Memorias de las Grandes Semanas Nacionales de la Matemática

Facultad de Ciencias Físico Matemáticas

Benemérita Universidad Autónoma de Puebla

EDITORES:
David Herrera Carrasco,
Fernando Macías Romero

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA
Enrique Agüera Ibáñez
*R*ector
José Ramón Eguíbar Cuenca
*S*ecretario General
Pedro Hugo Hernández Tejeda
*V*icerrector de Investigación y Estudios de Posgrado
Lilia Cedillo Ramírez
*V*icerrectora de Extensión y Difusión de la Cultura
Cupatitzio Ramírez Romero
*D*irector de la Facultad de Ciencias Físico Matemáticas
Carlos Contreras Cruz
*D*irector Editorial

Primera edición, 2009
ISBN:

Dirección de Fomento Editorial
2 Norte 1404, C. P. 72000
Puebla, Pue.
Teléfono y fax: 01 222 246 8559

Hecho en México
Made in Mexico

Las GRANDES SEMANAS NACIONALES DE LA MATEMÁTICA (GSNM) se realizan año con año entre los meses de agosto y septiembre desde el 2005.

Editores:
Fernando Macías Romero, David Herrera Carrasco.

Contenido

Presentación

Las GRANDES SEMANAS NACIONALES DE LA MATEMÁTICA se realizan cada año entre los meses de agosto y septiembre desde el 2005.
Éstas son un ejercicio anual del colectivo matemático nacional que, organizado bajo la dirección de la Academia de Matemáticas de la Facultad de Ciencias Físico Matemáticas, presenta una gran gama de actividades que forman parte del quehacer matemático.
En este volumen se recogen las memorias de la Primera a la Cuarta Gran Semana Nacional de la Matemática, agrupando los trabajos por orden alfabético en los apellidos de los autores.
Expresamos nuestro más sincero agradecimiento a todas las personas que hicieron posible la publicación de estas memorias, especialmente a Miguel Ángel García Ariza por su cuidadosa entrega en la edición.

Piedra de Sol, una experiencia sonora de divulgación de la ciencia.

Angoa Amador, José Juan, Aguilar Garduño, José Pablo. Aguilar Garduño, Mercedes
jangoa@fcfm.buap.mx, pabloaguilarg@att.net.mx, addhema@fcfm.buap.mx.

Piedra de sol es un poema. Piedra de sol es un concepto. Piedra de sol es una producción de audio resultante del proyecto "Divulgación de la ciencia a través de la radio" patrocinado por la Facultad de Ciencias Físico Matemáticas de la BUAP.

*

Ante todo, se trata de un proyecto serio y bien estructurado en el que se ha invertido mucho tiempo y un gran esfuerzo. Como todo trabajo creativo ha sido ha sido absorbente y placentero, sin embargo, no fue hecho para la simple diversión de un grupo de amigos sino que fue creado a partir objetivos claros y convicciones muy precisas.

El proyecto está integrado por varios elementos centrales, que han sido objeto de una constante revisión a lo largo de todo el proceso. El primero de ellos es la divulgación de la ciencia. Consideramos que junto con la enseñanza y la investigación es una de las labores medulares de la Facultad. Ésta, al ser tan rica en conocimientos ya que los crea y los transmite, tiene como siguiente paso natural el compartirlo para interesar a la comunidad en temas relativos a la ciencia o enriquecer su información científica. También consideramos que la divulgación de la ciencia es una forma de retribución a la comunidad ya que los impuestos que los ciudadanos pagan constituyen el subsidio que las universidades reciben.

En resumen, consideramos que la divulgación de la ciencia es una tarea fundamental y un compromiso social de la institución. Debido a ello, tenemos el firme convencimiento de que este proyecto debe ser auspiciado, protegido y amparado por la Facultad.

*

El poema

En 1957, Octavio Paz, premio nobel de literatura, escribió Piedra de sol. Pere Gimferrer en su libro "Octavio Paz. Prueba del nueve", declara: Piedra de sol es "… un itinerario… un recorrido hacia la fijación del instante… el instante de la plenitud liberadora y reveladora… tras haber comulgado con la naturaleza del instante nos arrebata de la firmeza y nos devuelve al fluir temporal".

El concepto

Ante todo habíamos decidido crear un concepto original y atractivo, con riqueza de contenido y amplio potencial de difusión. Para esto necesitábamos, ante todo, determinar un objetivo realista y asequible, específico para el proyecto y a continuación, el medio, el género, y el formato.

Después de muchas reuniones de trabajo nuestro objetivo quedó planteado de la siguiente forma: conseguir que el público se interese por la ciencia, inquietarlo, incitarlo a iniciar una búsqueda de conocimientos relativos a la ciencia. El conocimiento como producción humana tiene un valor incalculable y es patrimonio de todos, además, su búsqueda puede resultar altamente placentera.

El medio

En el país existen importantes instituciones que hacen divulgación de la ciencia a través de distintos medios como la operación de museos y la edición de libros y revistas. Otras, emplean los medios de comunicación masiva como los periódicos, la radio y la televisión.

En la radio encontramos el medio más idóneo para transmitir nuestros mensajes, ya que tiene como características esenciales:

- ✓ La magia. Estimula la imaginación para que las personas recreen en su mente las historias que la palabra, música, los efectos sonoros y silencios les presentan.
- ✓ Tiene un largo alcance, es posible llegar a los habitantes de las grandes ciudades y a pequeñas comunidades.
- ✓ Es un medio sólido y maduro que mueve grandes volúmenes de audiencia.
- ✓ Tiene una gran versatilidad por sus variados géneros y formatos.
- ✓ No requiere un equipo especial. Ni es necesario trasladarse a otros lugares. Es muy fácil adquirir un aparato de radio ya que está al alcance de todos.
- ✓ Sirve de compañía, descanso y diversión. Además, cuando las personas escuchan la radio, pueden a la vez, realizar otras actividades, como lavar ropa, preparar la comida, trabajar en el taller, etc. No requiere de toda su atención ni interfiere con sus actividades cotidianas.

Las decisiones que tuvimos que tomar fueron muchas, y pasamos largas horas considerando las posibilidades, ¿programas en vivo o grabados? ¿Revista radiofónica o informativo?, ¿radiodrama o radioarte?, ¿radio musical? ¿programa pora niños?

Nuestra elección se centró en el radiodrama con la opción de explorar el radioarte. En cuanto al formato preferimos la cápsula por las ventajas técnicas que nos presentaba.

Por su complejidad, la producción de audio en los géneros de radiodrama y radioarte resulta un gran desafío ya que el construir historias a través de voces, música, sonidos y silencio requiere no sólo de habilidades técnicas sino también de una sensibilidad tal que permita dirigirse al público de una forma directa, franca y amistosa.

A estas alturas nuestro panorama era muy claro: no queríamos presentar rollos aburridos, ni a petulantes académicos dirigiéndose a un público ignorante. Lo que queríamos era algo muy sencillo: hablar al público como un amigo que le habla a otro: "mira, esto es interesante, escúchalo" y todo ello con una producción de vanguardia.

La labor de divulgación de la ciencia por su naturaleza deber ser multidisciplinaria y para estar a la altura de nuestras expectativas, decidimos formar un equipo con especialistas en diversas áreas científicas y artísticas: enseñanza e investigación, producción de audio, actuación, locución y música.

El proceso.
Aunque para nosotros era muy claro que escribir textos de divulgación es una tarea ardua y que requiere de habilidades especiales, se nos ocurrió que resultaría muy atractivo tomar como punto de partida material escrito por integrantes de la Facultad. Así nos dimos a la tarea de invitar a profesores y estudiantes interesados en la divulgación de la ciencia y que tenía trabajos publicados.

Isaí Moreno Roque (doctorado en matemáticas y premio a primera novela Alfaguara) y Esaú Percino Zacarías (físico y con maestría en literatura mexicana) nos brindaron un material de base fantástico. Varios profesores que se interesaron en el proyecto decidieron escribir textos específicamente para la serie como Juan Angoa, Agustín Contreras, Alberto Cordero y otros más, nos proporcionaron materiales, que aunque no habían sido escritos por ellos, resultaron de gran utilidad.

El siguiente paso fue la adaptación literaria y la elaboración de los guiones. Esta etapa resultó abierta y muy creativa ya que nos sentimos con toda la libertad para apropiarnos de las historias que como ya dijimos eran muy buenas.

Ya con los guiones elaborados pasamos al diseño de audio, la selección de la música y los efectos sonoros. Después: la selección de los actores, ensayo con los actores, concentración de elementos de producción, concertación de agendas de grabación y producción, grabación de música original, grabación con actores, musicalización, ambientación y efectos de sonido, mezcla y masterización.

La temática
Piedra de sol presenta una temática muy variada, centrada principalmente en el área de las ciencias físico matemáticas y afines. Cada pieza es una aventura

que no cansa ni pierde actualidad. Algunas pueden conducirnos a lugares insospechados como a la comarca de los números reales a buscar al señor cero y al señor uno para encontrar una explicación a la ley de los signos; o a un mercado popular, en el que un perro llamada Pitágoras sorprende al público con sus notables habilidades mientras un pordiosero canta un blues. También y de manera sigilosa se atreve a atisbar detrás de los lujosos cortinajes de la alcoba de Scherezada, e incluso se remonta al año 2021 y nos cuenta como fue que unos traviesos niños encontraron un objeto muy extraño: ¡un libro!

Aunque algunas de las piezas tienen un toque de humor como el discurso a los graduados de Woody Allen; otras, presentan la crudeza de la irracionalidad humana como veredicto que la santa inquisición impone a Galileo o la historia de Hipasso de Metaponto.

Los personajes que aparecen en las piezas son tan diversos como un celebre científico que habla de su viaje a Pondicherry para caer en las trampas de Venus o un ingenuo beagle llamado Plusminus que cuenta como fue sometido a un estudio radiográfico por su célebre amo Roentgen.

Piedra de sol no resiste la tentación de explorar textos poéticos como el manual para sabios que presenta una visión de los conceptos fundamentales de Aristóteles, Galileo y Einstein o los encantos del péndulo que pueden verse en el balanceo de los monos, la cuerda del ahorcado o el espacio sideral.

En esencia, de Piedra de sol presenta una exploración de lo humano desenvolviéndose en ámbitos cercanos a la ciencia.

El producto
Como resultado tenemos un producto cultural formado por 30 piezas de audio que presentan historias, conceptos, textos poéticos y recreaciones:
1ª temporada, 14 programas, año 1998
2ª temporada, 14 programas, año 2000 aniversario L de la Facultad
3ª temporada, 2 programas, año 2001.

La difusión.
Piedra de sol, por su línea temática, riqueza de contenido y adaptación al lenguaje radiofónico tiene un alto potencial de difusión por lo que ha sido transmitida por diversas radiodifusoras culturales y comerciales. Ha llevado el nombre de la Facultad a lugares tan variados como la sierra norte de Puebla o Malargüe en Mendoza, Argentina. También ha cruzado el mar hasta Cataluña, España. Y aquí en la ciudad ha visitado Radio BUAP, la Radiante, la HR y Sicom. Y en el interior del país, Poza Rica, Veracruz y Tlaxcala, Tlax.

En los años 1998 y 2000 participó en la Bienal Latinoamericana de Radio. Esta actividad, además de premiar las mejores producciones radiofónicas de América Latina, España y Portugal es un lugar de encuentro que busca la actualización y el intercambio de conocimientos sobre las posibilidades creativas de la radio.

El material fue diseñado para la radio, pero congruente con el espíritu de servicio a la comunidad se ha conservado disponible para toda institución interesada, de tal modo que dos centros escolares del interior del estado, ubicados en Tlachichuca y San Miguel Canoa, han estado aplicando con éxito los programas en el salón de clase como elementos para motivar un mayor interés en la ciencia en sus alumnos de bachillerato.

Una gran dificultad

El éxito de una serie radiofónica, su aceptación y posicionamiento, depende de varios factores, entre ellos, su permanencia en el aire por periodos definidos. La continuidad determina su presencia en la mente del auditorio. Aquí es donde radica nuestra principal dificultad, el patrocinio cubrió sólo 30 programas.

Lo que queremos

a. Lograr que nuestro proyecto "Divulgación de la ciencia a través de la radio" sea auspiciado, protegido y amparado por la Facultad de Ciencias Físico Matemáticas de la BUAP.
b. Buscar la consolidación del proyecto y el equipo de trabajo.
c. Ampliar y mejorar el proyecto, experimentar más con el radioarte y a futuro, incursionar en otros géneros.
d. Establecer la continuidad, aumentar la cobertura y contactos con otras instituciones y radiodifusoras del planeta.
e. Tener presencia en internet.
f. Avanzar en cuestiones de evaluación y estudios de impacto.

Conclusiones

La labor de la divulgación de la ciencia, en lo general, es amplia, compleja y bastante ambiciosa. Poco a poco y con gran esfuerzo ha ido ganando terreno y cada vez son más las instituciones que reconocen su importancia y le brindan los espacios y los apoyos necesarios. Esperamos que los impulsores y creadores del conocimiento científico, muy pronto comprendan que hacer divulgación de la ciencia es parte de su labor y también de su compromiso social. Y aunque comprendemos que les resulta sumamente difícil escribir textos accesibles para un público amplio es un reto que deben tomar.

Convencidos de la validez de este proyecto y de que la divulgación de la ciencia es una aventura digna de vivirse, seguiremos con la difusión del material y explorando la alternativa de su empleo en el salón de clases.

Realización artística
J. Pablo Aguilar
Gisela Merchand

Coordinación y guiones
Mercedes Aguilar

Grabado y masterizado en
Idearte

Voces
Alan Arroyo
Néstor Vázquez
Pablo Aguilar
Mercedes Aguilar
Nuria Castells
Diego Rosas
Xóchitl Herena
Claudia Huerta.
Juan Angoa
Javier Albores
Humberto Moreno
Jorge Merchand
Blanca Hernández
Magali Lizaola
Flori Galván
Catalina Aguilar
Benito Lavalle
Brenda Morales
Francisco Javier Ruiz

Temas:
Los encantos del péndulo
Las trampas de Venus
El cielo como infierno
La matemática, eterno debate
El jabón en el baño y en las ciencias
Principios básicos del caminar
La aplicación de la reología es tan directa en el arte culinario
Manual para sabios
Resonancia en los muros de Jericó
Entrevista con Charles
Lo natural es lo más antinatural
Hipasso de Metaponto
Discurso a los graduados
El ceremonial
Contando
El elegido de los dioses
Severino Boecio
Isidoro de Sevilla, el primer enciclopedista medieval.
El asistente Plusminus
Rayos cósmicos
Cómo se divertían
Flavio Casiodoro

La música, misterio supremo
Instante, el tiempo puro
Ruidos misteriosos
Números reales
Galileo
Historia del ajedrez
Las embalsamadoras
San Agustín

Realizado durante la gestión del Dr. Mario Alberto Maya Mendieta
(1997-2000)

Control Estabilizante del Péndulo con Rueda de Reacción

Miguel Alvarado Flores, V. Vasilievich Alexandrov, W. Fermín Guerrero Sánchez
Facultad de Ciencias Fisico-Matemáticas, BUAP
18 sur y Av. San Claudio, Colonia San Manuel, Ciudad Universitaria
alvaradone@hotmail.com, valex@fcfm.buap.mx, willi@fcfm.buap.mx

Resumen

Se presenta un trabajo de investigación experimental en el contexto del control de los sistemas mecánicos subactuados basados en la estructura lagrangiana. Implementamos un control por retroalimentación de estado aplicando el método de asignación de polos y linealizando el modelo no lineal por el método de *Aproximación Lineal*. Para estabilizar en tiempo real el punto de equilibrio inestable del Péndulo con Rueda de Reacción (PCRR).

Palabras clave: subactuado, linealización, control, estabilidad, saturado.
Modalidad: Cartel.

1. Introducción

El $Péndulo\ con\ Rueda\ de\ Reaccion\ \ (PCRR)$ es un sistema subactuado, fue introducido por M. Spong. $et\ al$ [1]. Este péndulo (Fig. 1) puede girar libremente sobre su pivote, dispone de una rueda simétrica colocada en el extremo cuyo eje de rotación es paralelo al eje del péndulo. Controlamos el péndulo por la acción del torque generado por el giro de la rueda que a su vez tiene acoplado al actuador. Las variables θ y θ_r son las posiciones angulares medidas con respecto a la vertical que define el péndulo en reposo, en sentido contrario a las manecillas del reloj. El sistema está caracterizado por los siguientes parámetros: l_1 es la distancia al centro de masa del péndulo m_1, l_2 es la distancia al centro de masa de la rueda m_2, J_1 es el momento de inercia medido desde el centro de masa del péndulo, J_2 es el momento de inercia de la rueda de reacción medido desde el centro de masa de la rueda, g es la aceleración de la gravedad. La naturaleza $subactuada$ del dispositivo se debe a que tenemos más grados de libertad (θ y θ_r), que actuadores (motor acoplado en la rueda).

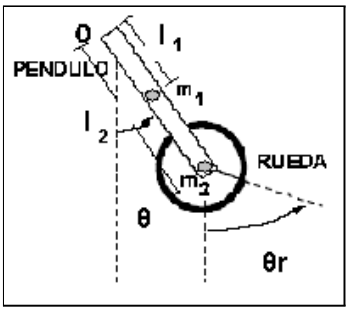


Fig. 1 Péndulo con Rueda de Reacción.

El PCRR está provisto con dos sensores ópticos (encoders) incrementales ubicados en el pivote del péndulo y en el eje de la rueda. Cuando el péndulo se desplaza partiendo de su estado de reposo hacia la derecha se mide un ángulo positivo y hacia la izquierda mide valores negativos. El origen o ángulo cero se determina con la primera medición que realiza el sensor óptico ubicado en el pivote del péndulo, por lo que es necesario que el PCRR esté en reposo al inicio de la operación.

2. Modelo Matemático

El primer paso en el diseño de un sistema de control es el desarrollo del modelo matemático del sistema que se va a controlar. Para el Péndulo con Rueda de Reacción se obtiene usando el formalismo de Euler-Lagrange, este modelo es no lineal. Calculamos la energía cinética por separado para el péndulo y la rueda. Para el péndulo la energía cinética es medida desde su centro de masa

$$T_1 = \frac{1}{2} m_1 v_1^2 + \frac{1}{2} J_1 \dot{\theta}^2,$$

$v_1 = l_1\dot{\theta}$ es la velocidad de desplazamiento de m_1, por lo tanto la energía cinética queda como

$$T_1 = \frac{1}{2}m_1 l_1^2 \dot{\theta^2} + \frac{1}{2}J_1\dot{\theta^2}.$$

Para la rueda, se calcula la energía cinética desde su centro de masa, la velocidad angular depende de la rotación de la rueda y de la rotación del péndulo, así que

$$T_2 = \frac{1}{2}m_2 l_2^2 \dot{\theta^2} + \frac{1}{2}J_2(\dot{\theta} + \dot{\theta_r})^2.$$

La energía cinética del PCRR está dada por

$$\begin{aligned} T &= T_1 + T_2 \qquad (1) \\ T &= \frac{1}{2}m_1 l_1^2 \dot{\theta^2} + \frac{1}{2}m_2 l_2^2 \dot{\theta^2} + \frac{1}{2}J_1\dot{\theta^2} + \frac{1}{2}J_2(\dot{\theta} + \dot{\theta_r})^2 \end{aligned}$$

renombrando $J = m_1 l_1^2 + m_2 l_2^2 + J_1$ obtenemos

$$T = \frac{1}{2}J\dot{\theta^2} + \frac{1}{2}J_2(\dot{\theta} + \dot{\theta_r})^2 \qquad (2)$$

Para el cálculo de la energía potencial consideremos que en el estado de reposo $U = 0$. Sea $M = m_1 + m_2$ ubicada en el centro de masa del eslabón y la rueda entonces de la definición de centro de masa $Ml = m_1 l_1 + m_2 l_2$. Donde l es la distancia del pivote O al centro de masa del péndulo y la rueda, h es la altura de M respecto de la posición de reposo.

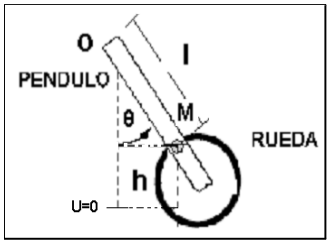


Fig. 2 Configuración para la energía potencial.

La energía potencial para el PCRR es $U = Mgh$ donde $h = l - l\cos\theta$ así que

$$U = Mgl(1 - \cos\theta) \qquad (3)$$

Finalmente el lagrangiano del PCRR es $L = T - U$ entonces

$$L = \frac{1}{2}J\dot{\theta^2} + \frac{1}{2}J_2(\dot{\theta} + \dot{\theta_r})^2 - Mgl(1 - \cos\theta) \qquad (4)$$

La posición angular de la rueda no aparece explícitamente en la ecuación (3), decimos entonces que es una variable cíclica. De las ecuaciones de Euler-Lagrange obtenemos las ecuaciones de movimiento,

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{\theta}}) - \frac{\partial L}{\partial \theta} = \tau_1, \qquad \frac{d}{dt}(\frac{\partial L}{\partial \dot{\theta_r}}) - \frac{\partial L}{\partial \theta_r} = \tau_2 \quad ,$$

tal que

$$\frac{\partial L}{\partial \dot{\theta}} = J\dot{\theta} + J_2(\dot{\theta} + \dot{\theta_r}); \qquad \frac{\partial L}{\partial \theta} = -Mgl(sen\theta).$$

$$\frac{\partial L}{\partial \dot{\theta_r}} = J_2(\dot{\theta} + \dot{\theta_r}); \qquad \frac{\partial L}{\partial \theta_r} = 0$$

Cuando θ_r=0 es decir que no hay acción del motor entonces no existe ninguna fuerza que mueva al PCRR por lo tanto $\tau_1 = 0$ obtenemos

$$J\ddot{\theta} + J_2(\ddot{\theta} + \ddot{\theta_r}) + Mglsen\theta = 0. \tag{5}$$

Si ahora $\theta = 0$ considerando que hay acción del motor obtenemos que para la rueda existe la acción del torque $\tau_2 = \tau$ por lo que obtenemos

$$J_2(\ddot{\theta} + \ddot{\theta_r}) = \tau \tag{6}$$

De la resta de (5) y (6) obtenemos

$$J\ddot{\theta} + Mglsen\theta = -\tau, \tag{7}$$

despejando $\ddot{\theta}$ de (5) y sustituyendo en (6) obtenemos

$$\frac{JJ_2}{J + J_2}\ddot{\theta_r} - \frac{J_2}{J + J_2}Mglsen\theta = \tau \tag{8}$$

Las ecuaciones (7) y (8) son las ecuaciones de movimiento del PCRR cuando el motor está accionado, despreciamos las fuerzas de fricción y la dinámica eléctrica del motor DC, entonces el torque está dado por:

$$\tau = kI, \tag{9}$$

donde k es la constante del motor, I es la corriente del motor, la ecuación de movimiento queda como

$$\begin{aligned} \ddot{\theta} + \frac{Mgl}{J}\sin\theta &= -I\frac{k}{J}, \\ \ddot{\theta}_r - \frac{Mgl}{J}\sin\theta &= kI\frac{J+J_2}{JJ_2} \end{aligned} \tag{10}$$

de aquí que los parámetros [9] del sistema son $a = \frac{Mgl}{J}$; $b = \frac{-k}{J}$; $br = k\left(\frac{J+J_2}{JJ_2}\right)$, donde

$$\begin{aligned} a &= \omega_p^2 = 78{,}4, \\ b &= 1{,}08, \\ b_r\square &= 198. \end{aligned} \tag{11}$$

3. Validación del Modelo Matemático

Para evaluar el modelo matemático realizamos un experimento sencillo en el kit mecatrónico y verificamos que la simulación de la ecuación (10) sea cualitativamente aproximada al resultado experimental. Tanto en la simulación como en el desarrollo del experimento se realizan en el ambiente de Matlab-Simulink. De la ecuación (10) y los parámetros (11)

$$\begin{aligned} \ddot{\theta} + 78{,}4\sin\theta &= -1{,}08u, \\ \ddot{\theta}_r - 78{,}4\sin\theta &= 198u, \end{aligned} \tag{12}$$

realizamos el cambio de variable $x_1 = \theta$, $x_2 = \dot{\theta}$, $x_3 = \dot{\theta}_r$ para llevar a (12) al modelo en el espacio de estados $\dot{x(t)} = f(x,u)$, tal que

$$\begin{aligned} \dot{x_1} &= x_2, \\ \dot{x_2} &= -1{,}08u - 78{,}4\sin x_1, \\ \dot{x_3} &= 198{,}0u + 78{,}4\sin x_1. \end{aligned} \tag{13}$$

Simulación. Realizamos Matlab-Simulink para realizar la simulación con la condición inicial $x_0 = [-1{,}7357, 0, 0]$. El tiempo de muestreo es de $0.005\ s$, se gráfican 2000 puntos que corresponden a un tiempo de 10 segundos.

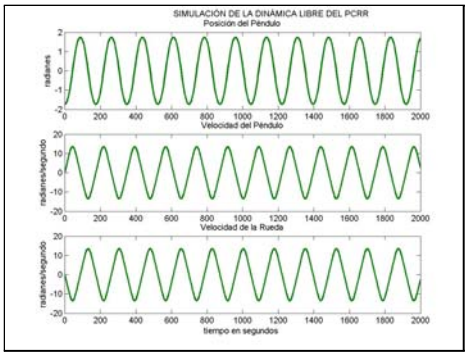


Fig. 3 Simulación del comportamiento libre del PCRR bajo una perturbación.

Arreglo experimental. Para el arreglo experimental utilizamos el Kit Mecatrónico [9] y ensamblamos el PCRR como se muestra en la Fig. (4). El Kit se conecta a través de la tarjeta de adquisición de datos $C6XDSK\_DIGIO$ a la PC a través del puerto paralelo con una interface diseñada en Matlab, el PWM envía la señal de control al motor. Los sensores ópticos registran variaciones en la posición del ángulo y la posición de la rueda. El tiempo de muestreo es de $\tau = 0.005\ s$.

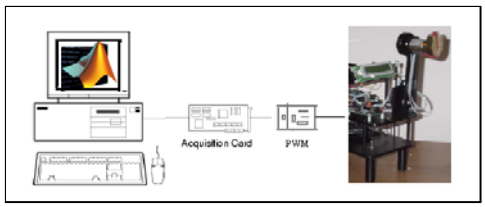


Fig. 4. Esquema del arreglo experimental.

El kit mecatrónico en su arreglo para el PCRR tiene dos sensores ópticos incrementales uno acoplado al pivote del péndulo y otro acoplado al eje del motor (actuador) que tiene acoplada a la rueda. Los sensores ópticos (encoders) envian la señales en tiempo real a través del *Real Time Work Shop* para estimar las variables del modelo matemático: *ángulo del péndulo, velocidad del péndulo, velocidad de la rueda*, la condición inicial experimental es $x_0 = [-1{,}7357, 0, 0]$. Comparando la Fig. (5) y la Fig. (6) cualitativamente podemos decir que el modelo matemático se aproxima al comportamiento real del PCRR.

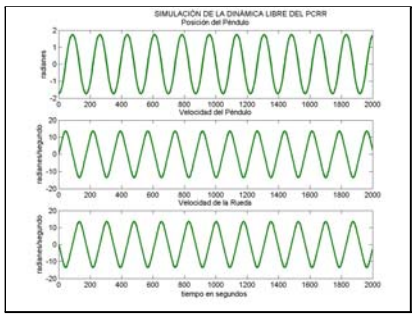


Fig. 5. Dinámica del PCRR bajo una perturbación.

4. Control del Péndulo con Rueda Inercial

El método de *Aproximación Lineal* también es conocido como el *Método de linealización por truncamiento de la serie de Taylor*, el control encontrado por este método es una ley de control por retroalimentación de estado ya que la ley de control depende del estado actual del sistema,

$$A = \frac{\partial f}{\partial x}(x,u)\mid_{(x=0,u=0)}, \quad B = \frac{\partial f}{\partial u}(x,u)\mid_{(x=0,u=0)}. \tag{14}$$

En el espacio de estados la ecuación

$$\dot{x(t)} = Ax(t) + Bu \tag{15}$$

$x(t)$ es el vector en el espacio de estados A es una matriz de $n \times m$ y B es una matriz de $n \times p$, u es la entrada de control. Si el sistema es estable los eigenvalores de A tienen parte real negativa en ese caso decimos que la matriz es Hurtwitz [2], si la matriz no es Hurtwitz y el sistema es controlable, entonces existe una ley de control $u = -Kx$, tal que

$$\dot{x(t)} = (A - BK)x, \tag{16}$$

es estable, es decir la matriz $(A - BK)$ tiene eigenvalores con parte real negativa.

Punto de equilibrio inestable $(\theta = \pi)$. Calculamos las ganancias utilizando el método de asignación de polos [2] para los eigenvalores $\lambda_1 = -5$, $\lambda_2 = -6$, $\lambda_3 = -7$.

$$K = [-171{,}9328 - 19{,}1907 - 0{,}0136] \tag{17}$$

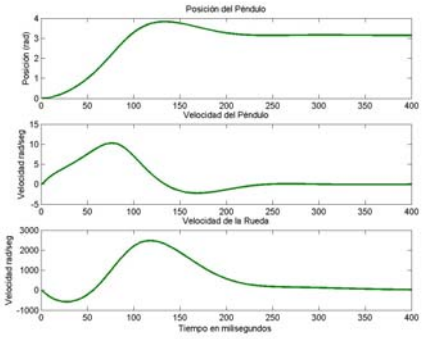


Fig. 6. Simulación de la estabilización en el punto de equilibrio inestable.

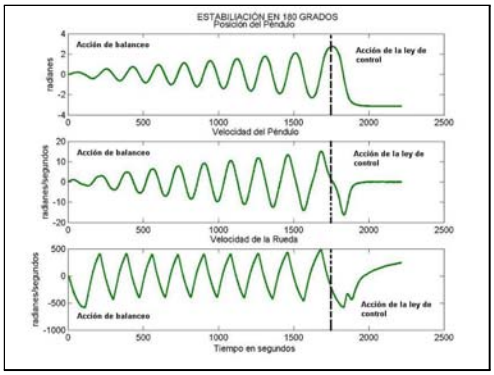


Fig. 7. Estabilización en 180 grados

Cuando realizamos la linealización por Aproximación Lineal la ley de control diseñada funciona para estabilizar en el origen del sistema. Por lo que para estabilizar en $x_1 = \pi$ tenemos que llevar al PCRR a una región cercana a este punto de equilibrio. Para llegar a esta región de Lyapunov utilizamos una estrategia de balanceo diseñada por $Spong\ et\ al$ $[10]$. En la figura (7), la parte transitoria corresponde a la acción de balanceo y la última parte corresponde a la acción de la ley de control propuesta por el método de aproximación. La gráfica muestra estabilidad en $x_1 = -\pi$, debido a que el encoder es incremental.

El sistema muestra una respuesta transitoria hasta que tiene energía suficiente para acercarse a la región de atracción donde es válida la ley de estabilidad encontrada. Esta respuesta transitoria puede disminuirse al utilizar técnicas más completas como la de pasividad o técnicas de saturación. Una vez que el sistema se encuentra cerca de la region de atracción x_2 (velocidad del péndulo) sera cero, x_3 (velocidad de la rueda) permanece constante, esto es debido a que el sistema compensa la acción de fuerzas externas en este caso debida al cableado del motor. En la ley de control calculada no está considerando las limitaciones físicas del actuador, para que la acción del control no fuerce al sistema cuando se estabiliza en el punto de equilibrio inestable, es conveniente aplicar una función de saturación [12], lo que de paso atenuará la respuesta transitoria. Sea la función de saturación

$$Sat[u(x)] = \frac{au(x)}{\sqrt[2]{a^2 + u^2(x)}}, \tag{18}$$

donde $u(x)$ es el control calculado por Taylor, tal que el sistema controlable $\dot{x} = f(x, u)$, cambia a la forma

$$\dot{x} = f\left(x, \frac{au(x)}{\sqrt[2]{a^2 + u^2(x)}}\right) \tag{19}$$

Implementamos la función de saturación para simular el comportamiento del sistema

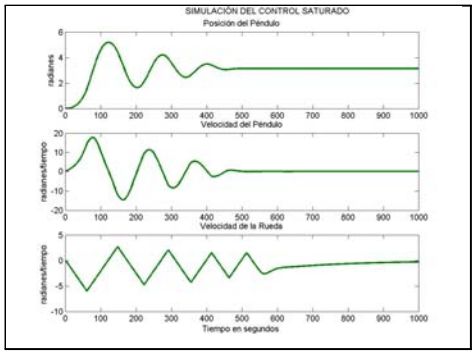


Fig. 8. Simulación del control saturado

La figura (8) nos muestra que la acción del control saturado limita la acción del actuador de tal manera que su influencia será más suave, esto aumentará la respuesta transitoria.

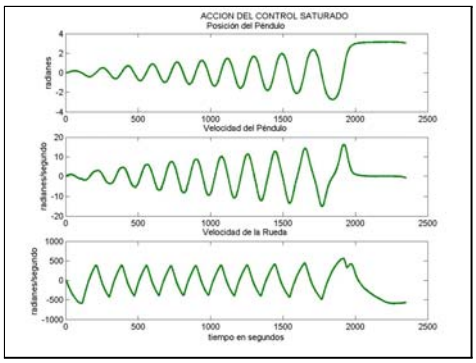


Fig. 9. Control Saturado

La función de saturación tarda más tiempo en estabilizar sin embargo es más estable ante perturbaciones externas.

5. Conclusiones

En este trabajo abordamos el problema de la estabilización en los puntos de equilibrio del Péndulo con Rueda de Reacción y la evaluación experimental del la ley de control estabilizante. En la Teoría de Control encontrar una ley estabilizante es una tarea fundamental. La implementación de leyes estabilizantes en un sistema dinámico implica modificar la configuración de cuerpo libre o pasiva a una configuración retroalimentada o activa. El PCCR es un sistema no lineal, para estos sistemas no hay una metodología que nos indique comó obtener la mejor ley de control, sino que ajustamos cada técnica según el sistema de estudio

y verificamos la validez de la ley encontrada, simulando el comportamiento en lazo cerrado y lo que es más importante evaluando experimentalmente. La literatura de los sistemas no lineales sugiere como primer paso analizar el comportamiento del sistema en una vecindad donde el comportamiento es semejante al de un sistema lineal, con la implementación del control por el método de *Aproximación Lineal* obtuvimos resultados satisfactorios a pesar de que el control no es el óptimo, esto nos permite en el futuro aplicar técnicas y metodologías más robustas para la estabilización.

La estabilización de los *Puntos de Equilibrio* muestra en términos prácticos un algoritmo destinado al rechazo de perturbaciones, útil en la estabilización de antenas móviles o en aplicaciones biomecánicas, aplicación se da cuando el sistema es un robot planar [7] este algoritmo puede utilizarse como respuesta a una falla mecánica.

Referencias

[1] V.V Alexandrov, S.I. Zlochevskii, S.S. Lemak, N.A. Parushnikov (Rusia), R. Reyes Sánchez, H. Salazar Ibargüen, I. Romero Medina (Mexico) *Introducción a la Modelación Matemática de Sistemas Controlables*, tomo I. Coordinado por V.V Alexandrov. Edición dedicaca al quincuagésimo aniversario de la FCFM-BUAP edit. UNIVERSIDAD AUTONOMA DE PUEBLA 2000.

[2] Katsuhiko Ogata. *Ingenieria de Control Moderna*. Prentice Hall 1998

[3] Sergio Dominguez, Pascual Campoy, Jose María Sebastián, Agustin Jiménez. *Control en el Espacio de Estados*. Sergio Dominguez et al. Edit Prentice Hall 2002

[4] Luis Manuel Hernández Gallardo. *Sobre el Metodo de Lyapunov*. http://valle.fciencias.unam.mx/~luism/smlyapunov.html

[5] Germund Dahlquist. *Numerical Methods*. Prentice Hall 1974

[6] Hassan Khalil. *Nonlinear systems*. Prentice Hall. 1996

[7] Mark W. Spong,M. Vidyasagar. *Robot Dynamics And Control*.John Wiley Sons. 1989

[8] Trinks, W. *Governors and the governing of Prime Movers.* Van Nostrand, Princeton N.J., 1919

[9] Mechatronics Control Kit *User's Manual*. Quanser Consulting Inc. 2001.

[10] M.W. Spong, P. Corke, and R. Lozano. *"Nonlinear control of the Inertia Wheel Pendulum"*. Automatica, submitted, September 1999.

[11] http://www.control-systems.net

[12] Landau. L.D. *Mecánica*. Editorial Reverté 1975

[13] Thomas L. Vincent, Walter J. Grantham. *Nonlinear and Optimal Control Systems*. Wiley-Interscience Publication. 1997

Aplicación del Teorema de Tikhonov para Simplificar el Modelo Matemático de un Sistema Dinámico Controlable

Guerrero Sánchez W. Fermín, Alexandrov Vladimir
Facultad de Ciencias Físico-Matemáticas BUAP
18 Sur y Av. San Claudio CU. CP 72570 Puebla Pue.
willi@fcfm.buap.mx-valex@fcfm.buap.mx

Resumen

El teorema de Tikhonov es un resultado fundamental de la teoría de perturbaciones singulares y proporciona un metodo de aproximación llamado método asintótico, donde el objetivo consiste en obtener fórmulas que describan cualitativamente el comportamiento en algún intervalo de los valores de las variables independientes. La precisión de tales fórmulas poseen limitaciones naturales. El modelo de perturbación singular de un sistema dinámico de dimensión finita fue extensamente estudiado por Tikhonov (1948-1952) Levinson (1950) Vasil'eva, en este trabajo el Teorema de Tikhonov será usado para implementar una metodología para simplificar el modelo matemático de un sistema dinámico controlable

Palabras claves: Parámetro pequeño, perturbación singular estándar, métodos asintóticos, control optimal, tiempos lentos y rápidos.

1. Introducción

La necesidad de desarrollar métodos para obtener soluciones aproximadas de las ecuaciones diferenciales a llevado a proponer metodos numéricos, asi cómo otra clase de métodos los llamados asintótivos, donde el objetivo principal es obtener fórmulas que describan el comportamiento cualitativo en algún intervalo de valores de la variable independiente, debe tenerse en cuenta que los metodos numéricos y los asintótivos no son excluyentes, se puede decir que se complementan. En una amplia variedad de aplicaciones prácticas los sistemas no-lineales en variables de estado envuelven diferentes escalas de tiempo i.e. escalas de tiempo lentas y rápidas. Este fenomeno es causado por una variedad de diferentes factores, el más común es la existencia de elementos pequeños

en los sistemas dinámicos. En este trabajo se presenta el llamado modelo de perturbación singular estándar el cual con la ayuda del Teorema de Tikhonov nos ayudará a simplificar y resolver una ecuación diferencial no lineal de manera aproximada

$$\begin{aligned} \dot{y} &= f(y, z, \mu, t) \\ \mu\dot{z} &= g(y, z, \mu, t) \end{aligned} \tag{1}$$

cuando se sustituye $\mu = 0$ en la segunda ecuación causa cambios fundamentales y abruptos en las propiedades dinámicas del sistema cómo en la ecuación diferencial $\mu\dot{z} = g$ que se degenera en una ecuación algebraica o trascendental

$$0 = g(y, z, \mu, t)$$

Se puede decir que la esencia de está teoría radica o se encuentra en la discontinuidad de la solución causada por la perturbación singular que puede ser evitada si se analiza con diferentes escalas de tiempo. Este enfoque de escala multitiempo es una característica fundamental del método de perturbaciones singulares [3],[5],[6],[1].

2. Normalización

En general, el proceso de normalización consiste en adimensionalizar las ecuaciones de movimiento de un sistema dinámico para facilitar la identificación de los parámetros pequeños.

Considere un sistema dinámico arbitrario, la ecuación de movimiento escrita en la forma de Cauchy es:

$$\begin{aligned} \frac{dX_1}{dt} &= F_1(X_1, X_2, ...T, A_1, A_2, ..B_1, B_2, ...) \\ \frac{dX_2}{dt} &= F_2(X_1, X_2, ...T, A_1, A_2, ..B_1, B_2, ...) \\ .. &= .... \end{aligned} \tag{2}$$

donde $X_1, X_2, ...$ son las variables de fase del problema y $A_1, A_2, .. B_1, B_2, ...$ son grupos de coeficientes con las mismas dimensiones.

La normalización de las ecuaciones (2) es satisfecha mediante algúnos pasos, la sucesión de los cuales puede ser diferente.

1 Rescribir el sistema en términos de las medidas numéricas que corresponden a todas las cantidades involucradas.

$$T = T_* t, \;\; X = X_* x, ..., A_1 = A_* a_1, ..., B_1 = B_* b, ... \tag{3}$$

2 Esquematizar la clase de movimientos para la cual será considerado el sistema (2).

Escoger $T_*, X_{1*}, ..., A_*, B_*, ...$ en (3) cómo iguales a algúnos valores que son característicos para las cantidades correspondientes en está clase de movimientos.

El valor característico para el tiempo T_* es principalmente determinado por los objetivos de la investigación, por el intervalo de tiempo en que la conducta del sistema es de interés para un investigador; o en el cual las variables del problema alcanzan sus valores límite; o por la razón de obtener a las ecuaciones en una forma más compacta; y así sucesivamente. Una elección apropiada de T_* en la clase de movimiento usualmente proporciona una condición:

$$T \leq T_* \tag{4}$$

Los valores característicos de las variables de fase son determinados por sus valores absolutos máximos en el intervalo de tiempo dado por (3). Así,

$$X_{1_*} = \max |X_1|, X_{1_*} = \max |X_1|, ... \tag{5}$$

Por analogia con (2.4) los valores característicos de los coeficientes son supuestos cómo iguales a los valores absolutos máximos en cada grupo de coeficientes.

$$A_* = \max_k \{|A_k|\}, \quad B_* = \max_h |B_h|, ... \tag{6}$$

Usar los valores característicos de (4)-(6) cómo unidades en (3) determina un sistema de unidades que es especifico para una clase dada de movimiento.

En este sistema de unidades los valores absolutos de las variables adimensionales $t, x_1, x_2, ...$ serán variados sobre intervalos del orden de la unidad de acuerdo con (4) y (5); y debido a (6), los valores absolutos de los coeficientes a_k, b_h no serán más grandes que los valores del orden de la unidad.

3 Dividamos cada ecuación de (2) transformado en concordancia con (3) y (6) por una combinacion de multiplicadores dimensionales $T_*, X_1, ...$, la cual tiene las mismas dimensiones que la ecuacion. Entonces (2) toma la forma:

$$\frac{T_1}{T_*}\frac{dx_1}{dt} = f_1(x_1, x_2, ..., t, \Delta_1, \Delta_2, ...) \tag{7}$$
$$\frac{T_2}{T_*}\frac{dx_2}{dt} = f_2(x_1, x_2, ..., t, \Delta_1, \Delta_2, ...)$$

Las ecuaciones (7) son adimensionales. Éstas son escritas en términos de medidas numéricas adimensionales: $t, x_1, x_2, ...$. Los Multiplicadores $\frac{T_1}{T_*}, \frac{T_2}{T_*}, ...$ en los lados izquierdos de las ecuaciones son también adimensionales, por consiguiente $T_1, T_2, ...$ tienen las dimensiones del tiempo y serán llamados a menudo "constantes temporales" para las variables correspondientes del sistema. Los grupos adimensionales $\Delta_1, \Delta_2, ...$ en el lado derecho serán expresados por $T_*, X_{1*}, ...$ Esto completa la normalizacion de las Ecs. (2).

En el sistema normalizado (7), las cantidades $\frac{T_1}{T_*}, \frac{T_2}{T_*}, ..., \Delta_1, \Delta_2, ...$ corresponden a la clase de movimiento elegido. Algunas de estas cantidades pueden ser lo suficientemente pequeñas en favor de jugar el papel de pequeños parámetros en un análisis de aproximación.

Debe notarse que la normalización de ecuaciones es la parte más importante y la menos formalizada del análisis fraccionario en un problema. está depende principalmente de la experiencia y habilidad de un investigador. Para escoger los valores característicos apropiados y las relaciones que los conectan, uno tiene que usar los datos experimentales, analogías, las estimaciones ásperas de solución de las ecuaciones iniciales, así cómo la busqueda de correspondencia con los objetivos de la investigación. No es necesario poner los valores característicos con una exactitud muy alta. Debe estar de acuerdo con un valor de un parámetro pequeño y con la exactitud requerida de las aproximaciones. Por ejemplo, si un parámetro pequeño es del orden de $10^{-3} - 10^{-4}$, la exactitud del valor-característico colocado debe ser suficiente que sea del orden del 10%.

3. Modelo de Perturbación Singular Estándar

El modelo de perturbación singular de un sistema Dinámico es un modelo en el espacio de estado en el cual la derivada de algúno de los estados está multiplicada por el parámetro pequeño positivo μ esto es

$$\dot{y} = f(y, z, \mu, t), \quad y(t_0) = y^0, \quad y \in R^n \tag{8}$$

$$\mu\dot{z} = g(y, z, \mu, t), \quad z(t_0) = z^0, \quad z \in R^m \tag{9}$$

suponemos que las funciones f y g son diferenciablemente continuas con respecto a sus argumentos $y, z, \mu, t \in D_y \times D_z \times [0, \mu_0] \times [0, t_1]$ donde

$D_y \subset R^n$ y $D_z \subset R^m$ son conjuntos conectados abiertos. Un sistema de este tipo es llamado sistema perturbado singularmente. El escalar μ representa todos los parámetros pequeños que serán ignorados. En control y teoría de sistemas, el modelo $(8), (9)$ es un paso hacia "la modelación de orden reducido". Cuando se coloca $\mu = 0$ se obtiene un sistema no perturbado de ecuaciones, o llamado un sistema degenerado donde el orden es menor que el orden del sistema $(8) - (9)$

$$0 = g(y, z, \mu, t), \quad \dot{y} = f(y, z, \mu, t) \tag{10}$$

Para encontrar la soluciòn de este sistema se resuelve la primera ecuaciòn de (10) está ecuaciòn es nolineal y puede poseer varias soluciones, suponemos que todas las soluciones (raices) $z = \varphi(y, t)$ de está ecuaciòn son reales y están aisladas en $\bar{D}$, una interpretaciòn geometrica de está funciòn $z = \varphi(y, t)$ se puede ver usado el teorema de la funciòn implicita[2]. Es necesario escoger una de las raices de $z = \varphi(y, t)$ y sustituir en la segunda ecuaciòn de (10) el criterio para escoger la raiz se da en los puntos 3 y 5 de las condiciones que debe satisfacer el Teorema de Tikhonov, despues de sustituir $z = \varphi(y, t)$ en la segunda ec. (10)se obtiene una ecuaciòn diferencial con respecto a y de las 2 condiciones iniciales que se dan en $(8) - (9)$ se reduce a una sola condiciòn inicial para la ecuaciòn

$$\frac{d\tilde{y}}{dt} = f(\tilde{y}, \varphi(\tilde{y}, t), \mu, t), \quad \tilde{y}(t_0) = y^0 \tag{11}$$

La barra es usada para indicar que las variables pertenecen a un sistema con $\mu = 0$. Se dice que el modelo (8)-(9) está en la forma estándar si y sólo si(10) tiene $k \geq 1$ raices reales aisladas

Definición 1 *La raiz $z = \varphi(y, t)$ de un sistema de ecuaciones*

$$g(y, z, \mu, t) = 0 \tag{12}$$

es llamado aislado en algún rango restringido de las variables y y t si las otra raices del sistema (12) no existen para cada valor fijo y y t en cualquier proximidad pequeña de la raiz.

Definición 2 *La raiz $z = \phi(y, t)$ se dice ser estable en el dominio $\bar{D}$ si, para todos los puntos $(y, t) \in \bar{D}$, se tiene la desigualdad*

$$\frac{\partial \varphi}{\partial z}(\phi(y, t), y, t) < 0$$

Teorema 3 *TEOREMA DE LA FUNCIÒN IMPLÌCITA. Sea $A \subset \mathbb{R}^m$ y $B \subset \mathbb{R}^m$ conjuntos abiertos. Sea $F: A \times B \to \mathbb{R}^n$ un mapeo C^∞. Denotemos por*

$$(x, y) = (x_1, \ldots, x_m, y_1, \ldots, y_n)$$

denota un punto de $A \times B \to \mathbb{R}^n$. *Suponga que para algún* $(x^0, y^0) \in A \times B$

$$F\left(x^0, y^0\right) = 0$$

y que la matriz

$$\frac{\partial F}{\partial y} = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} \cdots \frac{\partial f_1}{\partial y_n} \\ \cdots \cdots \cdots \\ \frac{\partial f_n}{\partial y_1} \cdots \frac{\partial f_n}{\partial y_n} \end{pmatrix}$$

es no singular en (x^0, y^0). *Entonces, existe una vecindad abierta* A_0 *de* x_0 *en* A *y* B_0 *de* y_0 *en* B *y un único mapeo* C^∞, $G : A_0 \to B_0$ *tal que*

$$F\left(x, G\left(x\right)\right) = 0$$

para toda $x \in A_0$.

El modelo (11) es algúnas veces llamado modelo de estado "casi-estacionario", cuando $\dot{z} = g/\mu$ puede ser muy grande cuando μ es pequeña, entonces $z\left(t\right)$ converge rápidamente a una raiz de (10), el cual es punto de equilibrio y corresponde a la forma cuasi-estacionaria de estado de (9). La forma de la singularidad (1) sugiere un cambio de la variable t, es decir sustituir t por un rescalamiento de la variable τ definida cómo

$$\tau = \frac{t}{\mu}$$

donde μ es el parámetro pequeño, las variable t y τ usualmente se conocen cómo el tiempo "lento" y "rápido" La conveniencia para usar un parámetro para lograr reducir el orden también tiene un inconveniente: no es siempre claro cómo escoger el parámetro que sera considerado pequeño, en este trabajo se usara la técnica de normalización [4], para un análisis geometrico local de la teoría de perturbaciones singulares ver [2]

En [5] se puede ver una versión un poco diferente del teorema de Tikhonov el cual está enfocado a problemas de control, aquí se dará la versión más conocida del teorema cómo se puede ver en [4],[6],[1].

Teorema 4 *Teorema de Tikhonov: Considere el problema de perturbaciones singulares*

$$\begin{aligned} \frac{dy}{dt} &= f\left(y, z, \mu, t\right) \quad y\left(0\right) = y_0, \quad y \in R^n \\ \mu\frac{dz}{dt} &= g\left(y, z, \mu, t\right) \quad z\left(0\right) = z_0, \quad z \in R^m, \quad \mu \ll 1 \end{aligned} \tag{13}$$

Suponga que las siguientes condiciones se satisfacen 1.- Las funciones $f\left(y, z, \mu, t\right)$ *y* $g\left(y, z, \mu, t\right)$ *son analíticas con respecto a* y, z, μ, t *en a* lg *ún*

dominio de las variables de estado. 2. La ecuación $g(y,z,\mu,t)=0$, *tiene una raiz* $z=\phi(y,t)$ *en a* $\lg$*ún dominio limitado* D *de variables y* y t, *y su raiz está aislada. 3. La función* $f(y,\phi(y,t),0,t)$ *es analítica con respecto a* y *y* t *4. Las condiciones iniciales* z_0 *están en el dominio de influencia de la raiz* $z=\phi(y,t)$ *para el sistema* $\frac{dz}{d\tau}=f(y,z,\mu,t)$. *5. El Punto estacionario* $z=\phi(y,t)$ *del sistema* $\frac{dz}{d\tau}=f(y,z,\mu,t)$ *es asintóticamente estable de acuerdo a Lyapunov para todo* y *y* t *, para el cual la raiz de* $g(y,z,\mu,t)=0$ *está definida. Si las condiciones del 1-5 se satisfacen la solución* $y(t,\mu)$, $z(t,\mu)$ *del problema* (13) *existe en* $[0,T]$ *y satisface el limite de las igualdades.*

$$\begin{aligned}&\lim_{\mu\to 0} y(t,\mu)=\tilde{y}(t) \quad para \quad 0\le t\le t'\\ &\lim_{\mu\to 0} z(t,\mu)=\phi(\tilde{y}(t),t)=\tilde{z}(t) \quad para \quad 0<t\le t'\end{aligned} \tag{14}$$

donde $\tilde{y}(t)$ *es la solución del problema degenerado*

$$d\tilde{y}(t)/dt=f(\tilde{y},\phi(\tilde{y},t),t,0)$$

Definición 5 *Modelo de capa frontera : Consiste en el modelo reducido que describe el movimiento de las variables rápidas de un sistema perturbado singular en una escala de tiempo rápido donde las variables lentas son tratadas cómo parámetros constantes.*

La condición esencial del teorema de Tikhonov es el requerimiento de la ¨región de atracción ¨ de las trayectorias del sistema inicial por la superficie $g(y,z,t,0)=0$.La interpretación de las condiciones del teorema de Tikhonov se puede ver en [5]

3.1 Planteamiento Matemático del Problema

Para la simplificación del modelo matemático usando el teorema de Tikhonov es importante que la ecuación diferencial no lineal que modela el sistema Dinámico se escriba en la forma estándar (1)y esto se logra aplicando el método de normalización [4], en está sección se presenta un algoritmo de cómo usar el Teorema de Tikhonov cuando el sistema dinámico involucra leyes de control.

El modelo matemático del objeto dinámico controlable se puede escribir cómo un sistema de ecuaciones sin dimensiones y con parámetro pequeño $0<\mu=cte\ll 1$

$$\mu\frac{dz}{dt}=g(z,y,t,u_1), \quad z(0,\mu)=z^0 \tag{15}$$

$$\frac{dy}{dt}=f(z,y,t,u_2), \quad y(0,\mu)=y^0 \tag{16}$$

donde $z(\tau,\mu)$- es el vector de coordenadas rápidas

$y(\tau,\mu)$- es el vector de coordenadas lentas

$u_1 = u_1(z,y,t)$ es el control escalar para construir el primer nivel de control (sistema de estabilización)

$u_2 = u_2(z,y,t)$ es el control escalar para construir el segundo nivel de control

De acuerdo al teorema de Tikhonov se pueden investigar dos subsistemas:

a) El primer subsistema adicional con tiempo rápido $\tau = \frac{t}{\mu}$ (tiempo computacional):

$$\frac{dz}{d\tau} = g(y,z,u_1), \quad z(0) = z^0 \tag{17}$$

donde la variable y está fija;

b) El segundo subsistema simplificado o degenerado que se obtiene considerando $(\mu = 0)$;

$$\begin{aligned} \frac{d\tilde{y}}{dt} &= f(\tilde{y},\tilde{z},u_2), \quad \tilde{y}(0) = y^0 \\ 0 &= g(\tilde{z},\tilde{y},u_1) \end{aligned} \tag{18}$$

donde $\tilde{y},\tilde{z}$ corresponde a las coordenadas simplificadas

$$\tilde{z}(0) \neq z(0), \quad \tilde{y}(0) = y^0$$

Para realizar el análisis primeramente tenemos que construir el primer nivel de control con dos objetivos principales los cuales son:

i) Verificar que se cumplan todas las condiciones del Teorema de Tikhonov

ii) Construir simultáneamente el sistema de estabilización.

3. 2 Síntesis del Sistema de Estabilización

Para investigar el primer subsistema adicional (17) se analiza el sistema no lineal algebraico

$$g(z,y,u_1) = 0 \tag{19}$$

Supongamos que el sistema (19) tiene solamente una raiz aislada, la cual podemos presentarla en forma directa cómo

$$z^0 = \varphi(y,u_1) \tag{20}$$

Para realizar está forma (20) representamos el control u_1 cómo una estrategia lineal

$$u_1 = u_1^0(y) + \Delta u_1 \tag{21}$$

donde $u_1^0(y)$ es el control principal para realizar la forma (20) y Δu_1 es el control adicional para estabilizar el punto de reposo $z^0(20)$.

Para construir el control adicional Δu_1 reescribiremos el subsistema (17) con ayuda de las desviaciones $\Delta z = z - z^0$ en forma lineal

$$\frac{d\Delta z}{d\tau} = A(y)\Delta z + b(y)\Delta u_1 \tag{22}$$

donde

$$A(y) = \frac{\partial g(\varphi(y, u_1^0(y)), y, u_1^0(y))}{\partial z}$$
$$b(y) = \frac{\partial g(\varphi(y, u_1^0(y)), y, u_1^0(y))}{\partial u_1}$$

Supongamos que el $\det(b(y)\ A(y)b(y)\ \ldots\ A^{n-1}(y)b(y)) \neq 0$, para $\forall y \in Y_0 \subset R^s$ es el espacio de coordenadas lentas.

Entonces existe el control adicional que se propone de la forma

$$\Delta u_1 = k^T(y)\Delta z \tag{23}$$

tal que la solucion trivial $\Delta z = 0$ del subsistema cerrado

$$\frac{d\Delta z}{d\tau} = \left(A(y) + b(y)k^T(y)\right)\Delta z \tag{24}$$

es estable asintóticamente y también la solución z^0 del subsistema no-lineal, el control Δu_1 estabiliza al sistema lineal alrededor del punto de equilibrio entonces también estabiliza al sistema no lineal alrededor del punto de equilibrio.

$$\frac{dz}{d\tau} = g\left(z, y, u_1^0(y) + \Delta u_1(y, \Delta z)\right) \tag{25}$$

es estable asintóticamente y la condición inicial z^0 pertenece a una región de atracción.

Por eso tenemos que se cumplen todas las condiciones de Tikhonov.

3.3 Diseño de Estimadores

Si no tenemos información exacta sobre las desviaciones Δz, entonces se debe construir un algoritmo de estimación y sintetizar el algoritmo (23) en la forma siguiente

$$\Delta u_1 = k^T(y)\Delta\tilde{z} \tag{26}$$

donde $\Delta\tilde{z}$ es la salida del algoritmo de estimación.

$$\frac{d\Delta\tilde{z}}{d\tau} = A(y)\,\Delta\tilde{z} + b(y)\,\Delta u_1 + \tilde{K}_e\left(\Delta\tilde{z}_1 - h^T(y)\,\Delta\tilde{z}\right) \tag{27}$$

$$\Delta\tilde{z}_1 = h^T(y)\,\Delta\tilde{z} \tag{28}$$

donde (28) es el modelo matemático de los sensores de estado

Para está situación se tiene el nuevo modelo del sistema adicional

$$\begin{aligned}
\frac{dz}{d\tau} &= g\left(z, y, u_1^0(y) + k^T(y)\,\Delta\tilde{z}\right), \qquad z(0) = z^0 \\
\frac{d\Delta\tilde{z}}{d\tau} &= A(y)\,\Delta\tilde{z} + \tilde{k}(y)\left(\Delta\tilde{z}_1 - h^T(y)\,\Delta\tilde{z}\right), \quad \Delta\tilde{z}(0) = 0
\end{aligned} \tag{29}$$

Para cumplir todas las condiciones del teorema de Tikhonov tenemos que suponer que la salida del sensor es una función suave dependiente del tiempo y que se satisface la condición de Kalman $\det(h, Ah, ..., A^{n-1}h) \neq 0$ para $\forall y \in Y_0$

En situacion asintótica para μ, cuando $\tau = t/\mu$ y para $t \in [0, t_1]$ tenemos que $\tau \to \infty$ para $\mu \to 0$. Por eso podemos usar los resultados de la teoría de control optimal para buscar los parámetros $k_1, k_2, ..., k_n,$.

Para el tiempo computacional $\tau \in [0, \infty]$ y para el funcional

$$\int_0^\infty (\Delta z^T \Im \Delta z + r_0 \Delta u_1^2) d\tilde{t} \qquad \text{donde } \Im^T = \Im > 0$$

y $r_0 = cte > 0$ segun el método de programaciòn dinámica de Bellman podemos calcular los parámetros optimales (esto es posible cuando $\det(b, Ab, ...A^{n-1}b) \neq 0$)

$$k^0 = -\frac{1}{r_0} L_0 b \tag{30}$$

donde L_0 es la soluciòn positiva $\left(L_0^T = L_0 > 0\right)$ de la ecuaciòn algebraica de Ricati

$$\frac{1}{r_0} L b b^T L - \left(A^T L + LA\right) - \Im = 0 \tag{31}$$

Despues de realizar la Síntesis del subsistema adicional tenemos que regresal al tiempo t. Ahora tenemos el siguiente subsistema (15).

$$\begin{aligned}
\mu\frac{dz}{dt} &= g\left(z, y, u_1^0(y) + k^T\Delta\tilde{z}\right), \quad z(0, \mu) = z^0 \\
\mu\frac{d\Delta\tilde{z}}{dt} &= A(y)\,\Delta\tilde{z} + b(y)\,k^T\Delta\tilde{z} + \tilde{k}(y)\left(\Delta\tilde{z}_1 - h^T(y)\,\Delta\tilde{z}\right), \quad \Delta\tilde{z}(0) = 0
\end{aligned}$$

donde $k = -\frac{1}{r_0} L_0 b$ de (31)

$\Delta\tilde{z}_1 = h^T(y)(z - z^0)$
$z^0 = \varphi(z, y, u_1^0(y))$

Entre todas las condiciones del teorema de Tikhonov hace falta verificar pertenencia de las condiciones iniicales z^0 en una regiòn de atracciòn Y_0. también tenemos que buscar los parámetros $\tilde{k}_1, \tilde{k}_2, ..., \tilde{k}_n,$. los cuales existen segun observabilidad completa del susbsistema adicional en desviaciones $(\det(h, Ah, ..., A^{n-1}h) \neq 0)$.

4. Aplicación del Teorema de Tikhonov a un Problema de Aeronáutica

En está parte se aplica la metodología mostrada anteriormente para simplificar un modelo matemático de un sistema dinámico controlable basado principalmente en el proceso de normalización para obtener la forma estándar (1), el siguiente tratamiento se desarrolla para el movimiento longitudinal de un avión en el espacio. Las ecuaciones dinámicas de movimiento longitudinal son:

$$\begin{aligned} M\frac{dV}{dT} &= -Mg\sin\theta - \frac{1}{2}PV^2Sc_x + P^T\cos\theta \\ MV\frac{d\theta}{dT} &= -Mg\cos\theta + \frac{1}{2}PV^2Sc_y + P^T\sin\theta \\ I_{zz}\frac{d\Omega}{dT} &= -\frac{1}{2}PV^2Sb_a\left(m_z^\alpha\alpha + m_z^\delta\delta\right) \\ \frac{d\varphi}{dT} &= \Omega, \quad \varphi = \theta + \alpha, P = P(Y), \\ \frac{dM}{dT} &= -U, \quad \frac{dX}{dT} = V\cos\theta, \quad \frac{dY}{dT} = V\sin\theta, \quad \frac{dI_{zz}}{dT} = -W \end{aligned} \tag{32}$$

$$\begin{aligned} c_y(\alpha) &= c_y^0 + c_y^\alpha\alpha \\ c_x(\alpha) &= c_x^0 + Bc_y^2 \end{aligned} \tag{33}$$

aquí M es la masa del avión, X y Y son las coordenadas del centro de masas; V es la velocidad; P es la densidad del aire; $\theta, \alpha, \vartheta$ y δ_z son el angulo de trayectoria, de ataque, de picada y la deflección del elevador; I_{zz}, S, y b_a son el momento de inercia , el área característica del sólido y la longitud; W la variacion del momento de inercia con respecto al tiempo, c_x, c_y son los coeficientes aerodinámicos de la fuerza longitudinal y normal, $m_z^\alpha + ...$ son los coeficientes aerodinámicos. Las ecuaciones (32) se basan en las siguientes suposiciones, la tierra es plana y no rota; la configuración del cuerpo del avión es fija; el eje del motor coincide con el eje longitudinal del avión. Estas suposiciones que se hacen no son importantes desde el punto de vista de la separación del movimiento, pero simplifican los cálculos.

4. 1 Ecuaciones de Movimiento para un Planeador

Del sistema de ecuaciones ecs. (32)se obtienen las ecuaciones de movimiento para el planeador que decendera por la glisada considerando que el unico control que se usara es el timon δ.

$$\begin{aligned} M\frac{dV}{dT} &= -Mg\sin\theta - \frac{1}{2}PV^2Sc_x \\ MV\frac{d\theta}{dT} &= -Mg\cos\theta + \frac{1}{2}PV^2Sc_y \\ I_{zz}\frac{d\Omega}{dT} &= -\frac{1}{2}PV^2Sb_a\left(m_z^\alpha\alpha + m_z^\delta\delta\right) \\ \frac{d\varphi}{dT} &= \Omega, \quad \varphi = \theta + \alpha \end{aligned} \tag{34}$$

El sistema de ecuaciones (34) serán normalizadas introduciendo cantidades analogas sin dimensiones

$$t = \frac{T}{T_*}, v = \frac{V}{V_*}, w = \frac{\Omega}{\Omega_*}$$

con respecto a las variables α, φ, θ se supone que se miden en diferentes sistemas de medición y que toman valores del orden de la unidad, por este motivo no se normalizarán, $M_*, V_*, U_*, W_*, X_*, Y_*, I_*$ se suponen son iguales al maximo a uno de los valores correspondientes para el tipo de avión especificado y el modo de vuelo. también suponemos que los valores característicos de las fuerzas aerodinámicas y las fuerzas de empuje son del orden del peso del avión.

Entonces P_* y P^T pueden encontrarse de la siguiente manera

$$\frac{1}{2}P_*V_*^2S = P_*^T = M_*g$$

aceptemos que $P = P_*, M = M_*, I_{zz} = I_*$. El valor de Ω_* se obtiene de un estimado dado por la ecuación simplificada que se obtiene del sistema de ecuaciones (32) considerando que $\delta = \theta = 0$

$$I_*\frac{d^2\varphi}{dT^2} = \frac{1}{2}P_*V_*^2Sb_am_z^\alpha\varphi$$

con $m_z^\alpha \sim -1$, el tiempo constante para este elemento de oscilación puede ser estimado por la expresion

$$T_1^2 = \frac{2I_*}{P_*V_*^2Sb_a} = \frac{M_*r_*^2}{M_*gb_a}$$

donde r_*es el radio central de inercia, entonces si $\varphi \sim 1$, $\Omega_* = \frac{1}{T_1}$

$$\begin{aligned}
\frac{T_2}{T_*}\frac{dv}{dt} &= -\sin\theta - v^2 c_x \\
\frac{T_2}{T_*}\frac{d\theta}{dt} &= -\frac{\cos\theta}{v} + v c_y \\
\frac{T_1}{T_*}\frac{dw_z}{dt} &= \{m_z^{\alpha}\alpha + m_z^{\delta_z}\delta\}v^2 \qquad (35) \\
\frac{T_1}{T_*}\frac{d\vartheta}{dt} &= w \qquad (36)
\end{aligned}$$

donde $T_2 = \frac{V_*}{g}, T_1$son las constantes de tiempo parciales del sistema.

las característica de vuelo para un avión :

$$\begin{aligned}
V_* &= 300\ m/seg, \\
T_2 &= \frac{V_*}{g} \approx 30\ seg, \quad T_2 \gg T_1 \\
b_a &\sim 10 mts.\ r_* = 10 mts.
\end{aligned}$$

4. 1. 1 Separación de las Componentes Dinámicas

Al separar las componentes dinámicas del movimiento del centro de masa tomando lugar en intervalos de tiempo del orden T_2, supongamos $T_* = T_2$ en el sistema (35) toma la forma

$$\begin{aligned}
\frac{dv}{dt} &= -\sin\theta - v^2 c_x, \quad v(t_0) = v_0 \qquad (37) \\
\frac{d\theta}{dt} &= -\frac{\cos\theta}{v} + v c_y, \quad \theta(t_0) = \theta_0 \\
\mu\frac{dw}{dt} &= -\{m_z^{\alpha}\alpha + m_z^{\delta}\delta\}v^2, \quad w(t_0) = w_0 \\
\mu\frac{d\varphi}{dt} &= w \quad \varphi(t_0) = \varphi_0
\end{aligned}$$

El sistema (37) es perturbado singularmente con respecto a μ. El sistema (37) involucra sólo un control que corresponde al timon δ.

