transitiva. Con esto respondemos la pregunta realizada en [2, p. 684].*

## 1. Introducción

El presente artículo, enmarcado en una subrama de la Topología llamada Teoría de Continuos e Hiperespacios, así como en los Sistemas Dinámicos Discretos, tiene como objetivo mostrar algunos de los resultados que aparecen en [1]. Dividimos este trabajo en cuatro secciones. En la Sección 2, presentamos las nociones de órbita y  $\omega$ -conjunto límite de un punto, bajo una función continua. En la Sección 3 damos las nociones de continuo y de hiperespacios de un continuo. Vemos que a los hiperespacios de un continuo se les puede dar una métrica, bajo la cual, éstos son continuos. Posteriormente consideramos algunos resultados sobre la convergencia de sucesiones en los hiperespacios.

También, en la Sección 3, mostramos que toda función continua  $f$ , definida entre dos continuos  $X$  y  $Y$ , induce de manera natural, dos funciones continuas entre los respectivos hiperespacios de  $X$  y  $Y$ . Como consecuencia de esto, la convergencia de una sucesión  $(A_n)_n$  en un hiperespacio de  $X$ , implica la convergencia de la sucesión  $(f(A_n))_n$ , en el respectivo hiperespacio de  $Y$ . Finalmente, en la Sección 4, consideramos la transitividad topológica, así como su relación con la órbita y el  $\omega$ -conjunto límite de un punto, bajo una función continua, cuyo dominio es un espacio métrico, compacto y sin puntos aislados. También mostramos que todo sistema dinámico, definido sobre un continuo, induce dos sistemas dinámicos definidos sobre los respectivos hiperespacios de dicho continuo. Además probamos que si un continuo  $X$  contiene un arco libre (ver Definición 4) y  $f: X \rightarrow X$  es una función continua, entonces la función inducida en el hiperespacio de los subcontinuos de  $X$ , no es transitiva (ver Teorema 12). Este es el teorema principal del artículo.

Para un espacio métrico  $X$ , con métrica  $d$ , si  $A \subset X$  entonces  $\text{int}_X(A)$  y  $\text{cl}_X(A)$  representan el interior y la cerradura de  $A$  en  $X$ , respectivamente. Si  $p \in X$  y  $\varepsilon > 0$ , entonces:

$$B_X(p, \varepsilon) = \{x \in X : d(x, p) < \varepsilon\}$$

es la bola abierta en  $X$ , con centro en  $p$  y radio  $\varepsilon$ . Las letras  $\mathbb{N}$  y  $\mathbb{R}$ , representan los conjuntos de los números naturales y de los números reales, respectivamente. Supongamos que  $X$  es un espacio topológico, que  $a \in X$  y que  $(a_n)_n$  es una sucesión en  $X$ . Entonces el símbolo  $a_n \rightarrow a$ , indica que la sucesión  $(a_n)_n$  converge al punto  $a$ . En el caso en que  $(a_n)_n$  sea una sucesión de números reales, el símbolo  $a_n \rightarrow \infty$ , indica que la sucesión  $(a_n)_n$  no está acotada superiormente.

## 2. El $\omega$ -conjunto límite

En la presente sección veremos una serie de nociones y resultados, propios de los Sistemas Dinámicos Discretos. Recordemos que un *sistema dinámico* es una pareja  $(X, f)$ , donde  $X$  es un espacio métrico y  $f$  es una función continua de  $X$  en  $X$ . Si  $X$  es un espacio topológico y  $f: X \rightarrow X$  es una función continua, entonces la pareja  $(X, f)$  se llama un *sistema dinámico topológico*.

Si  $X$  es un espacio topológico y  $f: X \rightarrow X$  es una función continua, consideramos que  $f^0$  es la función identidad sobre  $X$ ,  $f^1 = f$  y, para cada  $n \in \mathbb{N}$ , que  $f^{n+1} = f \circ f^n$ . Si  $p \in X$ , es natural investigar tanto el conjunto:

$$\text{orb}(p, f) = \{f^k(p) : k \in \mathbb{N} \cup \{0\}\}$$

como el subconjunto  $\omega(p, f)$  de los elementos  $x$  en  $X$ , para los que existe una sucesión  $(k_n)_n$  en  $\mathbb{N} \cup \{0\}$  tal que  $k_n \rightarrow \infty$  y  $f^{k_n}(p) \rightarrow x$ . Informalmente, decimos que  $\omega(p, f)$  es el conjunto de los puntos límites de  $\text{orb}(p, f)$ . En la siguiente definición, presentamos los nombres de los conjuntos anteriores.

**Definición 1.** Sean  $X$  un espacio topológico y  $f: X \rightarrow X$  una función continua. Si  $p \in X$ , entonces  $\text{orb}(p, f)$  es la **órbita de  $p$  bajo  $f$** , mientras que  $\omega(p, f)$  es el  **$\omega$ -conjunto límite de  $p$  bajo  $f$** .

En el siguiente resultado, presentamos las propiedades más importantes del  $\omega$ -conjunto límite. Su demostración se deja como ejercicio al lector.

**Teorema 1.** Sean  $X$  un espacio métrico y compacto,  $f: X \rightarrow X$  una función continua y  $p \in X$ . Entonces:

- 1)  $\omega(p, f) \neq \emptyset$ ;
- 2)  $\omega(p, f) = \bigcap_{m \geq 0} \text{cl}_X(\{f^k(p) : k \geq m\})$ ;
- 3)  $\omega(p, f)$  es un subconjunto cerrado de  $X$ ;
- 4)  $f(\omega(p, f)) = \omega(p, f)$ ;
- 5) si  $x \in \text{orb}(p, f)$ , entonces  $\omega(p, f) = \omega(x, f)$ .

Si  $X$  es un espacio topológico y  $f: X \rightarrow X$  es una función continua, entonces un subconjunto  $A$  de  $X$  se llama *f*-invariante o bien *invariante bajo f* si  $f(A) \subset A$ . En el caso en que  $f(A) = A$ , decimos que  $A$  es *fuertemente f*-invariante o bien *fuertemente invariante bajo f*. En éstos términos, las propiedades 1), 3) y 4) del Teorema 1 dicen que, para cada  $p \in X$ ,  $\omega(p, f)$  es un subconjunto cerrado y no vacío de  $X$  que, además, es fuertemente invariante bajo  $f$ .

### 3. Hiperespacios y funciones inducidas

En la presente sección veremos ahora una serie de nociones y resultados, propios de la Teoría de Continuos e Hiperespacios. Por un *continuo* entendemos un espacio métrico, compacto, conexo y no vacío. Si  $X$  es un continuo, consideramos las familias:

$$2^X = \{A \subset X : A \text{ es no vacío y cerrado en } X\}$$

y

$$C(X) = \{A \in 2^X : A \text{ es conexo}\}.$$

A  $2^X$  y  $C(X)$  les damos una métrica  $H$  que se llama la *métrica de Hausdorff*. Para definirla en  $2^X$  procedemos como sigue: si  $A \in 2^X$  y  $\varepsilon > 0$ , consideramos primero el conjunto

$$N_X(A, \varepsilon) = \bigcup_{a \in A} B_X(a, \varepsilon).$$

Notemos que, para cada  $A \in 2^X$  y toda  $\varepsilon > 0$ , el conjunto  $N_X(A, \varepsilon)$  es un abierto en  $X$  que contiene a  $A$ . Si  $A, B \in 2^X$  definimos ahora:

$$H(A, B) = \inf \{\varepsilon > 0 : A \subset N_X(B, \varepsilon) \text{ y } B \subset N_X(A, \varepsilon)\}.$$

En [8, Proposición 2.1] se prueba que  $H$  es una métrica en  $2^X$ . Como  $C(X) \subset 2^X$ , restringiendo la métrica  $H$  a  $C(X)$ , hacemos de éste un espacio métrico. Ahora bien, cuando a  $2^X$  y a  $C(X)$  les damos la topología inducida por la métrica de Hausdorff, decimos que  $2^X$  y  $C(X)$  son *hiperespacios de X*. A  $2^X$  le llamamos el *hiperespacio de los subconjuntos cerrados de X*, mientras que  $C(X)$  es el *hiperespacio de los subcontinuos de X*.

El siguiente es un resultado muy socorrido para determinar si dos elementos de  $2^X$  están a distancia menor que un número  $\varepsilon$  dado. Su demostración la dejamos como ejercicio al lector.

**Teorema 2.** Supongamos que  $X$  es un continuo. Si  $A, B \in 2^X$  y  $\varepsilon > 0$ , entonces  $H(A, B) < \varepsilon$  si y sólo si  $A \subset N_X(B, \varepsilon)$  y  $B \subset N_X(A, \varepsilon)$ .

Dejamos también como ejercicio al lector, la prueba del siguiente resultado.

**Teorema 3.** Supongamos que  $X$  es un continuo. Si  $A \in 2^X$  y  $U$  es un abierto en  $X$  tal que  $A \subset U$ , entonces existe  $\varepsilon > 0$  tal que  $N_X(A, \varepsilon) \subset U$ .

Intuitivamente, si  $A \in 2^X$  y  $\varepsilon > 0$  es pequeño, entonces  $N_X(A, \varepsilon)$  es un subconjunto abierto de  $X$  que se parece a  $A$ . El Teorema 2 dice, por tanto, que dos elementos  $A$  y  $B$  de  $2^X$  están muy cerca si básicamente uno se encuentra empalmado con respecto del otro. Como se menciona en [8, p. 22], “esta idea geométrica, de la métrica de Hausdorff, es buena pero tenemos que notar que, por ejemplo, si  $A$  es un disco en el plano, se pueden dar conjuntos finitos tan cercanos a  $A$  como se quiera, simplemente se toma una cuadrícula muy fina dentro del disco y se toma como conjunto finito al conjunto de los cruces de la cuadrícula”.

No es nuestra intención presentar un estudio detallado de los hiperespacios. Para una introducción a esta teoría, el lector interesado puede consultar [8]. Por ejemplo, en [8, Teorema 4.2 y Corolario 4.3] se prueba que si  $X$  es un continuo, entonces  $2^X$  y  $C(X)$  son compactos. En [8, Corolario 6.11 y Corolario 6.12] se prueba que si  $X$  es un continuo, entonces  $2^X$  y  $C(X)$  son conexos por trayectorias. Esto implica que los hiperespacios  $2^X$  y  $C(X)$  de un continuo  $X$  son, a su vez, continuos.

Si  $(A_n)_n$  es una sucesión de elementos en  $2^X$  y  $A \in 2^X$ , decimos que  $(A_n)_n$  converge a  $A$  y escribimos  $A_n \rightarrow A$ , si para cada  $\varepsilon > 0$  existe  $N \in \mathbb{N}$  tal que  $H(A_n, A) < \varepsilon$ , para cada  $n \geq N$ . En el siguiente teorema, presentamos una serie de propiedades de la convergencia en  $2^X$ . Su demostración la dejamos como ejercicio al lector.

**Teorema 4.** Supongamos que  $X$  es un continuo y que  $A, B \in 2^X$ . Sean  $(A_n)_n$  y  $(B_n)_n$  dos sucesiones de elementos en  $2^X$ , tales que  $A_n \rightarrow A$  y  $B_n \rightarrow B$ . Entonces:

- 1) si  $A_n \subset B_n$  para una infinidad de elementos  $n \in \mathbb{N}$ , entonces  $A \subset B$ ;
- 2)  $(A_n \cup B_n)_n$  es una sucesión en  $2^X$  tal que  $A_n \cup B_n \rightarrow A \cup B$ ;
- 3) si  $A_n \cap B_n \neq \emptyset$  para una infinidad de elementos  $n \in \mathbb{N}$ , entonces  $A \cap B \neq \emptyset$ ;
- 4) si  $A_{n+1} \subset A_n$  para cada  $n \in \mathbb{N}$ , entonces  $A = \bigcap_{n=1}^{\infty} A_n$ ;
- 5) si  $A_n \in C(X)$  para cada  $n \in \mathbb{N}$ , entonces  $A \in C(X)$ .

Una sucesión  $(A_n)_n$  tal que  $A_{n+1} \subset A_n$  para cada  $n \in \mathbb{N}$ , se llama *decreciente*. La parte 4) del Teorema 4, dice que si  $(A_n)_n$  es una sucesión decreciente en  $2^X$ , entonces  $(A_n)_n$  converge y su límite es  $\bigcap_{n=1}^{\infty} A_n$ . En tal situación, el límite de la sucesión está contenido en cada miembro de la sucesión.

Presentamos ahora el siguiente resultado, en cuya demostración, utilizamos el teorema anterior.

**Teorema 5.** Supongamos que  $X$  es un continuo con más de un punto y que  $A \in 2^X$ . Entonces  $\Lambda(A) = \{B \in 2^X : A \subset B\}$  es cerrado en  $2^X$ . Además si  $k \in \mathbb{N}$  y  $A_1, A_2, \dots, A_k \in 2^X$ , entonces  $\Lambda = \Lambda(A_1) \cup \Lambda(A_2) \cup \dots \cup \Lambda(A_k)$  no es denso en  $2^X$ .

**Demostración.** Para ver que el conjunto  $\Lambda(A)$  es cerrado en  $2^X$ , sea  $B \in \text{cl}_{2^X}(\Lambda(A))$ . Entonces  $B \in 2^X$  y, además, existe una sucesión  $(B_n)_n$  en  $\Lambda(A)$  tal que  $B_n \rightarrow B$ . Como cada  $B_n$  se encuentra en  $\Lambda(A)$ , tenemos que  $A \subset B_n$ , para cada  $n \in \mathbb{N}$ . Si definimos  $A_n = A$  para cada  $n \in \mathbb{N}$

entonces  $A_n \rightarrow A$  y, además,  $A_n \subset B_n$  para cada  $n \in \mathbb{N}$ . Luego, por la parte 1) del Teorema 4,  $A \subset B$ . Entonces  $B \in \Lambda(A)$ . Esto muestra que  $B \in \Lambda(A)$ . Con esto probamos que  $\Lambda(A)$  es cerrado en  $2^X$ .

Tomemos ahora  $k \in \mathbb{N}$  y  $A_1, A_2, \dots, A_k \in 2^X$ . Hagamos  $\Lambda = \bigcup_{i=1}^k \Lambda(A_i)$ . Como, para cada  $j \in \{1, 2, \dots, k\}$ ,  $\Lambda(A_j)$  es cerrado en  $2^X$ , tenemos que  $\Lambda$  es cerrado en  $2^X$ . Sea

$$M = \{j \in \{1, 2, \dots, k\} : A_j \text{ tiene un solo elemento}\}.$$

Por ser  $X$  es un espacio métrico, conexo y con más de un punto,  $X$  posee una cantidad no numerable de puntos. Entonces existe  $x \in X - \bigcup_{j \in M} A_j$ . Notemos que  $\{x\} \in 2^X - \Lambda = 2^X - \text{cl}_{2^X}(\Lambda)$ , así que  $\Lambda$  no es denso en  $2^X$ .  $\square$

Una demostración similar a la dada en el teorema anterior, puede aplicarse para probar el siguiente resultado.

**Teorema 6.** Supongamos que  $X$  es un continuo con más de un punto y que  $A \in C(X)$ . Entonces  $\Lambda(A) = \{B \in C(X) : A \subset B\}$  es cerrado en  $C(X)$ . Además si  $k \in \mathbb{N}$  y  $A_1, A_2, \dots, A_k \in C(X)$ , entonces el conjunto  $\Lambda = \Lambda(A_1) \cup \Lambda(A_2) \cup \dots \cup \Lambda(A_k)$  no es denso en  $C(X)$ .

El siguiente resultado indica que toda función continua, definida entre continuos, induce de manera natural dos funciones continuas entre los respectivos hiperespacios de los continuos en cuestión.

**Teorema 7.** Sean  $X$  y  $Y$  dos continuos y  $f: X \rightarrow Y$  una función continua. Para cada  $A \in 2^X$  definimos  $2^f(A) = f(A)$ , la imagen de  $A$  bajo  $f$ . Entonces:

- 1)  $2^f(A) \in 2^Y$  y la función  $2^f: 2^X \rightarrow 2^Y$  es continua;
- 2) si  $C(f) = (2^f)_{|C(X)}$ , entonces  $C(f)$  es una función continua de  $C(X)$  a  $C(Y)$ ;
- 3) si  $f$  es un homeomorfismo, entonces  $2^f$  y  $C(f)$  son homeomorfismos.

**Definición 2.** Sean  $X$  y  $Y$  dos continuos y  $f: X \rightarrow Y$  una función continua. A las funciones continuas  $2^f: 2^X \rightarrow 2^Y$  y  $C(f): C(X) \rightarrow C(Y)$  definidas en el Teorema 7 se les llama las **funciones inducidas** de  $f$ .

Supongamos que  $X$  y  $Y$  son dos continuos y que  $f: X \rightarrow Y$  una función continua. Si  $A \in 2^X$  y  $(A_n)_n$  es una sucesión en  $2^X$  tal que  $A_n \rightarrow A$  entonces, por el Teorema 7,  $(2^f(A_n))_n$  es una sucesión en  $2^Y$  que converge a  $2^f(A)$ . Como  $2^f(A) = f(A)$  y, para cada  $n \in \mathbb{N}$ ,  $2^f(A_n) = f(A_n)$ , tenemos que:

- a) si  $A \in 2^X$  y  $(A_n)_n$  es una sucesión en  $2^X$  tal que  $A_n \rightarrow A$ , entonces  $(f(A_n))_n$  es una sucesión en  $2^Y$  tal que  $f(A_n) \rightarrow f(A)$ .

De manera similar tenemos el siguiente resultado.

- b) si  $A \in C(X)$  y  $(A_n)_n$  es una sucesión en  $C(X)$  tal que  $A_n \rightarrow A$ , entonces  $(f(A_n))_n$  es una sucesión en  $C(Y)$  tal que  $f(A_n) \rightarrow f(A)$ .

Supongamos ahora que  $X$  es un continuo, que  $f: X \rightarrow X$  es una función continua, y que  $(A_n)_n$  es una sucesión decreciente en  $2^X$ . Si  $k \in \mathbb{N}$  entonces, para cada  $n \in \mathbb{N}$ , tenemos que  $f^k(A_{n+1}) \subset f^k(A_n)$ . Luego  $(f^k(A_n))_n$  es una sucesión decreciente en  $2^X$ . Por la parte 4) del Teorema 4,  $f^k(A_n) \rightarrow \bigcap_{n=1}^{\infty} f^k(A_n)$ . Entonces el límite de la sucesión  $(f^k(A_n))_n$  está contenido en cada miembro de dicha sucesión. Notemos que si la sucesión decreciente  $(A_n)_n$  está en  $C(X)$  entonces, de manera similar,  $(f^k(A_n))_n$  es una sucesión decreciente en  $C(X)$ , cuyo límite  $\bigcap_{n=1}^{\infty} f^k(A_n)$  está contenido en cada miembro de dicha sucesión.

#### 4. Transitividad y dinámica en los hiperespacios

Consideremos el sistema dinámico  $(X, f)$ , donde  $X$  es un continuo. Como  $f: X \rightarrow X$  es una función continua, por el Teorema 7, las funciones  $2^f: 2^X \rightarrow 2^X$  y  $C(f): C(X) \rightarrow C(X)$  son continuas. Entonces las parejas  $(2^X, 2^f)$  y  $(C(X), C(f))$  son dos sistemas dinámicos. De esta manera, un sistema dinámico  $(X, f)$  induce dos sistemas dinámicos  $(2^X, 2^f)$  y  $(C(X), C(f))$ . Notemos además que:

$$(2^f)^k(A) = f^k(A), \quad \text{para cada } (A, k) \in 2^X \times \mathbb{N}$$

y

$$(C(f))^k(A) = f^k(A), \quad \text{para cada } (A, k) \in C(X) \times \mathbb{N}.$$

Luego, para toda  $A \in 2^X$ , la órbita de  $A$  bajo  $2^f$  es el conjunto:

$$\text{orb}(A, 2^f) = \{f^k(A) : k \in \mathbb{N} \cup \{0\}\}$$

y, para cada  $A \in C(X)$ , la órbita de  $A$  bajo  $C(f)$  es el conjunto:

$$\text{orb}(A, C(f)) = \{f^k(A) : k \in \mathbb{N} \cup \{0\}\}.$$

Es natural preguntarse si una propiedad dinámica en el sistema  $(X, f)$ , determina una propiedad dinámica (posiblemente la misma) en los respectivos sistemas  $(2^X, 2^f)$  y  $(C(X), C(f))$ , y viceversa. Como se prueba en la Sección 7 de [1], la dinámica de los sistemas  $(2^X, 2^f)$  y  $(C(X), C(f))$  es, en general, mucho más compleja que la dinámica de  $(X, f)$ . Un sistema  $(X, f)$  cuya dinámica es muy poco interesante, puede generar un sistema dinámico  $(C(X), C(f))$  muy caótico e interesante.