4. 2 El Análisis de Estabilidad para Aterrizaje por Planeamiento

Procederemos a investigar el subsistema adicional con variables rápidas, analicemos el sistema algebraico que se obtiene al considerar $\mu = 0$

$$\begin{aligned}
0 &= -\{m_z^{\alpha}\alpha + m_z^{\delta}\delta\}v^2 \\
0 &= w
\end{aligned}$$

la solución del sistema de ecuaciones algebraicas, corresponde al punto de equilibrio y se obtiene $w = 0$, para determinar el control principal se presenta la siguiente estrategia $\delta = \delta_0 + \Delta\delta$ donde el control principal se determinar a partir de este sistema de ecuaciones algebraicas $\delta_0 = (m_z^\alpha / m_z^\delta)(\theta - \varphi_0)$ donde θ se mantiene constante

El análisis de estabilidad en el punto de equilibrio lo podemos realizar atravez del sistema rápido, haciendo un cambio en la escala del tiempo $\tau = \frac{t}{\mu}$ el subsistema rápido (17) toma la forma

$$\frac{dw}{d\tau} = -\{m_z^\alpha \alpha + m_z^\delta (\delta_0(\theta, \varphi_0) + \Delta\delta)\} v^2, \quad w(t_0) = w_0 \tag{38}$$

$$\frac{d\varphi}{d\tau} = w, \quad \varphi(t_0) = \varphi_0 \tag{39}$$

se propone que $\Delta\delta$ sea proporcional a w.

$$\Delta\delta = kw \tag{40}$$

escribiendo la ec. (38) y la (40) en desviaciones se obtiene la siguiente ec. diferencial

$$\frac{d^2\Delta\varphi}{d\tau^2} + m_z^\delta k v^2 \frac{d\Delta\varphi}{d\tau} + m_z^\alpha v^2 \Delta\varphi = 0 \tag{41}$$

la soluciòn trivial $\Delta\varphi = 0$ es estable asintóticamente, aplicando el teorema de Hurwitz se observa que $m_z^\delta k v^2 > 0$, $m_z^\alpha v^2 > 0$, entonces el subsistema rápido (41) tiene comportamiento asintóticamente estable y globalmente en el punto de equilibrio, $w = 0, \varphi_0$, que es también es estable asintóticamente para el subsistema nolineal

$$\frac{dw}{d\tau} = -\{m_z^\alpha \alpha + m_z^\delta (\delta_0(\theta, \varphi_0) + \Delta\delta)\} v^2$$

y la condiciòn φ_0 pertenece a una regiòn de atracciòn, se cumple el teorema de Tikhonov

4. 3 Reduccción al Sistema Simplicado

El subsistema simplificado con $\mu = 0$ nos proporciona el siguiente sistema de ecuaciones

$$\begin{aligned}
\frac{d\tilde{v}}{dt} &= -\sin\tilde{\theta} - \tilde{v}^2 c_x, \quad \tilde{v}(t_0) = \tilde{v}_0 \\
\frac{d\tilde{\theta}}{dt} &= -\frac{\cos\tilde{\theta}}{v} + \tilde{v} c_y, \quad \tilde{\theta}(t_0) = \tilde{\theta}_0 \\
0 &= -\{m_z^\alpha \alpha + m_z^\delta \delta\} v^2 \\
0 &= w
\end{aligned}$$

donde $\tilde{v}, \tilde{\theta}$ son coordenadas simplicadas donde $\tilde{v}(0) \neq v(0), \tilde{\theta}(0) \neq \theta(0)$

Se desea realizar un vuelo bajo las siguientes condiciones

$$\tilde{\theta} \equiv \tilde{\theta}_0 < 0, \tilde{v} = v(t_0)$$

de la primera ecuación de balance se obtiene que $\tilde{\theta}_0 = -arc\,\mathrm{Si}\,n(\tilde{v}^2 c_x) < 0$, los coeficientes aero dinámico están dados por (33) tomando $c_y^0 = 0$ entonces $c_y^\alpha(\varphi_0 - \theta_0) v_0 = \frac{\cos\theta_0}{v_0}$ se obtiene que $\varphi_0 = \theta_0 + \frac{\cos\theta_0}{c_y^\alpha v_0^2}$ de aquí que $\alpha_0 = \varphi_0 - \theta_0 = \frac{\cos\theta_0}{c_y^\alpha v_0^2} > 0$

4.3.1 Análisis de Estabilidad para el Sistema Simplicado

Usando desviaciones

$$\Delta v = \tilde{v} - \tilde{v}_0$$
$$\Delta\theta = \tilde{\theta} - \tilde{\theta}_0$$

Se tiene el siguiente sistema lineal

$$\Delta\dot{v} = -(c_x 2v_0)\Delta v - (\cos\theta_0)\Delta\theta$$
$$\Delta\dot{\theta} = \left(c_y^\alpha\alpha_0 + \frac{\cos\theta_0}{v_0^2}\right)\Delta v + \left(\frac{\sin\theta_0}{v_0} - c_y^\alpha v_0^2\right)\Delta\theta$$

donde

$$Det(A - \lambda E) = \begin{vmatrix} -c_x 2v_0 - \lambda & -\cos\theta_0 \\ c_y^\alpha\alpha_0 + \frac{\cos\theta_0}{v_0^2}\left(\frac{\sin\theta_0}{v_0} - c_y^\alpha v_0^2\right) - \lambda \end{vmatrix} =$$
$$= \lambda^2 + \left(c_x 2v_0 - \frac{\sin\theta_0}{v_0} + c_y^\alpha v_0\right)\lambda + \cos\theta_0\left(\frac{\cos\theta_0}{v_0} + c_y^\alpha\alpha_0\right) = 0$$

aplicando el teorema de Hurwitz al polinomio caracteristico el sistema es asintóticamente estable si los coeficientes del polinomio son positivos

$$\left(c_x 2v_0 - \frac{\sin\theta_0}{v_0} + c_y^\alpha v_0\right) > 0$$
$$\cos\theta_0\left(\frac{\cos\theta_0}{v_0} + c_y^\alpha\alpha_0\right) > 0$$

con esto se concluye que la trayectoria que se eligio para el aterrizaje es estable y se satisface el Teorema de Tikhonov.

5. Resultados Numéricos

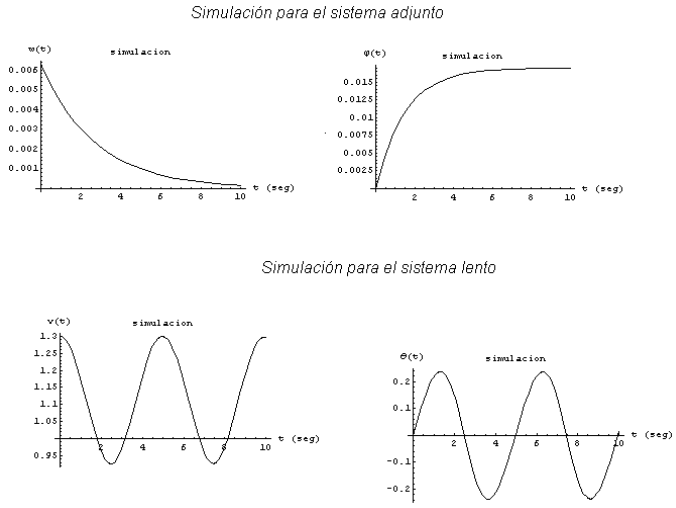
parámetros para el sistema adjunto

$c_x^0 = .3$

$c_y = .8$

$m = 500$

$v(0) = 1.3$
$\theta(0) = .001$
parámetros para el sistema con variables lentas
$m_z^{\alpha} = .5$
$m_z^{\delta} = .9$
$\theta = .017$
$k = 1.5$
$v = 300$
$\varphi_0 = .017$



Simulaciòn correspondiente al sistema rapido y lento

Conclusiones *El uso de la teoría de perturbaciones singulares muestra ser de gran ayuda para obtener soluciones bajo un método asintótivos, las cuales se obtienen si el modelo estándar cumple con las condiciones del Teorema de Tikhonov, la aplicación de este Teorema sirve para simplificar modelos matemáticos complejos y obtener una solución bajo condiciones deseadas.*

Referencias

[1] H. K. Khalil, Nonlinear Systems, Prentice Hall, New Jersey, 2202.

[2] A. Isidori, Nonlinear Control Systems, Springer-Verlag, 1995

[3] Shankar Sastry Grantham, "Nonlinear Systems Analysis Stability and Control" Springer-Verlag New York 1999.

[4] I. V. Novozhilov,"Fractional Analysis Methods of Motion Decomposition", Birkhäuser Boston Basel Berlin 1997

[5] A. N. Tikhonov, A. B. Vasil'eva, A. G. Sveshnikov, Differential Equations, Springer-Verlag Berlin Heidelberg 1985

[6] Peter Kokotovic, Hassan K. Khalil, John O ´Reilley, "Singular Perturbation Methods in Control analysis and Design" SIAM, Philadelphia 1999

Lanzamiento Optimal de un Avión Automático

Maribel Reyes Romero, V.V. Alexandrov, W. Fermín Guerrero Sánchez
Facultad de Ciencias Físico Matemáticas, BUAP
18 sur y Av. San Claudio, Colonia San manuel, Ciudad Universitaria
maribelrr@gmail.com, valex@fcfm.buap.mx, willi@fcfm.buap.mx

Resumen

En este trabajo se trata el problema del **lanzamiento vertical de un avión automático hasta alcanzar una altura dada h_k de manera que el combustible gastado sea mínimo**. Este problema pertenece a los llamados Problemas de Control Optimal, tales problemas consisten en hallar un control optimal, dado de antemano el criterio de optimalidad, que lleve al sistema al objetivo deseado. La propuesta para hallar tal control optimal es usar dos resultados fundamentales de la teoría de control optimal moderno llamados "**El Principio del Máximo de Pontryagin**" y la "**Condición necesaria de H. J. Kelly**". Así, en esta trabajo, se plantea el problema como un problema estándar de control óptimo y se aplica el El Principio del Máximo de Pontryagin y la Condición necesaria de H. J. Kelly para la solución del problema planteado.

Modalidad: cartel.

Palabras clave: control optimal, criterio de optimalidad, control optimal singular.

1. Introducción

En general, El Principio del Máximo de Pontryagin dá las condiciones necesarias para la optimización de un sistema. Éste trata el problema de optimización de maximizar o minimizar un funcional sujeto a ciertas constricciones, es decir

$$\begin{aligned} J(\mathbf{u}) &\rightarrow \textit{mín}, \text{ sujeto a} \\ \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)). \end{aligned}$$

Para aplicar el principio del máximo en la solución de problemas de control óptimo, el procedimento de diseño es iniciado con la maximización de la función de Pontriagyn $H(\mathbf{x}, \mathbf{\Psi}, \mathbf{u})$ con respecto al vector de control $\mathbf{u}$. Esto resulta en un vector de control óptimo, como función del vector $\mathbf{\Psi}$. El vector de control es sustituído entonces en el sistema adjunto y el problema resultante con valores en la frontera se resuelve para obtener la trayectoria optimal $\mathbf{x}^0(t)$ y así obtener $J(\mathbf{u}) \rightarrow \min$. El diseño del control óptimo tiene como objetivo la determinación de una ley de control óptima $\mathbf{u}^0$ o una función de control-óptimal $\mathbf{u}^0$.

Cuando el principio del Máximo se cumple pero no funciona bien para hallar el control optimal, entonces se usa la condición necesaria de H. J. Kelly para hallar el control, a este control se le llama control optimal singular.

Las ecuaciones diferenciales resultantes de estas dos condiciones necesarias se resuelven numéricamente para hallar la trayectoria optimal, pues se trata de ecuaciones diferenciales no lineales, difícil de obtener de ellas una solución analítica.

2. Planteamiento del Problema

Se quiere elevar un avión automático hasta la altura h_k con gasto mínimo de combustible.

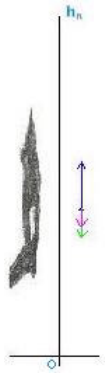


Figura1. Lanzamiento vertical hasta la altura h_k.

La propuesta es lanzarlo verticalmente con ayuda de cohetes adecuados unidos al avión y minimizar el gasto de combustible usando dos estrategias de control para lograr la altura deseada h_k. El procedimiento se ve posteriormente.

OBJETIVOS GENERALES:

1. Plantear al problema del lanzamiento vertical de un avión automático con gasto mínimo de combustible como un problema estándar de control optimal.
2. Hallar el control optimal que lleve al avión automático hasta la altura h_k.

OBJETIVOS PARTICULARES:
Usar las dos condiciones necesarias:

1. Principio del Máximo de Pontriagyn, y
2. Condiciòn necesaria de H. J. Kelly

para hallar el control optimal.

3. Desarrollo del Problema

Las ecuaciones de movimiento para el movimiento vertical del avión son:

$$\begin{aligned} \dot{y_1} &= y_2; \quad y_1(0) = 0 \\ \dot{y_2} &= -g - \frac{\rho S C_x}{2}\frac{y_2^2}{y_3} + \mu\frac{u}{y_3}; \quad y_2(0) = 0 \\ \dot{y_3} &= -u; \quad y_3(0) = M_0 \end{aligned} \tag{1}$$

donde y_1 es la altura del avión, y_2 es la velocidad del avión, y_3 es la masa del avión, u es la velocidad con la que se está gastando el combustible, μ es la velocidad de los gases (resultado de la quema del combustible en los cohetes) relativa al avión, g es la aceleración de la gravedad, ρ es la densidad del aire, A es el área de la sección transversal del avión y C_x es una cantidad empírica adimensional conocida como coeficiente de resistencia. Aunque ρ y g varían con la altura, vamos a considerar alturas para las cuales podemos consider ρ y g constantes. Las ecuaciones (1) pueden escribirse en forma vectorial como

$$\begin{aligned} \dot{\mathbf{y}} &= \mathbf{f} \\ \mathbf{y}(\mathbf{0}) &= \mathbf{y}_0 \end{aligned}$$

donde

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}; \ \dot{\mathbf{y}} = \begin{pmatrix} \dot{y_1} \\ \dot{y_2} \\ \dot{y_3} \end{pmatrix}; \ \mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} y_2 \\ -g - \frac{\rho S C_x}{2}\frac{y_2^2}{y_3} + \mu\frac{u}{y_3} \\ -u \end{pmatrix}.$$

Entonces, ahora se plantea el problema como un problema estándar de control optimal.

Queremos minimizar el funcional $\mathcal{J}(u)$

$$\mathcal{J}(u) = \varphi_0(\mathbf{y}(t_k)) = [M_0 - y_3(t_k)] \longrightarrow \underset{u(\cdot)\in U}{mín} \tag{2}$$

tal que $\mathbf{y}$ satisface el sistema de ecs. diferenciales :

$$\begin{aligned} \dot{y_1} &= y_2; \quad y_1(0) = 0 \\ \dot{y_2} &= -g - \frac{\rho S C_x}{2}\frac{y_2^2}{y_3} + \mu\frac{u}{y_3}; \quad y_2(0) = 0 \\ \dot{y_3} &= -u; \quad y_3(0) = M_0 \end{aligned} \tag{3}$$

con

$$0 \leq u \leq u_{\max} \tag{4}$$

y el objetivo o conjunto terminal dado por

$$M = \{y_1 - h_k = 0\} \tag{5}$$

es decir, queremos hallar el **control optimal** u^0 que conduzca al avión hacia el **objetivo** h_k. Aquí, $\mathcal{J}(u)$ es el criterio de optimalidad, es decir, u^0 es el control optimal en el sentido de que minimiza a $\mathcal{J}(u)$.

Ahora escribimos el sistema adjunto asociado al vector adjunto $\mathbf{\Psi} = \begin{pmatrix} \Psi_1 & \Psi_2 & \Psi_3 \end{pmatrix}^T$, dado por:

$$\dot{\mathbf{\Psi}} = -\left(\frac{\partial \mathbf{f}}{\partial \mathbf{y}}\right)^T \mathbf{\Psi}$$

entonces, el sistema adjunto para el problema (2), (3), (4), (5) es

$$\dot{\mathbf{\Psi}} = \begin{pmatrix} \dot{\Psi}_1 \\ \dot{\Psi}_2 \\ \dot{\Psi}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -\Psi_1 + 1\rho SC_x \frac{y_2}{y_3}\Psi_2 \\ \Psi_2\left(\frac{\mu u}{y_3^2} - \frac{\rho SC_x y_2^2}{2y_3^2}\right) \end{pmatrix} \tag{6}$$

entonces $\Psi_1 = \lambda = cte$. La función de Pontriagyn H está dada por

$$H = \mathbf{\Psi}^T \mathbf{f}$$

entonces La función de Pontriagyn H para el problema (2), (3), (4), (5) es

$$H = H_0 + H_1 u \tag{7}$$

donde

$$\begin{aligned} H_0 &= \Psi_1 y_2 + \Psi_2\left(-g - \frac{\rho SC_x y_2^2}{2y_3}\right) \\ H_1 &= \left(\Psi_2 \frac{\mu}{y_3} - \Psi_3\right) \end{aligned} \tag{8}$$

Ahora aplicamos el principio del máximo de Pontryagin el cual se establece como sigue:

Consideremos el sistema

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}[\mathbf{x}, \mathbf{u}], \quad \mathbf{x}(t_0) = x_0 \\ \mathbf{u}(\cdot)\, \epsilon U &= \{\mathbf{u}(\cdot)\, \epsilon L_2^s \epsilon \Omega \subset \mathbb{R}^s\} \\ &\quad s \text{ es el número de controles} \end{aligned}$$

El objetivo es alcanzar una variedad M en algún instante de tiempo t_k.

$\mathbf{x}(t_k)\, \epsilon M\, (suave) \subset \mathbb{R}^n$, es decir $\mathbf{x}(t_k) \notin M \ \forall\ t_k \epsilon (t_0, t_k)$.

Nuestro criterio de calidad $\mathcal{J}(\mathbf{u}) = \varphi_0\left(y\left(t_k^0\right)\right) \rightarrow \min$.

Teorema 1. Suponga que:

1. existe $u(\cdot)\, \epsilon U$ tal que $x(t_k)\, \epsilon M$.
2. existe $u^0(\cdot)\, \epsilon U$ tal que $J\left(\mathbf{u}^0\right) = \min\limits_{\mathbf{u}(\cdot)\epsilon U} J(\mathbf{u})$.

Si $u^0(t)$ es control optimal (es decir, hace mínimo a $J(\mathbf{u})$), entonces existe un par $\{\lambda_0 \geq 0, \mathbf{\Psi}(\cdot)\}$ no trivial tal que

$\alpha)$ $max\ H\left(\mathbf{\Psi}\left(t\right),y^{0}\left(t\right),\mathbf{u}\right)=H\left(\mathbf{\Psi}\left(t\right),y^{0}\left(t\right),\mathbf{u}^{0}\right)$

$\beta)$ $\Psi\left(t_{k}^{0}\right)+\lambda_{0}\frac{\partial\varphi_{0}\left(y\left(t_{k}^{0}\right)\right)}{\partial y}$, es ortogonal a M en el punto $y\left(t_{k}^{0}\right)$.

$\gamma)$ $H=H\left(\mathbf{\Psi}\left(t\right),y^{0}\left(t\right),\mathbf{u}^{0}\right)\equiv 0.$

Ahora aplicamos el principio del máximo de Pontryagin.
Supongamos que $\mathbf{u}^{0}(t)$ es control optimal
De $\alpha)$ tenemos que si $H_1\neq 0$, entonces $u_1^0=\left\{\begin{array}{ll} 0 & \text{si } H_1<0 \\ u_{\max} & \text{si } H_1>0 \end{array}\right\}$.
De $\beta)$ tenemos : $\Psi_2(t_k^0)=0;\quad \Psi_3(t_k^0)=\lambda_0$.
Y por último tenemos de $\gamma)$ $\mathcal{H}(t)=H(\mathbf{\Psi}(t),y^0(t),u^0(t))\equiv 0$.
Además de $\beta)$ y $\gamma)$ tenemos que $\lambda\neq 0$ y $\lambda_0\neq 0$.

Observemos que si $H_1=0$, no podemos hallar el control optimal al maximizar la función de Pontriagyn, pues éste ya no aparece, cuando sucede esto, al control optimal se le llama control optimal singular. Para este caso usamos la condición necesaria de H. J. Kelly, la cual se establece como sigue:
Considere el sistema de la forma

$$\begin{array}{rcl} \dot{\mathbf{x}}(t) & = & \mathbf{f}_0(\mathbf{x}\left(t\right))+\mathbf{f}_1(\mathbf{x}\left(t\right))u\left(t\right), \\ x & \in & R^n,\quad u(t)\in U\subset R^1,\quad \mathbf{x}\left(0\right)=\mathbf{x}_0, \end{array} \tag{9}$$

Teorema 2. Suponga que $u^{0}\left(t\right)$ es un control singular optimal del problema (9), las funciones f_0 y f_1 son continuamente diferenciables al menos 5 veces, y la siguiente igualdad se tienen a lo largo del control $u\left(t\right)$:

$$\left(-1\right)^{k}\frac{\partial}{\partial u}\frac{d^{2k}}{dt^{2k}}\frac{\partial H}{\partial u}=0,\quad k=0,1 \tag{10}$$

Entonces para que el control $u^{0}\left(t\right)$ sea optimal, es necesario que la siguiente igualdad sea válida:

$$\frac{\partial}{\partial u}\frac{d^{2}}{dt^{2}}\frac{\partial H}{\partial u}\geq 0, \tag{11}$$

dado el control $u_{0}\left(t\right)$.

Ahora vamos a aplicar la condición de H.J. Kelly a nuestro problema.

Supongamos ahora que $H_1(t)\equiv 0$ en $t\in[\tau,\overline{\tau}]$, entonces

$$\frac{dH_1(t)}{dt}\equiv 0$$

entonces

$$-\dot{\widetilde{\Psi}}_3+\dot{\widetilde{\Psi}}_2\frac{\mu}{y_3}-\frac{\widetilde{\Psi}_2\mu}{y_3^2}\dot{y}_3=\frac{d\widetilde{H}_1(t)}{dt}=0 \tag{12}$$

donde

$$\begin{aligned} \widetilde{\Psi}_1 &= \frac{\Psi_1}{\lambda}; \quad \widetilde{\Psi}_2 = \frac{\Psi_2}{\lambda}; \quad \widetilde{\Psi}_3 = \frac{\Psi_3}{\lambda} \\ \text{entonces } \widetilde{\Psi}_1 &= 1; \quad \widetilde{\Psi}_2 = \frac{\Psi_2}{\lambda}; \quad \widetilde{\Psi}_3 = \frac{\Psi_3}{\lambda} \end{aligned}$$

sustituyendo $\dot{y}_3$, $\dot{\widetilde{\Psi}}_2$, $\dot{\widetilde{\Psi}}_3$ en (12) obtenemos:

$$\widetilde{\Psi}_2 = \frac{\mu y_3}{\left(\frac{\rho S C_x y_2^2}{2y_3} + \mu\rho S C_x y_2\right)} \tag{13}$$

$\mu > 0$, $y_3 > 0$, $\rho > 0$, $S > 0$, $C_x > 0$ y $y_2 > 0$, ya que no consideramos el vuelo de regreso, entonces $\widetilde{\Psi}_2 > 0$.

Ahora veamos $\frac{d^2\widetilde{H}_1(t)}{dt}$

$$\frac{d^2\widetilde{H}_1(t)}{dt} = \frac{1}{y_3^2}\left\{\begin{array}{c} -\mu\dot{y}_3 + \dot{\widetilde{\Psi}}_2\left[\frac{\rho S C_x y_2^2}{2y_3} + \mu\rho S C_x y_2\right] + \\ +\widetilde{\Psi}_2\left[\rho S C_x y_2\dot{y}_2 + \mu\rho S C_x\dot{y}_2\right] - \\ -\left[\widetilde{\Psi}_2\left(\mu\rho S C_x y_2 + \frac{\rho S C_x y_2^2}{2y_3}\right) - \mu y_3\right]\frac{2}{y_3^2}\dot{y}_3 \end{array}\right\} = 0 \tag{14}$$

sustituyendo $\dot{y}_3$, $\dot{\widetilde{\Psi}}_2$ y (13) en (14) obtenemos:

$$\begin{aligned} 0 \equiv & \left(-1 + \frac{\widetilde{\Psi}_2\rho S C_x y_2}{y_3}\right)\left[\frac{\rho S C_x y_2^2}{2y_3} + \mu\rho S C_x y_2\right] + \\ & +\widetilde{\Psi}_2\left[\rho S C_x y_2 + \mu\rho S C_x\right]\left[-g - \frac{\rho S C_x y_2^2}{2y_3}\right] + \\ & \left(\frac{\mu\widetilde{\Psi}_2}{y_3}\left[\rho S C_x y_2 + \mu\rho S C_x\right] + \mu\right)u \end{aligned} \tag{15}$$

entonces

$$\frac{\partial}{\partial u}\frac{d^2\widetilde{H}_1(t)}{dt} = \frac{\mu\widetilde{\Psi}_2}{y_3}\left[\rho S C_x y_2 + \mu\rho S C_x\right] + \mu$$

y como, $\mu > 0$; $y_3 > 0$; $\rho > 0$; $S > 0$; $C_x > 0$, $y_2 > 0$, $\widetilde{\Psi}_2 > 0$

$$\left(\frac{\mu\widetilde{\Psi}_2}{y_3}\left[\rho S C_x y_2 + \mu\rho S C_x\right] + \mu\right) = \frac{\partial}{\partial\mu}\left(\frac{d^2 H_1}{dt^2}\right) > 0 \tag{16}$$

lo que quiere decir, por la condición de H.J. Kelly que podemos tener un control optimal singular $u_{\text{opt.singular}}$.

Además de la condición γ) del Principio del máximo, $H_0 = 0$. De esta condición y de (13) obtenemos

$$y_3 = \frac{\rho S C_x}{2\mu g}\left[y_2^3 + \mu y_2^2\right] \tag{17}$$

Y de (15), podemos despejar $u_{\text{opt.singular}}$, obteniendo:

$$u_{\text{opt.singular}} = -\frac{\left(-1+\frac{\widetilde{\Psi}_2\rho SC_x y_2}{y_3}\right)\left[\frac{\rho SC_x y_2^2}{2y_3}+\mu\rho SC_x y_2\right]}{\left(\frac{\mu\widetilde{\Psi}_2}{y_3}\left[\rho SC_x y_2+\mu\rho SC_x\right]+\mu\right)} - \\ -\frac{\widetilde{\Psi}_2\left[\rho SC_x y_2+\mu\rho SC_x\right]\left[-g-\frac{\rho SC_x y_2^2}{2y_3}\right]}{\left(\frac{\mu\widetilde{\Psi}_2}{y_3}\left[\rho SC_x y_2+\mu\rho SC_x\right]+\mu\right)} \tag{18}$$

Y como $\widetilde{\Psi}_2$ y y_3 están en función de y_2 solamente, entonces $u_{\text{opt.singular}}$ depende sólo de y_2.

Entonces finalmente, hemos hallado el control que conducirá al sistema hacia el objetivo, con un gasto mínimo de combustible.

$$u^0 = \left\{ \begin{array}{ll} u_{\max} & \text{si } t \in [0,\tau] \\ u_{\text{opt.singular}} & \text{si } t \in [\tau, t_k] \end{array} \right\}$$

4. Resultados Numéricos

Vamos a considerar que la altura que se desea alcanzar es $h_k = 10,000$ metros.

Para realizar la simulación se usan los siguientes parámetros tomados de [**?**] .

$$\begin{aligned} C_x &= 0.05 \\ g &= 9.8\frac{m}{s} \\ \rho &= 1.22\frac{kg}{m^3} \\ M_0 &= M_{estructura} + M_{0combustible} = 150kg + 300kg = 450kg \\ S &= 0.16m^2 \\ U_{\max} &= 8.8235\frac{kg}{s} \\ \mu &= 1700\frac{m}{s} \end{aligned}$$

Con estos valores

$$\begin{aligned} \frac{\rho SC_x}{2} &= 0.0048\frac{kg}{m} \\ \frac{\rho SC_x}{2\mu} &= 2.823\times 10^{-6}\frac{kgs^3}{m^3} \\ \frac{\rho SC_x}{2\mu g} &= 2.87\times 10^{-7}\frac{kgs^3}{m^3} \end{aligned}$$

Primero resolvemos numéricamente las ecuaciones diferenciales (3) usando $u_{\max}$ y hagamos una gráfica de masa contra velocidad y veamos en qué momento

τ, aproximadamente, se intersecta con $y_3 = \frac{\rho S C_x}{2\mu g}\left[y_2^3 + \mu y_2^2\right]$, la masa que corresponde al control $u_{\text{opt.singular}}$, pues éste será el momento en que comienza actuar $u_{\text{opt.singular}}$. Las ecuaciones (3) usando $u_{\max}$ se resuelven numéricamente en MATLAB.

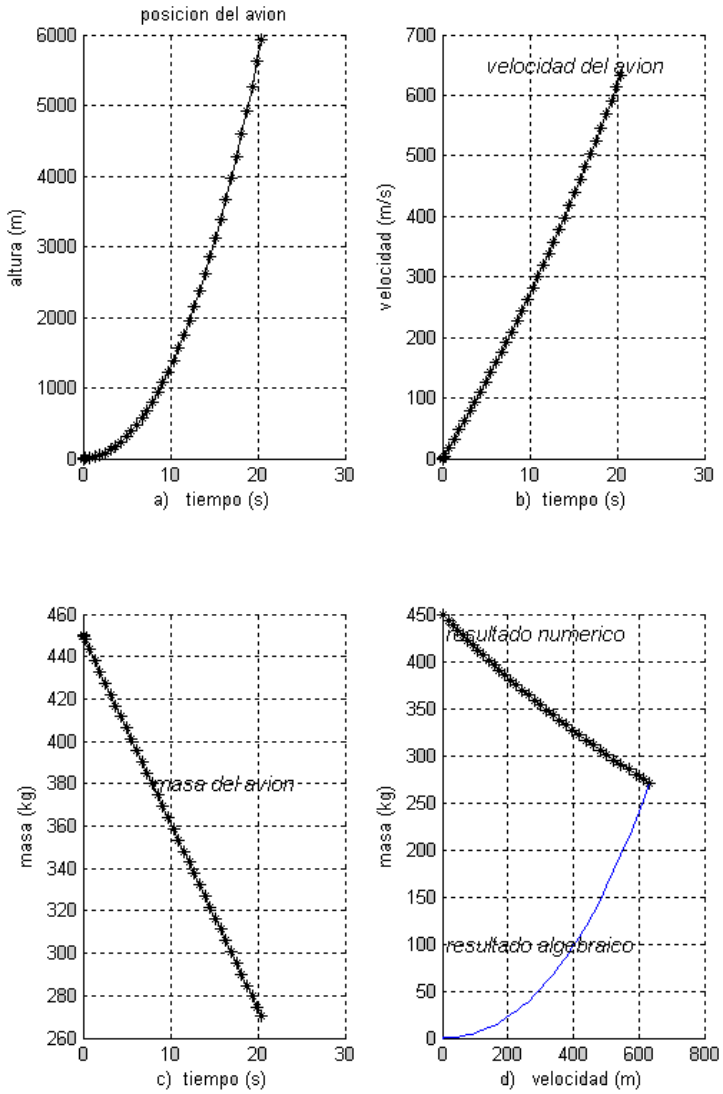


Figura2. Estado del sistema

Como podemos ver de la Figura 2, el momento en que comienza actuar $u_{\text{opt.singular}}$ es $\tau = 20.44$ segundos.

Ahora resolviendo las ecuaciones (3) usando $u_{\text{opt.singular}}$, pues ya tenemos

$\tau = 20.44$, y tambièn $y_1(\tau)$, $y_2(\tau)$, $y_3(\tau)$. En la Figura 3, se muestran las gráficas desde que comienza el movimiento hasta que se alcanza la altura $h_k = 10,000m$.

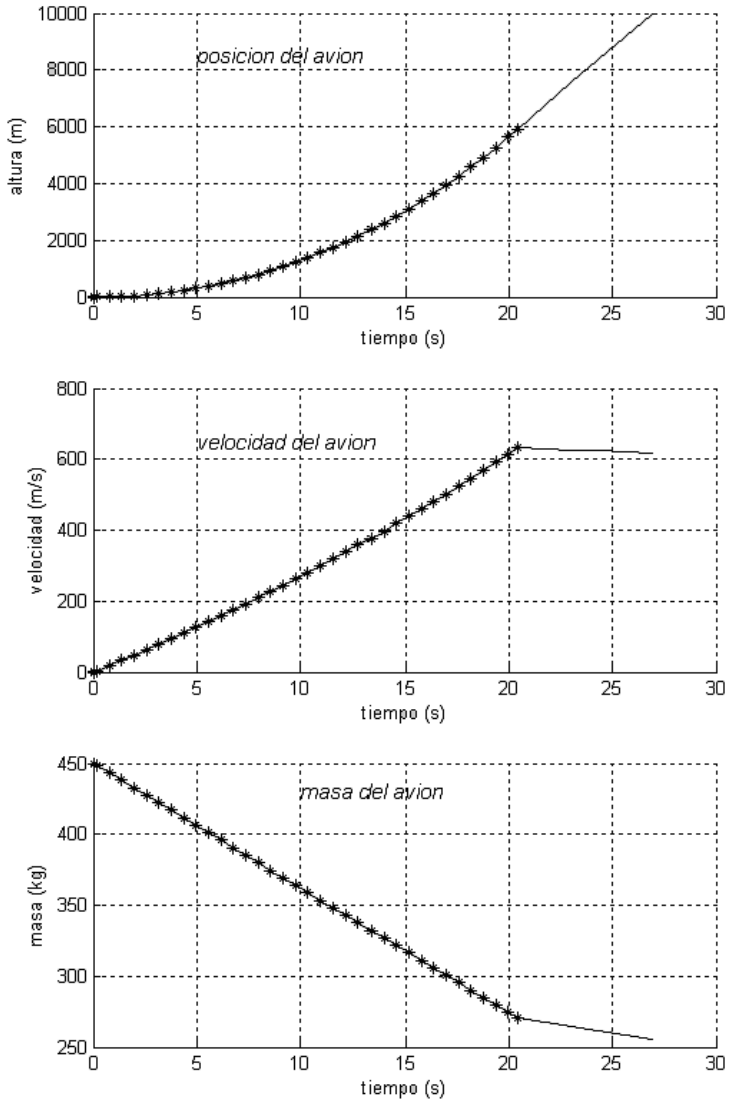


Figura 3. La trayectoria marcada con * corresponde a $u_{\max}$. La trayectoria sólida corresponde a $u_{\text{opt.singular.}}$

Observamos de la figura 3 que el objetivo se alcanza en $t = 27$ segundos aproximadamente.

Una gráfica de la masa como función de la velocidad se muestra en la Figura 4.

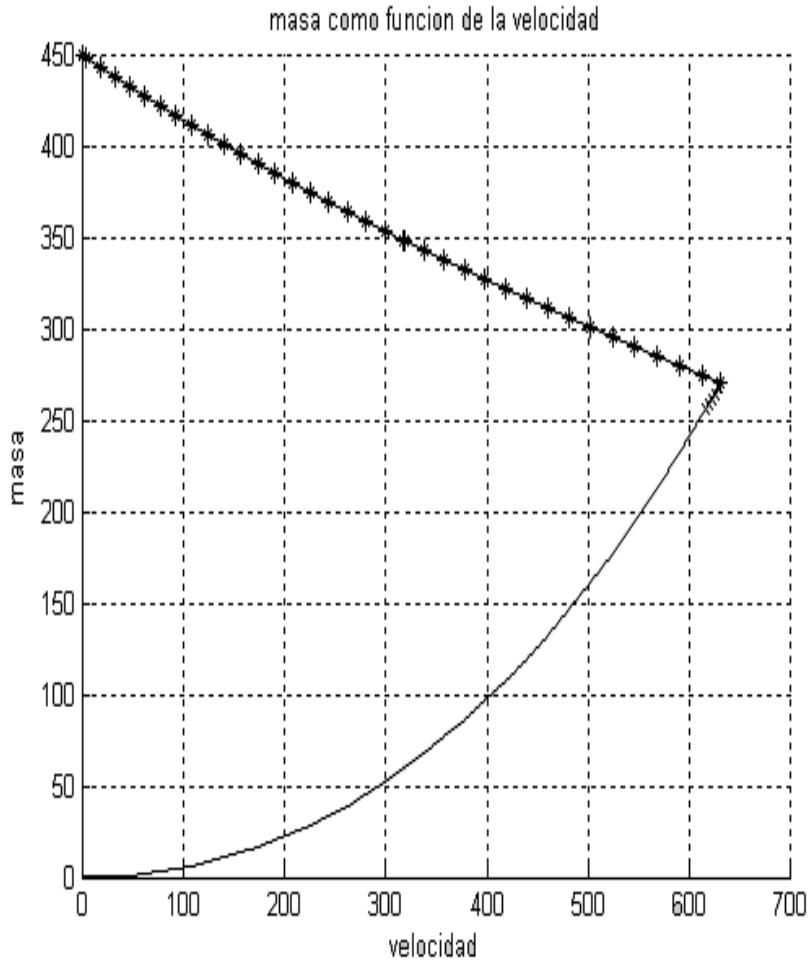


Figura 4. La trayectoria marcada con * corresponde a $u_{\max}$. La trayectoria sólida corresponde al resultado algebraico correspondiente a $u_{\text{opt.singular}}$. La trayectoria marcada con $\times$ corresponde al resultado numérico correspondiente a $u_{\text{opt.singular}}$.

Una gráfica del control como función del tiempo se muestra en la figura 5.

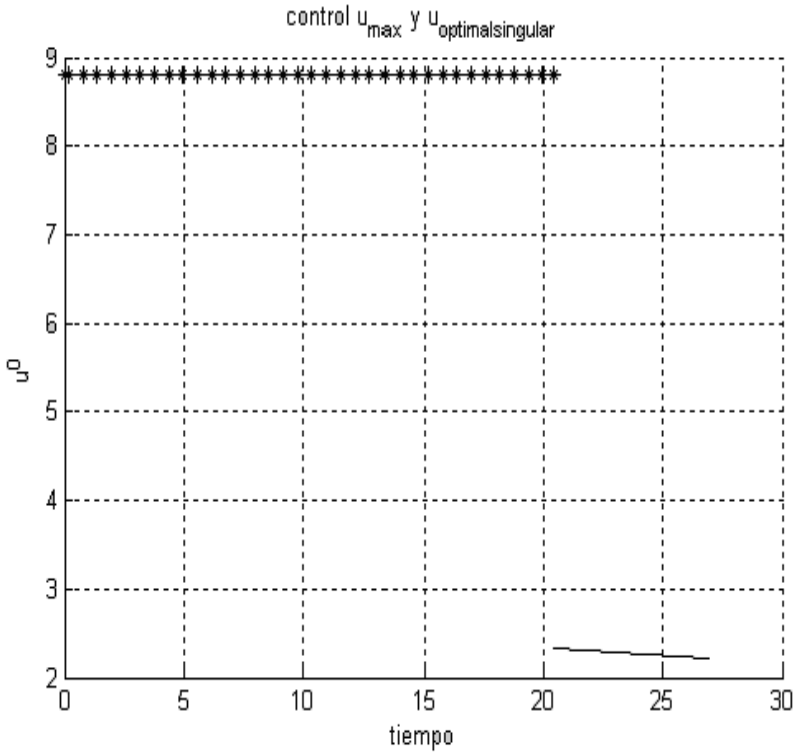


figura 5. Control como función del tiempo. La curva marcada con * corresponde a $u_{\max}$. La línea sólida corresponde a $u_{\text{opt.singular}}$.

5. Conclusiones

Los resultados obtenidos al aplicar **El Principio del Máximo de Pontryagin** y la **Condición necesaria de H. J. Kelly** al problema de lanzamiento vertical de un avión automático se listan como sigue:

1. La aplicación de **El Principio del Máximo de Pontryagin**" y la **Condición necesaria de H. J. Kelly** ayudaron a encontrar satisfactoriamente el control optimal que conduce al avión a la altura h_k con mínimo de combustible gastado.

2. El control optimal que se encontró consiste de dos controles $u_{\max}$ y $u_{\text{opt.singular}}$ que actúan en diferentes intervalos de tiempo.

3. El control optimal como función del tiempo, como se ve en la figura 5, contiene una singularidad.

4. Los resultados numéricos coinciden con el resultado algebraico como lo muestra la gráfica de la masa como función de la velocidad.

5. Si la altura que se desea alcanzar es menor que 5000 metros, $u_{\text{opt.singular}}$ no se tiene.

Referencias

[1] ROBERT C. NELSON. Flight Stability and Automatic Control, Mc Graw-Hill, segunda edición.

[2] T. L. VINCENT, W. J. GRANTHAM, Nonlinear and Optimal Control Systems, Jhon Wiley & Sons,Inc. 1997.

[3] N. F. KRASNOV. Aerodinámica en Preguntas y problemas, editorial Mir Moscú.

[4] J. MACKI, A. STRAUSS, Introduction to Optimal Control Theory, Springer-Verlag 1982.

SOBRE UN PROBLEMA INVERSO PARA UNA ECUACIÓN PARABÓLICA FUERTEMENTE DEGENERADA.

S. BERRES[A], R. BÜRGER[B], A. CORONEL[C], AND M. SEPÚLVEDA[B]

[A]Institut für Angewandte Analysis und Numerische Simulation, Universität Stuttgart,
Pfaffenwaldring 57, D-70569 Stuttgart, Germany.
E-mail: berres@mathematik.uni-stuttgart.de.

[B]Departamento de Ingeniería Matemática, Universidad de Concepción,
Casilla 160-C, Concepción, Chile.
E-mail: rburger@ing-mat.udec.cl, mauricio@ing-mat.udec.cl

[C]Departamento de Ciencias Básicas, Facultad de Ciencias, Universidad del Bío-Bío,
Casilla 447, Campus Fernando May, Chillán, Chile.
E-mail: acoronel@roble.fdo-may.ubiobio.cl

Resúmen. En este trabajo se estudia un problema inverso para una ecuación parabólica fuertemente degenerada que modela la separación de una mezcla de sólido y fluido por centrifugación. El problema inverso (PI) consiste en la determinación de los coeficientes en la ecuación diferencial que gobierna el proceso a partir de mediciones de la concentración de sólidos que es la variable cuya evolución es descrita por el modelo o problema directo. El PI es formulado como un problema de minimización para una adecuada función de costo que compara la solución del modelo con las observaciones. Se demuestra un resultado de continuidad de la solución entrópica con respecto a los coeficientes que implica la existencia de soluciones del PI respectivo. La obtención de los puntos estacionarios de la función costo son obtenidos por un método de descenso donde el gradiente es formalmente calculado a través de una formulación Lagrangiana que lleva a la introducción de un estado adjunto dado por un problema retrógrado con valores en la frontera para una ecuación diferencial parabólica lineal fuertemente degenerada y con coeficientes discontinuos. Haciendo un cálculo similar a esta deducción formal del gradiente, se obtiene un método que permite calcular de manera eficiente el gradiente exacto para el modelo discretizado. Este método es utilizado para la identificación de los parámetros que intervienen en el flujo de densidad y el coeficiente de difusión, a través de las relaciones constitutivas propias del modelo.

Palabras Clave: identificación de parámetros, problemas inversos, leyes de conservación, parabólicas fuertemente degeneradas.

Trabajo a ser presentado en Segunda Gran Semana Nacional de la Matemática, a llevarse a cabo en Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla, Puebla-Oaxaca, México, 23 al 27 de octubre de 2006. Esta investigación fue parcialmente financiada por Conicyt a través del Fondap en Matemáticas Aplicadas, Fondecyt 1030718, Fondecyt 1050728 y la Universidad del Bío- Bío a través del proyecto interno 055409 1/R..

1. Introducción

La sedimentación es un proceso mecánico para la separación de una mezcla. Su gran utilidad en los procesos industriales lo convierten en un fenómeno relevante para la investigación científica (ver [11, 8]). Los principales aspectos históricos de la evolución de esta teoría aparecen detallados en [6] y [12]. En estos trabajos, se establece como modelo matemático para la descripción del fenómeno de separación sólido-fluido, una ecuación parabólica fuertemente degenerada. Los términos convectivos y difusivos que interviene en este modelo parabólico degenerado, se determinan en la práctica a través de hipótesis constitutivas dadas por expresiones que dependen de un número finito de parámetros. Sin embargo, su identificación y análisis matemático requieren el estudio de un problema inverso que consiste en la determinación de los coeficientes del modelo basándose en un conocimiento de las condiciones iniciales, de frontera y perfiles observados de la solución.

El caso de interés para el presente trabajo es el modelo matemático para la sedimentación por efecto de fuerzas centrífugas o centrifugación. En la Figura 1 se muestran dos dispositivos considerados. Para distinguir los dos casos, se introduce el parámetro σ que toma los valores $\sigma = 0$ y $\sigma = 1$. La única coordenada espacial es el radio r, que varía entre el radio interior $R_0 > 0$ y el radio exterior $R > R_0$, correspondiendo al menisco de la suspensión y a la pared exterior, respectivamente. Resumiendo, la derivación detallada dada en [1] se obtiene la siguiente ecuación parabólica fuertemente degenerada:

$$\partial_t \phi + \frac{1}{r^\sigma}\partial_r\left(-\frac{\omega^2}{g} r^{1+\sigma} f_{\mathrm{bk}}(\phi)\right) = \frac{1}{r^\sigma}\partial_r\big(r^\sigma \partial_r A(\phi)\big), \quad (r,t) \in Q_T, \tag{1.1}$$

donde ϕ es la concentración del sólido como una función del radio y el tiempo, $Q_T := (R_0, R) \times (0, T)$ y g es la aceleración de la gravedad. Las funciones $f_{\mathrm{bk}}(\phi)$ y $A(\phi)$ son la función de flujo de Kynch y el coeficiente de difusión integrada, respectivamente, que modelan el transporte y la compresibilidad de la suspensión, respectivamente. Se asume que $f_{\mathrm{bk}}(\phi)$ es una función Lipschitz continua que satisface $f_{\mathrm{bk}}(\phi) = 0$ para $\phi \leqslant 0$ y $\phi \geqslant \phi_{\max}$, donde $\phi_{\max}$ es la concentración máxima del sólido, y $f_{\mathrm{bk}}(\phi) < 0$ para $0 < \phi < \phi_{\max}$. La función $A(\phi)$ es dada por

$$A(\phi) = \int_0^\phi a(s)\, ds, \quad a(\phi) := -\frac{f_{\mathrm{bk}}(\phi)\sigma_{\mathrm{e}}'(\phi)}{\Delta\varrho g \phi},$$

donde $\Delta\varrho$ es la diferencia de la densidad del sólido y el fluido, σ_{e} es el esfuerzo efectivo de sólidos, y σ_{e}' es su derivada. Se asume que el esfuerzo efectivo de sólidos es cero en la zona de transporte y las partículas no están en contacto, es decir cuando ϕ es menor que la concentración crítica ϕ_{c}, y es una función estrictamente creciente de ϕ para $\phi > \phi_{\mathrm{c}}$, es decir, se tiene

$$\sigma_{\mathrm{e}}(\phi) \begin{cases} = 0 & \text{para } \phi \leqslant \phi_{\mathrm{c}}, \\ > 0 & \text{para } \phi > \phi_{\mathrm{c}}, \end{cases} \qquad \sigma_{\mathrm{e}}'(\phi) \begin{cases} = 0 & \text{para } \phi \leqslant \phi_{\mathrm{c}}, \\ > 0 & \text{para } \phi > \phi_{\mathrm{c}}. \end{cases}$$

Combinando las hipótesis sobre f_{bk} y sobre σ_{e}, se observa que $a(\phi) = 0$ para $\phi \leqslant \phi_{\mathrm{c}}$ y $\phi \geqslant \phi_{\max}$ and $a(\phi) > 0$ para $\phi_{\mathrm{c}} < \phi < \phi_{\max}$. Así, (1.1) es una ecuación diferencial parcial hiperbólica de primer orden para $\phi \leqslant \phi_{\mathrm{c}}$ y $\phi \geqslant \phi_{\max}$ y una ecuación diferencial parcial parabólica para $\phi_{\mathrm{c}} < \phi < \phi_{\max}$. Debido a que la degeneración de parabólica en hiperbólica sucede en intervalo de la solución de longitud positiva, (1.1) es llamada *parabólica fuertemente degenerada*.

En este trabajo, se limita la discusión a dos formas paramétricas de funciones que modelan f_{bk} y σ_{e}. De acuerdo a las fórmulas introducidas por Michaels y Bolger [28] (donde $0 < \phi_{\mathrm{m}} \leqslant \phi_{\max}$) y Richardson-Zaki [29] (donde $\phi_{\mathrm{m}} = 1$), f_{bk} es dado por

$$f_{\mathrm{bk}}(\phi) = \begin{cases} u_\infty \phi (1 - \phi/\phi_{\mathrm{m}})^C & \text{para } 0 < \phi < \phi_{\max}, \\ 0 & \text{para } \phi \leqslant 0 \text{ and } \phi \geqslant \phi_{\max}, \end{cases} \qquad u_\infty < 0,\ C \geqslant 1. \tag{1.2}$$

En tanto que σ_{e} es definida por una modelo potencial [30]

$$\sigma_{\mathrm{e}}(\phi) = \begin{cases} 0 & \text{para } \phi \leqslant \phi_{\mathrm{c}}, \\ \sigma_0\big((\phi/\phi_{\mathrm{c}})^k - 1\big) & \text{para } \phi > \phi_{\mathrm{c}}, \end{cases} \qquad \sigma_0 > 0,\ k \geqslant 1. \tag{1.3}$$

Para completar el modelo de centrifugación de una suspensión de concentración inicial $\phi_0 = \phi_0(r)$ dada por (1.1) junto con la condición inicial

$$\phi(r,0) = \phi_0(r), \quad r \in [R_0, R], \tag{1.4}$$

donde se asume que $\phi_0(r) \in [0, \phi_{\max}]$ para todo $r \in [R_0, R]$, y las condiciones de frontera cinemáticas

$$\left(\frac{\omega^2 r_{\rm b}}{g} f_{\rm bk}(\phi) + \partial_r A(\phi)\right)(r_{\rm b}, t) = 0, \quad t > 0,\ r_{\rm b} \in \{R_0, R\}, \tag{1.5}$$

que expresan que el flujo a través de $r = R_0$ y $r = R$ es cero.

El análisis matemático de las ecuaciones parabólicas fuertemente degeneradas, ha recibido en las últimas décadas una atención especial ya que su nolinealidad y cambio de comportamiento de parabólico a hiperbólico hacen que la solución se comporte como en las Leyes de Conservación nolineales, es decir se observa la formación de discontinuidades (choques) o una regularización (ondas de rarefacción) en la solución independientemente de la regularidad de las condiciones iniciales y de frontera impuestas (ver [9, 20, 25]). De esta manera las soluciones para (1.1) deben ser entendidas en el sentido de la entropía de Kružkov.

Un análisis para el problema de Cauchy asociado a la ecuación (1.1) fue hecho por Carrillo en un extenso y profundo trabajo (ver [9]). Este artículo establece aspectos intrínsicos del comportamiento de la solución bajo hipótesis generales sobre los coeficientes. A pesar de la generalidad en la que se enuncian estos resultados su adaptación al problema con valores en la frontera no es directa, apareciendo complejidades meritorias de un análisis especial. Bürger y coautores en [2], siguiendo el trabajo de Carrillo (ver [9]), realizaron el estudio del modelo de sedimentación bajo efectos de la gravedad solamente. El estudio de existencia, unicidad y estabilidad con respecto a la condición inicial del problema de centrifugación fue hecho en [4], siguiendo también la técnica de [9, 2].

El método de Volúmenes Finitos es la herramienta natural para la simulación numérica de Leyes de Conservación (ver [21]). Así, en el caso de la ecuación (1.1) resulta también ser la técnica numérica mas adecuada. La discretización de (1.1) considerada en este artículo son un esquema explícita y un implícito. En ambos casos se considera la discretización en el interior del dominio físico y se incorporan adecuadamente las condiciones de frontera. El flujo numérico (para la convección) utilizando en las simulaciones numéricas es el de Engquist-Osher (ver [18]), cuya evidencia de buen comportamiento y coherencia para reflejar el fenómeno modelado es presentada en los trabajos de R. Bürger y colaboradores, principalmente en [5], donde fue introducido.

En el proceso de simulación es imprescindible determinar valores numéricos para los distintos parámetros físicos que intervienen en los términos convectivo y difusivo de la ecuación parabólica degenerada. A pesar que experimentalmente o mediante tablas se puede obtener aproximaciones, la determinación de estos datos, o valores numéricos de los parámetros, es por lo general un problema difícil. Así, la calibración y validación de estos parámetros se hace a partir de cuán cerca está la solución obtenida mediante simulación numérica del modelo, de una determinada observación experimental. El tratamiento matemático se traduce en un problema de identificación de parámetros que es una situación particular del siguiente problema inverso :

PI. *Dadas las funciones ϕ_0, las condiciones de frontera y un conjunto de mediciones experimentales $\hat{\phi}(r,t)$ encontrar $f_{\rm bk}, A$ de tal manera que la solución entrópica de (1.1) esté lo "mas cercana posible" de $\hat{\phi}(x,t)$.*

La formulación PI establecida para el problema inverso es equivalente al problema de optimización

$$\min_{f_{\rm bk}, A} \mathscr{J}(\phi, \hat{\phi}), \quad \text{sujeto a que } u \text{ sea solución débil de (1.1)}, \tag{1.6}$$

donde la función costo $\mathscr{J}$ se escoge para dar precisión matemática (analítica y numérica) del término ambiguo "mas cercana posible". El funcional natural $\mathscr{J}$ y que será considerado a lo largo de este trabajo es el de mínimos cuadrados, el cual compara las funciones ϕ y $\hat{\phi}$ en la norma L^2. El planteamiento (1.6) de PI permite estudiar su existencia-unicidad y formular técnicas para su solución.

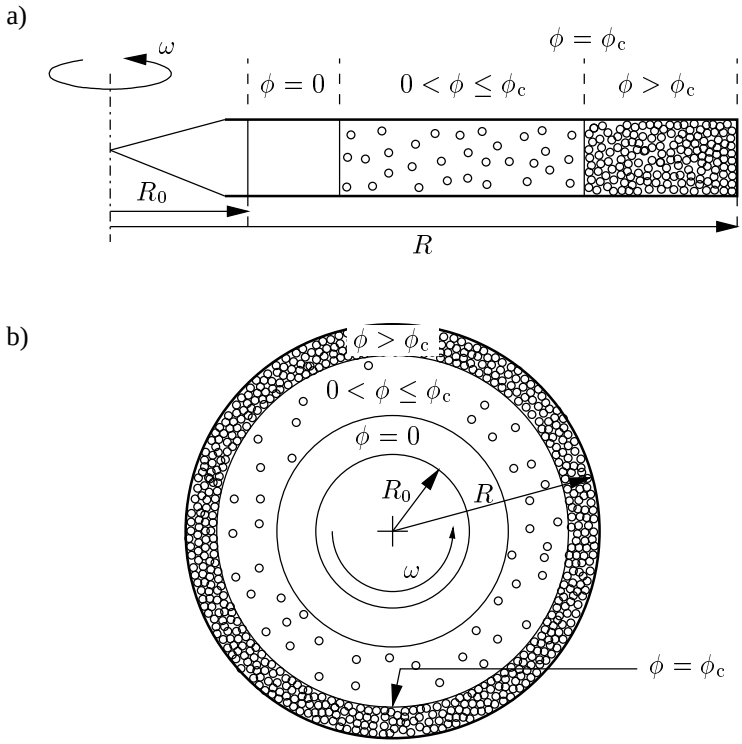


FIGURA 1. (a) Tubo giratorio con sección transversal constante. ($\sigma = 0$), (b) Cilindro giratorio axisimétrico ($\sigma = 1$). Las zonas de concentración son el líquido claro ($\phi = 0$), la zona de la mescla ($0 < \phi \leqslant \phi_c$) y la zona de compresión $\phi > \phi_c$.

La existencia de soluciones para (1.6) está basado en la dependencia continua de la solución entrópica con respecto a los coeficientes. Sin embargo, la unicidad de soluciones para (1.6), es un problema mal-puesto, pudiéndose construir ejemplos explícitos para el caso de las Leyes de Conservación (ver [24]).
La técnica utilizada en esta trabajo, para la solución numérica de (1.6) está basada en el método del gradiente. La nolinealidad de la ecuacion (1.1) implica que su solución no dependa explícitamente de los coeficientes, dificultando de este modo la obtención del gradiente. Sin embargo, haciendo cálculos muy similares a los realizados por James y Sepúlveda para el caso de la Leyes de Conservación (ver [24]) es posible la deducción formal de un gradiente continuo para $\mathscr{J}$. Este procedimiento es detallado en las sección 4 y fue recientemente introducido por los autores y colaboradores en [13, 14, 15, 16, 17].
El trabajo es organizado como sigue: En la sección 2 se presenta el análisis del buen planteamiento del problema directo, en la sección 3 se analiza la existencia de soluciones para el problema inverso, en la sección 4 se presenta el método de aproximación del problema inverso y finalmente en la sección 5 se presenta un ejemplo numérico para el esquema de identificación.

2. Soluciones entrópicas del problema directo

Para el análisis del problema con valores iniciales y en la frontera (1.1), (1.4), (1.5), es conveniente estudiar la siguiente forma general

$$\partial_t \phi + \partial_r f(\phi, r) = \partial_r^2 A(\phi) + g(\phi, r), \tag{2.1}$$

es decir, ecuaciones en forma conservativa y con término fuente, que contienen a (1.1) si se escoge

$$f(\phi, r) := -\frac{\omega^2 r}{g} f_{\mathrm{bk}}(\phi) - \frac{\sigma}{r} A(\phi), \quad g(\phi, r) := \sigma \left[\frac{\omega^2}{g} f_{\mathrm{bk}}(\phi) + \frac{A(\phi)}{r^2} \right]. \tag{2.2}$$

Para el análisis que se presenta es necesario observar que $f(\phi, r)$ y $g(\phi, r)$ se escriben como

$$f(\phi, r) = k_1(r) f^1(\phi) + k_2(r) f^2(\phi), \tag{2.3}$$

$$k_1(r) := -\omega^2 r/g,\ k_2(r) := -\sigma/r,\ f^1(\phi) := f_{\mathrm{bk}}(\phi),\ f^2(\phi) := A(\phi), \tag{2.4}$$

$$g(\phi, r) = g^1(\phi) + k_3(r) g^2(\phi),$$

$$k_3(r) := \sigma/r^2,\ g^1(\phi) := \frac{\sigma\omega^2}{g} f_{\mathrm{bk}}(\phi),\ g^2(\phi) := A(\phi). \tag{2.5}$$

En esta notación, las condiciones de frontera (1.5) toman la siguiente forma

$$\big(f(\phi, r_{\mathrm{b}}) - \partial_r A(\phi)\big)(r_{\mathrm{b}}, t) = -\frac{\sigma}{r_{\mathrm{b}}} A\big(\phi(r_{\mathrm{b}}, t)\big), \quad t > 0,\ r_{\mathrm{b}} \in \{R_0, R\}. \tag{2.6}$$

Es conocido que las soluciones del problema con valores iniciales y en la frontera (1.1), (1.4), (1.5) (o equivalentemente, el problema con valores iniciales y en la frontera (1.4), (2.1), (2.6)) presentan discontinuidades debido a la nolinealidad del flujo y a la degeneración del término de difusión, y tienen que ser caracterizadas como soluciones débiles. Para garantizar la unicidad de soluciones débiles se tienen que definir las soluciones entrópicas.

Definición 2.1. *Una función $\phi \in L^\infty(Q_T) \cap BV(Q_T)$ es una* solución entrópica *del problema con valores iniciales y en la frontera (1.4), (2.1), (2.6) si se satisfacen las siguientes condiciones:*

(1) *El coeficiente integrado tiene la siguiente regularidad $\partial_r A(\phi) \in L^2(Q_T)$.*
(2) *La condición de frontera (2.6) es válida en el siguiente sentido:*

$$\gamma(r_{\mathrm{b}}, t)\bigg(f(\phi, r_{\mathrm{b}}) - \partial_r A(\phi) + \frac{\sigma}{r_{\mathrm{b}}} A\big(\phi(r_{\mathrm{b}}, t)\big)\bigg) = 0, \quad t > 0,\ r_{\mathrm{b}} \in \{R_0, R\}. \tag{2.7}$$

donde $\gamma(\cdot, t)$ es el operador de trace.
(3) *La condición inicial (1.4) es válida en el siguiente sentido:*

$$\lim_{t \downarrow 0} \phi(r, t) = \phi_0(r) \quad \text{para casi todo } r \in (R_0, R).$$

(4) *La siguiente condición de entropía es válida:*

$$\forall \varphi \in C_0^\infty(Q_T),\ \varphi \geqslant 0, \forall k \in \mathbb{R} : \int\int_{Q_T} \Big\{ |\phi - k| \partial_t \varphi + \operatorname{sgn}(\phi - k)$$
$$\times \Big[\big(f(\phi, r) - f(k, r) - \partial_r A(\phi)\big) \partial_r \varphi + \big(f_r(k, r) - g(\phi, r)\big) \varphi \Big] \Big\}\, dt\, dr \geqslant 0.$$

La prueba de existencia y unicidad una solución entrópica del problema directo, siguiendo el método de pseudoviscocidad es presentado en [4]. En dicho trabajo también se presenta un esbozo de la prueba de unicidad, que en particular está relacionado con los resultados de Carrillo [9] que permiten aplicar la técnica de duplicación de variables ("doubling of the variables") introducida por Kružkov (ver [27]) a ecuaciones parabólicas fuertemente degeneradas. Ambas pruebas, existencia y unicidad, son obtenidas con pequeñas modificaciones de las ideas presentadas en [2]. Como resultado se tiene el siguiente teorema.

Teorema 2.1. *El problema con valores iniciales y en la frontera (1.4), (2.1), (2.6) tiene un única solución entrópica.*

3. Existencia de soluciones para el problema inverso.

En esta sección se establece las condiciones suficientes para la existencias de soluciones del problema inverso. La existencia es consecuencia de la dependencia continua de la solución entrópica del problema directo con respecto a las nolinealidades. La dependencia continua para el problema de valores iniciales con flujo que depende de la variable espacial fue estudiado en [19, 26]. Su extensión al al presente problema con valores iniciales y en la frontera es casi directa, siguiendo los trabajos de Carrillo [9] y Cockburn-Gripenberg [10]. La diferencia con el análisis realizado en [13]

consiste en la condición de frontera y la presencia de un término fuente en (2.1). Primeramente se establece el siguiente lema, donde se utiliza la siguiente aproximación de la función signo:

$$\operatorname{sgn}_\varepsilon(x) = \begin{cases} \operatorname{sgn}(x) & \text{para } |x| > \varepsilon, \\ x/\varepsilon & \text{para } x \leqslant \varepsilon. \end{cases}$$

Lema 3.1. *Suponiendo que la función $A(\cdot)$ es suave y satisface $A'(s) > 0$. Entonces la siguiente desigualdad es válida para $\varphi \in C_0^\infty(Q_T)$ con $\varphi \geqslant 0$ y $k \in \mathbb{R}$:*

$$\int\int_{Q_T} \Big\{ |\phi - k| \partial_t \varphi + \operatorname{sgn}(\phi - k)\big(f(\phi, r) - f(k, r) - \partial_r A(\phi)\big)\partial_r \varphi - \operatorname{sgn}(\phi - k) \\ \times \big(\partial_r f(\phi, r) - g(\phi, r)\big)\varphi \Big\}\, dtdr = \lim_{\varepsilon \downarrow 0} \iint_{Q_T} A'(\phi)\big(\partial_r \phi\big)^2 \operatorname{sgn}'_\varepsilon(\phi - k)\varphi \, dtdr. \tag{3.1}$$

Pueba. Como en [13], se define

$$\psi_\varepsilon(z) := -\operatorname{sgn}_\varepsilon\big(A^{-1}(z) - k\big), \quad A_{\psi_\varepsilon}(\phi) := \int_k^\phi \psi_\varepsilon\big(A(s)\big)\, ds.$$

En la prueba de este lema, $\langle \cdot, \cdot \rangle$ denota la usual pariedad entre $H^{-1}(a, b)$ y $C_0^1(R_0, R)$. Entonces la "regla de la cadena débil" ("weak chain rule", ver [9, 26]) implica

$$-\int_0^T \big\langle \partial_t \phi, -\operatorname{sgn}_\varepsilon(\phi - k)\varphi \big\rangle\, dt = \iint_{Q_T} A_{\psi_\varepsilon} \partial_t \varphi \, dt\, dr. \tag{3.2}$$

Por otro lado, de la Definición 2.1 se obtiene

$$-\int_0^T \big\langle \partial_t \phi, \operatorname{sgn}_\varepsilon(\phi - k)\varphi \big\rangle dt + \iint_{Q_T} \Big\{ \big(f(\phi, r) - f(k, r) - \partial_t A(\phi)\big)\partial_r\big(\operatorname{sgn}_\varepsilon(\phi - k)\varphi\big) \\ -\big(\partial_r f(\phi, r) - g(\phi, r)\big) \operatorname{sgn}_\varepsilon(\phi - k)\varphi \Big\} dtdr = 0. \tag{3.3}$$

La desigualdad (3.1) es consecuencia de combinar (3.2) y (3.3) con $\varepsilon \downarrow 0$. □

Teorema 3.1. *Sean u y v dos soluciones entrópicas de los siguientes problemas de valores iniciales y en la frontera*

$$\begin{aligned} \partial_t u + \partial_r f_1(u, r) &= \partial_r^2 A(u) + g_1(u, r), \quad (r, t) \in Q_T, \\ u(r, 0) &= u_0(r), \quad r \in (R_0, R), \\ \big(f_1(u, r_{\mathrm{b}}) - \partial_r A(u)\big)(r_{\mathrm{b}}, t) &= -\frac{\sigma}{r_{\mathrm{b}}} A\big(u(r_{\mathrm{b}}, t)\big), \quad r_{\mathrm{b}} \in \{R_0, R\},\ t > 0 \end{aligned} \tag{3.4}$$

y

$$\begin{aligned} \partial_t v + \partial_r f_2(v, r) &= \partial_r^2 B(v) + g_2(v, r), \quad (r, t) \in Q_T, \\ v(r, 0) &= v_0(r), \quad r \in (R_0, R), \\ \big(f_2(v, r_{\mathrm{b}}) - \partial_r B(v)\big)(r_{\mathrm{b}}, t) &= -\frac{\sigma}{r_{\mathrm{b}}} B\big(v(r_{\mathrm{b}}, t)\big), \quad r_{\mathrm{b}} \in \{R_0, R\},\ t > 0, \end{aligned} \tag{3.5}$$

respectivamente, donde $f_i(u, r) = k_1(r) f_i^1(u) + k_2(r) f_i^2(u)$, $i = 1, 2$, $f_1^2(u) = A(u)$ y $f_2^2(u) = B(u)$, y $k_1(r)$ y $k_2(r)$ son especificados en 2.4. Entonces existen constantes C_1 y C_2 tal que la desigualdad

$$\big\|u(\cdot, t) - v(\cdot, t)\big\|_{L^1} \leqslant \exp(\tilde{C}_3 t)\Big\{ \|u_0 - v_0\|_{L^1} + t\Big[C_1\Big(\big\|f_1^1 - f_2^1\big\|_{\mathrm{Lip}} + \big\|f_1^2 - f_2^2\big\|_{\mathrm{Lip}} \Big) \\ + C_2\Big(\big\|g_1^1 - g_2^1\big\|_{L^\infty[0, \phi_{\max}]} + \big\|g_1^2 - g_2^2\big\|_{L^\infty} \Big) \Big] + C_{\mathrm{D}} \sqrt{t} \big\|\sqrt{a} - \sqrt{b}\big\|_{L^\infty} \Big\} \tag{3.6}$$

es válida para todo $t \in [0, T]$, donde $L^1 = L^1(R_0, R)$, $L^\infty = L^\infty[0, \phi_{\max}]$, $a(u) = A'(u)$, $b(u) = B'(u)$, y

$$\tilde{C}_3 := \big\|g_2^1\big\|_{\mathrm{Lip}} + \|k_3\|_{L^\infty[R_0, R]} \big\|g_2^2\big\|_{\mathrm{Lip}}.$$

Prueba. La prueba es una adaptación del análisis desarrollado por Evje, Karlsen y Risebro [19] y una aplicación de la desigualda de Gronwall. □

El siguiente corolario es una consecuencia del Teorema 3.1.

Corolario 3.1. *La aplicación*

$$\tilde{\mathscr{J}} : \left[\mathrm{Lip} \cap L^{\infty}_{[0,\phi_{\max}]}\right] \times L^{\infty}_{[0,\phi_{\max}]} \times \left[\mathrm{Lip} \cap L^{\infty}_{[0,\phi_{\max}]}\right] =: \mathscr{M} \ni (f, A, g) \mapsto \mathscr{J} \in \mathbb{R}$$

es continua. Además, si $(f, g, A) \in \mathscr{F}$*, donde* $\mathscr{F}$ *es un conjunto compacto de* $\mathscr{M}$*, entonces existe al menos una solución del problema inverso.*

4. Problema inverso como un problema de optimización. Aproximación numérica.

El problema inverso es interpretado como un problema de optimización con restricciones. En particular, para la identificación suponemos que los parámetros de las funciones constitutivas, son agrupados en el vector de parámetros denotado por $\mathbf{e}$, el cual depende de las propiedades del material considerado. El conjunto de datos observados se denota por $\hat{\phi}(r,t)$ y se supone que es una función constantes a trozos de longitud $\Delta r \times \Delta t$, y que son dados sobre la malla estructurada con

$$(r,t) \in \hat{Q} := \{r_1, \ldots, r_{\hat{J}}\} \times \{t_1, \ldots, t_{\hat{N}}\} \subset \overline{Q}_T := [R_0, R] \times [0, T].$$

El objetivo es determinar el vector de parámetros $\mathbf{e}$ para el cual la solución del problema directo, $\phi(r,t)$, da la mejor aproximación de $\hat{\phi}(r,t)$ (es el sentido que se precisará). La solución $\phi = \phi(\mathbf{e})$ depende de los parámetros que se escogen, dado que $f = f(\mathbf{e})$ y $A = A(\mathbf{e})$. Sin embargo, por comodidad notacional, la dependencia de los parámetros tanto en la solución como en las funciones constitutivas no se escribirá explícitamente. Así el problema de identificación de parámetros se puede escribir como el problema de optimización, donde las restricciones son dadas por el problema directo, es decir el problema con valores iniciales y en la frontera (1.1), (1.4), (1.5) en su formulación débil, ver Definición 2.1, esto es:

$$\text{minimizar } \mathscr{J}(\phi) \text{ bajo la restricción } \phi = \phi(\mathbf{e}), \tag{4.1}$$

donde la 'función costo' $\mathscr{J} = \mathscr{J}(\phi)$ mide la calidad de la aproximación. La función costo depende del vector de parámetros $\mathbf{e}$ a través de la solución del modelo. Una elección natural es la distancia L^2 entre el dato observado $\hat{\phi}$ y la solución $\phi = \phi(\mathbf{e})$ de la fución modelo, que da origen a la función costo

$$\mathscr{J}\big(\phi(\mathbf{e})\big) := \frac{1}{2}\int_{\hat{Q}} \big(\phi(r,t) - \hat{\phi}(r,t)\big)^2 \, dt\, dr. \tag{4.2}$$

Dado que las ecuaciones parabólicas fuertemente degeneradas generalmente tienen discontinuidades, la ecuación modelo (1.1) como restricción sobre $\phi = \phi(\mathbf{e})$ es reemplazada por su formulación débil

$$\begin{aligned} E(\phi, p; \mathbf{e}) := &- \int\int_{Q_T} \big(\phi \partial_t p + f(\phi, r)\partial_r p + A(\phi)\partial_r^2 p + g(\phi, r)p\big)\, dt\, dr \\ &+ \int_{R_0}^{R} \phi p\Big|_{t=0}^{T} dr + \int_0^T A(\phi)\left(\partial_r p - \sigma\frac{p}{r}\right)\Big|_{r=R_0}^{R} dt = 0, \end{aligned} \tag{4.3}$$

donde p es la función test.

4.1. Cálculo formal del gradiente.

Este procedimiento presentado con detalle por los autores en [24], se resume en los siguientes tres pasos :

Paso1. Se introduce una formulación lagrangiana de (4.1)-(4.2). En este caso está dada por

$$\mathscr{L}(\phi, p; \mathbf{e}) := \mathscr{J}(\phi) - E(\phi, p; \mathbf{e}), \tag{4.4}$$

donde $\mathbf{e}$ denota el vector de los parámetros en las funciones f, A, que se van a identificar y E la formulación variacional de (1.1) dada en (4.3). La función test p puede ser considerado como un multiplicador de Lagrange generalizado (ver [24]).

Paso2. Se introduce la ecuación adjunta para (1.6). El gradiente de la función costo en (4.4) está dado por

$$\begin{aligned} \frac{d\mathscr{J}(\phi(\mathbf{e}))}{d\mathbf{e}} &= \frac{d\mathscr{L}(\phi(\mathbf{e}), p; \mathbf{e})}{d\mathbf{e}} + \frac{dE(\phi(\mathbf{e}), p; \mathbf{e})}{d\mathbf{e}} \\ &= \left\langle \partial_\phi \mathscr{L}\big(\phi(\mathbf{e}), p; \mathbf{e}\big), \frac{d\phi(\mathbf{e})}{d\mathbf{e}} \right\rangle + \frac{d\mathscr{L}(\phi(\mathbf{e}), p; \mathbf{e})}{d\mathbf{e}}, \end{aligned}$$

debido a que $E(\phi, p; \mathbf{e}) = 0$. En esta derivación formal de la derivada total de la función costo $\partial_\phi \mathscr{L}\big(\phi(\mathbf{e}), p; \mathbf{e}\big)$ no puede ser calculado, dado que la solución del problema directo ϕ no puede ser calculado explícitamente en términos de $\mathbf{e}$. Así, se hace necesaria la introducción de una función test p de modo que $\partial_\phi \mathscr{L}(\phi(\mathbf{e}), p; \mathbf{e}) = 0$, es decir que anula el primer término del lado derecho de (4.5). La derivada de $\mathscr{L}$ en la dirección $\delta\phi$ está dada formalmente por

$$\begin{aligned}
\big\langle \partial_\phi \mathscr{L}(\phi, p; \mathbf{e}), \delta\phi \big\rangle &= \big\langle \partial_\phi \mathscr{J}\,(\phi) - \partial_\phi E(\phi, p; \mathbf{e}), \delta\phi \big\rangle \\
&= \int\int_{Q_T} \delta\phi\big(\phi(r,t) - \hat{\phi}(r,t)\big)\delta_{(r,t)\in\hat{Q}}\, dt\, dr \\
&\quad + \int\int_{Q_T} \delta\phi\big(\partial_t p + \partial_\phi f(\phi, r)\partial_r p + \partial_\phi A(\phi)\partial_r^2 p + \partial_\phi g(\phi, r) p\big)\, dt\, dr \\
&\quad - \int_{R_0}^{R} \delta\phi p(r, T)\, dr + \int_0^T \delta\phi \partial_\phi A(\phi)\left(\partial_r p - \sigma\frac{p}{r}\right)\Big|_{r=R_0}^{R} dt.
\end{aligned}$$

A fin de anular esta expresión, basta que p satisfaga la denominada ecuación adjunta, dada por siguiente problema retrógrado con la condición final y de frontera

$$\partial_t p + \partial_\phi f(\phi, r)\partial_r p + \partial_\phi A(\phi)\partial_r^2 p = -(\phi - \hat{\phi})\delta_{(r,t)\in\hat{Q}} - \partial_\phi g(\phi, r) p \quad \text{para } (r,t) \in Q_T, \tag{4.5}$$

$$p(r, T) = 0 \quad \text{para } r \in [R_0, R] \tag{4.6}$$

$$\left(\partial_r p - \sigma\frac{p}{r_{\mathrm{b}}}\right)(r_{\mathrm{b}}, t) = 0 \quad \text{para } t < T, r_{\mathrm{b}} \in \{R_0, R\}. \tag{4.7}$$

Paso3 Se obtiene el gradiente. Siendo p solución del problema adjunto (4.5)-(4.7) se sigue que el gradiente de la $\mathscr{J}$ viene dado por

$$\frac{d\,\mathscr{J}\,(\phi(\mathbf{e}))}{d\mathbf{e}} = \int\int_{Q_T} \left(\frac{df(\phi, r)}{d\mathbf{e}}\partial_r p + \frac{dA(\phi, r)}{d\mathbf{e}}\partial_r^2 p + \frac{dg(\phi, r)}{d\mathbf{e}} p\right) dt\, dr, \tag{4.8}$$

Los Pasos 1 a 3 describen un cálculo formal que permite resolver el problema de minimización asociado al problema inverso. Sin embargo, aparecen dos dificultades. La primera, se presenta a nivel continuo, y es acerca de la validez de este cálculo. En efecto, debido a las discontinuidades de la solución del problema directo, y de las eventuales discontinuidades del problema adjunto, la diferenciabilidad de la función costo en un sentido clásico, no es un problema fácil de determinar, tal como se observa por ejemplo para el caso puramente hiperbólico (ver [24]). La segunda dificultad, es acerca de la discretización del problema (4.5)-(4.7), para el cálculo del gradiente discretizado. El problema (4.5)-(4.7) corresponde a una ecuación lineal de convección-difusión fuertemente degenerada con coeficientes discontinuos, cuya existencia y unicidad de soluciones es hasta ahora un problema abierto. Si bien se puede hacer una analogía al caso hiperbólico, en que se tiene un conocimiento parcial de existencia y unicidad, a través de las soluciones reversibles, el problema de escoger un método numérico correcto que aproxime a la solución, sigue siendo un problema difícil (ver [7, 22]). En la práctica, subsanamos momentáneamente estos dos problemas, repitiendo el cálculo formal para el caso discreto, es decir calculando un esquema adjunto asociado al esquema del problema directo y luego un gradiente de la función costo discretizada que resulta ser un gradiente exacto. La convergencia del gradiente exacto al gradiente del problema continuo (o quizás a algún subgradiente) sigue siendo hasta ahora un problema abierto.

4.2. **Esquema numérico para la identificación.** Se introduce la notación usual de mallas rectangulares sobre Q_T seleccionando $J, N \in \mathbb{N}$ y definiendo $\Delta r := (R - R_0)/J$, $\Delta t := T/N$, $r_j := R_0 + j\Delta r$ y $t_n := n\Delta t$. El esquema numérico para la solución del problema directo es dado por una aproximación por Volúmenes Finitos de (1.1) dada por

$$\phi_j^{n+1} = \phi_j^n - \lambda_j\big(\mathbf{F}_{j+1/2}^n(\mathbf{e}) - \mathbf{F}_{j-1/2}^n(\mathbf{e})\big) + \mu_j\big(\mathscr{A}_{j+1/2}^n(\mathbf{e}) - \mathscr{A}_{j-1/2}^n(\mathbf{e})\big), \tag{4.9}$$

donde $\lambda_j = \mu_j := \Delta t/(r_j^\sigma \Delta r)$, con la siguiente discretización de la condición inicial

$$\phi_j^0 = \phi_j^{\mathrm{init}}, \quad j = 0, \ldots, J \tag{4.10}$$

y las siguientes versiones discretas de las condiciones de frontera (1.5):

$$\lambda_0 \mathbf{F}^n_{-1/2}(\mathbf{e}) - \mu_0 \mathscr{A}^n_{-1/2}(\mathbf{e}) \quad = 0, \tag{4.11}$$

$$\lambda_J \mathbf{F}^n_{J+1/2}(\mathbf{e}) - \mu_J \mathscr{A}^n_{J+1/2}(\mathbf{e}) \quad = 0. \tag{4.12}$$

Se continúa con una formulación Lagrangiana discreta que permite introducir el estado adjunto discreto

$$\begin{aligned} p_j^n \quad = \quad & p_j^{n+1} - \sum_{k=-K}^{K-1} \partial_{\phi_j^n} \mathbf{F}^n_{j+k+1/2}(\mathbf{e}) \big(\lambda_{j+k} p_{j+k}^{n+1} - \lambda_{j+k+1} p_{j+k+1}^{n+1}\big) \\ & + \sum_{\ell=-\overline{K}}^{\overline{K}-1} \partial_{\phi_j^n} \mathscr{A}^n_{j+\ell+1/2}(\mathbf{e}) \big(\mu_{j+\ell} p_{j+\ell}^{n+1} - \mu_{j+\ell+1} p_{j+\ell+1}^{n+1}\big) - \big(\phi_j^n(\mathbf{e}) - \hat{\phi}_j^n\big) \delta_{(j,n)\in \hat{Q}_\Delta} \\ & \qquad \text{para } j = 0, 1, \ldots, J \text{ and } n = N-1, N-2, \ldots, 0 \end{aligned} \tag{4.13}$$

con la condición final $p_j^N = 0$ para $j \in \{0, \ldots, \max(K, \overline{K})\} \cup \{J - \max(K, \overline{K}) + 1, \ldots, J\}$, y se considera la notación convencional $\mathbf{F}^n_{k+1/2} = \mathscr{A}^n_{\ell+1/2} = 0$ para $\ell, k \leqslant -1$ y $\ell, k \geqslant J$. En el siguiente paso se calcula el siguiente gradiente discreto para la función costo

$$\nabla_{\mathbf{e}} \mathscr{J}_\Delta(\mathbf{e}) = \Delta r \Delta t \nabla_{\mathbf{e}} \mathscr{L}_\Delta\big(\phi_\Delta(\mathbf{e}), p_\Delta; \mathbf{e}\big) = -\Delta r \Delta t \nabla_{\mathbf{e}} E_\Delta\big(\phi_\Delta(\mathbf{e}), p_\Delta; \mathbf{e}\big)$$

donde

$$\nabla_{\mathbf{e}} E_\Delta = \sum_{(j,n)\in Q_\Delta} \nabla_{\mathbf{e}} \mathbf{F}^n_{j+1/2}(\mathbf{e}) \big(\lambda_j p_j^{n+1} - \lambda_{j+1} p_{j+1}^{n+1}\big) - \nabla_{\mathbf{e}} \mathscr{A}^n_{j+1/2}(\mathbf{e}) \big(\mu_j p_j^{n+1} + \mu_{j+1} p_{j+1}^{n+1}\big).$$

5. Ejemplo Numérico

En esta sección se presenta un ejemplo numérico, para otros ejemplos ver [13, 14, 15]. Se consideran como observación un perfil de concentración en el tiempo $t = T$ como una función de r. para el caso del tubo giratorio ($\sigma = 0$). Las observaciones son generadas por una simulación del problema directo con los parámetros utilizados por Bürger y Concha [1]. La concentración inicial es homogénea $\phi_0 = 0.07$ sobre el dominio $r \in [R_0, R] = [0.06\,\mathrm{m}, 0.3\,\mathrm{m}]$; y la función de flujo es escogida de acuerdo a (1.2), donde $\phi_\mathrm{m} = 1$, $u_\infty = 0.0001\,\mathrm{m/s}$, $C = 5$, $\phi_{\max} = 0.66$, y la velocidad angular ω es tal que $R\omega^2 = 10000\mathrm{g}$. Además, se considera la ley potencial (1.3) con $\sigma_0 = 5.7\,\mathrm{Pa}$, $k = 9$ y $\phi_\mathrm{c} = 0.1$ para el esfuerzo efectivo de solidos ; y finalmente, la densidad $\Delta\varrho = 1660\,\mathrm{kg/m^3}$ y la aceleración de la gravedad $g = 9.81\,\mathrm{m/s^2}$. El esquema utilizado para la simulación del problema directo es el esquema explícito de Engquist-Osher y de segundo orden y con una parámetros de discretización $J = 200$ y N tal que se cumpla la siguiente condición CFL (cf. [1]):

$$\frac{R\omega^2}{g} \max_{\phi\in[0,\phi_{\max}]} \big|f'_{\mathrm{bk}}(\phi)\big| \frac{\Delta t}{\Delta r} + 2 \max_{\phi\in[0,\phi_{\max}]} a(\phi) \frac{\Delta t}{(\Delta r)^2} < 1. \tag{5.1}$$

Se recuerda que el flujo numérico del esquema de Engquist-Osher [1, 5, 18] viene dado por

$$\mathbf{F}^{\mathrm{EO}}(u, v, r) := f(0) + \int_0^u \max\big\{\partial_s f(s, r), 0\big\} ds + \int_0^v \min\big\{\partial_s f(s, r), 0\big\} ds.$$

Los requerimientos de estabilidad impuestos por (5.1) al esquema explícito implican la necesidad de un refinamiento extremo, es decir con valores de Δt ($\approx (\Delta x)^2$), lo cual incrementa considerablemente el tiempo computacional. Esta desventaja es superada considerando un esquema totalmente implícito, el cual resulta incondicionalmente estable. Asi para la identificación presentad se considera la siguiente descritización implícita y de primer orden de (4.8):

$$\begin{aligned} \phi_j^{n+1} = \phi_j^n - \lambda_j \big(\mathbf{F}^{n+1}_{j+1/2}(\mathbf{e}) - \mathbf{F}^{n+1}_{j-1/2}(\mathbf{e})\big) + \mu_j \big(\mathscr{A}^{n+1}_{j+1/2}(\mathbf{e}) - \mathscr{A}^{n+1}_{j-1/2}(\mathbf{e})\big), \\ \phi_j^0 = \phi_j^{\mathrm{init}}, \quad \lambda_0 \mathbf{F}^{n+1}_{-1/2}(\mathbf{e}) - \mu_0 \mathscr{A}^{n+1}_{-1/2}(\mathbf{e}) = \lambda_J \mathbf{F}^{n+1}_{J+1/2}(\mathbf{e}) - \mu_J \mathscr{A}^{n+1}_{J+1/2}(\mathbf{e}) = 0. \end{aligned} \tag{5.2}$$

La formulación débil $E_\Delta = E_\Delta\big(\phi_\Delta(\mathbf{e}), p_\Delta; \mathbf{e}\big)$ es dada por

$$\begin{aligned} E_\Delta &= \sum_{(j,n)\in Q_\Delta} \Big\{ \phi_j^n (p_j^n - p_j^{n+1}) + \mathbf{F}_{j+1/2}^n (\lambda_j p_j^n - \lambda_{j+1} p_{j+1}^n) - \mathscr{A}_{j+1/2}^n (\mu_j p_j^n - \mu_{j+1} p_{j+1}^n) \Big\} \\ &\quad + \sum_{j=0}^{J} \Big\{ \big[\phi_j^N + \lambda_j \big(\mathbf{F}_{j+1/2}^N - \mathbf{F}_{j-1/2}^N\big) - \mu_j \big(\mathscr{A}_{j+1/2}^N - \mathscr{A}_{j-1/2}^N\big)\big] p_j^N \\ &\quad - \big[\phi_j^0 + \lambda_j \big(\mathbf{F}_{j+1/2}^0 - \mathbf{F}_{j-1/2}^0\big) - \mu_j \big(\mathscr{A}_{j+1/2}^0 - \mathscr{A}_{j-1/2}^0\big)\big] p_j^0 \Big\}. \end{aligned} \tag{5.3}$$

El esquema adjunto y el gradiente en cada uno de los ejemplos se obtienen con la metodología desarrollada en las sección 4.2 y se obtiene el esquema lineal implícito

$$\mathbf{A}_n P^n = P^{n+1}, \quad \text{for } n = N-1, \ldots, 0,$$

con la condición final

$$p_j^N = \frac{\phi_j^N - \hat{\phi}_j^N}{a_{j,j}^N} \qquad \text{para } j \in \{0, 1, \ldots, J\},$$

donde $\mathbf{A}_n$ es la matríz tridiagonal $J \times J$ con entradas

$$\begin{aligned} a_{j,j-1}^n &= \lambda_{j-1} \partial_{\phi_j^n} F_{j-1/2}^n - \mu_{j-1} \partial_{\phi_j^n} \mathscr{A}_{j-1/2}^n, \quad j = 2, \ldots, J, \\ a_{j,j}^n &= 1 + \lambda_j \big(\partial_{\phi_j^n} F_{j+1/2}^n - \partial_{\phi_j^n} F_{j-1/2}^n\big) - \mu_j \big(\partial_{\phi_j^n} \mathscr{A}_{j+1/2}^n - \partial_{\phi_j^n} \mathscr{A}_{j-1/2}^n\big), \\ &\qquad\qquad j = 1, \ldots, J-1, \\ a_{j,j+1}^n &= -\lambda_{j+1} \partial_{\phi_j^n} F_{j+1/2}^n + \mu_{j+1} \partial_{\phi_j^n} \mathscr{A}_{j+1/2}^n, \quad j = 1, \ldots, J-1, \\ a_{0,0}^n &= 1 + \lambda_0 \partial_{\phi_0^n} F_{1/2}^n - \mu_0 \partial_{\phi_0^n} \mathscr{A}_{1/2}^n, \\ a_{J,J}^n &= 1 - \lambda_J \partial_{\phi_J^n} F_{J-1/2}^n + \mu_J \partial_{\phi_J^n} \mathscr{A}_{J+1/2}^n. \end{aligned}$$

El gradiente de la función costo discreta es dada por

$$\nabla_{\mathbf{e}} \mathscr{J}_\Delta(\mathbf{e}) = -\Delta r \nabla_{\mathbf{e}} E_\Delta\big(\phi_\Delta(\mathbf{e}), p_\Delta; \mathbf{e}\big),$$

donde el gradiente de la formulación débil discreta se evalúa mediante

$$\begin{aligned} \nabla_{\mathbf{e}} E_\Delta &= \sum_{(j,n)\in Q_\Delta} \Big\{ \nabla_{\mathbf{e}} \mathbf{F}_{j+1/2}^n (\lambda_j p_j^n - \lambda_{j+1} p_{j+1}^n) - \nabla_{\mathbf{e}} \mathscr{A}_{j+1/2}^n (\mu_j p_j^n - \mu_{j+1} p_{j+1}^n) \Big\} \\ &\quad + \sum_{j=0}^{M} \Big\{ \big[\lambda_j \big(\nabla_{\mathbf{e}} \mathbf{F}_{j+1/2}^N - \nabla_{\mathbf{e}} \mathbf{F}_{j-1/2}^N\big) - \mu_j \big(\nabla_{\mathbf{e}} \mathscr{A}_{j+1/2}^N - \nabla_{\mathbf{e}} \mathscr{A}_{j-1/2}^N\big)\big] p_j^N \\ &\quad - \big[\lambda_j \big(\nabla_{\mathbf{e}} \mathscr{A}_{j+1/2}^0 - \nabla_{\mathbf{e}} \mathscr{A}_{j-1/2}^0\big) - \mu_j \big(\nabla_{\mathbf{e}} \mathscr{A}_{j+1/2}^0 - \nabla_{\mathbf{e}} \mathscr{A}_{j-1/2}^0\big)\big] p_j^0 \Big\}. \end{aligned}$$

El método del gradiente conjugado que se utiliza para la identificación es el de Polak-Ribière y que tiene como vector inicial $\mathbf{e} = (5.5, 6.5, 9.5, 0.08)$. Para resolver el paso de la minimización unidimensional con el algoritmo del gradiente conjugado se emplea el algoritmo de búsqueda lineal de Wolfe tal como es descrito en [23].

Los resultados de la identificación son presentados en la Tabla 1. Se muestran los perfiles identificados para tres distintos tiempos de observación ($\{0.1s, 0.3s, 1.2s\}$) cuyos gráficos son presentados el las Figuras 2, 3 y 4. En dichas figuras se presentan los resultados con varios tamaños de resolución y así se evidencia la convergencia del esquema de identificación cuando la aproximación se incrementa.

Bibliografía

[1] R. Bürger and F. Concha. Settling velocities of particulate systems: 12. batch centrifugation of flocculated suspensions. *Int. J. Mineral Process.*, 63:115–145, 2001.

[2] R. Bürger, S. Evje, and K. H. Karlsen. On strongly degenerate convection-diffusion problems modeling sedimentation-consolidation processes. *J. Math. Anal. Appl.*, 247:517–556, 2000.

[3] R. Bürger, S. Evje, K. H. Karlsen, and K. A. Lie. Numerical methods for the simulation of the settling of flocculated suspensions. *Chem. Eng. J.*, 80:91–104, 2000.

T	J	C	σ_0	k	ϕ_c	L^2 error	rate
0.1	IG	5.500000	6.500000	9.500000	0.080000	1.437×10^{-2}	–
	100	5.500019	6.499972	9.499787	0.103858	3.074×10^{-3}	–
	150	5.499859	6.499976	9.499794	0.106275	2.294×10^{-3}	0.722
	200	5.500190	6.499977	9.499799	0.107336	2.188×10^{-3}	0.164
0.3	IG	5.500000	6.500000	9.500000	0.080000	2.567×10^{-2}	–
	100	5.500066	6.499970	9.499746	0.109268	2.800×10^{-3}	–
	150	5.499963	6.499959	9.499700	0.108991	2.766×10^{-3}	0.030
	200	5.500081	6.499966	9.499729	0.108849	2.766×10^{-3}	0.000
1.2	IG	5.500000	6.500000	9.500000	0.080000	3.933×10^{-2}	–
	100	5.500434	6.499983	9.499801	0.110288	1.427×10^{-3}	–
	150	5.500173	6.499975	9.499766	0.109467	7.190×10^{-4}	1.691
	200	5.500364	6.499981	9.499813	0.109645	6.116×10^{-4}	0.562

TABLA 1. Ejemplo 1: Perfil para el vector de parámetros inical (IG) y perfiles identificados para $T = 0.1$, $T = 0.3$ $T = 1.2$ y la norma L^2 de los errores estimados (donde sea aplicable).

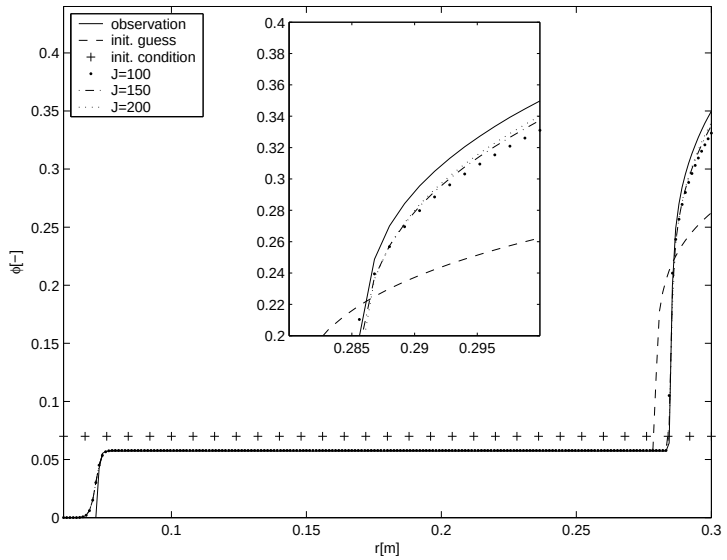


FIGURA 2. Example 1: Comparación entre los perfiles identificados y observados en $T = 0.1$.

[4] R. Bürger and K. H. Karlsen. A strongly degenerate convection-diffusion problem modeling centrifugation of flocculated suspensions. In *Hyperbolic Problems: Theory, Numerics, Applications, Vol. I, II (Magdeburg, 2000)*, volume 141 of *Internat. Ser. Numer. Math., 140*, pages 207–216. Birkhäuser, Basel, 2001.

[5] R. Bürger and K. H. Karlsen. On some upwind difference schemes for the phenomenological sedimentation-consolidation model. *J. Engrg. Math.*, 41:145–166, 2001.

[6] R. Bürger and W. L. Wendland. Sedimentation and suspension flows: historical perspective and some recent developments. *J. Engrg. Math.*, 41:101–116, 2001.

[7] F. Bouchut and F. James. One-dimensional transport equations with discontinuous coefficients. *Nonlinear Anal. TMA*, 32(7):891–933, 1998.

[8] M. C. Bustos, F. Concha, R. Bürger, and E. M. Tory. *Sedimentation and thickening*, volume 8 of *Mathematical Modelling: Theory and Applications*. Kluwer Academic Publishers, Dordrecht, 1999.

[9] J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Rat. Mech. Anal.*, 147:269–361, 1999.

[10] B. Cockburn and G. Gripenberg. Continuous dependence on the nonlinearities of solutions of degenerate parabolic equations. *J. Diff. Ens.*, 151:231–251, 1999.

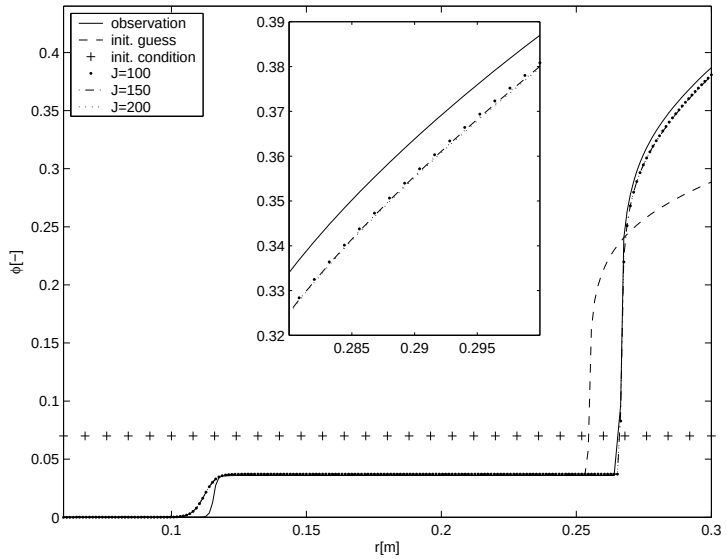


FIGURA 3. Example 1: Comparación entre los perfiles identificados y observados en $T = 0.3$.

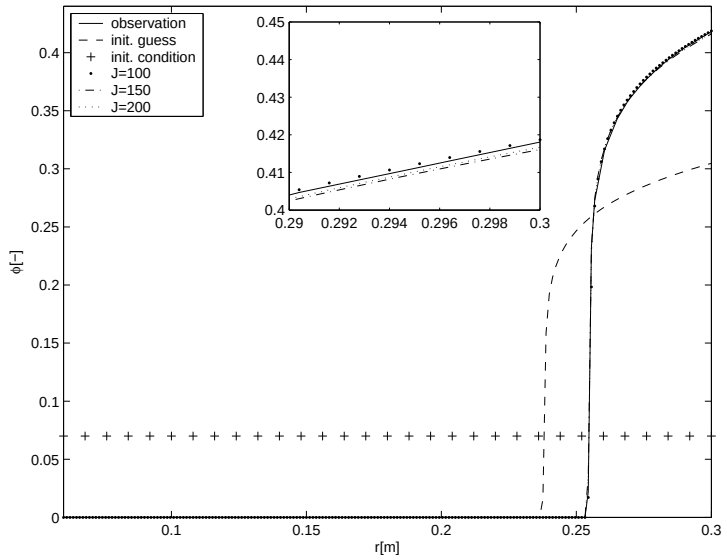


FIGURA 4. Example 1: Comparación entre los perfiles identificados y observados en $T = 1.2$.

[11] F. Concha. *Manual de Filtración y Separación*. Centro de Tecnología Mineral, CETTEM, fconcha@udec.cl, Concepción, Chile, pp. 184, 2001.

[12] F. Concha and R. Bürger. Thickening in the 20th century: a historical perspective. *Minerals & Metallurgical Process.*, 20(2):57–67, 2003.

[13] A. Coronel, F. James, and M. Sepúlveda. Numerical identification of parameters for a model of sedimentation processes. *Inverse Problems*, 19(4):951–972, 2003.

[14] S. Berres, R. Bürger, A. Coronel, and M. Sepúlveda. Numerical identification of parameters for a strongly degenerate convection-diffusion problem modelling centrifugation of flocculated suspensions. *Appl. Numer. Math.*, 52 (2005), pp. 311-337.

[15] S. Berres, R. Bürger, A. Coronel, and M. Sepúlveda. Numerical identification of parameters for flocculated suspension from concentration measurements during batch centrifugation. *Chem. Eng. J.*, 111 (2005), pp. 91-103.

[16] R. Bürger, A. Coronel, and M. Sepúlveda. A semi-implicit monotone difference scheme for an initial- boundary value problem of a strongly degenerate parabolic equation modelling sedimentation-consolidation processes *Math. Comp.*, 75 (2006), 91-112.

[17] R. Bürger, A. Coronel, and M. Sepúlveda. On an upwind difference scheme for strongly degenerate parabolic equations modelling the settling of suspensions in centrifuges and non-cylindrical vessels. *Appl. Numer. Math.*, 56 (2006), 1397-1417.

[18] B. Engquist and S. Osher. One-sided difference approximations for nonlinear conservation laws. *Math. Comp.*, 36:321–351, 1981.

[19] S. Evje, K. H. Karlsen, and N. H. Risebro. A continuous dependence result for nonlinear degenerate parabolic equations with spatially dependent flux function. In *Hyperbolic problems: Theory, Numerics, Applications, Vol. I, II (Magdeburg, 2000)*, volume 141 of *Internat. Ser. Numer. Math., 140*, pages 337–346. Birkhäuser, Basel, 2001.

[20] S. Evje and K. H. Karlsen. Monotone difference approximations of BV solutions to degenerate convection-diffusion equations. *SIAM J. Numer. Anal.*, 37(6):1838–1860 (electronic), 2000.

[21] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 713–1020. North-Holland, Amsterdam, 2000.

[22] L. Gosse and F. James. Numerical approximations of one-dimesional linear conservation equations with discontinuous coefficients *Math. Comp.*, 69:987–1015, 2003.

[23] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms. I*, volume 305 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993. Fundamentals.

[24] F. James and M. Sepúlveda. Convergence results for the flux identification in a scalar conservation law. *SIAM J. Control Optim.*, 37:869–891 (electronic), 1999.

[25] K. H. Karlsen and M. Ohlberger. A note on the uniqueness of entropy solutions of nonlinear degenerate parabolic equations. *J. Math. Anal. Appl.*, 275:439–458, 2002.

[26] K. H. Karlsen and N. H. Risebro. On the uniqueness and stability of entropy solutions of nonlinear degenerate parabolic equations with rough coefficients. *Discrete Contin. Dyn. Syst.*, 9(5):1081–1104, 2003.

[27] S. N. Kružkov. First order quasilinear equations in several independent space variables. *Math. USSR Sb.*, 10:217–243, 1970.

[28] A. S. Michaels and J. C. Bolger. Settling rates and sediment volumes of flocculated Kaolin suspensions. *Ind. Engrg. Chem. Fund.*, 1:24–33, 1962.

[29] J. F. Richardson and W. N. Zaki. Sedimentation and fluidization: Part I. *Trans. Instn. Chem. Engrs. (London)*, 32:35–53, 1954.

[30] F. M. Tiller and W. F. Leu. Basic data fitting in filtration. *J. Chin. Inst. Chem. Engrs.*, 11:61–70, 1980.

Construcciones Geométricas Básicas y un Poco de Arte Óptico

Carrión Cadena Michael Marisela.
Torres Ramírez Luis Alberto.
alberto_tr79@hotmail.com
Instituto ARNAIZ
Blvd. De las Torres No. 631
Colonia Loma Linda
Puebla, Pue.

Resumen

Uno de los principales temas a desarrollar en segundo año de Secundaria, son las construcciones de figuras geométricas (planas) usando regla y compás. Se propone realizar –además de las construcciones indicadas en los libros de referencia– como actividades adiciones, algunas figuras especiales que pueden obtenerse a través de líneas rectas y curvas, por ejemplo imágenes sencillas de arte óptico (*op art*), o bien usando varias líneas rectas que sean tangentes a curvas, para construir diversas imágenes. Se pueden realizar estas actividades de forma manual, en papel cartulina y papel cascarón, o bien, de forma electrónica, en la plática, se mostrarán algunos ejemplos creados en el aula de nuestra institución, así como ejemplos más complejos de reconocidos artistas.

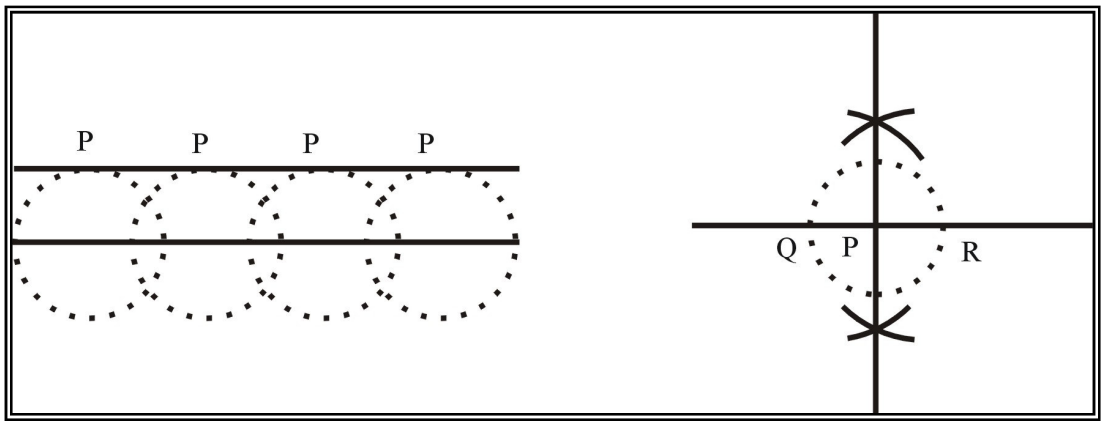
Palabras clave: Secundaria, Figuras geométricas, regla y compás, arte óptico, tangentes.
Modalidad: Enseñanza de las Matemáticas (Secundaria)

Construcciones Geométricas Básicas.

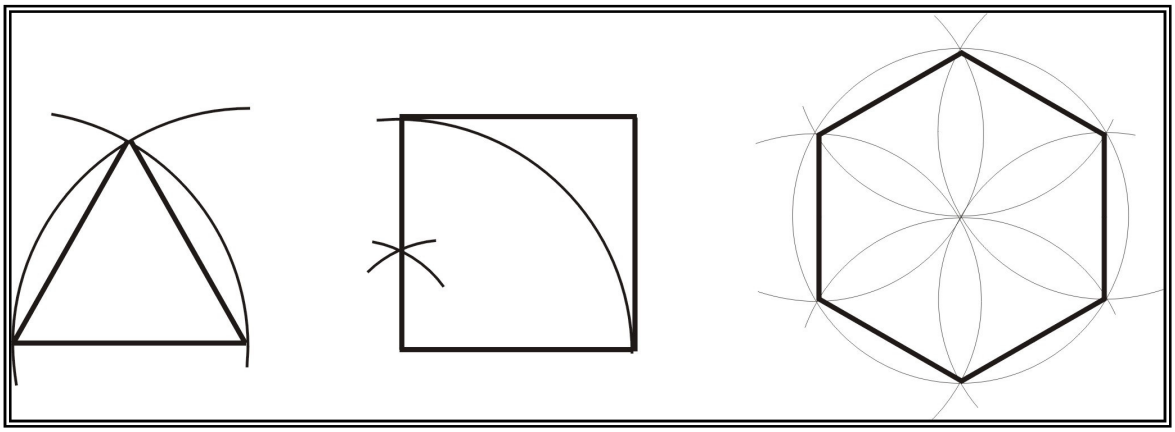
El origen de la geometría se encuentra relacionado con la actividad humana más antigua de la historia: la agricultura. Nace hacia el año 4000 a. C. con los egipcios y babilónicos. La palabra geometría proviene del griego *geo* que significa *tierra* y *metren* que significa *medir*, nos indica el estudio de las propiedades y relaciones formales de las figuras del plano y del espacio. Los primeros vestigios sobre el estudio de las figuras geométricas básicas y la medición de áreas y volúmenes se remontan a la civilización egipcia y su desarrollo y aplicación se debieron a que, con las crecidas anuales del río Nilo, frecuentemente desaparecían los límites de los campos de cultivo. Posteriormente, los griegos le otorgaron un alto grado de formalización matemática que culminó en *Los Elementos* de Euclides. Rene Descartes, aunó la geometría con el álgebra con lo que se desarrollo lo que hoy se

conoce como geometría analítica. A partir del siglo XIX desaparece la creencia de una sola geometría.

Al comenzar con el estudio de la geometría, nos encontramos en la necesidad de construir figuras usando solamente dos instrumentos: regla y compás. En el segundo grado de nivel medio, se da una breve introducción a este tema. Se comienza mostrando los métodos de trazo de paralelas y perpendiculares.



Una vez que el alumno domina los trazos básicos, se continua con la construcción de polígonos regulares, como el triángulo, cuadrado, pentágono, hexágono, etc.



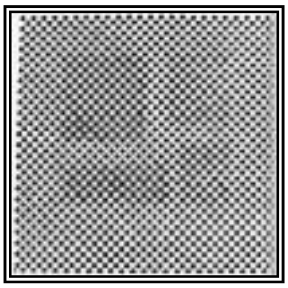
Adicionalmente, se muestra como construir óvalos, ovoides y algunas otras figuras sencillas, estas figuras se usan posteriormente para la creación de figuras a escala y proyecciones.

El tema se debe desarrollar de tal forma que el alumno sea capaz de construir dichas figuras y mostrar la utilidad de las construcciones geométricas sencillas.

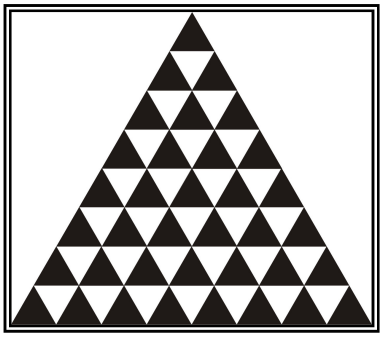
Aplicaciones.

Es común, en la creación de un dibujo o un cuadro de arte, usar como base figuras geométricas sencillas, pero un buen dibujo depende de la capacidad del artista. Otra opción, la cual sólo depende de trazos básicos, como círculos y rectas, puede ser usado para obtener un verdadero cuadro de arte. A continuación mostramos dos opciones que pueden llevarse a cabo en el salón de clase.

1. Op art



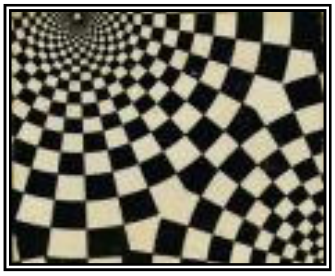
El interés de un grupo de artistas por plasmar sensaciones de movimiento en una superficie bidimensional, engañando al ojo humano mediante ilusiones ópticas, dio lugar a una corriente artística que se denominó *Arte Óptico*, conocida comúnmente por su acepción en inglés: Op Art; abreviación de Optical Art.



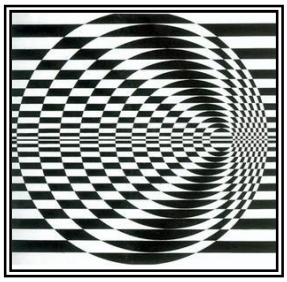
El arte óptico apareció hacia finales de los años 50, consolidándose en la primera mitad de la década siguiente. En 1964, Time Magazine publicó un artículo sobre un grupo de artistas cuyo objetivo principal era crear ilusiones ópticas a través de sus obras. La revista bautizó la tendencia con el nombre de Op Art. En 1965, el Museo de Arte Moderno de Nueva York albergó una exposición de artistas de todo el mundo consolidando el estilo.

Destacaron entre los representantes del Op Art: Victor Vasarely, Bridget Riley, Frank Stella, Josef Albers, Lawrence Poons, Kenneth Noland y Richard Anuszkiewicz.

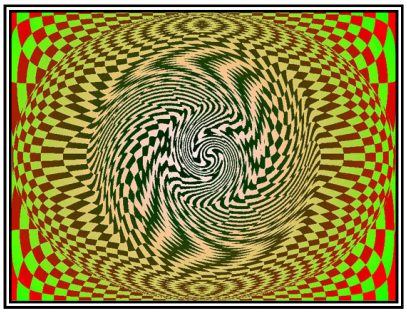
Caracterizaron al Arte Óptico las creaciones compuestas por patrones de repetición de líneas, cubos y círculos concéntricos, en los que predominaban el blanco y negro y la contraposición de colores complementarios.



Mediante la repetición de formas simples y un habilidoso uso de colores, luces y sombras, los artistas ópticos lograban en sus obras amplios efectos de movimiento, brindándole total dinamismo a superficies planas, las cuales terminaban siendo ante el ojo humano espacios tridimensionales llenos de vibración, movimiento y oscilación.



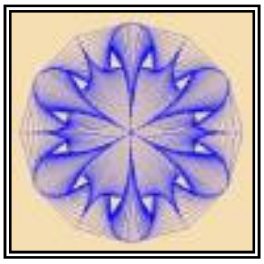
El Op Art generó una amplia polémica sobre su relación con el Arte Cinético. Numerosos fueron los personajes que enmarcaron la tendencia óptica dentro de la cinética, sin embargo, otra opinión diferenciaba una de otra, considerando que en el Op Art existía una ausencia total de movimiento real.

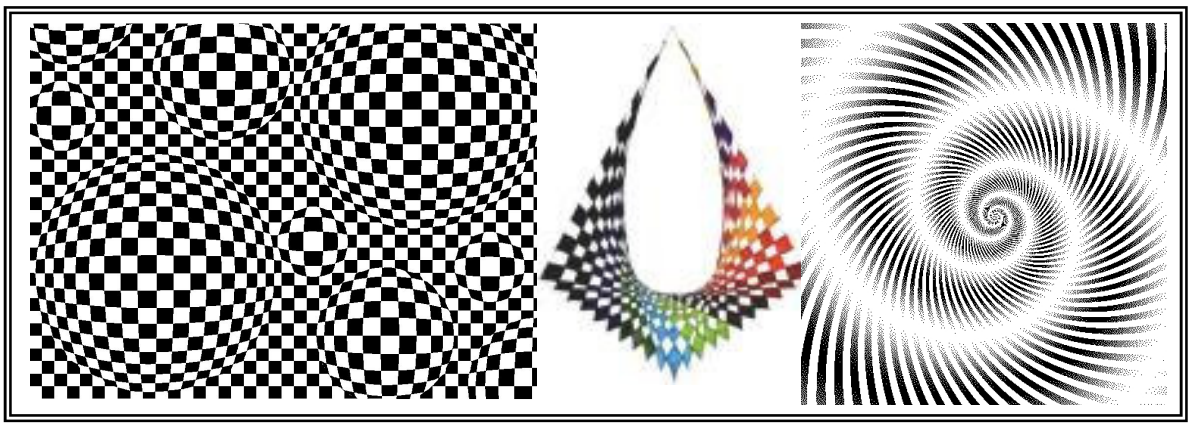


Con el auge del Popular Art, la industria de la moda textil tomó planteamientos de diversos movimientos artísticos para llevarlos a las prendas de vestir, poniendo al alcance de las masas elementos de sofisticado diseño. Entre las tendencias que protagonizaron esta popularización se encontró el Arte Óptico.

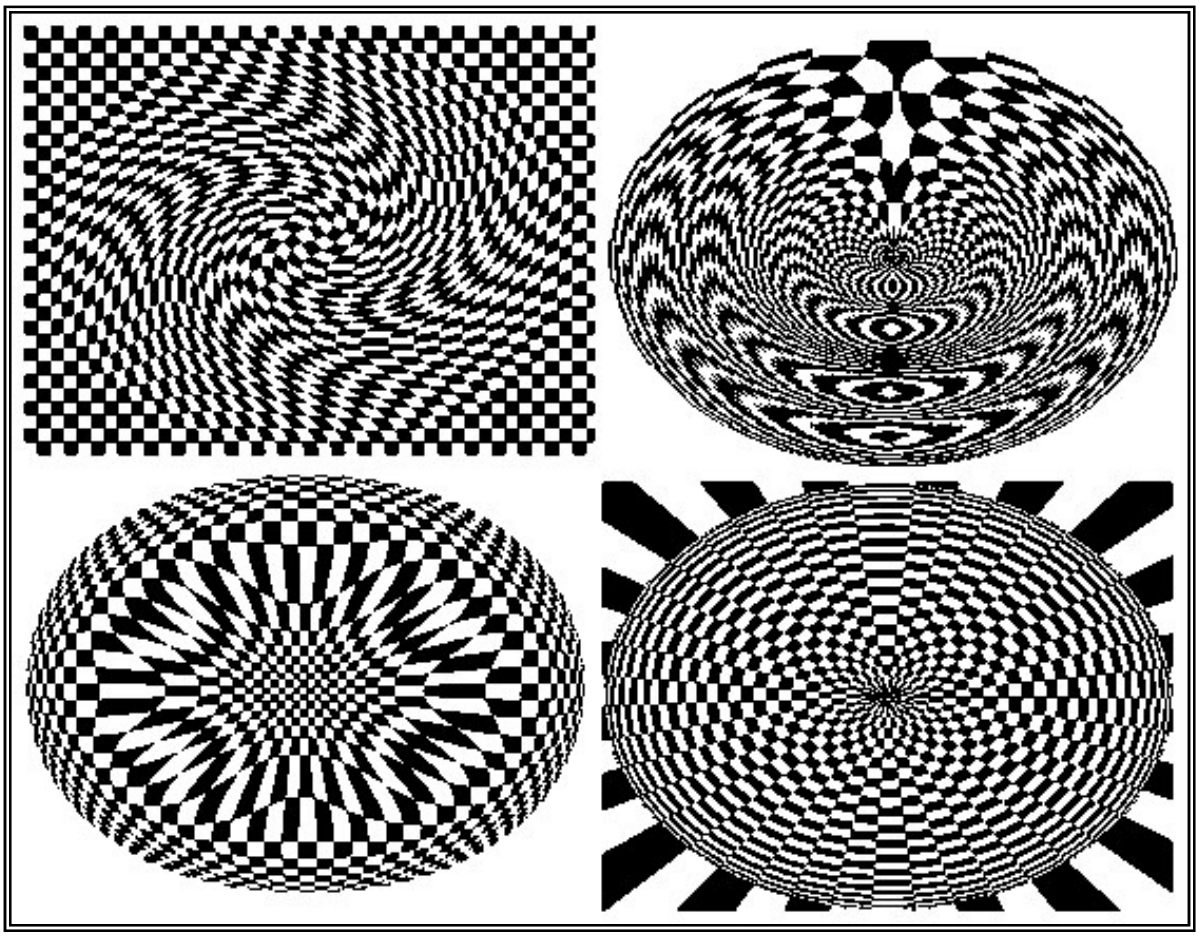
Sus máximos exponentes

Cuando se habla de Op Art es difícil que la primera referencia no apunte hacia el trabajo del artista de origen húngaro Víctor Vasarely (1908-1997) cuyos estudios sentaron los principios del arte óptico.





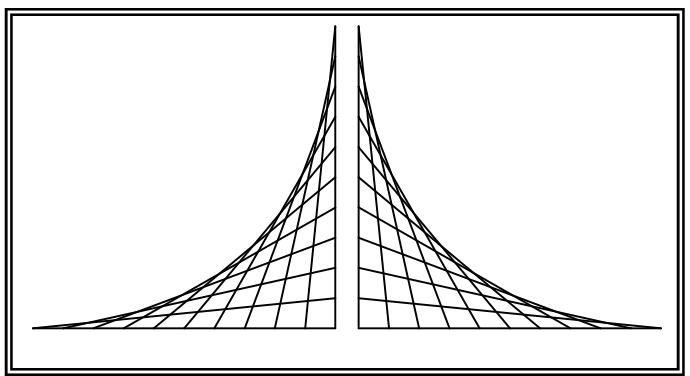
El uso de las figuras geométricas básicas en este tipo de arte, nos permite crearlas, con algo de trabajo, en el aula. Es un ejercicio entretenido y que da una mayor experiencia en las construcciones marcadas en los libros de texto. Claro, no complicarlas demasiado como algunas de las figuras mostradas anteriormente o como los siguientes ejemplos.



Este tipo de actividad puede llevarse a cabo usando solamente papel cartulina y plumón de color negro. También puede hacerse de tal forma que sea repetitiva cierta figura geométrica, por ejemplo, cuadrados, circunferencias, triángulos, pentágonos, etc., concéntricos, alternándolos de tal forma que quede uno en blanco y el otro en negro. Todo esto muestra al alumno una gran relación entre la matemática y el arte.

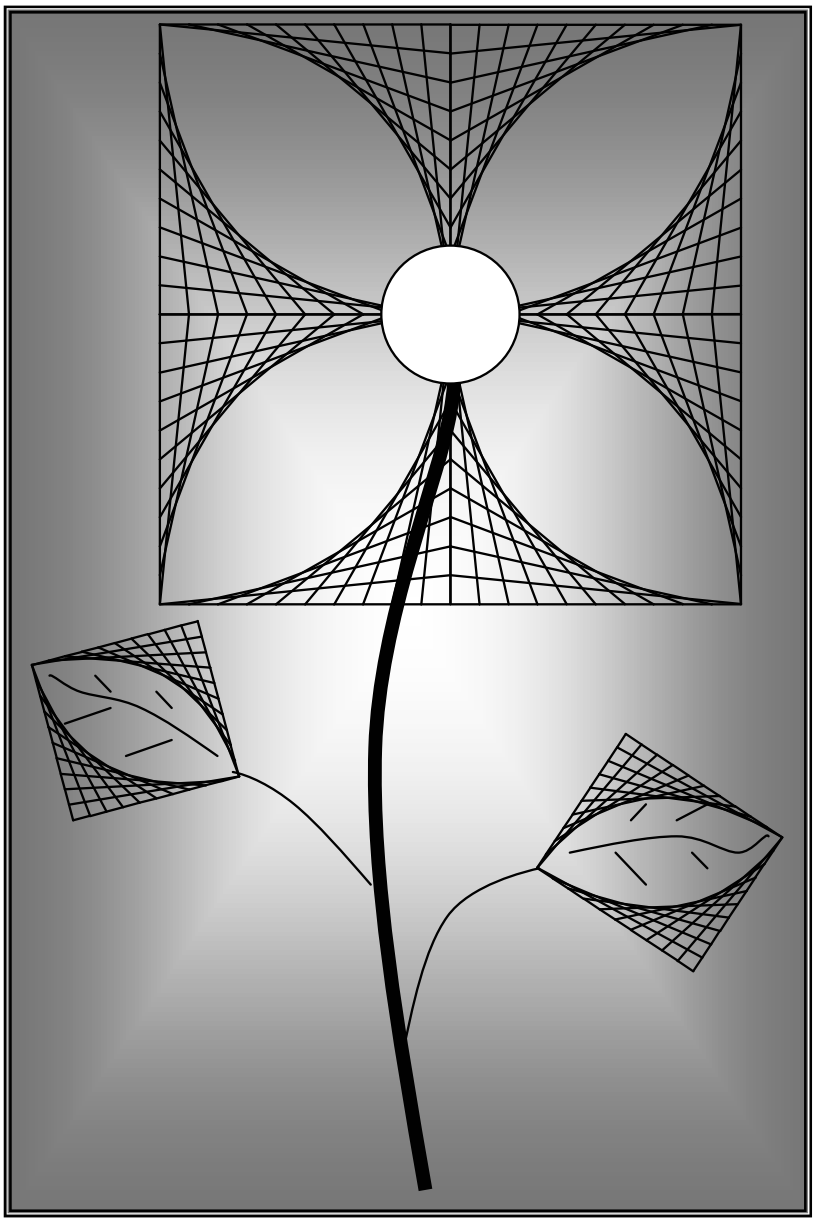
La siguiente opción, la hemos creado en el salón de clases usando papel cascarón, alfiles e hilo de diversos colores.

2. Tangentes a curvas



Recordemos la definición que se usa comúnmente para la

tangente a una curva cerrada: es una recta que sólo toca a la curva en un punto. Esta definición nos da una idea para construir figuras a través de líneas. Es claro que una pequeña dificultad es la cantidad de tangentes que debe de colocarse, pero como se muestra en las siguientes figuras, lo que se obtiene es muy práctico.



CONOS DE CONTINUOS CON LA PROPIEDAD DEL PUNTO FIJO

Florencio Corona y Raúl Escobedo
Facultad de Ciencias Físico Matemáticas (FCFM)
Benemérita Universidad Autónoma de Puebla (BUAP)
Puebla, Pue., México, 72570.
caos@fcfm.buap.mx, escobedo@fcfm.buap.mx

Resumen

Demostraremos que si $Y \in Clase(U)$, entonces para cada continuo X y para cada función continua y suprayectiva $f : X \longrightarrow Y$, se tiene que la función inducida $Cono(f) : Cono(X) \longrightarrow Cono(Y)$, es universal. Como una consecuencia obtenemos que $Cono(Y)$ tiene la propiedad del punto fijo.

1. Introducción

En 1967, Holsztyński introdujo la noción de función universal [2]. Una función continua $f : X \to Y$ entre espacios topológicos X y Y se dice que es una *función universal* si para cada función continua $g : X \to Y$, existe un punto $x \in X$ tal que $f(x) = g(x)$.

Posteriormente, en 1997, Marsh introdujo la noción de clase universal [4]. La *clase universal*, $Clase(U)$, es la colección de todos los continuos Y con la propiedad de que para cualquier continuo X y cualquier función continua y suprayectiva $f : X \to Y$, se tiene que f es una función universal. Es bien sabido que la $Clase(U)$ y la colección de los continuos con semimargen suprayectivo cero son la misma, para esto, consulte [1, Theorem 25].

En 1983, Marsh demostró, en [3, Theorem 5], que el cono de cualquier continuo con semimargen suprayectivo cero (*i.e.* en la $Clase(U)$), tiene la propiedad del punto fijo. El proposito de este trabajo es dar una generalización al resultado de Marsh antes mencionado, mostrando que, la función inducida, a los conos, de una función de un continuo sobre un continuo que pertenece a la $Clase(U)$, es universal.

Como antecedente a estos resultados, se conoce que el cono sobre un continuo encadenable (tipo-arco), tiene la propiedad del punto fijo, se sabe que la colección de los continuos encadenables es una subcolección de la $Clase(U)$, así, los continuos encadenables heredan las propiedades de los continuos en la $Clase(U)$.

2. Definiciones

Un *continuo* es un espacio métrico, compacto, conexo y con más de un punto. El *cono* sobre un continuo X, denotado por $Cono(X)$, es el espacio cociente $(X\times[0,1])/(X\times\{1\})$ obtenido del producto cartesiano $X\times[0,1]$ identificando $X\times\{1\}$ a un punto, v_X, llamado el *vértice* del $Cono(X)$. La *base* del $Cono(X)$ es el conjunto $\{(x,0) : x \in X\}$, el cual denotamos por B_X. Un punto en el $Cono(X)$ lo denotamos por $\mathbf{x} = (x,t) \in X\times[0,1]$, donde para cada $x \in X$, $v_X = (x,1)$. La función $\pi_1 : Cono(X) \setminus \{v_X\} \longrightarrow X$ dada por $\pi_1(x,t) = x$ es la proyección sobre X y la función $\pi_2 : Cono(X) \longrightarrow [0,1]$ dada por $\pi_2(x,t) = t$ es la proyección sobre $[0,1]$, note que $\pi_2(v_X) = 1$ y $\pi_2(B_X) = 0$.

Dada una función continua $f : X \longrightarrow Y$, se induce de manera natural una función a los conos, $Cono(f) : Cono(X) \longrightarrow Cono(Y)$, la cual resulta ser continua y está dada por $Cono(f)(x,t) = (f(x),t)$. Note que, si f es suprayectiva, entonces $Cono(f)$ también lo es. Además, $Cono(f)(v_X) = v_Y$ y $Cono(f)(B_X) \subset B_Y$.

Un espacio X tiene la *propiedad del punto fijo*, si para cada función continua $f : X \to X$ existe un punto $x \in X$ tal que $f(x) = x$.

Sean A, B y H subconjuntos cerrados de un continuo X, donde A y B son ajenos. Se dice que H *corta débilmente entre* A y B en X, si para cada continuo C en X tal que $C \cap A \neq \emptyset$ y $C \cap B \neq \emptyset$, se tiene que $C \cap H \neq \emptyset$. Se dice que X es *s-conexo entre* A y B, si para cada subconjunto cerrado H que corta débilmente entre A y B en X, existe una componente K de H que corta débilmente entre A y B en X. Un continuo X es *s-conexo*, si para cada par de subconjuntos cerrados y ajenos A y B en X, se tiene que X es s-conexo entre A y B.

3. Propiedades generales

Los siguientes resultados sobre funciones universales son bien conocidos y los puede consultar en [2, Proposition 1, Proposition 2, Proposition 8].

1. **Proposición.** Si $f : X \longrightarrow Y$ es una función universal, entonces f es suprayectiva.

2. **Proposición.** Si $f : X \longrightarrow Y$ es una función universal, entonces Y tiene la propiedad del punto fijo.

3. **Proposición.** Si X es un continuo y $f : X \longrightarrow [0, 1]$ es una función continua y suprayectiva, entonces f es universal.

Como consecuencia a la Proposición 2, tenemos que si $Y \in Clase(U)$ entonces Y tiene la propiedad del punto fijo.

De los resultados que aparecen en el artículo de Marsh, referente a s-conexidad y la propiedad del punto fijo [3], se deduce la siguiente:

4. **Proposición.** Si X es un continuo, entonces $Cono(X)$ es s-conexo.

4. El teorema

En esta sección mostraremos que, la función inducida, a los conos, de una función de un continuo sobre un continuo que pertenece a la $Clase(U)$ es universal. Como consecuencia tenemos que, el $Cono(Y)$ tiene la propiedad del punto fijo cuando Y pertenece a la $Clase(U)$ [3, Theorem 5].

5. **Teorema.** Si $Y \in Clase(U)$, entonces para cada continuo X y para cada función $f : X \longrightarrow Y$ continua y suprayectiva, se tiene que la función inducida $Cono(f) : Cono(X) \longrightarrow Cono(Y)$ es universal.

Prueba.
Sea $f : X \longrightarrow Y$ como en el teorema y sea $g : Cono(X) \longrightarrow Cono(Y)$ una función continua. Debemos probar que:

$$\text{Existe } \mathbf{p} \in Cono(X) \text{ tal que } Cono(f)(\mathbf{p}) = g(\mathbf{p}). \ (*)$$

Supongamos que $Cono(f)(\mathbf{x}) \neq g(\mathbf{x})$, para cada $\mathbf{x} \in Cono(X)$.

Sea

$$\mathcal{H} = \{\mathbf{x} \in Cono(X) : \pi_2(Cono(f)(\mathbf{x})) = \pi_2(g(\mathbf{x}))\}.$$

Notamos que $\mathcal{H} \neq \emptyset$ y $\mathcal{H}$ es cerrado.

En efecto, que $\mathcal{H}$ es cerrado, se sigue por estar definido con funciones continuas.

Veamos que $\mathcal{H} \neq \emptyset$:

Note que por la Proposición 3, $\pi_2 \circ Cono(f) : Cono(X) \longrightarrow [0,1]$ es una función universal, y como $\pi_2 \circ g : Cono(X) \longrightarrow [0,1]$ es continua, existe $\mathbf{x_0} \in Cono(X)$ tal que $\pi_2(Cono(f)(\mathbf{x_0})) = \pi_2(g(\mathbf{x_0}))$, es decir, $\mathbf{x_0} \in \mathcal{H}$. Así, $\mathcal{H} \neq \emptyset$.

Como estamos suponiendo que (*) no se cumple, se tiene que $v_X \notin \mathcal{H}$, de esto se sigue que $v_Y \notin Cono(f)(\mathcal{H})$. También ocurre que $v_Y \notin g(\mathcal{H})$.

Note que si $\mathcal{C}$ es un continuo en $Cono(X)$ tal que $\mathcal{C} \cap (B_X) \neq \emptyset$ y $\mathcal{C} \cap \{v_X\} \neq \emptyset$, entonces $\mathcal{C} \cap \mathcal{H} \neq \emptyset$. Es decir, $\mathcal{H}$ corta débilmente entre B_X y $\{v_X\}$ en $Cono(X)$. Luego, como por la Proposición 4, $Cono(X)$ es s-conexo, existe $\mathcal{K}$ componente de $\mathcal{H}$ tal que $\mathcal{K}$ corta débilmente entre B_X y $\{v_X\}$ en $Cono(X)$.

Note que, como $v_Y \notin Cono(f)(\mathcal{K})$ y $v_Y \notin g(\mathcal{K})$, podemos considerar las funciones, $\pi_1 \circ Cono(f) \mid_{\mathcal{K}} : \mathcal{K} \longrightarrow Y$ y $\pi_1 \circ g \mid_{\mathcal{K}} : \mathcal{K} \longrightarrow Y$. Se tiene que $\pi_1 \circ Cono(f) \mid_{\mathcal{K}}$ es suprayectiva y como $Y \in Clase(U)$ se tiene que $\pi_1 \circ Cono(f) \mid_{\mathcal{K}}$ es universal. Luego, existe $\mathbf{p} \in \mathcal{K}$ de tal forma que

$$\pi_1(Cono(f)(\mathbf{p})) = \pi_1(g(\mathbf{p})). \quad (1)$$

Finalmente, como $\mathbf{p} \in \mathcal{K} \subset \mathcal{H}$, se sigue

$$\pi_2(Cono(f)(\mathbf{p})) = \pi_2(g(\mathbf{p})). \quad (2)$$

Concluimos de (1) y (2) que, $Cono(f)(\mathbf{p}) = g(\mathbf{p})$, lo cual es una contradicción, por tanto (*) es cierto, con todo se tiene la prueba del teorema.