El estudio de la dinámica en los hiperespacios comenzó en 1975, con el artículo [3], aunque los autores restringieron su estudio al caso de sistemas dinámicos  $(X, f)$  donde  $f$  es un homeomorfismo. A partir del 2003 el estudio de la dinámica en los hiperespacios, particularmente para sistemas dinámicos  $(X, f)$ , donde  $X$  es un espacio métrico y localmente compacto y  $f: X \rightarrow X$  es una función continua, se intensificó. El lector interesado puede consultar los artículos [2], [5], [6], [7], [11], [12], [10], [13], [14], [15], [16], [17], [19], [20], [21], [22], [23], [24] y [25].

Una propiedad dinámica es la transitividad. De acuerdo al libro *Topological Dynamics* que en 1995 publicaron W. H. Gottschalk y G. A. Hedlund, dicha noción se debe a G. D. Birkhoff, quien la utilizó en 1920.

**Definición 3.** Sean  $X$  un espacio topológico y  $f: X \rightarrow X$  una función continua. Decimos que  $f$  es **topológicamente transitiva** si para cada par  $U$  y  $V$  de subconjuntos abiertos y no vacíos de  $X$ , existe  $k \in \mathbb{N}$  tal que  $f^k(U) \cap V \neq \emptyset$ .

Mientras no haya lugar a confusión, diremos que  $f$  es transitiva en lugar de topológicamente transitiva. Intuitivamente una función continua es transitiva, si posee puntos que eventualmente se mueven, bajo iteración, de una vecindad arbitrariamente pequeña a cualquier otra. Por tanto, el sistema dinámico no se puede descomponer como la unión de dos abiertos, ajenos e invariantes bajo la función.

Conviene mencionar que, en algunos textos, a las funciones transitivas se le llama *regionalmente transitivas*, *topológicamente ergódicas*, *topológicamente indescomponibles*, *topológicamente irreducibles* o bien *nomádicas*. En lo que respecta a esta sección, vamos a presentar primero dos propiedades de las funciones transitivas, cuando se definen en un espacio métrico y compacto. Para un estudio más profundo de la transitividad, recomendamos al lector el artículo [9].

**Teorema 8.** Si  $X$  es un espacio métrico y compacto y  $f: X \rightarrow X$  es transitiva, entonces  $f$  es suprayectiva.

**Demostración.** Como  $f$  es continua y  $X$  es compacto,  $f(X)$  es un subconjunto compacto del espacio  $X$ , que es  $T_2$ . Luego  $f(X)$  es cerrado en  $X$ . Supongamos ahora que  $f(X) \neq X$ . Entonces  $V = X - f(X)$  es un subconjunto abierto y no vacío de  $X$ . Como  $X$  es un subconjunto abierto y no vacío de  $X$  y  $f$  es transitiva, existe  $k \in \mathbb{N}$  tal que  $f^k(X) \cap V \neq \emptyset$ . Tomemos un punto  $y \in f^k(X) \cap V$ . Entonces existe  $v \in X$  tal que  $f^k(v) = y$ . Sea  $x \in f^{k-1}(v)$ . Notemos que  $x$  es un elemento de  $X$  tal que  $y = f^k(v) = f(f^{k-1}(v)) = f(x)$ . Luego  $f(x) = y \in V = X - f(X)$ , es decir,  $f(x) \notin f(X)$ . Como esto es un absurdo, deducimos que  $f(X) = X$ .  $\square$

**Teorema 9.** Supongamos que  $X$  es un espacio métrico, compacto y sin puntos aislados. Si  $f: X \rightarrow X$  es una función continua, entonces las siguientes afirmaciones son equivalentes:

- 1)  $f$  es transitiva;
- 2) existe  $p \in X$  tal que  $\text{orb}(p, f)$  es denso en  $X$ ;
- 3) existe  $q \in X$  tal que  $\omega(q, f) = X$ .

**Demostración.** La equivalencia entre las afirmaciones 1) y 2) se sigue de [18, Proposición 1.1], mientras que la equivalencia entre las afirmaciones 1) y 3), aparece probada en [4, Proposición 39, p. 155].  $\square$

Supongamos que  $X$  es un continuo y que  $f: X \rightarrow X$  es una función continua. Como ya indicamos, los hiperespacios  $2^X$  y  $C(X)$  de  $X$ , también son continuos. Entonces, por el Teorema 9, la transitividad de  $f$ ,  $2^f$  y  $C(f)$  equivale a encontrar un punto con órbita densa, o bien un punto cuyo  $\omega$ -conjunto límite sea  $X$ ,  $2^X$  o  $C(X)$ , según corresponda.

Consideremos ahora la siguiente definición.

**Definición 4.** Supongamos que  $X$  es un continuo y que  $A$  es un arco en  $X$ , con puntos extremos  $p$  y  $q$ . Decimos que  $A$  es un **arco libre en  $X$** , si el conjunto  $A - \{p, q\}$  es abierto en  $X$ .

Como indicamos en la Sección 1, nuestro objetivo es mostrar que si un continuo  $X$  contiene un arco libre, entonces para ninguna función continua  $f: X \rightarrow X$ , la función inducida  $C(f): C(X) \rightarrow C(X)$  es transitiva. Para probar esto el siguiente resultado, que aparece originalmente en [1, Teorema 4.3], es importante.

**Teorema 10.** Sean  $X$  un continuo y  $f: X \rightarrow X$  una función continua tal que  $2^f: 2^X \rightarrow 2^X$  es transitiva. Sea  $B \in 2^X$  tal que  $\omega(B, 2^f) = 2^X$ . Entonces, para cada  $n \in \mathbb{N} \cup \{0\}$ , tenemos que  $\text{int}_X(f^n(B)) = \emptyset$ .

**Demostración.** Supongamos, por el contrario, que existe  $n_0 \in \mathbb{N} \cup \{0\}$  tal que  $\text{int}_X(f^{n_0}(B)) \neq \emptyset$ . Para simplificar la notación, hagamos  $D = f^{n_0}(B)$ . Notemos que  $D = (2^f)^{n_0}(B)$ . Luego  $D \in \text{orb}(B, 2^f)$  y, como  $\omega(B, 2^f) = 2^X$ , por la propiedad 5) del Teorema 1,  $\omega(D, 2^f) = 2^X$ .

Como  $\text{int}_X(D) \neq \emptyset$ , existen  $x \in \text{int}_X(D)$  y  $\varepsilon > 0$  tales que  $B_X(x, \varepsilon) \subset D$ . Notemos que  $\{x\} \in 2^X = \omega(D, 2^f)$ , así que existe  $k \in \mathbb{N}$  tal que  $H(f^k(D), \{x\}) < \varepsilon$ . Por el Teorema 2, tenemos que:

$$f^k(D) \subset N_X(\{x\}, \varepsilon) = B_X(x, \varepsilon) \subset D.$$

Por tanto  $f^k(D) \subset D$ . Aplicando  $f^{kn}$  en ambos miembros de la contención anterior, sucede que:

$$f^{k(n+1)}(D) = f^{kn}(f^k(D)) \subset f^{kn}(D).$$

Entonces  $(f^{kn}(D))_n$  es una sucesión decreciente en  $2^X$ . Luego, por la parte 4) del Teorema 4, la sucesión  $(f^{kn}(D))_n$  converge en  $2^X$  al elemento

$$C = \bigcap_{n=1}^{\infty} f^{kn}(D).$$

Como  $f^{kn}(D) \rightarrow C$ , para cada  $s \in \mathbb{N} \cup \{0\}$  tenemos, por la continuidad de la función  $(2^f)^s$ , que

$$f^{kn+s}(D) \rightarrow f^s(C). \quad (1)$$

También, como  $f^{kn}(D) \rightarrow C$ , sucede que  $\omega(D, 2^{f^k}) = \{C\}$  y, por la propiedad 4) del Teorema 1,

$$\{f^k(C)\} = 2^{f^k}(\{C\}) = 2^{f^k}(\omega(D, 2^{f^k})) = \omega(D, 2^{f^k}) = \{C\}.$$

Luego  $f^k(C) = C$ . Consideremos ahora la órbita de  $D$  bajo  $2^f$ , a partir de su  $k$ -ésimo término. Es decir, consideremos el conjunto:

$$\mathcal{D} = \{f^m(D) : m \in \mathbb{N} \text{ y } m \geq k\}.$$

De dicho conjunto, los elementos que dejan residuo cero módulo  $k$ , constituyen la sucesión decreciente  $(f^{kn}(D))_n$ , que converge a  $C$ . Notemos que  $C \subset f^{kn}(D)$ , para cada  $n \in \mathbb{N}$ . Ahora bien, los elementos de  $\mathcal{D}$  que dejan residuo uno módulo  $k$ , constituyen la sucesión decreciente  $(f^{kn+1}(D))_n$

que, por (1), converge a  $f(C)$ . Por tanto  $f(C) \subset f^{kn+1}(D)$ , para cada  $n \in \mathbb{N}$ . Continuando, los elementos de  $\mathcal{D}$  que dejan residuo dos módulo  $k$ , constituyen la sucesión decreciente  $(f^{kn+2}(D))_n$  que, por (1), converge a  $f^2(C)$ . Además  $f^2(C) \subset f^{kn+2}(D)$ , para cada  $n \in \mathbb{N}$ . Después de aplicar  $k - 1$  veces este razonamiento, tenemos que los elementos de  $\mathcal{D}$  que dejan residuo  $k - 1$  módulo  $k$ , constituyen la sucesión decreciente  $(f^{kn+k-1}(D))_n$  que, por (1), converge a  $f^{k-1}(C)$ . Además  $f^{k-1}(C) \subset f^{kn+k-1}(D)$ , para cada  $n \in \mathbb{N}$ . Al llegar al paso  $k$ , tenemos que los elementos de  $\mathcal{D}$  que dejan residuo  $k$  módulo  $k$ , constituyen la sucesión decreciente  $(f^{kn+k}(D))_n$  que, por (1), converge a  $f^k(C)$ . Como  $f^k(C) = C$  y, además, dejar residuo  $k$  módulo  $k$  es equivalente a dejar residuo cero módulo  $k$ , luego de aplicar  $k$  veces el razonamiento anterior, obtenemos la sucesión decreciente  $(f^{kn}(D))_n$  con la que iniciamos. Con esto hemos probado que:

- a) si  $m$  es un número natural tal que  $m \geq k$  y  $m \equiv i \pmod{k}$  para alguna  $i \in \{0, 1, \dots, k - 1\}$ , entonces  $f^i(C) \subset f^m(D)$ .

Ahora bien, para cada  $i \in \{0, 1, \dots, k - 1\}$ , sea:

$$\Lambda_i = \{A \in 2^X : f^i(C) \subset A\}.$$

Hagamos también  $\Lambda = \Lambda_0 \cup \Lambda_1 \cup \dots \cup \Lambda_{k-1}$ . Por la segunda parte del Teorema 5,  $\Lambda$  no es denso en  $2^X$ . Afirmamos ahora que:

- b)  $\omega(D, 2^f) \subset \Lambda$ .

Para ver esto tomemos un elemento  $A \in \omega(D, 2^f)$ . Entonces  $A \in 2^X$  y, además, existe una sucesión  $(m_s)_s$  en  $\mathbb{N} \cup \{0\}$  tal que  $m_s \rightarrow \infty$  y  $f^{m_s}(D) \rightarrow A$ . Como  $m_s \rightarrow \infty$  podemos suponer, sin perder generalidad, que  $m_s \geq k$ , para cada  $s \in \mathbb{N}$ . Como todo número natural es congruente, módulo  $k$ , con alguno de  $0, 1, \dots, k - 1$ , existen una subsucesión  $(m_{s_l})_l$  de  $(m_s)_s$  y un número  $i \in \{0, 1, \dots, k - 1\}$ , de forma que  $m_{s_l} \equiv i \pmod{k}$ , para toda  $l \in \mathbb{N}$ . De esta manera  $(f^{m_{s_l}}(D))_l$  es una sucesión en  $2^X$  que converge a  $A$  y, por a),  $f^i(C) \subset f^{m_{s_l}}(D)$ , para cada  $l \in \mathbb{N}$ . Ahora bien, por la primera parte del Teorema 5,  $\Lambda_i$  es cerrado en  $2^X$ . Como  $(f^{m_{s_l}}(D))_l$  es una sucesión en  $\Lambda_i$  que converge a  $A$ , tenemos que  $A \in \Lambda_i \subset \Lambda$ . Esto prueba b).

Como  $\omega(D, 2^f) = 2^X$  tenemos, por b), que  $\Lambda = 2^X$ . Entonces  $\Lambda$  es denso en  $2^X$ . Con todo, tenemos que  $\Lambda$  es denso y, a la vez, no es denso en  $2^X$ . En vista de que esto es una contradicción, que provino de haber supuesto que existe  $n_0 \in \mathbb{N} \cup \{0\}$  tal que  $\text{int}_X(f^{n_0}(B)) \neq \emptyset$ , deducimos que  $\text{int}_X(f^n(B)) = \emptyset$ , para cada  $n \in \mathbb{N} \cup \{0\}$ .  $\square$

Se puede efectuar una prueba similar a la dada en el teorema anterior, para probar el siguiente resultado.

**Teorema 11.** Sean  $X$  un continuo y  $f: X \rightarrow X$  una función continua tal que  $C(f): C(X) \rightarrow C(X)$  es transitiva. Sea  $B \in C(X)$  tal que  $\omega(B, C(f)) = C(X)$ . Entonces, para cada  $n \in \mathbb{N} \cup \{0\}$ , tenemos que  $\text{int}_X(f^n(B)) = \emptyset$ .

Estamos en condiciones de probar el teorema principal de este artículo.

**Teorema 12.** Supongamos que  $X$  es un continuo y que  $f: X \rightarrow X$  es una función continua. Si  $X$  contiene un arco libre, entonces la función inducida  $C(f): C(X) \rightarrow C(X)$  no es transitiva.

**Demostración.** Supongamos, por el contrario, que  $C(f)$  es transitiva. Entonces existe  $B \in C(X)$  tal que  $\omega(B, C(f)) = C(X)$ . Supongamos que  $A$  es un arco libre en  $X$ , con puntos extremos  $p$  y  $q$ . Tomemos  $x, y \in A - \{p, q\}$  tales que  $x \neq y$ , así como el subarco  $C$  de  $A$ , con puntos extremos  $x$  y  $y$ . Notemos que  $A - \{p, q\}$  es un subconjunto abierto de  $X$ , que contiene al elemento  $C$  de  $2^X$ . Entonces, por el Teorema 3, existe  $\varepsilon > 0$  tal que  $N_X(C, \varepsilon) \subset A - \{p, q\}$ . Como  $x \neq y$ , podemos tomar  $\varepsilon$  de forma que, además, satisface que  $B_X(x, \varepsilon) \cap B_X(y, \varepsilon) = \emptyset$ . Ahora bien, en vista de que  $C \in C(X) = \omega(B, C(f))$ , existe  $k \in \mathbb{N}$  tal que  $H(f^k(B), C) < \varepsilon$ . Por el Teorema 2, tenemos que  $f^k(B) \subset N_X(C, \varepsilon)$ . Por tanto  $f^k(B) \subset A - \{p, q\}$ . Esto implica que  $f^k(B)$  es un subcontinuo de  $A$ . Como  $A$  es un arco, tenemos que  $f^k(B)$  es un arco o bien un conjunto con un solo punto. Vamos a mostrar ahora que  $f^k(B)$  tiene más de un punto. Para ver esto notemos que, por el Teorema 3,  $C \subset N_X(f^k(B), \varepsilon)$ . Entonces  $x, y \in N_X(f^k(B), \varepsilon)$ . Esto implica que  $f^k(B) \cap B_X(x, \varepsilon) \neq \emptyset$  y  $f^k(B) \cap B_X(y, \varepsilon) \neq \emptyset$ . Como  $B_X(x, \varepsilon) \cap B_X(y, \varepsilon) = \emptyset$ , se tiene que  $f^k(B)$  tiene más de un punto. Entonces  $f^k(B)$  es un arco contenido en el arco libre  $A$  de  $X$ . Luego  $\text{int}_X(f^k(B)) \neq \emptyset$ . Como esto contradice el Teorema 11, deducimos que  $C(f)$  no es transitiva.  $\square$

Del Teorema 12 se infiere que si  $X$  es un continuo y  $f: X \rightarrow X$  es una función continua, entonces para que  $C(f)$  tenga la esperanza de ser transitiva, se requiere que  $X$  no contenga arcos, o bien que ningún arco en  $X$  sea libre.

**Teorema 13.** Existen un continuo  $X$ , de dimensión infinita, y un homeomorfismo  $f: X \rightarrow X$ , tales que la función inducida  $C(f): C(X) \rightarrow C(X)$  es transitiva.

**Demostración.** Para cada  $n \in \mathbb{Z}$ , hagamos  $J_n = [0, 1]$ . Consideraremos que cada  $J_n$  pose la topología usual que hereda de  $\mathbb{R}$ . Sea

$$X = \prod_{n \in \mathbb{Z}} J_n = \{ \hat{t} = (t_n)_{n \in \mathbb{Z}} : t_n \in [0, 1], \text{ para cada } n \in \mathbb{Z} \}.$$

Consideramos en  $X$  la métrica  $D$  definida, para  $\hat{t} = (t_n)_{n \in \mathbb{Z}}$  y  $\hat{s} = (s_n)_{n \in \mathbb{Z}}$  en  $X$ , como:

$$D(\hat{t}, \hat{s}) = \sum_{n \in \mathbb{Z}} \frac{|t_n - s_n|}{2^{|n|}}.$$

Es conocido que la topología en  $X$ , inducida por la métrica  $D$ , es la topología producto de  $X$ . Entonces  $(X, D)$  es un continuo de dimensión infinita. Consideremos ahora la función  $f: X \rightarrow X$  definida, para  $\hat{t} = (t_n)_{n \in \mathbb{Z}}$  en  $X$ , como:

$$f(\hat{t}) = \hat{s} = (s_n)_{n \in \mathbb{Z}},$$

donde  $s_n = t_{n-1}$ , para cada  $n \in \mathbb{Z}$ . No es difícil probar que  $f$  es un homeomorfismo. En la Sección 7 de [1], se prueba que  $f$  satisface las siguientes propiedades:

- 1)  $f$  es transitiva ([1, Teorema 7.3]);
- 2) el conjunto de puntos periódicos de  $f$ , es denso en  $X$  ([1, Teorema 7.4]);
- 3) la entropía topológica de  $f$  es infinita ([1, Teorema 7.6]);

4) la función inducida  $C(f): C(X) \rightarrow C(X)$  es transitiva ([1, Teorema 7.18]).

□

En vista del resultado anterior, es natural considerar las siguientes preguntas, que permanecen abiertas.

**Pregunta 1.** ¿Existen un continuo  $X$ , de dimensión finita, y una función continua  $f: X \rightarrow X$ , de forma que la función inducida  $C(f): C(X) \rightarrow C(X)$  sea transitiva?

**Pregunta 2.** ¿Existe una función continua  $f: [0, 1]^2 \rightarrow [0, 1]^2$  tal que la función inducida  $C(f)$  es transitiva?

## 5. Referencias

- [1] G. Acosta, A. Illanes y H. Méndez-Lango, *The transitivity of induced maps*, a publicarse en *Topology Appl.*
- [2] J. Banks, *Chaos for induced hyperspace maps*, *Chaos Solitons & Fractals*, 25 (3) (2005), 681–685.
- [3] W. Bauer y K. Sigmund, *Topological dynamics of transformations induced on the space of probability measures*, *Monatsh. Math.* 79 (1975), 81–92.
- [4] L. S. Block y W. A. Coppel, *Dynamics in One Dimension*, *Lecture Notes in Math.*, Vol. 1513, Springer-Verlag, New York, 1992.
- [5] J. S. Cánovas-Peña y G. S. López, *Topological entropy for induced maps*, *Chaos Solitons & Fractals*, 28 (2006), 979–982.
- [6] A. Fedeli, *On chaotic set-valued discrete dynamical systems*, *Chaos Solitons & Fractals*, 23 (4)(2005), 1381–1384.
- [7] R. Gu y W. Guo, *On mixing property in set-valued discrete systems*, *Chaos Solitons & Fractals* 28 (2006), 747–754.
- [8] A. Illanes, *Hiperespacios de Continuos*, *Aportaciones Matemáticas, Serie Textos Nivel Medio* Núm. 28. Sociedad Matemática Mexicana. México, 2004.
- [9] S. Kolyada, *Some Aspects of Topological Transitivity—A Survey*, *Grazner Math. Ber.*, Vol. 334 (1997), 3–35.
- [10] D. Kwietniak y P. Oprocha, *Topological entropy and chaos for maps induced on hyperspaces*, *Chaos Solitons & Fractals* 33 (2007), 76–86.
- [11] G. Liao, L. Wang y Y. Zhang, *Transitivity, mixing and chaos for a class of set-valued mappings*, *Sci. China Ser. A*, 49 (1)(2006), 1–8.