Como una consecuencia del Teorema 5 y la Proposición 2, obtenemos el resultado de Marsh [3, Theorem 5]:

6. **Corolario.** Si $Y \in Clase(U)$, entonces $Cono(Y)$ tiene la propiedad del punto fijo.

Referencias

[1] J.J. Charatonik, R. Escobedo, On semi-universal mappings, in: A. Illanes, I. W. Lewis, S. Macías (Eds.), Continuum Theory: Proceedings of the Special Session in Honor of Professor Sam B. Nadler, Jr's 60th birthday, in: Lecture Notes Pure Appl. Math., vol. 230, Marcel Dekker, New York, 2002.

[2] W. Holsztyński, *Universal Mappings and Fixed Point Theorems*, Bull. Pol. Acad. Sci., 15 (1967), 433-438.

[3] M. M. Marsh, *s-connected Spaces and the Fixed Point Property*, Top. Proc., 8 (1983), 85-97.

[4] M. M. Marsh, *Some Generalizations of Universal Mappings*, Rocky Mountain J. Math., 27 (1997), 1187-1198.

PGSM
Puebla, MÉXICO
Noviembre de 2005

Programación del Método Bootstrap

Armando Vargas L., Miguel Ángel Vargas L., Martín Estrada A., Rogelio González V.
Facultad de Ciencias de la Computación (FCC)
Benemérita Universidad Autónoma de Puebla (BUAP)
Puebla, Pue., México, 72570
alomeli@puebla.megared.net.mx, mavlomeli@yahoo.com, mestrada@cs.buap.mx, rgonzalez@cs.buap.mx

Resumen

En este trabajo, se presentará el Método Bootstrap que nos permite hallar de una manera aproximada el error estándar de la media y la mediana; así como su algoritmo, complejidad computacional y los resultados que ofrece la implementación del método, comparando éstos con las soluciones obtenidas de manera analítica, siempre y cuando sea posible obtenerlas.

Palabras clave: programación Bootstrap; error estándar de media y mediana.

1. Introducción.

Bradley Efron es el creador del Método Bootstrap, actualmente es profesor y presidente del departamento de estadística en la universidad de Stanford y ha sido galardonado con la medalla Wilks.

Existen problemas que no son solubles usando los estimadores clásicos, por lo que se buscan otras técnicas para hallar resultados. Efron en 1979 desarrolló el Método Bootstrap [1], basándose en el método para calcular el error estándar de $\hat{\theta}$. Sus usos actuales incluyen áreas tan diversas como la bioestadística y la astrofísica. Lo sorprendente es que los métodos usados son similares en ambas áreas.

En [5] propone el uso del método Bootstrap para resolver algunos ejercicios que plantea y en [9] utiliza el método para problemas en Intervalos de Confianza. En particular nosotros utilizamos el método para hallar el error estándar de la media y

la mediana aunque el software elaborado también aproxima al error estándar de la varianza y la desviación estándar.

2. El Método Bootstrap

La idea del método Bootstrap es un muy sencilla, dada una distribución de muestreo, generar *B* nuevas muestras aleatorias a partir de la original, entonces se elige el tipo de estadística a utilizar (media, mediana, varianza o desviación estándar) para evaluar cada muestra en la función electa, generando así *B* resultados y finalmente calcular el error estándar de los *B* valores obtenidos con anterioridad.

A continuación se presenta el **algoritmo del Método Bootstrap:**

1. Dada $x_1, x_2, \ldots, x_n$ se define la distribución de muestreo.
2. Generar *B* muestras aleatorias $x^*_1, x^*_2, \ldots, x^*_B$ de la distribución de muestreo. (Efron recomienda *B* = 100) Esto se le llama Bootstrap simple.
3. Se obtienen *B* valores del bootstrap $\hat{\theta}*$ donde

   $$\hat{\theta}*\,(b) = S(x^{*b}) \qquad b = 1, 2, \ldots, B$$

   donde $S(x^{*b})$ es la estadística a usar (media, mediana, varianza, desviación estándar).

   - Caso media: $S(x^{*b}) = \bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$
   - Caso mediana:

     Ordenar los datos x_i

     - Si *n* es impar, entonces $S(x^{*b}) = \tilde{x} = x_{(n \text{ div } 2 + 1)}$.
     - Si *n* es par, entonces $S(x^{*b}) = \tilde{x} = \{x_{(n \text{ div } 2)} + x_{(n \text{ div } 2 + 1)}\} / 2$

     NOTA: `div` denota división entera y ésta a su vez tiene mayor prioridad que la suma.

   - Caso varianza: $S(x^{*b}) = S^2 = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

- Caso desviación estándar: $S(x^{*b})$ = S = $\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$

4. Se calcula el error estándar del bootstrap

$$\hat{se}_B = \left\{ \frac{\sum_{b=1}^{B}\left[\hat{\theta}^*(b) - \hat{\theta}^*(\cdot)\right]^2}{(B-1)} \right\}^{1/2}$$

donde

$$\hat{\theta}^*(\cdot) = \frac{\sum_{b=1}^{B}\hat{\theta}^*(b)}{B}$$

El esquema del Bootstrap se muestra en la figura 1.

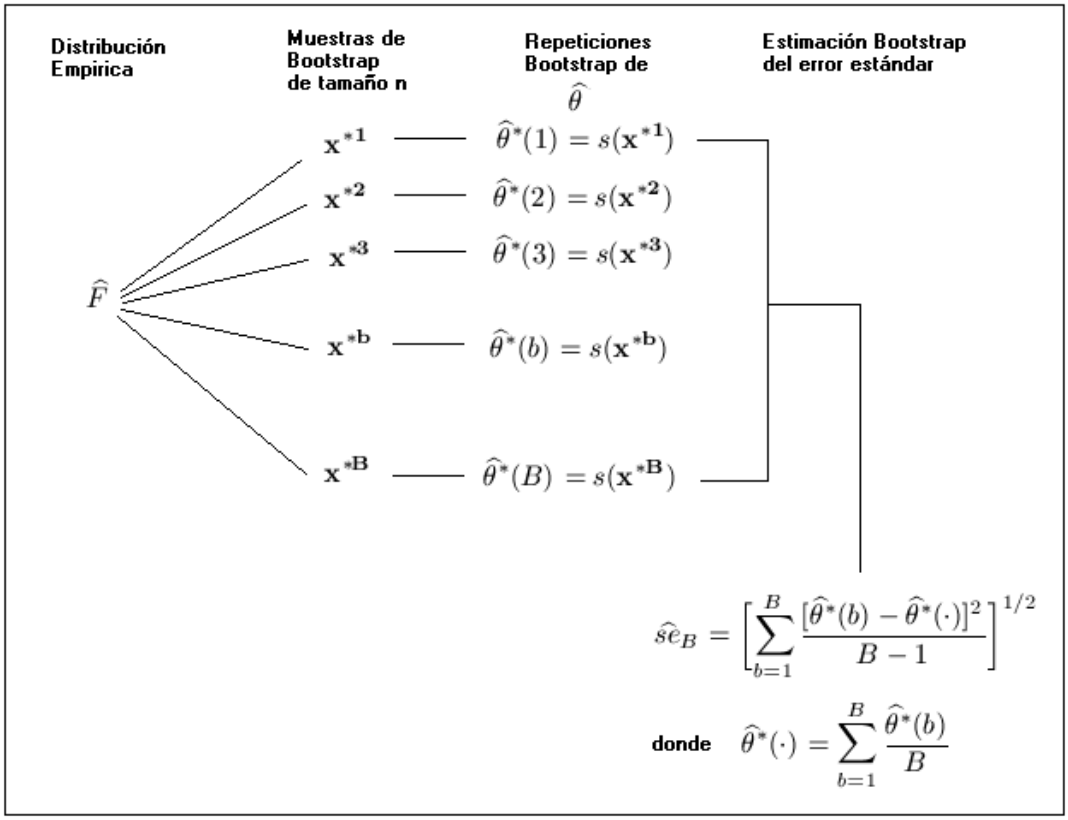


Figura 1: Esquema del Bootstrap

3. Análisis Computacional del Bootstrap

Al hablar de la eficiencia o complejidad computacional de un algoritmo, nos estamos refiriendo a determinar matemáticamente la cantidad de recursos necesarios para su almacenamiento y ejecución en una computadora. Los recursos de la computadora son el tiempo que el algoritmo tarda en ejecutarse y la memoria utilizada. Decimos que un algoritmo es más eficiente o de menor complejidad que otro cuando utiliza menos recursos. En la eficiencia de tiempo de un algoritmo, vamos a usar funciones matemáticas sin unidades concretas.

En computación existen problemas verdaderamente difíciles porque son "tecnológicamente imposibles de resolver", a este tipo de problemas se les llama *indecidibles*. Por debajo de éstos existen los problemas de la clase NP, que significa que son problemas de "tiempo polinómico no determinista", donde una máquina no determinista tiene una variedad de pasos siguientes alternativos. Existe un subconjunto de problemas de la clase NP que son los NP completos (que contiene a los más difíciles). La complejidad exacta de los problemas NP completos todavía tiene que ser determinada y sigue siendo el problema abierto más notable en las ciencias de la computación [6 – 8]. Otro subconjunto de la clase NP, son los problemas polinomiales (determinísticos) que en este caso particular cae el algoritmo Bootstrap.

Para estudiar la complejidad del algoritmo Bootstrap, se introducirán varios conceptos, uno de ellos, cuando decimos que la complejidad de un algoritmo es $T(n) = cf(n)$ (por ejemplo $f(n) = n^2$) no estamos haciendo referencia a una determinada unidad de tiempo, sino a que el tiempo de ejecución del algoritmo es proporcional a cierta función $f(n)$, donde la constante de proporcionalidad c depende de la computadora.

En general se tienen dos criterios para calcular $T(n)$:

- Considerar $T(n)$ como el tiempo para el peor caso de los posibles.
- Hacer $T(n)$ igual al tiempo medio de todos los casos posibles.

Notación O( )

Sea $f(n) = O(g(n))$ si existen constantes c y $n0$ tales que

$$f(n) \leq cg(n) \qquad \forall\ n \geq n0$$

Si $f(n)$ denota el consumo de recursos de un algoritmo en función de la entrada n, decidir que f tiene el orden de g supone que $cg(n)$ es una cota superior del tiempo de ejecución del algoritmo.

Propiedades de la notación O( )

Para cualquier par de funciones $f(n)$ y $g(n)$ se verifican las siguientes propiedades:

- $O(cf(n))$ es $O(f(n))$.
- $O(f(n) + g(n))$ es $max(O(f(n)), g(n)))$.
- $O(f(n)) + O(g(n))$ es $O(f(n) + g(n))$.
- $O(f(n))\ \ O(g(n))$ es $O(f(n)\ \ g(n))$.
- $O(O(f(n)))$ es $O(f(n))$.

Eficiencia del Método Bootstrap

Analizando el algoritmo Bootstrap para estimación de error estándar, podemos notar que se usan diferentes estadísticas como la media, mediana, varianza y desviación estándar, por lo que tenemos cuatro algoritmos diferentes en cuestión de lo que tienen que calcular y por consecuencia el análisis de eficiencia puede variar entre cada uno. Básicamente el algoritmo es el mismo en los pasos 1, 2 y 4 pero en el paso 3 la forma de calcular y programar cada estadística es diferente a las demás.

Para calcular la eficiencia, vamos a tomar el peor de los casos, cuando n el número de datos de la muestra es muy grande y B el número de iteraciones también lo es. Por simplicidad podemos hacer $n = B$ (ambos valores por hipótesis son grandes y no es de suma importancia saber cuál es más grande, para hacer una diferencia entre estos).

La operación en el paso 1, es la lectura de *n* muestras, que se realiza en T(n)=n unidades o pasos, entonces su complejidad es de *O*(*n*).
Las operaciones en el paso 2, son la generación de *n* = *B* iteraciones con *n* muestras aleatorias cada una, es decir, por una iteración se generan *n* muestras y como son *n* iteraciones, entonces tenemos que *T*(*n*) = *n*+*n*+. . . +*n* (n veces) = n^2 pasos, es decir, su complejidad es *O*(n^2).

Para el paso 3 tenemos cuatro casos:
1. Media: Para calcular el valor de la media se requieren *n* pasos (debido a la sumatoria que va de 1 hasta *n*), pero se necesitan *n* = *B* iteraciones para obtener *B* valores de $\hat{\theta}*$, entonces *T*(*n*) = n^2, por lo tanto su complejidad es *O*(n^2).

2. Mediana: Primero se tienen que ordenar los datos, donde los mejores algoritmos de ordenación tienen una complejidad de *O*(*T*(*n*)) = *O*(*n log n*) (puede consultar [6], [7] y [8]), después se tiene una condición y su complejidad es *O*(2) (*T*(*n*) = 2 (condición + el cálculo de la mediana)), entonces por propiedades de la notación *O*, la complejidad del proceso para hallar la mediana es *O*(*n log n*) + *O*(2) = *O*(*n log n* + 2) = *O*(*n log n*). Aunque existe otro algoritmo mejor para hallar la mediana[1], el cual tiene una complejidad de *O(n)*, dicho algoritmo se especifica en [6] y [7], por lo que se utilizará éste en nuestra implementación; entonces por último se tienen que realizar *n* = *B* iteraciones para obtener los valores de $\hat{\theta}*$, por lo que la complejidad final es *O*(*n*) · *O*(*n*) = *O*(n^2).

3. Varianza: Primero hay que calcular la media, pero ya sabemos que su complejidad es de *O*(*n*), entonces ahora veamos el cálculo de complejidad de la función $\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$. Observemos que es una sumatoria, entonces realiza *n* iteraciones con una división al final, por lo que realiza *T*(*n*) = *n*+1 pasos, y su complejidad es de *O*(*n*), pero nos falta sumar la complejidad de la media, que realmente no afecta el resultado debido que *O*(*n*)+*O*(*n*) = *O*(*n*+*n*) = *O*(*n*).

[1] El algoritmo de la mediana se resuelve utilizando la técnica “Divide y Vencerás” [6] y [7].

Finalmente se deben realizar las $n = B$ iteraciones para los valores de $\hat{\theta}*$, entonces la complejidad final es $O(n) \cdot O(n) = O(n \cdot n) = O(n^2)$.

4. Desviación estándar: Es exactamente el procedimiento de la varianza (salvo una raíz cuadrada), entonces su complejidad es $O(n^2)$.

En el paso 4 tenemos que hallar la complejidad del error estándar del bootstrap, que implica calcular la eficiencia de la media y la varianza (que van implícitas dentro de la fórmula del error estándar), como $n = B$, entonces la complejidad de la media y la varianza es O(n), por lo que la complejidad del error estándar es $O(n) + O(n) = O(n + n) = O(n)$.

Finalmente, podemos calcular la complejidad de cada algoritmo Bootstrap para estimación de error estándar, es decir, se suman las complejidades de los pasos 1, 2 y 4 más la complejidad respectiva del paso 3:

1. Media:

$$O(n) + O(n^2) + O(n^2) + O(n) = \mathbf{O(n^2)}$$

2. Mediana:

$$O(n) + O(n^2) + O(n^2) + O(n) = \mathbf{O(n^2)}$$

3. Varianza:

$$O(n) + O(n^2) + O(n^2) + O(n) = \mathbf{O(n^2)}$$

4. Desviación estándar:

$$O(n) + O(n^2) + O(n^2) + O(n) = \mathbf{O(n^2)}$$

4. Pruebas del Software Bootstrap

La programación del método fue hecha en Delphi 6.0 para Windows, por lo que probaremos el software con los siguientes experimentos:

Experimento 1:

Un modelo teórico sugiere que *X*, el tiempo de falla de un líquido aislante entre electrodos en un voltaje particular, tiene $f(x, \lambda) = \lambda e^{-\lambda x}$, que es una distribución exponencial. Una muestra aleatoria de $n = 10$ tiempos de falla (en minutos) da los siguientes datos:

41.53, 18.73, 2.99, 30.34, 12.33, 114.52, 73.02, 223.63, 26.78, 4.0

Solución por el método clásico:

Sabemos que el modelo tiene una distribución exponencial, y que tiene una función de probabilidad, por lo que haciendo un cambio de variable tenemos:

$$\lambda = \frac{1}{\beta}$$

$$f(x) = \frac{1}{\beta} e^{\frac{-x}{\beta}}, \ \beta > 0 \qquad 0 < x < \infty$$

que tiene una media (consultar [2, 3, 4 y 5])

$$f(x) = \beta$$

y una varianza

$$f(x) = \beta^2$$

Se calcula la media de la muestra para obtener la varianza.

$$\bar{x} = x_1 + x_2 + \ldots + x_{10}/10$$

entonces,

$$\bar{x} = 55.087$$

Sabemos que para calcular el error estándar de la media [2 - 5], podemos aplicar la siguiente fórmula:

$$\sqrt{\frac{S^2}{n}}$$

$$\Rightarrow \sqrt{\frac{\beta^2}{n}} \Rightarrow \sqrt{\frac{(55.087)^2}{10}} = 17.421$$

$\therefore$ el error estándar de la media es 17.421

Los resultados obtenidos del programa son los siguientes:

Iteraciones:	50	100	250	500	1000	10000
Resultado:	21.1431	20.6808	21.4709	22.5207	21.2395	20.6836

Por lo que los resultados se aproximan al valor real.

Experimento 2:

Se lanza un dado 6 veces, con los siguientes datos $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$, $x_5 = 5$ y $x_6 = 6$. Calcular el error estándar de la media aplicando el método clásico y el sistema Bootstrap.

Se calcula la media de la muestra, para hallar la varianza.

$$\bar{x} = x_1 + x_2 + \ldots + x_6 / 6$$

Ahora se calcula la varianza de la muestra

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

$$S^2 = 3.5$$

Calculando la media del error estándar

$$\sqrt{\frac{S^2}{n}}$$

$$\Rightarrow \sqrt{\frac{(3.5)^2}{6}} = 0.7637$$

Insertando los datos al programa, éste arrojó los siguientes resultados:

Iteraciones:	50	100	250	1000	10000
Resultados:	0.73036660	0.70236731	0.77548268	0.73899990	0.71109649

Los resultados tuvieron una buena aproximación con respecto al valor real.

Experimento 3:
Considere el siguiente grupo de valores, 94, 197, 16, 38, 99, 141 y 23. Calcule el error estándar de la mediana utilizando el sistema Bootstrap.

La fórmula para calcular el error estándar de la mediana de manera clásica es complicada, porque hay que hacer derivaciones complejas y no es tan trivial obtener su solución.

Observamos la solución que da [1] en este problema.

Iteraciones	50	100	250	500	1000	10000
Error est.	32.21	36.35	34.46	36.72	36.48	37.83
Bootstrap	34.115	37.157	37.960	37.157	36.670	37.837

Comparando la solución de [1] con el sistema Bootstrap, el resultado es muy aproximado, y con esto se comprueba la veracidad del sistema.

5. Conclusiones

Hay problemas que no se pueden resolver de manera clásica y aplicando el sistema Bootstrap sí se pueden resolver, por lo que el software puede ser utilizado por investigadores. En el caso de estudiantes al intentar resolver problemas de manera clásica implica mucha teoría, y a veces se necesita una solución rápida, por lo que al aplicar el sistema Bootstrap, el tiempo es mínimo. La comparación de la resolución de problemas, tanto de manera clásica como aplicando el sistema Bootstrap, tiene una muy buena aproximación y se satisface la necesidad establecida.

La complejidad computacional del método Bootstrap es $O(n^2)$ por lo que la implementación en una computadora es factible.

Referencias

[1] Efron Bradley: *An introduction to the Bootstrap*, Chapman & Hall/CRC, USA, 1993.

[2] Kalbfeisch J. G.: *Probabilidad e Inferencia Estadística 2*, AC, España, 1984.

[3] Mendenhall III Willian: *Estadística Matemática con Aplicaciones*, 6a. Edición, Thomson, México, 2002.

[4] Freund, Walpole: *Estadística Matemática con Aplicaciones*, 4a. Edición, Prentice Hall, México, 1992.

[5] Devore, Jay L. : *Probabilidad para ingeniería y ciencias*, 4a. Edición, Thomson Editores, México, 1998.

[6] Weiss Mark Allen, *Estructuras de Datos y Algoritmos*, Addison-Wesley Iberoamericana, México, 1995.

[7] Brassard G., Bratley T., *Fundamentos de Algoritmia*, Prentice Hall, España, 1996.

[8] Galve J., González J., Sánchez A., Velázquez J., *Algorítmica*, Addison-Wesley Iberoamericana, E. U. A., 1993.

[9] Saavedra P.: "Estudio del Bootstrap", España, saavedra@dma.ulpgc.es, 2003.

Reconstrucción de atractores determinados del análisis de SME

I. Flores-Nava, H. G. González-Hernández, D. Mocencahua-Mora.
Facultad de Ciencias de la Electrónica. BUAP
Avenida San Claudio y 18 Sur. San Manuel.

Email: {flornav_ i,hgonz,dmocencahua}@ece.buap.mx

***Resumen*— La señal superficial EMG (electromiográfica) es una señal no lineal con una razón de ruido de pequeña señal. El objetivo de este trabajo es identificar el comportamiento de las SME(señal mioeléctrica) superficial de acuerdo al tiempo de retardo y la dimensión de empotramiento, y con estos realizar la reconstrucción del atractor y proyectarlo en tres dimensiones para identificar su forma y comportamiento. Las señales se tomaron del antebrazo al realizar tres tipos de agarre: Cilíndrico, Esférico y de Precisión.**

***Palabras Clave*— Señal mioelèctrica, Dinámica no lineal, Atractores, Caos.**

1. Introducción

La señal mioeléctrica (SME) es la grabación de la suma de trenes de potencial de acción asíncronos generados por los músculos de las extremidades en el proceso de movimiento del mismo. Esta actividad puede ser monitoreada por electrodos colocados encima de la piel (superficiales). La señal recibida provee información concerniente a la actividad local asociada con la contracción del músculo conocida como el potencial de acción de la unidad motora [5].

La adquisición de SME por medio de electrodos superficiales es una forma práctica y segura de obtenerla sin ser invasiva. Sin embargo sólo permite obtener señales burdas del músculo que se desea investigar. La amplitud de la señal está limitada de 0 a $10\ mV$(pico-pico) o $0-1.5\ mV$(rms). La energía útil está limitada de 0 a 500 Hz en el rango de frecuencia, con una energía dominante de $50-150Hz$ [8].

La SME es constantemente aplicada en rehabilitación y para controlar dispositivos protéticos para personas con amputaciones o con mal-formaciones congénitas en alguna extremidad del cuerpo [1], [9]. Pero las aplicaciones dependen de las características obtenidas de la SME.

La SME puede ser analizada por diversos métodos. Los dos tipos de métodos más empleados son los que evalúan en el dominio del tiempo y los que evalúan en el domino de la frecuencia. Dentro de los del dominio del tiempo están los estadísticos y los geométricos, mientras que por parte del dominio de la frecuencia están los espectrales [10]. Estos métodos han demostrado utilidad clínica y diagnóstica importante.

Muchos parámetros de la señal han sido usados para representar estas características. En algunos sistemas de control creados para prótesis de miembro superior [1], se han utilizado ciertos parámetros en el dominio del tiempo, tales como cruces por cero, el valor absoluto promedio, y como proceso estocástico.

Entre las técnicas usadas para determinar el comportamiento de sistemas no lineales, se cuenta el de coordenadas retrasadas, el cual reconstruye el atractor de la señal que se examina, para realizar éste, se necesitan dos parámetros como los son el tiempo de retardo y la dimensión de empotramiento, los cuales se determinan por medio de los métodos de información mutua promedio y falsos vecinos cercanos respectivamente [3]. Estos se aplicaron a las series de tiempo que describen señales mioeléctricas generadas por el movimiento de el dedo pulgar.

Este trabajo describe los resultados obtenidos usando la extracción de características de tres modos de agarre: cilíndrico, esférico y de precisión. La búsqueda en la identificación de los movimientos de la mano, se realiza mediante las señales electromiográficas obtenidas a partir de diversos músculos que intervienen en el movimiento.

El documento está organizado como sigue: en la siguiente sección se describe el equipo empleado para la obtención de SME y se mencionan las características del software que se utilizó para la obtención de las mismas. En la sección III se describen los métodos no-lineales empleados en el análisis de las señales. En la IV se presentan los resultados obtenidos y finalmente la sección V contiene las conclusiones.

2. Obtención de la SME

A. Electrodos usados

La actividad muscular puede ser monitoreada a través de electrodos superficiales colocados encima de la piel. La señal adquirida provee información concerniente al total de la actividad eléctrica asociada con la contracción muscular.

En la medición de SME se usaron electrodos superficiales para detectar la señal asociada al movimiento del dedo

pulgar, los músculos que son asociados con los movimientos de éste son los músculos de flexión y abducción en la parte inferior del antebrazo. La configuración que se utiliza para obtener dicha señal es de forma diferencial, en la cual se colocan los electrodos en los músculos donde se tiene la mayor actividad durante el movimiento del pulgar: abductor largo del pulgar y flexor corto del pulgar (más activo en la extensión y abducción del pulgar) y el flexor largo del pulgar (más activo en la flexión del pulgar) [6] y el tercer electrodo es una referencia a tierra (1). Los electrodos fueron verificados y clasificados para poder obtener la señal en tiempo real.

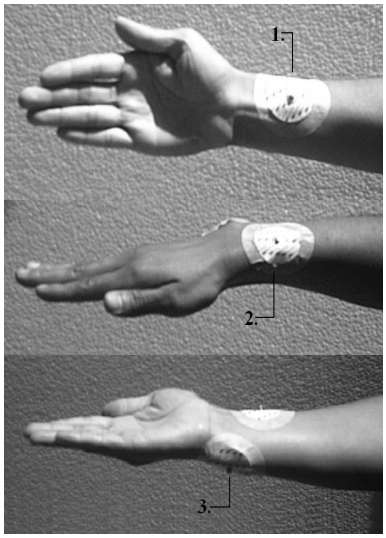


Fig. 1. Localización de electrodos utilizado para la detección del movimiento del pulgar. 1) Abductor Largo del Pulgar y Flexor corto del pulgar. 2) Flexor largo del Pulgar y 3) Base de cúbito (ulna)

La configuración diferencial contiene un electrodo a tierra sobre tejido eléctricamente inactivo cerca de la parte baja del cúbito.

B. Equipo de medición y procesamiento de datos

Para la medición de SME se utilizaron electrodos superficiales de la marca $3M$ modelo $No.2239$, y un amplificador de instrumentación que cuenta con tres electrodos, dos electrodos conectados en configuración diferencial y el tercero como referencia a tierra. Además se utilizó una tarjeta de adquisición de datos de la marca Advantech, modelo $PCI-1716$, que cuenta con conversión A/D de 16-bits con una razón de muestreo de 250kHz.

Los objetos utilizados para realizar los tres agarres mencionados fueron un cilindro grande de: $6cm$ de diámetro, una pelota de: $5.2cm$ de diámetro y un cilindro pequeño de diámetro: $0.5cm$, con los cuales se desarrollaron los movimientos seleccionados.

Los sujetos (8 personas) fueron guiados para realizar el movimiento deseado, donde la posición inicial es de reposo, para luego realizar el movimiento y finalmente regresar a la posición inicial. Se tomó de cada sujeto 5 series de cada movimiento, teniendo un número de muestras total de 50000 en 5 segundos, las cuales fueron tomadas consecutivamente, esperando entre cada serie 4 min. Para no tener un aprendizaje neuronal y perder información que nos pudiera ser útil.

Las series de mediciones fueron procesadas y analizadas fuera de línea usando matlab 7.0, con el cual se ha podido graficar las señales entrantes así como aplicar los filtros y sobre la señal ya filtrada realizar los diferentes métodos de dinámica no lineal mencionados con anterioridad.

Se aplicaron tres tipos de filtros a la señal EMG: filtro pasa-altas, pasa bajas y notch. El filtro pasa-altas es un filtro de 4to. orden Butterworth digital con una frecuencia de corte de $15Hz$, para quitar ruido del movimiento de mecanismos, el filtro pasa-bajas con una frecuencia de corte de $500Hz$, ya que la señal EMG tiene una energía útil limitada de 0 a $500Hz$, y finalmente se aplicó un filtro notch para remover las señales dentro de los $55-66Hz$, eliminando el ruido causado por la linea de alimentación.

C. Agarres

Hay un gran número de formas de agarre, que la mano humana realiza y hay muchos estudios con sus propias definiciones de agarres básicos, la clasificación elegida fue la dada por Vuskovic [7], de la cual sólo se tomarán tres tipos: cilíndrico, esférico y de precisión, como se ilustra en la figura (2).

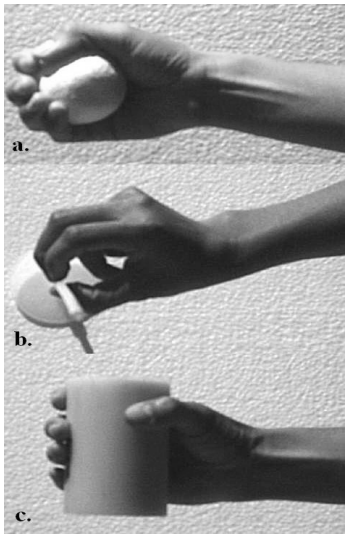


Fig. 2. Agarres: a. Esférico, b. Precisión c. Cilíndrico

3. Análisis de sistemas caóticos

Gran parte de los métodos de estudio de series de tiempo no lineales están basados en la teoría de los sistemas

dinámicos; desde este punto de vista se entiende que la evolución del tiempo está definida en el espacio de estados. Si deseamos modelar la dinámica del sistema, es de gran ayuda establecer un espacio de estados de manera que al especificar un punto en este espacio se esté especificando un estilo del sistema que estamos estudiando y viceversa. Luego, es posible estudiar la dinámica del sistema, estudiando la dinámica de los puntos correspondientes en el espacio de estados, ya que esta es la manera óptima de estudiar sus propiedades dinámicas y geométricas.

A. Reconstrucción del atractor en el espacio de estados

Existen distintos métodos y técnicas por los cuales a partir de la serie de tiempo podemos crear un conjunto de vectores que representen al sistema en el espacio de estados; esta tarea se llama reconstrucción del espacio de estados.

Uno de los problemas al comenzar la reconstrucción del espacio de estados consiste en convertir las mediciones en vectores de estados y uno de los métodos que existe para solucionarlo es el de coordenadas retrasadas (delays).

El verdadero atractor del sistema es construido graficando la evolución del vector en el espacio de estados.
Se suele contar con una única serie de tiempo de mediciones $x(n)$, que es el registro de la variable. Esta serie de tiempo de observaciones es el resultado de una función arbitraria de medición $M(\bullet)$ que opera sobre el vector de estado $x(n)$ de la siguiente manera:

$$s(n) = M(x(n)) \tag{1}$$

Para reconstruir el atractor tenemos que recurrir al teorema de coordenadas retrasadas. En esencia, el teorema garantiza que aplicando el método de retardos sobre las mediciones de la serie de tiempo, se puede reconstruir la dinámica original, pero bajo ciertas condiciones. En el método de coordenadas retrasadas se crea un vector de coordenadas para cada valor de la serie utilizando los mismos valores de la serie de tiempo pero retrasados en tiempo:

$$y(n) = [x(n), x(n+T)], x(n+2T), ..., s(n+(d_E-1)T)), \tag{2}$$

Donde d_E se llama dimensión de empotramiento y T es conocido como el tiempo de retardo. Este último es un múltiplo del tiempo de muestreo T_S. Mediante este empotramiento, es posible reconstruir la existencia de d_E y T de manera que el mapeo de $x(n)$ a $x(n+T)$ sea posible.

Como resultado, tenemos una serie de vectores que representan los distintos estados del sistema en el espacio de estados:

$$Y = y(1), y(2), y(3), ..., y(N-(d_E-1)T)), \tag{3}$$

donde N es la longitud de la serie original. La idea de esta reconstrucción es capturar los estados originales del sistema, para cada uno de los momentos que realizamos las observaciones del sistema [4], [3].

B. Parámetros de empotramiento

El proceso de búsqueda de los parámetros de empotramiento comienza estimando el tiempo de retardo, ya que es necesario, para después calcular la dimensión de empotramiento del atractor.

1) Tiempo de retardo: El tiempo de retardo τ es definido como retardo necesario para alcanzar la independencia de los elementos del vector de coordenadas retrasado de $s(n)$, es decir, τ da noción de la calidad de tiempo mínimo necesario para que el atractor del sistema pueda desplegarse en toda su magnitud. Si τ es demasiado pequeño, entonces los elementos de este vector son muy similares y son más propensos a los efectos de ruido, ya que varios elementos del vector se superponen unos a otros ocultando la forma real del atractor, por otro lado, si tau es demasiado grande entonces los elementos son demasiado disímiles y los vectores tenderían a ocupar todo el espacio de estados. Esto también causa dificultades en la reconstrucción del atractor.

2) Dimensión de empotramiento: Al realizar el análisis de la dinámica del sistema, nuestro deseo es poder determinar la dimensión del espacio donde se tenga las coordenadas suficientes como para poder desplegar las órbitas del atractor del sistema sin traslapes, ya que estos ocurren al proyectar el atractor en las dimensiones más bajas del espacio. Es necesario examinar el conjunto de datos, identificando cuándo ocurren estos traslapes no deseados. La dimensión menor, en donde se despliega el atractor libre de traslapes es llamada la dimensión del espacio de reconstrucción (empotramiento), d_E.

C. Método de información mutua promedio

Para determinar el tiempo de retardo se utilizó el método de información mutua promedio. Este método se basa en la idea de Shanon sobre información el cual se basa en la idea del cálculo de la información mutua entre dos valores (mediciones), x_n y x_{n+T}, la cual es la cantidad aprendida por x_{n+T} acerca de x_n para algún n. El promedio de esta información calculada entre el conjunto de mediciones x_n y el conjunto de mediciones x_{n+T} sobre los valores del conjunto es llamado información mutua promedio [4], [3].

El promedio ponderado de toda esta información estadística está dado por:

$$I(T) = \sum_{x_n, x_{n+T}} P(x_n, x_{n+T}) log_2 \left[\frac{P(x_n, x_{n+T})}{P(x_n))P(x_{n+T})} \right] \tag{4}$$

donde $P(\bullet)$ es la densidad de probabilidad de una medición y $P(\bullet, \bullet)$ la densidad de probabilidad conjunta de dos

mediciones. La prescripción para determinar si los valores de $x(n)$ y $x(n+T)$ son suficientemente independientes tal que se puedan usar para reconstruir el atractor $y(n)$, es tomar T donde el primer mínimo de $I(T)$ ocurre.

Dado que $I(T) \geq 0$ y a medida que T es mayor el comportamiento caótico de la señal hace que $x(n)$ y $x(n+T)$ sean cada vez más independientes hasta tender a cero. Así al calcular los valores de $I(T)$ para valores crecientes de T y tomar el primer mínimo dado en $I(T)$ es tomado como un buen estimado del tiempo de retardo.

D. Método de falsos vecino cercanos

Podemos definir d_E como la dimensión necesaria para poder desplegar el atractor en su plenitud, de manera que no existan falsos vecinos. Este método estima d_E observando la estructura geométrica del atractor a medida que incrementa la dimensión de empotramiento. Este método analiza el atractor del sistema completamente y mide el porcentaje de falsos vecinos para cada dimensión. Un buen estimador de d_E es encontrado donde el porcentaje de falsos vecinos cercanos se aproxima a cero.

El método trabaja de la siguiente manera; se supone que se ha realizado la reconstrucción del vector de estados para la dimensión d con valores de datos.

$$Y = y(1), y(2), y(3), ..., y(N-(d_E-1)\tau)), \qquad (5)$$

Utilizando el tiempo de retardo sugerido por el método de información mutua promedio. Se examinan los vecinos cercanos en el espacio de estados del vector $y(k)$ para el momento k. Estos vectores tendrán la forma:

$$y^{NN}(k) = \left[S^{NN}(k), S^{NN}(k+\tau), ..., S^{NN}(k+(d-1)\tau)\right] \qquad (6)$$

si el vector $y^{NN}(k)$ es realmente un vecino de $y(k)$, entonces éste llegó allí debido a la dinámica original del sistema. Este es el vector anterior o posterior a $y(k)$ siguiendo la órbita, siempre que los intervalos del tiempo sobre la órbita sean lo suficientemente pequeños o bien este vector llegó a la velocidad de $y(k)$ tras haber evolucionado a lo largo de la órbita y alrededor del atractor.

Si el vector $y^{NN}(k)$ es un falso vecino de $y(k)$ que llegó a este vecindario debido a la proyección desde una dimensión mayor porque la presente dimensión d no llega a desplegar el atractor en su totalidad, entonces al moverse a la próxima dimensión $d+1$ es posible que este falso vecino se mueva fuera del vecindario de $y(k)$. Observando el punto $y(k)$ y preguntando para qué dimensión han desaparecido todos los falsos vecinos, iremos eliminando las órbitas más bajas del sistema, hasta finalmente identificar la dimensión d_E donde el atractor es desplegado.

Es necesario tener un criterio para decidir cuándo dado un punto $y(k)$ y su vecino próximo $y^{NN}(k)$, visto desde la dimensión d, es cercano o lejano en la dimensión $d+1$. Al pasar de la dimensión d a la dimensión $d+1$, el componente adicional en el vector $y(k)$ es justamente $s(k+d\tau)$ y en el vector $y^{NN}(k)$ el componente agregado es $s_{NN}(k+d\tau)$. Al comparar la distancia entre los vectores $y(k)$ y $y^{NN}(k)$ en la dimensión d, con la distancia entre estos mismos vectores en la dimensión $d+a$, se puede establecer fácilmente cuál es un falso vecino. Para lograr esto, es necesario comparar $\mid s(k+d\tau) - s_{NN}(k+d\tau) \mid$ con la distancia Euclideana $\mid y(k) - y^{NN}(k) \mid$ entre vecinos cercanos de la misma dimensión d. Si la distancia adicionada es grande, comparada con la distancia de dimensión d entre vecinos cercanos, entonces tenemos un falso vecino. Si en cambio, esta distancia no es grande, tenemos un vecino real [4], [3].

El cuadrado de la distancia Euclideana, entre puntos vecinos próximos, vista desde la dimensión d es:

$$R_d(k)^2 = \sum_{m=1}^{d} [s(k+(m-1)\tau) - s^{NN}(k+(m-1)\tau)]^2,$$

mientras que para $d+1$ es:

$$\begin{aligned} R_{d+1}(k)^2 &= \sum_{m=1}^{d+1} [s(k+(m-1)\tau) - s^{NN}(k+(m-1)\tau)]^2 \\ &= R_d(k)^2 + \mid s(k+d\tau) - sNN(k+d\tau) \mid^2 \end{aligned}$$

La distancia entre puntos tomado desde $d+1$ relativa a la distancia en la dimensión d es:

$$\sqrt{\frac{R_{d+1}^2 - R_d(k)^2}{R_d(k)^2}} = \frac{\mid s(k+d\tau) - sNN(k+d\tau) \mid}{R_d(k)}$$

Cuando este valor es mayor a un umbral determinado, estamos ante un falso vecino.

4. Resultados experimentales

A. Obtención de señales

Las SME que se seleccionaron para ser analizadas corresponden a los movimientos cilíndrico, esférico y de precisión del dedo pulgar, colocando los electrodos como se menciona previamente (Fig. 1), las cuales se pueden observar en las figuras: 3, 4 y 5, en las cuales se describen los movimientos realizados por el dedo que representan el reposo, agarre y reposo, en total son 5 segundos de cada serie que contiene 50000 muestras.

Para realizar el análisis de cada serie y obtener características significativas de cada movimiento, se dividió cada serie en tres partes, dos de las cuales corresponden a la mano en reposo (p1 y p3) y la última con la mano realizando el movimiento (p2). Con las series separadas se tiene que calcular el tiempo de retardo y la dimensión de empotramiento, para eso se utilizaron los métodos de promedio de

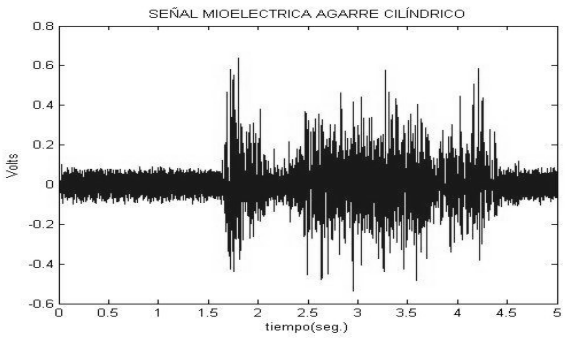


Fig. 3. Agarre de cilíndrico

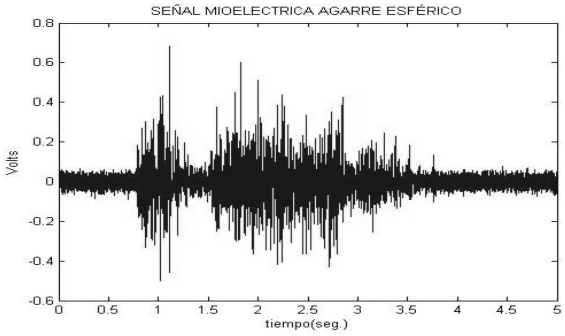


Fig. 4. Agarre esférico

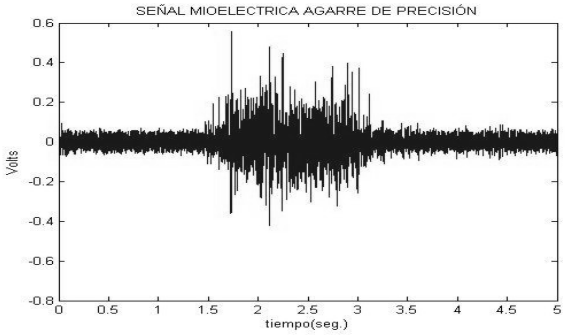


Fig. 5. Agarre de precisión

Movimiento	Serie	Tiempo de retardo	Dim. empotramiento
Cilíndrico	p1	19	14
	p2	27	26
	p3	19	14
Esférico	p1	13	14
	p2	26	16
	p3	14	24
Precisión	p1	18	15
	p2	22	25
	p3	23	16

TABLA I
TABLA DE VALORES DEL TIEMPO DE RETARDO Y DIMENSIÓN DE EMPOTRAMIENTO.

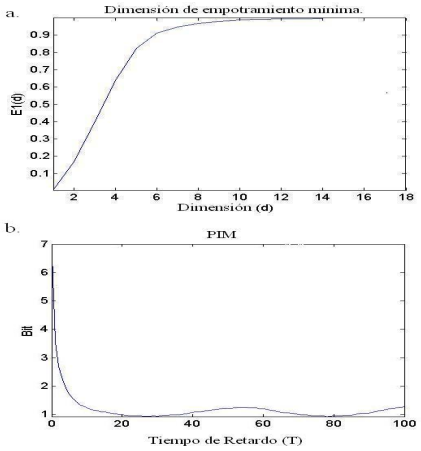


Fig. 6. Gráfica que muestra a. falsos vecinos cercanos, b. promedio de información mutua, de mano en reposo.

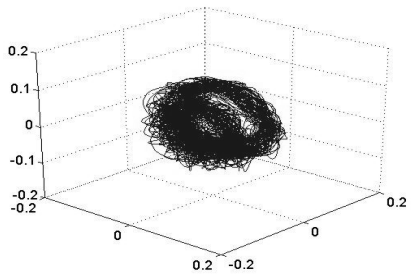


Fig. 7. Gráfica que muestra la proyección de la reconstrucción del atractor, en reposo.

información mutua y falsos vecinos cercanos [3]. Estos dos parámetros son necesarios para poder calcular la dimensión de correlación y el máximo exponente de Lyapunov.

En la Tabla I, se muestran los valores para el tiempo de retardo y la dimensión de empotramiento respectivos para cada serie. Los algoritmos para calcular estos valores se implementaron en Matlab 7.0.

Las gráficas correspondientes a cada caso se muestran a continuación:

5. Conclusión

Se presenta un análisis de series de tiempo en los cuales se determinan la dimensión donde se encuentra empotrado el atractor de dicha serie, y el tiempo de retardo óptimo, estos se han calculado aplicando dos métodos los cuales son el promedio de información mutua y falsos vecinos cercanos, aplicados a las señales filtradas. Para realizar la reconstrucción del atractor se utilizó el método de coordenadas retrasadas, con el cual podemos observar la proyección del mismo en tres dimensiones.

Las gráficas muestran el comportamiento no lineal que sugiere que pertenecen a un atractor extraño en el cual se muestra el comportamiento dinámico del sistema. Se puede observar que los atractores del dedo pulgar en movimiento difieren de los del dedo en reposo.

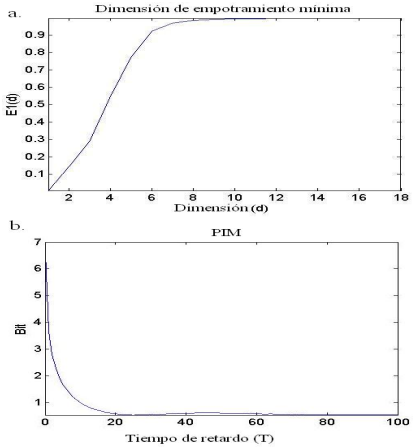


Fig. 8. Gráfica que muestra a. falsos vecinos cercanos, b. promedio de información mutua, agarre cilíndrico.

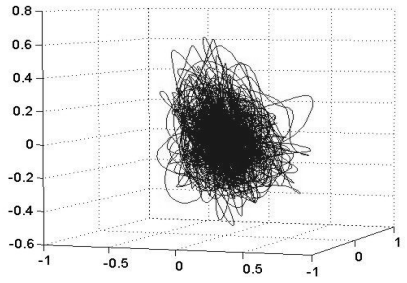


Fig. 9. Gráfica que muestra la proyección de la reconstrucción del atractor, agarre cilíndrico.

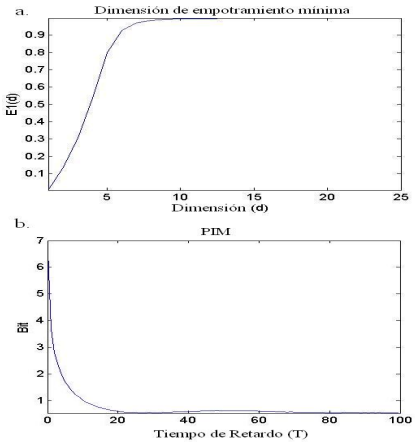


Fig. 10. Gráfica que muestra a. falsos vecinos cercanos, b. Promedio de información mutua, agarre esférico.

Las características obtenidas servirán para que posteriormente se realice un análisis usando técnicas para determinar la distancia entre órbitas y poder realizar una clasificación propia de cada agarre.

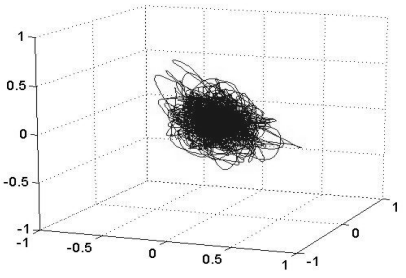


Fig. 11. Gráfica que muestra la proyección de la reconstrucción del atractor, agarre esférico.

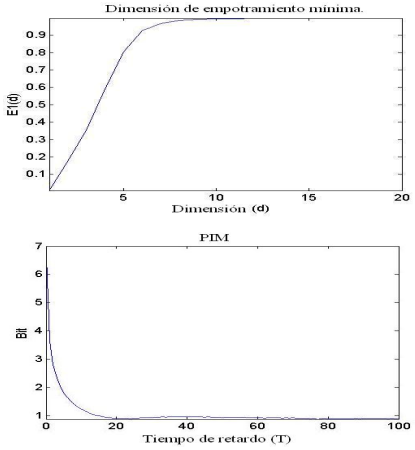


Fig. 12. Gráfica que muestra a. falsos vecinos cercanos, b. promedio de información mutua, agarre de precisión.

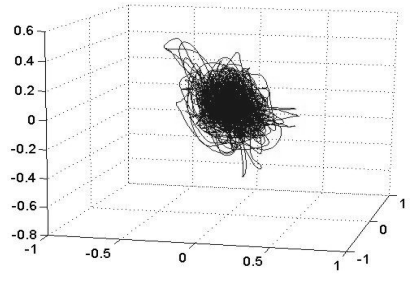


Fig. 13. Gráfica que muestra la proyección de la reconstrucción del atractor, agarre de precisión.

6. Agradecimientos

Este trabajo fue apoyado parcialmente por CONACYT y PROMEP.

Referencias

[1] B. Hudgins and P. Parker. *A New Strategy for Multifunction Myoelectric Control*. IEEE Transactions on biomedial Engineering. Vol. 40. No.1. Enero 1993.

[2] R.S. Parker and L.O. Chua. *Practical Numerical Algoritms for Chaotic Systems*, Springer-Verlag, New York, 1989.

[3] J.C. Sprott. *Chaos and Time Series Analysis*, Oxford University Press, 2003.

[4] H.D.I. Abarbanel. *Analysis of observed chaotic data*, Springer-Verlag, 1995, pp.13-65.
[5] H.P. Laudin. *Electromyography*, Elsevier, Vol. 5 Reiss Joshua D.
[6] H. Rouviere and A. Delmas. *Anatomía Humana. Descriptiva, Topográfica y funcional*, 9°Ed., Masson, 1991, pp. 29-39,134,273-290.
[7] M.I. Vuskovic, A.L. Pozos and R. Pozos. *Classification of Grasp Modes Based on Electromyographic Patterns of Preshaping Motions*. In Proceedings of the 1995 IEEE International Conference on Systems, Man and Cybernetics, 89-95, 1995.
[8] C.J. De Luca. *The use of Surface Electroyography in Biomechanics*, The international Society for Biomechanics, DelSys Inc., 1993.
[9] K. Englehart, B. Hudgins, P.A. Parker, M. Stevenson. *Classification of the myoelectric signal using time-frequency based representations*. Med. Eng. Phys., 21:431-438, 1999.
[10] Ed. Akay M. *Nonlinear Biomedical Signal Processing, Dynamic Analysis and Modeling*, IEEE Press., Vol.II, 2001.

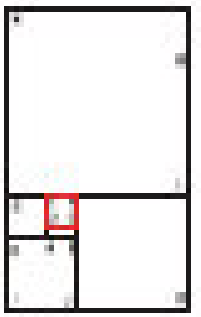
Las Matemáticas De Dali

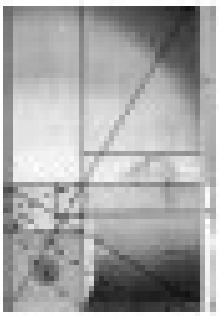
Autor: Mat. María de Lourdes Hernández Campos
Coautor: Mat. Esteban Rubén Hurtado Cruz

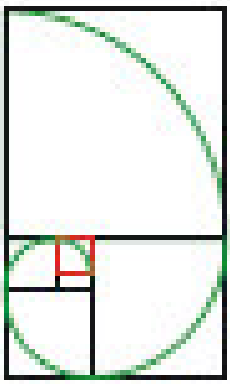
En este trabajo se muestran algunas expresiones artísticas de Dalí, en las que plasma su sentir sobre la Topología, fractales, Geometría Proyectiva, Espiral Logarítmica, Hipercubo y Triángulo Áureo, reflejando su conocimiento de los mismos.

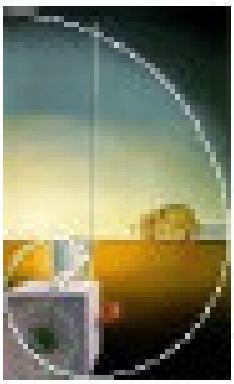
1. Espiral áurea

Partiendo del cuadrado CDEF construimos un rectángulo áureo ABEF. Si a éste le añadimos sobre su lado mayor un cuadrado, obtenemos otro rectángulo áureo BFGH.
Si después, con este rectángulo repetimos el proceso, obtenemos otro rectángulo áureo. Este proceso se puede reproducir indefinidamente, obteniendo así una sucesión de rectángulos áureos encajados.
Una vez construida esta sucesión de rectángulos áureos encajados, si unimos mediante un arco de circunferencia dos vértices opuestos de cada uno de los cuadrados obtenidos, utilizando como centro de la misma, otro de los vértices del mismo cuadrado, obtenemos una curva muy similar a una espiral logarítmica. Es la famosa Espiral áurea (o espiral de Durero).

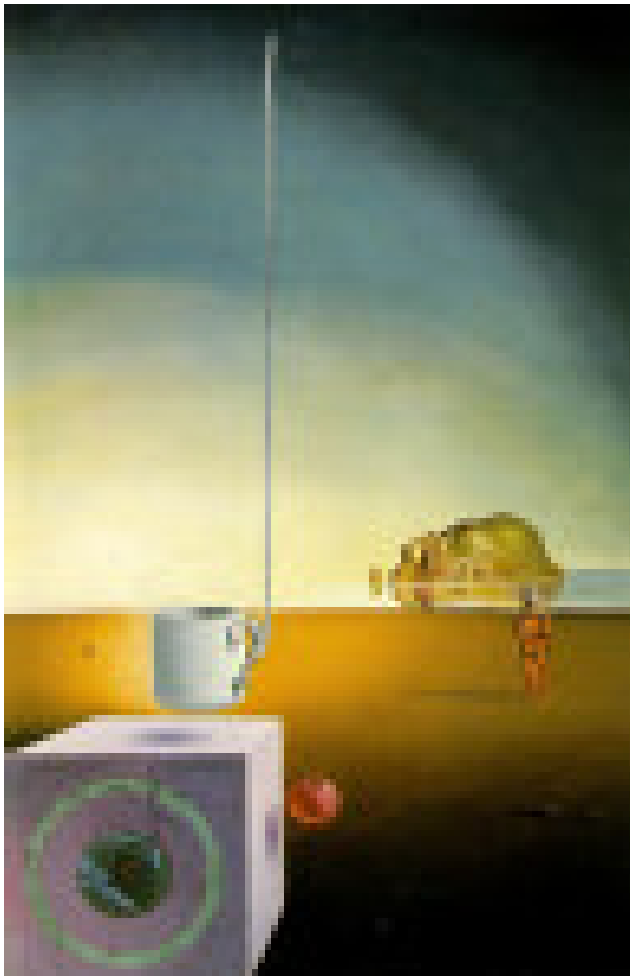






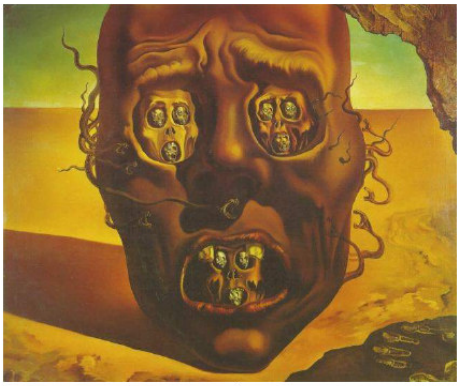


Esta obra se puede considerar como un homenaje, no carente de humor, al Rectángulo de Oro. No sólo se puede descomponer el cuadro en una serie de rectángulos áureos sino que, además, los diferentes elementos del cuadro, son la llave que permite reconstruir estos rectángulos. A partir de la “taza”, se obtiene una sucesión de rectángulos áureos que nos llevan a una espiral áurea que acaba en la sombra negra de la parte alta del c u a d r o .
Por otra parte, ese “anexo inexplicable” del título que sale del “asa da taza” y que obliga a prolongar el cuadro hacia arriba, es en realidad totalmente explicable: resulta que las dimensiones del cuadro están en proporción áurea, siendo el anexo el elemento que justifica tales dimensiones.



2. Fractales

Técnicamente, un fractal es un objeto que no pierde su definición formal a medida que es ampliado, manteniendo su estructura idéntica a la original. (Por ejemplo, una circunferencia parece perder su curvatura a medida que ampliamos una de sus partes).
Hay dos características importantes que ayudan a comprender la estructura y concepción de un fractal: su área o superficie es finita, es decir, tiene límites, y por el contrario, su perímetro o longitud es infinita.



"El rostro de la guerra"; 1940. Óleo sobre lienzo; 64 x 79 cm.
Rotterdam, Museo Boymans van Beuningen

Dalí parece ser el primer artista que pintó un fractal: era su visión de la guerra.
En esta obra los ojos y la boca contienen una cara, cuyos ojos y boca contienen, a su vez, una cara cuyos ojos y boca contienen una cara. Es un ejemplo obvio de fractal en el arte.
Un análisis del trabajo revela que el fractal representado es el llamado "polvo de Cantor", generado por tres contracciones con factor de contracción aproximado de 0.21, y de dimensión Hausdorff 0,705. Pertenece a los triángulos de Sierpinsky.

3. Topología

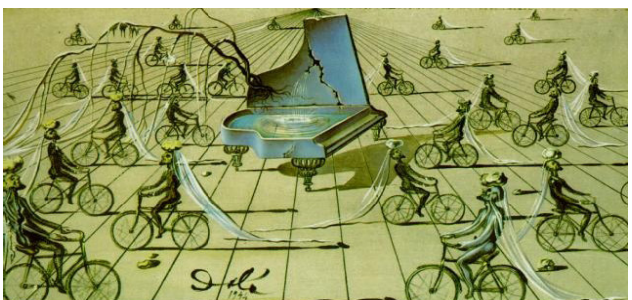
La topología es una rama de las matemáticas que estudia las propiedades de las figuras geométricas o los espacios que no se ven alterados por transformaciones continuas y biyectivas, y de inversa continua (homeomorfismos).
Es decir, en topología está permitido doblar, estirar, encoger, retorcer… los objetos para pasar de uno a otro, pero no se permiten transformaciones que puedan provocar una discontinuidad como por ejemplo romper ni separar lo que estaba unido (la transformación debe ser continua) ni pegar lo que estaba separado (la inversa también debe ser continua).
Por ejemplo, en topología un triángulo es lo mismo que un cuadrado, ya que podemos transformar uno en otro sin romper ni pegar, o una taza es lo mismo que una rosquilla, para la topología estos cuerpos son iguales y les llamamos homeomorfos.

Pero una circunferencia no es lo mismo que un segmento, ya que habría que partirla por algún punto.
Podemos referirnos a la Topología como una “geometría cualitativa” en la que no se trabaja con nociones cuantitativas como: longitud, ángulo, área, volumen, etc. sino que se centra en cuestiones cualitativas como por ejemplo, si tiene agujeros o no, si tiene borde, o si se puede partir en componentes conexas. Atendiendo a estas características se hace una clasificación topológica de las superficies.

4. Geometría proyectiva

La geometría proyectiva surgió en el Renacimiento como una necesidad de los pintores de dar rigor matemático al dotar a sus cuadros de perspectiva.
En la geometría del sistema visual las paralelas no existen, por lo tanto, necesitamos una geometría en la que dos líneas se corten. El lugar donde las paralelas parecen cortarse está en el infinito, dado que en el plano euclidiano la infinitud no existe, es necesario añadir puntos “ideales” en la infinitud del plano. Esos puntos “del infinito” forman una línea adicional que también tenemos que añadir al plano euclídeo.
En la perspectiva, las líneas de profundidad en una obra, cuando son prolongadas, se encuentran en un determinado punto. Este punto se llama **punto de fuga** y se encuentra, casi siempre, en la **línea del horizonte**, esto es, una línea horizontal, a la altura de nuestra vista.

“..Dalí conocía perfectamente la geometría en muchos sentidos, era un maestro de las formas precisas y de la geometría descriptiva, y podía realizar precisos estudios arquitectónicos basados en estructuras matemáticas.
Conocía perfectamente la perspectiva, razón por la cual después la podía distorsionar tan bien...”.
Dalí se presentó a los surrealistas en París con esta composición. Es una obra organizada alrededor de una plataforma en perspectiva sobre la que se sitúan unas formas blandas.
Él comenta de esta obra: “...La numeración en las figuras probablemente se corresponde con mi inconsciente interés en el sistema métrico. (...) En aquella época yo estaba ensimismado con los sistemas de pesos y medidas, y los números iban apareciendo por todas partes. También estaba absorto con el sistema métrico, la división numérica de las cosas mundanas...”



5. Hipercubo

Imaginemos un punto en el espacio: tiene dimensión cero porque no tiene anchura, longitud o altura y es infinitamente pequeño.
Tomamos un punto y lo trasladamos sobre una línea recta, una distancia de una unidad, por ejemplo, consiguiendo así un segmento. Todo segmento es unidimensional, sólo tiene una dimensión: la longitud, y todos los segmentos tienen el mismo ancho y altura que es infinitamente pequeña. Si se ampliara infinitamente el segmento cubriría el espacio unidimensional
Movamos ese segmento una unidad en dirección perpendicular a él. Generamos un cuadrado unitario. Todos los cuadrados son de dos dimensiones, se diferencian entre ellos por la anchura y la longitud y todos tienen la misma altura, que es infinitamente pequeña. Los bordes son de la misma longitud, y todos los ángulos son rectos. Si se ampliara el cuadrado infinitamente, cubriría el espacio de dos dimensiones.
Tomemos el cuadrado y trasladémoslo una unidad en una tercera dirección, perpendicular a las dos primeras, creando así un cubo unitario. Todos los cubos son tridimensionales porque se diferencian entre ellos por tres medidas anchura, longitud, y altura. Como en el cuadrado, todos los bordes de un cubo tienen la misma longitud, y todos los ángulos son rectos. Si se ampliara el cubo infinitamente en todas las direcciones, cubriría el espacio tridimensional.
Ahora, el paso final. Tomamos el cubo y lo estiramos en otra dirección perpendicular a las tres primeras. ¿Pero como podemos hacer esto? Es imposible hacerlo dentro de las restricciones de la tercera dimensión. Ahora bien , dentro de la cuarta dimensión es posible. el espacio generado por este movimiento es un hipercubo (tesseract) con cuatro aristas perpendiculares cortándose en cada vértice.

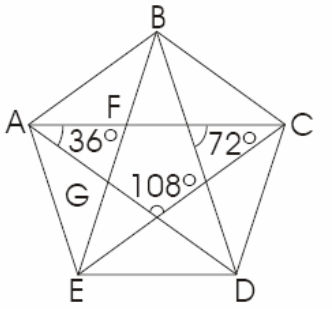
En 1953, inspirado por su viaje a Nueva York, Dalí anunció que: iba a pintar un Cuadro, que él mismo calificó de sensacional: “Un Cristo explosivo, nuclear e hipercúbico”. Este podría ser el primer cuadro que reconcilia una fórmula neoclásica en la técnica con un contenido compuesto por elementos cúbicos.

Los distintos elementos del cuadro y su disposición, dan lugar a discusiones sobre a su intencionalidad:
! la composición de la cruz, (concebida como ocho cubos adosados por una cara);
! El suelo embaldosado, en donde vemos la proyección en forma de cruz latina, (como ilustración del paso a dos dimensiones); y ! la posición del Cristo (desplazado para que la sombra se sitúe en el centro).
El cuadro de Dalí, representa una construcción inusual para las artes religiosas pero conocida en geometría como un hipercubo desplegado ou tesseract, una imagen tridimensional de la figura cuadrimensional inimaginable en una realidad física, pero que tiene un lugar legítimo en construcciones matemáticas abstractas.
De todas las formas, lo único seguro es la fascinación de Dalí por combinar la espiritualidad y la técnica expresada como geometría o matemáticas.

En esta obra, Dalí cuadruplica la inscripción visigótica “Silo princeps fecit” (o príncipe Silo me hizo) de una iglesia asturiana en el mismo formato laberíntico que permite leerla de 2024 formas distintas. Pura combinatoria.

6. Triángulo Áureo

Consideremos un pentágono regular en el cual dibujamos las diagonales. En esta figura sólo aparecen tres ángulos diferentes, miden 36º, 72º e 108º. Hay varios tipos diferentes de triángulos isósceles, de los cuales seleccionamos tres: los triángulos ABE, ABF y AFG (el resto de triángulos son semejantes a alguno de estos). Finalmente, hay cuatro segmentos diferentes en estos triángulos, que llamaremos: BE=a, AB=AE=b, AF=BF=AG=c y GF=d. Las longitudes de estos segmentos cumplen: a>b>c>d.

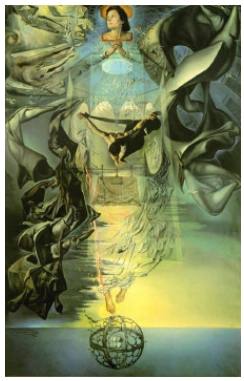


Si consideramos cada uno de estos triángulos por separado y aplicamos el teorema del seno, obtenemos las siguientes proporciones : a/b = b/c = c/d = ϕ.
Es decir, una vez ordenadas las longitudes de los cuatro segmentos de mayor a menor la razón entre cada una de ellas y la siguiente, es constante e igual al número de oro.

Por tanto, en los triángulos BED y DEF sus lados están en proporción áurea reciben el nombre de “triángulos áureos”.

La perspectiva desde la que se contempla la cruz determina en buena medida la insólita composición de esta pintura.
Dalí la explicó así: “Primeramente, en 1950, tuve un 'sueño cósmico' en el que vi esta imagen de colores que, en mi sueño, representaba el 'núcleo del átomo'. Este núcleo tomó más tarde un sentido metafísico, yo lo consideró la unidad misma del universo, ¡Cristo! En segundo lugar, gracias a las indicaciones del padre carmelita Bruno, vi el Cristo dibujado por San Juan de la Cruz y concebí de forma geométrica un triángulo y un círculo que resumían estéticamente todas mis experiencias anteriores, e inscribí mi Cristo en ese triángulo”.
De hecho, en el análisis de la obra, se encuentran al menos dos triángulos áureos, uno que enmarca el Cristo y otro que enmarca la cruz.
La obra se divide claramente en dos zonas, la etérea (parte superior) y la cotidiana (paisaje inferior), ambas separadas por distinta iluminación, pero unidas en el ojo del observador por la coincidencia del punto de fuga.

Dalí pinta la segunda versión del “Cristo de San Juan de la Cruz” como detalle central de su cuadro “Asunción corpuscularia lapislazulina”, pintado durante el verano de 1952.



ALGUNAS COMPACTACIONES DEL RAYO CON LA PROPIEDAD DEL PUNTO FIJO

Jesús Fernando Tenorio Arvide y María de Jesús López Toriz
Facultad de Ciencias Físico Matemáticas (FCFM)
Benemérita Universidad Autónoma de Puebla (BUAP)
Puebla, Pue., México, 72570.
jtenorio@fcfm.buap.mx, mtoriz@fcfm.buap.mx

Resumen

Demostraremos que si $X = Y \cup S$ es una compactación del rayo $S = [0, \infty)$ con residuo un espacio métrico compacto y arco conexo Y con la propiedad del punto fijo, entonces X también tiene la propiedad del punto fijo.

1. Definiciones, ejemplos, notaciones y el teorema.

1. Definición. Un espacio métrico compacto X *es una compactación de un espacio métrico* Z, si X contiene un subespacio Z' el cual es homeomorfo a Z y Z' es denso en X (i. e. $\overline{Z'} = X$). El conjunto $Y = X \setminus Z'$ se llama *residuo de la compactación*.