- [12] L. Liu, Y. Wang y G. Wei, *Topological entropy of continuous functions on topological spaces*, por aparecer en *Chaos Solitons & Fractals*.
- [13] X. Ma, B. Hou y G. Liao, *Chaos in hyperspace system*, por aparecer en *Chaos, Solitons & Fractals*.
- [14] J. S. C. Peña y G. S. López, *Topological entropy for induced hyperspace maps*, *Chaos, Solitons & Fractals* 28 (2005), 979–982.
- [15] A. Peris *Set-valued discrete chaos*, *Chaos Solitons & Fractals*, 26 (1) (2005), 19–23.
- [16] H. Román-Flores, *A note on transitivity in set-valued discrete systems*, *Chaos Solitons & Fractals*, 17 (1) (2003), 99–104.
- [17] H. Román-Flores y Y. Chalco-Cano, *Robinson's chaos in set-valued discrete systems*, *Chaos Solitons & Fractals*, 25 (1) (2005), 33–42.
- [18] S. Silverman, *On maps with dense orbits and the definition of chaos*, *Rocky Mountain Jour. Math.*, 22 (1992), 353–375.
- [19] Y. Wang y G. Wei, *Conditions ensuring that hyperspace dynamical systems contain subsystems topologically (semi-)conjugate to symbolic dynamical systems*, por aparecer en *Chaos, Solitons & Fractals*.
- [20] Y. Wang y G. Wei, *Properties of induced hyperspace dynamical systems of symbolic dynamical systems*, por aparecer en *Nonlinear Anal.*
- [21] Y. Wang y G. Wei, *Equivalence of topological weak mixing:  $s$ -set*, por aparecer en *Top. Methods Nonlin. Anal.*
- [22] Y. Wang y G. Wei, *Characterizing mixing, weak mixing and transitivity on induced hyperspace dynamical systems*, *Topology Appl.* 155 (2007), 56–68.
- [23] Y. Wang y G. Wei, *On metrization of the hit-or-miss topology using Alexandroff compactifications*, H. T. Nguyen y E. Miranda (Editores), *Random Sets and Imprecise Probabilities*, *Internat. J. Aprox. Reason.* 46 (1) 82007), 47–64.
- [24] Y. Wang, G. Wei y W. H. Campbell, *Compact and locally compact hyperspace dynamical systems induced from locally compact dynamical systems*, por aparecer en *Chaos Solitons & Fractals*.
- [25] G. Zhang, F. Zeng y X. Liu, *Devaney's chaotic on induced maps of hyperspaces*, *Chaos Solitons & Fractals*, 27 (2) (2006), 471–475.

# Existencia de funciones de Whitney

Genaro Hernández Mada y Carlos Alberto Robles Corbalá

Departamento de Matemáticas

Universidad de Sonora

genaro.hernandezm@correoa.uson.mx, crobles@gauss.mat.uson.mx

## Resumen

*Las funciones de Whitney juegan un papel importante en el estudio de los Hiperespacios de continuos. Intuitivamente, una función de Whitney es una manera de medir el tamaño de subconjuntos cerrados de un continuo. En este trabajo se enuncian algunas de sus propiedades y algunas de las aplicaciones que tienen dentro de la teoría de continuos e hiperespacios.*

## 1. Introducción

El objetivo del presente trabajo es enunciar algunas de las propiedades de las funciones de Whitney y algunas de las aplicaciones que tienen. Esto se hace dentro de la teoría de continuos e hiperespacios.

En la primera parte, se presentan las definiciones básicas para el estudio de funciones de Whitney, algunos ejemplos, y algunas de sus propiedades. Se presenta también la demostración de que uno de los ejemplos presentados es efectivamente una función de Whitney. En la segunda parte, se presentan algunas de las aplicaciones que tienen las funciones de Whitney, dentro de la teoría de continuos e hiperespacios.

## 2. Definiciones básicas

Las siguientes definiciones forman un punto de partida para nuestro estudio.

**Definición 1.** Un continuo es un espacio métrico compacto, conexo y no vacío.

Para evitar trivialidades, en ocasiones, también se pide que el continuo tenga más de un punto. A estos se les llama continuos no degenerados.

A menos que se especifique lo contrario, dado un continuo  $X$ , le llamaremos  $d$  a la métrica de  $X$ .

**Definición 2.** Si  $X$  es un continuo y  $Y \subset X$  es a su vez un continuo, diremos que  $Y$  es un *subcontinuo* de  $X$ .

Dado un continuo  $X$ , los *hiperespacios* son ciertas familias de subconjuntos de  $X$ , con alguna característica particular. Entre los más estudiados se encuentran:

$$2^X = \{A \subset X : A \text{ es cerrado y no vacío}\},$$

$$C(X) = \{A \in 2^X : A \text{ es conexo}\},$$

$$F_n(X) = \{A \in 2^X : A \text{ tiene a lo más } n \text{ puntos}\} \quad (n \in \mathbb{N}),$$

$$C_n(X) = \{A \in 2^X : A \text{ tiene a lo más } n \text{ componentes}\} \quad (n \in \mathbb{N}).$$

Para estudiar a estos hiperespacios como espacios métricos, basta con dotar de una métrica a  $2^X$ , pues los demás están contenidos en él. Para lograr esto, necesitaremos las siguientes definiciones:

**Definición 3.** Sea  $X$  un continuo. Dados  $\epsilon > 0$  y  $p \in X$ , se le llama la bola de radio  $\epsilon$  centrada en  $p$  al conjunto

$$B_\epsilon(p) = \{q \in X : d(p, q) < \epsilon\}.$$

**Definición 4.** Sea  $X$  un espacio métrico. Dados  $\epsilon > 0$  y  $A \in 2^X$ , se le llama la nube de radio  $\epsilon$  centrada en  $A$  al conjunto

$$N(\epsilon, A) = \{q \in X : \text{existe } x \in A \text{ tal que } d(x, q) < \epsilon\}.$$

Obsérvese que la nube de radio  $\epsilon$  centrada en  $A$  puede representarse como

$$N(\epsilon, A) = \bigcup_{a \in A} B_\epsilon(a).$$

Con las definiciones 3 y 4 podemos definir la métrica en  $2^X$ . Dados  $A, B \in 2^X$ , se define

$$H(A, B) = \inf\{\epsilon > 0 : A \subset N(\epsilon, B) \text{ y } B \subset N(\epsilon, A)\}.$$

Se puede verificar que  $H$  efectivamente es una métrica en  $2^X$ , y se le llama *métrica de Hausdorff*. Para ver una demostración de que  $H$  es una métrica puede verse, por ejemplo, la referencia [3] de la página 22 a la 24.

El siguiente resultado es útil porque nos brinda una manera de pasar de la métrica en  $2^X$  a la métrica en  $X$  y viceversa.

**Lema 1.** Sea  $X$  un continuo y  $A, B \in 2^X$ . Entonces,  $H(A, B) < r$  si y sólo si  $A \subset N(r, B)$  y  $B \subset N(r, A)$ .

Una demostración de este útil resultado puede encontrarse en la referencia [1] en las páginas 3 y 4.

### 3. Funciones de Whitney

Habiendo definido una métrica para  $2^X$ , nos gustaría tener una manera de medir el tamaño de los elementos de  $2^X$ , mediante una función. A estas funciones les llamaremos *funciones de Whitney*:

**Definición 5.** Una función de Whitney es una función continua  $\mu : 2^X \rightarrow [0, \infty)$  que satisface las siguientes condiciones:

1.  $\mu(\{p\}) = 0$  para cada  $p \in X$ ,
2. Si  $A \subsetneq B$ , entonces  $\mu(A) < \mu(B)$ .

Cabe mencionar que, como  $X$  es compacto, y  $\mu$  es continua,  $\mu$  alcanza su valor máximo. Si definimos la función  $\mu'$ , de manera que para cada  $A \in 2^X$

$$\mu'(A) = \frac{\mu(A)}{\mu(X)},$$

Es fácil notar que  $\mu'$  también es una función de Whitney y toma valores en el intervalo  $[0, 1]$ . Es decir, podríamos definir una función de Whitney como una función que toma valores en  $[0, 1]$  y que cumple las propiedades de la definición 5.

Se pueden encontrar ejemplos de funciones de Whitney para continuos en específico. Sin embargo, para probar resultados se necesita asegurar la existencia de funciones de Whitney para continuos arbitrarios, así que es útil determinar si para cualquier continuo  $X$  puede construirse una función de Whitney. La respuesta es afirmativa, y básicamente se conocen tres construcciones.

**Ejemplo 1.** Sea  $X$  un continuo. Si en lugar de considerar la métrica  $d$  de  $X$ , consideramos la métrica  $\tilde{d}$  definida, para cada par de puntos  $x, y \in X$  por

$$\tilde{d}(x, y) = \frac{d(x, y)}{\text{diam}(X)},$$

tenemos que  $\tilde{d}(x, y) \leq 1$  para todos los puntos  $x, y \in X$ , y se puede demostrar que la topología inducida por  $d$  y por  $\tilde{d}$  son la misma. Por lo tanto, podemos suponer que la métrica  $d$  en  $X$  satisface  $d(x, y) \leq 1$  para cada par de puntos  $(x, y)$ .

Como  $X$  es compacto, tiene un subconjunto denso y numerable, digamos  $D = \{p_1, p_2, \dots\}$ . Para cada  $n \in \mathbb{N}$ , definimos  $\mu_n : 2^X \rightarrow [0, 1]$  por:

$$\mu_n(A) = \text{máx}\{d(a, p_n) : a \in A\} - \text{mín}\{d(a, p_n) : a \in A\}.$$

Geoméricamente,  $\mu_n(A)$  puede interpretarse como el ancho del anillo mínimo centrado en  $p_n$  que contiene a  $A$ .

Ahora definimos  $\mu : 2^X \rightarrow [0, 1]$  por:

$$\mu(A) = \sum_{n=1}^{\infty} \frac{\mu_n(A)}{2^n}.$$

La función  $\mu$  resulta estar bien definida, y es una función de Whitney.

**Ejemplo 2.** Para  $n \in \mathbb{N}$ , definimos  $\mu_n : 2^X \rightarrow [0, \infty)$  dada por:

$$\mu_n(A) = \inf\{\epsilon > 0 : \text{existen puntos } p_1, p_2, \dots, p_n \in X \text{ tales que } A \subset N(\epsilon, \{p_1, \dots, p_n\})\}.$$

La función  $\mu : 2^X \rightarrow [0, \infty)$  dada por

$$\mu(A) = \sum_{n=1}^{\infty} \frac{\mu_n(A)}{2^n}$$

también es una función de Whitney.

**Ejemplo 3.** Al igual que en el Ejemplo 1, supongamos que la métrica  $d$  en  $X$  satisface que  $d(x, y) \leq 1$  para todo par de puntos  $x, y \in X$ . Dados  $A \in 2^X$  y  $n \in \mathbb{N}$ , definimos

$$L_n(A) = \{\epsilon \geq 0 : \text{existen puntos (no necesariamente distintos) } a_1, \dots, a_{n+1} \in A \\ \text{tales que si } i \neq j, \text{ entonces } B_\epsilon(a_i) \cap B_\epsilon(a_j) = \emptyset\}.$$

Cabe mencionar que definimos  $B(0, p) := \emptyset$ . Luego, se definen

$$\mu_n(A) = \sup L_n(A),$$

y

$$\mu(A) = \sum_{n=1}^{\infty} \frac{\mu_n(A)}{2^n}.$$

Esta última función  $\mu$  también resulta ser una función de Whitney.

Las demostraciones de que las funciones  $\mu$  en los ejemplos son funciones de Whitney son largas, por lo que aquí solamente se presenta la demostración para el Ejemplo 3.

**Teorema 2.** La función  $\mu$  en el Ejemplo 3 es una función de Whitney.

*Demostración.* Analizaremos algunas propiedades de las funciones  $\mu_n$  que se usan para definir  $\mu$ . Estas propiedades implican que  $\mu$  es una función de Whitney, y se encuentran a continuación:

1. Para cada  $n \in \mathbb{N}$ , la función  $\mu_n$  está bien definida y  $\mu_n(A) \leq 1$  para todo  $A \in 2^X$ .

En efecto, notemos que  $0$  siempre pertenece a  $L_n(A)$ , por lo que  $L_n(A) \neq \emptyset$ . Luego, si  $\epsilon > 0$  y  $\epsilon \in L_n(A)$ , entonces existen puntos  $a_1, \dots, a_{n+1} \in A$  tales que las bolas de la forma  $B_\epsilon(a_i)$  son ajenas entre sí. Por lo tanto,  $\epsilon \leq d(a_1, a_{n+1}) \leq \text{diam}(X) \leq 1$ . Por lo tanto,  $L_n(A)$  está acotado superiormente por 1, y por lo tanto su supremo existe y es menor o igual que 1.

2. Para cada  $n \in \mathbb{N}$ , la función  $\mu_n$  es continua.

Sea  $\epsilon > 0$  y  $\delta = \epsilon$ . Supongamos que  $H(A, B) < \delta$ . Demostraremos que  $|\mu_n(A) - \mu_n(B)| \leq \delta$ .

Sea  $r \in L_n(A)$ . Entonces, existen elementos  $a_1, \dots, a_{n+1} \in A$  tales que las bolas de la forma  $B_r(a_i)$  son ajenas entre sí. Como  $H(A, B) < \delta$ , por el lema 1, se sigue que  $A \subset N(\delta, B)$ , por

lo que existen elementos  $b_1, \dots, b_{n+1} \in B$  tales que  $d(a_i, b_i) < \delta$  para cada  $i = 1, \dots, n + 1$ . Esto último es consecuencia de la definición de nube de radio  $\delta$  centrada en  $B$ .

Tomemos  $\epsilon_0 = \max\{0, r - \delta\}$ . Dada  $1 \leq i \leq n + 1$ , veamos que  $B_{\epsilon_0}(b_i) \subset B_r(a_i)$ . Si  $\epsilon_0 = 0$ , entonces  $B_{\epsilon_0}(b_i) = \emptyset$  y es claro que  $B_{\epsilon_0}(b_i) \subset B_r(a_i)$ . En otro caso,  $\epsilon_0 > 0$ , y podemos tomar  $p \in B_{\epsilon_0}(b_i)$ . Luego,

$$d(p, a_i) \leq d(p, b_i) + d(b_i, a_i) < \epsilon_0 + \delta = r,$$

de donde  $p \in B_r(a_i)$ , y se sigue que  $B_{\epsilon_0}(b_i) \subset B_r(a_i)$ .

Como las bolas de la forma  $B_r(a_i)$  son ajenas entre sí, entonces las bolas de la forma  $B_{\epsilon_0}(b_i)$  también lo son. Esto implica que  $\epsilon_0 \in L_n(B)$ . Por lo tanto,  $\mu_n(B) = \sup L_n(B) \geq \epsilon_0 \geq r - \delta$ , es decir, que  $\mu_n(B) + \delta$  es una cota superior de  $L_n(A)$ , pues  $r$  era arbitrario. Como es una cota superior, se sigue que  $\mu_n(B) + \delta \geq \sup L_n(A) = \mu_n(A)$ , o equivalentemente,  $\mu_n(A) - \mu_n(B) \leq \delta$ .

De manera similar se demuestra que  $\mu_n(B) - \mu_n(A) \leq \delta$ . Por lo tanto,  $|\mu_n(A) - \mu_n(B)| \leq \delta$ , que es lo que queríamos demostrar.

3. Si  $A$  tiene al menos  $n + 1$  puntos distintos, entonces  $\mu_n(A) > 0$ .

Esto es claro, ya que hay números positivos en  $L_n(A)$ . Como hay  $n + 1$  puntos distintos, tomamos  $r$  como la mitad de la mínima distancia entre dos de esos  $n + 1$  puntos, y es claro que  $r \in L_n(A)$ .

4. Si  $A$  tiene exactamente  $n$  puntos, entonces para  $k \geq n$ , se tiene  $\mu_k(A) = 0$ .

Esto es cierto pues si  $A$  tiene exactamente  $n$  puntos, al tomar  $a_1, \dots, a_{n+1} \in A$ , por el principio de las casillas, existen  $i, j \in \{1, \dots, n + 1\}$  tales que  $a_i = a_j$ . Es decir, uno de ellos debe repetirse. Luego, la única manera de que  $B_\epsilon(a_i) \cap B_\epsilon(a_j) = \emptyset$  es si  $\epsilon = 0$ . Por lo tanto,  $L_n(A) = \{0\}$ . El argumento para cualquier  $k \geq n$  es el mismo.

5. Para cada  $p \in X$  y cada  $n \in \mathbb{N}$ ,  $\mu_n(\{p\}) = 0$ .

Esta es una consecuencia inmediata de la propiedad anterior.

6. Para cada  $A \in 2^X$  y cada  $n \in \mathbb{N}$ ,  $\mu_n(A) \geq \mu_{n+1}(A)$ .

Sea  $\epsilon \in L_{n+1}(A)$ . Entonces, se pueden encontrar  $n + 2$  bolas ajenas por parejas, centradas en puntos de  $A$ . Como se pueden encontrar  $n + 2$  de estas bolas, también pueden encontrarse  $n + 1$ , por lo que  $\epsilon \in L_n(A)$ . Por lo tanto,  $L_{n+1}(A) \subset L_n(A)$  y  $\sup L_{n+1}(A) \leq \sup L_n(A)$ , que es lo que queríamos demostrar.

7. Para cada  $A \in 2^X$ , se cumple  $\lim_{n \rightarrow \infty} \mu_n(A) = 0$ .

Esto se demuestra usando la definición de límite de una sucesión. Sea  $\epsilon > 0$ . Como  $A$  es compacto, existen  $m \in \mathbb{N}$  y puntos  $a_1, \dots, a_m \in A$  tales que

$$A \subset B_{\frac{\epsilon}{2}}(a_1) \cup \dots \cup B_{\frac{\epsilon}{2}}(a_m).$$

Sean  $p_1, \dots, p_{m+1} \in A \subset B_{\frac{\epsilon}{2}}(a_1) \cup \dots \cup B_{\frac{\epsilon}{2}}(a_m)$ . Por el principio de las casillas, al menos dos de los puntos  $p_i$  están en la misma bola. Supongamos, sin pérdida de generalidad, que  $p_1, p_2 \in$

$B_{\frac{\epsilon}{2}}(a_1)$ . Notemos que  $d(p_1, p_2) \leq d(p_1, a_1) + d(p_2, a_1) < \epsilon$ , por lo que  $p_1 \in B_{\frac{\epsilon}{2}}(p_2) \cap B_{\frac{\epsilon}{2}}(p_1)$ . Esto último implica que ni  $\epsilon$  ni ningún número mayor pueden estar en  $L_m(A)$ , por lo que  $\mu_m(A) \leq \epsilon$ .

Por último, la Propiedad 6 implica que  $\mu_n(A) \leq \epsilon$  para todo  $n \geq m$ , que es lo que queríamos demostrar.

8. Si  $A, B \in 2^X$  y  $A \subset B$ , entonces  $\mu_n(A) \leq \mu_n(B)$  para toda  $n \in \mathbb{N}$ .

Esta es una consecuencia inmediata de la definición de  $\mu_n$ .

Para demostrar que  $\mu$  es una función de Whitney, solamente falta probar que si  $A, B \in 2^X$  y  $A \subsetneq B$ , entonces  $\mu(A) < \mu(B)$ . De acuerdo a la última propiedad de las funciones  $\mu_n$ , basta probar que existe  $m \in \mathbb{N}$  tal que  $\mu_m(A) < \mu_m(B)$ . Para demostrar esto último, consideraremos dos casos.

Si  $A$  es finito, con  $m$  elementos, entonces  $B$  tendrá al menos  $m + 1$  puntos. Entonces, por las Propiedades 3 y 4,  $\mu_m(B) > 0 = \mu_m(A)$ , que es lo que queríamos demostrar.

Si  $A$  es infinito, por la Propiedad 3, tenemos que  $\mu_n(A) > 0$  para toda  $n \in \mathbb{N}$ . Sea  $b \in B - A$  y  $\delta > 0$  tal que  $B_{2\delta}(b) \cap A = \emptyset$ . Por la Propiedad 7, sabemos que existe un número  $m \in \mathbb{N}$  tal que  $0 < \mu_m(A) < \delta$ , y una infinidad de números  $n$  para los cuales  $\mu_n(A) < \mu_m(A)$ .