2. Ejemplo. El intervalo $[0, 1]$ es una compactación de $(0, 1)$. En efecto, ponemos $X = [0, 1]$, $Z = (0, 1)$, $Z' = (0, 1)$ y $\overline{Z'} = [0, 1]$. Por otro lado, también el intervalo $[0, 1]$ es una compactación de $[0, \infty)$ y de $\mathbb{R}$ (aquí, $Z = [0, \infty) \approx [0, 1) = Z'$ o bien $Z = \mathbb{R} \approx (0, 1) = Z'$, respectivamente).

3. Ejemplo. Sea $f(x) = sen\frac{1}{x}$ y denotemos por $G(f)$ la gráfica de la función f, es decir, $G(f) = \{(x, f(x)) : x \in (0, 1]\}$. Sea X la cerradura en $\mathbb{R}^2$ de $G(f)$. Tenemos que X es una compactación de $[0, \infty)$ (ponemos $Z = [0, \infty) \approx G(f) = Z'$).

4. Notación. $X = Y \cup Z$ denota una compactación de Z con residuo Y.

5. Definición. Un espacio métrico X *es arco conexo* si cada dos puntos de X pueden unirse mediante un arco en X, es decir, si $p, q \in X$ entonces existe un homeomorfismo $h : [0,1] \to X$ tal que $h(0) = p$ y $h(1) = q$. Una *arco componente* en un espacio es un subconjunto arco conexo maximal.

6. Definición. Un espacio métrico X *tiene la propiedad del punto fijo (p.p.f.)*, si toda función continua $f : X \to X$ tiene un punto fijo, es decir, existe un punto $x_0 \in X$ tal que $f(x_0) = x_0$.

7. Ejemplo. -El intervalo $[0,1]$ tiene la p.p.f.
-Todo arco tiene la p.p.f.
-La circunferencia unitaria en $\mathbb{R}^2$ no tiene la p.p.f.

8. Definición. Decimos que S *es un rayo* si $S \approx [0, \infty)$.

9. Teorema. *Sea $X = Y \cup S$ una compactación del rayo $S \approx [0, \infty)$ con residuo un espacio métrico compacto y arco conexo Y. Si Y tiene la p.p.f., entonces X tiene la p.p.f.*

Demostración. Sea $f : X \to X$ una función continua. Se tiene que Y y S son las arco componentes de X (pues S es arco conexo, Y es arco conexo, $Y \cap S = \emptyset$ y $X = Y \cup S$). Dado que la arco conexidad se preserva bajo funciones continuas, se tiene que $f(Y)$ es arco conexo. Así, ocurren dos casos:

(1) $f(Y) \subset Y$, o bien
(2) $f(Y) \subset S$.

Si ocurre (1), consideremos la función $f|_Y : Y \to Y$. Entonces, como $f|_Y$ es continua y Y tiene la p.p.f., existe un punto $y_0 \in Y$ tal que $f(y_0) = y_0$. Así, f tiene un punto fijo y terminamos.

Si ocurre (2), entonces resulta que $f(S) \subset S$.

En efecto, dado un punto $y \in Y$, se tiene, por la densidad de S en X, que existe una sucesión de puntos en S, $\{s_n\}_{n=1}^{\infty}$, que converge al punto y. Puesto que $f(y) \in S$ (por el caso (2)), S es un conjunto abierto y la sucesión $\{f(s_n)\}_{n=1}^{\infty}$ converge al punto $f(y)$, se tiene que existe $N \in \mathbb{N}$ tal que $f(s_N) \in S$. Luego, $f(s_N) \in f(S) \cap S$. De aquí, como S es arco componente y $f(S)$ es arco conexo, concluimos que $f(S) \subset S$.

Así, $f(Y) \cup f(S) \subset S$. Por lo que $f(X) \subset S$. Puesto que $f(X)$ es compacto y conexo, entonces existe un arco $A \subset S$ tal que $f(X) \subset A$.

Ahora, consideremos la función $f|_A : A \to A$. Luego, como $f|_A$ es continua y A tiene la p.p.f., obtenemos un punto $a_0 \in A$ tal que $f(a_0) = a_0$. Así, f tiene un punto fijo.

Por lo tanto X tiene la p.p.f. ⊡

10. Nota. El Teorema 9 sigue siendo válido si cambiamos $S \approx [0, \infty)$ por $S \approx (-\infty, \infty)$.

11. Nota. La misma técnica se aplica para la demostración del siguiente hecho:

12. Teorema. *Si $X = S \cup Y$ es una compactación del rayo $S \approx [0, \infty)$ con residuo el espacio métrico compacto y arco conexo Y tal que $Y \times [0, 1]$ tiene la p.p.f., entonces $X \times [0, 1]$ tiene la p.p.f.*

PGSM
Puebla, MÉXICO
Noviembre de 2005

Biorritmo, una Aplicación de la Trigonometría

Mtro. Juan Carlos Macías Romero
Facultad de Ciencias Físico Matemáticas, BUAP
18 sur y Av. San Claudio, Colonia San Manuel, Ciudad Universitaria

jcmacias@fcfm.buap.mx

Resumen

En esta plática daremos algunas ideas que pueden servir para que la enseñanza de las matemáticas en el nivel medio deje a un lado sus viejos hábitos y al mismo tiempo se modernice. La intención es dar un ejemplo de una aplicación de la trigonometría. Más específicamente, hablaremos sobre la teoría del biorritmo y su relación con la función seno. Brevemente esta teoría sostiene que nuestros estados físicos, emocionales e intelectuales, son periódicos y se pueden representar mediante funciones senoidales.

Antes que todo, debe quedar claro, que en esta plática no nos proponemos dictar reglas para enseñar mejor ni queremos proveer al maestro de una fórmula mágica para facilitar la comprensión de las matemáticas por parte del alumno, pero sí daremos algunas ideas que pueden servir para motivar al alumno a interesarse en el estudio de esta disciplina y para facilitar la transmisión de algunos conceptos matemáticos por parte del profesor.

Los problemas de la enseñanza de las matemáticas que se presentan a menudo en las escuelas secundarias, desafortunadamente son muchos y sin duda graves. Entre ellos nos parece necesario considerar dos puntos:

a) La lección de matemáticas resulta en general aburrida, pesada y difícil. Ciertos conceptos no son afirmados, aun cuando el profesor se afane en repetirlos y busque aclararlos con numerosas explicaciones; de algunas propiedades no se entiende inmediatamente el sentido. Es notable “la incomprensión por la matemática” y el “temor a la matemática”.

b) Los jóvenes que actualmente salen de nuestras escuelas secundarias tienen la idea de que las matemáticas consisten, por una parte, en un puro mecanismo, y por la otra, que se trata de una construcción perfecta y completamente terminada, ignorando si se puede hacer o no algún descubrimiento con esta disciplina.

Si reflexionamos sobre la importancia que tiene hoy una cultura matemática, entendiéndola como un hábito mental matemático más que una suma de conocimientos, nos podemos dar cuenta de la responsabilidad que tienen el redactor de programas, los profesores de matemáticas a cualquier nivel, y la escuela toda. Este problema tiene ya sus años, basta remontarse al congreso de 1958 de la sociedad Matemática Belga donde el tema principal fue “La responsabilidad humana del profesor de matemáticas”.

Y claro, la primera responsabilidad recae en el ciclo de la escuela secundaria; aquí, los muchachos de 11 a 14 años no deben ser rechazados de un estudio no muy adecuado a su edad, las inquietudes típicas del preadolescente no deben ser sofocadas, sino servir de impulso para un desarrollo activo del programa. El país tiene necesidad de estos muchachos y nosotros tenemos el deber de transmitirles el lenguaje apasionante y el patrimonio de ideas que encierren las matemáticas.

Por lo antes expuesto, es urgente pensar en algo que nos pueda ayudar para dejar a un lado los viejos hábitos que utilizamos en la enseñanza y que estemos dispuestos de verdad a innovar y a cambiar por el bien de nuestros alumnos y de nosotros mismos.

Estamos realmente convencidos que la actitud de los alumnos puede ser fácilmente dirigida en sentido positivo si nosotros tomamos una actitud positiva también.

De este modo, consideremos las siguientes preguntas

¿Cómo motivar a nuestros alumnos?

¿Cómo hacer para que la clase no sea tan aburrida?

¿Se podrá hacer que la clase de matemáticas sea divertida?

¿Se podrá cambiar alguna actividad del programa por otra que implique el mismo conocimiento?

A continuación daremos un ejemplo que da respuesta a las preguntas anteriores.

El ejemplo tiene que ver sobre una aplicación interesante de la función seno que se encuentra en la llamada teoría del biorritmo.

En la naturaleza hay muchos fenómenos que se rigen por ciclos como el clima, las estaciones, la reproducción de los animales, las cosechas, las mareas, las fases de la luna, etc.

Cada uno de estos ciclos se produce con una periodicidad diferente.

En el caso de los seres vivos estos ritmos se denominan Biorritmos, y existen diferentes biorritmos que afectan nuestro comportamiento en distintas maneras.

El Biorritmo se basa en la idea de los ciclos de la vida aplicados al ser humano, si tenemos en cuenta que en el cuerpo humano se producen distintos tipos de alteraciones biológicas producidas cíclicamente como la respiración, el ritmo cardiaco, el sueño y la vigilia, la menstruación en la mujer, etc.

Al nacer, según la teoría del biorritmo, cada ciclo comienza desde cero y empieza a subir en una fase positiva, durante la cual las energías y las capacidades son altas.

En el caso de los humanos, se ha comprobado estadísticamente que el ciclo físico se repite cada 23 días, el emocional cada 28 días y el ciclo intelectual cada 33 días.

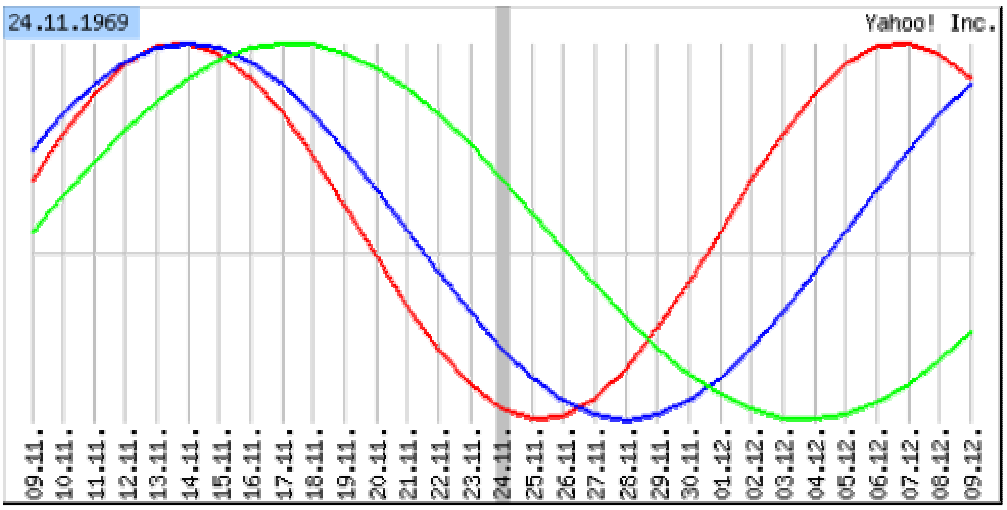
Entonces, los biorritmos son los ciclos fluctuantes en la vida de una persona, en el aspecto físico, emocional e intelectual.

El ciclo físico controla nuestra energía, vitalidad, aguante e iniciativa. Cuando este ciclo se encuentra en la fase alta, nos sentimos mejor, estamos llenos de energía y es menos probable que nos enfermemos. Cuando este ciclo se encuentra en la fase baja, tendemos a cansarnos más y estamos más propensos a enfermarnos.

El ciclo emocional controla nuestra sensibilidad y temperamento. Cuando este ciclo se encuentra en la fase alta, es más probable que estemos alegres, creativos y sensuales.
Cuando este ciclo se encuentra en la fase baja, es más probable que estemos de mal humor, irritables, tristes o deprimidos.

El ciclo intelectual controla nuestra capacidad de pensar. Cuando este ciclo se encuentra en la fase alta, somos capaces de pensar más rápido, resolvemos mejor los problemas, no concentramos más y memorizamos mejor. Cuando este ciclo se encuentra en la fase baja, no tenemos buena memoria, nos resulta difícil concentrarnos y no tomamos las mejores decisiones.

Los tres biorritmos comienzan en una fase positiva en el momento del nacimiento, y continúan con regularidad a lo largo de la vida. Cada ritmo consiste en días "altos", "bajos" y "críticos" (cuando pasa por el punto cero).
(ver figura)



Para leer tu biorritmo, fíjate en la columna gris que representa el día actual. Las líneas representan los diferentes ciclos detallados más abajo y pueden tener tendencias ascendentes o descendentes. La línea divisoria central representa

la media y, por lo tanto, una línea ascendente por encima de la línea central representa que será buen día para determinado ciclo.

El biorritmo consiste de 3 ciclos:

Ciclo físico: fuerza, energía, potencia sexual y física

Ciclo intelectual: facultades mentales, lógica, funciones cerebrales

Ciclo emocional: sentimientos y emociones

Es importante tener en cuenta que la curva no se divide en una mitad "buena" y en otra "mala", sino más bien en una mitad "activa" (arriba) y otra más bien "pasiva" (abajo).

Cada vez que el ciclo cruza el punto cero al pasar de la fase activa a la pasiva, se dice que la persona está en un "día crítico" o en "estado crítico".
En un día crítico, nuestro sistema se encuentra en un estado de desorden y confusión. Las habilidades asociadas están inestables, y la persona debe ser particularmente cuidadosa. Es más probable que las cosas nos salgan mal y que nos ocurran accidentes.

Sin embargo podemos prevenir futuros problemas estando alertas en nuestros días críticos.

La relación entre el biorritmo y la función seno reside en las siguientes fórmulas:

Para calcular nuestro estado físico en cualquier día, t, desde nuestro nacimiento usamos:

$$F = sen\left(\frac{2\pi}{23}t\right)$$

para nuestro estado emocional usamos:

$$E = sen\left(\frac{2\pi}{28}t\right)$$

y para nuestro estado intelectual usamos:

$$I = sen\left(\frac{2\pi}{33}t\right)$$

Usando estas funciones, caracterizamos nuestros días buenos como aquellos para los cuales *F, E* e *I* son positivos y nuestros días malos como aquellos para

los cuales son negativos. Cuanto más cercano el estado sea a +1, tanto mejor el día para esa fase particular de su bienestar. Su estado global se obtiene usualmente promediando los tres valores.

Para usar las fórmulas para el estado del biorritmo, debe usted calcular el numero de días a partir de la fecha de nacimiento dada hasta el día actual.

Un método conveniente para que los alumnos calculen el número de días que han vivido hasta el día actual es recordar que los años bisiestos ocurren cada cuatro años y considerar que el 2004 fue bisiesto.

Así por ejemplo, si una persona nació el 25 de febrero de 1990 y la fecha actual es 24 de noviembre de 2005 entonces este alumno ha vivido 15(365) + días de años bisiestos +días transcurridos a partir del 25 de febrero de 2005 al 24 de noviembre de 2005.

Veamos: ha vivido 5475 + 4 + 271= 5750 días.

Luego, el número de días vividos es t = 5750

Así que, su estado físico está dado por:

$$F = sen\left(\frac{2\pi}{23}(5750)\right) = sen\left(1570.7963\right) = 0$$

su estado emocional está dado por:

$$E = sen\left(\frac{2\pi}{28}(5750)\right) = sen\left(1290.2969\right) = .7818$$

y su estado intelectual está dado por:

$$I = sen\left(\frac{2\pi}{33}(5750)\right) = sen\left(1094.7974\right) = .9988$$

Compruebe que su calculadora esté en modo radianes para obtener los números indicados.

El promedio de estos tres números es .5935. Todas las indicaciones son de que este alumno podrá disfrutar de este día.

Cabe señalar que la esta teoría no está comprobada científicamente. Por lo que se debe tener cuidado con la interpretación de los resultados. Uno como profesor debe recalcarle al alumno que esta teoría se basa sólo en estadísticas, sin embargo, es una aplicación interesante de la función seno.

Consideramos que es una bonita motivación para que el alumno use las matemáticas con relación a su vida y cuerpo. Si bien la interpretación de los resultados no tiene bases matemáticas si que a los estudiantes les parece divertido. Algunos hasta se asombran de saber cuantos días han vivido pues tal vez nunca se lo habían preguntado.

La experiencia de aplicar este ejemplo con los alumnos es satisfactoria y como profesores le podemos sacar bastante provecho.

Consideraciones Sobre Continuidad y el Teorema de Darboux

Francisco Javier Mendoza Torres y Maria Guadalupe Morales Macías
Facultad de Cs. Físico Matemáticas
Benémerita Universidad Autónoma de Puebla
Puebla, Pue., México, 72570
jmendoza@fcfm.buap.mx, lupittah@hotmail.com

Resumen. En este trabajo se presentan las ideas básicas sobre funciones continuas. Se aclara que el que una función cumpla el Teorema del Valor Intermedio no es garantía de continuidad. Se exhiben funciones que cumplen este treorema y no son continuas, para ésto se presenta el Teorema de Darboux, el cual ayuda a entender esta situación.

Palabras clave: función continua, valor intermedio, teorema de Darboux.

1. Introducción

Las funciones continuas constituyen la clase básica de funciones para las operaciones del Análisis Matemático. Los primeros matemáticos que intentaron definir una función continua la consideraban como aquella cuya gráfica no tenía huecos. Por ejemplo, Euler la definió como una curva descrita por el libre deslizamiento de la mano y Cauchy como "una función para la que un cambio infinitesimal en una variable produce un cambio infinitesimal en el valor, es decir, tiene ausencia de saltos". Actualmente la definición básica de función continua se hace sobre espacios topológicos y de estos se particulariza sobre otros tipos de espacios. Nosotros trataremos el tema de las funciones continuas reales de variable real. Así, tenemos que la función $f : I \subset \mathbb{R} \to \mathbb{R}$ será continua en $x_0 \in I$ si

$$\lim_{x \to x_0} f(x) = f(x_0),$$

y será continua en I si lo es en cada punto de I.

Sabemos que la Propiedad del Valor Intermedio es consecuencia de la continuidad de una función y durante algún tiempo se consideró que esta propiedad y la continuidad eran equivalentes. Sin embargo, como expondremos, hay funciones que cumplen la primera y no la segunda. En este sentido, el teorema de Darboux nos proporciona un buen número de funciones que cumplen ésto.

2. El Teorema del Valor Intermedio

Comentemos el siguiente problema: un ermitaño sale de su cabaña a las nueve de la mañana y se dirige a una cueva en la montaña, a la cual llegó a las 5 de la tarde, en donde pernoctó. Durante todo el camino fue marcando el camino recorrido de tal forma que al regresar pise sus propias huellas. Al día siguiente, salió a las 9 de la mañana de regreso a su cabaña, a la cual llego a las 5 de la tarde. El problema consiste en probar que en un mismo momento estuvo, en ambos días, en el mismo lugar. Para resolver este problema podemos usar el **teorema del valor intermedio,** el cual enunciamos a continuación:

Sea $f : [a, b] \to \mathbb{R}$ **una función continua, sea** y_0 **que esta entre** $f(b)$ **y** $f(a)$**, entonces existe** $t_0 \in [a, b]$ **tal que** $f(t_0) = y_0$.

Así, para nuestro problema tenemos lo siguiente: representemos por $f(t)$ el camino recorrido en la ida, de tal forma que en el tiempo t se ha recorrido hasta el

camino $f(t)$, y sea $g(t)$ el camino de regreso con la idea anterior. Ambas funciones tienen como dominio $[0,8]$, que es el período de tiempo que dura el recorrido, además $f(8) = g(0)$ y $f(0) = g(8)$. Al considerar la función $h(t) = f(t) - g(t)$, tenemos que como $h(0) = f(0) - f(8) < 0 < f(8) - f(0) = h(8)$ y como $h(t)$ es continua pues el ermitaño no hace magia, entonces por el TVI, existirá un $t_0 \in [0,8]$ tal que $h(t_0) = f(t_0) - g(t_0) = 0$. Esto es, en el tiempo t_0 el ermitaño estuvo en el mismo lugar.

3. Darboux y las funciones que cumplen la propiedad del Valor Intermedio sin ser continuas

En el siglo diecinueve, los matemáticos consideraron durante mucho tiempo que la propiedad del Valor Intermedio y la continuidad de una función eran equivalentes. En 1875 el matemático francés Jean Gaston Darboux (1842-1917) probó que esta equivalencia no era cierta. Por ejemplo, las funciones

$$f:\mathbb{R}\to\mathbb{R} \qquad\qquad g:\mathbb{R}\to\mathbb{R}$$
$$f(x)=\begin{cases} 2\text{sen}\left(1/x^2\right)-2x^{-1}\cos\left(1/x^2\right) & \text{si } x\neq 0 \\ 0 & \text{si } x=0 \end{cases} \quad \text{y} \quad g(x)=\begin{cases} 2x & \text{si } x<1 \\ \frac{1}{2x^{1/2}} & \text{si } x\geq 1, \end{cases}$$

cuyas gráficas se exhiben abajo, cumplen la propiedad del valor intermedio, sin embargo, por ejemplo en $[-1,1]$ y en $[0,2]$, f y g, respectivamente, no son continuas.

gráfica de $f(x)$

gráfica de $g(x)$

Estas funciones tienen la característica de ser derivadas de funciones. La función f es derivada de

$$h : \mathbb{R} \to \mathbb{R}$$
$$h(x) = \begin{cases} x^2 \text{sen}\left(1/x^2\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

y g es derivada de la función

$$i : \mathbb{R} \to \mathbb{R}$$
$$i(x) = \begin{cases} x^2 & \text{si } x < 1 \\ x^{1/2} & \text{si } x \geq 1. \end{cases}$$

Darboux dió ejemplos de funciones derivadas discontinuas con la propiedad del Valor Intermedio y resolvió esta situación en un teorema que es el siguiente.

Teorema de Darboux. Sea I un intervalo, y sea $f : I \to \mathbb{R}$ una función diferenciable. Si a y b son puntos de I con $a \leq b$ y si y esta entre $f'(a)$ y $f'(b)$, entonces existe un número x en $[a, b]$ tal que $f'(x) = y$.

Así, todas las funciones definidas sobre intervalos que son derivadas de otras, sean o no continuas, cumplen la propiedad del valor intermedio.

4. Darboux y sus demostraciones

En el 2004 aparece en *The American Mathematical Monthly, vol. 111, número 8,* un artículo de Lars Olsen en donde presenta una demostración de este teorema, señala que ésta es alterna a las que hasta la fecha han aparecido en los libros de texto, argumentando que esas no son tan faciles de asimilar por los estudiantes. En lo que sigue presentaremos tres demostraciones de este teorema, la segunda aparece en el libro de Análisis Matemático de Apostol y la tercera es la que propone Olsen, la idea es hacer notar que la tercera es una ligera modificación de la segunda. Por

otro lado, la primera demostración nos parece la más sencilla de las tres, también es la que aparece con mayor frecuencia en los libros de texto.

Demostración 1. Supongamos que $f'(a) < y < f'(b)$ y sea $\mu : I \to \mathbb{R}$ definida como $\mu(t) = f(t) - yt$.

Tenemos que $\mu'(t) = f'(t) - y$. Como $f'(a) < y < f'(b)$, entonces $\mu'(a) < 0$ y $\mu'(b) > 0$, lo cual nos dice que ni a ni b son puntos donde μ alcanza un máximo local. Como μ es continua en $[a, b]$ entonces debe alcanzar ese máximo en un punto $x \in (a, b)$. Por lo tanto: $\mu'(x) = f'(x) - y = 0$, esto es: $f'(x) = y$. ■

Demostración 2. Sea y tal que $f'(a) < y < f'(b)$. Y sea $g : I \to \mathbb{R}$ tal que

$$g(t) = \begin{cases} \frac{f(t)-f(a)}{t-a} & \text{si } t \neq a \\ f'(a) & \text{si } t = a \end{cases},$$

esta función es continua en I. Por el teorema del Valor Medio, para cada $t \in (a, b]$ existe $x_t \in (a, b)$ tal que $g(t) = \frac{f(t)-f(a)}{t-a} = f'(x_t)$. Esto nos dice que el intervalo que va de $f'(a)$ a $\frac{f(b)-f(a)}{b-a}$ esta contenido en $f'((a, b))$. ...(1)

Por otro lado, definiendo $h : I \to \mathbb{R}$ como

$$h(t) = \begin{cases} \frac{f(t)-f(b)}{t-b} & \text{si } t \neq b \\ f'(b) & \text{si } t = b \end{cases},$$

tenemos, por las mismas razones que para $g(t)$, que el intervalo cerrado que va de $\frac{f(b)-f(a)}{b-a}$ a $f'(b)$ esta contenido en $f'((a, b))$. ...(2).

Por (1) y (2), y como $f'(a) < y < f'(b)$, entonces existe $x_y \in (a, b)$ tal que $f'(x_y) = y$.■

Demostración 3. Supongamos que y esta entre $f'(a)$ y $f'(b)$. Sean las funciones continuas $f_a, f_b : I \to \mathbb{R}$ definidas como:

$$f_a(t) = \begin{cases} \frac{f(a)-f(t)}{a-t} & \text{si t} \neq a \\ f'(a) & \text{si t} = a \end{cases}$$

y

$$f_b(t) = \begin{cases} \frac{f(t)-f(b)}{t-b} & \text{si } t \neq b \\ f'(b) & \text{si } t = b \end{cases}.$$

Se tiene que $f_a(a) = f'(a)$, $f_a(b) = f_b(a)$ y $f_b(b) = f'(b)$. Por lo cual, y esta entre $f_a(a)$ y $f_a(b)$, o esta entre $f_b(a)$ y $f_b(b)$. Si y esta entre $f_a(a)$ y $f_a(b)$, entonces, por la continuidad de f_a y por el teorema del valor intermedio, existe $s \in (a, b]$ tal que

$$y = f_a(s) = \frac{f(a) - f(s)}{a - s}.$$

Por el teorema del Valor Medio existe $x \in [a, s]$ tal que

$$y = \frac{f(a) - f(s)}{a - s} = f'(x).$$

Si y esta entre $f_b(a)$ y $f_b(b)$, seguimos un razonamiento análogo al anterior para llegar al resultado.■

5. Bibliografía

[1] Apostol, T. M., *Análisis Matemático,* Segunda edición, Reverte, México

[2] Alexandrov A. D., Kolmogorov A. N., Laurientev M. A. y otros,.*La Matemática: su contenido, método y significado, vol. 1,* Alianza Universidad, Madrid, 1979.

[3] Stewart I., *Conceptos de matemática moderna,* Alianza Editorial, Madrid, 1977.

[4] Olsen Lars, *A New Proof of Darboux´s Theorem,* American Mathematical Monthly, vol 11, num. 8, 2004.

Lógica difusa y aplicaciones

Daniel Mocencahua Mora
Facultad de Ciencias de la Electrónica. BUAP.
Avenida San Claudio y 18 Sur. San Manuel.

dmocencahua@ece. buap. mx.

Resumen

Se revisan los conceptos elementales de la lógica difusa y se mencionan líneas de desarrollo tanto en la matemática como en la tecnología.

Palabras clave: Lógica difusa, control difuso.

Modalidad: Conferencia de divulgación en Matemáticas Aplicadas.

1. Introducción.

En 1965, Zadeh publicó el primer artículo sobre una novedosa forma de caracterizar incertidumbres no probabilísticas, a las cuales llamó conjuntos difusos [1]. Aunque la lógica difusa cumple hoy cuarenta años es una ciencia joven y vigorosa que comprende varias disciplinas como el cálculo de reglas if-then difusas, gráficas difusas, interpolación difusa, topología difusa, resonancia difusa, sistemas de interface difusa y modelado difuso. Las aplicaciones, las cuales son multidisciplinarias por naturaleza, incluyen control automático, aparatos electrónicos, procesamiento de señales, predicción de series de tiempo, recuperación de información, administración de bases de datos, visión por computadora, clasificación de datos, toma de decisiones y más.

Esta corriente de pensamiento de la ingeniería afecta también a la Matemática y su historia demuestra ciertos puntos sobresalientes:

1973. Mamdani: Control de máquina de vapor

1977. Ostergaard: Molino de cemento

1980. Tong: Tratamiento de aguas residuales

1983. Hirota, Predrycz: Conjuntos borrosos probabilísticos

1983. Takagi y Sugeno: Derivación de reglas

1984. Sugeno y Murakami. Estacionamiento de un trailer

1985. Kiszka y Gupta: Estabilidad de sistemas borrosos

1985. Togai y Watanabe: Chip borroso

1986. Yamakawa: Hardware de un controlador borroso

1987. Operación Automática del Metro de Sendai (Hitachi)

1988. Dubois y Prade: Razonamiento aproximado

Es importante resaltar que en el metro de Sendai el pasajero no siente tirones de arranque o frenado ya que el metro usa esta lógica para que de acuerdo a la cantidad de pasajeros y las condiciones ambientales estas dos variables sean perfectamente controladas. A partir de los 90´s esta lógica se aplica principalmente en los sistemas expertos y en la investigación de la robóti-

ca, como se menciona adelante. Se debe notar que esta tecnología no se aprovecha en sus inicios en EU debido, según algunos [5], a que el vocablo "difuso" no es atractivo para dispositivos de alta precisión: no es deseable una cámara con "enfoque difuso de imagen". Sin embargo los ingenieros orientales no tienen problema con este tipo de "contradicciones" y llegan al mercado antes volviéndose líderes del mismo.

2. Qué es la Lógica Difusa

La lógica clásica impone a sus enunciados únicamente los valores falso o verdadero y de esta manera han modelado satisfactoriamente una gran parte del razonamiento "natural"de las computadoras: es la lógica boleana. No hay más que dos opciones: falso verdadero, blanco y negro, alto y bajo. Lo cual permite trabajar muy bien a los circuitos electrónicos binarios.

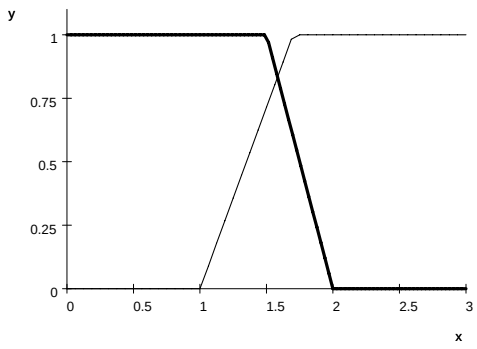
Por medio de la lógica difusa pueden formularse matemáticamente nociones como "un poco caliente" o "muy frío", de forma que sean procesadas por computadoras y cuantificar expresiones humanas vagas, tales como "muy alto" o "luz brillante". No siempre estemos de acuerdo cuando algo deja de estar frío y cuando comienza a estar caliente.

Se permiten grados de certidumbre en los valores de verdad, buscando con esto acercarse al pensamiento humano, por lo cual se dice que esta lógica forma parte del razonamiento aproximado.

Por ejemplo al describir el conjunto A de personas altas la lógica clásica partiría al conjunto de las personas en altas y bajas de acuerdo a un umbral convenido, por ejemplo 1.65 m. La función que describe a este conjunto es la función característica de A:

$$\chi_A(x) = \begin{cases} 1 \text{ si } x \in A \\ 0 \text{ si } x \notin A \end{cases}$$

Sin embargo en la lógica difusa se describe al conjunto en base a grados de pertenencia al conjunto. De hecho se generan dos conjuntos: el de las personas bajas y el de las personas altas. Se define bajo a lo que sea parecido al intervalo de 0 a 1.5m, tomando en cuenta que otros valores se parecen a los del intervalo, por lo que su valor es menor a 1 pero no cero. Por otra parte los valores a partir de 1.5 m se acercan a lo que nos parece alto por lo que tienen ciertos valores menores que 1, pero conforme se acercan al valor de 1.7 son valores más altos.



Conjuntos difusos **bajo** y **alto** (punteado)

Estos valores se conocen como valores de pertenencia al conjunto y la gráfica de la función resultante se conoce como función de membresía (mf). Se formaliza este concepto con la siguiente definición.

Definición 1 *Sea X una colección de objetos, denotados por x, un conjunto difuso A en X es definido por un conjunto de pares ordenados*

$$A = \{(x, \mu_A(x)) : x \in X\},$$

donde $\mu_A(x)$ es la función de membresía de A y $0 \leq \mu(x) \leq 1$.

Note que la función característica de un conjunto puede ser una función de membresía, con lo que la lógica clásica es generalizada por la difusa, lo cual fue uno de los objetivos de Zadeh para definirla así.

Otra manera de denotar a los conjuntos difusos es:

$$A = \begin{cases} \sum_{x_i \in X} \mu_A(x_i)/x_i, & \text{si } X \text{ es discreto,} \\ \int_X \mu_A(x)/x, & \text{si } X \text{ es continuo.} \end{cases}$$

La sumatoria y el signo de la integral denotan la unión de los pares $(x, \mu_A(x))$, no indican suma o integración, y el símbolo "/" tan solo es notación y de ninguna manera indica división.

Es importante resaltar el hecho de que los elementos que se estudien (el conjunto X es conocido como universo del discurso) pueden estar al mismo tiempo en distintos conjuntos difusos con distintos grados de membresía.

2.1. Operaciones difusas

Ahora se definen las operaciones difusas análogas a las de los conjuntos clásicos.

Definición 2 *Sean A, B conjuntos difusos en X. Decimos que A* ***es subconjunto de*** *(o está contenido en) B, lo cual se denota $A \subseteq B$, si $\mu_A(x) \leq \mu_B(x)$, $\forall x \in X$.*

Definición 3 *Sean A, B conjuntos difusos en X. $A = B$ si $A \subseteq B$ y $B \subseteq A$. Esto es equivalente a decir que $A = B$ si y sólo si $\mu_A(x) = \mu_B(x)$.*

Definición 4 *La* ***intersección difusa*** *de A y B está definida por medio de la función de membresía siguiente:*

$$\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\}, \ \forall x \in X.$$

Definición 5 *La* ***unión difusa*** *de A y B está definida por medio de la función de membresía siguiente:*

$$\mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\}, \ \forall x \in X.$$

Ejemplo 1 *Sean A y B conjuntos difusos de $X = \{-2, -1, 0, 1, 2, 3, 4\}$.*

$$\begin{aligned} A &= 0{,}6/(-2) + 0{,}3/(-1) + 0{,}6/0 + 1{,}0/1 \\ &\quad + 0{,}6/2 + 0{,}3/3 + 0{,}4/4, \\ B &= 0{,}1/(-2) + 0{,}3/(-1) + 0{,}9/0 \\ &\quad + 1{,}0/1 + 1{,}0/2 + 0{,}3/3 + 0{,}2/4. \end{aligned}$$

Luego, $A \cup B$ tiene la forma:

$$\begin{aligned} A \cup B &= 0{,}6/(-2) + 0{,}3/(-1) + 0{,}9/0 \\ &\quad + 1{,}0/1 + 1{,}0/2 + 0{,}3/3 + 0{,}4/4. \end{aligned}$$

Note que la función de membresía de A no es una función de distribución de probabilidad ya que la suma de los valores es mayor que 1.

Definición 6 *El* ***complemento*** *de un conjunto difuso A es definido como*

$$\mu_{\overline{A}}(x) = 1 - \mu_A(x).$$

El siguiente par de proposiciones nos indican qué tanto generaliza la lógica difusa a la clásica [6].

Proposición 1 *Las leyes de Idempotencia, Distributividad, Conmutatividad, Asociatividad, Absorción, Neutro, Identidad, Ley de la doble negación, y las leyes de De Morgan, son válidas en la lógica difusa.*

Proposición 2 *Las leyes del tercero excluido y la ley de no contradicción no son satisfechas en la lógica difusa.*

La intersección y unión definidas son casos particulares de las llamadas **T-normas y T-conormas,** respectivamente, y son las únicas que cumplen la ley distributiva [4].

2.2. Funciones de membresía usuales

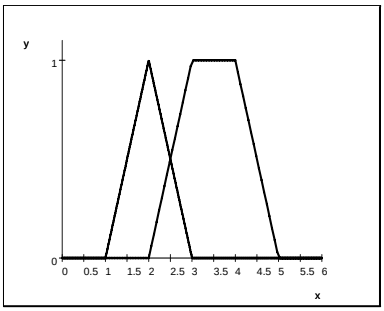
Se definen ahora diversas clases de funciones parametrizadas que se usan comúnmente para definir funciones de membresía.

Definición 7 *Una función de membresía* ***triangular*** *se define por tres parámetros que cuales determinan las coordenadas de tres vértices:*

$$trimf(a,b,c) = max\left(\text{mín}\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right)$$

Definición 8 *Una función de membresía* ***trapezoidal*** *está definida por cuatro parámetros como sigue:*

$$trapmf(a,b,c,d) = max\left(\text{mín}\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right)$$



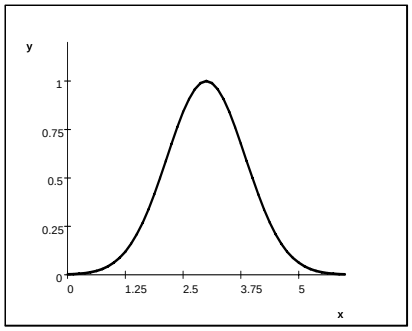
Funciones $trianmf\ (1,2,3)$ y $trapmf\ (2,3,4,5)$

Obviamente, una función de membresía triangular es un caso especial de una trapezoidal. Estos dos tipos son los más usados en las implementaciones de tiempo real debido a la simplicidad de sus fórmulas y la eficiencia computacional al programarlas. Sin embargo debido que estas funciones son compuestas por segmentos de líneas, el switcheo en los puntos especificados por los parámetros no es suave. Por lo que se hace necesario tener funciones de membresía suaves y no lineales para algunas aplicaciones

Definición 9 *Una función de membresía* ***Gaussiana*** *está definida por dos parámetros como sigue:*

$$gausmf(\sigma, c) = e^{-\left(\frac{x-c}{\sigma}\right)^2}$$

donde c representa el centro de la función y σ determina el ancho de la misma.

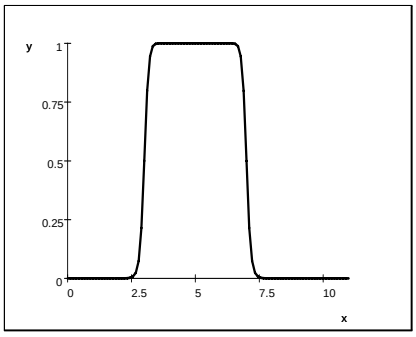


Función de membresía $gausmf(3,1,2)$

Definición 10 *Una función de membresía de* ***campana*** *se define por:*

$$gbellmf(a,b,c) = \frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}}$$

donde el parámetro b es usualmente positivo.



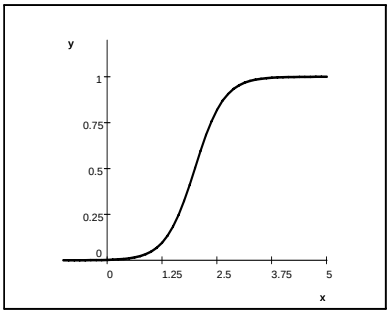
Función de membresía $gbellmf(2,4,5)$

Debido a su suavidad y a su notación concisa, las funciones de membresía Gaussianas y las de campana han venido incrementando la popularidad de los métodos para especificar conjuntos difusos.

Definición 11 *Una función de membresía* ***sigmoidal*** *es definida por*

$$sigmf(a,c) = \frac{1}{1+\exp\left[-a(x-c)\right]}$$

Donde a controla la pendiente en los puntos de cruce $x = c$.



Función $sigmf(2,3)$

Dependiendo del signo del parámetro a una función de membresía sigmoidal abre hacia la izquierda o derecha y por lo tanto es apropiada para representar conceptos como "muy largo", o "muy negativo".

Se hace notar que la lista de funciones de membresía introducida en esta sección no es exhaustiva, otras funciones especializadas pueden ser creadas para aplicaciones específicas si es necesario. En particular, cualquier función de distribución probabilística continua puede ser usada como función de membresía.

2.3. Implicación difusa

La implicación difusa "si x está en A entonces y está en B", donde A y B son conjuntos difusos, se formaliza por medio de **relaciones difusas**, que son generalizaciones de las relaciones cartesianas. Las implicaciones difusas usuales son: Mamdani, Larsen, producto acotado, producto drástico, max-min (Zadeh), implicación booleana, Lukasiewicz, Kleene-Dienes y Yager [10].

La manera en que la lógica es útil en la tecnología es a través de reglas difusas, estas son de la forma: *IF* **Condiciones** *THEN* **Acciones**. Donde las **Condiciones** son aquello que el sistema observa del mundo a través de sus sensores y **Acciones** es la forma en que el sistema responde a estas condiciones a través de sus actuadores.

La diferencia con la lógica clásica es que aquí se permiten expresiones como: IF la temperatura es alta THEN la válvula del gas se cierra mucho.

Una forma usual en la práctica es tener varias condiciones y varias acciones: IF x1 es A AND x2 es B AND x3 es C THEN u1 es D AND u2 es E, donde A, B, C, D, E son conjuntos difusos.

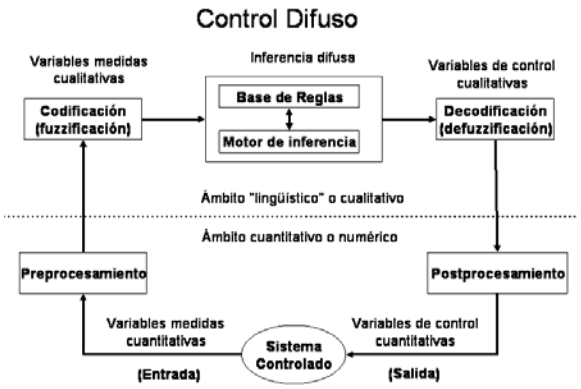
Se busca entonces generar las reglas correspondientes al sistema y para ello se puede hacer uso de la experiencia del usuario del mismo, ya que se le puede pedir que indique que haría en cada situación y eso se escribe en la forma anterior para poder implementarlo en un sistema experto o en un control difuso.

Es útil observar el comportamiento de estos sistemas, por lo cual diversos paquetes de software se han diseñado, destacando fuzzytech y el fuzzy logic toolbox de Matlab [3].

3. Control difuso

Un controlador difuso (FLC) es una ley de control descrita por un sistema de reglas difusas, con predicados no precisos, y con un mecanismo de inferencia difusa.

Este puede tener la siguiente forma:



Se nota que la señal de entrada va a una etapa de codificación (*fuzzificación*). Esto es necesario porque esta señal estima la diferencia entre la señal de referencia y los sensores del sistema controlado, por lo que es una señal no difusa. Así que esta señal se cambia a difusa por algún método definido por el diseñador.

En el mecanismo de inferencia se toma la decisión de la acción de acuerdo a lo que dicten las reglas y su salida es difusa por lo que se necesita decodificar (*defuzzificar*).

El mecanismo de inferencia más común es el de Mamdani, un ingeniero italiano que lo diseñó para controlar una máquina de vapor y que es la primera aplicación documentada del control difuso. En este tipo de mecanismo de inferencia se toman la intersección de las entradas (mínimo) y se calcula la acción con la unión (máximo) de las salidas, y se defuzzifica con el centro de masas de el conjunto difuso resultante [2].

Otro mecanismo muy usado es del de Takagi-Sugeno (TSK) que tiene la forma "si x está en A y y está en B entonces $z = f(x, y)$". Donde A y B son difusos y f es una función, comúnmente un polinomio en la variables del antecedente, aunque se puede usar una función que sea más adecuada al sistema.

Ambos modelos de control contrastan en el siguiente sentido: si bien Mamdani es fácil de generar por su tratamiento lingüístico del sistema a controlar, no es fácil realizar un análisis de estabilidad formal. Los modelos TKS, sin embargo, tienen pruebas de estabilidad, pero definir la mejor función para el consecuente no siempre es tan evidente.

Una aplicación interesante de los controladores difusos es que en algunos casos particulares se puede asegurar que son aproximadores universales, o sea que se cumple el siguiente teorema.

Teorema 1 *Dada una función continua de valores reales g definida sobre el conjunto compacto U y dado $\varepsilon > 0$ cualquiera, existe un sistema de control difuso que tiene como salida la función f tal que*

$$\sup_{x \in U} \|g(x) - f(x)\| \leq \varepsilon.$$

Wang demostró este teorema usando el teorema de Stone-Weierstrass y un tipo muy particular de sistemas de control difuso ([11] citado en [10]) y Castro lo hizo con sistemas de tipo Mamdani [12].

4. ¿Para qué sirve la lógica difusa?

Anteriormente las aplicaciones sólo se daban en control de procesos o de máquinas en la industria pero actualmente han llegado al alcance del hombre común. El baño difuso, por ejemplo, mantiene la temperatura del agua para que esté del agrado del usuario. Actualmente marcas como Matsuhita, Daewoo y otras, fabrican electrodomésticos con tecnología difusa. Destacan las lavadoras que con sus sensores detectan

el color, el tipo y la cantidad de ropa y seleccionan la cantidad apropiada de agua, la temperatura, la cantidad de detergente y tipo de lavado de entre 600 combinaciones (reglas) existentes en su memoria. Mitsubishi tiene aparatos de aire acondicionado difusos que controlan la temperatura de acuerdo a índices que el usuario define. Casi todas las marcas de electrónica deben de haber incorporado en sus cámaras de video controles difusos para el enfoque y estabilización de imagen. Las primeras en hacerlo fueron Fisher y Sanyo. En las PDAs existen sistemas expertos basados en lógica difusa que permiten el reconocimiento de la letra manual y la convierten a datos para la máquina. Ciertos autos tienen frenos antibloqueo con esta lógica y la Nissan incorporó la transmisión automática difusa. Se conocen elevadores difusos, ajuste de brillo y contraste en monitores de tv, y estacionamiento automático de autos, además de que se buscan autos que se conduzcan automáticamente por medio de visión artificial.

La última década ha visto aplicaciones de inteligencia artificial por medio de la lógica difusa en sistemas expertos fuera del diseño tecnológico. Existen aplicaciones en control de calidad, supervisión y detección de fallos; previsión en terremotos (ingeniería civil); ecología: modelos de población y análisis toxicológico de químicos; planeación estratégica, toma de decisiones, actuaría, planeación de portafolios; ergonomía; procesamiento de lenguaje y aplicaciones sicológicas [9].

Una de las disciplinas no tecnológicas que ha explotado esta tecnología ha sido la medicina: extracción de características específicas en imágenes, clasificación, monitoreo de anestesia, procesamiento de señales electrofisiológicas, reconocimiento de patrones aplicados al diagnóstico y toma de decisiones de acuerdo a la evolución del paciente [13]. La aplicación biomédica que más éxito parece tener es la de controlar prótesis por medio de señales mioeléctricas como en [17].

5. ¿Qué queda por hacer?

Reconocimiento de patrones en audio y video, identificación de ADN, portafolios de inversión, diagnóstico médico, sistemas de concesión (o negación) de créditos, son muchas de las aplicaciones que ya existen pero que pueden ser mejoradas.

La lógica difusa ha demostrado ser una opción en el control de sistemas complejos, sin embargo la *fuzzificación* y la elección de los distintos conjuntos difusos es más un arte que una ciencia. Una forma de resolver esto es aplicar redes neuronales o algoritmos genéticos que optimizan estas elecciones, esta es una vía de trabajo en desarrollo, tanto matemático como tecnológico.

La robótica es un campo de investigación que día a día aprovecha los frutos del cómputo suave, que tiene a la lógica difusa como una de sus herramientas: control de robots móviles [8], coordinación de manipuladores en ambientes coperativos, control adaptivo, arbitraje de comportamientos [16], son técnicas con las cuales los robots más sofisticados interactúan con el mundo [7].

El desarrollo formal de la lógica difusa es una vía de investigación en vigor: programación difusa, topología difusa, teoría de la medida difusa, caos, ecuaciones diferenciales, teoremas de punto fijo, son temas actuales en la literatura (ver, por ejemplo, los últimos números de [14]). Inclusive el problema de los aproximadores universales sigue siendo investigado [15].

6. Conclusión

Se han presentado los conceptos básicos de la lógica difusa, un bosquejo histórico y las distintas vías de desarrollo tanto en la matemática aplicada, la pura o la tecnología, conmemorando los cuarenta años de existencia de esta disciplina y como invitación a su estudio.

Referencias

[1] Zadeh, L.; "Fuzzy sets". Information and Control 8, 338-353 (1965).

[2] Jerry M. Mendel. Fuzzy Logic Systems for Engineering: A Tutorial. IEEE Proc., vol. 83, no. 2, pp. 345-377, March (1995)

[3] The Mathworks Inc. Fuzzy Logic Toolbox. For use with Matlab. (2000)

[4] George J. Klir, Bo Yuan. Fuzzy Sets and Fuzzy Logic. Theory and Applications. Prentice Hall. (1995)

[5] Timothy J. Ross. Fuzzy Logic with Engineering Applications, Mac Graw Hill International Editions, New York. (1997).

[6] Ching-Teng Lin, C. S. George Lee, Neural Fuzzy Systems: A neuro-fuzzy synergism to intelligent systems. Prentice Hall (1996)

[7] Clarence W. de Silva. Inteligent Control. Fuzzy Logic Applications. CRC Press. (1995)

[8] Margarita Olaya C., Daniel Mocencahua M. et al. Control difuso para un robot móvil. Memorias del Congreso Interuniversitario de Electrónica Computación y Eléctrica 2005, Puebla, 9 marzo de 2005.

[9] Hans-Jürgen Zimmermann (Editor). Practical applications of fuzzy tecnologies. Kluwer Academic Publishers. USA. (1999)

[10] Robert Fúller. Introduction to Neuro-Fuzzy Systems. Advances in Soft Computing. Physica Verlag. Heidelberg. (2000)

[11] L. X. Wang. Fuzzy zyztems are universal approximators. Proc. IEEE 1992. Int. Conference Fuzzy Systems. San Diego. 1163-1170. (1992)

[12] J. L. Castro. Fuzzy logic controllers are universal approximators. IEEE transactions on Syst. Man Cybernet. 25. 629-635. (1995)

[13] Senén Barro, Roque Marín (editors). Fuzzy logic in medicine. Physica-Verlag. Germany. (2002)

[14] D. Dubois. H. Prade (editors). Fuzzy Sets and Systems. Elsevier.

[15] Ronald R. Yager. Vadik Kreinovich. Universal approximation theorem for uninorm-based fuzzy systems modeling. Fuzzy sets and systems. Vol. 140. Issue 2. Dec. 331-339. (2003)

[16] M. Cuesta Hernández, D. Mocencahua Mora. Arbitraje de comportamientos para robots móviles con lógica difusa. Memorias del Congreso Interuniversitario de Electrónica Computación y Eléctrica 2005, Puebla, 9 marzo de 2005.

[17] Oscar A. Morales Pizarro, J. E. Flores-Mena, D. Mocencahua M. Control difuso de una prótesis de tres dedos. Memorias del Congreso Interuniversitario de Electrónica Computación y Eléctrica 2005, Puebla, 9 marzo de 2005.

UN CASO ESPECIAL DE MATRIZ DE TRANSICION: MODELO MATRICIAL DE LESLIE

FRANCISCO SOLANO TAJONAR SANABRIA
(ftajonar@fcfm.buap.mx)
LUCILA MUÑIZ MERINO
(lucymerino74@hotmail.com)

Resumen

El estudio de la dinámica de poblaciones usualmente se refiere al análisis de las fluctuaciones de la abundancia. Sin embargo, este planteamiento lleva implícito una serie de conocimientos sobre las propiedades particulares de la población y sobre las variables que actúan en la expresión cuantificable de esas propiedades. El cálculo de estos parámetros poblacionales es básico para la realización de estudios de manejo y la implementación de medidas de conservación de estudios ecológicos posteriores.

En este trabajo se presenta el modelo matricial de Leslie que resulta ser un caso especial de una matriz de transición y el cual se ha utilizado para modelar problemas de crecimiento poblacional clasificados por a edad y/o tamaño.

Palabras Claves: Población**,** Natalidad, Densidad, Modelo matricial de Leslie.

I.- Introducción

La dinámica de las poblaciones es uno de los temas de mayor importancia para entender el desarrollo temporal y espacial de grupos de organismos de la misma especie que se desarrollan en distintos ambientes. En el nivel práctico, el interés se centra en el manejo de plagas agrícolas, para comprender la epidemiología de numerosas enfermedades, para

estimar densidades pesqueras y cuotas de extracción, para manejar poblaciones silvestres y para entender los aspectos demográficos de la población humana.

Una **población** desde el punto de vista ecológico es un grupo de organismos de la misma especie, que habitan un lugar determinado, en el cual utilizan recursos y se reproducen. Este grupo de organismos esta caracterizado por una serie de propiedades que son propias y únicas de ese nivel de organización. Así, si un organismo se reproduce decimos que tiene una cría, pero algún organismo vecino puede no reproducirse. El promedio de nacimientos que se den en el grupo constituirá la **natalidad** del grupo. La natalidad no es una propiedad de los individuos, solo emerge cuando se tiene una población.

Las propiedades emergentes de una población son: densidad, tasas de crecimiento, tasas de mortalidad y natalidad, distribución espacial, distribución por sexos, por edades, tipos de crecimiento, frecuencias génicas, variabilidad genética y otras. Así, para realizar el análisis de una población se debe investigar en sus propiedades emergentes, ya que estas proporcionaran información sobre la fluctuación de la abundancia y de la distribución del hábitat.

La **densidad** es la representación de la abundancia de la población y se expresa como el número de individuos o biomasa en función del espacio o volumen ocupado. La densidad puede ser absoluta cuando se considera todo el hábitat sin importar si se ocupa o no por la especie investigada. La densidad será ecológica cuando se toma en cuenta el sitio ocupado en forma efectiva. La abundancia determina algunos efectos a nivel de la población, por ejemplo, si esta decrece se tiene que disminuye la probabilidad de dejar descendencia debido a que la probabilidad de encuentros entre sexos es baja. Cuando se llega a este nivel se habla de **un tamaño crítico poblacional y de un efecto de grupo.**

La densidad es producto del balance entre **natalidad** y **mortalidad poblacional,** y también del balance entre **inmigración** y **emigración**. Estos dos últimos factores suelen adscribirse por comodidad a la natalidad (b) y mortalidad (d) respectivamente. Estos parámetros poblacionales indican un cambio en el tamaño de la población en relación de los que nacen como los que mueren. Estas tasas pueden expresarse como tasas brutas, que es tomar la diferencia producida en el total de la población, por la adición o sustracción de especimenes en un lapso de tiempo o como tasas específicas cuando se considera la edad o sexo de los organismos. Cuando se lleva el intervalo de tiempo al límite más pequeño se habla de **tasas instantáneas de mortalidad y natalidad,** que se utilizan para determinar el crecimiento poblacional en cualquier instante. En este trabajo se analiza el **modelo matricial de Leslie** que se utiliza en ecología para determinar el crecimiento de una población y los porcentajes de distribución a lo largo del tiempo**.** Otro problema importante en ecología son las consecuencias ambientales que se pueden sufrir con respecto a los cambios climáticos, y estos puedan repercutir en la producción agrícola y en la vegetación natural, un modelo que analiza este tipo de problemas es el llamado modelo **de estado y transición**.

II.- Modelos de Crecimiento

Iniciamos con el modelo más simple que permite determinar el crecimiento de una población, llamado modelo de crecimiento. Suponga que un organismo del cual se posee inicialmente 2 individuos, tiene una capacidad de reproducción constante de 2 especimenes. Así tenemos que en la primera reproducción se tendrá 2*2, en la siguiente 4*2 y así sucesivamente. En general, si la población inicial es N_0, al final del primer periodo reproductivo t , se tiene que

$$N_1 = N_0 \bullet \lambda$$

$$N_2 = N_1 \bullet \lambda = N_0 \bullet \lambda^2$$

y de esta forma, se obtiene que

$$N_t = N_{t-1} \bullet \lambda = N_0 \bullet \lambda^t \ . \qquad (1)$$

En este caso, λ denota un crecimiento finito, es decir, un crecimiento por pulsos discreto. Por ello se denomina **tasa discreta de crecimiento poblacional** o **tasa finita de crecimiento,** este parámetro informa de cómo crece o decrece una población entre periodos de tiempo. Esto se puede usar para definir tasas de extracción de especies silvestres. Tomando el crecimiento en función de d y b, se tiene que: $b - d = \lambda = \Delta N / N \bullet \Delta t$, si el crecimiento es por pulsos.

En ausencia de factores limitantes, esto es, con alimento suficiente y adecuado, con espacio suficiente y adecuado, una población crecerá exponencialmente, un modelo con estas características se denomina de **crecimiento exponencial** y no tiene un límite pre-establecido.

En muchas situaciones el crecimiento definido por periodos de tiempo no permite realizar comparaciones entre poblaciones que tienen diferentes periodos reproductivos, ni tampoco estimar con precisión las variaciones del desarrollo poblacional en cada instante, para resolver esto se utiliza la **tasa instantánea de crecimiento** o **tasa de crecimiento específico,** que es el parámetro de mayor importancia relativa en la dinámica de cualquier población. En este caso tenemos que

$$dN / Ndt = b - d = r \quad \text{o} \quad dN / dt = rN \qquad (2)$$

donde r se considera constante para cada especie y se le denomina **tasa intrínseca de crecimiento**. Para obtener una expresión del crecimiento poblacional en función del tiempo se obtiene integrando (2), dando como resultado

$$N_t = N_0 \bullet e^{rt}. \qquad (3)$$

Esto indica que se puede conocer el crecimiento o tamaño de una población en cualquier instante, si se conoce la población inicial N_0 y el valor de r. De (2) también se puede decir que si:

- $b > d$ entonces $r > 0$ y la población crece,
- $b < d$ entonces $r < 0$ y la población decrece,
- $b = d$ entonces $r = 0$ y la población se mantiene estable.

III.- Matrices de Leslie

El modelo matricial de Leslie es una herramienta usada para determinar el crecimiento de una población así como la distribución por edad a lo largo del tiempo. Esta descripción fue hecha por P.H. Leslie en 1945. Se ha usado para estudiar la dinámica de poblaciones de una amplia variedad de organismos, como truchas, conejos, escarabajos, piojos, orcas, humanos y también se ha usado para predecir distribuciones de clases de edad estable en pinos.

El modelo matricial de **Leslie** es modelo de transición donde para su uso **s**e considera un conjunto de supuestos, tales como: que tipo de objetos o sujetos son considerados, edad máxima de los objetos o sujetos, como se agrupan (edad o tamaño), la probabilidad de sobrevivir, la tasa de supervivencia, la fecundidad y la distribución inicial. Con estos supuestos el modelo de Leslie está definido por la ecuación

$$X_k = L^k X_0, \tag{4}$$

donde X_0 es el vector inicial de distribución de la población, y X_k el vector de distribución de la población en el instante k. Si la matriz de Leslie L es diagonalizable, entonces $L = VDV^{-1}$, donde D es una matriz diagonal formada por los eigenvalores de la matriz L. Las columnas de V son los eigenvectores correspondientes. En este caso, el modelo de Leslie se puede escribir como

$$X_k = c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \ldots + c_n\lambda_n^k v_n, \tag{5}$$

donde λ_i, v_i son el eigenvalor y eigenvector asociados. Si λ_1 es el eigenvalor estrictamente dominante de L, entonces para valores grandes de k se tiene que

$$X_k \approx c_1\lambda_1^k v_1, \tag{6}$$

y la proporción de objetos o sujetos en cada clase de edad tiende a una constante. Estas proporciones límites se pueden determinar a partir de las componentes de v_1. Por último, el eigenvalor dominante λ_1 determina la tasa de cambio de un año para otro. Como

$$X_k \approx \lambda_1 X_{k-1}, \tag{7}$$

para valores grandes de k, el vector de población en el instante k es un múltiplo del vector de población en el instante $k-1$. Si $\lambda_1 > 1$ entonces la población tendrá un crecimiento indefinido. Si $\lambda_1 < 1$ entonces la población se extinguirá.

Observaciones:

1. La ecuación (4) nos indica que si conocemos el vector de distribución inicial X_0 y la matriz de Leslie L podemos determinar el vector de distribución de la población en cualquier instante o tiempo posterior, con la multiplicación de una potencia apropiada de la matriz de Leslie L por el vector de distribución inicial X_0. En general, la matriz de Leslie L es un caso especial de una matriz de transición y usualmente no tiene un vector de probabilidades estacionarias, sin embargo, una proporción estable límite de clases edad/tamaño es alcanzada y esta dada por (5), al vector v_1 se le llama vector de probabilidades pseudo-estacionarias.
2. Se puede notar que la ecuación (4) tiene una expresión semejante a una ecuación en diferencias.

Ejemplo: Aplicamos el modelo al estudio de una especie de salmón. El modelo de Leslie para este caso parte de las siguientes hipótesis:

- Solo se consideran las hembras en la población de salmones.
- La máxima edad alcanzada por un individuo son tres años.
- Los salmones se agrupan en tres intervalos de un año cada uno.
- La probabilidad de sobrevivir un salmón de un año para otro depende de su edad.
- La tasa de supervivencia P_i en cada grupo es conocida.
- La fecundidad (tasa de reproducción) F_i en cada grupo es conocida.
- La distribución inicial es conocida.

Con esto, es posible construir un modelo determinista con matrices. Como la edad máxima de un salmón es 3 años, la población entera puede dividirse en tres clases de un año. La

clase 1 contiene los salmones en su primer año de vida, la clase 2 a los salmones entre 1 y 2 años, y la clase 3 a los salmones de más de 2 años.

Suponga que se conoce el número de hembras en cada una de las tres clases en el momento $t = t_0$. Sea $X_1^{(0)}$ el número de hembras en la primera clase, $X_2^{(0)}$ el número de hembras en la segunda clase y $X_3^{(0)}$ el número de hembras en la tercera clase, con estos números se puede formar el vector

$$X_0 = \begin{pmatrix} X_1^{(0)} \\ X_2^{(0)} \\ X_3^{(0)} \end{pmatrix}, \qquad (8)$$

que denota el vector inicial de distribución por edad, o vector de distribución de edad en el instante $t = t_0$.

Es claro que conforme pasa el tiempo, el número de hembras en cada una de las tres clases cambia por la acción de tres procesos biológicos: nacimiento, muerte y envejecimiento. Mediante la descripción cuantitativa de estos procesos se puede estimar el vector de distribución por edad en tiempos futuros.

Suponga que la población es observada en instantes de tiempo discretos de un año, denotados por t_0; t_1;…. Los procesos de nacimiento y muerte entre dos observaciones sucesivas se pueden describir a través de los parámetros tasa media de reproducción y tasa de supervivencia.

Sea F_i el número de medio de hembras nacidas de una hembra en la i-ésima clase, esto es, es la tasa media de reproducción de la i-ésima clase $i = 1,2,3$. Sea P_1 la fracción de hembras en la primera clase que sobreviven el año para pasar a la segunda clase, P_2 la fracción de hembras en la segunda clase que sobreviven el año para pasar a la tercera clase.

No hay P_3, ya que al cumplir 3 años, el salmón muere tras desovar, y ninguno sobrevive para llegar a una cuarta clase. En general tenemos que

- F_i es la tasa media de reproducción de una hembra en la clase i .
- P_i es la tasa de supervivencia de hembras en la clase i .

Por definición $F_i \geq 0$, ya que la descendencia no puede ser negativa. Para este ejemplo tenemos que $F_1 = 0$, $F_2 = 0$, porque el salmón solamente produce huevos en su último año de vida. Por esto, solo F_3 tiene un valor positivo. También se tiene que, $0 < P_i \leq 1$ para $i = 1,2,$ al suponer que algunos salmones deben sobrevivir para llegar a la siguiente clase, excepto para la ultima clase, donde el salmón muere.

Defina el vector de distribución por edad en el instante t_k por

$$X_k = \begin{pmatrix} X_1^{(k)} \\ X_2^{(k)} \\ X_3^{(k)} \end{pmatrix}, \tag{9}$$

donde $X_i^{(k)}$ es el número de salmones hembras en la clase i en el instante t_k .

En el instante t_k, el número de salmones en la primera clase, $X_1^{(k)}$, es igual a los salmones nacidos entre los instantes t_{k-1} y t_k. El número de descendientes producidos por cada clase se puede calcular multiplicando la tasa media de reproducción de la clase por el número de hembras en la clase de edad. La suma de todos estos valores proporciona el total de descendientes, esto es

$$X_1^{(k)} = F_1 X_1^{(k-1)} + F_2 X_2^{(k-1)} + F_3 X_3^{(k-1)}, \tag{10}$$

que indica que el número de hembras en la clase 1 es igual al número de hijas nacidas de hembras en la clase 1 entre los instantes t_{k-1} y t_k, más el número de hijas nacidas de hembras en la clase 2 entre t_{k-1} y t_k, más el número de hijas nacidas de hembras en la clase 3 entre t_{k-1} y t_k. Como los salmones solamente producen huevos en su último año de vida, tenemos

$$X_1^{(k)} = 0 \bullet X_1^{(k-1)} + 0 \bullet X_2^{(k-1)} + F_3 X_3^{(k-1)}. \tag{11}$$

El número de hembras en la segunda clase de edad en el instante t_k se obtiene a partir de las hembras de la primera clase en el instante t_{k-1} que sobreviven al instante t_k, es decir;

$$X_2^{(k)} = P_1 X_1^{(k-1)}. \tag{12}$$

Análogamente el número de hembras en la tercera clase de edad en el instante t_k se obtiene a partir de las hembras de la segunda clase en el instante t_{k-1} que sobreviven al instante t_k, es decir;

$$X_3^{(k)} = P_2 X_2^{(k-1)}. \tag{13}$$

De (12), (14) y (15), tenemos el siguiente sistema de ecuaciones

$$\begin{aligned} X_1^{(k)} &= F_1 X_1^{(k-1)} + F_2 X_2^{(k-1)} + F_3 X_3^{(k-1)}, \\ X_2^{(k)} &= P_1 X_1^{(k-1)}, \\ X_3^{(k)} &= P_2 X_2^{(k-1)}, \end{aligned} \tag{14}$$

que en notación matricial (14), se puede escribir de la siguiente forma

$$\begin{pmatrix} X_1^{(k)} \\ X_2^{(k)} \\ X_3^{(k)} \end{pmatrix} = \begin{pmatrix} F_1 & F_2 & F_3 \\ P_1 & 0 & 0 \\ 0 & P_2 & 0 \end{pmatrix} \begin{pmatrix} X_1^{(k-1)} \\ X_2^{(k-1)} \\ X_3^{(k-1)} \end{pmatrix}, \tag{16}$$

o equivalentemente

$$X_k = LX_{k-1}, \tag{17}$$

a (18) se le llama el modelo matricial de Leslie y L la matriz de Leslie. Para el ejemplo tenemos que $F_1 = F_2 = 0$, así que la matriz de Leslie para la población de salmones es

$$L = \begin{pmatrix} 0 & 0 & F_3 \\ P_1 & 0 & 0 \\ 0 & P_2 & 0 \end{pmatrix}. \tag{18}$$

De (17) se puede calcular el vector de distribución por edad en cualquier instante.