Notemos que como la sucesión  $(\mu_n(A))_{n=1}^{\infty}$  converge a 0, y todos los elementos son positivos, la sucesión no puede ser constante indefinidamente. Por la observación anterior, podemos suponer que  $\mu_{m+1}(A) < \mu_m(A)$ .

Sea  $\epsilon \in L_m(A)$ . Entonces, existen  $a_1, \dots, a_{m+1} \in A$  tales que las bolas  $B_{\epsilon}(a_1), \dots, B_{\epsilon}(a_{m+1})$  son ajenas por parejas. Como  $\epsilon \leq \mu_m(A) < \delta$ , entonces  $B_{\epsilon}(b) \cap N(\epsilon, A) = \emptyset$ . De aquí se sigue que las bolas  $B_{\epsilon}(b), B_{\epsilon}(a_1), \dots, B_{\epsilon}(a_{m+1})$  son ajenas por parejas, de donde  $\epsilon \in L_{m+1}(B)$ . Por lo tanto,  $L_m(A) \subset L_{m+1}(B)$ . Luego,  $\mu_m(A) \leq \mu_{m+1}(B)$ . Por lo tanto,  $\mu_{m+1}(A) < \mu_{m+1}(B)$ , que es lo que faltaba demostrar.  $\square$

Hemos demostrado que en cualquier continuo  $X$  se pueden definir funciones de Whitney. Los tres ejemplos presentados son los que se presentan comúnmente en la literatura, y a partir de ellos se pueden generar más ejemplos, gracias a los siguientes teoremas que enunciamos sin demostración.

**Teorema 3.** Sea  $X$  un continuo. Si  $\mu_1$  y  $\mu_2$  son funciones de Whitney para  $2^X$ , y  $a_1, a_2 > 0$ , entonces la función  $\mu : 2^X \rightarrow [0, \infty)$  definida por

$$\mu(A) = a_1\mu_1(A) + a_2\mu_2(A)$$

es una función de Whitney.

**Teorema 4.** Si  $\mu$  es una función de Whitney para  $2^X$  y  $A \in 2^X$ , entonces la función  $\mu' : 2^X \rightarrow [0, \infty)$  definida por

$$\mu'(A) = (\mu(A \cup B)\mu(B))^{\frac{1}{2}}$$

es una función de Whitney.

## 4. Aplicaciones de las funciones de Whitney

Un hecho conocido al estudiar los hiperespacios de continuos es que  $2^X$  y  $C(X)$  son compactos. También resultan ser conexos por arcos, y para demostrar esto se utiliza el siguiente concepto:

**Definición 6.** Dados  $A, B \in C(X)$  con  $A \subsetneq B$ , diremos que una función continua  $\alpha : [0, 1] \rightarrow C(X)$  es un *arco ordenado de A a B* en  $C(X)$  si  $\alpha(0) = A$ ,  $\alpha(1) = B$  y  $\alpha(s) \subsetneq \alpha(t)$  cuando  $0 \leq s < t \leq 1$ .

La idea de un arco ordenado es inflar continuamente un continuo  $A$  hasta llegar a un continuo  $B$ .

**Teorema 5.** Sea  $X$  un continuo. Dados  $A, B \in C(X)$  con  $A \subsetneq B$ , existe un arco ordenado de  $A$  a  $B$  en  $C(X)$ .

Este teorema se demuestra a partir de la existencia de una función de Whitney definida en  $C(X)$ . La demostración es larga y la omitiremos en este trabajo. La demostración puede encontrarse en la referencia [3], en las páginas 90 a 92. La existencia de arcos ordenados tiene dos implicaciones importantes.

**Corolario 6.** El hiperespacio  $C(X)$  es conexo por arcos.

*Demostración.* Sea  $A \in C(X)$  tal que  $A \neq X$ . Por el Teorema 4, existe un arco que conecta a  $A$  con  $X$ . Como  $A$  es arbitrario, todos los elementos se pueden conectar por arcos con  $X$ , lo que implica que  $C(X)$  es conexo por arcos.  $\square$

**Corolario 7.** El hiperespacio  $2^X$  es conexo por arcos.

La demostración de este segundo corolario es también una consecuencia de la existencia de arcos ordenados en  $C(X)$ , pero también será omitida en este trabajo.

Una vez conocido que los hiperespacios  $C(X)$  y  $2^X$  son conexos por arcos, podemos concluir que son conexos y por lo tanto son continuos.

La última aplicación de las funciones de Whitney, que mencionaremos aquí, es sobre la contractibilidad en hiperespacios. Para esto, se necesitan las siguientes tres definiciones.

**Definición 7.** Un espacio topológico  $X$  es *contraíble* si existe una función continua  $K : X \times [0, 1] \rightarrow X$  y existe un punto  $x_0 \in X$  tal que para todo  $x \in X$ , se satisface que  $K(x, 0) = x$  y  $K(x, 1) = x_0$ . A una función  $K$  con estas características se le llama *contracción*.

**Definición 8.** Sea  $X$  un continuo. Diremos que  $X$  es *indescomponible* si no se puede escribir como la unión de dos subcontinuos propios. Si además, cualquier subcontinuo de  $X$  es indescomponible, diremos que  $X$  es *hereditariamente indescomponible*.

**Definición 9.** Un continuo  $X$  se dice tener la *propiedad de Kelley en un punto*  $p \in X$  si para toda sucesión  $(p_n)_{n=1}^{\infty} \subset X$  tal que  $p_n \rightarrow p$  y todo subcontinuo  $A$  de  $X$  tal que  $p \in A$ , existe una sucesión de subcontinuos  $(A_n)_{n=1}^{\infty}$  de  $X$  tal que  $A_n \rightarrow A$  y  $p_n \in A_n$  para cada  $n \in \mathbb{N}$ .

Nótese, en la última definición, que la convergencia de la sucesión  $(p_n)_{n=1}^{\infty}$  es con la métrica de  $X$ , y la convergencia de la sucesión  $(A_n)_{n=1}^{\infty}$  es con la métrica de Hausdorff en  $2^X$ . Los siguientes teoremas relacionan estas definiciones.

**Teorema 8.** Si un continuo  $X$  tiene la propiedad de Kelley e todos sus puntos, entonces  $C(X)$  y  $2^X$  son contraíbles.

**Teorema 9.** Si  $X$  es un continuo hereditariamente indescomponible, entonces  $X$  tiene la propiedad de Kelley, y por lo tanto,  $C(X)$  y  $2^X$  son contraíbles.

La demostración de este último teorema utiliza la existencia de una función de Whitney.

## 5. Conclusiones

En nuestro desarrollo, las funciones de Whitney comenzaron a estudiarse como una manera de medir el tamaño de subconjuntos cerrados de un continuo. Después de presentar algunas de las aplicaciones que tiene la existencia de funciones de Whitney, en la teoría de continuos hiperespacios, puede notarse que teoremas de gran utilidad son consecuencia de esto. Esta es sólo una de las razones por las cuales las funciones de Whitney son objeto de estudio.

## 6. Referencias

- [1] Aguilera Miranda, María Ilena y García Velásquez, Luis Miguel (2003). *Hiperespacios de Continuos Métricos y No Métricos*. México.
- [2] Escobedo, R., Macías, S., Méndez, H. (2006). *Invitación a la Teoría de Continuos Hiperespacios*. México. Sociedad Matemática Mexicana.
- [3] Illanes Mejía, Alejandro (2004). *Hiperespacios de continuos*. México. Sociedad Matemática Mexicana.
- [4] Munkres, James R. (2000). *Topología*. España. Prentice-Hall.
- [5] Nadler, Sean B. Jr. (2006). *Hyperspaces of sets. A text with Research Questions*. México. Sociedad Matemática Mexicana.

# Dimensión en niveles de Whitney de una gráfica finita

Roberto Torres Hernández  
Licenciatura en Matemáticas Aplicadas  
Universidad Autónoma de Querétaro

## 1. Introducción

Un *continuo*  $X$  es un espacio métrico compacto, conexo y con más de un punto. Un *hiperespacio* es un espacio constituido por una clase específica de subconjuntos de un espacio dado. Los hiperespacios más comunes de un continuo  $X$  son:

$$2^X = \{A \subset X : A \text{ es cerrado y } A \neq \emptyset\}$$

$$C(X) = \{A \in 2^X : A \text{ es conexo}\}$$

Para cada  $n \in \mathbb{N}$ , definimos:

$$F_n(X) = \{A \in 2^X : A \text{ tiene a lo más } n \text{ puntos}\}$$

En el hiperespacio  $2^X$ , se puede definir una métrica como sigue:

Si  $A \in 2^X$  y  $\varepsilon > 0$ , la *nube de radio*  $\varepsilon$  y *centro en*  $A$  es el conjunto

$$N(\varepsilon, A) = \{x \in X : \text{Existe } a \in A \text{ tal que } d(x, a) < \varepsilon\}$$

donde  $d$  denota la métrica del continuo  $X$ .

Observemos que este conjunto es la unión de todas las bolas abiertas de radio  $\varepsilon$  con centro en algún elemento de  $A$ , es decir, si  $B_\varepsilon(a) = \{x \in X : d(x, a) < \varepsilon\}$  entonces

$$N(\varepsilon, A) = \bigcup_{a \in A} B_\varepsilon(a)$$

Si  $A, B \in 2^X$ , la *métrica de Hausdorff*  $H$  para  $2^X$  se define como

$$H(A, B) = \inf \{\varepsilon > 0 : A \subset N(\varepsilon, B) \text{ y } B \subset N(\varepsilon, A)\}$$

No es difícil probar que efectivamente,  $H$  es una métrica, por lo que  $2^X$  se convierte en un espacio métrico. Más aún, los primeros resultados que se demuestran respecto a los hiperespacios son que  $2^X$  es compacto y conexo (de hecho, conexo por trayectorias), de donde  $2^X$  es un continuo.

Puede probarse también que el hiperespacio  $C(X)$  es cerrado en  $2^X$  y por lo tanto, es un continuo. Es en este espacio en el que se trabajará en el presente artículo. Para iniciar, mostraremos los dos ejemplos más simples.

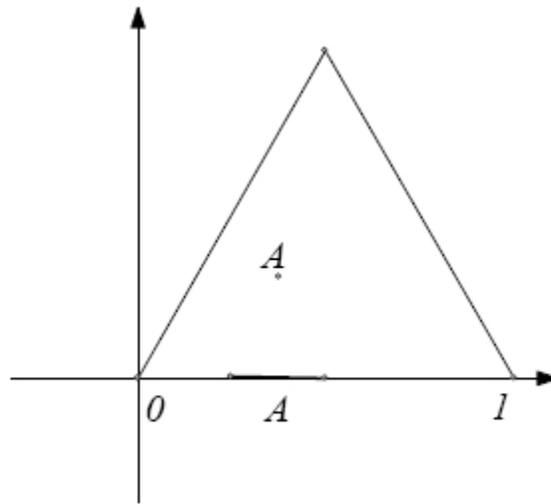
Sea  $X$  un arco, digamos  $X = [0, 1]$ . Un subconjunto cerrado, conexo y no vacío de  $X$  es un subintervalo, de donde

$$C(X) = \{[a, b] : 0 \leq a \leq b \leq 1\}$$

Considerese la función  $F : C(X) \rightarrow \mathbb{R}^2$  dada por

$$F([a, b]) = \left( \frac{a+b}{2}, b-a \right)$$

Geoméricamente estamos pensando a  $X$  como el intervalo  $[0, 1]$  del eje  $x$  del plano, y a cada elemento de  $C(X)$  se le asocia el punto compuesto por su punto medio y su longitud. La imagen de esta función es el triángulo en el plano  $\mathbb{R}^2$  con vértices en los puntos  $(0, 0)$ ,  $(1, 0)$  y  $(\frac{1}{2}, 1)$ :



**Figura 1.**

De esto se puede deducir que si  $X$  es un arco,  $C(X)$  es homeomorfo al disco.

Sea  $X$  una curva cerrada simple. Sin perder generalidad podemos pensar que  $X = S^1$ , es decir

$$X = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

Observemos que cada  $A \in C(X)$  con  $A \neq X$  es un subarco de  $X$  que queda totalmente determinado por su punto medio  $m(A)$  y su longitud  $\delta(A)$ .

Si  $D$  es el disco acotado por  $X$  descrito en coordenadas polares y  $\theta(m(A))$  es el ángulo formado por el eje polar y el segmento de recta que une al polo con el punto medio  $m(A)$ , podemos definir la función  $F : C(X) \rightarrow D$  dada por

$$F(A) = \left( \theta(m(A)), 1 - \frac{\delta(A)}{2\pi} \right) \text{ si } A \neq X \text{ y } F(X) = 0.$$

De esta manera, se tiene que si  $X$  es una curva cerrada simple,  $C(X)$  es homeomorfo al disco.

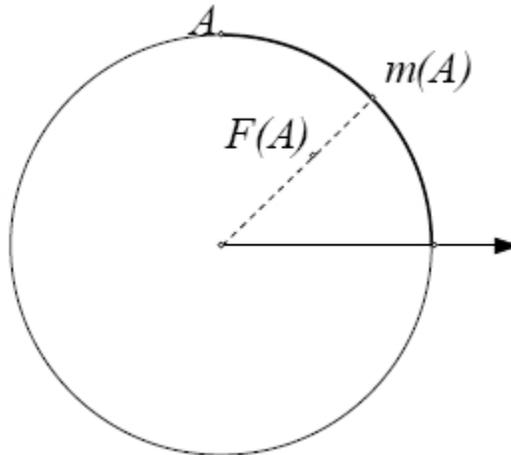


Figura 2.

## 2. Las funciones y los niveles de Whitney

Una herramienta muy importante para el estudio del hiperespacio  $C(X)$  es el concepto de función de Whitney, que se define a continuación.

Una *función de Whitney* es una función continua  $\mu : C(X) \rightarrow \mathbb{R}$  tal que:

1.  $\mu(\{x\}) = 0$
2. Si  $A \subseteq B \neq A$ , entonces  $\mu(A) < \mu(B)$ .

Hablando intuitivamente, podemos pensar que una función de Whitney mide el “tamaño” de un elemento de  $C(X)$ .

Aunque se conocen pocos ejemplos explícitos de funciones de Whitney para espacios arbitrarios, se sabe que el conjunto de estas funciones es homeomorfo a  $l_2$ .

Dada una función de Whitney  $\mu$ , un nivel de Whitney es un conjunto de la forma

$$\mu^{-1}(t) = \{A \in C(X) : \mu(A) = t\} \text{ con } 0 < t < \mu(X).$$

Observemos que los niveles de Whitney son los conjuntos que “agrupan” a los elementos de  $C(X)$  con la misma medida.

En México, el estudio del hiperespacio  $C(X)$  ha usado como herramienta frecuente y muy fuerte, las ideas de función y nivel de Whitney, algunos de estos trabajos inspirados por la fructífera labor del Dr. Alejandro Illanes Mejía del Instituto de Matemáticas de la UNAM.

A continuación se describirá de manera muy intuitiva, algunos resultados sobre la dimensión de los hiperespacios de las gráficas finitas. Anticipamos también que por comodidad no se definirán todos los conceptos que se utilizan, pero aclaramos que todos ellos tendrán el significado matemático usual.

### 3. Dimensión y gráficas finitas

Intuitivamente podemos pensar que las gráficas finitas conexas son unión de segmentos (espacios homeomorfos al intervalo  $[0, 1]$ ) por puntos a los que llamaremos vértices. El orden de un vértice es el número de segmentos que inciden en él. Un vértice  $v$  es extremo si su orden es 1 ( $ord(v) = 1$ ), de otra manera se le llama interno ( $ord(v) \geq 3$ , ya que puede despreciarse el caso  $ord(v) = 2$  pues se tendría un segmento). También pensemos que se tiene definido una función de Whitney en el espacio que estamos tratando.

Entenderemos por dimensión de un espacio en un punto como la cantidad de “maneras” o “libertades” que tiene el punto de “moverse” dentro del espacio. En particular, confrontaremos la dimensión de un punto en el hiperespacio  $C(X)$  con la dimensión de un nivel de Whitney en el mismo punto. Para ilustrar estas ideas consideremos el siguiente ejemplo:

Sea  $X = [0, 1]$  y sea  $A \in C(X)$  el subintervalo  $A = [\frac{1}{4}, \frac{1}{2}]$ . Una posible función de Whitney para este caso puede ser  $\mu : C(X) \rightarrow \mathbb{R}$  dada por  $\mu([a, b]) = b - a$ , es decir, a cada subintervalo se le asocia su longitud.

Ahora,  $\mu^{-1}(\frac{1}{4})$  es el conjunto de subintervalos de longitud  $\frac{1}{4}$  (nuestro  $A$  entre ellos) y si echamos mano de la representación geométrica que tenemos para  $C(X)$  dada en la primera sección, vemos que este nivel de Whitney es el segmento horizontal dentro del triángulo que está a altura  $\frac{1}{4}$ .

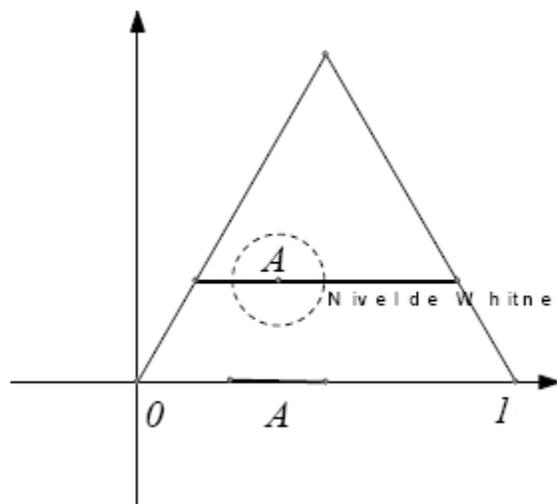


Figura 3.

Si pensamos a  $A$  en  $C(X)$ , se puede decir que la dimensión de  $C(X)$  en  $A$  es dos, pues existe una

vecindad de  $A$  en  $C(X)$  homeomorfa al plano (el disco punteado en la figura anterior). Por otro lado, la dimensión de  $\mu^{-1}(\frac{1}{4})$  en  $A$  es uno, ya que la vecindad de  $A$  en su nivel de Whitney es un subintervalo abierto y por tanto es homeomorfo a la recta real.

Estos mismos hechos pueden observarse sin tener la representación geométrica de  $C(X)$  como sigue:

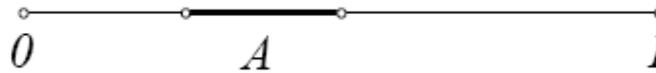


Figura 4.

Como  $A \subset X$ , podemos pensar que  $A$  tiene dos maneras de “moverse” en  $X$ , una en dirección a cero y otra en dirección a uno, pero solo una forma en  $\mu^{-1}(\frac{1}{4})$  ya que si a partir de  $A$  se “creciera” (o “decreciera”) en ambas direcciones se llegaría a subcontinuos que contienen (o están contenidos) propiamente en  $A$  y por la segunda condición de función de Whitney, estos subintervalos no pertenecerían a  $\mu^{-1}(\frac{1}{4})$ .

Otro ejemplo puede ayudar a clarificar estas ideas. Considerese la gráfica  $X$  y el subconjunto  $A \in C(X)$  ilustradas a continuación:

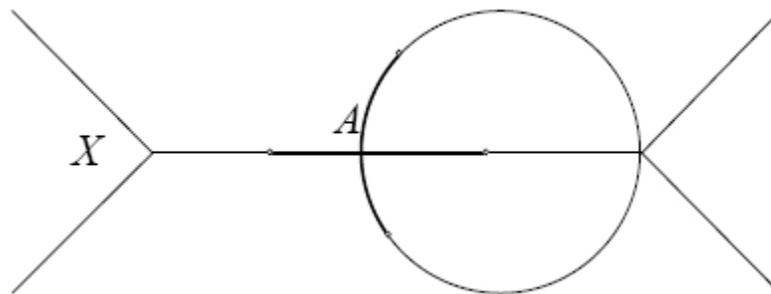
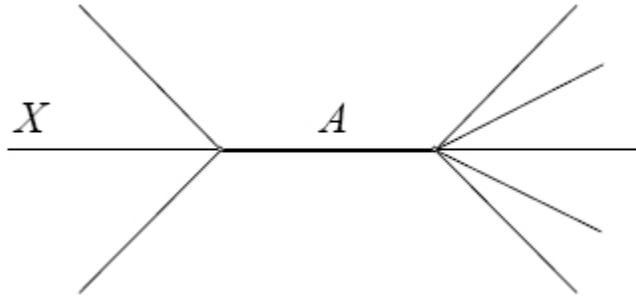


Figura 5.