En general, la matriz de Leslie tiene la siguiente representación

$$\begin{bmatrix} X_1^{(k)} \\ X_2^{(K)} \\ X_3^{(K)} \\ X_4^{(K)} \\ \bullet \\ \bullet \\ \bullet \\ X_n^{(K)} \end{bmatrix} = \begin{bmatrix} F_{1,1} & F_{1,2} & F_{1,3} & \bullet & \bullet & \bullet & F_{1,n-1} & F_{1,n} \\ p_{2,1} & 0 & 0 & \bullet & \bullet & \bullet & 0 & 0 \\ 0 & p_{3,2} & 0 & \bullet & \bullet & \bullet & 0 & 0 \\ 0 & 0 & p_{4,3} & & & & & 0 \\ \bullet & & & \bullet & & & & \bullet \\ \bullet & & & & \bullet & & & \bullet \\ \bullet & & & & & \bullet & & \bullet \\ 0 & 0 & 0 & \bullet & \bullet & \bullet & p_{n,n-1} & 0 \end{bmatrix}^k \times \begin{bmatrix} X_1^{(0)} \\ X_2^{(0)} \\ X_3^{(0)} \\ X_4^{(0)} \\ \bullet \\ \bullet \\ \bullet \\ X_n^{(0)} \end{bmatrix} \tag{19}$$

IV.- EJEMPLO NUMERICO

Población con 3 clases de edad

Descendencia femenina x hembra, clase 2=4

Descendencia femenina x hembra, clase 3=3

Hembras sobrevivientes de clase 1=50%

Hembras sobrevivientes de clase 2=25%

La matriz de Leslie y el vector inicial de esta población son:

$$L = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} .. X^0 = \begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix}$$

Distribución de población para un periodo de 10 años

X =

Columns 1 through 4

10	**70**	**27.5**	**143.75**
10	**5**	**35**	**13.75**
10	**2.5**	**1.25**	**8.75**

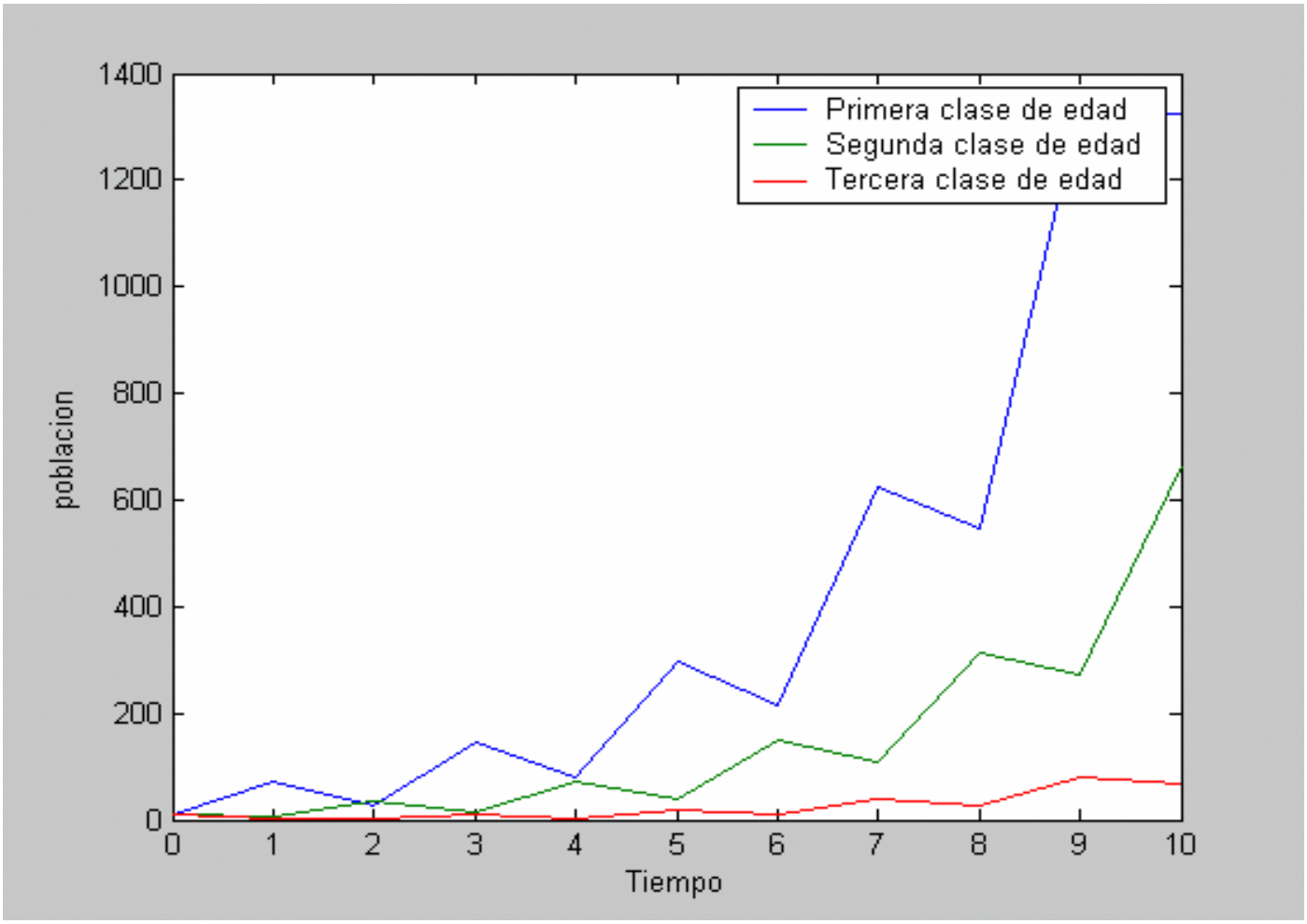
Columns 5 through 8

81.25	**297.81**	**216.41**	**626.09**
71.875	**40.625**	**148.91**	**108.2**
3.4375	**17.969**	**10.156**	**37.227**

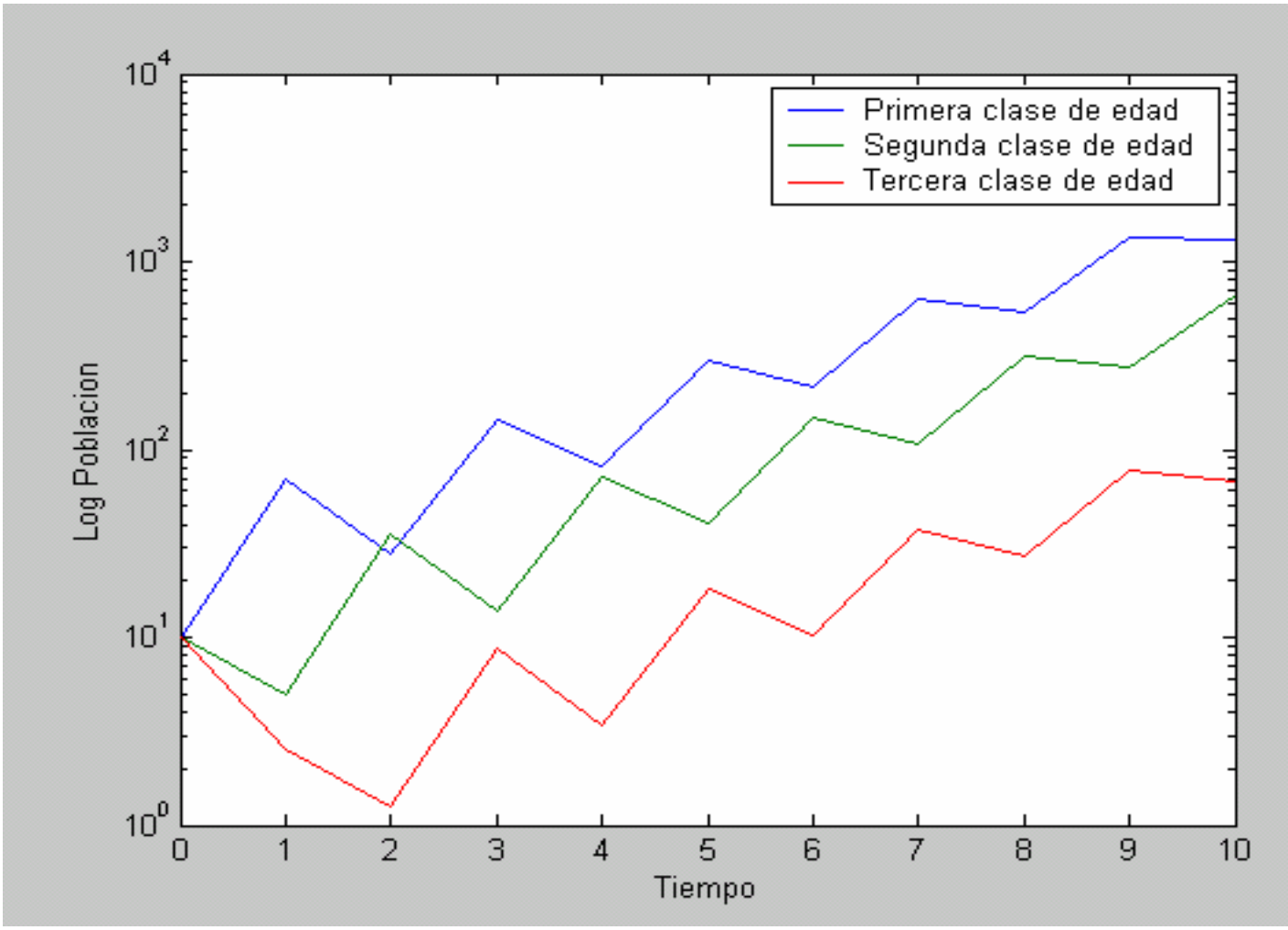
Columns 9 through 11

544.49	**1333.3**	**1323.8**
313.05	**272.25**	**666.67**
27.051	**78.262**	**68.062**

Evolución de la población. El número de hembras en cada grupo de edad se incrementa con el tiempo, con cierto comportamiento oscilatorio.



Evolución de la población (escala logarítmica).



Matriz ortogonal de vectores propios y matriz diagonal de valores propios

>> [V,D]=eig(L)

V =

-0.94737 0.93201 0.22588

-0.31579 -0.356 -0.59137

-0.052632 0.067989 0.77412

D =

1.5 0 0

0 -1.309 0

0 0 -0.19098

El autovalor dominante λ1 =1.5 nos dice cómo cambia el vector de población de un año para otro

>>x100=L^100*x0

x100 =

1.0e+019

1.1555

0.3851

0.0642

>> x99=L^99*x0

x99 =

7.7033e+018

2.5678e+018

4.2796e+017

>> 1.5*x99

ans =

1.0e+019

1.1555

0.3851

0.0642

Comportamiento límite del porcentaje de población en cada clase de edad.

v1 =

-0.94737

-0.31579

-0.052632

Porcentaje de edad

>> v1/sum(v1)

ans =

0.72

0.24

0.04

Porcentaje de edad tras 100 años

>>x=x100/sum(x100)

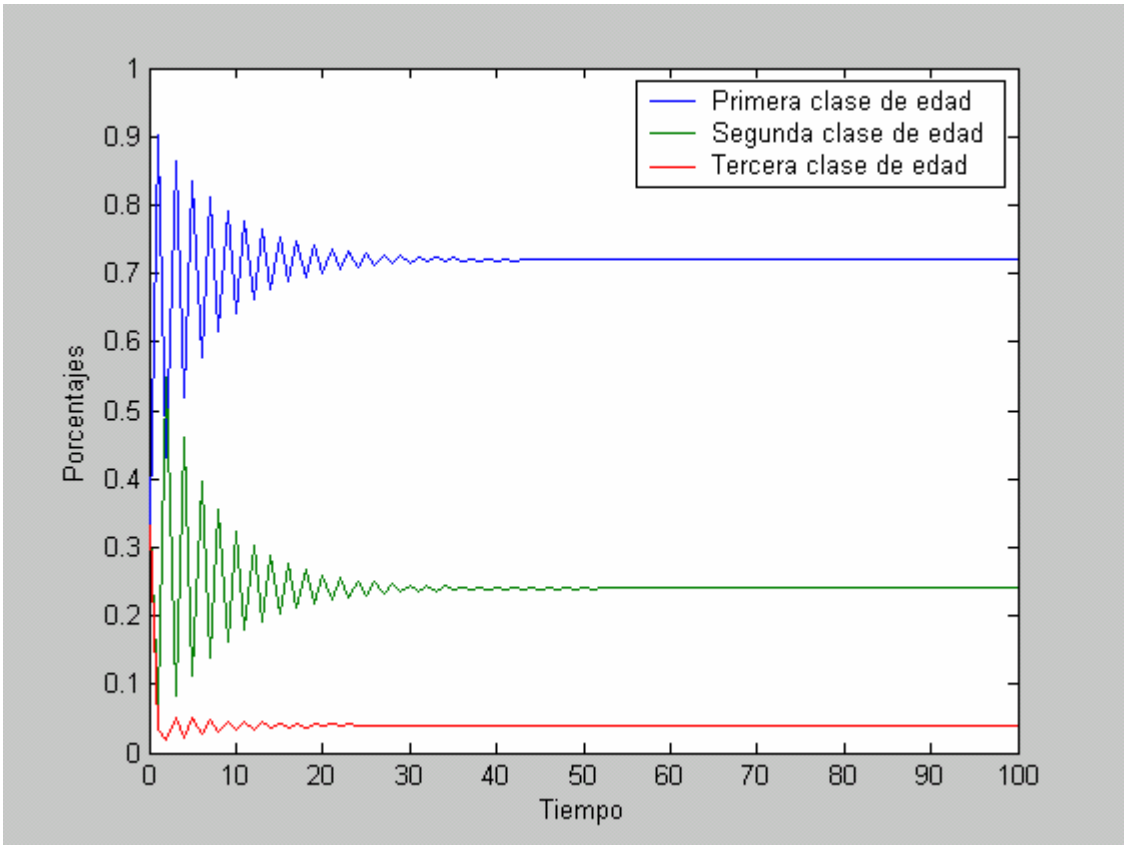
x =

0.72

0.24

0.04

Se puede apreciar que después de un número suficiente de años. El porcentaje de organismos en cada clase se aproxima a 74%, 24% y 4%.



Bibliografía

- **Leslie P.H. (1945). On the use of matrices in certain population mathematics. Biometrika, Volume 33, pp. 183-212**
- **Scanlan J. C. (1994). The use of state and transition model for predicting vegetation change in rangelands. Tropical Grasslands, Volume 28, pp. 229-240**

Inferencia del Coeficiente de Ajuste de Lundberg Para El modelo Clásico de Riesgo

Dr. Francisco Sergio Salem Silva.
Lic. Víctor Hugo Vázquez Guevara.

Primera Gran Semana de las Matemáticas

Resumen.

Se presenta el proceso clásico de riesgo y se discuten una cota y una aproximación para la probabilidad de ruina. Sin embargo, ambas presentan la dificultad de tener un parámetro que debe ser estimado. Se dan condiciones para la existencia de este parámetro, así como un método para estimarlo y un Teorema para validar su comportamiento asintótico.

1. El proceso Clásico de Riesgo.

Suponer que una compañía de seguros con un cierto capital inicial u debe pagar ciertas cantidades aleatorias de dinero a sus asegurados en caso de sufrir algún percance, los cuales ocurren también de manera aleatoria. Así mismo, la compañía recibe el pago de primas por parte de sus clientes a una taza p>0 por unidad de tiempo determinísticamente. Se define el proceso de riesgo R(t), t≥0 de la compañía al tiempo t por:

$$R_t = u + pt - \sum_{k=1}^{N_t} U_k$$

En dónde:
N_t Denota el número de reclamaciones en [0,t]. (Poisson (β)), y
U_n Denota el tamaño de la n-ésima reclamación (con distribución *B).*

2. Probabilidad de Ruina.

Se define ahora la probabilidad de ruina por:

$$\Psi(u) = P\left[\inf_{t\geq 0}\left(R_t < 0\right)\middle| R_0 = u\right]$$

Sin embargo, ésta no siempre puede calcularse de manera exacta, debido a la presencia de convoluciones de grados altos.

Sin embargo se cuenta con ciertas herramientas. La primera de éllas es una cota superior, la cual se llama Desigualdad de Lundberg:

$$\Psi(u) \leq e^{-\gamma u}$$

Mientras que la segunda es una aproximación, llamada de Crámer-Lundberg

$$\Psi\left(u\right) \sim Ce^{-\gamma u},\ u \to \infty$$

En dónde

$$C = \frac{1-\rho}{\beta \hat{B}'\left[\gamma\right] - 1}$$

Con

$$\rho = \lim_{t\to\infty}\frac{1}{t}\sum_{k=1}^{N_t} U_k$$
$$= \beta E\left[U\right]$$

Y γ (llamado coeficiente de ajuste de Lundberg) satisface

$$\beta\left(\hat{B}\left[\gamma\right] - 1\right) - \gamma = 0$$

Que es una de las versiones de la Ecuación de Lundberg.

3. Existencia del Coeficiente de Ajuste de Lundberg.

El coeficiente de ajuste no siempre existe, por tanto es importante señalar en qué casos está garantizada su existencia.

El coeficiente de ajuste de Lundberg existe cuando:

1. $\hat{B}[\alpha] < \infty,$ para toda $\alpha < \infty$.

2. Existe $\alpha' < \infty$ tal que $\hat{B}[\alpha] < \infty$ para toda $\alpha < \alpha'$ y $\hat{B}[\alpha] = \infty$ para toda $\alpha \geq \alpha'$

3. Existe $\alpha' < \infty$ tal que $\hat{B}[\alpha] < \infty$ para toda $\alpha \leq \alpha'$ y $\hat{B}[\alpha] = \infty$ para toda $\alpha > \alpha'$

4. Estimación del Coeficiente de Ajuste de Lundberg.

El coeficiente de ajuste de Lundberg no siempre es fácil de obtener (cuando existe) así que es necesario hacer una estimación de él. Esto se hará por medio de la solución empírica γ_T de la ecuación de Lundberg. Para ello, sean:

$$\beta_T = \frac{N_T}{T}, \quad \hat{B}_T[\alpha] = \frac{1}{N_T}\sum_{i=1}^{N_T} e^{\alpha U_i}, \quad \kappa_T(\alpha) = \beta_T\left(\hat{B}_T[\alpha] - 1\right) - \alpha$$

Y sea γ_T tal que $\kappa_T(\gamma_T)=0$.

Se presenta ahora un resultado que afirma que esta solución empírica converge en algún sentido al coeficiente de ajuste.

Teorema: Conforme $T \to \infty$ se cumple que $\gamma_T \xrightarrow{c.q.} \gamma$. Y si además se cumple que

$$\hat{B}[2\gamma] < \infty$$

entonces

$$\gamma_T - \gamma \approx N\left(0, \frac{1}{T}\sigma_\gamma^2\right)$$

en dónde

$$\sigma_\gamma^2 = \frac{\beta_\kappa(2\gamma)}{\kappa'(\gamma)^2}$$

5. Conclusiones.

- Se obtuvo una aproximación de la probabilidad de Ruina, sin embargo es necesario estimar el Coeficiente de Ajuste para hacer uso de élla.

- Con el fin de obtener otro enfoque se planea reducir el modelo de riesgo a una suma geométrica para acotar la probabilidad de ruina con la herramienta existente para las sumas geométricas.

- Se pretende aplicar el método de solución empírica para estimar parámetros presentes en otras desigualdades

6. Bibliografía.

- Ruin Probabilities

Soren Asmussen, Wolrd Sceintific.

- Essentials of Stochastic Processes

Rick Durrett, Springer.

- Lundberg Approximations for Compound Distributions with Insurance Applications

Gordon E. Willmot, X. Sheldon Lin, Springer

BLACK & SCHOLES SIN LÁGRIMAS

Liliana Santamaría Barrera.
Víctor Hugo Vázquez Guevara

Primera Gran Semana de las Matemáticas.

Resumen.

En este trabajo se deduce la fórmula de Black & Scholes para valuar opciones call europeas manteniendo la teoría al mínimo.

1. Introducción.

Existen contratos que son instrumentos para controlar el riesgo subyacente en los mercados y son utilizado por los especuladores en los mercados de acciones, paridad interbancaria de divisas, servicios y energía; por citar algunos. Sin embargo, éstos exigen el pago de una cierta cantidad (que sea justa para ambas partes) en compensación de una inequidad inherente a la volatilidad del mercado, esto es precisamente lo que la fórmula de Black & Scholes resuelve.

2. Un poco sobre las tazas de interés.

Suponga que tenemos una cantidad P de dinero que invertimos un periodo de tiempo t a una tasa de interés compuesto (de manera continua) r *en un ambiente libre de riesgo*, entonces el monto al final del periodo es

$$Pe^{rt}$$

De manera similar si sabemos que tendremos el monto P en el tiempo t, entonces el valor en el tiempo cero (valor presente) es

$$Pe^{-rt}$$

3. Definiciones preliminares.

OPCIÓN. Contrato que da a quien lo posee el derecho (pero no la obligación) de comprar o vender un bien a un precio fijo (precio de ejercicio).

OPCIÓN CALL. Esta opción da a la persona que la tiene, el derecho a comprar un bien a un precio acordado previamente.

OPCIÓN PUT. Esta opción da a la persona que la tiene el derecho de vender el bien por el precio establecido que se estableció en tiempo futuro.

Las opciones mas comunes son la Europea, la Americana, la Asiática y las Exóticas.

4. **Notación.**

De aquí en adelante se harán las siguientes convenciones:

S(t) Valor del bien en el mercado en el instante *t*
q Precio de ejercicio
u Fecha de expiración
r Tasa de interés en un ambiente libre de riesgos
S_0 Precio actual del bien en estudio

5. **Sobre el valor de las Opciones Europeas.**

Las posibles situaciones derivadas de la definición de opción call se pueden resumir en dos casos:

a) $q < S(u)$
b) $q > S(u)$

En el caso a) el tenedor de la opción ejerce su derecho a comprar el bien, y la ganancia que obtiene será $C= S(u)-q$ y en el caso b) no lo ejerce y por tanto no gana nada.

El suscriptor no gana en la fecha de expiración, es más puede perder un monto ilimitado.

Para compensar esta "inequidad", al momento de firmar el contrato el tenedor paga por el derecho que da la opción, a este monto se le denomina valor de la opción.

5.1 Valor de una Opción Call.

Como consecuencia del comentario anterior, el valor de la opción call europea en la fecha de expiración es:

$$C=máx(S(u)-q,0)$$

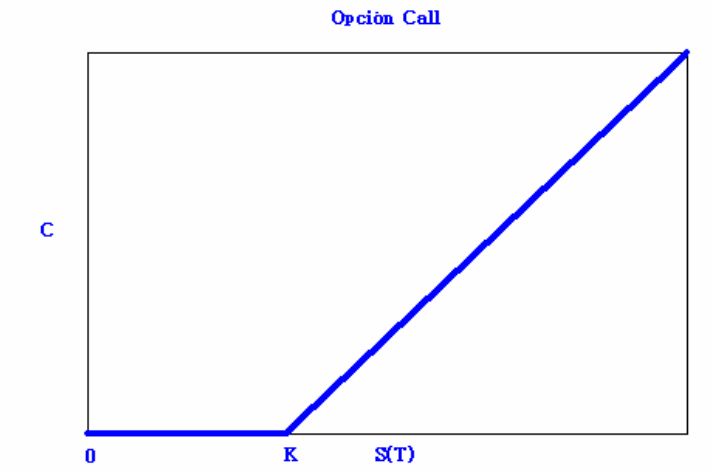


Diagrama de ganancia para una Opción Europea Call.

5.2 Valor de una Opción Put.

Mientras tanto, el valor de la opción put europea en la fecha de expiración es:

$$P=máx(q-S(u),0)$$

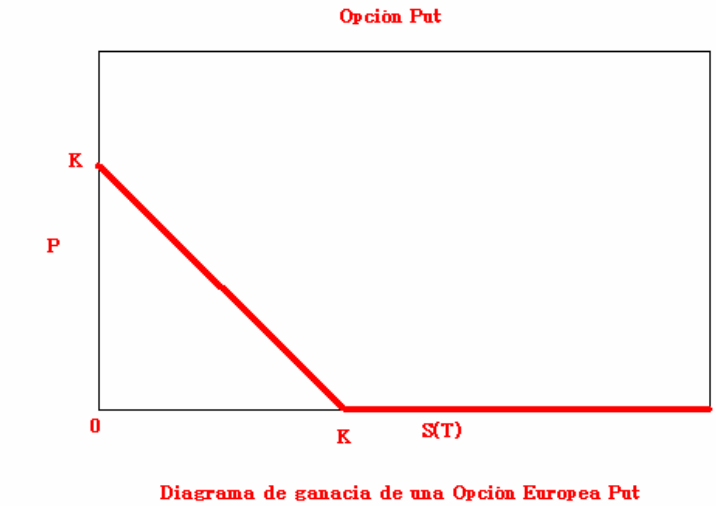


Diagrama de ganacia de una Opción Europea Put

6. Deducción de la fórmula de Black & Scholes.

Un modelo popular para modelar el cambio del precio de un activo al término de un periodo de longitud *u* es [1]:

$$S(u) = S_0 e^{Z_u}$$

En dónde Z_u se distribuye normalmente con media μu y varianza $\sigma^2 u$ y S_0 es el valor presente del activo.

A σ se le llama la volatilidad del precio del activo.

Tenemos además que [2]:

$$E\left(S(u)\right) = S_0 e^{\mu u + \frac{\sigma^2 u}{2}}$$

Notar que $S(u) = S_0 e^{Z_u} = e^{Z_u + \log(S_0)}$

Entonces *S(u)* tiene una distribución *log-normal* con parámetros $\mu u + log(S_0)$ y $\sigma^2 u$.

Black y Scholes (1973) desarrollaron un esquema para valuar opciones de activos cuyos precios tienen una distribución *log-normal.*

De aquí en adelante consideraremos un tiempo *u* fijo y escribiremos el precio del activo como:

$$S(u) = S_0 e^{\mu u + \sigma u^{\frac{1}{2}} z}$$

$$Z \sim N(0,1)$$

Supongamos que necesitamos valuar una opción call europea, con precio de ejercicio *q* y fecha de expiración *u*.

Supongamos que estamos en un ambiente libre de riesgo. Esto es, forzamos a que valor presente de $E[S_u]$ $(e^{-ru} E[S_u])$ sea igual a S_0.

Por otro lado $E\left[S_u\right] = S_0 e^{\mu u + \frac{\sigma^2 u}{2}}$

Igualando, obtenemos:

$$S_0 = e^{-ru} S_0 e^{\mu u + \frac{\sigma^2 u}{2}}$$

De aquí que:

$$\mu = r - \frac{\sigma^2}{2}$$

Ahora podemos determinar el precio de la opción.

El valor de la opción al tiempo *u* será $h(S_u)$, dónde

$$h(S) = \begin{cases} s - q & s > q \\ 0 & o.f. \end{cases}$$

Tenemos ahora que $h(S_u)>0$ sí y sólo sí

$$Z > \frac{\log\left(\frac{q}{S_0}\right) - \left(r - \frac{\sigma^2}{2}\right)u}{\sigma\sqrt{u}} = c$$

Entonces el precio de la opción en un ambiente libre de riesgo es el valor presente de E[$h(S_u)$], que es igual a:

$$e^{-ru} E\left[h(S_u)\right] = e^{-ru} \int_c^{\infty} \left[S_0 e^{\left[r - \frac{\sigma^2}{2}\right]u + \sigma u^{\frac{1}{2}} z} - q \right] \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

$$= e^{-ru} e^{-\frac{\sigma^2}{2}u} \int_c^{\infty} \left[S_0 e^{ru + \sigma u^{\frac{1}{2}} z} \right] \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz - e^{-ru} q \int_c^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

$$= S_0 \left[1 - \Phi(c - \sigma\sqrt{u}) \right] - e^{-ru} q \left[1 - \Phi(c) \right]$$

$$= S_0 \left[\Phi(\sigma\sqrt{u} - c) \right] - q e^{-ru} \left[\Phi(-c) \right]$$

El lado derecho de la igualdad

$$e^{-ru}E\left[h(S_u)\right] = S_0\left[\Phi(\sigma\sqrt{u}-c)\right] - qe^{-ru}\left[\Phi(-c)\right]$$

Es la famosa fórmula de Black & Scholes para valuar una opción call Europea.

En la cual, todos los parámetros son conocidos salvo quizás σ, sin embargo, éste puede ser estimado por métodos estadísticos.

Bibliografía.

1. Simulación de Ecuaciones Diferenciales Estocásticas Utilizando el Método Éuler-Murayama. Vázquez Guevara Víctor Hugo, 2004. Tesis de Licenciatura en Matemáticas. Colegio de Matemáticas, Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla.

2. Probability and Statistics. Morris H. DeGroot, Mark J. Schervish. Addison Wesley. Third Edition. 2002.

3. An Introduction to Financial Option Valuation. Desmond J. Higham. Cambridge. 2004.

4. Derivatives in Financial Markets with Stochastic Volatility. Jean-Pierre Fouque, George Papanicolaou. Cambridge. 2001.

Descomposición de Continuos

Alicia Santiago Santos.
Facultad de Ciencias Físico-Matemático. BUAP
es20707@alumnos.fcfm.buap.mx

Resumen

En este trabajo vemos un método de costrucción de continuos, este método es el de descomposiciones semicontinuas superiormente. Primero recordaremos la definición de espacio de descomposición en general y después veremos cuando el espacio de descomposición de un continuo es un continuo.

0.1 Definición. Sean (X, τ) es un espacio topológico y $\mathcal{D}$ una partición de X. Sea

$$\tau_{\mathcal{D}} = \{U \subset \mathcal{D} : \bigcup U \in \tau\}$$

Note que $\tau_{\mathcal{D}}$ es una topología para $\mathcal{D}$; de hecho, si $\pi : X \to \mathcal{D}$ definida para cada $x \in X$ por

$$\pi(x) = D \in \mathcal{D} \text{ donde } D \text{ es el único elemento de } \mathcal{D} \text{ tal que } x \in D$$

vemos que $\tau_{\mathcal{D}}$ es la topología más grande para $\mathcal{D}$ tal que π es continua. El espacio $(\mathcal{D}, \tau_{\mathcal{D}})$ es llamado espacio de descomposición de X, o más simplemente una descomposición de X. La topología $\tau_{\mathcal{D}}$ es llamada la topología de descomposición. Recalcamos que el término descomposición significa una descomposición con la topología descomposición.

Intuitivamente, una descomposición es el espacio obtenido del espacio original al identificar todos los puntos de cada elemento de la partición dada. Por esta razón las descomposiciones frecuentemente son llamadas espacios de identificación. Algunas veces también son llamados espacios cocientes.

Dado un continuo X su descomposicón no necesariamente es un continuo, aún cuando los elementos de la partición sean subconjuntos cerrados de X (tal partición es llamada una partición cerrada).

El siguiente ejemplo muestra lo dicho anteriormente.

0.1 Ejemplo. Sea $X = [-1, 1]$ y $\mathcal{D}$ la partición de X dada por

$$\mathcal{D} = \{\{x, -x\} : -1 < x < 1\} \bigcup \{\{-1\}, \{1\}\}$$

Entonces el espacio de descomposición $\mathcal{D}$ no es un continuo ya que $\mathcal{D}$ no es de Hausdorff y así, no es metrizable. En efecto, si U y V son abiertos en D tales que $\{-1\} \in U$ y $\{1\} \in V$. Entonces $\pi^{-1}(U)$ y $\pi^{-1}(V)$ son abiertos en $[-1,1]$, $-1 \in \pi^{-1}(U)$ y $1 \in \pi^{-1}(V)$. Sea $w \in X$ con $0 < w < 1$ tal que $[-1,-w) \in \pi^{-1}(U)$ y $(w,1] \in \pi^{-1}(V)$. Entonces $\forall a \in (w,1)$ se tiene que $\{a,-a\} \in \pi((-1,-w)) = \pi((w,1)) \subset U \cup V$ por lo tanto $U \cap V \neq \varnothing$.

Más adelante veremos la condición bajo la cual una descomposición de un continuo resulta ser un continuo. Pero antes veamos unos resultados

0.1 Lema. Sean X un espacio métrico compacto y $f : X \to Y$ una función continua y suprayectiva. Si Y es de Hausdorff entonces Y es metrizable.

Prueba. Como X es compacto y f es una función continua y suprayectiva, se tiene que Y es compacto. Ahora, como Y es compacto y Hausdorff, se tiene que Y es Normal. Así, por el Teorema de Metrización de Urysohn, es suficiente probar que Y tiene una base numerable. Para esto, sea C una base numerable para X. Para cada subconjunto finito $\mathcal{L}$ de C definimos

$$E_{\mathcal{L}} = Y \setminus f\left(X \setminus \bigcup \mathcal{L}\right)$$

Pongamos $B = \{E_{\mathcal{L}} : \mathcal{L} \text{ es un subconjunto finito de } C\}$. Como la colección de subconjuntos finitos de un conjunto numerable es numerable, tenemos que B es numerable. Así, basta demostrar que B es una base para Y. Primero note que, $f(X \setminus \bigcup \mathcal{L})$ es cerrado en Y (ya que $\bigcup \mathcal{L}$ es abierto y f es cerrada). Luego, cada elemento de B es un subconjunto abierto de Y. Ahora, sea U un subconjunto abierto de Y y $q \in U$. Observe que $f^{-1}(q)$ es cerrado en X y que $f^{-1}(U)$ es abierto en X. Entonces, $f^{-1}(q)$ es compacto en X y por lo tanto en $f^{-1}(U)$. Así, como C es base, existe un subconjunto finito $\mathcal{L}$ de C tal que

$$f^{-1}(q) \subset \bigcup \mathcal{L} \subset f^{-1}(U),$$

se sigue que facilmente que $q \in E_{\mathcal{L}} \subset U$. En resumen, hemos probado que B es una base numerable para Y. $\square$

0.1 Teorema. Una descomposición $\mathcal{D}$ de un espacio métrico compacto X es metrizable si y sólo si es de Hausdorff.

Prueba. Supongamos que el espacio de descomposición $\mathcal{D}$ es de Hausdorff. Como X es compacto y la función $\pi : X \to D$ es continua y suprayectiva, se sigue del Lema anterior que $\mathcal{D}$ es metrizable. La otra implicación se tiene porque todo espacio métrico es de Hausdorff. $\square$

Usaremos las descomposicones de espacios metricos compactos para construir otros espacios metricos compactos o continuos. Puede ser inconveniente tener que estar checando el Teorema 0.1 cada vez que querramos asegurar que una descomposici ón es de Hausdorff y así metrizable. Por lo tanto, queremos una condición que nos permita verificar esto con más facilidad. Como veremos en el Teorema 0.3, la siguiente definición da tal condición.

0.2 Definición. Sea (X, τ) es un espacio topológico. Una partición $\mathcal{D}$ de X se dice semi-continua superiormente (usc) si para cada $D \in \mathcal{D}$ y $U \in \tau$ tales que $D \subset U$, existe $V \in \tau$ con $D \subset V$ tal que si $A \in \mathcal{D}$ y $A \cap V \neq \varnothing$ entonces $A \subset U$.

0.1 Nota. Esta condición no toma en cuenta la topología de descomposición, es la partici ón la que es usc. Sin embargo, cuando la partición es usc, siempre se dice que la descomposición es usc teniendo presente que la descomposición todavía tiene la topología de descomposición.

0.3 Definición. Sea $\mathcal{D}$ una descomposición de X, un subconjunto de X es $\mathcal{D}-$ saturado si es la unión de una subcolección de $\mathcal{D}$.

Observaciones:

1. Si $\pi : X \to \mathcal{D}$ es la función natural, entonces $\pi^{-1}(C)$ es $\mathcal{D}-$saturado, para $C \subset \mathcal{D}$.

2. $A \subset X$ es $\mathcal{D}-$saturado si y sólo si $A = \pi^{-1}(\pi(A))$.

3. Si V es $\mathcal{D}-$saturado y abierto en X, $\pi(A)$ es abierto en $\mathcal{D}$.

La siguiente proposición da dos formas de ver a las descomposiciones usc.

0.1 Proposición. Sea (X, τ) un espacio topológico. Si $\mathcal{D}$ es una descomposición de X y $\pi : X \to \mathcal{D}$ es la función natural, entonces las siguientes condiciones son equivalentes.

1. $\mathcal{D}$ es una descomposición usc.

2. π es una función cerrada.

3. Si $D \in \mathcal{D}$, $U \in \tau$, y $D \subset U$, entonces existe $V \in \tau$ tal que $D \subset V \subset U$ y V es $\mathcal{D}-$saturado.

Prueba. [1 $\Rightarrow$ 2] Sea C un subconjunto cerrado de X. Por definición tenemos que $\pi(C)$ es cerrado si y sólo si $\pi^{-1}[\mathcal{D} - \pi(C)]$ es abierto en X. Sea $p \in \pi^{-1}[\mathcal{D} - \pi(C)]$. Entonces, $\pi(p) \in \mathcal{D} - \pi(C)$, luego $\pi(p) \subset X - C$ [puesto que si $y \in \pi(p) \cap C$, entonces $\pi(y) \cap \pi(p) \neq \varnothing$ así $\pi(y) = \pi(p)$, y en consecuencia $\pi(p) \in \pi(C)$]. Puesto que $X - C \in \tau$ y $\mathcal{D}$ es una descomposición usc, existe $V \in \tau, \pi(p) \subset V$ tal que si $x \in V$, entonces $\pi(x) \subset X - C$. Claramente $p \in V$, además $\pi(V) \subset \mathcal{D} - \pi(C)$ ya que si $\pi(x) \in C$ se tiene que $\pi(x) = \pi(y)$ para alguna $y \in C$, así $y \in \pi(x) \cap C$ y de aquí $\pi(x) \not\subseteq X - C$, por lo tanto $x \notin V$. Por lo anterior,

$$V \subset \pi^{-1}[\mathcal{D} - \pi(C)].$$

Como $p \in V \in \tau$, tenemos probado que $\pi^{-1}[\mathcal{D} - \pi(C)]$ es abierto. Con esto, $\pi(C)$ es cerrado.

[2 $\Rightarrow$ 3] Sea $D \in \mathcal{D}$ y $U \in \tau$ tal que $D \subset U$ entonces, poniendo

$$V = \pi^{-1}[D - \pi(X - U)]$$

se tiene que V satisface la condición 3.

[3 $\Rightarrow$ 1] Es consecuencia inmediata de la definición de descomposición usc. $\square$

0.2 Lema. Sea (X, τ) un espacio topológico T_1. Si $\mathcal{D}$ es una descomposición usc del espacio X, entonces D es una partición cerrada de X.

0.2 Nota. No toda partición cerrada es necesariamente usc.

0.2 Teorema. Toda descomposición usc de un continuo es un continuo.

Prueba. Sean X un continuo, $\mathcal{D}$ una descomposición de X y $\pi : X \to D$ la función natural. Usando la invarianza de la compacidad y la conexidad bajo la función continua π tenemos que $\mathcal{D}$ es un espacio compacto y conexo. Además, por el Teorema anterior es metrizable. Con todo se tiene que $\mathcal{D}$ es un continuo. $\square$

0.3 Teorema. Toda descomposición usc de un espacio métrico compacto es metrizable.

Prueba. Por el Teorema previo basta probar que el espacio de descomposisción es de Hausdorff. Para esto, sean (X, τ) un espacio topológico métrico compacto, $\mathcal{D}$ una descomposición de X y $\pi : X \to \mathcal{D}$ la función natural. Tomemos $D_1, D_2 \in \mathcal{D}$ tales que $D_1 \neq D_2$. Por el Lema previo D_1 y D_2 son subconjuntos cerrados y disjuntos de X. Luego, como X es normal, existen U_1 y $U_2 \in \tau$ tales que $U_1 \cap U_2 = \varnothing$ y $D_i \subset U_i$ para cada i. Además, ya que $\mathcal{D}$ es usc, por la proposición existen $V_1, V_2 \in \tau$ tales que $D_i \subset V_i \subset U_i$ y V_i son $\mathcal{D}$-saturados para cada i. Note que $D_i \in \pi(V_i)$ para cada i. Dado que $D_i \subset V_i$ y $D_i \in \mathcal{D}$, por la observación 3) se tiene que $\pi(V_1)$ y $\pi(V_2)$ son abiertos en $\mathcal{D}$. Resta demostrar que $\pi(V_1) \cap \pi(V_1) = \varnothing$. Primero note que $V_1 \cap V_2 = \varnothing$ ya que $U_1 \cap U_2 = \varnothing$ y $V_i \subset U_i$. Además, de la Observación 2 tenemos que

$$\pi^{-1}[\pi(V_1)] = V_i \text{ para cada } i$$

Así,

$$\pi(V_1) \cap \pi(V_1) = \varnothing.$$

Con todo, hemos probado que $\mathcal{D}$ es de Hausdorff. □

De este teorema se sigue que

0.1 Corolario. Toda descomposición usc de un continuo es un continuo.

Por último damos algunos ejemplos de descomposiciones usc.

0.2 Ejemplo. Banda de Möebius.

Sea X el cuadrado sólido $[0,1] \times [0,1]$ y sea $\mathcal{D}$ la partición de X cuyos elementos son

$$\{(x,0),(1-x,1)\} \text{ para } 0 \leq x \leq 1, \{(x,y)\} \text{ para } 0 < y < 1$$

Se puede ver que $\mathcal{D}$ es usc. Así, por el Corolario 0.1, el espacio descomposición es un continuo. El cual es llamado la Banda de Möebius.

0.3 Ejemplo. M- continuo.

Sea X el continuo seno-$(\frac{1}{x})$ y sea $\mathcal{D}$ la partición de X cuyos elementos no degenerados son

$$\{(0,y),(0,1-y)\} \text{ para cada } y \text{ tal que } 0 \leq y < \tfrac{1}{2}.$$

Se puede ver que $\mathcal{D}$ es usc. Así, por el Corolario 0.1, el espacio de descomposición es un continuo. El cual es llamado el M- continuo.

Referencias

1. S. B. Nadler, Jr. Continuum Theory: An introduction, Pure and Applied Mathematics Series, Vol. 158, Marcel Dekker, Inc., New York, Basel and Hong Kong, 1992.

2. J.R. Munkres, Topolología....

Propiedades Curiosas del Triángulo de Pascal y las Torres de Hanoi

Alicia Santiago Santos
Facultad de Ciencias Físico Matemáticas (FCFM)
Benemérita Universidad Autónoma de Puebla (BUAP)
18 sur y Av. San Claudio, Colonia San Manuel, Ciudad Universitaria

es20707@alumnos.fcfm.buap.mx

Resumen

¿Alguna vez se ha puesto a hacer algo aparentemente repetitivo? Por ejemplo que se ponga a pasar objetos de un lado a otro. Este proceso tiene que ver con un concepto matemático conocido como proceso de recurrencia. La recurrencia se presenta cuando tomas el número de eventos anteriores para obtener el actual número de eventos. En esta exposición abordamos dos nociones que tienen que ver con dicho concepto: las torres de Hanoi y el triángulo de Pascal. Además presentamos algunas propiedades curiosas que tiene este triángulo.

1. Introducción

Las matemáticas pueden ser tan divertidas como uno se lo proponga, por ejemplo el asociar la siguiente leyenda al juego de las torres de Hanoi: “En el gran templo de Benarés, bajo la cúpula que señala el centro del mundo, reposa una bandeja de cobre sobre la cual están colocadas tres agujas de diamante colocadas en forma vertical. Se cuenta que una mañana lluviosa el rey mandó colocar en una de las agujas sesenta y cuatro discos de oro puro, ordenados por tamaños; desde el mayor, que reposa en la bandeja, hasta el más pequeño, en lo alto de la aguja. Se llama la torre de Brahma.

Incansablemente, día tras día, los sacerdotes del templo mueven los discos haciéndolos pasar de una aguja a otra, de acuerdo a las leyes de Brahma, que dictan que el sacerdote en turno no mueva más de un disco a la vez, ni lo sitúe encima de un disco de menor tamaño..."

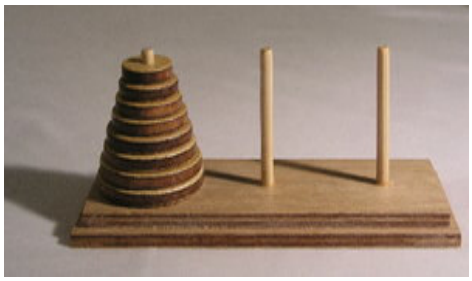
Suponiendo que cada movimiento tardara en realizarse un minuto.

¿Cuánto tiempo tardarían los sacerdotes en completar su trabajo? Por lo tanto *¿En qué tiempo se desintegraría el mundo?*

En esta exposición damos respuesta a las preguntas anteriores, además de que vemos que las torres de Hanoi y el triángulo de Pascal tienen una propiedad en común.

El juego de las torres de Hanoi o torres de diamante, es un juego oriental muy antiguo que hoy se conoce en todo el mundo. Este juego consta de tres columnas y una serie de discos de distintos tamaños. Los discos están acomodados de

mayor a menor en una de las columnas. La figura que sigue nos muestra un modelo del juego, aunque es importante señalar que hay varias formas de representar el juego, una forma es representarlo en forma triangular.



El juego consiste en pasar todos los discos a otra de las columnas y dejarlos acomodados como estaban: de mayor a menor. No importa la columna a la que se pasen.

Las reglas del juego son las siguientes:

1. Sólo se puede mover un disco a la vez.
2. Para pasar los discos de lugar se pueden usar las tres columnas del juego; es decir, que los distintos discos se pueden ir acomodando en las columnas según convenga.
3. Nunca deberá quedar un disco grande sobre uno chico.

Realizando el juego para el caso de 3 discos podemos ver que es necesario llevar a cabo 7 movimientos, para el caso de 4 discos 15 movimientos, etc. entonces al ir aumentando el número de discos el número mínimo de movimientos crece de manera exponencial. Así:
Para 1 disco hace falta 1 movimiento.
Para 2 discos hacen falta 3 movimientos.
Para 3 discos hacen falta 7 movimientos.
Para 4 discos hacen falta 15 movimientos.

NOTA: Llamaremos evento a cada paso realizado.

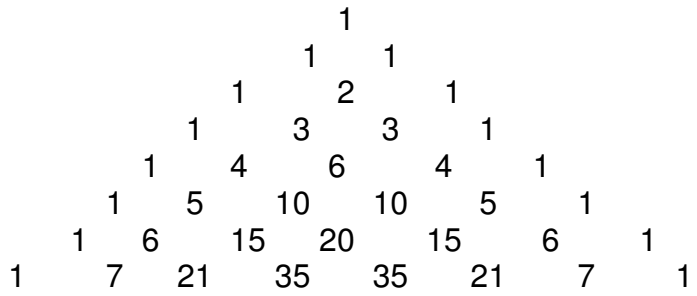
Entonces para encontrar la respuesta a la pregunta hecha al principio tendríamos que construir una torre con 64 discos y ejecutar cada uno de los pasos. Como es de imaginarse esto sería un poco cansado, así que hay que intentarlo de otra manera.

El estudio del triángulo de Pascal deja asombrado a cualquiera por todo lo que se puede hacer con él, vemos como nos ayuda a resolver el problema plantado en la leyenda. Para eso primero recordemos quien fue y que hizo Blaise Pascal.

Blaise Pascal fue un matemático y físico francés que vivió de 1623 a 1662. Construyó una máquina sumadora a la que llamo "Pascalina" y hoy en día es

considerada la primera máquina sumadora de la historia, Inventó la jeringa y la prensa hidráulica y descubrió lo que hoy se conoce como "la Ley de la Presión de Pascal". Trabajó en distintas áreas de las matemáticas pero uno de sus descubrimientos más famosos es el conocido "triángulo de Pascal".

Como su nombre nos lo dice, construiremos un triángulo formado por números de la siguiente manera. Se empieza por el « 1 » de la cumbre. De una línea a la siguiente se conviene escribir los números dejando media casilla. Así, las casillas tendrán cada una dos casillas justo encima, en la línea anterior. El valor que se escribe en una casilla es la suma de los valores de las dos casillas encima de ella. El valor cero no se escribe. Las primeras líneas han sido representadas en la siguiente figura.



El triángulo de Pascal es un triángulo de números enteros, infinito, simétrico y debe su nombre al célebre matemático Blaise Pascal quien estudió algunas propiedades del mismo. No obstante hay que recordar que el triángulo de Pascal era conocido desde mucho antes. Las primeras referencias del triángulo corresponden a China, donde está constatado que el triángulo era conocido alrededor de 1100. En relación con el triángulo de Pascal se suelen citar al matemático chino Yang Hui, del siglo XIII, conocido por haber estudiado algunas de sus propiedades, y al matemático persa Omar Khayyam, del siglo XI-XII, cuyo descubrimiento del triángulo se presume que fue independiente del descubrimiento por parte de los matemáticos chinos.
El interés de dicho triángulo se debe a múltiples razones a continuación vemos algunas propiedades interesantes que se pueden encontrar en él.

- Los números que aparecen en cada fila son los coeficientes que se obtienen al desarrollar

  $$(a + b)^n.$$

  Por ejemplo, si nos fijamos en la fila-3 observamos que los números 1, 3, 3, 1 son precisamente los coeficientes del desarrollo de

  $$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

- Aparecen los números poligonales (los triangulares, los cuadrados, los

pentagonales, etc.) los cuales fueron inventados por los pitagóricos.

Recordemos que un número es triangular cuando se puede representar como un triángulo equilátero formado por puntos de modo que se vayan alternando los puntos desde el vértice superior formado por un solo punto hasta la base del triángulo contando en cada fila un punto más que en la inmediata superior.

Los números triangulares *(*1, 3, 6, 10, 15, ...) son enteros del tipo

$$N = 1 + 2 + 3 + ... + n$$

Los números cuadrados (1, 4, 9, 16, 25, ...) son enteros del tipo

$$N = 1 + 3 + 5 + 7 + ... + (2n-1)$$

Los números pentagonales (1, 5, 12, 22, ...) son enteros del tipo

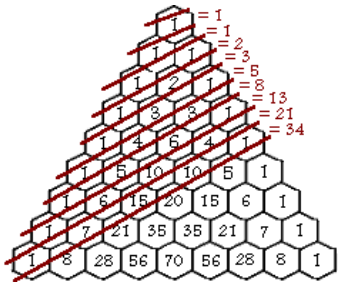
$$N = 1 + 4 + 7 + ... +(3n-2)$$

Los números hexagonales (1, 6, 15, 28, ...) son enteros del tipo

$$N = 1 + 5 + 9 + ... + (4n-3)$$

Y así sucesivamente.
Los números triangulares los podemos encontrar en la tercera diagonal del triángulo y los números cuadrados se encuentran en el triángulo de Pascal recurriendo a la misma diagonal que en el caso anterior, pero cada uno es construido sumando dos números triangulares consecutivos. Eso nos proporciona: 1, 4, 9, 16, 25, ...

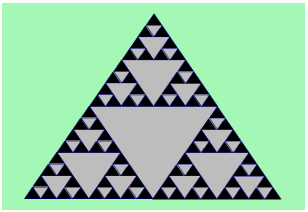
- Cualquier diagonal que empiece en un extremo del triángulo, y de la longitud que sea, cumple la siguiente propiedad: La suma de todos los números que la integran se encuentran justo debajo del último de ellos, en la diagonal contraria.
- La serie de Fibonacci (1, 1, 2, 3, 5, 8, 13, 21, …) puede ser encontrada dividiendo al mismo según las líneas que mostramos en el diagrama, los números atrapados entre ellas suman cada uno de los elementos de esta sucesión.



La suma de los elementos de cualquier fila es el resultado de elevar 2 al número que define a esa fila. Así:
$2^0 = 1$
$2^1 = 1+1 = 2$
$2^2 = 1+2+1 = 4$
$2^3 = 1+3+3+1 = 8$
$2^4 = 1+4+6+4+1 = 16$

- Si el primer elemento de una fila es un número primo, todos los números de esa fila serán divisibles por él (menos el 1, claro). Así, en la fila 7: (1 7 21 35 35 21 7 1) los números 7, 21 y 35 son divisibles por 7.

- Cualquier diagonal que empiece en un extremo del triángulo, y de la longitud que sea, cumple la siguiente propiedad: La suma de todos los números que la integran se encuentran justo debajo del último de ellos, en la diagonal contraria.
- El triangulo de sierpinski se trata de un fractal determinístico que se puede generar de diversas formas. La más usual consiste en partir de un triángulo equilátero, marcar los puntos medios de sus lados y extraer el triángulo interior (considerado como conjunto abierto). Se repite el proceso con los tres triángulos que quedan y así sucesivamente (formalmente el triángulo de Sierpinski se define como la intersección de los conjuntos cerrados que van apareciendo en cada etapa)**.**
  Las siguientes figuras representa la tercera etapa y en la segunda la cuarta etapa.

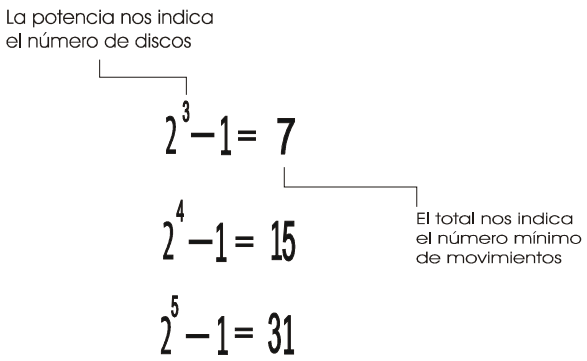


Entonces si elegimos colorear los números pares se observa perfectamente que al ir aumentando el número de filas el objeto resultante se va aproximando al triángulo de Sierpinski.

Ahora veamos como obtener las respuestas de las preguntas planteadas al principio a partir del triángulo de Pascal.

Como la suma de los elementos de cualquier fila es el resultado de elevar 2 al número que define a la fila entonces las potencias nos indican el número de discos que se tienen puestos en la primera columna. Si a cada uno de estos resultados les restamos la unidad, tenemos el número mínimo de movimientos en el que

logramos pasar todos los discos a la tercera columna, siguiendo las reglas del juego.



En general tenemos que para n discos hacen falta 2n - 1 (2 a la n menos 1) movimientos.
Volviendo a la leyenda el tiempo que necesitarían los sacerdotes para terminar el juego, es el siguiente:

264 - 1, o sea **18,446,744,073,709,551,615.**

Que es el número de movimientos necesarios.
Suponiendo que los sacerdotes realicen un movimiento por segundo y trabajen las 24 horas del día, durante los 365 días del año, tomando en cuenta los años bisiestos tardarían 58,454,204,609 siglos más 6 años en concluir la obra siempre que no se equivoquen, pues un pequeño descuido podría echar por tierra todo lo hecho.

Además de que el triángulo de Pascal nos ayuda en la solución del juego de las torres de Hanoi, este triángulo y el juego tienen algo en común y es el ser procesos recurrentes. La recurrencia se presenta cuando tomamos el número de eventos anterior para obtener el actual número de eventos. Por ejemplo para el juego observemos la siguiente tabla

# Discos		# Eventos
1	0 + 0 + 1	1
2	1 + 1 + 1	3
3	3 + 3 + 1	7
4	7 + 7 + 1	15
5	15 + 15 + 1	31
6	31 + 31 + 1	63

En ella podemos ver que para el evento siguiente necesitamos del evento anterior y en el caso del triángulo de Pascal, necesitamos de la fila anterior para conseguir la fila siguiente. Otro ejemplo de un proceso de recurrencia es la construcción de los números de Fibonacci.

Sugerencias

Inventa tus propios juegos de recurrencia e investiga más sobre el fabuloso triángulo de Pascal y sus aplicaciones.

Geometría Olímpica

Hugo Villanueva Méndez
Facultad de Ciencias Físico Matemáticas (FCFM)
Benemérita Universidad Autónoma de Puebla (BUAP)
Puebla, Puebla, 72570
Est061@fcfm.buap.mx

Resumen

En esta plática se propondrán varios problemas de Geometría que han aparecido en exámenes de las diferentes etapas de la olimpiada de matemáticas, tanto en el estado de Puebla como a nivel nacional. Al resolver estos problemas intentamos d un panorama general de los conocimientos necesarios de esta materia que se estudia en los cursos usuales de matemáticas en secundaria y bachillerato.

Introducción.

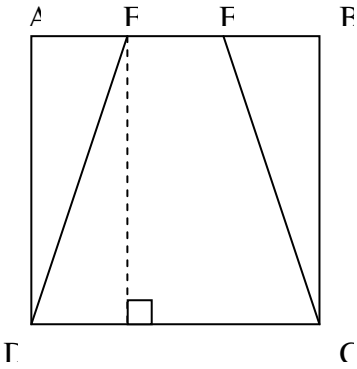
La Olimpiada de Matemáticas es un concurso donde los problemas no son del estilo escolar donde sólo hay que aplicar fórmulas para resolverlos. En este tipo de problemas se necesita de creatividad, de pensar, analizar los problemas e ir formando una solución bien argumentada de manera lógica. La geometría es una de las áreas que forman parte de la olimpiada, en ésta uno puede visualizar los problemas mediante un dibujo, lo cual ayuda mucho a imaginar y manejar los datos para la solución del problema. Veamos ejemplos de problemas de diferentes niveles de dificultad. Se invita al lector a intentar resolverlos antes de ver la solución.

Problema 1. Considere un cuadrado ABCD de área 9. Sean E y F dos puntos sobre el lado AB tales que AE=EF=FB. Calcula el área del trapecio EFCD.

Solución. Como cualquier problema de geometría, lo primero que hay que hacer es una figura que cumpla con las condiciones del problema.

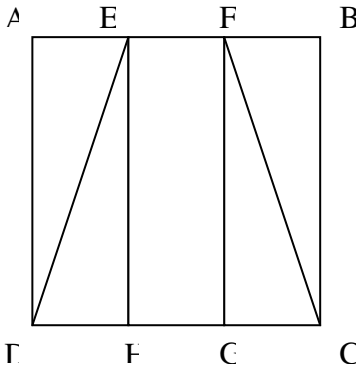
Ahora, la solución inmediata es darse cuenta que, como el área de un cuadrado es lado por lado, entonces el lado mide 3 de longitud. Después observar que se puede aplicar la fórmula del área de un trapecio pues la base mayor es DC que mide 3, la base menor es EF que mide una tercera parte del lado, es decir 1, y la altura coincide con el lado del cuadrado. Se tiene que el área buscada es

$$\frac{(3+1)3}{2} = 6.$$



Esta no es una solución olímpica, más bien es una solución escolar. Se necesita conocer la fórmula y aplicarla para obtener el resultado. Pero en este tipo de concurso en general las soluciones no son de este estilo. Puede suceder que alguien no recuerde dicha fórmula y necesite hacer algo distinto. Veamos lo que es una solución olímpica.

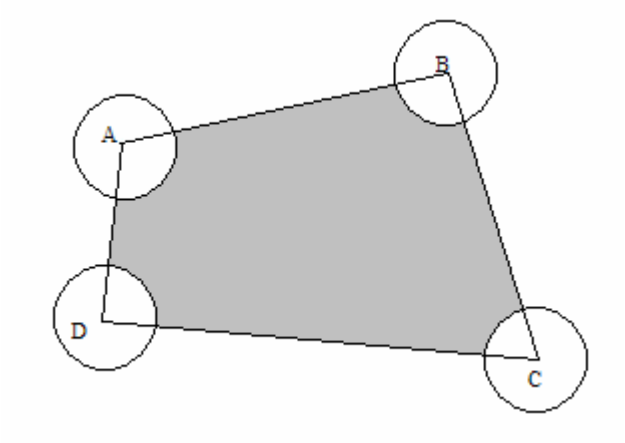
Tomemos los puntos G y H en CD tales que CG = GH = HD. Es fácil observar que los segmentos EH y FG dividen al cuadrado en tres rectángulos iguales y por tanto, de misma área. También es claro que los triángulos FGC y EDA son iguales, ambos son la mitad de un rectángulo. Entonces podemos reacomodar el área del trapecio al considerar al triángulo ADE en lugar del triángulo FGC para obtener el rectángulo AFGD que está formado por dos rectángulos. Como cada rectángulo es la tercera parte del cuadrado, su área es 3. De aquí que el área del trapecio que es igual al área del rectángulo AFGD, es igual a dos veces al área de un rectangulito, es decir 6, que es lo que anteriormente obtuvimos.



Podemos observar que en esta solución no hace falta conocer la fórmula, ni saber la longitud del lado del cuadrado a diferencia de la primera solución. Fue una solución más creativa que el simple hecho de hacer cuentitas. Veamos otro ejemplo.

Problema 2. En la siguiente figura, ABCD es un cuadrilátero de área 100. Con centro en los vértices se construyen circunferencias de radio 3. Calcular el área sombreada.

Solución. A diferencia del problema anterior donde teníamos una fórmula que nos ayudara, para esta figura no la tenemos. Más aún, no sabemos siquiera qué tipo de cuadrilátero es. Entonces hay que intentarlo de otra manera.

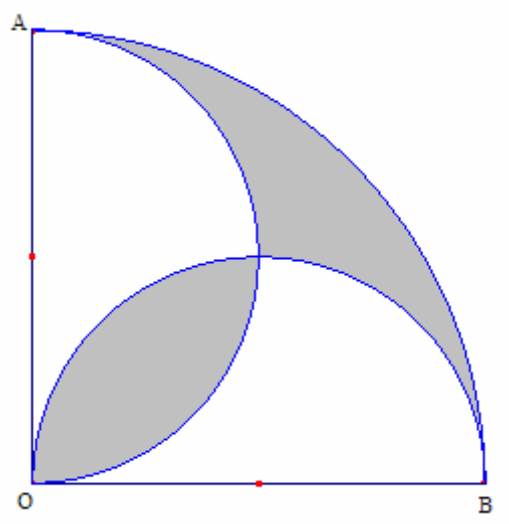


No tenemos una manera de calcular el área de manera directa. Muchas veces como en este caso hay que darle la vuelta al problema. Es decir, si pudiéramos calcular el área que no queremos y se la restamos al área total, que conocemos es de 100, obtendríamos el área deseada. Nuestro problema es ahora encontrar el área restante.

Observemos que el área restante es la suma de las áreas de cuatro sectores de círculos de radio 2. Como los ángulos internos de un cuadrilátero es 360º, al unir los cuatro sectores, se completaría un círculo. Por lo que el área restante es igual al área de un círculo de radio 2, es decir, 4π. De aquí que el área deseada es $100 - 4\pi$.

Aumentemos el nivel de los problemas.

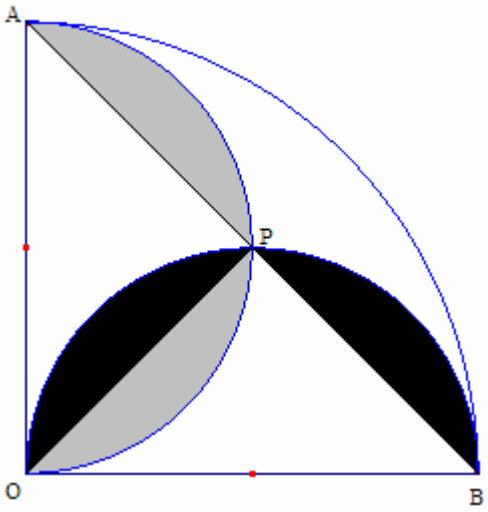
Problema 3. El arco AB es un cuarto de una circunferencia de centro O y radio 10. Los arcos OA y OB son semicircunferencias. ¿Cuál es el área de la región sombreada?



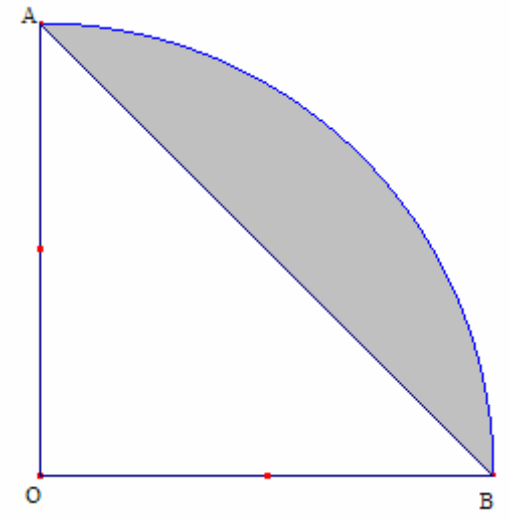
Solución. Nuevamente, la figura sombreada no es alguna para la cual conocemos una fórmula para obtener su área. Entonces podríamos intentar lo hecho en el problema anterior. Puesto que podemos calcular el área total (un cuarto de círculo de radio 10), solo nos faltaría calcular el área blanca.

¿Y cómo? El área blanca no se ve muy fácil de calcular. Una manera de intentarlo es utilizar nuevamente la idea de encontrar el resto, pues podemos encontrar el área del sector OB sabiendo que es un semicírculo de radio 5; pero tendríamos que calcular la región común a los dos semicírculos y regresamos de alguna manera al problema original. Entonces hay que utilizar algo más.

Algo que nos funcionó en la solución olímpica del primer problema fue reacomodar la región deseada de manera que nos quedara una región con la misma área pero más fácil de calcular. Esta misma idea la podemos seguir de la siguiente manera:



Si P es el punto (distinto de O) donde se intersectan los dos semicírculos, entonces la simetría de la figura nos permite afirmar que el segmento AB pasa por el punto P. Además, las regiones grises y las regiones negras tienen la misma área. Por lo tanto, podemos reacomodar la región formada por la intersección de los dos semicírculos y cambiarla por estas dos nuevas regiones formadas por las cuerdas AP y PB.

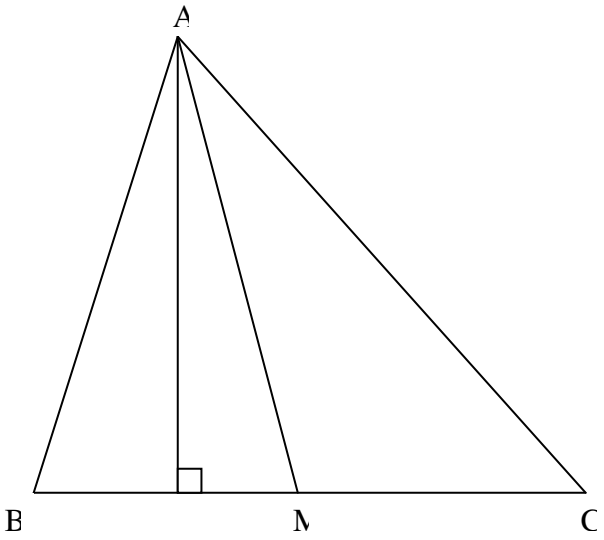


Entonces la región buscada tiene la misma área que la cuerda AB, la cual la podemos calcular de la siguiente manera: al área total le restamos el área del triángulo AOB, la cual podemos calcular. El área buscada es:

$$\frac{10^2\pi}{4}-\frac{10x10}{2}=25\pi-50.$$

Los anteriores problemas prácticamente no necesitan de teoría fuera de lo usual para resolverlos, pero llega el momento en que es necesaria cierta teoría para poder dar con la solución. Por eso, conforme los participantes van avanzando, se les va entrenando y se les dan nuevas herramientas y estrategias para resolver problemas.

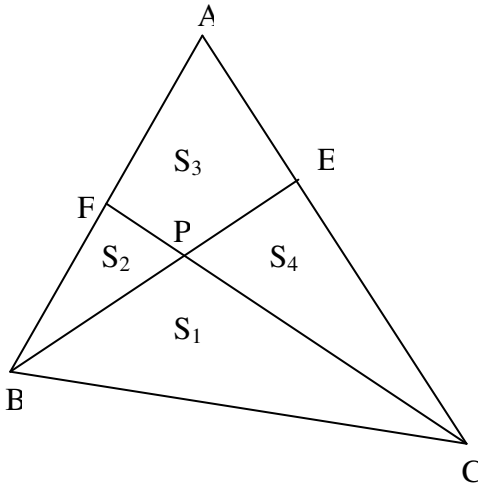
Por ejemplo, el segmento que une un vértice con el punto medio del lado opuesto de un triángulo (mediana), divide a éste en dos triángulos con la misma área. Es fácil de ver pues los dos triángulos tendrán la misma altura desde su vértice común A y la misma base al ser ambas la mitad de la longitud del lado BC del triángulo. Más aún, se cumple el recíproco, si un segmento que une un vértice con un punto del lado opuesto de un triángulo divide a éste en dos triángulos con la misma área, entonces dicho punto es el punto medio del lado. También es fácil de ver pues los dos triángulos tienen la misma área y la misma altura, entonces deben tener la misma base, es decir BM = MC, lo cual nos dice que M tiene que ser el punto medio de BC.



El siguiente problema es importante, no sólo porque ya es necesaria cierta teoría (aún no difícil) para resolverlo, sino porque con él se puede dar una idea de una estrategia importante en la solución de problemas.

Problema 4. El triángulo ABC se ha dividido en regiones más pequeñas como se muestra en la figura, cuyas áreas son S_1, S_2, S_3 y S_4. ¿Cuándo es posible que $S_1 = S_2 = S_3 = S_4$?

Solución. La pregunta "¿Cuándo es posible que $S_1 = S_2 = S_3 = S_4$?", se refiere a que hay que decir para que tipo de triángulo y qué líneas particulares se cumple. Podríamos comenzar analizando al triángulo equilátero, el más sencillo, pero una vez elegido, tenemos que analizar qué líneas lo cumplen; hay una infinidad de casos para analizar. Y esa cantidad hay que analizarla para cada triángulo. Lo cual se ve muy difícil.



Parece que no hay una manera directa de atacar el problema, entonces hay que entrarle de manera indirecta. Uno debe tratar de responder a la pregunta: "¿Qué necesito para verificar lo que me piden?".

Comencemos entonces suponiendo que $S_1 = S_2 = S_3 = S_4 = S$. Observemos ahora al triángulo EBC. P es un punto sobre el lado BE tal que divide al triángulo en dos triángulos con la misma área. Lo mencionado antes de este problema nos dice entonces que P es punto medio de EB. Visto de otra manera, diríamos que para que los triángulos EPC y PBC tengan la misma área, P debe ser punto medio de EB.

Ahora, en el triángulo ABC, el área del triángulo FBC es la suma de las áreas de los triángulos BPC y BPF, es decir $S_1 + S_2 = 2S$; y el área del triángulo AFC es la suma de las áreas del triángulo EPC y del cuadrilátero AFPE, es decir $S_4 + S_3 = 2S$. Entonces, ambos triángulos AFC y FBC tienen la misma área 2S; de aquí que F es punto medio de AB.

Hemos visto que si $S_1 = S_2 = S_3 = S_4$, entonces P es punto medio de EB y F es punto medio de AB. Si observamos al triángulo ABE, los puntos E y F son puntos medio de dos de sus lados. Sabemos además que el segmento que une los puntos medios de dos lados de un triángulo es paralelo al tercer lado. De aquí que FP es paralelo a AC. Pero esto es imposible pues estas líneas se intersectan en el punto C. ¿Qué está mal? Lo que nos llevó a concluir algo falso fue la suposición hecha al principio. Entonces, si suponiendo eso llegamos a algo no cierto, la suposición inicial debe estar mal, debe ser falsa. Concluimos finalmente que NUNCA es cierto que $S_1 = S_2 = S_3 = S_4$.

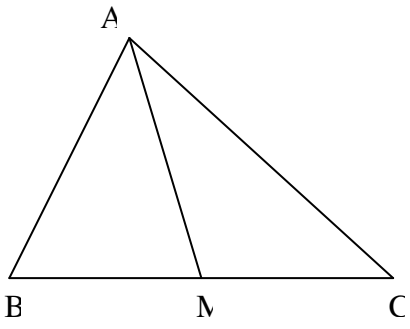
Como dijimos antes, este problema nos daría una manera de pensarlos. Cuando se ve difícil la manera directa de atacarlo, esta es una buena opción. Este problema sirve entonces para introducir a los participantes a lo que se llama método de "reducción al absurdo" o "contradicción", el cual es muy útil en muchos problemas no solo de geometría ya que se ocupa en todas las áreas.

Hasta aquí, los problemas han sido del tipo: "calcula…", "¿cuándo sucede que…?". Cuando los participantes presentan el examen de la etapa regional de la olimpiada de matemáticas se pueden enfrentar a otro tipo de problemas.

Problema 5. En un triángulo, la longitud de una mediana es igual a la mitad de la longitud del lado a la que fue trazada. Demuestra que el triángulo es un triángulo rectángulo.
(Nota: Una mediana es el segmento que une un vértice con el punto medio del lado opuesto)

Solución. ¿Qué diferencia hay entre este problema con los anteriores? La primera diferencia que se puede observar con tres de los cuatro problemas anteriores es que este problema no tiene un dibujo.

Esto se puede solucionar haciendo la figura que cumpla con las condiciones del problema, no más no menos. No hay que suponer de entrada que el triángulo es de algún tipo, en este caso rectángulo. Ya que el problema solo nos habla de "un triángulo". Lo único que nos da el problema es una condición acerca de una mediana del triángulo.



Por lo tanto sólo hay que suponer esto y hacer nuestra figura, un triángulo ABC, M el punto medio de BC y AM es la mediana tal que mide la mitad de BC.

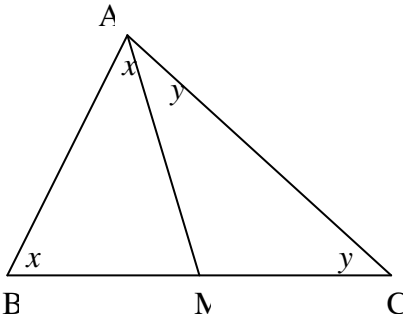
Otra diferencia y quizá la más importante es lo que el problema pide que hagamos: "Demuestre que…". Cuando un estudiante comienza a enfrentarse con

problemas de olimpiada no sabe qué significa “demostrar” y puede hacer toda clase de soluciones incorrectas. Demostrar no es otra cosa que argumentar de manera lógica por qué algo es cierto Veamos cómo se resuelve el problema.

Primero, ¿cómo verificamos que el triángulo es rectángulo?, es decir, ¿cómo atacamos el problema? Tenemos dos maneras, por un lado sabemos que los triángulos rectángulos tienen un ángulo recto o de 90º; por otro lado, sabemos que un triángulo rectángulo cumple con el teorema de Pitágoras. Entonces para verificar que el triángulo del problema es rectángulo hay que ver que uno de sus ángulos es de 90º o bien que sus lados cumplen con el teorema de Pitágoras.

¿Cuál escogemos? El problema no nos dice nada acerca de ángulos ni tampoco nos da una relación que involucre los tres lados del triángulo. Una manera de comenzar el problema es obtener más información a partir de la dada.

¿Qué tenemos? Que AM es la mitad de BC, es decir AM = ½ BC. ¿Qué más? Como M es punto medio de BC, entonces BM = MC = ½ BC. Entonces AM = BM = CM = ½ BC. Por lo tanto los triángulos ABM y AMC son isósceles. De aquí que los ángulos respectivos son iguales. Llamémosle x y y a dichos ángulos como se muestra en la figura. ¡Ya tenemos ángulos involucrados! ¿Ahora?



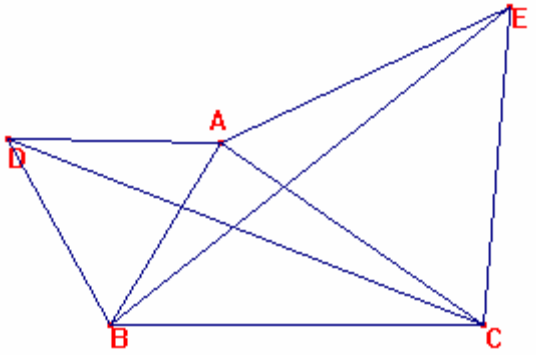
Sabemos que los ángulos internos en un triángulo suman 180º, entonces $180º = \angle B + \angle A + \angle C = x + (x + y) + y = 2x + 2y = 2(x + y)$. De aquí que (dividiendo entre 2) $x + y = 90º$. Pero $\angle A = x + y$. Concluimos que $\angle A = 90º$. ¡Ya acabamos! Entonces el triángulo ABC es rectángulo.

Se puede calcular la mediana de un triángulo respecto a los tres lados. Con esto y cuentitas simples se puede demostrar que los lados del triángulo del problema cumplen con el teorema de Pitágoras lo cual nos daría otra solución. Pero la solución presentada es más ilustrativa. Este problema apareció en la etapa regional del estado de Puebla. Pasemos al siguiente nivel.

Problema 6. Sobre los lados AB y AC de un triángulo ABC se construyen los triángulos equiláteros ABD y ACE. Demuestra que DC = EB.

Solución. Nuevamente primero hagamos un dibujo que ilustre el problema.

Ahora, ¿cómo demostrar lo que se pide? Si los dos segmentos a demostrar que son iguales tienen un extremo común, la estrategia es simple, tratar de demostrar que el triángulo formado es isósceles. Pero éste no es el caso. Aquí los segmentos parecen no tener nada que ver uno con el otro.



Entonces la estrategia, aunque básica, no es obvia. Si encontramos dos triángulos, cada uno con un lado igual a uno de los segmentos deseados, y demostramos que son iguales, entonces sus tres lados son iguales, en particular los segmentos buscados.

¿Qué triángulos parecen cumplirlo en este problema? Se invita al lector a buscarlos antes de pasar al siguiente párrafo.

Observemos los triángulos ADC y ABE. Uno tiene a DC como lado y el otro tiene a EB. Como el triángulo ADB es equilátero, AD = AB y como AEC es equilátero, AE = AC. Para ver que los triángulos son congruentes (iguales), según el criterio LAL, nos falta ver que $\angle DAC = \angle BAE$, pues ya tenemos los dos lados iguales.

Pero $\angle DAC = \angle DAB + \angle BAC = 60^\circ + \angle BAC = \angle CAE + \angle BAC = \angle BAE$. Concluimos que los triángulos DAC y BAE son congruentes. Por lo tanto sus lados son iguales, en particular CD = EB, que es lo que se quería demostrar.

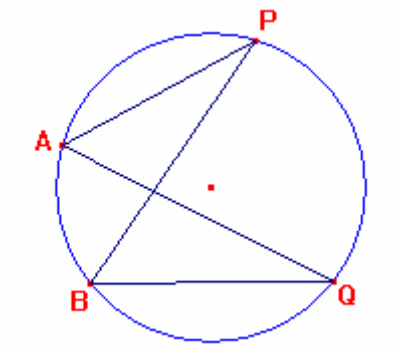
Para este problema no fue necesario comenzar con obtener información adicional pues teníamos una estrategia que pudimos seguir. En el anterior teníamos qué queríamos pero no sabíamos cómo obtenerlo.

Es bueno que quienes inician comiencen a atacar el problema obteniendo toda la información adicional que el problema proporcione, puesto que uno no sabe que va a utilizar y que no sirve. Este problema es de nivel de etapa estatal.

Pasemos al último problema, de mayor nivel. Apareció en un examen selectivo de Puebla, es decir, el examen que sirvió para elegir a los seis estudiantes que representarían a Puebla en el concurso Nacional. Antes unos comentarios.

Ya se dijo que conforme avanzan se les proporcionan nuevas herramientas y estrategias. Es decir, se les enseñan y demuestran diversos teoremas. Ahora bien, no se les enseña como una fórmula que van a aplicar e inmediatamente el problema está resuelto, más bien como una herramienta para obtener información que ayude a resolverlo. Uno no sabe cuando y cómo va a ocupar los teoremas, tiene que tenerlos presentes siempre para cuando se necesiten.

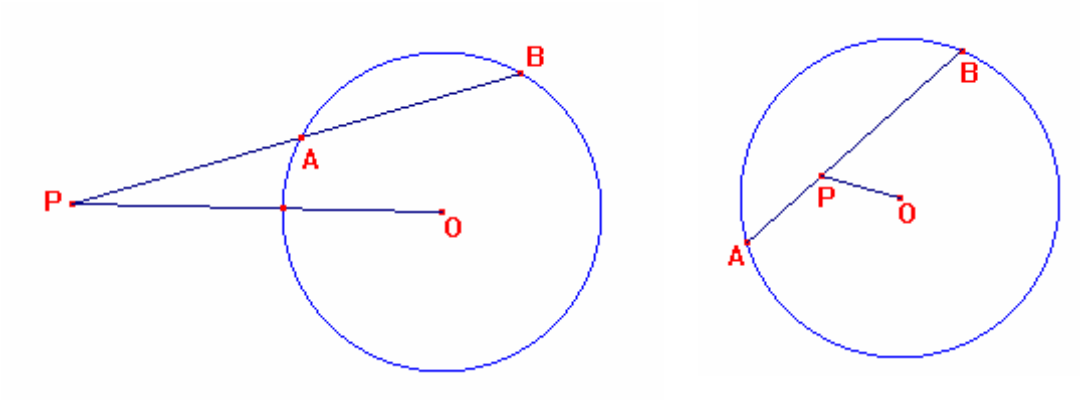
Dos resultados que usaremos en la solución del problema son los siguientes.



El primero es básico, de los primeros que se muestran, dice que si dos ángulos inscritos en una circunferencia abren el mismo arco, entonces son iguales. En la figura, $\angle APB = \angle AQB$, pues abren el arco AB.

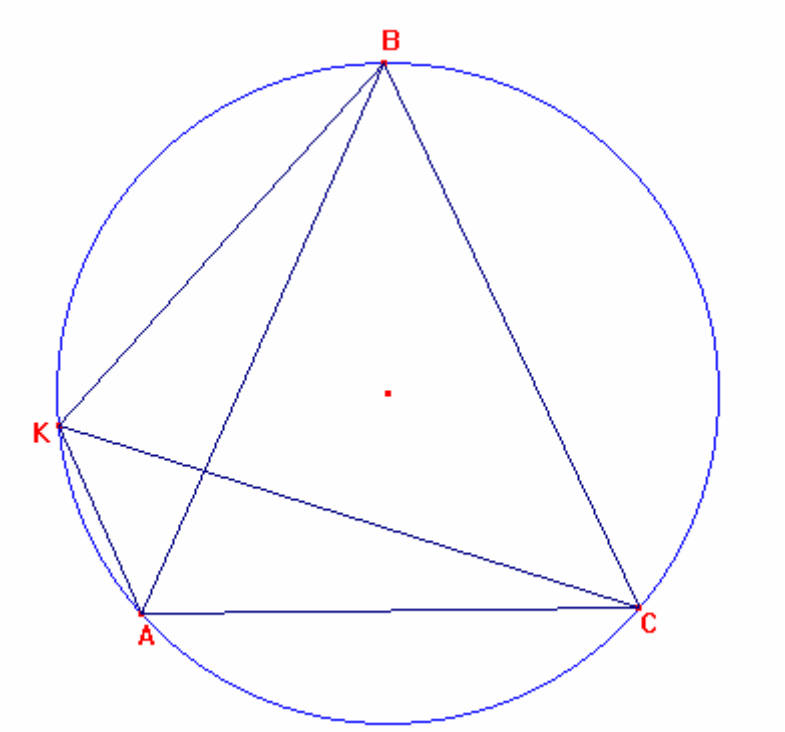
El segundo es más importante, recibe el nombre de Potencia de Punto. Consideremos una circunferencia de centro O y radio r. Sea P un punto del plano. Una recta que pasa por P interfecta a la circunferencia en los puntos A y B. Entonces el producto (PA)(PB) es constante, es decir, es independiente de la elección de la recta; si

P está afuera de la circunferencia la constante es igual a $(OP)^2 - r^2$; si P está dentro de ella la constante es ahora $r^2 - (OP)^2$.



Problema 7. En una circunferencia está inscrito el triángulo isósceles ABC (AB = BC). En el arco AB se toma al azar el punto K y se une por medio de cuerdas a los vértices del triángulo. Muestra que $(AK)(KC) = (AB)^2 - (KB)^2$.

Solución. La figura que ilustra el problema es la siguiente.

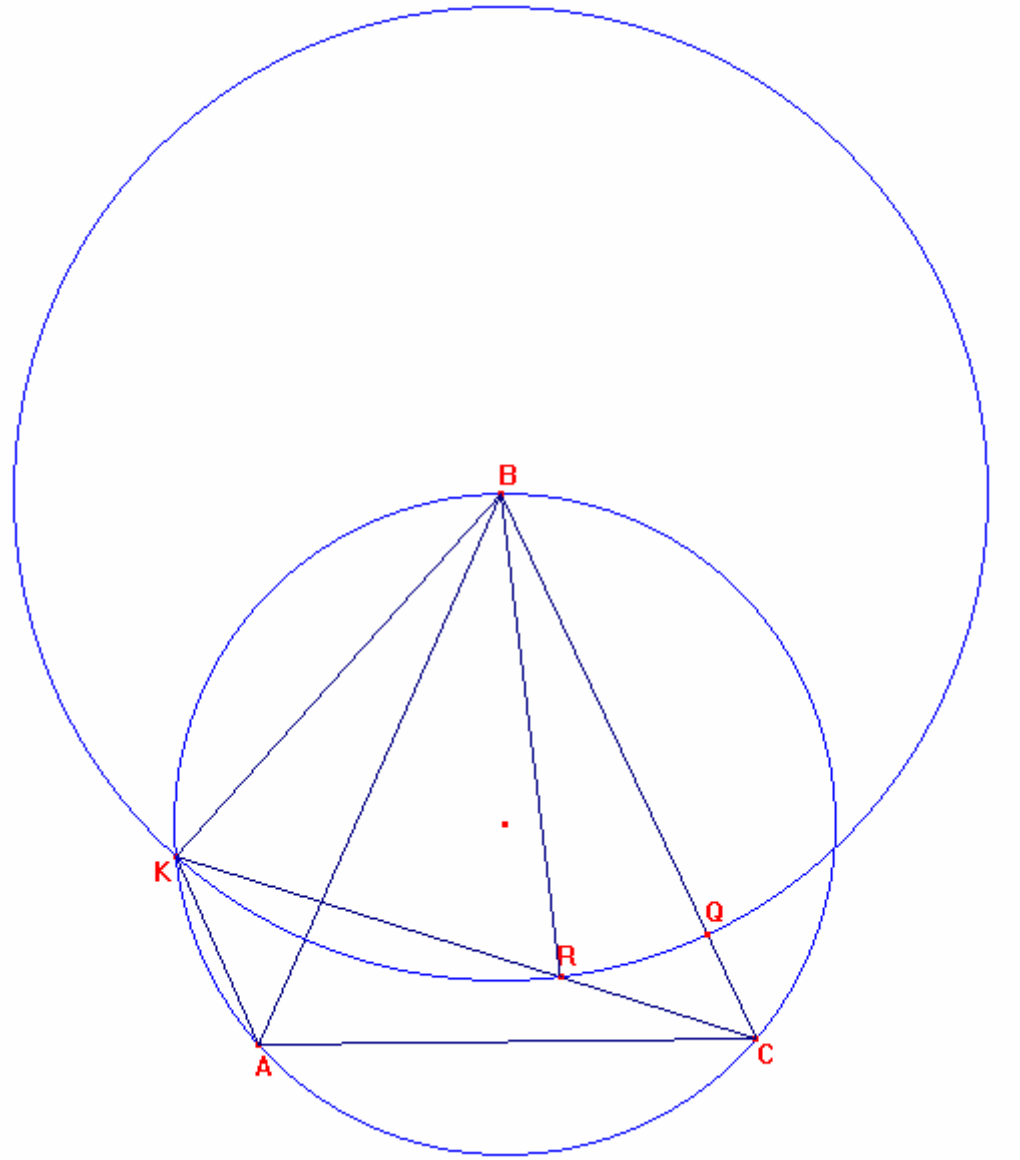


Ahora la eterna pregunta, ¿cómo comenzar a atacar el problema?

Si observamos bien lo que queremos demostrar, se parece mucho a lo que dice el teorema de potencia de punto, pero no lo es, al menos no a simple vista pues en el teorema de potencia de punto se tienen tres puntos colineales, lo que no tenemos aquí. Pero podemos hacer que se parezca aún más.

Uno de los dos, AB o KB debería funcionar como radio de una circunferencia para aplicar dicho teorema. Tomemos una circunferencia con centro en B y radio KB, la cual interfecta a BC y KC en los puntos Q y R respectivamente.

Entonces aplicando la potencia del punto C respecto a esa circunferencia tenemos que el producto (CR)(CK) es constante y puesto que C está fuera del círculo, la constante es $(BC)^2 - (KB)^2$. Pero BC = AB. Por lo tanto la igualdad queda, $(CR)(KC) = (AB)^2 - (KB)^2$. Que es casi lo que queremos mostrar, solo nos hace falta ver que CR = KA para tener la igualdad deseada.



Entonces el problema se reduce a demostrar una igualdad de segmentos la cual sabemos la estrategia a seguir, buscar triángulos congruentes. Se invita nuevamente al lector a encontrarlos.

Observemos los triángulos BKA y BRC. Se tiene que BR = BK por ser radios de la circunferencia. Además BC = AC. Entonces según el criterio LAL solo falta ver que $\angle KBA = \angle RBC$. Para ver esto, primero observemos que $\angle BAC = \angle BKC$ por ser ángulos inscritos que abren el arco BC, por esto y por ser ABC y BKR triángulos isósceles, se tiene que dichos triángulos son semejantes y por tanto $\angle KBR = \angle ABC$. De aquí que $\angle KBA = \angle KBR - \angle ABR = \angle ABC - \angle ABR = \angle RBC$. Concluimos que los triángulos BKA y BRC son congruentes, entonces sus tres lados son iguales. En particular AK = RC. Finalmente, sustituyendo AK por RC en la igualdad antes obtenida se tiene $(AK)(KC) = (AB)^2 - (KB)^2$. Lo cual termina la prueba.

Observaciones que el teorema de potencia de punto se utilizó sólo para reducir el problema a otro más sencillo, no para solucionarlo directamente. La manera en que a uno se le puede ocurrir utilizarlo se da después de experiencia en la solución de

problemas de manera que el teorema se vuelve de uno mismo, tenerlo de esa manera, no solo tenerlo de memoria, sino saberlo utilizar, sirve mucho.

También es importante el hecho de realizar una construcción extra al dibujo original, la cual debe ser adecuada para que nos proporcione más datos que sirvan a la solución de problema. El hacer trazos auxiliares que sí ayuden también forma parte del análisis que se hace al problema al comenzar a resolverlo. Las referencias presentadas proporcionan la teoría que uno puede aprender en olimpiada de matemáticas así como problemas para entrenar. En [3] se presentan los resultados básicos no sólo de geometría, también de combinatoria y teoría de números.

Referencias:

[1] Bulajich, M., Gómez, J. *Geometría.* Cuadernos de Olimpiadas de Matemáticas, Instituto de Matemáticas, UNAM, 2002.

[2] Bulajich, M; Gómez, J. *Geometría- Ejercicios y Problemas.* Cuadernos de Olimpiadas de Matemáticas, Instituto de Matemáticas, UNAM, 2002.

[3] Illanes, A. *Principios de Olimpiada.* Cuadernos de Olimpiadas de Matemáticas, Instituto de Matemáticas, UNAM, 2001.

ASP and Agents

Fernando Zacarías Flores

Benemérita Universidad Autónoma de Puebla
Department of computer science
Postal Code 72570, Puebla, México
fzflores@siu.buap.mx

Abstract. In this proposal, we present ASP and its applications in rational agents. We consider Answer Set Programming and rational agent paradigms. Considering these two paradigms, we can develop systems to approach in a direct way to rational behavior.

1 Introduction and Motivation

Nowadays, when we want to begin a serious research in the computer science, we should consider all the opinions that impact in a direct way in our research. Mainly if these opinions are make by connoted researchers. Also, is very important to consider as many current opinions as we can and classic opinions as well. In this sense we approach the logic firstly. Is logic a theory sufficiently robust to think that it can allow us model intelligence behavior? If we want to design an entity capable to be rational in some environment, then we need an unambiguous language, clear, simple. This language should allow us draw inferences, knowledge representation and updating knowledge and the behavior desired [BG94]. About 1960, McCarthy [McC59] proposed for the first time the use of logical formulas as a basis for a knowledge representation language of this type. This is how he explains the advantages of such representation:

"Expressing information in declarative sentences is far more modular than expressing it in segments of computer programs or in tables. Sentences can be true in a much wider context than specific programs can be used. The supplier of a fact does not have to understand much about how the receiver functions or how or whether the receiver will use it. The same fact can be used for many purposes, because the logical consequences of collections of facts can be available".

This idea has been the start point for many researchers, who have worked it strongly with various backgrounds and interests. Logic theory has demonstrated to be a paradigm with a huge level of abstraction to generalize from a problem domain to another [LAP01].

The remainder of the paper is structured as follows: In section 2 we briefly recap the basic background used throughout the paper. In section 3, we present our new proposal on modeling of agents. Next, we present our application based in rational agents presented in section 4. Finally, in section 5, we give our conclusions and future work.

2 Logic as Universal Language

Logic has demonstrated to be a paradigm with a huge level of abstraction to generalize from a dominion of problem to another [LAP01]. The classical logic of predicate calculus served as the main technical tool for the knowledge representation. Moreover, it's a good language of representation for static knowledge [LAP01], Also it has a well defined semantics with a good and power inference mechanism. However, there are paradigms more opened and dynamic. Then, it's necessary to consider languages or representation and integration of knowledge to advance along the time. Even, for the knowledge commonsense this tool is inadequate. For this reason, many researchers have focused their efforts to explore and develop new logical formalisms. Some of the most important proposals presented in [BG94] are circumscription [MD80, McC86, and Lif85b], default logic [Rei80b] and non monotonic modal logics [MD80, McD82 and Moo85].

We can stand out that the language of the logic is the language of the science. In fact, in mathematics is considered the universal language. We consider that the logic has been, it is and it will continue being our universal language. However, when we think in real applications we should also think in tools that allow us to apply the theory in the construction of them. In this sense Kowalski and Colmerauer created the logic programming [Llo87] and the development of the first logic programming language, Prolog [CKPR73].

3 Answer Set Programming

The stable model semantics is one of the most commonly accepted approaches to provide semantics to logic programs with NAF. Stable model semantics relies on the idea of accepting multiple minimal models as a description of the meaning of a program. In spite of this wide acceptance and its extensive mathematical foundations, stable model semantics have only recently found its way into mainstream "practical" logic programming. The recent successes have been a substantial effort towards understanding how to write programs under stable model semantics, recently referred as Answer Set Programming (ASP). ASP is a computation paradigm in which logical theories (Horn clauses with NAF) serve as problem specifications and solutions are represented by collection of model. ASP has been concretized in a number of related formalism - e.g., disjunctive logic programming and Datalog with constraints [ET00, and Eiter98].

Besides, it exist some researching groups in Europe: Pereira, Alferes, Eiter, Leite, and U.S.A Gelfond, Lifschitz, Przymusinski, Przymusinska, Baral, Dix that are developing theory and applications related to the semantic of the stable models [LAP01]. Nowadays, the semantic of stable models is called as answer sets programming (ASP), and using the concept of negation as failed, it lets us to solve problems with default knowledge.

Finally, we think that all the mentioned points in this article allow us approach to the software modeling with rational behavior in an efficient way. When we mention rational software, actually we are talking about software based on agents. In a concrete way, we wish to reach our investigation in the beliefs revision process, it means, the process that let an agent to manage its knowledge and beliefs.

For this reason, we think that logic is a good theory to model belief revision. The beliefs revision is a process that let the agents have behaviour closer to the human. Some researchers have even proposed to model agents based on mental state concept (Pereira), this represent the internal state of an intentional system. Generally, this build a set of beliefs, desires (goals), intentions, etc. that characterize the agent every instant.

4 Applications based on Agents

In this section we present different applications based in rational agents. These proposals integrate both ASP and Java, eliminating the traditional high gap between formal theory and practice. We use java as front-end, in which we develop the interfaz for user and the database administration. While in the declarative part, we use answer set programming through DLV. This paradigm gives the formal support to our agents.

4.1 Electronic board

Nowadays, the agent paradigm has recently increased its influence in the research and development of computational logic-based systems. However, the development of rational agents is a hard task, due to this involves both the updates and beliefs revision process. Furthermore, when we think in rational agents, we should consider some human characteristics like: reasoning, planning, and acting in a dynamic world. Our application [ZT2000, OZ2003] consists in a work environment (figure 1) that integrates necessary tools for correct performance of our scientists, incorporating a new formalization about update process. Our application consists of: an intelligent calendar that has as main ingredient the negotiation of meetings among multiple members of our scientific community via rational agents; a chat, that allows virtual meetings reducing the big distances; electronic mail and short messages via mobile telephone, used as our general communication channel in the negotiation meetings by

our agents; an rational editor, that counts with a rational agent that assists in papers writing and a knowledge base on the scientists' belief. In the figure 1, we present the main interfaz of our application. As we can observe one of the main tasks is the presentation of the schedule of each one of the researchers in our community. All this with the following purpose: that any member of the community can request a meeting with any of the researchers. The acceptance or not of a meeting depends on the beliefs that our agent has with respect to the required researcher. These knowledge bases are formed for beliefs, knowledge and preferences. Obviously, these knowledge bases change in a dynamic way according to the researcher's dynamic behavior.

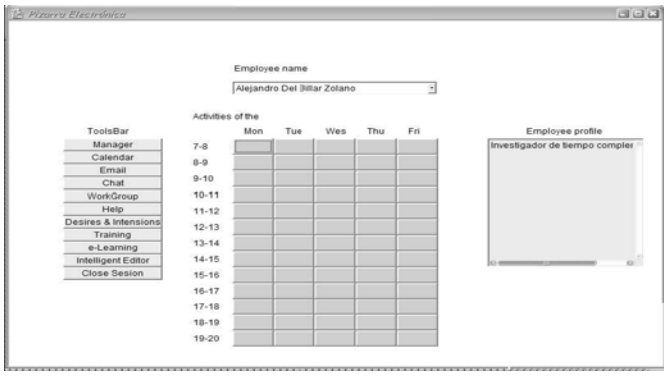


Figure 1. Electronic board

4.2 Financial Mobile System

The globalisation of the economy and the acceleration of technological change have made knowledge-based activities and the use of new technologies a priority issue for the growth of a country. Consequently, the new mobile technologies play a vital role in the correct acting of our activities. New information technology can help us achieve our perspectives of growth. So, we present another application based on agents and ASP paradigms called Financial Mobile System [ZSZMC06]. This system provides to the user services and applications for personal administration. Also, the system has an intelligent autonomous agent that carries out the search and purchase of those demanded products for its user in an autonomous way. Furthermore, we incorporate to our agent, supervision capacities in the control of the daily expense, these capacities allow the agent to learn of their user.

The information and communication technologies enter more facets of our lives. Society and the economy are adapting to the wave of innovation which is associated with this change resulting in new activities and new ways of doing things. In today's increasingly “knowledge-based” society the new raw material for economic activity and social interaction is information. The efficient use of information and communications technologies can provide a competitive edge to economic activity and have the potential to fundamentally transform society. The Information Society affects particular individuals, communities, sectors of economic activity, regions and even

countries in different ways. In this context, an innovative action is designed to investigate how to maximize the benefits which the Information Society can provide.

So, we describe a novel intelligent agent that allows to the user of mobile telephony to have a control in the personal expenses. This agent is a tool based on mobile technology in their users' benefit. Also, we have incorporate another intelligent agent whose capacity is to be able to carry out the purchases of those products that the user needs and that they are considered inside their budget, for that which makes use of a connection with “amazon.com” through internet. Additionally, we have incorporated our agent the capacity to learn of its user through a daily supervision.

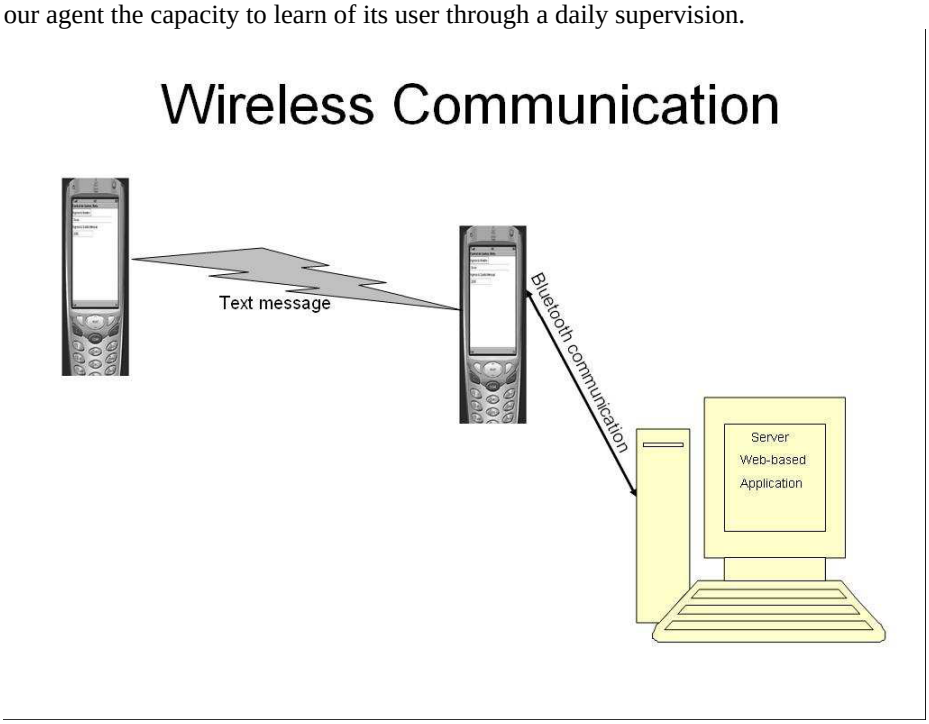


Figure 2. Architecture of Financial Mobile System

Our system makes the search and negotiation in internet of the product requested by user. The cellular phone (A) sends a message to another cellular telephone (Telephone B) located to a side of the applications server. Immediately, telephone B forwards the message to server through a port Bluetooth. Once the application located in the server receives the message, it begins the search and negotiation of the best offer.

4.3 Mobile Clinic

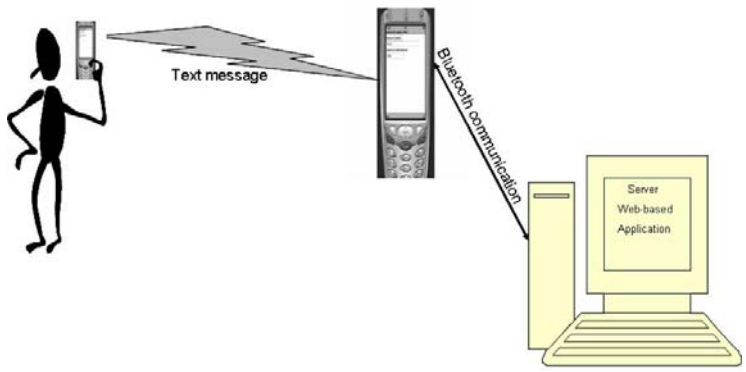


Figure 3. Mobile Clinic

Nowadays, the technological development grows to enlarged steps and the mobile technology is not the exception. Business users of today demand and home users of tomorrow will demand that they can continue their usual way of using computing and communication services wherever and whenever in an easy way. One of the technologies that have grown vertiginously in the last years are the cellular telephones. Moreover, cards PCMCIA has been developed for cellular phones so that laptops and palmtops can exploit connectivity provided by mobile phones. This type of developments faces us to the development of technologies of mobile software based on agents. In the same time agent technology has gained attention both among software houses and among research groups. However, agent technology is not yet mature enough for worldwide exploitation in telecommunication applications. There are still problems in designing software systems that integrate the concepts and functionality of personal mobility, and agent technology. So, the intelligent telephones have arrived at the market and they are here to remain for a lot of time. Moreover, if we consider that the own necessities of the mobile life and the digital world in which we live demand devices that are versatile, i.e., that in oneself tool multiple services can be had and therefore bigger benefits.

In this proposal, we propose to development an application based on agents called “Mobile clinic”. The main idea is, to take a picture of each one of the patient's irises with the camera of the cellular phone. Next to send these pictures through a multimedia message. Later, these pictures are processed using the irisdiologia paradigm. Finally, our system emits an diagnose clinical that is sent through a text message among cellular phones.

5 Conclusions and future work

We have presented several applications for mobile cellular telephone. These applications are based on mobile systems" and are based on intelligent agents. This it is a first step in the development of real and useful applications for the users of mobile telephone. Furthermore, these applications incorporate agents specialized in: e-commerce, learning about its user and control in the daily expenses. There is a lot of work to develop in this application, this is owed mainly to the limitations that a cellular telephone has still. However, the basic operation this fact and in a second version we will incorporate more services offered through of communication with a applications server.

With respect to future work, we will continue our research in mobile systems. In particular form, we will continue our analyses about e-commerce and e-business.

References

[BG94] Chitta Baral and Michael Gelfond . “Logic Programming and Knowledge Representation -- A-Prolog perspective”. *Journal of Logic Programming*, 19,20:73-148, 1994. (Survey paper).

[CKPR73] A. Colmerauer, H. Kanoui, R. Pasero, and P. Roussel. Un systeme de Communication Homme-Machine en Français. Technical report, Groupe de Intelligence Artificielle Universitae de Aix-Marseille II, Marseille, 1973.

[EFLP99] Thomas Eiter, Wolfgang Faber, Nicola Leone, and Gerald Pfeifer. The Diagnosis Frontend of the `dlv` System. *AI Communications -- The European Journal on Artificial Intelligence*, 12(1-2):99-111, 1999.

[Eiter98] T. Eiter et al. The KR System dlv : Progress Report, Comparisons, and Benchmarks. In International Conference on Principles of Knowledge Representation and Reasoning, 1998.

[ET00] D.East and M.Truszczynski. Datalog with Constraints. In National Conference on Artificial Intelligence, pages 163-168. AAAI/MIT Press, 2000.

[LAP01] J. A. Leite, J. J. Alferes and L. M. Pereira, MINERVA -Combining Societal Agents Knowledge. *International Workshop on Agent Theories, Architectures, and Languages (ATAL'01), pages 133-145*, Seattle, USA, August 2001

[Lif85b] V. Lifschitz. Computing circumscription. In Proc. Of IJCAI-85, pages 121-127, 1985.

[Llo87] J. Lloyd. “Foundations of Logic Programming”. Springer-Verlag, second edition, 1987.

[McC59] J. MacCarthy. Programs with common sense. In Proc. Of the Teddington Conference on the Mechanization of Thought Proccesses, pages 75-91, London, 1959. Her Majesty's Stationery Office.

[McD82] D. McDermontt. Nonmonotonic logic II: Nonmonotonic modal theories. Journal of the ACM., 29(1):33-57, 1982.

[MD80] D. McDermott and J. Doyle. Nonmonotonic logic I. Artificial Intelligence, 13(1, 2):41=72, 1980.

[Moo85] R. Moore. Semantical considerations on Nonmonotonic Logic. Artificial Intelligence, 25(1):75-94, 1985.

[Rei80b] R.Reiter. A logic for default reasoning. Artificial Intelligence, 13(1, 2):81-132, 1980.

[ZSZMC06] F. Zacarias, A. Sanchez, D. Zacarias, A. Méndez and R. Cuapa. Financial Mobile System, Austrian Computer Society book series, Austria 2006

[ZT2000] F. Zacarias and F. Tobon. Pizarra electronica basada en agents. Jornadas Chilenas de Computación, Copiapó, Chile, 2000

[OZ2003] M. Osorio and F. Zacarias, Irrelevance of Syntax in updating answer set programs, Proceedings Of Fourth Mexican International Conference On Computer Science Enc'03, pp.183-188, Eds. J. H. Sossa, and E. Perez, México, 2003.

RESUMEN SOBRE LÓGICA POSIBILISTA

JOSÉ ARRAZOLA
IVAN CORTÉS
JESÚS LAVALLE

Resumen. Este trabajo pretende más que introducir al lector en el conocimiento sobre Lógica Posibilista, invitarlo a la lectura de esta intesante temática.

1. Introducción

En la vida real existen situaciones en donde sólamente se cuenta con información incompleta o parcialemente inconsistente. La Lógica Posibilista ofrece una herramienta para hacer inferencias con este tipo de información.

El trabajo con la incertidumbre en un contexto lógico no es un tema nuevo; Uno de los intentos por tratar con información incompleta utiliza el cálculo de probabilidades, pero éste presenta algunas dificultades al tratar de hacer un análisis en un contexto lógico:

- Un conjunto de proposiciones no es cerrado deductivamente, esto es, a partir de las restricciones $P(\neg p \vee q) \geq \alpha$ y $P(p) \geq \alpha$, sólamente se puede deducir que $P(q) \geq \max\{0, 2\alpha - 1\}$, en donde $\max\{0, 2\alpha - 1\} < \alpha$.
- Exite una gran discrepancia entre $P(\neg p \vee q)$ y $P(p|q)$ (la disyunción exclusiva), lo cual lleva a la pregunta sobre el correcto modelado de una regla del tipo $p \rightarrow q$.

En 1976 Rescher propuso que la *fuerza* de una deducción es igual a la *fuerza* del argumento más debil usado en su deducción; La contribución de la Lógica Posibilista es relacionar ésta idea con las medidas de necesidad difusas, dentro de la teoría de Zadeh, ya que se cumple:

$$N(\neg p \vee q) \geq \alpha \text{ y } N(p) \geq \alpha, \text{ implica } N(q) \geq \min\{\alpha, \beta\}$$

La Lógica Posibilista puede ser utilizada para modelar conocimiento lleno de incertidumbre, cuando éste se representa en el contexto de la teoría de la posibilidad.

En [11], Zadeh introdujo las medidas de posibilidad como un índice escalar que evalua la consistencia de una proposición difusa con respecto al estado del conocimiento expresado por medio de una *restricción difusa*. Una restricción difusa es un conjunto difuso de valores *posibles* y su función de membresía se llama distribución de posibilidad. Fue entonces cuando se hizo patente que la "Lógica Difusa" de Zadeh no era otra lógica multivaluada, sino una forma de razonar bajo incertidumbre o conocimiento incompleto descrito por restricciones difusas (lo que Zadeh llamó Razonamiento *aproximado*).

En [1] se incluye una axiomatización y un método de refutación basado en *resolución extendida* que es viable de ser implementado en una computadora y que soporta inconsistencia parcial.

La lógica posibilista está estrechamente relacionadad con la teoría *belief revision*.

En [7] se presentan algunos algoritmos para el *problema de la deducción* en la Lógica Posibilista Estándar , así como algunos algoritmos para encontrar *modelos* en Lógica Posibilista Estándar .

Entre las aplicaciones de la lógica posibilística se encuentran:

(1) Razonamiento No-Monótono [3, 5], razonamiento utilizando *reglas con negación por falla*[8].
(2) Belief revision [4].
(3) Se puede usar la Lógica Posibilista para crear la teoría *Possibilistic Logic Programming*, la cual es particularmente útil cuando se trata con incertidumbre o con optimización min-max. Los detalles formales sobre la semántica declarativa y *procedural* de los programas lógicos posibilistas se pueden encontrar en [6] y algunas extensiones que incorporan la negación por falla se pueden encontrar en los trabajos de Wagner [10].

En [10] se demuestra que los programas lógicos normales, bajo la semántica de modelos estables de Gelfond y Lifschitz, se pueden *sumergir* en programas lógicos difusos (bajo su semántica estable) y los programas lógicos extendidos, bajo la semántica de Answer Set de Gelfond y Lifschitz, se pueden sumergir en programas lógicos posibilistas, bajo la semántica estable posibilista.

En [9] se muestra que los conceptos de Answet Set Programming y Lógica Difusa se pueden combinar en un sólo esquema llamado *Fuzzy Answer Set Programming (FASP)*. Su propuesta muestra que este enfoque es una extensión de la teoría de Answer Set Programming tradicional, a diferencia de otros enfoques.

En la Lógica Posibilista Completa, el conocimiento incierto se expresa en términos de oraciones *certainty-qualified* o *possibility-qualified*. La Lógica Posibilista Completa trata con objetos sintácticos que expresan desigualdades que resultan de éste tipo de oraciones. En lo que sigue hablaremos sobre la llamada Lógica Posibilista Estándar la cual es un fragmento de la Lógica Posibilista Completa, que sólo trata con oraciones *certainty-qualified*.

2. Lo básico de Lógica Posibilista

Una fórmula posibilista estándar es un par $(\varphi\ \alpha)$, donde φ es una fórmula proposicional clásica y $\alpha \in (0, 1]$. $(\varphi\ \alpha)$ expresa que φ es cierta al menos hasta un grado α, esto es, $N(\varphi) \geq \alpha$, donde N es una *medida de necesidad* que modela el estado de conocimiento. Al escalar α se conoce como la *valuación* de la fórmula y se denota como $val(\varphi)$.

Una base de conocimiento *necessity-valued* (o tambien llamada base de conocimiento posibilista estándar) $\mathcal{F}$ se define entonces como un conjunto finito de fórmulas necessity-valued. $\mathcal{F}^*$ denota el conjunto de fórmulas clásicas que se obtiene de

$\mathcal{F}$ al ignorar los *pesos*, esto es, si $\mathcal{F} = \{(\varphi_i\ \alpha_i)\ :\ i = 1, \ldots, n\}$ entonces, $\mathcal{F}^* = \{\varphi_i\ :\ i = 1, \ldots, n\}$. A $\mathcal{F}^*$ se le llama la *proyección clásica* de $\mathcal{F}$.

Una base de conocimiento posibilista estándar también puede ser vista como una colección anidada de fórmulas clásicas: si α es cualquier valuación, definimos el $\alpha-$corte $\mathcal{F}_\alpha$ y el $\alpha-$corte estricto $\mathcal{F}_{\overline{\alpha}}$ como:

$$\begin{aligned} \mathcal{F}_\alpha &= \{(\varphi\ \beta) \in \mathcal{F}\ :\ \beta \geq \alpha\} \\ \mathcal{F}_{\overline{\alpha}} &= \{(\varphi\ \beta) \in \mathcal{F}\ :\ \beta > \alpha\} \end{aligned}$$

Sus proyecciones clásicas son

$$\begin{aligned} \mathcal{F}^*_\alpha &= \{\varphi\ :\ (\varphi\ \beta) \in \mathcal{F},\ \beta \geq \alpha\} \\ \mathcal{F}^*_{\overline{\alpha}} &= \{\varphi\ :\ (\varphi\ \beta) \in \mathcal{F},\ \beta > \alpha\} \end{aligned}$$

En la Lógica Posibilista Estándar los conceptos de *satisfacción* y *consecuencia lógica* estan definidos en términos de distribuciones de posibilidad sobre el conjunto de "mundos clásicos". Una distribución de posibilidad π es una función de Ω (el conjunto de todos los mundos posibles) a $[0, 1]$. $\pi(\omega)$ refleja que tan posible es que ω sea el "mundo real". Cuando $\pi(\omega) = 1$ (resp. $\pi(\omega) = 0$) entonces es completamente posible (resp. completamente imposible) que ω sea el mundo real. Una distribución de posibilidad está *normalizada* si, y sólo si $\exists\omega$ tal que $\pi(\omega) = 1$.

La *medida de posibilidad* Π inducida por π es una función de $\mathcal{L}$ (un lenguage lógico de primer orden o proposicional) a $[0, 1]$ definida por

$$\Pi(\varphi) = \sup\{\pi(\omega)\ :\ \omega \models \varphi\}$$

La *medida de necesidad* (dual) N inducida por π se define como

$$N(\varphi) = 1 - \Pi(\neg\varphi) = \inf\{1 - \pi(\omega)\ :\ \omega \models \neg\varphi\}$$

Se cumplen las siguientes propiedades:

$$\begin{array}{rcl} N(\top) & = & 1 \\ N(\varphi \wedge \psi) & = & \min\{N(\varphi), N(\psi)\} \\ N(\varphi \vee \psi) & \geq & \max\{N(\varphi), N(\psi)\} \\ \text{Si } \varphi \models \psi \quad \text{entonces} & & N(\psi) \geq N(\varphi) \end{array}$$

Decimos que una distribución de posibilidad π satisface a la fórmula posibilista estándar $(\varphi\ \alpha)$ si, y sólo si, $N(\varphi) \geq \alpha$, donde N es la medida de necesidad inducida por π. Usaremos la notación $\pi \models (\varphi\ \alpha)$. Una distribución de posibilidad π satisface una base de conocimiento posibilista estándar $\mathcal{F} = \{(\varphi_i\ \alpha_i) \mid i = 1, \ldots, n\}$ si, y sólo si, para toda i, $\pi \models (\varphi_i\ \alpha_i)$. Esto lo denotaremos por $\pi \models \mathcal{F}$.

Una fórmula posibilista estándar $(\varphi\ \alpha)$ es una *consecuencia lógica* de una base de conocimiento posibilista estándar $\mathcal{F}$ si, y sólo si, para cualquier π que satisfaga a $\mathcal{F}$, se tiene que también π satisface a $(\varphi\ \alpha)$.

Tenemos el siguiente problema de deducción: Sea $\mathcal{F}$ una base de conocimiento posibilista estándar y sea φ una fórmula clásica que queremos deducir de $\mathcal{F}$ hasta cierto grado; tenemos que calcular la valuación más alta α (esto es, la mejor cota inferior de una medida de necesidad) de manera que $(\varphi\ \alpha)$ sea una consecuencia lógica de $\mathcal{F}$, es decir, debemos calcular

$$Val(\varphi, \mathcal{F}) = \sup\{\alpha \in (0, 1]\ :\ \mathcal{F} \models (\varphi\ \alpha)\}$$

Un resultado fundamental de la deducción a partir de bases de conociemiento posibilista estándar es que siempre existe una distribución de posibilidad menos *específica*[1] que satisfaga una base de conocimiento posibilista estándar $\mathcal{F}$. A saber, si $\mathcal{F} = \{(\varphi_i\ \alpha_i)\ :\ i = 1, \ldots, n\}$ entonces la distribución de posibilidad menos específica $\pi_{\mathcal{F}}$ que satisface $\mathcal{F}$ se define como

$$\pi_{\mathcal{F}}(\omega) = \begin{cases} 1 \text{ si } \omega \models \varphi_1 \wedge \varphi_2 \wedge \cdots \wedge \varphi_n \\ \min\{1 - \alpha_i\ :\ \omega \models \neg\varphi,\ i = 1, \ldots, n\} \text{ otro caso} \end{cases}$$

2.1. **Proposition.** [1] Para cualquier distribución de posibilidad π, π satisface a $\mathcal{F}$ si, y sólo si, $\pi \leq \pi_{\mathcal{F}}$

Como consecuencia tenemos el siguiente corolario,

2.2. **Corollary.** [1]

$$\mathcal{F} \models (\varphi\ \alpha) \text{ si, y sólo si, } \pi_{\mathcal{F}} \models (\varphi\ \alpha)$$

O en otros términos, $Val(\varphi, \mathcal{F}) = N_{\mathcal{F}}(\varphi)$, donde $N_{\mathcal{F}}$ es la medida de necesidad inducida por $\pi_{\mathcal{F}}$.

Lo anteriormente expuesto nos da una idea somera de lo que trata y de que manera la Lógica Posibilista, reiteramos que con ello pretendemos hacer una invitación a los lectores para que investiguen respecto al tema.

Referencias

[1] Didier Dubois-Jérôme Lang-Henri Prade, *Possibilistic Logic*, Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3, 439-513, Oxford University Press.

[2] Didier Dubois-Henri Prade, *Possibilistic Logic: a retrospective and prospective view*, Fuzzy Sets and Systems 144: 3-23,2004.

[3] Didier Dubois-Henri Prade, *Possibilistic Logic, preferential models, nonmonotonicity and related issues*. Proc. of IJCAI'91, 419-424.

[4] Didier Dubois-Henri Prade, *Belief Change and Possibility Theory*. In P. Gärdenfors, ed., Belief Revision, 142-182, Cambridge University Press, 1992.

[5] Didier Dubois-Jérôme Lang-Henri Prade, *Automated Reasoning Using Possiblisitic Logic: Semantics, Belief Revision, and Variable Certainty Weights*. IEEE Trans. on Data and Knowledge Engineering, 1994.

[6] Didier Dubois-Jérôme Lang-Henri Prade, *Towards Possibilistis Logic Programming*. Proc. of ICLP'91, 581-595.

[7] Jérôme Lang, *Possibilistic Logic: complexity and algorithms*, in: J. Kholas, S. Moral (Eds.), Algorithms for Uncertainty and Defeasible Reasoning, Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 5, Kluwer Academic Publishers, Dordrecht, 2001, pp. 179-220.

[8] Salem Benferhat-Didier Dubois-Henri Prade, *Default Rules and Possibilistic Logic*. Proceedings of KR'92, 673-684.

[9] Davy Van Nieuwenborgh, *Fuzzy Answer Set Programming*,(**impreso**)

[10] Gerd Wagner, *Negation in Fuzzy and Possibilistic Logic Programs*, (**impreso**)

[11] Zadeh L.A., *Fuzzy Sets as a Basis for a Theory of Possibility*, Fuzzy Sets and systems, 1(1),3-28, 1978.

arrazola@fcfm.buap.mx
icoc526328@mail.cs.buap.mx
jlavallenator@gmail.com

[1]Decimos que la distribución π es más específica que la distribución π' si, y sólo si, $\pi < \pi'$

MODELOS PSEUDO P-ESTABLES

JOSÉ ARRAZOLA
JESÚS LAVALLE
FELIPE MAZÓN

Resumen. Una lógica de tres valores llamada G'_3 ha sido introducida recientemente para definir una nueva semántica para el razonamiento no monótono[2]. Esta lógica ha sido axiomátizada en [3]. Se presenta las nociones de modelo *pseudo-pestable* y *grado de contradicción* para generalizar la de modelo p-estable cuando se presenta información contradictoria. El resultado principal de este trabajo es: Sea P es un programa normal entonces, M es un modelo p-estable de P sii M es un modelo pseudo p-estable de P.

1. La Lógica G'_3

Una lógica de tres valores llamada G'_3 ha sido introducida recientemente para definir una nueva semántica para el razonamiento no monótono[2]. En este trabajo sólo se hará mensión de los resultados necesarios y se remite al lector al artículo original para sus demostraciones.

2. Definiendo G'_3

La lógica G'_3 es una lógica de tres valores con valores de verdad 0, 1 y 2 donde 2 es el único valor designado. Definimos las tablas de verdad para los conectivos $\rightarrow$ y $\neg$ de la lógica G'_3 en el Cuadro 1.

Cuadro 1. Tablas de verdad para los conectivos en G'_3

x	$\neg x$
0	2
1	2
2	0

$\rightarrow$	0	1	2
0	2	2	2
1	0	2	2
2	0	1	2

La conjunción se define como la función mín. En [1], G'_3 se introduce sólo para probar que $a \vee (a \rightarrow b)$ no es un teorema de C_ω.

Observese que esta lógica es paraconsistente. Basta notar que la fórmula $a \wedge \neg a \rightarrow b$ no es un teorema de G'_3 y por tanto no se trivializa ante la presencia de fórmulas contradictorias.

3. Axiomatización of G'_3

En [3], se presenta una axiomatización tipo Hilbert de G'_3. Esta lógica tiene tres conectivos lógicos primitivos, a saber $\rightarrow$, $\wedge$, y $\neg$. También tiene otros conectivos que se definen como:

1. $a \vee b := ((a \rightarrow b) \rightarrow b) \wedge ((b \rightarrow a) \rightarrow a)$.

2. $\sim a := a \rightarrow (\neg a \wedge \neg\neg a)$.

A partir de ahora, el símbolo $\vdash$ denotará $\vdash_{G'_3}$, a menos que se diga otra cosa. La lógica G'_3 tiene todos los axiomas de la lógica C_ω más los siguientes axiomas:

E1 $\neg\neg\mathscr{B} \rightarrow \sim \neg\mathscr{B}$
E2 $\neg\mathscr{B} \leftrightarrow \neg\neg\neg\mathscr{B}$
E3 $\sim \mathscr{B} \rightarrow \neg\mathscr{B}$
E4 $(\neg\neg\mathscr{B} \wedge \neg\neg\mathscr{C}) \leftrightarrow \neg\neg(\mathscr{B} \wedge \mathscr{C})$
E5 $(\neg\mathscr{B} \wedge \neg\mathscr{C}) \rightarrow \neg(\mathscr{B} \wedge \mathscr{C})$
E6 $(\sim\sim \mathscr{B} \wedge \neg\mathscr{B}) \rightarrow (\neg\neg\mathscr{C} \rightarrow (\sim\sim (\mathscr{B} \wedge \mathscr{C}) \wedge \neg(\mathscr{B} \wedge \mathscr{C})))$
E7 $\sim \mathscr{B} \rightarrow \neg\neg(\mathscr{B} \rightarrow \mathscr{C})$
E8 $\neg\neg(\mathscr{B} \rightarrow \mathscr{C}) \leftrightarrow ((\mathscr{B} \rightarrow \mathscr{C}) \wedge (\neg\neg\mathscr{B} \rightarrow \neg\neg\mathscr{C}))$
E9 $(\sim\sim \mathscr{B} \wedge \neg\mathscr{B}) \rightarrow ((\sim\sim \mathscr{C} \wedge \neg\mathscr{C}) \rightarrow \neg\neg(\mathscr{B} \rightarrow \mathscr{C}))$
E10 $\neg\neg\mathscr{B} \rightarrow ((\sim\sim \mathscr{C} \wedge \neg\mathscr{C}) \rightarrow (\sim\sim (\mathscr{B} \rightarrow \mathscr{C}) \wedge \neg(\mathscr{B} \wedge \mathscr{C})))$

Un resultado útil, es el siguiente.

Teorema 3.1. *Sea Γ y Δ dos conjuntos de fórmulas. Sea A, A_1, A_2, B, y C fórmulas arbitrarias. Entonces se cumplen las siguientes propiedades:*

1. $\Gamma \vdash B$ *implica* $\Gamma \cup \Delta \vdash B$
2. $\Gamma, A \vdash B$ *si, y sólo si,* $\Gamma \vdash A \rightarrow B$
3. $\Gamma \vdash A_1 \wedge A_2$ *si, y sólo si,* $\Gamma \vdash A_1$ *y* $\Gamma \vdash A_2$
4. $\Gamma, A \vdash B$ *y* $\Gamma, \neg A \vdash B$ *si, y sólo si,* $\Gamma \vdash B$
5. $\Gamma \vdash B$ *y* $\Delta, B \vdash C$ *entonces* $\Gamma \cup \Delta \vdash C$

En [3] se demuestra el siguiente Teorema de Robustéz y Completéz:

Teorema 3.2. *Una fórmula φ es un teorema en G'_3 si, y sólo si, es una G'_3-tautología.*

Es importante observar que esta lógica es diferente de otras lógicas paraconsistentes, por ejemplo de la definida por da Costa: C_1. Esto se desprende del hecho de que el esquema de axioma $\neg(\alpha \wedge \neg\alpha)$ no es válido en C_1 , pero es un teorema en G'_3. Más aún, la Ley de Pierce es válida en C_1 pero no en G'_3. También, se ha visto en [2] que G'_3 es diferente de Pac. Finalmente, G'_3 también es diferente de la lógica paraconsistente de cuatro valores introducida por Arnon Avron.

4. Programas Lógicos

La idea de la programación Lógica es utilizar a la lógica como un lenguaje de programación[4].

Definición 4.1. En nuestro contexto un *programa lógico* es únicamente una teoría y una *clase de programas lógicos* es un conjunto de programas lógicos.

De hecho podemos pensar que las palabras teoría y programa son sinónimos, usualmente empleamos la primera cuando estamos en el contexto de lógica y la segunda cuando se trata de programación.
La sintaxis de las fórmulas usualmente se define en términos de algunas fórmulas especiales conocidas como *cláusulas*.

Definición 4.2. Una *cláusula* es una fórmula de la forma $\mathscr{H} \leftarrow \mathscr{B}$ donde la implicación es el conectivo principal y $\mathscr{H}$, $\mathscr{B}$ representan fórmulas que tiene sólamente los conectivos $\neg, \wedge, \vee$.

- Las fórmulas $\mathscr{H}$ y $\mathscr{B}$ se conocen como *cabeza* y *cuerpo* de la cláusula respectivamente.
- La cláusula $\bot \leftarrow \mathscr{B}$ se le llama *constraint* y se dice que la cabeza está vacía.
- Una cláusula $\mathscr{H} \leftarrow \bot$, se llamará un *hecho*. Por simplicidad se escribirá $\mathscr{H}$.

En la literatura referente a la programación lógica podemos encontrar diferentes tipos de cláusulas:

1. Una cláusula es una *cláusula definite* si $\mathscr{H} = \{a\}$ y $\mathscr{B} = \mathscr{B}^+ = \bigwedge_{i=1}^n b_i$, donde a, b_i son átomos.
2. Una cláusula es una *cláusula normal* si $\mathscr{H} = \{a\}$ y $\mathscr{B} = \mathscr{B}^+ \wedge \mathscr{B}^-$, donde $\mathscr{B}^+ = \bigwedge_{i=1}^n b_i$ y $\mathscr{B}^- = \bigwedge_{k=1}^n \neg d_k$, con a, b_i, d_k átomos.
3. Una cláusula es una *cláusula disyuntiva* si $\mathscr{H} = \bigvee_{j=1}^m c_j$ y $\mathscr{B} = \mathscr{B}^+ \wedge \mathscr{B}^-$, donde $\mathscr{B}^+ = \bigwedge_{i=1}^n b_i$ y $\mathscr{B}^- = \bigwedge_{k=1}^n \neg d_k$, con a, b_i, d_k átomos..

Así, una programa definite (normal, disyuntivo) es un conjunto de cláusuluas definite (normales, disyuntivas).

5. Semántica Estable

En el corazón de la semántica estable se encuentra el concepto de *reducto* de un programa P con respecto a un conjunto de átomos, el cual se define a continuación.

Definición 5.1. Dado un programa lógico P y un conjunto de átomos M definimos el reducto de P con respecto a M como el programa lógico

$$P^M = \left\{ \mathscr{H} \leftarrow \mathscr{B}^+ \mid \mathscr{H} \leftarrow \mathscr{B}^+ \wedge \neg \mathscr{B}^- \in P, \mathscr{B}^- \cap M = \emptyset \right\}$$

Definición 5.2. Sea P un programa normal y M un conjunto de átomos. Decimos que M es un *modelo estable* de P si, y sólo si, M es el modelo mínimo de P^M.

6. Semántica P-estable

Ahora tenemos la siguiente definición.

Definición 6.1. Sea P un programa normal, sea M un conjunto de átomos. Decimos que M es un modelo p-estable de P si $RED(P, M) \Vdash_C M$, es decir, si M es un modelo clásico de $RED(P, M)$ y $RED(P, M) \vdash_C M$.

Definición 6.2. Sea P un programa normal y sea M un conjunto de átomos. Definimos

$$RED(P, M) = \left\{ a \leftarrow \mathscr{B}^+ \wedge \neg(\mathscr{B}^- \cap M) \mid a \leftarrow \mathscr{B}^+ \wedge \neg \mathscr{B}^- \in P \right\}$$

Los tres lemas siguientes (6.3, 6.4 y 6.5) pueden encontrarse en [2].

Lema 6.3. *Sea P un programa normal y sea a un átomo. Sean X, Y dos lógicas tales que $C_\omega \subseteq X, Y \subseteq C$. Entonces, $P \vdash_X a$ si, y sólo si, $P \vdash_Y a$.*

Lema 6.4. *Sea P un programa normal y $M \subseteq \mathscr{L}_P$. Entonces se tiene que: $RED(P, M) \vdash_{G'_3} M$ si, y sólo si, $P \cup \neg\widetilde{M} \vdash_{G'_3} M$.*

Lema 6.5. *Sea P un programa normal y sea $M \subseteq \mathscr{L}_P$. Entonces se tiene que M es un modelo clásico de $P \cup \neg\widetilde{M}$ si, y sólo si, M es un modelo clásico de $RED(P, M)$.*

7. Modelos Pseudo p-estables

Dado que en G_3' la fórmula $\neg\mathscr{A} \wedge \mathscr{A}$ no es una fórmula contradictoria, se tiene que no se puede deducir cualquier cosa a partir de ella. Con esto en mente, se definirá la nocion de *grado de contradicción* de un programa.

Definición 7.1. Decimos que un programa lógico P es *contradictorio* respecto a $\neg$ si existe una fórmula φ tal que $P \vdash_{G_3'} \varphi$ y $P \vdash_{G_3'} \neg\varphi$.

Definición 7.2. Sea P un programa lógico. Definimos el *grado de contradicción* de P, denotado por $\mathrm{G}(P)$, como

$$\mathrm{G}(P) = \mathrm{Card}\left\{a \in \mathscr{L}_P \mid P \vdash_{G_3'} a \ \text{ y } \ P \vdash_{G_3'} \neg a\right\}$$

Proposición 7.3. *Sea P un programa lógico. Si P tiene un modelo clásico, entonces $G(P) = 0$.*

Demostración. Sea M un modelo clásico de P y supongamos que $G(P) > 0$. Como P tiene un modelo clásico, entonces es consistente (en el sentido clásico). Como $G(P) > 0$, entonces existe un átomo a tal que $P \vdash_{G_3'} a$ y $P \vdash_{G_3'} \neg a$. Ahora bien, por el Lema 6.3 se tiene que $P \vdash_C a$ y $P \vdash_C \neg a$, violando la consistencia (clásica) de P. □

Proposición 7.4. *Sean P_1 y P_2 programas normales, $M_1 \subseteq \mathscr{L}_{P_1}$ y $M_2 \subseteq \mathscr{L}_{P_2}$ con $\mathscr{L}_{P_1} \cap \mathscr{L}_{P_2} = \emptyset$, entonces*

$$RED(P_1 \cup P_2, M_1 \cup M_2) = RED(P_1, M_1) \cup RED(P_2, M_2)$$

Demostración. Sea $r \in RED(P_1, M_1) \cup RED(P_2, M_2)$. Entonces $r = \alpha_1 \leftarrow \mathscr{B}_1^+ \wedge \neg(\mathscr{B}_1^- \cap M_1)$ ó $r = \alpha_2 \leftarrow \mathscr{B}_2^+ \wedge \neg(\mathscr{B}_2^- \cap M_2)$. Como $\mathscr{L}_{P_1} \cap \mathscr{L}_{P_2} = \emptyset$, entonces $\mathscr{B}_1^- \cap M_2 = \emptyset$ y, $\mathscr{B}_2^- \cap M_1 = \emptyset$.