Observemos que la dimensión de  $C(X)$  en  $A$  es cuatro, pues  $A$  tiene cuatro “libertades” para moverse en  $X$  ( $\dim_{C(X)} A = 4$ ), pero si  $A \in \mu^{-1}(t)$ , la dimensión de  $\mu^{-1}(t)$  en  $A$  es tres ( $\dim_{\mu^{-1}(t)} A = 3$ ) ya que una dirección o “manera” de moverse de  $A$  en  $X$  se pierde para conservar el tamaño de los subcontinuos que pertenecen al nivel de Whitney en el cual también está  $A$ .

A partir de los ejemplos anteriores, se podría pensar que la idea de confrontar las dimensiones de  $C(X)$  y  $\mu^{-1}(t)$  en un subcontinuo  $A$  no es muy interesante, ya que bastará restar uno a la dimensión de  $C(X)$  para obtener la de  $\mu^{-1}(t)$ . Sin embargo, esto no es cierto en general como podemos observar en el siguiente ejemplo:


**Figura 6.**

Aquí tenemos que  $\dim_{C(X)} A = 8$ , pues  $A$  puede “crecer” en  $X$  a lo largo de los ocho segmentos que inciden en sus extremos, pero  $\dim_{\mu^{-1}(t)} A = 5$ , pues si  $A$  crece hacia la derecha, ya no puede crecer a la izquierda, porque se saldría del nivel de Whitney, y si crece a la izquierda, ya no puede hacerlo a la derecha.

En este caso tenemos la desigualdad estricta

$$\dim_{C(X)} A - 1 > \dim_{\mu^{-1}(t)} A$$

Una cuestión interesante aquí, sería caracterizar a los subcontinuos  $A$  de  $X$  tales que cumplan la desigualdad anterior. Para ello, notemos que  $A$  tienen las siguientes características:

1.  $A$  es subgráfica (Subconjunto de una gráfica que es también gráfica)
2.  $A$  es interna (No contiene vértices internos)
3.  $A$  es acíclica (No contiene espacios homeomorfos al círculo)

De hecho, estas condiciones son necesarias y suficientes:

**Teorema 1.**  $A$  es una subgráfica interna y acíclica si y sólo si

$$\dim_{C(X)} A - 1 > \dim_{\mu^{-1}(t)} A.$$

Otra idea interesante parte de la definición de araña. Una gráfica es una *araña* si tiene uno y sólo un vértice interno. Luego, las arañas deben de tener la forma:

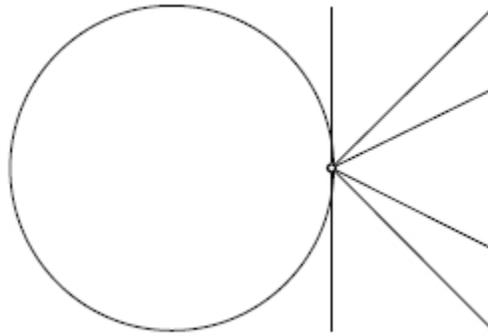


Figura 7.

Ahora, para cada  $A \in C(X)$ , se tienen dos posibilidades:

1.  $A$  contiene al único vértice interno.
2.  $A$  no lo contiene.

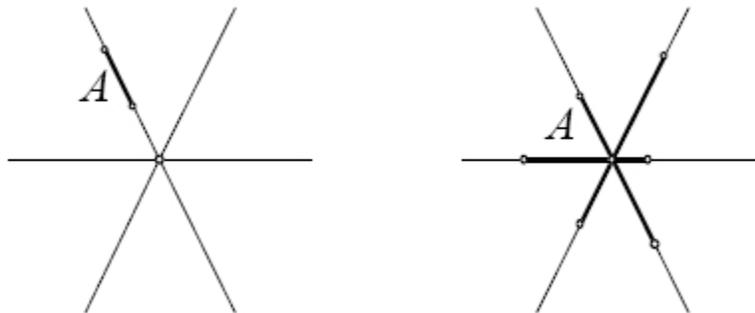


Figura 8.

En el primer caso tenemos que  $A$  debe de estar totalmente contenido en un segmento o en un lazo, de donde  $\dim_{\mu^{-1}(t)} A = 1$ , y esto no contribuye al cálculo de la dimensión.

En el segundo caso, un elemento típico de cualquier nivel de Whitney se ilustra en la figura de la derecha. Aquí es claro que no importa si los “brazos” de  $A$  son cortos o largos, de cualquier modo, las maneras de “crecer” son las mismas para cada subcontinuo  $A$ . Lo anterior se resume en el siguiente:

**Teorema 2.**  $X$  es araña si y sólo si todos sus niveles de Whitney tienen la misma dimensión.

Sea  $\mu$  una función de Whitney. Diremos que  $\mu^{-1}(t)$  es un *nivel grande* si  $t_0 < t < \mu(X)$  donde

$$t_0 = \text{máx} \{ \mu(S) : S \text{ s subgráfica propia no vacía de } X \}$$

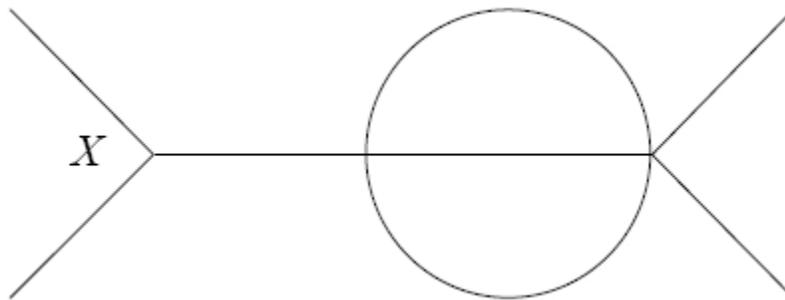
También diremos que  $\mu^{-1}(t)$  es un *nivel chico* si  $t < t_0 < \mu(S)$  donde

$$t_0 = \min \{ \mu(S) : S \text{ es segmento de } X \}$$

Intuitivamente hablando, se puede decir que los niveles grandes agrupan a los elementos de  $C(X)$  que son “casi”  $X$ , y que un nivel chico consta de los subcontinuos de  $X$  que son más “pequeños” que un intervalo.

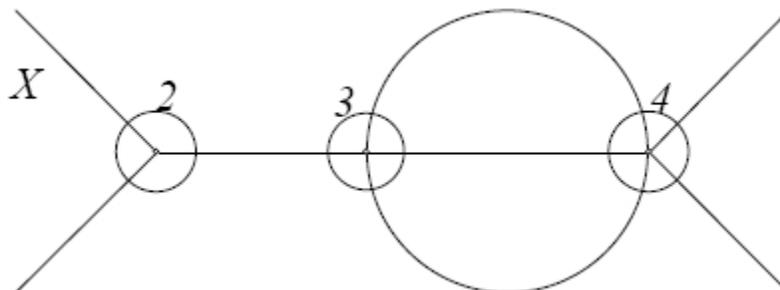
Analizando a los niveles chicos, podemos observar que un elemento de estos niveles, si está totalmente contenido en un intervalo, sólo tiene una manera de “moverse” dentro del nivel de Whitney, por lo que estos elementos pueden representarse por un punto, exactamente como en el intervalo donde está. Si contiene a un vértice interno (no puede contener dos, por que contendría a un segmento completo), entonces las maneras de moverse que tiene son todos los segmentos que inciden en ese vértice menos uno, para cuidar el “tamaño”. Esta idea nos proporciona una manera geométrica de visualizar los niveles chicos.

Sea  $X$  la gráfica siguiente:



**Figura 9.**

Los niveles chicos deben entonces de tener la misma forma de  $X$ , salvo que en cada vértice interno de la gráfica se sustituye este por un cubo de dimensión el orden del vértice menos uno:



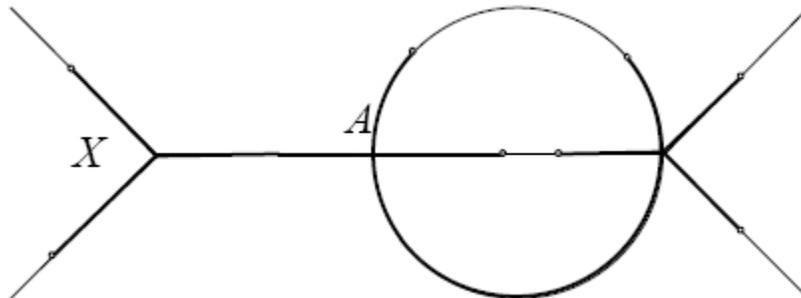
**Figura 10.**

Los números en la gráfica indican la dimensión del cubo representado por los círculos pequeños. Más formalmente:

**Teorema 3.** Si  $\mu^{-1}(t)$  es un nivel chico, entonces

$$\dim \mu^{-1}(t) = \{ord(v) - 1 : v \text{ es vértice interno de } X\}$$

Con respecto a los niveles grandes, podemos decir que si  $A$  pertenece a uno de estos, entonces  $A$  “cubre” casi a todo  $X$ . En el siguiente dibujo, observemos que todo vértice interno es *punto de corte* ( $p \in X$  es de corte si  $X - \{p\}$  no es conexo), luego, como  $A$  es conexo, debe de contener a todo punto de corte y en consecuencia a todo vértice interno, es decir, los pedazos de  $X - A$  deben de estar contenidos en los segmentos de  $X$ .



**Figura 11.**

Luego, la manera que tiene cualquiera de estos  $A$  de “moverse” en su nivel es la misma, de donde la dimensión de  $\mu^{-1}(t)$  es siempre igual. Formalmente:

**Teorema 4.** Sea  $\mu^{-1}(t)$  un nivel grande. Son equivalentes:

1.  $\mu^{-1}(t)$  tiene dimensión homogénea (la misma en cada uno de sus puntos).
2. Todo vértice interno de  $X$  es de corte.

Finalmente, podemos mencionar un par de preguntas abiertas como motivación en esta dirección de investigación:

- ¿Qué condición debe de cumplir la gráfica  $X$  para que sus niveles grandes sean contraíbles?
- Dada una gráfica acíclica ¿cuántos niveles de Whitney diferentes (topológicamente) existen en  $C(X)$ ?

## 4. Referencias

- [1] Illanes, A. y Nadler, S. *Hyperspaces: Fundamentals and Recent Advances*. Marcel Dekker. (1999).
- [2] Kelley, J. L. *Hyperspaces of a continuum*. Trans. Amer. Math. Soc. 52 (1942).

- [3] Duda, R. *On the hyperspace of subcontinua of a finite graph I*. Fundamenta Mathematicae 62 (1968).
- [4] Duda, R. *On the hyperspace of subcontinua of a finite graph II*. Fundamenta Mathematicae 63 (1968).
- [5] Duda, R. Corrections to the paper *On the hyperspace of subcontinua of a finite graph I*. Fundamenta Mathematicae 69 (1970).
- [6] Illanes A. y Puga, I. *Determining finite graphs by their large Whitney levels*. Topology and its applications 60 (1994).
- [7] Illanes A. y Torres, R. *The dimension of Whitney levels of a finite graph*. Glasnik Matematicki Vol. 32 (52) (1997).
- [8] Illanes, A. *The space of Whitney levels*. Topology and its applications. 40 (1991).
- [9] Illanes, A. *The space of Whitney levels is homeomorphic to  $l_2$* . Colloquium Mathematicum. LXV (1993).

# Funciones entre hiperespacios de continuos y relación de orden

Gloria Andablo-Reyes  
Facultad de Ciencias Físico Matemáticas  
Universidad Michoacana de San Nicolás de Hidalgo  
gloria.andablo@gmail.com

## Resumen

*Un continuo  $X$  es un espacio métrico, compacto, conexo y con más de un punto. Un subconjunto no vacío, conexo y cerrado de  $X$  se llama subcontinuo. Trabajaremos con colecciones de subconjuntos de continuos que satisfacen propiedades específicas, las cuales son llamadas hiperespacios de continuos. Sea  $X$  un continuo, nos concentraremos en dos de sus hiperespacios:  $2^X$ , que consta de los subconjuntos no vacíos y cerrados de  $X$  y,  $C(X)$  cuyos elementos son los subcontinuos de  $X$ . Sean  $H$  y  $K$  hiperespacios de  $X$  y  $Y$  respectivamente, estamos interesados en encontrar condiciones necesarias y/o suficientes bajo las cuales exista una función  $f$  de  $H$  en  $K$  tal que si  $A, B \in H$  y  $A \subset B$ , entonces  $f(A) \subset f(B)$ . La función  $f$  se llama encaje ordenado y aquí presentamos ejemplos de ellos.*

## 1. Introducción

La Teoría de Continuos e Hiperespacios son áreas de la topología relativamente jóvenes, sus orígenes se remontan a la década 1910-1920. La Teoría de Continuos se interesa por estudiar a los espacios métricos, compactos, conexos y con más de un punto, esta clase de espacios son conocidos como *continuos*. Llamamos *subcontinuo* a un subconjunto no vacío, conexo y cerrado de un continuo. Observe que un subcontinuo puede constar de un único punto. Por otra parte, la Teoría de Hiperespacios estudia colecciones de subconjuntos de un espacio  $X$  que satisfacen ciertas propiedades específicas, llamamos a estas colecciones *hiperespacios*. De este modo, los elementos de un hiperespacio del espacio  $X$  son subconjuntos de  $X$ . En este trabajo nos enfocaremos en hiperespacios de continuos, por lo que a lo largo de este trabajo todos nuestros espacios serán continuos. En particular trabajaremos con dos hiperespacios: el hiperespacio  $2^X$  que consta de los subconjuntos no vacíos y cerrados de  $X$  y, el hiperespacio  $C(X)$  que está formado por los subcontinuos de  $X$ . Note que el hiperespacio  $C(X)$  está contenido en el hiperespacio  $2^X$ .

Dotamos a nuestros hiperespacios de una métrica, la métrica de Hausdorff, con la que se convierten en espacios métricos (Proposición 2.1 de [4]) y de hecho en continuos, como se muestra en los teoremas 2.4 y 4.2 de [4]. Intuitivamente, dos elementos de un hiperespacio están cerca con la métrica de Hausdorff si están casi empalmados.

Los hiperespacios también poseen una estructura de orden, la cual es inducida de manera natural por la contención. Dicha estructura ha sido muy poco explotada, sólo algunos autores la han

aprovechado (ver [6] y [7]). La existencia de esta estructura de orden da origen de manera natural a la siguiente cuestión. Sean  $H$  y  $K$  hiperespacios de los continuos  $X$  y  $Y$  respectivamente, nos preguntamos cuáles son las condiciones necesarias y/o suficientes, bajo las que podemos definir una función continua e inyectiva  $f$  de  $H$  en  $K$  tal que si  $A, B \in H$  y  $A \subset B$ , entonces  $f(A) \subset f(B)$ .

Como ya mencionamos antes,  $H$  y  $K$  son también continuos, en particular,  $H$  es un espacio Hausdorff y  $K$  es un espacio compacto, entonces una función continua e inyectiva  $f$  de  $H$  en  $K$  es un homeomorfismo sobre su imagen, esto es, un *encaje topológico*. Así que una función  $f$  como la descrita arriba, que además de ser un encaje es compatible con el orden inducido por la contención en  $H$ , será llamada *encaje ordenado*. Si es posible definir un encaje ordenado de  $H$  en  $K$ , diremos que  $H$  puede encajarse ordenadamente en  $K$ . Los encajes ordenados de hiperespacios de continuos fueron introducidos en [1].

En la Sección 2 presentaremos, para continuos específicos, modelos para el hiperespacio  $C(X)$ , los cuales serán útiles cuando, en la Sección 3, discutamos mediante la construcción de ejemplos, el concepto de encaje ordenado para los hiperespacios  $2^X$  y  $C(X)$ .

## 2. Modelos para hiperespacios de continuos

En ocasiones nos resulta muy conveniente trabajar con un espacio en el cual sus elementos sean puntos y no subconjuntos de otro espacio, como sucede en el caso de los hiperespacios y, si además dicho espacio nos resulta familiar es aún mejor. Por lo que cuando es posible y la situación así lo amerita, trabajamos con un espacio homeomorfo al hiperespacio del continuo en cuestión, al cual llamamos *modelo del hiperespacio*. En el Teorema 14.12 de [5] se demuestra que para cualquier continuo  $X$ , el hiperespacio  $2^X$  contiene un cubo de Hilbert, esto es, contiene un espacio homeomorfo al producto  $\prod_{n \in \mathbb{N}} [0,1]$ , debido a ello no tenemos modelos para el hiperespacio  $2^X$ .

Por otro lado, para ciertos continuos  $X$ , es posible construir modelos para el hiperespacio  $C(X)$ . Primero introducimos una lista de continuos que utilizaremos en la siguiente sección y posteriormente presentaremos, sin demostración, modelos para los hiperespacios  $C(X)$  de algunos de estos continuos.

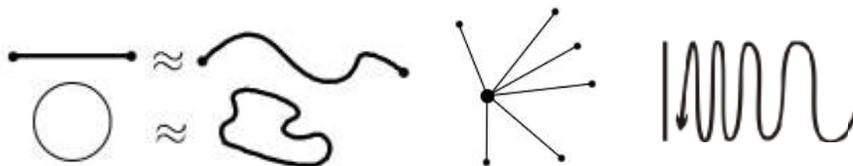


Figura 1. El arco, la curva cerrada simple, el  $n$ -odo y la curva sinoidal del topólogo.

El arco es el continuo más sencillo pues es homeomorfo al intervalo  $[0,1]$ , así que cuando hablemos de un arco podemos tener presente a este intervalo, al cual denotamos por  $I$ .

La curva cerrada simple es homeomorfa a la circunferencia unitaria  $\{(x,y) \in \mathbb{R}^2: x^2 + y^2 = 1\}$ , de este modo, al igual que sugerimos con el arco, cuando hablemos de una curva cerrada simple podemos tener en mente a la circunferencia unitaria, a la cual denotamos por  $S^1$ .

El  $n$ -odo simple es el continuo que resulta al unir  $n$  intervalos que comparten un único punto común  $v$  (al cual llamamos vértice del  $n$ -odo y es punto extremo de cada uno de los intervalos) de manera que la intersección de dos de estos intervalos es únicamente  $v$ . El 3-odo simple es mejor conocido como *triodo simple* y lo denotamos por  $T$ .

La curva sinoidal del topólogo se obtiene al considerar la cerradura en  $\square^2$  del subconjunto  $W = \{(x, \text{sen}(1/x)) \in \square^2: 0 < x \leq 1\}$ . Notemos que  $\text{Cl}_{\square^2}(W) = W \cup \{(0,y): -1 \leq y \leq 1\}$ .

Los continuos que hemos enumerado aparecen en la Figura 1.

Como prometimos, ahora presentamos los modelos de los hiperespacios  $C(X)$  para tres de los continuos que enumeramos arriba. Un modelo para  $C(I)$  es el triángulo  $\Delta = \{(x,y) \in \square^2: 0 \leq y \leq 1 \text{ y } y/2 \leq x \leq 1 - y/2\}$ , para  $C(S^1)$  es el disco  $D = \{(x,y) \in \square^2: x^2 + y^2 \leq 1\}$ , para  $C(T)$  es el cubo con alitas que aparece en la Figura 2 junto a los modelos que se describieron antes. El lector que esté interesado en la construcción de estos modelos puede consultar [2] o [4].

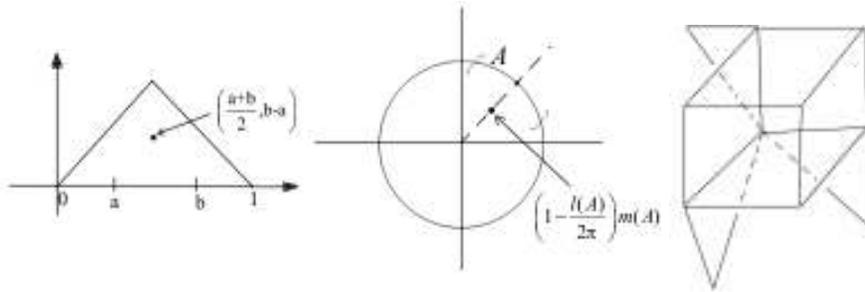


Figura 2. Modelos para los hiperespacios  $C(I)$ ,  $C(S^1)$  y  $C(T)$ .

### 3. Encajes ordenados para los hiperespacios $2^X$ y $C(X)$

En esta sección explicaremos por qué los encajes ordenados resultan más interesantes para los hiperespacios  $C(X)$  que para los hiperespacios  $2^X$ . Además construiremos ejemplos de encajes ordenados de hiperespacios para continuos específicos.

Iniciamos revisando las condiciones de existencia de los encajes ordenados para los hiperespacios  $2^X$ . Para ello necesitamos demostrar primero el siguiente teorema.

**Teorema 1.** Sea  $f: X \rightarrow Y$  una función continua. Si para cada  $A \in 2^X$  definimos  $2^f(A) = f(A)$ , entonces  $2^f(A) \in 2^Y$  y la función  $2^f: 2^X \rightarrow 2^Y$  así definida es continua.

**Demostración.** Como  $2^f(A) = f(A) \subset Y$ ,  $2^f(A)$  es un subconjunto no vacío de  $Y$ . Por otra parte,  $f(A)$  es la imagen continua de un subconjunto compacto de  $X$ , así que  $2^f(A) = f(A)$  resulta ser un subconjunto compacto de  $Y$ . Por lo que  $2^f(A) \in 2^Y$ . La continuidad de  $2^f$  se sigue del Lema 13.3 de [5].

La demostración del siguiente resultado la dejamos al lector.

**Corolario 2.** Si  $f: X \rightarrow Y$  es un encaje, entonces  $2^f: 2^X \rightarrow 2^Y$  es un encaje ordenado.

De hecho tenemos un resultado aún más poderoso, el cual nos dice bajo qué condiciones puede definirse un encaje ordenado entre los hiperespacios  $2^X$ .

**Teorema 3.** Sean  $X$  y  $Y$  continuos cualesquiera, entonces siempre es posible definir un encaje ordenado de  $2^X$  en  $2^Y$ . No incluimos la demostración de este resultado pues no es parte del objetivo de este trabajo, los interesados en consultarla en [1].

Debido a este teorema, el estudio de los encajes ordenados definidos para los hiperespacios  $2^X$  no es muy interesante. Qué pasa cuando consideramos el hiperespacio  $C(X)$ ? Resulta que los encajes ordenados definidos en estos hiperespacios no son triviales, esto es, existen continuos  $X$  y  $Y$  para los que a pesar de que  $C(X)$  puede encajarse en  $C(Y)$ ,  $C(X)$  no puede encajarse ordenadamente en  $C(Y)$ . Nuestro siguiente ejemplo muestra la certeza de esta afirmación.

**Ejemplo 4.**  $C(S^1)$  puede encajarse en  $C(I)$ , sin embargo  $C(S^1)$  no puede encajarse ordenadamente en  $C(I)$ .

Como vimos en la sección anterior, el disco  $D$  y el triángulo  $\Delta$  son modelos para  $C(S^1)$  y  $C(I)$  respectivamente. Así que de hecho  $C(S^1)$  es homeomorfo a  $C(I)$ , por lo que  $C(S^1)$  puede encajarse en  $C(I)$ . Veamos ahora por qué no es posible definir un encaje ordenado de  $C(S^1)$  en  $C(I)$ . Para ello supongamos lo contrario. Sea  $f: C(S^1) \rightarrow C(I)$  un encaje ordenado, entonces  $f(S^1)$  es un subintervalo cerrado de  $I$  tal que  $f(A) \subset f(S^1)$  para cada  $A \in C(S^1)$ . De este modo  $f$  también es un encaje ordenado de  $C(S^1)$  en  $f(C(S^1))$ . Por lo que, sin pérdida de generalidad, podemos asumir que  $f(S^1) = I$ . Por tanto,  $f$  es un encaje del disco  $D$  en el triángulo  $T$  que aplica un punto que no pertenece a la frontera de  $C(S^1)$  en un punto que pertenece a la frontera de  $C(I)$ . Esta contradicción demuestra que  $C(S^1)$  no puede encajarse ordenadamente en  $C(I)$ .

Debido a la naturaleza de los modelos para  $C(S^1)$  y  $C(T)$ , es claro que  $C(S^1)$  puede encajarse en  $C(T)$ . Veremos en nuestro siguiente ejemplo que de hecho podemos definir un encaje de  $C(S^1)$  en  $C(T)$  que respete el orden inducido por la contención,

**Ejemplo 5.**  $C(S^1)$  puede encajarse ordenadamente en  $C(T)$ .

Usaremos el siguiente modelo específico para  $T$ : Sea  $\{e_1, e_2, e_3\}$  la base canónica de  $\square^3$ . Dados dos puntos  $p, q \in \square^3$ , denotemos por  $pq$  al segmento que los une y,  $pq = p$  en el caso  $p = q$ . Definimos entonces  $T = \Theta[10e_1] \cup \Theta[10e_2] \cup \Theta[10e_3]$ , donde  $\Theta$  denota el origen de  $\square^3$ . Usaremos la letra  $d$  para denotar la distancia en  $\square^2$ . Enseguida llevaremos a cabo la tarea de definir nuestro encaje ordenado. Sea  $l: C(S^1) \rightarrow [0, 2\pi]$  la función que a cada subcontinuo de  $S^1$  le asigna su longitud. Sea  $m: C(S^1) \setminus \{S^1\} \rightarrow \square^2$  la función dada por  $m(A) = \text{punto medio de } A$ . Notemos que  $A \in C(S^1) \setminus \{S^1\}$ . Entonces  $A$  está completamente determinado por su punto medio  $m(A)$  y su longitud  $l(A)$ . Escribamos  $m(A) = (x_A, y_A)$ . Entonces definimos  $g: C(S^1) \setminus \{S^1\} \rightarrow C(T)$  por

$$g(A) = \Theta[(1 + x_A + l(A))e_1] \cup \Theta[(1 + y_A + l(A))e_2] \cup \Theta[l(A)e_3].$$

La continuidad de  $g$  se sigue de la continuidad de las funciones  $l$  y  $m$ .

Veamos que  $g$  es inyectiva. Para esto supongamos que  $g(A) = g(B)$ . Entonces,

$$\begin{aligned} 1 + x_A + l(A) &= 1 + x_B + l(B), \\ 1 + y_A + l(A) &= 1 + y_B + l(B), \\ l(A) &= l(B). \end{aligned}$$

Así que,  $l(A) = l(B)$  y  $m(A) = m(B)$ . Por tanto,  $A = B$ . Lo que demuestra que  $g$  es inyectiva.

Ahora vamos a demostrar que, si  $A \subset B$ , entonces  $g(A) \subset g(B)$ . Notemos que si  $A \subset B$  entonces  $l(A) \leq l(B)$ . Sea  $l'$  la longitud del arco más pequeño en  $S^1$  que une a  $m(A)$  con  $m(B)$ . Sea  $C$  la

mitad del arco  $B$  que contiene a  $m(A)$ . En el caso que  $m(A) = m(B)$ , tomamos como  $C$  a cualquiera de las dos mitades de  $B$ . Entonces,  $C$  contiene al arco que une a  $m(A)$  con  $m(B)$  y a una de las dos mitades del arco  $A$ . De aquí que  $l' + l(A)/2 \leq l(B)/2$ .

Entonces

$$2|x_A - x_B| \leq 2d(m(A), m(B)) \leq 2l' \leq l(B) - l(A).$$

De modo que  $1 + x_A + l(A) \leq 1 + x_B + l(B)$ . Análogamente,  $1 + y_A + l(A) \leq 1 + y_B + l(B)$ .

Además, puesto que  $l(A) \leq l(B)$ , concluimos que  $g(A) \subset g(B)$ .

Por tanto, hemos demostrado que  $g: C(S^1) \setminus \{S^1\} \rightarrow C(T)$  es un encaje ordenado.

Ahora estamos listos para definir nuestro encaje ordenado. Definimos  $f: C(S^1) \rightarrow C(T)$  por

$$\begin{aligned} f(A) &= \Theta[(2 + 2\pi)e_1] \cup \Theta[(2 + 2\pi)e_1] \cup \Theta[2\pi e_3], \text{ si } A = S^1 \text{ y,} \\ f(A) &= \Theta[((2 + 2\pi)l(A)/2\pi + (1 - (l(A)/2\pi))(1 + x_A + l(A))) e_1] \cup \Theta[((2 + 2\pi)l(A)/2\pi + (1 - \\ &\quad (l(A)/2\pi))(1 + y_A + l(A))) e_2] \cup \Theta[l(A)e_3], \text{ si } A \neq S^1. \end{aligned}$$

Puesto que  $f$  es el resultado de ajustar linealmente  $g$ , entonces  $f$  resulta ser continua e inyectiva.

En efecto, puesto que el ajuste es lineal,  $f$  es claramente continua en  $C(S^1) \setminus \{S^1\}$ . Ahora, si  $\{A_n\}_{n \in \mathbb{N}}$  es una sucesión en  $C(S^1) \setminus \{S^1\}$  tal que  $\lim A_n = S^1$ , entonces  $\lim l(A_n) = 2\pi$ . De manera que  $\lim f(A_n) = f(S^1)$ . Lo que demuestra que  $f$  es continua en  $C(S^1)$ .

Por otra parte, sean  $A, B \in C(S^1)$  tales que  $f(A) = f(B)$ . Supongamos que  $A = S^1$ . Entonces, como  $l(A) = l(B)$ , tenemos que  $B = S^1$ . Así que  $A = B$ .

Supongamos ahora que  $A \neq S^1$ . Entonces, puesto que  $l(A) = l(B)$ , tenemos que  $B \neq S^1$  y que

$$1 + x_A + l(A) \leq 1 + x_B + l(B).$$

$$1 + y_A + l(A) \leq 1 + y_B + l(B).$$

Por lo que  $g(A) = g(B)$ . Entonces, como  $g$  es inyectiva, tenemos que  $A = B$ . Por tanto  $f$  es inyectiva.

De este modo, sólo resta demostrar que  $A \subset B$  implica que  $f(A) \subset f(B)$ . Puesto que

$$(1 - l(B)/2\pi)(1 + x_B + l(B)) - (1 + x_A + l(A)) \geq 0,$$

tenemos que

$$(1 + x_B + l(B)) - (l(B)/2\pi)(1 + x_B + l(B)) \geq (1 + x_A + l(A)) - (l(B)/2\pi)(1 + x_A + l(A)).$$

De modo que

$$\begin{aligned} (2 + 2\pi)l(A)/2\pi + (1 - (l(A)/2\pi))(1 + x_A + l(A)) &= (1 + x_A + l(A)) + \\ &\quad [(2 + 2\pi) - (1 + x_A + l(A))] l(A)/2\pi \\ &\leq (1 + x_A + l(A)) + \\ &\quad [(2 + 2\pi) - (1 + x_A + l(A))] l(B)/2\pi \\ &= (1 + x_A + l(A)) - \\ &\quad (1 + x_A + l(A)) l(B)/2\pi + (2 + 2\pi)l(B)/2\pi \\ &\leq (1 + x_B + l(B)) - \\ &\quad (1 + x_B + l(B)) l(B)/2\pi + (2 + 2\pi)l(B)/2\pi \\ &= (1 + x_B + l(B)) + \\ &\quad [(2 + 2\pi) - (1 + x_B + l(B))] l(B)/2\pi \\ &= (2 + 2\pi)l(B)/2\pi + (1 - (l(B)/2\pi))(1 + x_B + l(B)) \end{aligned}$$

De manera análoga,

$$(2 + 2\pi)l(A)/2\pi + (1 - (l(A)/2\pi))(1 + y_A + l(A)) \leq (2 + 2\pi)l(B)/2\pi + (1 - (l(B)/2\pi))(1 + y_B + l(B))$$

Puesto que  $l(A) \leq l(B)$ , entonces  $f(A) \subset f(B)$ . Por tanto, hemos demostrado que  $C(S^1)$  puede encajarse ordenadamente en  $C(T)$ .

El siguiente ejemplo muestra que en ocasiones el encaje ordenado de hiperespacios se da de manera muy natural y sencilla. Para construir el encaje ordenado deseado necesitamos introducir el concepto de función de Whitney.

**Definición 6.** Una función de Whitney en el hiperespacio  $2^X$  es una función  $\mu: 2^X \rightarrow [0,1]$  que satisface:

- i)  $\mu(\{p\}) = 0$  para cada  $p \in X$ .
- ii)  $\mu(A) < \mu(B)$  siempre que  $A$  esté contenido propiamente en  $B$ ,
- iii)  $\mu(X) = 1$

En el Teorema 5.3 de [4], se demuestra que  $2^X$  admite funciones de Whitney para cualquier continuo  $X$ , llevando a cabo la construcción de dicha función.

Intuitivamente, una función de Whitney “mide el peso” de los subcontinuos de un continuo dado.

**Ejemplo 7.** Denotemos por  $X$  a la curva sinoidal del topólogo. Entonces  $C(X)$  puede encajarse ordenadamente en  $C(T)$ .

Recordemos que  $X \cong W \cup J$ , donde  $W = \{(x, \sin(1/x)) \in \mathbb{R}^2: 0 < x \leq 1\}$  y  $J = \{(0, y): -1 \leq y \leq 1\}$ . Sea  $\mu: C(X) \rightarrow [0,1]$  una función de Whitney y  $r: X \rightarrow J$  dada por  $r((z, w)) = (0, w)$ . Entonces definimos  $\varphi: X \rightarrow [0,1]$  dada por

$$\begin{aligned}\varphi((z, w)) &= 0, \text{ si } (z, w) \in J \text{ y,} \\ \varphi((z, w)) &= z, \text{ si } (z, w) \in W.\end{aligned}$$

Entonces  $\varphi$  es una función continua.

Sea  $\phi: J \rightarrow [0,1]$  un homeomorfismo. Ahora, para  $A \in C(X)$ , tomamos  $x_A = \max\varphi(A)$  y  $y_A = \max\phi(r(A))$ . Notemos que en  $C(X)$  tenemos tres tipos de subcontinuos, a saber, si  $A \in C(X)$  entonces,  $A \subset W$  o bien  $A \subset J$  o bien  $A$  es de la forma  $\varphi^{-1}([0, a])$  para alguna  $a = x_A$ .

Si  $A \subset W$ , entonces  $A$  está completamente determinada por  $\mu(A)$  y  $x_A$ . Si  $A \subset J$  entonces  $A$  queda determinada por  $\mu(A)$  y  $y_A$ . Por último si  $A$  es de la forma  $\varphi^{-1}([0, x_A])$ , entonces  $A$  queda determinada por  $\mu(A)$  y  $x_A$ . Entonces por lo discutido arriba definimos  $f: C(X) \rightarrow C(T)$  por

$$f(A) = \Theta[x_A e_1] \cup \Theta[y_A e_2] \cup \Theta[\mu(A) e_3].$$

La continuidad de  $f$  se sigue de la continuidad de las funciones  $x_A$ ,  $y_A$  y  $\mu(A)$ .

Probemos ahora que  $f$  es inyectiva. Sean  $A, B \in C(X)$  tales que  $f(A) = f(B)$  entonces  $x_A = x_B$ ,  $y_A = y_B$  y  $\mu(A) = \mu(B)$  entonces, por la discusión de arriba,  $A = B$ .

Finalmente, si  $A \subset B$ , entonces  $x_A \leq x_B$ ,  $y_A \leq y_B$  y  $\mu(A) \leq \mu(B)$ . Por lo que  $f(A) \subset f(B)$ . Por tanto hemos demostrado que  $C(X)$  puede encajarse ordenadamente en  $C(T)$ .

## 4. Conclusiones

Como se habrá podido apreciar, el concepto de encaje ordenado para hiperespacios de continuos es sencillo, pero el definir uno de estos encajes entre dos hiperespacios puede no ser una tarea sencilla. Una de las aplicaciones de los encajes ordenados dentro de la Teoría de hiperespacios es una caracterización de los continuos  $X$  para los cuales  $C(X)$  puede encajarse ordenadamente en  $C(Y)$ , cuando  $Y$  es un continuo hereditariamente indescomponible (ver [1]). En [3] se estudian encajes ordenados entre productos simétricos, hiperespacios cuyos elementos son subconjuntos no vacíos, cerrados y con a lo más  $n$  elementos. Actualmente estudiamos encajes ordenados

definidos entre el  $n$ -ésimo producto simétrico de un continuo  $X$  y el hiperespacio  $C(Y)$ , a este respecto ya se han obtenido resultados muy interesantes.

## 5. Referencias

- [1] Andablo G. (2002). Ordered embeddings of hyperspaces. En Illanes, A.; Macías, S.; Lewis, W. (Ed), *Continuum Theory* (Denton, TX, 1999), Lecures notes in Pure and Applied Mathematics, 230, New York, Marcel Dekker, 51-65.
- [2] Andablo G. y Castañeda E. (2008). Un breve espacio para los hiperspacios. *Ciencia Ergo Sum*, 15, No. 003, 317-324.
- [3] Andablo G. and Neumman V. (2008). Ordered embeddiings of symmetric products. *Houston J. Math.*, 34, No. 1, 115-122.
- [4] Illanes A. *Hiperspacios de continuos*. México, Aportaciones Matemáticas: Textos, 28. Sociedad Matemática Mexicana.
- [5] Illanes A. and Nadler, Jr. S. B. (1999). *Hyperspaces: Fundamentals and Recent Advances*. New York and Basel. Marcel Dekker, Inc.
- [6] Ward Jr. L. (1980). A note on Whitney maps. *Canad. Math. Bull.*, 23, 373-374.
- [7] Ward Jr. L. (1981). Extending Whitney maps. *Pacific. J. Math.*, 93, 465-469.



# La conexidad del hiperespacio $2^X$

Alejandra Fonseca Morales y Carlos Alberto Robles Corbalá

Departamento de Matemáticas

Universidad de Sonora

alejandra.fonsecam@correoa.uson.mx, crobles@gauss.mat.uson.mx

## Resumen

*En este trabajo se mostrará que para un continuo  $X$ , su hiperespacio  $2^X$  es conexo. Para esto, empezaremos definiendo algunos conceptos básicos, siguiendo con la prueba de la densidad de  $F(X)$  en  $2^X$ , para finalizar con la prueba de la conexidad del hiperespacio  $2^X$ .*

## 1. Introducción

En este trabajo se definirá el concepto de continuo, además mencionaremos los hiperespacios más conocidos:  $2^X$ ,  $C(X)$ ,  $F_n(X)$  y  $C_n(X)$ . Como todos los hiperespacios mencionados anteriormente están contenidos en  $2^X$ , es suficiente dotar de una métrica a  $2^X$ , y así  $C(X)$ ,  $F_n(X)$  y  $C_n(X)$  dispondrán de esta métrica. Se define entonces la métrica de Hausdorff para el hiperespacio  $2^X$ .

Una vez dada la métrica para los hiperespacios, se probará que la unión de los hiperespacios  $F_n(X)$ , es un conjunto denso en  $2^X$ . Además se mostrará que si  $g: X^n \rightarrow F_n(X)$  es la función definida por  $g((x_1, \dots, x_n)) = \{x_1, \dots, x_n\}$ , entonces  $g$  es una función continua y suprayectiva.

Con lo dicho anteriormente quedará probada la conexidad de  $2^X$ .

## 2. Definiciones básicas y la métrica de Hausdorff

Para iniciar nuestro estudio veamos las siguientes definiciones y ejemplos.

**Definición 1.** Diremos que  $X$  es un continuo si es un espacio métrico, compacto, conexo y no vacío.

**Ejemplo 1.** Algunos continuos son:

1. Dados  $a, b \in \mathbb{R}$ , el intervalo  $[a, b]$  es un continuo.
2. Dados  $x \in \mathbb{R}^n$  y  $\epsilon > 0$ , la bola cerrada  $B_\epsilon(x)$ , es un continuo.
3. Si  $W = \{(x, \sin(1/x)) \in \mathbb{R}^2 : 0 < x \leq (1/2\pi)\}$  entonces  $X = W \cup \{(0, y) \in \mathbb{R}^2 : -1 \leq y \leq 1\}$  es un continuo llamado *la curva sinoidal del topólogo*.

4. Tomemos el continuo  $X$  del ejemplo anterior y consideremos un arco  $Z$  del punto  $(0, -1)$  al punto  $(2\pi, 1)$ , de tal forma que la intersección  $X \cap Z = \{(0, -1), (2\pi, 1)\}$ . Entonces  $V = X \cup Z$  es un continuo llamado *el círculo de Varsovia*.

Una de las técnicas más importantes para obtener ejemplos interesantes de continuos es el uso de *intersecciones anidadas*.

**Teorema 1.1**

Si  $\{X_i\}_{i=1}^{\infty}$  es una sucesión de subcontinuos de un espacio métrico  $(Y, d)$ , tal que para toda  $n \in \mathbb{N}$ ,  $X_{n+1} \subset X_n$  entonces  $\bigcap_{n=1}^{\infty} X_n$  es un continuo.

La demostración la podemos encontrar en [6], en la página 13.

5. La curva universal de Sierpinski. Empezamos dividiendo el cuadrado  $S_0 = \{[0, 1] \times [0, 1]\}$  en nueve cuadrados congruentes y tomamos  $S_1 = S_0 - \{(1/3, 2/3) \times (1/3, 2/3)\}$ . Análogamente, dividimos cada uno de los restantes ocho cuadrados en nueve cuadrados congruentes, y llamamos  $S_2$  al continuo que se obtiene al quitar el interior de cada uno de los ocho cuadrados centrales. Continuando de esta manera, definimos  $S_3, S_4$ , etc. Sea  $S = \bigcap_{n=1}^{\infty} S_n$ , entonces por el Teorema 1.1,  $S$  es un continuo, llamado *la curva universal de Sierpinski*.
6. La curva universal de Menger. Consideremos primero el cubo  $M = \{[0, 1] \times [0, 1] \times [0, 1]\}$ . Dividimos cada una de las caras de  $M$  en nueve cuadrados congruentes y hagamos un agujero a través del interior de cada cuadrado central, esto nos da un continuo  $M_1$ . Dividamos cada uno de los restantes cuarenta y ocho cuadrados en nueve cuadrados congruentes y hagamos un agujero a través del interior de los cuadrados centrales, de esta manera obtenemos un continuo  $M_2$ . Procedemos inductivamente para obtener continuos  $M_n$ , *la curva universal de Menger*, es por definición  $M = \bigcap_{n=1}^{\infty} M_n$ , y por el Teorema 1.1,  $M$  es un continuo.

**Definición 2.** Un *hiperespacio* es una familia de subconjuntos de  $X$ , a la que le asociamos alguna propiedad en particular.

Los que la literatura registra como los más estudiados son:

1.  $2^X = \{A \subset X : A \text{ es cerrado y no vacío}\}$
2.  $C(X) = \{A \in 2^X : A \text{ es conexo}\}$
3.  $F_n(X) = \{A \in 2^X : A \text{ tiene a lo más } n \text{ puntos}\} (n \in \mathbb{N})$
4.  $C_n(X) = \{A \in 2^X : A \text{ tiene a lo más } n \text{ componentes}\} (n \in \mathbb{N})$

Nosotros nos enfocaremos a probar la conexidad de  $2^X$ . Los siguientes conceptos son básicos para nuestro objetivo.

**Definición 3.** Sea  $\epsilon > 0$ ,  $p \in X$  y  $A \in 2^X$ , se define la bola de radio  $\epsilon$  y centro  $p$  por  $B(\epsilon, p) = \{q \in X : d(p, q) < \epsilon\}$ . Y la nube de radio  $\epsilon$  y centro en  $A$  por  $N(\epsilon, A) = \{q \in X : \exists p \in A \text{ tal que } d(p, q) < \epsilon\}$ .

Es fácil convencerse de que  $N(\epsilon, \{p\}) = B_\epsilon(p)$  y  $N(\epsilon, A) = \cup \{B_\epsilon(a) : a \in A\}$ .

**Definición 4.** Dados  $A, B \in 2^X$  definimos  $H : 2^X \times 2^X \rightarrow \mathbb{R}^+$  por  $H(A, B) = \inf \{\epsilon > 0 : A \subset N(\epsilon, B) \wedge B \subset N(\epsilon, A)\}$ .  
H se conoce como la *métrica de Hausdorff*.

**Lema 1.**  $H(A, B) < \kappa \Leftrightarrow B \subset N(\kappa, A) \wedge A \subset N(\kappa, B)$ .

Para la demostración consultar [5].

**Proposición 2.** Dados  $A, B, C \in 2^X$ , se cumple que:

1.  $H(A, B)$  está bien definida,
2.  $H(A, B) \geq 0$ ,
3.  $H(A, B) = H(B, A)$ ,
4.  $H(A, B) = 0$  si y sólo si  $A = B$ ,
5.  $H$  satisface la desigualdad del triángulo, es decir:  $H(A, B) \leq H(A, C) + H(C, B)$ .

*Demostración.* La demostración la podemos encontrar en [2], en las páginas 22, 23, 24. □

### 3. Densidad en $2^X$

Recordemos que un subconjunto  $A$  de un espacio topológico  $(X, \tau)$  es denso en  $X$ , si y sólo si, para todo abierto  $B$  en  $X$ , se tiene que  $B \cap A \neq \emptyset$ .

**Lema 3.** El conjunto  $F(X) = \bigcup \{F_n(X) : n \in \mathbb{N}\}$  es denso en  $2^X$ .

*Demostración.* Tenemos que probar que toda bola centrada en cualquier  $A \in 2^X$ , se intersecta con  $F(X)$ . Para esto tomemos  $\epsilon > 0$  y  $A \in 2^X$ , entonces  $\{B(\epsilon, a) : a \in A\}$  es una cubierta por abiertos para el compacto  $A$ , ( $A$  es un conjunto cerrado, no vacío contenido en el compacto  $X$ , por tanto  $A$  es compacto). Entonces existen  $a_1, a_2, \dots, a_n \in A$ , tales que  $A \subset \bigcup_{i=1}^n B(\epsilon, a_i) = N(\epsilon, \{a_1, a_2, \dots, a_n\})$ . Por otra parte,  $\{a_1, a_2, \dots, a_n\} \subset A \subset N(\epsilon, A)$ . Entonces  $\{a_1, a_2, \dots, a_n\} \in B(\epsilon, A) \cap F(X)$ , por lo tanto  $F(X)$  es denso en  $2^X$ . □

**Lema 4.** Sea  $X^n$  el producto topológico de  $n$  copias del continuo  $X$  por si mismo. Si  $g: X^n \rightarrow F_n(X)$  es la función definida por  $g((x_1, \dots, x_n)) = \{x_1, \dots, x_n\}$ , entonces  $g$  es una función continua y suprayectiva.

*Demostración.* Sea  $D((x_1, \dots, x_n), (y_1, \dots, y_n)) = \max \{d(x_1, y_1), \dots, d(x_n, y_n)\}$  la métrica para  $X^n$ , esta métrica induce la topología producto.

\*Probemos la continuidad de  $g$ .

Sea  $\epsilon > 0$  y  $\delta = \epsilon$ , tal que  $D((x_1, \dots, x_n), (y_1, \dots, y_n)) = \max \{d(x_1, y_1), \dots, d(x_n, y_n)\} < \epsilon$ . Entonces  $d(x_i, y_i) < \epsilon \forall i \in \{1, \dots, n\}$ . De modo que  $\{x_1, \dots, x_n\} \subset N(\epsilon, \{y_1, \dots, y_n\})$  y  $\{y_1, \dots, y_n\} \subset N(\epsilon, \{x_1, \dots, x_n\})$ , entonces  $H(\{x_1, \dots, x_n\}, \{y_1, \dots, y_n\}) < \epsilon$ . Por lo tanto  $g$  es continua.

\* $g$  es suprayectiva.

Cada elemento de  $F_n(X)$  se puede escribir en la forma  $\{x_1, \dots, x_n\}$ , pues si tomamos un conjunto de menos de  $n$  elementos, bastará con repetir algunos y esto no cambia al conjunto.

Por lo tanto  $g$  es continua y suprayectiva.  $\square$

## 4. Conexidad de $2^X$

**Teorema 5.** El hiperespacio  $2^X$  es conexo.

*Demostración.* Notemos que para cada  $n \in \mathbb{N}$ ,  $F_1(X) \subset F_n(X)$  con  $F_n(X)$  conexo, dado que la imagen de un conexo bajo una función continua  $g$ , es conexo. De manera que el conjunto  $F(X) = \bigcup \{F_n(X) : n \in \mathbb{N}\}$ , es la unión de conexos con intersección común  $F_1(X)$ . Así que  $F(X)$  es conexo, por lo que su cerradura es conexa. Por lo tanto  $2^X$  es conexo.  $\square$

## 5. Conclusiones

Con la demostración de que el hiperespacio  $2^X$  es conexo y probando adicionalmente que este espacio también es compacto, tendremos un resultado importante, a saber:

El espacio  $2^X$  con la métrica de Hausdorff, es un espacio métrico, conexo, compacto y no vacío, es decir, el hiperespacio  $2^X$  de un continuo  $X$ , es un continuo.

## 6. Referencias

- [1] Escobedo R., Macías S. y Méndez H. (2006). *Invitación a la teoría de los continuos y sus hiperespacios*. México, D.F. Aportaciones Matemáticas SMM.
- [2] Illanes M., A. (2004). *Hiperespacios de Continuos*. México, D.F. Aportaciones Matemáticas SMM.

- [3] Sam B. Nadel, Jr.(2006). *Hyperspaces of sets. A text with research questions*. México, D.F. Aportaciones Matemáticas SMM.
- [4] Munkres, J. (2002). *Topología*. Ed. Prentice Hall.
- [5] Aguilera M., M. y García V., L. (2003). *Hiperespacios de Continuos Métricos y No Métricos*. Tesis.
- [6] Macías, S. (2000). *Introducción a la teoría de los continuos*. Instituto de Matemáticas, UNAM.



# La propiedad de compacidad en el hiperespacio $2^X$

Carmen Geraldí Higuera Chan y Carlos Alberto Robles Corbalá

Departamento de Matemáticas

Universidad de Sonora

carg\_hch@hotmail.com, crobles@gauss.mat.uson.mx

## Resumen

*A partir de un continuo  $X$  arbitrario, se probará que  $2^X$ , el hiperespacio de cerrados no vacíos de  $X$ , es un conjunto compacto. Para hacer esto se demostrará que toda sucesión  $(A_n)_{n=1}^{\infty}$  en  $2^X$  tiene una subsucesión de Cauchy, y por otra parte, que toda sucesión de Cauchy en  $2^X$  converge, es decir  $2^X$  es completo. Con esto terminaríamos, ya que toda sucesión en  $2^X$  tendrá una subsucesión convergente en  $2^X$ , esto es,  $2^X$  es compacto.*

## 1. Introducción

El presente trabajo tiene como finalidad la demostración de la compacidad del hiperespacio  $2^X$  de un continuo  $X$ . Una de las propiedades para la demostración de que el hiperespacio de un continuo, resulta a su vez, ser otro continuo.

Para esto se incluyen secciones previas al contenido central del trabajo, con la finalidad de tener un trabajo en el que no sea necesario el recurrir a otras fuentes para entenderlo.

Se incluye una sección de definiciones básicas, donde se definen los principales objetos de estudio, continuo e hiperespacio. Además la definición de una métrica en el hiperespacio, llamada la métrica de Hausdorff.

En la otra sección se mostrarán algunas características de la convergencia con esta métrica en el hiperespacio, que posteriormente serán utilizadas. Para finalizar, en las últimas dos secciones, se estará en condiciones de demostrar la completez del hiperespacio  $2^X$ , con lo que se tendrá la compacidad de  $2^X$ .

## 2. Definiciones Básicas

En esta parte se definirán conceptos básicos para la continuación del objetivo principal del trabajo.

**Definición 1.** Un continuo es un espacio métrico conexo, compacto y no degenerado (con más de dos puntos).

**Definición 2.** Dado un espacio topológico  $X$ . Un hiperespacio de  $X$  es una familia de subconjuntos de  $X$  con una propiedad “ $p$ ”.

En nuestro caso, estudiaremos el hiperespacio de un continuo  $X$  (la topología de  $X$  es la inducida por la métrica  $d$  de  $X$ ) y  $p$  es la propiedad de ser un conjunto cerrado y no vacío. Esto es:

$$2^X = \{A \subset X \mid A \text{ es cerrado y } A \neq \emptyset\}$$

Ahora definiremos una métrica para este hiperespacio, la cual nos permitirá analizar la cercanía entre dos puntos del hiperespacio  $2^X$ .

**Definición 3.** Sea  $X$  un continuo con métrica  $d$  y sea  $\epsilon > 0$ . Definimos el conjunto:

$$N(\epsilon, A) = \{x \in X : \text{existe } a \in A \text{ tal que } d(a, x) < \epsilon\}$$

llamada la nube de radio  $\epsilon$  con centro en  $A$ .

**Lema 1.** Sean  $A, B \in 2^X$ .

- (a)  $N(\epsilon, A) = \bigcup_{a \in A} B_\epsilon(a)$ .
- (b) Si  $A \subseteq B$ , entonces  $N(\epsilon, A) \subseteq N(\epsilon, B)$ .
- (c) Si  $0 < \delta < \epsilon$ , entonces  $N(\delta, A) \subseteq N(\epsilon, A)$ .

*Demostración.* (a) Sea  $x \in N(\epsilon, A)$ , entonces existe  $a \in A$  tal que  $d(a, x) < \epsilon$ , o sea que  $x \in B_\epsilon(a) \subseteq \bigcup_{a \in A} B_\epsilon(a)$ .

Si  $x \in \bigcup_{a \in A} B_\epsilon(a)$ , entonces existe  $a \in A$  tal que  $x \in B_\epsilon(a)$ , es decir,  $d(x, a) < \epsilon$  por lo que  $x \in N(\epsilon, A)$ .

(b) Sea  $x \in N(\epsilon, A)$ , entonces existe  $a \in A$  tal que  $d(a, x) < \epsilon$ , y como  $A \subseteq B$ , tenemos  $a \in B$  y  $d(a, x) < \epsilon$ , entonces  $x \in N(\epsilon, B)$ .

(c) Sea  $x \in N(\delta, A)$ , entonces existe  $a \in A$  tal que  $d(a, x) < \delta$ , y como  $0 < \delta < \epsilon$ , tenemos  $a \in A$  y  $d(a, x) < \delta < \epsilon$ , por lo que  $x \in N(\epsilon, A)$ .  $\square$

**Definición 4.** La métrica de Hausdorff para  $2^X$  se define por:

$$H(A, B) = \inf \{\epsilon > 0 : A \subset N(\epsilon, B) \text{ y } B \subset N(\epsilon, A)\}$$

Con esta métrica podemos notar de una manera intuitiva, dos conjuntos están cercanos, si y sólo si, están casi empalmados uno con el otro.

Efectivamente,  $H$  es una métrica.

**Teorema 2.** Sea  $X$  un continuo y  $H$  como en la definición, entonces:

1.  $H$  está bien definida.
2.  $H(A, B) \geq 0$ .
3.  $H(A, B) = H(B, A) \forall A, B \in 2^X$ .
4.  $H(A, B) = 0 \iff A = B$ .
5.  $H(A, C) \leq H(A, B) + H(B, C)$  para todo  $A, B, C \in 2^X$ .

*Demostración.* La prueba la podemos encontrar en [2] págs. 22, 23 y 24 o también en [5] págs. 2 y 3.  $\square$

**Proposición 3.** Sean  $A, B \in 2^X$ . Entonces  $H(A, B) < r$ , si y sólo si,  $A \subset N(r, B)$  y  $B \subset N(r, A)$ .

*Demostración.* ( $\Rightarrow$ ) Sea  $\epsilon(A, B) = \{\epsilon > 0 : A \subset N(\epsilon, B) \text{ y } B \subset N(\epsilon, A)\}$ . Por la definición de ínfimo, existe  $r_0 \in \epsilon(A, B)$ , con  $H(A, B) < r_0 < r$ .

Dado que  $r_0 \in \epsilon(A, B)$  y  $r_0 < r$ , tenemos que  $A \subset N(r, B)$  y  $B \subset N(r, A)$ .

( $\Leftarrow$ ) Como  $A \subset N(r, B)$ , tenemos que para todo elemento  $a \in A$  existe  $b_a \in B$  con  $a \in B_r(b_a)$  (esto es, si y sólo si,  $d(a, b_a) < r$ ).

Entonces existe  $\delta_a > 0$  tal que  $d(a, b_a) < \delta_a < r$ . Así  $A \subset \bigcup_{a \in A} B_{\delta_a}$ .

Como  $A$  es compacto (es un cerrado dentro de un compacto). Existe  $n \in \mathbb{N}$  y  $a_1, a_2, \dots, a_n \in A$ , tales que  $A \subset \bigcup_{i=1}^n B_{\delta_{a_i}}(b_{a_i}) \subset \bigcup_{i=1}^n N(\delta_{a_i}, \{b_{a_i}\}) \subset \bigcup_{i=1}^n N(\delta_{a_i}, B)$ .

Análogamente existen  $b_1, b_2, \dots, b_m \in B$  tales que  $B \subset \bigcup_{i=1}^m N(\delta_{b_i}, A)$ .

Sea  $s = \text{máx}\{\delta_{a_1}, \dots, \delta_{a_n}, \delta_{b_1}, \dots, \delta_{b_m}\}$ . Entonces  $A \subset N(s, B)$  y  $B \subset N(s, A)$ .

De esta manera,  $H(A, B) = \text{ínf } \epsilon(A, B) \leq s < r$ . □

### 3. Convergencia en $2^X$

El objetivo de esta sección es la caracterización de la convergencia con la métrica de Hausdorff de una sucesión en  $2^X$  en términos de subconjuntos de  $X$  que a continuación vamos a definir.

**Definición 5.** Sea  $\{A_n\}_{n=1}^{\infty}$  una sucesión de elementos en  $2^X$ .

1.  $\text{lím inf}(A_n) = \{x \in X \mid \text{para todo } \epsilon > 0; B_{\epsilon}(x) \cap A_n \neq \emptyset \text{ para casi toda } n\}$  (todas excepto un número finito).
2.  $\text{lím sup}(A_n) = \{x \in X \mid \text{para todo } \epsilon > 0; B_{\epsilon}(x) \cap A_n \neq \emptyset \text{ para una infinidad de } n\}$ .

Inmediatamente tenemos la siguiente proposición.

**Proposición 4.** Sea  $X$  un continuo y  $\{A_n\}_{n=1}^{\infty}$  una sucesión de elementos en  $2^X$ .

Entonces:

1.  $\text{lím inf}(A_n) \subseteq \text{lím sup}(A_n)$ .
2.  $\text{lím inf}(A_n)$  y  $\text{lím sup}(A_n)$  son subconjuntos cerrados de  $X$ .
3.  $\text{lím sup}(A_n) \neq \emptyset$ .

*Demostración.* La prueba la podemos encontrar en [6] págs. 8 y 9. □

**Proposición 5.** Sea  $X$  un continuo y  $\{A_n\}_{n=1}^{\infty} \subseteq 2^X$  una sucesión de subconjuntos de  $X$ .

$x \in \text{lím sup}(A_n)$ , si y sólo si, existe una sucesión de números naturales  $n_1, n_2, \dots$  y existen puntos  $x_{n_k} \in A_{n_k}$  para todo  $k \in \mathbb{N}$  tal que  $\text{lím}_{k \rightarrow \infty} x_{n_k} = x$ .

*Demostración.* ( $\Leftarrow$ ) Supongamos que existe una sucesión de números naturales  $n_1, n_2, \dots$  y existen puntos  $x_{n_k} \in A_{n_k}$  para todo  $k \in \mathbb{N}$  tales que  $x_{n_k} \rightarrow x$ .

Por demostrar que  $x \in \text{lím sup}(A_n) = \{x \in X \mid \text{para todo } \epsilon > 0; B_{\epsilon}(x) \cap A_n \neq \emptyset \text{ para una infinidad de } n\}$ .

Sea  $\epsilon > 0$ , entonces existe  $K \in \mathbb{N}$  tal que  $d(x, x_{n_k}) < \epsilon$  para toda  $k \geq K$  (porque  $\text{lím}_{k \rightarrow \infty} x_{n_k} = x$ ), con esto tenemos que  $x_{n_k} \in A_{n_k} \cap B_{\epsilon}(x)$  para toda  $k \geq K$ .

De aquí que  $B_{\epsilon}(x) \cap A_n \neq \emptyset$  para una infinidad de  $n$ 's ( $n_k, n_{k+1}, \dots$ ). Por lo tanto  $x \in \text{lím sup}(A_n)$ .

( $\Rightarrow$ ) Tomemos  $x \in \text{lím sup}(A_n)$ . Entonces, para toda  $\epsilon > 0$ ,  $B_{\epsilon}(x) \cap A_n \neq \emptyset$  para una infinidad de

$n$ 's. En particular si tomamos  $\epsilon = 1$  se tiene que  $B_1(x) \cap A_n \neq \emptyset$  para una infinidad de  $n$ 's.

Elegimos  $n_1 \in \mathbb{N}$  y  $x_{n_1} \in B_1(x) \cap A_{n_1}$ , entonces  $d(x, x_{n_1}) < 1$  y  $x_{n_1} \in A_{n_1}$ .

Ahora para  $\epsilon = 1/2$  se tiene que  $B_{1/2}(x) \cap A_n \neq \emptyset$  para una infinidad de  $n$ 's, entonces podemos elegir  $n_2 > n_1$  tal que  $B_{1/2}(x) \cap A_{n_2} \neq \emptyset$ . Escogemos  $x_{n_2} \in B_{1/2}(x) \cap A_{n_2}$ , entonces  $d(x, x_{n_2}) < 1/2$  y  $x_{n_2} \in A_{n_2}$ .

Procediendo inductivamente, construimos una sucesión de números naturales  $n_1 < n_2 < \dots$  y de puntos  $x_{n_k} \in A_{n_k}$  tales que  $d(x, x_{n_k}) < 1/k$ .

De esto concluimos que  $x_{n_k} \rightarrow x$  cuando  $k \rightarrow \infty$ .  $\square$

Ahora probemos el siguiente resultado. Donde veremos la relación de los límites inferiores y superiores con la convergencia en la métrica de Hausdorff.

**Teorema 6.** Sea  $\{A_n\}_{n=1}^{\infty} \subseteq 2^X$  entonces  $\{A_n\}_{n=1}^{\infty}$  converge con la métrica de Hausdorff a un  $A \in 2^X$ , si y sólo si,  $\limsup(A_n) = \liminf(A_n) = A$ .

*Demostración.* ( $\Rightarrow$ ) Supongamos que  $A_n \rightarrow A$  cuando  $n \rightarrow \infty$  en  $2^X$  con la métrica de Hausdorff. Pd.  $\liminf(A_n) = A = \limsup(A_n)$ .

Sabemos de la proposición anterior que  $\liminf(A_n) \subseteq \limsup(A_n)$ .

Así que es suficiente probar:

- (1)  $A \subseteq \liminf(A_n)$ .
- (2)  $\limsup(A_n) \subseteq A$ .

Dem (1).

Sea  $a \in A$ , Pd.  $a \in \liminf(A_n)$ .

Dado  $\epsilon > 0$ . Como  $\lim_{n \rightarrow \infty} A_n = A$  con la métrica de Hausdorff. Entonces existe  $N \in \mathbb{N}$  tal que  $H(A, A_n) < \epsilon$ , si  $n \geq N$ .

Observe que  $H(A, A_n) = \inf \{ \lambda > 0 : A \subseteq N(\lambda, A_n) \text{ y } A_n \subseteq N(\lambda, A) \} < \epsilon$ , para  $n \geq N$ . Entonces para  $n \geq N$ ,  $A \subseteq N(\epsilon, A_n)$  y  $A_n \subseteq N(\epsilon, A)$ . Así,  $a \in A \subseteq N(\epsilon, A_n)$  tenemos que  $a \in N(\epsilon, A)$  para  $n \geq N$ .

Entonces existe  $x_n \in A_n$  tal que  $d(a, x_n) < \epsilon$ , si  $n \geq N$ .

Por lo tanto para los  $n$ 's  $\geq N$ ,  $x_n \in A_n \cap B_\epsilon(a)$ .

De aquí que para  $\epsilon > 0$ ,  $A_n \cap B_\epsilon(a) \neq \emptyset$  para todas excepto un número finito de  $n$ 's ( $n \geq N$ ).

Por lo tanto  $a \in \liminf(A_n)$ . Con esto  $A \subseteq \liminf(A_n)$ .

Dem (2).

Supongamos lo contrario, es decir,  $\limsup(A_n) \not\subseteq A$ .

Así existe  $x \in \limsup(A_n)$  tal que  $x \notin A$ .

Como  $A$  es cerrado. Existe  $\epsilon > 0$  tal que  $B_\epsilon(x) \cap A = \emptyset$ .

Como  $x \in \limsup(A_n) = \{x \in X \mid \text{para todo } \epsilon > 0; B_\epsilon(x) \cap A_n \neq \emptyset \text{ para una infinidad de } n\}$ .

Entonces, para  $\epsilon > 0$ , que existe, se cumple

$$B_\epsilon(x) \cap A_n \neq \emptyset \text{ para una infinidad de } n\text{'s}$$

Como  $\lim_{n \rightarrow \infty} A_n = A$ , entonces para  $\frac{\epsilon}{2} > 0$ , existe  $N \in \mathbb{N}$  tal que  $H(A_n, A) < \frac{\epsilon}{2}$ , para  $n \geq N$ .  
Es decir  $A \subseteq N(\frac{\epsilon}{2}, A_n)$  y  $A_n \subseteq N(\frac{\epsilon}{2}, A)$  para  $n \geq N$ .

Tomemos  $M \geq N$  tal que  $B_{\frac{\epsilon}{2}} \cap A_M \neq \emptyset$  y sea  $z \in B_{\frac{\epsilon}{2}} \cap A_M \neq \emptyset$ .

Entonces  $d(x, z) < \frac{\epsilon}{2}$  y  $z \in A_M \subseteq N(\frac{\epsilon}{2}, A)$ .

Así  $d(x, z) < \frac{\epsilon}{2}$  y existe  $a \in A$  tal que  $d(a, z) < \frac{\epsilon}{2}$ .

Por lo tanto,  $d(x, a) \leq d(x, z) + d(z, a) = \epsilon$  implica que  $d(x, a) < \epsilon$ , para  $a \in A$ .

Con lo cual tenemos  $B_\epsilon(x) \cap A \neq \emptyset$ . Pues  $a \in B_\epsilon(x) \cap A \neq \emptyset$ . Pero esto es una contradicción a la elección de  $\epsilon$  tal que  $B_\epsilon(x) \cap A = \emptyset$ .

Por lo tanto  $\limsup(A_n) = A$ .

De (1) y (2)  $\lim(A_n)_{n \rightarrow \infty} = A$ .

( $\Leftarrow$ ) Supongamos que  $\liminf(A_n) = \limsup(A_n)$ . Pd.  $A_n \in 2^X$  si  $n \rightarrow \infty$ .

Sea  $A = \limsup(A_n)$  y  $A \in 2^X$  ( $A \neq \emptyset$  y  $A$  cerrado).

Probemos que  $A_n \rightarrow A$  cuando  $n \rightarrow \infty$  con la métrica de Hausdorff.

Sea  $\epsilon > 0$  probemos que

- (a) Existe  $M_1 \in \mathbb{N}$  tal que  $A \subseteq N(\epsilon, A_n)$  para todo  $n \geq M_1$ .
- (b) Existe  $M_2 \in \mathbb{N}$  tal que  $A_n \subseteq N(\epsilon, A)$  para todo  $n \geq M_2$ .

Dem.(a).

Observemos que la familia  $\{B_{\frac{\epsilon}{2}}(a) : a \in A\}$  es una cubierta abierta para  $A$ .

$A$  es cerrado, no vacío en el compacto  $X$ , entonces  $A$  es compacto.

Por lo que existe una subcubierta finita, esto es, existe  $m \in \mathbb{N}$  y  $a_1, a_2, \dots, a_m \in A$  tal que

$$A \subseteq B_{\frac{\epsilon}{2}}(a_1) \cup B_{\frac{\epsilon}{2}}(a_2) \cup \dots \cup B_{\frac{\epsilon}{2}}(a_m).$$

Ahora, como  $A = \liminf(A_n) = \{x \in X \mid \text{para todo } \epsilon > 0; B_\epsilon(x) \cap A_n \neq \emptyset \text{ para casi toda } n\}$  (todas excepto un número finito) y  $a_i \in A$  para toda  $i = 1, \dots, m$ .

Luego para  $\lambda > 0$ ,  $B_\lambda(a_i) \cap A_n \neq \emptyset$  para casi toda  $n$  (todas excepto un número finito), así para cada  $i \in \{1, 2, \dots, m\}$  existe  $N_i \in \mathbb{N}$ , tal que si  $n \geq N_i$  entonces  $B_{\frac{\epsilon}{2}}(a_i) \cap A_n \neq \emptyset$ .

Sea  $M_1 = \max\{N_1, \dots, N_m\}$ . Así dada  $n \geq M_1$  y  $i \in \{1, 2, \dots, m\}$ ,  $B_{\frac{\epsilon}{2}}(a_i) \cap A_n \neq \emptyset$ .

Con esto tenemos la siguiente afirmación

$$A \subseteq N(\epsilon, A_n) \text{ para todo } n \geq M_1.$$

En efecto, sea  $n \geq M_1$  y  $a \in A \subseteq \bigcup_{i=1}^m B_{\frac{\epsilon}{2}}(a_i)$ , entonces existe  $i_0 \in \{1, 2, \dots, m\}$  tal que  $a \in B_{\frac{\epsilon}{2}}(a_{i_0})$ , es decir,  $d(a, a_{i_0}) < \epsilon/2$ .

Además para los  $n \geq M_1$ , existe  $x \in B_{\frac{\epsilon}{2}}(a_{i_0}) \cap A_n$ .

Luego  $d(a, x) \leq d(a, a_{i_0}) + d(a_{i_0}, x) < \epsilon/2 + \epsilon/2 = \epsilon$ .

Así  $d(a, x) < \epsilon$ , implica que  $a \in N(\epsilon, A_n)$  para  $n \geq M_1$ .

Por lo tanto  $A \subseteq N(\epsilon, A_n)$  para  $n \geq M_1$ .

Dem. (b).

Supongamos que (b) es falso, esto es:

Para toda  $N \in \mathbb{N}$  existe  $n's \geq N$  tal que  $A_n \subsetneq N(\epsilon, A)$

Así:

Para  $N = 1$  existe  $n_1 \geq 1$  tal que  $A_{n_1} \subsetneq N(\epsilon, A)$   
 Para  $N = n_1 + 1$  existe  $n_2 > n_1$  tal que  $A_{n_2} \subsetneq N(\epsilon, A)$   
 Para  $N = n_2 + 1$  existe  $n_3 > n_2$  tal que  $A_{n_3} \subsetneq N(\epsilon, A)$

Procediendo de manera inductiva, existe una sucesión de números naturales  $n_1 < n_2 < n_3 < \dots$  tal que  $A_{n_k} \subsetneq N(\epsilon, A)$  para todo  $k \in \mathbb{N}$ .

Luego para cada  $k \in \mathbb{N}$  tomemos  $x_{n_k} \in A_{n_k} - N(\epsilon, A) \subseteq X$ .

Como  $X$  es compacto, existe  $x_0 \in X$  y una subsucesión  $\{x_{n_{k_i}}\}_{i=1}^{\infty}$  de  $\{x_{n_k}\}_{k=1}^{\infty}$  tal que  $\lim_{i \rightarrow \infty} x_{n_{k_i}} = x_0$ .

Observemos que para todo  $i \in \mathbb{N}$ ,  $x_{n_{k_i}} \in X - N(\epsilon, A)$ , conjunto cerrado en  $X$  (pues es complemento de un abierto).

De aquí que  $x_0 \in X - N(\epsilon, A)$ , por lo que  $x_0 \notin A$ .

Ahora  $\lim_{i \rightarrow \infty} x_{n_{k_i}} = x_0$  y  $\{A_{n_{k_i}}\}_{i=1}^{\infty}$  es una subsucesión de  $\{A_n\}_{i=1}^{\infty}$ .

Entonces por la caracterización de  $\limsup(A_n)$  vista anteriormente se tiene que  $x_0 \in \limsup(A_n) = A$ . Por lo que  $x_0 \in A$ . Y tenemos una contradicción, por lo que (b) es verdadera.

Por lo que existe  $M_2 \in \mathbb{N}$  tal que  $A_n \subseteq N(\epsilon, A)$  para todo  $n \geq N$ .

Ahora con (a) y (b) probados, tomemos  $N = \max\{M_1, M_2\}$ , entonces:

$$A \subseteq N(\epsilon, A_n) \text{ y } A_n \subseteq N(\epsilon, A) \text{ para } n \geq N.$$

De aquí que  $H(A, A_n) < \epsilon$ , si  $n \geq N$ .

Por lo tanto  $\lim_{n \rightarrow \infty} A_n = A$  con la métrica de Hausdorff. □

#### 4. Completez de $2^X$

Un espacio  $X$  se dice ser completo si toda sucesión de Cauchy en  $X$  converge en  $X$ . En este apartado se mostrará la completez de  $2^X$ .

**Teorema 7.** Sea  $\{A_n\}_{n=1}^{\infty}$  una sucesión de Cauchy en  $2^X$ .

Entonces  $\{A_n\}_{n=1}^{\infty}$  converge con la métrica de Hausdorff a un  $A_0 \in 2^X$ .

*Demostración.* Por el teorema anterior:

El único candidato para  $\lim(A_n)$  cuando  $n \rightarrow \infty$  es  $A_0 = \limsup(A_n)$  que es un conjunto no vacío y cerrado de  $X$ . Y para ver la convergencia, igual por el teorema anterior, solo tenemos que ver que  $\limsup(A_n) \subset \liminf(A_n)$ .

Sea  $x \in A_0 = \limsup(A_n)$  y  $\epsilon > 0$ , entonces  $B_{\frac{\epsilon}{2}}(x) \cap A_n \neq \emptyset$  para una infinidad de  $n$ 's.

Por otra parte, como  $\{A_n\}_{n=1}^{\infty}$  es una sucesión de Cauchy, entonces existe  $N \in \mathbb{N}$  tal que

$$H(A_n, A_m) < \epsilon, \text{ para todo } n, m \geq N$$

En particular tomemos  $M_0 \in \mathbb{N}$ ,  $M_0 > N$ , entonces  $B_{\frac{\epsilon}{2}}(x) \cap A_{M_0} \neq \emptyset$  y además  $H(A_{M_0}, A_n) < \epsilon/2$  pues ambos índices  $M_0$  y  $n$  son mayores o iguales a  $N$ . Lo cual implica que  $A_{M_0} \subset N(\epsilon/2, A_n)$ .

Sea  $y \in B_{\frac{\epsilon}{2}}(x) \cap A_{M_0}$ , entonces  $y \in A_{M_0} \subset N(\epsilon/2, A_n)$ . Así existe  $z_n \in A_n$  tal que  $d(y, z_n) < \epsilon/2$  y por tanto  $d(x, z_n) \leq d(x, y) + d(y, z_n) < \epsilon/2 + \epsilon/2 = \epsilon$ .

Con esto probamos que  $B_\epsilon(x) \cap A_n \neq \emptyset$  para toda  $n \geq N$ . Por lo que  $x \in \liminf(A_n)$  y concluimos que toda sucesión de Cauchy converge en el espacio. Por lo tanto  $2^X$  es completo.  $\square$

## 5. Compacidad de $2^X$

Ahora estamos en condiciones de probar la compacidad de  $2^X$ .

**Teorema 8.**  $2^X$  es compacto.

*Demostración.* Usando el resultado anterior, es suficiente probar que toda sucesión tiene una sub-sucesión de Cauchy. La cual construiremos inductivamente.

*Afirmación.* Dado  $\epsilon > 0$  y un subconjunto infinito  $J$  de  $\mathbb{N}$ , existe otro subconjunto infinito  $J_1$  de  $J$  tal que  $H(A_n, A_r) \leq \epsilon$ , para todo  $n, r$  elemento de  $J_1$ .

Sean  $\epsilon > 0$  y  $J$  un subconjunto infinito de  $\mathbb{N}$ . Como  $X$  es compacto (pues  $X$  es continuo), existen  $m \in \mathbb{N}$  y  $x_1, x_2, \dots, x_m \in X$  tales que  $X = \bigcup_{i=1}^m B_{\frac{\epsilon}{2}}(x_i)$ .

Ahora para cada  $n \in J$  definimos  $K_n = \{i \in \{1, \dots, m\} : A_n \cap B_{\frac{\epsilon}{2}}(x_i) \neq \emptyset\}$ , observamos que  $K_n$  está bien definidos para toda  $n \in \mathbb{N}$ .

Sea  $\Lambda = \{F : F \subseteq \{1, 2, \dots, m\}\}$ , notemos que  $\Lambda$  tiene cardinalidad finita ( $2^m$ ).

Definamos la función  $f : J \rightarrow \Lambda$  tal que  $f(n) = K_n$ . Entonces  $J = \bigcup \{f^{-1}(K) \subset J : K \in \Lambda\}$ .

Como  $J$  es infinito y  $\Lambda$  es finito, entonces algún conjunto  $f^{-1}(K)$  debe ser infinito.

Definamos  $J_1 = f^{-1}(K)$ . Es claro que  $J_1 \subset J$  y es infinito. Falta ver que  $H(A_n, A_m) \leq \epsilon$ , si  $n, m$  es elemento de  $J_1$ .

Sea  $G = \{x_i : i \in K\} \in 2^X$  (es cerrado y no vacío). Veamos que dada  $n \in J_1$ , tenemos  $H(A_n, G) < \epsilon/2$ . Sea pues  $n \in J_1$ , entonces  $f(n) = K_n = K$ .

Así  $K_n = \{i \in \{1, \dots, m\} : A_n \cap B_{\frac{\epsilon}{2}}(x_i) \neq \emptyset\} = K$ .

De modo que  $G = \{x_i \in X : A_n \cap B_{\frac{\epsilon}{2}}(x_i) \neq \emptyset\}$ .

Si  $x_i \in G$  entonces  $A_n \cap B_{\frac{\epsilon}{2}}(x_i) \neq \emptyset$ . De manera que existe  $a \in A_n$  tal que  $d(a, x_i) < \epsilon/2$ . Así que  $x_i \in N(\epsilon/2, A_n)$ . Esto prueba que  $G \subseteq N(\epsilon/2, A_n)$ .

Por otro lado tomemos  $x \in A_n$ . Entonces  $x \in X = \bigcup_{i=1}^m B_{\frac{\epsilon}{2}}(x_i)$ , así que existe  $i_0 \in \{1, \dots, m\}$  tal que  $x \in B_{\frac{\epsilon}{2}}(x_{i_0})$ . De manera que  $A_n \cap B_{\frac{\epsilon}{2}}(x_{i_0}) \neq \emptyset$  y  $d(x, x_{i_0}) < \epsilon/2$ . Con esto  $A_n \subseteq N(\epsilon/2, G)$ .

Podemos asegurar que  $H(A_n, G) < \epsilon/2$ , para toda  $n \in J_1$  y como  $G$  es fijo, tenemos que  $H(A_n, A_r) < \epsilon$ , para  $n, r \in J_1$ .

Probada la afirmación. Ahora procedemos a construir la subsucesión de Cauchy, repitiendo el proceso que se hizo anteriormente, obteniendo una sucesión  $\{J_n\}_{n=1}^{\infty}$  de subconjuntos de  $\mathbb{N}$  infinitos, tal que  $J_1 \supseteq J_2, \supseteq \dots$

Y además, para todo  $k \in \mathbb{N}, H(A_n, A_r) \leq 1/k$  para todo  $n, r \in J_k$ .

Construcción de la subsucesión:

Tomemos  $n_1 \in J_1$ . Como  $J_2$  es infinito, existe  $n_2 \in J_2$  de manera que  $n_1 < n_2$ . Como  $J_3$  es infinito, existe  $n_3 \in J_3$  de manera que  $n_2 < n_3$ , etc.

Por lo tanto, existe  $n_1, n_2, \dots$  tal que  $n_i \in J_1$  y  $n_i < n_{i+1}$  para  $i \in \mathbb{N}$ .

Sea  $\epsilon > 0$ . Por la propiedad arquimediana existe  $k \in \mathbb{N}$  tal que  $\frac{1}{k} < \epsilon$ .

Luego afirmamos que si  $r, s \geq K$ , entonces  $H(A_{n_r}, A_{n_s}) < \epsilon$ .

En efecto, sea  $r, s \geq K$ , entonces  $J_s, J_r \subseteq J_K$ , de manera que  $n_s \in J_s, n_r \in J_r$  implican que  $n_s, n_r \in J_K$ , y por la elección de  $J_K$ , tenemos que  $H(A_{n_s}, A_{n_r}) < 1/K < \epsilon$ . Con esto probamos y terminamos la prueba de que  $2^X$  es compacto. □

## 6. Conclusiones

Con esto, tenemos probada una de las propiedades más deseables en un conjunto por su utilidad e importancia en diversos resultados, tanto del análisis como de la topología, la compacidad. Además, agregando que el hiperespacio  $2^X$  es conexo, tenemos que el hiperespacio  $2^X$  es continuo, pues ya es metrizable, no vacío y compacto. A partir de esto, se le puede aplicar a  $2^X$  toda la teoría de continuos conocida.

## 7. Referencias

- [1] Escobedo R., Macías S. y Méndez H. (2006). *Invitación a la teoría de los continuos y sus hiperespacios*. México, D.F. Aportaciones Matemáticas SMM.
- [2] Illanes M., A. (2004). *Hiperespacios de Continuos*. México, D.F. Aportaciones Matemáticas SMM.
- [3] Sam B. Nadel, Jr.(2006). *Hyperspaces of sets. A text with research questions*. México, D.F. Aportaciones Matemáticas SMM.
- [4] Munkres, J. (2002). *Topología*. Ed. Prentice Hall.
- [5] Aguilera M., M. y García V., L. (2003). *Hiperespacios de Continuos Métricos y No Métricos*. Tesis.
- [6] Illanes M., A. (1990). *Notas de clase de Hiperespacios de Continuos*.

Memorias de la XIX Semana  
Regional de Investigación y  
Docencia en Matemáticas

Mosaicos Matemáticos No. 32

Se terminó de imprimir en el mes de  
noviembre de 2009 en los Talleres  
Gráficos de la Universidad de Sonora y  
consta de 250 ejemplares