Así, si $r = \alpha_1 \leftarrow \mathscr{B}_1^+ \wedge \neg(\mathscr{B}_1^- \cap M_1)$, con $\alpha_1, \mathscr{B}_1^+, \mathscr{B}_1^-, M_1 \subseteq \mathscr{L}_{P_1}$, se tiene que

$$\begin{aligned} r &= \alpha_1 \leftarrow \mathscr{B}_1^+ \wedge \neg(\mathscr{B}_1^- \cap M_1) \\ &= \alpha_1 \leftarrow \mathscr{B}_1^+ \wedge \neg(\mathscr{B}_1^- \cap M_1) \wedge \top \\ &= \alpha_1 \leftarrow \mathscr{B}_1^+ \wedge \neg(\mathscr{B}_1^- \cap M_1) \wedge \neg(\emptyset) \\ &= \alpha_1 \leftarrow \mathscr{B}_1^+ \wedge \neg(\mathscr{B}_1^- \cap M_1) \wedge \neg(\mathscr{B}_1^- \cap M_2) \\ &= \alpha_1 \leftarrow \mathscr{B}_1^+ \wedge \neg((\mathscr{B}_1^- \cap M_1) \cup (\mathscr{B}_1^- \cap M_2)) \\ &= \alpha_1 \leftarrow \mathscr{B}_1^+ \wedge \neg(\mathscr{B}_1^- \cap (M_1 \cup M_2)) \end{aligned}$$

Y así, $r \in RED(P_1 \cup P_2, M_1 \cup M_2)$.

Análogamente, si $r = \alpha_2 \leftarrow \mathscr{B}_2^+ \wedge \neg(\mathscr{B}_2^- \cap M_2)$, con $\alpha_2, \mathscr{B}_2^+, \mathscr{B}_2^-, M_2 \subseteq \mathscr{L}_{P_2}$, se tiene que

$$\begin{aligned} r &= \alpha_2 \leftarrow \mathscr{B}_2^+ \wedge \neg(\mathscr{B}_2^- \cap M_2) \\ &= \alpha_2 \leftarrow \mathscr{B}_2^+ \wedge \neg(\mathscr{B}_2^- \cap M_2) \wedge \top \\ &= \alpha_2 \leftarrow \mathscr{B}_2^+ \wedge \neg(\mathscr{B}_2^- \cap M_2) \wedge \neg(\emptyset) \\ &= \alpha_2 \leftarrow \mathscr{B}_2^+ \wedge \neg(\mathscr{B}_2^- \cap M_2) \wedge \neg(\mathscr{B}_2^- \cap M_1) \\ &= \alpha_2 \leftarrow \mathscr{B}_2^+ \wedge \neg((\mathscr{B}_2^- \cap M_2) \cup (\mathscr{B}_2^- \cap M_1)) \\ &= \alpha_2 \leftarrow \mathscr{B}_2^+ \wedge \neg(\mathscr{B}_2^- \cap (M_2 \cup M_1)) \end{aligned}$$

Y así, $r \in RED(P_1 \cup P_2, M_1 \cup M_2)$. Por lo tanto, en cualquiera de los dos casos, $r \in RED(P_1 \cup P_2, M_1 \cup M_2)$. Por lo tanto,

$$RED(P_1, M_1) \cup RED(P_2, M_2) \subseteq RED(P_1 \cup P_2, M_1 \cup M_2)$$

Sea $r \in RED(P_1 \cup P_2, M_1 \cup M_2)$. Así,

$$\begin{aligned} r &= \alpha \leftarrow \mathscr{B}^+ \wedge \neg(\mathscr{B}^- \cap (M_2 \cup M_1)) \\ &= \alpha \leftarrow \mathscr{B}^+ \wedge \neg((\mathscr{B}^- \cap M_2) \cup (\mathscr{B}^- \cap M_1)) \end{aligned}$$

Como $\mathscr{L}_{P_1} \cap \mathscr{L}_{P_2} = \emptyset$, entonces si $\mathscr{B}^- \cap M_1 \neq \emptyset$ entonces $\mathscr{B}^- \cap M_2 = \emptyset$. O, por otro lado, si $\mathscr{B}^- \cap M_2 \neq \emptyset$ entonces $\mathscr{B}^- \cap M_1 = \emptyset$. De manera que $r \in RED(P_1, M_1)$ o $r \in RED(P_2, M_2)$, es decir, $r \in RED(P_1, M_1) \cup RED(P_2, M_2)$. Por lo tanto,

$$RED(P_1 \cup P_2, M_1 \cup M_2) \subseteq RED(P_1, M_1) \cup RED(P_2, M_2)$$

□

Definimos ahora la noción de modelo pseudo p-estable.

Definición 7.5. Sea P un programa lógico y $M \subseteq \mathscr{L}_P$. Decimos que M es un *Modelo pseudo p-estable* si se cumplen las siguientes condiciones:

1. G(P)=G($P \cup \neg\widetilde{M}$)
2. $P \cup \neg\widetilde{M}$ es literal completo (*i.e.*, para toda $a \in \mathscr{L}_P$ se tiene que $P \cup \neg\widetilde{M} \vdash_{G'_3} a$ o $P \cup \neg\widetilde{M} \vdash_{G'_3} \neg a$)

Observación 7.6. Note que puede suceder que G(P)=0, pero al mismo tiempo tener que G($P \cup \neg\widetilde{M}$) > 0. Por ejemplo,

$$\begin{aligned} P: \quad a &\leftarrow \neg a \\ c. & \end{aligned}$$

Se tiene que si $M = \{c\}$ entonces $\neg\widetilde{M} = \{\neg a\}$, y por lo tanto

$$\begin{aligned} P \cup \neg\widetilde{M}: \quad a &\leftarrow \neg a. \\ c. & \\ \neg a. & \end{aligned}$$

De donde se ve que $P \cup \neg\widetilde{M} \vdash_{G'_3} a$ y $P \cup \neg\widetilde{M} \vdash_{G'_3} \neg a$, y por tanto G($P \cup \neg\widetilde{M}$) > 0.

Necesitaremos el siguiente lema.

Lema 7.7. *Sea P un programa normal y $M \subseteq \mathscr{L}_P$ un modelo pseudo p-estable, entonces $RED(P, M) \vdash_{G'_3} M$*

Demostración. Supongamos que $RED(P, M) \nvdash_{G'_3} M$, entonces $P \cup \neg\widetilde{M} \nvdash_{G'_3} M$, y por tanto existe $a \in M$ tal que $P \cup \neg\widetilde{M} \nvdash_{G'_3} a$; Como M es literal completo, entonces se debe tener que $P \cup \neg\widetilde{M} \vdash_{G'_3} \neg a$. De aqui que $P \vdash_{G'_3} \neg a$, y así, $P \vdash_C \neg a$ lo cual es imposible ya que P es normal y no demuestra átomos negativos. □

Enunciamos ahora el resultado principal.

Teorema 7.8. *Sea P un programa normal. $M \subseteq \mathscr{L}_P$ es un modelo p-estable de P si, y sólo si, M es un modelo pseudo p-estable.*

Demostración. Sea M un modelo p-estable de P. Entonces, por definición, se tiene que $RED(P, M) \Vdash_C M$, es decir, que M es un modelo de $RED(P, M)$ y $RED(P, M) \vdash_C M$.
Por el Lema 6.5 se tiene que M es un modelo clásico de $P \cup \neg\widetilde{M}$, y por tanto, $P \cup \neg\widetilde{M}$ es consistente (en el sentido clásico). Ahora bien, si G($P \cup \neg\widetilde{M}$) >0 entonces existe φ tal que $P \cup \neg\widetilde{M} \vdash_{G'_3} \varphi$ y $P \cup \neg\widetilde{M} \vdash_{G'_3} \neg\varphi$. Pero esto implica que $P \cup \neg\widetilde{M} \vdash_C \varphi$

y $P \cup \neg\widetilde{M} \vdash_C \neg\varphi$, violando la consistencia de $P \cup \neg\widetilde{M}$. Por lo tanto $\mathrm{G}(P \cup \neg\widetilde{M}) = 0 = G(P)$.

Además, por el Lema 6.3, se tiene que $RED(P, M) \vdash_{G'_3} M$. Así, por el Lema 6.4 se tiene que $P \cup \neg\widetilde{M} \vdash_{G'_3} M$. De aqui que $P \cup \neg\widetilde{M}$ sea literal completo y por tanto M es un modelo Pseudo p-estable.

Inversamente, supongamos que M es un modelo Pseudo p-estable, esto es, $G(P) = G(P \cup \neg\widetilde{M})$ y para todo $a \in \mathscr{L}_P$ se tiene que $P \cup \neg\widetilde{M} \vdash_{G'_3} a$ o $P \cup \neg\widetilde{M} \vdash_{G'_3} \neg a$. Por el Lema 7.7 se tiene que $RED(P, M) \vdash_{G'_3} M$. Como G'_3 es más débil, tenemos que $RED(P, M) \vdash_C M$.
Mostraremos que M modela clásicamente a $RED(P, M)$. Para esto, es suficiente con probar que M modela a P, ya que esto implica que M modela clásicamente a $P \cup \neg\widetilde{M}$, y por el Lemma 6.5, se obtiene la conclusión.
Supongamos que M no modela clásicamente a P, entonces existe una cláusula $a \leftarrow \mathscr{B}^+ \wedge \mathscr{B}^-$ que es falsa, es decir, $a \in \widetilde{M}$, $\mathscr{B}^+ \subseteq M$ y $\mathscr{B}^- \subseteq \widetilde{M}$. Como $P \cup \neg\widetilde{M} \vdash_{G'_3} M$, entonces $P \cup \neg\widetilde{M} \vdash_{G'_3} \mathscr{B}^+$, y además $P \cup \neg\widetilde{M} \vdash_{G'_3} \neg\mathscr{B}^-$. Así, por Modus Ponens, se obtiene que $P \cup \neg\widetilde{M} \vdash_{G'_3} a$, y por tanto, $P \cup \neg\widetilde{M} \vdash_C a$. Por otro lado, $P \cup \neg\widetilde{M} \vdash_C \neg a$, violando la consistencia de $P \cup \neg\widetilde{M}$. Por lo tanto, M modela a P. □

Corolario 7.9. *Sea M un modelo clásico de un programa normal P tal que $RED(P, M) \vdash_C M$, entonces M es tambien un modelo pseudo p-estable de P.*

Demostración. Sea M un modelo clásico de P, entonces M es un modelo clásico de $P \cup \neg\widetilde{M}$ (ya que $\neg\widetilde{M}$ contiene sólamente átomos verdaderos). Ahora bien, como M es un modelo clásico de $P \cup \neg\widetilde{M}$, se tiene por el Lemma 6.5 que M es un modelo clásico de $RED(P, M)$. Por hipótesis, tenemos que $RED(P, M) \vdash_C M$ y así, se tiene por el Teorema 7.8 que M es un modelo Pseudo-pestable. □

Debemos observar la importancia de la hipótesis $RED(P, M) \vdash_C M$ en el Corolario 7.9. Consideremos los siguientes ejemplos:

- M es modelo clásico de P, pero no es un modelo Pseudo p-estable de P.

  Sea P el siguiente programa:

  $$P: \quad b \leftarrow a \wedge \neg a.$$

  Observe que $M = \{a, b\}$ es un modelo para P. Ahora bien, $\neg\widetilde{M} = \emptyset$, y por lo tanto

  $$P \cup \neg\widetilde{M}: \quad b \leftarrow a \wedge \neg a.$$

  Se tiene que $M = \{a, b\}$ no es un modelo Pseudo p-estable, ya que $P \cup \neg\widetilde{M}$ no es literal completo. De hecho, el programa P no tiene modelos Pseudo p-estables.
- M es un modelo clásico de P, y tambien es un modelo Pseudo p-estable de P.

  Sea P el siguiente programa:

  $$P: \quad b \leftarrow a \wedge \neg a.$$
  $$a.$$

Observe que $M = \{a\}$ es un modelo para P. Ahora bien, $\neg\widetilde{M} = \{\neg b\}$, y por lo tanto

$$P \cup \neg\widetilde{M}: \quad b \leftarrow a \wedge \neg a.$$
$$a.$$
$$\neg b.$$

Además, dado que $P \cup \neg\widetilde{M}$ es literal completo y que $\mathrm{G}(P) = \mathrm{G}(P \cup \neg\widetilde{M})$=0, se tiene que M es un modelo Pseudo p-estable de P.

Corolario 7.10. *Sean P_1 y P_2 programas normales con $\mathscr{L}_{P_1} \cap \mathscr{L}_{P_2} = \emptyset$, $M_1 \subseteq \mathscr{L}_{P_1}$ y $M_2 \subseteq \mathscr{L}_{P_2}$. Entonces M_1 y M_2 son modelos Pseudo pestables de P_1 y P_2 (resp.) si, y sólo si, $M_1 \cup M_2$ es un modelo Pseudo pestable de $P_1 \cup P_2$.*

Demostración. Supongamos que M_1 y M_2 son modelos Pseudo p-estables de P_1 y P_2, respectivamente. Por el Teorema 7.8, se tiene que M_1 es un modelo p-estable de P_1 y que M_2 es un modelo p-estable de P_2, es decir, M_1 modela clásicamente a P_1, $RED(P_1, M_1) \vdash_C M_1$, M_2 modela clásicamente a P_2 y $RED(P_2, M_2) \vdash_C M_2$.

Es claro entonces que $M_1 \cup M_2$ modela clásicamente a $P_1 \cup P_2$. Además, observe que por monotonía se tiene que $RED(P_1, M_1) \cup RED(P_2, M_2) \vdash_C M_1$ y $RED(P_1, M_1) \cup RED(P_2, M_2) \vdash_C M_2$, de aqui que $RED(P_1, M_1) \cup RED(P_2, M_2) \vdash_C M_1 \cup M_2$. Así, por la Proposición 7.4, se tiene que $RED(P_1 \cup P_2, M_1 \cup M_2) \vdash_C M_1 \cup M_2$. Por lo tanto, $M_1 \cup M_2$ es un modelo p-estable del programa normal $P_1 \cup P_2$, y por el Teorema 7.8 se tiene que $M_1 \cup M_2$ es un modelo pseudo p-estable de $P_1 \cup P_2$.

Inversamente, supongamos que $M_1 \cup M_2$ es un modelo pseudo p-estable de $P_1 \cup P_2$. Como $P_1 \cup P_2$ es un programa normal, entonces por el Teorema 7.8 se tiene que $M_1 \cup M_2$ es un modelo p-estable de $P_1 \cup P_2$, es decir, $M_1 \cup M_2$ modela clásicamente a $P_1 \cup P_2$ y $RED(P_1 \cup P_2, M_1 \cup M_2) \vdash_C M_1 \cup M_2$.

Supongamos que M_1 no modela clásicamente a P_1, entonces $M_1 \cup M_2$ no modela clásicamente a P_1, y por tanto, $M_1 \cup M_2$ no modela clásicamente a $P_1 \cup P_2$. De forma análoga, se demuestra que M_2 modela clásicamente a P_2. □

Referencias

[1] W. A. Carnielli and J. Marcos.: *A taxonomy of C-Systems.* In Paraconsistency: The Logical Way to the Inconsistent, Proceedings of the Second World Congress on Paraconsistency (WCP 2000), number 228 in Lecture Notes in Pure and Applied Mathematics, pages 1–94. Marcel Dekker, Inc., 2002.

[2] M. Osorio, J. A. Navarro, J. Arrazola, and V. Borja.: *Logics with common weak completions.* Accepted in Journal of Logic and Computation, 2006.

[3] Mauricio Osorio, José Arrazola, José L. Carballido, Oscar Estrada, *An axiomatization of G'_3*, Workshop in Logic, Language and Computation 2006, http://SunSITE.Informatik.RWTH-Aachen.DE/Publications/CEUR-WS/, Vol 220, ISSN 1613-0073, Noviembre 2006.

[4] J. W. Lloyd, *Foundations of Logic Programming.* Springer Verlag, Berlin. 1984.

arrazola@fcfm.buap.mx
jlavallenator@gmail.com
cambron99@hotmail.com

¿QUÉ ES ANSWER SET PROGRAMMING (ASP)?

JOSÉ ARRAZOLA
JESÚS LAVALLE
FELIPE MAZÓN

Resumen. Se presenta una breve introdución a Answer Set Programming (ASP), por medio de la solución de tres problemas concretos usando ASP: El problema de las N Reinas, el juego de Sudoku y el Coloreo de Nodos. Las soluciones se implementan en el buscador de modelos estables Smodels.

1. Introducción

Answer Set Programming (también conocida como *programación lógica estable* o *A-Prolog*) es un tipo especial de programación declarativa, su sintaxis es parecida a la utilizada en programación lógica. En programación lógica tradicional, la *negación por falla* indica la ausencia de evidencia, o la falla en la derivación de una literal; En Answer Set Programming, indica la consistencia de una literal, es decir, la literal puede ser asumida falsa sin que se produzca una inconsistencia. Así, Answer Set Programming consiste en codificar una problema como un conjunto de reglas (o cláusulas) de manera que su(s) solución(es) sean *capturadas* por los modelos estables de dichas reglas [1].
Generalemente, el crédito por la creación de la Programación Lógica se les dá a Robert Kowalski y a Alain Colmerauer, cuyo trabajo en esta área se realizó a mediados de los setentas[2].

Los lenguages de Programación Lógica tiene una base lógica formal, lo cual permite que el programador se concentre en el "qué" resolver en lugar de en el "cómo" resolverlo. Esto permite que los programas hechos en los lenguages de programación lógica sean mucho mas sencillos que los realizados en lenguages de *programacion imperativa* (programación convencional). A pesar de la sencillez obtenida en los programas lógicos, éstos no han sido muy populares debido principalmente a una razón: existe la creencia de que los programas realizados en lenguajes de programación lógica(Prolog, Aurora, etc.) no son tan eficientes como aquellos realizados en lenguajes de programación imperativa (Pascal, C, C++, etc.). Existen algunos estudios comparativos entre programación lógica y programación imperativa que muestran que realmente no existe mucha diferencia en cuanto a la eficiencia (vea por ejemplo [3]).

La mayoría de los lenguajes de programación lógica trabajan de manera *interactiva* por medio de "preguntas y respuestas", es decir, el usuario hace una pregunta y el lenguaje de programación utiliza el programa lógico para tratar de dar la respuesta, que en la mayoría de los casos consiste simplemente en un "si", un "no" o (si hay variables involucradas en la pregunta) los valores de la variable para los

cuales la pregunta tendría una respuesta afirmativa. La diferencia entre la programación lógica y la programación imperativa se entiende mejor mediante un ejemplo. Supongamos que queremos resolver el siguiente problema:

"Escribir un programa que tome dos números y calcule su suma".

La solución a este problema usando C pudiera darse como:

```
#include <stdio.h>

main (void){

int a, b;

printf ("Da un numero:\n");
scanf ("%d", &b);
printf ("Da el segundo numero:\n");
scanf ("%d", &a);
printf("La suma es: %d\n\n", a+b);

return 1;
}
```

mientras que una solución en Prolog se vería simplemente como:

```
suma(X,Y,Z)  :- Z is X + Y
```

Al ejecutar el programa en Prolog obtenemos:

```
| ?- suma(2,3,V).
V=5
```

Observemos que en la solución dada en C, tenemos que preocuparnos por definir el tipo de valor que tomarán las variables y debemos decirle al compilador qué valor debemos esperar cuando se pide el valor; más aun, debemos decirle explícitamente que operaciones realizar para llegar al resultado. Esto no sucede en el programa lógico, aquílo único que nos preocupa es cuándo el consecuente de la cláusula dada es verdadero (puede ser deducida), porque cuando lo es, tenemos precisamente la respuesta.

Es cierto que el ejemplo mostrado es trivial, pero esto nos da una idea de la reducción en el tamaño del código y de la sencillez de la lectura del mismo.

Como ya se dijo, Answer Set Programming consiste en codificar una problema como un conjunto de reglas (o cláusulas) de manera que su(s) solución(es) sean *capturadas* por los modelos estables de dichas regla. Así pues, veamos algunos de los conceptos involucrados en la teoría de Answer Sets.

2. Answer Sets (Modelos Estables)

Necesitaremos definir algunos conceptos.

2.1. **Definición.** Una cláusula es una fórmula de la forma $\mathcal{H} \leftarrow \mathcal{B}$, en donde $\mathcal{H}$ y $\mathcal{B}$ son conjunciones o disyunciones de átomos. A la fórmula $\mathcal{H}$ se le llama la *cabeza* del programa y a la fórmula $\mathcal{B}$ se le llama el *cuerpo* del programa.

2.2. **Definición.** Una cláusula normal es una fórmula de la forma

$$A_0 \longleftarrow A_1 \wedge A_2 \wedge \cdots \wedge A_k \wedge \neg A_{k+1} \wedge \cdots \wedge \neg A_n$$

en donde A_i son literales (*i.e.*, son átomos o las negaciones de átomos).

2.3. **Definición.** Una *cláusula aumentada* se define como una fórmula de la forma $\mathcal{H} \longleftarrow \mathcal{B}$, donde $\mathcal{H}$ y $\mathcal{B}$ son fórmulas que no contienen el conectivo $\longleftarrow$.

2.4. **Definición.** Una restricción o *constraint* es una fórmula de la forma

$$\bot \longleftarrow A_1 \wedge A_2 \wedge \cdots \wedge A_k \wedge \neg A_{k+1} \wedge \cdots \wedge \neg A_n$$

2.5. **Definición.** Una cláusula definite es una fórmula de la forma

$$A_0 \longleftarrow A_1 \wedge A_2 \wedge \cdots \wedge A_k$$

en donde A_i son literales. Observe que no hay literales negadas.

Un programa es un conjunto de fórmulas de alguno de los tipos anteriores. Así, un *programa normal* es un conjunto de cláusulas normales, un *programa definite* es un conjunto de cláusulas definite y un *programa aumentado* es un conjunto de cláusulas aumentadas. Denotaremos por $\mathfrak{L}_P$ a el conjunto de átomos que aparecen en las fórmulas del programa P.

2.6. **Definición.** Un *modelo* de un programa normal P es un subconjunto de $\mathfrak{L}_P$ tales que, considerando a sus elementos como verdaderos y como falsos a los elementos de $\widetilde{M} = \mathfrak{L}_P \setminus M$, se tiene que cada cláusula del programa P es verdadera,

Por ejemplo, consideremos el siguiente programa normal:

$$\begin{aligned} P: & \\ a &\leftarrow \neg b \\ b &\leftarrow \neg c \\ c &\leftarrow \neg a \\ c &\leftarrow \neg b \end{aligned}$$

Como se puede comprobar, un modelo para P es $M = \{a, c\}$. Observemos que otros modelos para P son $M' = \{a, b\}$ y $M'' = \{a, b, c\}$. La observacion anterior nos puede sugerir dos preguntas:

(1) ¿Todo programa normal tiene algún modelo?
(2) ¿Si tiene modelo, entonces tiene un modelo mínimo? (en el sentido de la contención)

La respuesta a la primer pregunta es afirmativa, ya que en el peor de lo casos, el modelo para un programa normal son todos los átomos que aparecen en la cabeza de cada cláusula. La respuesta a la segunda pregunta es negativa, como lo muestra el ejemplo anterior en donde los modelos M y M' no son comparables. Lloyd da una respuesta afirmativa para programas definite.

Al tratar de generalizar a programas normales los resultados obtenidos para programas *definite*, Gelfond y Lifschitz propusieron la siguiente transformacion: Dado $M \subseteq \mathfrak{L}_P$, definimos para cada cláusula de un programa normal P:

$$\left(A_0 \longleftarrow \bigwedge_{i=1}^{k} B_i \wedge \bigwedge_{i=k+1}^{n} \neg B_i \right)^M =$$

$$\begin{cases} \top & \text{si} \quad B_i \in M \text{ para algún } i = k+1, \ldots, n \\ A_0 \longleftarrow \bigwedge_{i=1}^{k} B_i & \text{si} \quad B_i \notin M \text{ para todo } i = k+1, \ldots, n \end{cases}$$

Así, definimos:

$$P^M = \left\{ \left(A_0 \longleftarrow \bigwedge_{i=1}^{k} B_i \wedge \bigwedge_{i=k+1}^{n} \neg B_i \right)^M \; : \; A_0 \longleftarrow \bigwedge_{i=1}^{k} B_i \wedge \bigwedge_{i=k+1}^{n} \neg B_i \in P \right\}$$

Observe que si P es un programa normal, entonces P^M es un programa definite y por tanto tiene un modelo mínimo.

2.7. **Definición.** Dado un programa normal P, decimos que $M \subseteq \pounds_P$ es un *Modelo Estable (Answer Set)* de P si M es un modelo para P^M.

Al igual que antes, nos pudieramos preguntar si todo programa normal P tiene un modelo estable. El siguiente ejemplo, muestra que la respuesta es negativa.

1. **Ejemplo.** Considere el programa normal P:

$$P:$$
$$p \quad \leftarrow \quad \neg p$$

Aquí, $\pounds_P = \{p\}$. Por tanto los únicos candidatos a ser answer set son $\pounds_P$ y $\emptyset$.

Si $M = \emptyset$, entonces

$$P^M:$$
$$p \quad \leftarrow \quad \top \qquad (\text{ o simplemente, } p)$$

pero $\emptyset$ no es un modelo de P^M, por tanto no es answer set de P.

Si $M = \pounds_P$, entonces

$$P^M:$$
$$\top$$

pero $\pounds_P$ no es un modelo de P^M y por ende no es un answer set de P.
Por lo tanto, el programa P no tiene answer sets, pero si tiene un modelo (clásico): $\{p\}$.

2. **Ejemplo.** Considere el siguiente programa normal

$$P:$$
$$p \quad \leftarrow \quad \neg q$$

En este caso $\pounds_P = \{p, q\}$. Luego, los posibles candidatos para ser answer set de P son los subconjuntos $\pounds_P = \{p, q\}, \{p\}, \{q\}, \emptyset$. Veamos cual de ellos es un answer set:

Si $M = \emptyset$, entonces

$$P^M:$$
$$p$$

pero entonces M no es un modelo de P^M, así $M = \emptyset$ no es un answer set de P.

Si $M = \{q\}$, entonces

$$P^M : \\ \top$$

pero entonces M no es un modelo de P^M. Por tanto, $M = \{q\}$ no es un answer set de P.

Si $M = \{p\}$, entonces

$$P^M : \\ p \longleftarrow \top \qquad (\text{ o simplemente, } p)$$

En este caso, M si es un modelo de P^M, y por tanto $M = \{p\}$ si es un answer set de P.

Finalmente, si $M = \{p, q\}$, entonces

$$P^M : \\ \top$$

pero entonces M no es un modelo de P^M, consecuentemente $M = \{p, q\}$ no es un answer set de P.

Observe que los modelos (clásicos) de P son $\{p\}, \{q\}$ y $\{p, q\}$, y que $\{q\}$ no es un answer set de P. Esto muestra que los conceptos de Modelo y Modelo Estable no necesariamente coinciden.

Intuitivamente, un modelo es estable si todo átomo en él tiene "alguna razón" para estar ahi: para cada átomo en el modelo tiene que existir alguna regla que tiene a ese átomo en la cabeza y tal que el cuerpo de la regla es verdadera en el modelo.

3. Smodels

Básicamente existen cuatro tipos de objetos en los lenguajes de programación lógica: *átomos*, *constantes*, *variables* y *reglas*. Las constantes son objetos individuales que existen en el universo del problema, éstas pueden ser números o constantes simbólicas. Se utiliza las variables para generalizar; a diferencia de la programación imperativa, en donde usualmente se le asigna algún valor a las variables, en los programas lógicos se encuentra el valor correcto para ellas. Un átomo consiste de un símbolo de predicado seguido por una lista de constantes o variables entre paréntesis. Los átomos se utilizan para expresar relaciones entre constantes, por ejemplo el átomo *padre(juan,maría)* pudiera decirnos que juan es el padre de maría. Una regla nos permite hacer inferencias basadas en los predicados, por ejemplo la regla "*hermanos(X,Y) :- padre(Z,X),padre(Z,Y)*" nos dice que X y Y son hermanos si ambos tienen el mismo padre. Las reglas tienen dos partes: la cabeza(que se encuentra de lado izquierdo de ":-" y el cuerpo, que se encuentra a la derecha de ":-").

También existen *literales de restricción* (constraint literals), las cuales tienen la forma " $lower\{l_1, l_2, \ldots, l_n\}upper$", donde *lower* y *upper* son expresiones aritméticas. Una literal de restricción es satisfecha, si el número de literales satisfechas en el cuerpo de la restricción está entre *lower* y *upper*.

Una *literal condicional* es de la forma: "$p(X) \ : \ q(X)$". Si la extensión de q es $\{q(a_1), q(a_2), \ldots, q(a_n)\}$, la condición de arriba es semánticamente equivalente a escribir $p(a_1), p(a_2), \ldots, p(a_n)$, en el lugar de la condición. Así, por ejemplo

```
q(1..2).
a :- 1 \set{ p(X) : q(X)}.
```

dará

```
q(1). q(2).
a :- 1 \set{p(1), p(2)}.
```

Regularmente un programa lógico puede dividirse en dos partes: un conjunto de reglas de inferencia y una base de datos considerados verdaderos denominados"hechos" con ello se realiza la inferencias. Por ejemplo, el siguiente programa codifica una base de datos familiar y simple:

```
hermano(X,Y) :- padre(Z,X), parent(Z,Y).
madre(X,Y) :- padre(X,Y),mujer(X).
tio(X,Y) :- padre(Z,Y),hermano(Z,X),hombre(X).

mujer(joan). mujer(jill). hombre(jack).
padre(joan, jack). padre(joan, jill).
```

4. Ejemplos

4.1. **N-Reinas.** En el ajedréz, una reina puede moverse tan lejos y en la dirección que desee. Un tablero de ajedréz tiene 8 renglones y 8 columnas. El problema clásico pide que se acomoden 8 reinas en un tablero de ajedréz ordinario de manera que ninguna de ellas pueda atacar a alguna otra en un movimiento. Una solución se muestra en la figura 1.

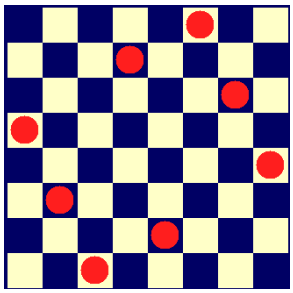


Figure 1. Solución para el caso n=8

Una solución escrita en lenguaje C es la siguiente:

```
include <stdio.h>
define TRUE 1
define FALSE 0

int *board;
```

```
int main()
{
board=(int*)calloc(8+1,sizeof(int));
board++;
//here goes the userinput no of queens doesnÂ´t matter in this code
int col;
int queens=2;//example from userinput
for(col=1;col<=8;col++)
     placeQueens(1,col,queens);
}
void placeQueens(int row,int col,int queens)
{
       int i;
        for(i=1;i<row;i++)
       {
        if((board[i] == col) || ((row+col)==(i+board[i]))
                    ||((row-col)==(i-board[i])))
        {
            check= FALSE;
        }
        else
         check=TRUE;
     }
      if(check==TRUE)
       {
         board[row]=col;

        if(row==8)
        {

            for(i=1;i<=queens;i++)
            printf("(%d,%d)",i,board[i]);
            printf("\n");
        }
        else
        {
            for(i=1;i<=8;i++)
            placeQueens(row+1,i);
        }
    }

}
```

En Niemelä [1] se presenta la siguiente solución al problema de Las N-Reinas:

```
1 { q(X,Y) : d(X) } 1 :- d(Y).
1 { q(X,Y) : d(Y) } 1 :- d(X).
:- d(X), d(Y), d(X1), d(Y1),
   q(X,Y), q(X1,Y1),
```

```
   X != X1, Y != Y1,
   abs( X - X1) == abs( Y - Y1).
d(1..n).
```

Lo anterior nos brinda una idea de la competencia entre ambos tipos de programación.

4.2. **Sudoku.** Sudoku es un juego muy popular que consiste en completar con los números del 1 al n, un cuadrado de tamaño $n \times n$ de forma que no se repita ninguna cifra en cada fila, ni en cada columna, ni en cada *subcuadrado*. Por ejemplo, en la siguiente figura se ve el problema en un cuadrado de tamaño 4×4 junto a su solución:

1			
		2	
	3		
			4

Solución $\Longrightarrow$

1	2	4	3
3	4	**2**	1
4	**3**	1	2
2	1	3	**4**

En Jiménez en [5] se presenta la siguiente solución al juego del Sudoku en un tablero de tamaño 9×9:

```
const n=9.
posicion(1..n).
#domain posicion(X; Y; Z; XA; XB; YA; YB).
valor(1..n).
#domain valor(V).
1 { estado(X,Y,M) : valor(M)} 1.

:- estado(XA,Y,V),
   estado(XB,Y,V),
   XA != XB.

:- estado(X,YA,V),
   estado(X,YB,V),
   YA != YB.

:- estado(XA,YA,V),
   estado(XB,YB,V),
   mismo_cuadrado(XA,XB),
   mismo_cuadrado(YA,YB),
   XA != XB,
   YA != YB.

mismo_cuadrado(X,Y) :-
   raiz_de_n(M),
   div(X-1,M) == div(Y-1,M).

raiz_de_n(X) :-
   X*X == n.

#hide.
#show estado(X,Y,V).
```

4.3. **Coloreo de Nodos.** Un problema clásico en programación es el de colorear los nodos de un grafo de manera que dos nodos adyacentes no tengan el mismo color. En [4] se muestra el siguiente programa que resuelve el problema del coloreo de nodos.

```
color(rojo). color(azul). color(amarillo).
col(X,rojo) :- nodo(X), not col(X,azul), not col(X, amarillo).
col(X,azul) :- nodo(X), not col(X,rojo), not col(X, amarillo).
col(X,amarillo) :- nodo(X), not col(X,azul), not col(X, rojo).
fail :- arista(X,Y), color(C),col(X,C), col(Y,C).

nodo(a). nodo(b). nodo(c). nodo(d).
arista(a,b) arista(b,c). arista(c,d). arista(d,a).
compute 1 { not fail }.
```

4.4. **Conclusión.** Como el lector habrá notado este trabajo lo introduce a la teoría de ASP mediante el uso de ejemplos, permitiendole comparar a ASP con algún tipo de programación imperativa

Referencias

[1] Ilkka Niemelä, Patrick Simmons and Tommi Syrjänen, *Smodels: A System for Answer Set Programming*, arXiv:cs.AI/0003033 v1 6 Mar 2000.
[2] Valdimir Lifschitz, *Foundations of Logic Programming*
[3] V. M. Calegario and I. C. Dutra. *Performance Comparison between Conventional and Logic Programming Systems*. Technical Report ES-478/98, COPPE/Systems Engineering and Computer Science, Setembro 1998. http://citeseer.ist.psu.edu/calegario98performance.html
[4] Tommi Syrjänen, *Lparse 1.0 User's Manual*, http://www.tcs.hut.fi/Software/smodels.
[5] José A. Alonso Jiménez, *Solución declarativa del Sudoku mediante ASP*,

arrazola@fcfm.buap.mx
jlavallenator@gmail.com
cambron99@hotmail.com

ESTABILIDAD EN PROGRAMACIÓN LINEAL

SORAYA GÓMEZ Y ESTRADA
LIDIA HERNÁNDEZ REBOLLAR
ARTURO LANCHO ROMERO
BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA
UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA

Resumen. En programación lineal (PL) el estudio de sensibilidad es uno de los temas importantes porque facilita los cálculos cuando se modifican algunos de los datos del problema. Con este estudio conocemos, por ejemplo, cual es el valor óptimo después de cambiar ciertos datos del problema original. La estabilidad, en cambio, es un concepto que nos da información cualitativa sobre el comportamiento de un problema cuando modificamos los datos. La estabilidad en programación lineal semi-infinita (PLSI) ha sido estudiada por grupos de investigadores en España, Alemania, Bulgaria y recientemente en Puebla. El objetivo del proyecto de investigación que presentamos en esta ocasión, es estudiar la estabilidad en problemas particulares de la programación lineal ordinaria. El primer problema que hemos abordado es el del problema del transporte.

1. Introducción

En muchas aplicaciones de la programación lineal se desea resolver un problema, llamado primal, algunos de cuyos datos son resultado de estimaciones imprecisas, son redondeados o pueden variar con el paso del tiempo. Si los coeficientes de la función objetivo representan, por ejemplo, los costos de las materias primas, sabemos que estos están sujetos a pequeñas, o algunas veces, grandes variaciones. De igual forma pueden variar los coeficientes de las restricciones y el lado derecho, inclusive se contempla la posibilidad de aumentar o disminuir restricciones. Las posibles perturbaciones dependen del problema concreto que se está resolviendo.

El análisis de sensibilidad nos da información cuantitativa sobre el conjunto factible, el conjunto óptimo y el valor optimal sin necesidad de volver a hacer todos los cálculos cuando modificamos algunos datos del problema.

Los paquetes comerciales de software que resuelven problemas de PL por el método simplex suelen proporcionar, además de la solución óptima primal y dual de un problema, un análisis de sensibilidad. Este análisis individual, para cada costo c_i y para cada lado derecho b_j, consiste del intervalo dentro del cual pueden variar dichos parámetros sin que cambie el punto óptimo, pero la interpretación de este intervalo de variación puede ser engañosa cuando la solución del problema, primal o dual, es múltiple. Cuando decimos que un problema es estable podemos referirnos a que su conjunto factible, el optimal o su función valor óptimo no tiene "grandes" cambios cuando se cambian los datos, y estaremos hablando, en este caso, de una propiedad cualitativa de estos conjuntos, en el sentido de que al hacer pequeñas perturbaciones en algunos o todos los datos, el conjunto factible u optimal

se mantienen diferentes del vacío o "cercanos" a los respectivos conjuntos factible u optimal del problema original.

En la primera parte de este trabajo presentamos algunas definiciones y revisamos algunos teoremas generales sobre sensibilidad en PL y sobre estabilidad en PLSI. En la segunda parte presentamos el problema del transporte y algunas de sus características que esperamos nos sean útiles en el estudio de la estabilidad de este problema, en particular de su conjunto factible y del conunto de puntos extremos. Para denotar a los reales positivos usaremos el símbolo $\mathbb{R}_{++}$, y para los reales mayores o iguales que cero usaremos $\mathbb{R}_{+}$.

2. Resultados sobre sensibilidad y estabilidad

Consideremos al problema primal (P) en su forma general, con su correspondiente vector de datos (c^T, a_i^T, b_i)

$$\begin{aligned}(P)\qquad \min\ & c^T x \\ \text{s.a}\ & a_i^T x \geq b_i,\ i \in I \\ & a_i^T x \leq b_i,\ i \in J \\ & a_i^T x = b_i,\ i \in K\end{aligned}$$

donde los conjuntos de índices I, J, K son finitos, disjuntos dos a dos, y donde también puede ocurrir que alguno de ellos sea vacío pero no los tres. Al conjunto factible de (P) lo denotaremos por F. El problema perturbado $(\widetilde{P})$ puede expresarse como

$$\begin{aligned}(\widetilde{P})\qquad \min\ & \widetilde{c}^T x \\ \text{s.a}\ & \widetilde{a}_i^T x \geq \widetilde{b}_i,\ i \in I \\ & \widetilde{a}_i^T x \leq \widetilde{b}_i,\ i \in J \\ & \widetilde{a}_i^T x = \widetilde{b}_i,\ \ i \in K\end{aligned}$$

cuyo conjunto factible es $\widetilde{F} \subset \mathbb{R}^n$ y cuyo vector de datos es $(\widetilde{c}^T, \widetilde{a}_i^T, \widetilde{b}_i),\ i \in I \cup J \cup K$. Llamaremos tamaño de la peturbación a la distancia euclidea entre los vectores de datos de (P) y de $(\widetilde{P})$, que representaremos como $d((P), (\widetilde{P}))$. Diremos que (P) es estable si existe $\varepsilon > 0$ tal que $\widetilde{F} \neq \varnothing$ si $d((P), (\widetilde{P})) < \varepsilon$.

2.1. Teorema. [5] (**De sensibilidad respecto de** c). Si la única solución óptima de (P) es $\overline{x}$, entonces existe un entorno de c donde $v(z) = \overline{x}^T z$, para todo z perteneciente al mismo. Por tanto, una perturbación suficientemente pequeña de c, Δc, incrementa el valor de (P) en $\Delta v = \overline{x}^T \Delta c$.

2.2. Teorema. [5] (**De sensibilidad respecto de** b). Si $\overline{y}$ es la única solución óptima del problema dual de (P), entonces existe un entorno de b donde $v(w) = \overline{y}^{\ T} w$, para todo w perteneciente al mismo. Por tanto, una perturbación suficientemente pequeña de b, Δb, incrementa el valor del problema en $\Delta v = \overline{y}^{\ T} \Delta b$.

2.3. Teorema. [5] (**De estabilidad**). (P) es estable si y sólo si $\{a_i, i \in K\}$ es LI, cuando $K \neq \varnothing$, y existe $\overline{x} \in \mathbb{R}^n$ tal que $a_i^T \overline{x} > b_i$ para todo $i \in I$, $a_i^T \overline{x} < b_i$ para todo $i \in J$ y $a_i^T \overline{x} = b_i$ para todo $i \in K$. En particular, la condición de Slater caracteriza la estabilidad de los problemas en forma canónica.

En programación lineal semi infinita el problema primario suele presentarse de la manera siguiente

$$(P) \qquad \inf \quad c'x$$

$$\text{s.a.} \quad a_t'x \geq b_t, \qquad t \in T,$$

donde $c \in \mathbb{R}^n$, T es un conjunto de índices, $a_t = a(t) = (a_1(t), ..., a_n(t))' \in \mathbb{R}^n$, y $b_t = b(t) \in \mathbb{R}$. Las funciones a y b son funciones del conjunto T en $\mathbb{R}^n$ y $\mathbb{R}$ respectivamente.

Cada problema (P) está representado por la triada $\pi = (c, a., b.)$, la cual pertenece al espacio de parámetros Π definido como sigue

$$\Pi = \mathbb{R}^n \times (\mathbb{R}^n)^T \times \mathbb{R}^T.$$

En este espacio hemos fijado la dimensión n y el conjunto de índices T.

En los teoremas que se presentan más adelante se ha caracterizado la estabilidad del problema (P) con respecto a la consistencia de su conjunto factible. Si llamamos Π_c al conjunto de parámetros π tales que su conjunto factible F es no vacío, entonces diremos que (P) es estable si y sólo si $\pi \in \text{int } \Pi_c$. Dicha caracterización se ha hecho por la semicontinuidad inferior de la multifunción conjunto factible y también se han usado otros conceptos de continuidad para funciones multivaluadas o multifunciones que definiremos enseguida.

2.4. Definición. Si Y y Z son dos espacios topológicos y $S : Y \to Z$ es un mapeo multivaluado, entonces, diremos que S es **semicontinuo inferiormente según Berge** (B-lsc), en un punto $y \in Y$, si para todo abierto $W \subset Z$ tal que $W \cap S(y) \neq \emptyset$ existe un abierto $U \subset Y$, que contiene a y, tal que $W \cap S(y^1) \neq \emptyset$ para cada $y^1 \in U$.

2.5. Definición. S es **semicontinuo superiormente según Berge** (B-usc) en un punto $y \in Y$, si para todo abierto $W \subset Z$ tal que $S(y) \subset W$ existe una vecindad abierta U de y en Y, tal que $S(y^1) \subset W$ para cada $y^1 \in U$.

2.6. Definición. Si S es tanto semicontinuo inferiormente como superiormente según Berge en $y \in Y$, entonces es llamado **B-continuo** en y.

La multifunción conjunto factible $\mathcal{F} : \Pi \rightrightarrows \mathbb{R}^n$ es tal que para cada $\pi \in \Pi$, $\mathcal{F}(\pi) = F$.

2.7. Teorema. [2, Teorema 1] [3, Teorema 3.1 y Teorema 6.2] Para $\pi = (c, a., b.) \in \Pi$ dado, son equivalentes:

i) $\pi \in \text{int } \Pi_c$.
ii) $\mathcal{F}$ es B-lsc en $\pi \in \Pi_c$.
iii) Existe $\bar{x} \in \mathbb{R}^n$ y $\varepsilon > 0$ tal que $a_t'\bar{x} \geq b_t + \varepsilon$ para todo $t \in T$ ($\bar{x}$ es llamado strong Slater element, o SS-element).
iv) $(c, a., b.^1) \in \Pi_c$ para b^1 suficientemente cercano a b (en la pseudométrica uniforme).

2.8. Teorema. [2, Teorema 2] Sea $\pi \in \Pi_c$.

i) Si F es acotado, entonces $\mathcal{F}$ es B-usc y uniformemente acotado en π.
ii) Si F es acotado y $\pi \in \text{int } \Pi_c$, entonces $\mathcal{F}$ es B-continuo y H-continuo en una cierta vecindad de π.

3. El problema del transporte

El problema del transporte de Hitchcock con m-fuentes y n-destinos puede formularse como sigue: Sea $c = (c_{ij})$ con i=$1, 2, ...m$ y j=$1, 2, ..., n$, el vector renglón de costos de tamaño $m \times n$, b=$(b_1, b_2, ...b_m, b_{m+1}, b_{m+2}, ..., b_{m+n})'$ el vector columna de tamaño $m + n$, donde b_i representa la oferta de la fuente i-ésima, i=$1, 2, ...m$, y b_{m+j} representa la demanda del destino j-ésimo, j=$1, 2, ..., n$. El problema consiste en minimizar el costo total por el envio de x_{ij} unidades de mercancía (de cada fuente i a los destinos j), suponiendo que la suma de la oferta de todas las fuentes es igual a la suma de la demanda de todos los destinos. Esto es,

$$\begin{aligned} & \min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \\ \text{s.a.} \quad & \sum_{j=1}^{n} x_{ij} = b_i \qquad i=1, 2, ...m \\ & \sum_{i=1}^{m} x_{ij} = b_{m+j} \quad j=1, 2, ..., n \\ & x_{ij} \geq 0 \quad (\ i=1, 2, ...m, \ \ j=1, 2, ..., n) \end{aligned}$$

de la representación anterior: $\sum_{j=1}^{n} b_{m+j} = \sum_{i=1}^{m} b_i$.

Haciendo perturbaciones sólo en el vector b, deseamos determinar cómo cambia el conjunto factible F_b. Fijando m y n, y el vector de costos c, el espacio de parámetros es:

$$B = \left\{ b \in \mathbb{R}^{m+n}_{++} \colon \sum_{j=1}^{n} b_{m+j} = \sum_{i=1}^{m} b_i \right\}.$$

El conjunto $B \subset \mathbb{R}^{m+n}_{++}$ es un cono relativamente abierto que no contiene al origen. Por ejemplo para $m = 1$, $n = 2$,

$$B = \left\{ b \in \mathbb{R}^{3}_{++} \colon b_1 = b_2 + b_3 \right\}.$$

Luego, para medir las perturbaciones consideraremos la norma euclidea restringida a B. De manera más general, como $\mathbb{R}^{m+n}$ es un espacio topológico y $B \subset \mathbb{R}^{m+n}$, entonces, al hablar de conjuntos abiertos en B, lo haremos considerando la topología usual inducida por $\mathbb{R}^{m+n}$ en B.

A cada parámetro b le asociaremos su conjunto factible F_b, su conjunto de puntos óptimos F_b^* y su conjunto de puntos extremos E_b.

Para cada $b \in B$, nuestro problema (P_b) puede formularse también en forma matricial:

$$\begin{aligned} & \min cx \\ & \text{s.a. } Ax = b \\ & x \geq 0. \end{aligned}$$

La matriz A es de tamaño $(m + n) \times mn$ y tiene una forma muy particular:

$$A = \begin{pmatrix} \mathbf{1}, 0, ..., 0 \\ 0, \mathbf{1}, ..., 0 \\ 0, 0, ..., \mathbf{1} \\ I, I, ..., I \end{pmatrix}$$

donde $\mathbf{1} = (1, 1, ..., 1)$ es un $n-$vector renglón de unos e I es la matriz identidad $n \times n$. Se puede ver facilmente que debido a las restricciones del problema la matriz A tiene rango $m + n - 1$.

Con esta nueva notación, el conjunto factible para cada $b \in B$ puede escribirse como sigue:

$$F_b = \left\{x \in \mathbb{R}_+^{mn} \colon Ax = b\right\}.$$

Puesto que el vector $\bar{x}$ con componentes $\bar{x}_{ij} = \frac{b_i b_{m+j}}{\sum_{i=1}^{m} b_i}$ satisface la ecuación $Ax = b$ y $\bar{x}_{ij} \geq 0$ para toda i=$1, 2, ...m,\ j = 1, 2, ..., n$, tenemos que $F_b \neq \varnothing$. Por otro lado es fácil ver también que si $x \in F_b$ entonces se cumplen las desigualdades

$$0 \leq x_{ij} \leq \min\{b_i, b_{m+j}\}$$

para cada una de las componentes de x. Luego F_b es acotado, y por lo tanto, nuestro problema tiene solución, esto es, que $F_b^* \neq \varnothing$. Las conclusiones que hemos obtenido para el conjunto factible y el optimal se cumplen para cualquier $b \in B$, luego, si denotamos por B_c y B_s los espacios de problemas de transporte que tienen solución factible y solución óptima, respectivamente, podemos afirmar que $B = B_c = B_s$. Además, int $B =$ int $B_c = B$, por ser B un conjunto abierto en el espacio $\mathbb{R}^{m+n-1}$. Las igualdades anteriores nos indican que cualquier problema de transporte con m fuentes y n destinos es estable, en el sentido de que, bajo pequeñas perturbaciones en cada $b \in B$, el conjunto factible se conserva diferente del vacío, y lo mismo ocurre con el conjunto optimal.

3.1. **Multifunciones factible, optimal y de puntos extremos.** Para estudiar más ampliamente la estabilidad del conjunto factible, del conjunto optimal y del conjunto de puntos extremos definiremos las siguientes multifunciones o funciones multivaluadas, $\mathcal{F}$, $\mathcal{F}^*$ y $\mathcal{E}$, respectivamente, todas del conjunto B a un subconjunto de $\mathbb{R}_{++}^{mn}$, y tales que, asignan, a cada parámetro b, su correspondiente conjunto factible, conjunto optimal y conjunto de puntos extremos, respectivamente. Es decir, para cada $b \in B$, $\mathcal{F}(b) = F_b$, $\mathcal{F}^*(b) = F_b^*$, y $\mathcal{E}(b) = E_b$.

Si la función multivaluada $\mathcal{F}$ es continua en un determinado $b \in B$, significa que para parámetros b' muy cercanos a b su correspondiente conjunto factible $F_{b'}$ estará también muy cercano a F_b, esto nos define otro tipo de estabilidad para el problema (P_b), pues nos dice que cuando nos aproximamos a b, los puntos factibles de los problemas perturbados también se aproximan a los puntos factibles del problema original. Para la continuidad de las funciones multivaluadas mencionadas antes consideraremos la semicontinuidad inferior y la semicontinuidad superior según Berge, de tal forma que entenderemos por función continua en un punto a aquella que es tanto semicontinua inferiormente como superiormente en el mismo punto.

3.2. **Acerca del conjunto factible.** Para estudiar la continuidad de la multifunción conjunto factible hemos visto la necesidad de tener un representación o caracterización de los puntos factibles del problema del transporte. Puesto que $F_b = \left\{x \in \mathbb{R}_+^{mn} \colon Ax = b\right\}$, no ha sido dificil obtener dicha representación para varios casos, veamos dos de ellos.

CASO 1. $m = 1, n = 2$

En el caso una fuente y dos destinos $b = (b_1, b_2, b_3)$, se satisface $b_1 = b_2 + b_3$, y $x = (x_1, x_2) \in F$ si y sólo si

$$\begin{aligned} x_1 + x_2 &= b_1 \\ x_1 \quad\quad &= b_2 \\ x_2 &= b_3 \\ x_1, x_2 &\geq 0. \end{aligned}$$

Por lo tanto $F_b = \{(b_2, b_3)\}$.

CASO 2. $m = 2, n = 3$

En este caso tenemos 2 fuentes y 3 destinos, por lo que $b = (b_1, b_2, b_3, b_4, b_5)$, el cual satisface la ecuación $b_1 + b_2 = b_3 + b_4 + b_5$. Además, $x = (x_{13}, x_{14}, x_{15}, x_{23}, x_{24}, x_{25}) \in F_b$ si y sólo si

$$\begin{aligned} x_{13} + x_{14} + x_{15} \quad\quad\quad\quad\quad\quad\quad &= b_1 \\ \quad\quad\quad\quad\quad\quad x_{23} + x_{24} + x_{25} &= b_2 \\ x_{13} \quad\quad\quad\quad + x_{23} \quad\quad\quad &= b_3 \\ \quad x_{14} \quad\quad\quad\quad + x_{24} \quad\quad &= b_4 \\ \quad\quad x_{15} \quad\quad\quad\quad + x_{25} &= b_5 \\ x_{13}, x_{14}, x_{15}, x_{23}, x_{24}, x_{25} &\geq 0. \end{aligned}$$

Este sistema tiene infinitas soluciones. Si $x_{24} = t$ y $x_{25} = s$, obtenemos que $x \in F_b$ si y sólo si

$$x = \begin{pmatrix} b_3 - b_2 + t + s \\ b_4 - t \\ b_5 - s \\ b_2 - t - s \\ t \\ s \end{pmatrix}$$

$$0 \leq t \leq b_4, 0 \leq s \leq b_5, b_2 - b_3 \leq t + s \leq b_2.$$

De los casos 1 y 2 deducimos que los puntos factibles del problema del transporte dependen linealmente de las componentes del vector b. Esperamos obtener una representación general para los puntos factibles que nos ayude a demostrar la continuidad o no de la multifunción conjunto factible.

3.3. **Acerca de los puntos extremos.** Debido a la acotación y la cerradura del cunjunto factible F_b, sabemos que este es un polítopo, luego $E_b \neq \varnothing$ y coincide con su conjunto de vértices.

De ahora en adelante consideraremos a la matriz A^0 y al vector b^0 como los originales pero sin el último renglón y llamaremos S_k a una submatriz cuadrada de A^0 de tamaño $m + n - 1$, de tal forma que $A^0 = [S_k, N]$ y donde N representa el bloque restante de A^0, entonces rank $(S_k) = m + n - 1$ y $\det(S_k) \neq 0$. En consecuencia, el sistema de ecuaciones $S_k x = b^0$, tiene una única solución $x_k = S_k^{-1} b^0$. Sea

$$\overline{x}_k = \begin{bmatrix} x_k \\ x_N \end{bmatrix}$$

donde x_N es un vector de ceros de tamaño $mn - (m + n - 1)$, y cuyas componentes ocupan el lugar de las variables que no fueron elegidas para formar la submatriz S_k. Si cada componente de x_k, es mayor o igual que cero, entonces a $\overline{x}_k$ se le llama

una solución básica factible. Es conocido de la literatura, ver por ejemplo [4], que cada solución básica factible es un vértice de F_b.

Sea S_{kl} la matriz S_k a la cual se le ha sustituido la columna $l-ésima$ por el vector b^0, entonces, de acuerdo a la Regla de Cramer:

$$x_{kl} = \frac{\det(S_{kl})}{\det(S_k)} \text{ y } \bar{x}_k = \begin{bmatrix} \frac{\det(S_{kl})}{\det(S_k)} \\ x_N \end{bmatrix}$$

con $l = 1, 2, ..., (m+n-1)$.

Ahora observemos que como S_k es una matriz de ceros y unos con determinante diferente de cero, $\det(S_{kl})$ es una combinación lineal de componentes del vector b, luego, podemos concluir que cada elemento del conjunto de puntos extremos de F_b depende linealmente de las componentes del parámetro b, es decir, dado $b \in B$, con conjunto factible F_b,

$$E_b = \left\{ x_k \in \mathbb{R}^{mn}_{+} \colon x_k = \begin{bmatrix} \frac{\det(S_{kl})}{\det(S_k)} \\ x_N \end{bmatrix} \right\}.$$

Por lo tanto, pequeñas perturbaciones en las componentes de b, generan puntos extremos muy cercanos. Actualmente buscamos más información sobre el comportamiento de estos puntos extremos cuando perturbamos b para deducir la continuidad o no de la multifunción conjunto de puntos extremos.

3.1. Ejemplo. Sea $m = 2, n = 2$, y $b = (b_1, b_2, b_3, b_4)$ un vector en $\mathbb{R}^4_{++}$ el cual satisface $b_1 + b_2 = b_3 + b_4$. La matriz del sistema es:

$$A = \begin{pmatrix} 1,1,0,0 \\ 0,0,1,1 \\ 1,0,1,0 \\ 0,1,0,1 \end{pmatrix}, \; A^0 = \begin{pmatrix} 1,1,0,0 \\ 0,0,1,1 \\ 1,0,1,0 \end{pmatrix}$$

$$S_1 = \begin{pmatrix} 1,1,0 \\ 0,0,1 \\ 1,0,1 \end{pmatrix}, \; S_2 = \begin{pmatrix} 1,1,0 \\ 0,0,1 \\ 1,0,0 \end{pmatrix}, \; S_3 = \begin{pmatrix} 1,0,0 \\ 0,1,1 \\ 1,1,0 \end{pmatrix} \text{ y } S_4 = \begin{pmatrix} 1,0,0 \\ 0,1,1 \\ 0,1,0 \end{pmatrix}$$

$$\bar{x}_1 = \begin{bmatrix} b_3 - b_2 \\ b_4 \\ b_2 \\ 0 \end{bmatrix}, \; \bar{x}_2 = \begin{bmatrix} b_3 \\ b_1 - b_3 \\ 0 \\ b_2 \end{bmatrix}, \; \bar{x}_3 = \begin{bmatrix} b_1 \\ 0 \\ b_3 - b_1 \\ b_4 \end{bmatrix}, \; \bar{x}_4 = \begin{bmatrix} 0 \\ b_1 \\ b_3 \\ b_2 - b_3 \end{bmatrix}$$

Los vectores $\bar{x}_k$ con $k = 1, 2, 3, 4$ son los puntos extremos de F_b si y sólo si, cada una de sus componentes es mayor o igual que cero. Por ejemplo, si $b = [20, 10, 15, 15]'$, tenemos dos vértices de F_b, $\bar{x}_1 = [5, 15, 10, 0]'$ y $\bar{x}_2 = [15, 5, 0, 10]$, los otros dos tienen una componente negativa y por lo tanto quedan desechados.

4. Conclusiones

La caracterización de los puntos factibles y extremos que se ha obtenido hasta ahora es muy importante para conocer estos dos conjuntos e intuir el cambio que sufrirán cuando perturbemos el parámetro b. En el caso del conjunto factible nos hace falta tener una representación general. Sin embargo, para ambos conjuntos de puntos hemos observado que estos guardan una dependencia lineal con las componentes del vector b, por lo que conjeturamos que las multifunciones conjunto factible y conjunto de puntos extremos serán continuas. La demostración de esta conjetura será objeto de un trabajo próximo.

Referencias

[1] Goberna M.A. and López M.A., *Linear Semi-infinite Optimization,* Wiley, Chichester, 1998.
[2] Goberna M.A., López M.A. and Todorov M.I., *On the stability of the feasible set in linear optimization,* Set-Valued Analysis (2001), 75-99.
[3] Goberna M.A., López M.A. and Todorov M.I., *Stability theory for linear inequality systems,* SIAM J. Matrix Anal. Appl., 17 (1996), 730-743.
[4] Bazaraa M., *Programación Lineal y Flujo en Redes,* 2da. edición, Limusa, México, 2005.
[5] Goberna M.A., Jornet V. and Puente R., *Programación Lineal. Teoría, Métodos y Modelos,* McGraw Hill, 2004.
[6] Szware W., *The Stability of the Transportation Problem,* Mathematica, Cluj 4(27), 1962.

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y Río Verde, Ciudad Universitaria
San Manuel, CP 72570, Puebla Pue., México.
sgomez@fcfm.buap.mx
lhernan@fcfm.buap.mx
alancho06@hotmail.com

A CONDITION TO DECIDE WHEN THE MEMBERS OF A CLASS OF CONTINUA ARE C-DETERMINED

DAVID HERRERA-CARRASCO
FERNANDO MACÍAS-ROMERO

Abstract. For a metric continuum X, let $C(X)$ denote the hyperspace of subcontinua of X. The members of a class Γ of continua are said to be C-determined provided that if $X, Y \in \Gamma$ and $C(X) \approx C(Y)$, then $X \approx Y$. In this paper we prove the next result: Let Γ be a class of continua such that for each $X \in \Gamma$, we have $cl_{C(X)}(\Omega(X)) \approx X$, then the members of the class Γ are C-determined. Moreover we obtain some applications of this result

1. INTRODUCTION

[1]A *continuum* is a nonempty, compact, connected, metric space. We use the term *subcontinuum* when referring to a continuum as a subset of a given topological space.

Throughout this paper the letter I represents the closed unit interval $[0, 1]$ in the real line $\mathbb{R}$. The letter $\mathbb{N}$ denotes the set of positive integers. The letter X denotes a continuum. The *hyperspace of subcontinua* of X is denoted by $C(X)$ and the *hyperspace of singletons* of X by $F_1(X)$. The hyperspaces $C(X)$ and $F_1(X)$ are metrized by the *Hausdorff metric*. If two continua X and Y are homeomorphic, we write $X \approx Y$. Note that $X \approx F_1(X)$.

For $n \in \mathbb{N}$, an n-cell is a topological space that is homeomorphic to cartesian product $I^n = \prod_{j=1}^{n} I_j$, where $I_j = I$ and we tacitly assume that the cartesian products have the Tychonoff (product) topology. A 1-cell is called an *arc*; an *end point* of an arc, A, is either one of the two points of A that are the image of the end points of I under any homeomorphism of I onto A. Whenever we say A is an arc $[p, q]$ we

[1]*2000 Mathematics Subject Classification. Primary: 54B20; Secondary: 54D35, 54F15, 54F50.*

Key words and phrases. continuum, hyperspace, contour, C-determined, n-cell, n-od.

This is the final form of the paper

mean A is an arc with end points p and q. A *finite graph* is a continuum that can be written as the union of finitely many arcs, any pair of which can intersect in at most one or both of their end points.

It is clear that if $X \approx Y$, then $C(X) \approx C(Y)$. The converse is not true. For example, if $S^1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}$, then $C(I)$ and $C(S^1)$ are homeomorphic to a 2-cell ([8, 3.1 and 3.2]) However there exist classes of continua for which we can prove that if $C(X) \approx C(Y)$, then $X \approx Y$. With respect to this, S. B. Nadler, Jr. introduce in [10, Definition 0.61] the following concept: The members of a class Γ of continua are said to be *C-determined* provided that if $X, Y \in \Gamma$ and $C(X) \approx C(Y)$, then $X \approx Y$. The members of the following classes of continua are known to be C-determined:

1. *Hereditarily indecomposable* continua ([10, Theorem 0.60]).
2. *Smooth fans* ([3, Corollary 3.3]).
3. *Indecomposable* continua such that all their proper *nondegenerate* subcontinua are arcs ([9, Theorem 3]).
4. *Compactifications* of the interval $[0, \infty)$ ([1]).
5. Finite graphs G such that G is not an arc and G is not a circle ([2]). Here we give a different proof of the same fact.

It is also known that the members of the following classes of continua are not C-determined:

1. *Chainable* continua ([5, Theorem 2]).
2. *Fans* ([6]).

The following is an open problem: Are the members of the class of *circle-like* continua, C-determined? ([10, Question 0.62]).

The main results we obtain in this paper are:

(1) Theorem 3.3. *Let a class of continua Γ such that for each $X \in \Gamma$, we have $X \approx cl_{C(X)}(\Omega(X))$, then the members of the class Γ are C-determined.*

(2) Theorem 4.6. *The members of the class G, are C-determined.*

(3) Corollary 4.9. *If $\Gamma = \mathfrak{D} \cup \mathfrak{G}$, then the members of the class Γ are C-determined.*

(4) Theorem 4.10. *If Γ is the union of classes of continua such that for each $X \in \Gamma$, we have $cl(\Omega(X)) \approx X$. Then the members of the class Γ are C-determined.*

2. PRELIMINARIES

For $n \in \mathbb{N}$, an *n-manifold* is a metric separable space such that each point of which has a closed neighborhood that is homeomorphic to I^n. The *manifold interior* of an n-manifold K (denoted by $IntM(K)$) consists of all those points of K that have neighborhoods in K that are homeomorphic to Euclidean n-space $\mathbb{R}^n$. The *manifold boundary* of an n-manifold K (denoted by ∂K) consists of all the points of K that are not in the manifold interior of M. We note that if h is a homeomorphism of an n-manifold K onto K, then $h(IntM(K)) = IntM(K)$ and $h(\partial K) = \partial K$. A topological space that is homeomorphic to ∂I^2 is called a *simple closed curve*.

Let T be a topological space, $H \subset T$. Then, the *boundary* of H (in T) is $Bd_T(H) = cl_T(H) - int_T(H)$, where we use the symbols $cl_T(H)$ and $int_T(H)$ to denote the closure and the interior of H in T, respectively.

The following lemma is significative to give the Definition 3.1 of the paper.

: Lemma 2.1 (*a*) [11, 19.33] *If I^n is an $n-$cell in $\mathbb{R}^n$, then $\partial I^n = Bd_{\mathbb{R}^n}(I^n)$ (the boundary manifold of each n-cell is called a $(n-1)$-sphere and the interior manifold of every n-cell is homeomorphic to $\mathbb{R}^n$).*
(*b*) [11, 19.34] *If V and W are n-cells such that $V \subset W$, then $V - \partial V$ is an open set in W. In particular, if V is a $2-$cell and $h : I^2 \to V$ is a homeomorphism, then $h(Bd_{\mathbb{R}^n}(I^2)) = h(\partial I^2) = \partial V$ (is a simple closed curve).*

3. CONDITION FOR C-DETERMINED

We define the next subset of a hyperspace:

: Definition 3.1. Let Z be a continuum, V be a 2-cell in Z and $h : I^2 \to V$ be a homeomorphism, then the contour of V is the set $Contour(V) = \{y \in V : y = h(x)$ for some $x \in Bd_{\mathbb{R}^n}(I^2)\}$. Hence $Contour(V) = h(Bd_{\mathbb{R}^2}(I^2))$. Given a continuum X, we define $\Omega(X) = \{A \in C(X) :$ there exists a neighborhood V of A in $C(X)$ such that V is a 2-cell and $A \in Contour(V)\}$.

By a *simple triod* we mean a continuum homeomorphic to the union of three arcs each two of which intersect at a single common end point of the three arcs.

: Example 3.2. (a) Let $J = [a, b]$. It is known that $C(J)$ is a 2-cell. The contour of $C(J)$ is
$Contour(C(J)) = C(\{a\}, J) \cup C(\{b\}, J) \cup F_1(J)$ ([7, 5.1]). Hence $S^1 \approx cl_{C(J)}(\Omega(J))$. So, $cl_{C(J)}(\Omega(J)) \not\approx J$.
(b) Let T be a simple triod, it is known that $C(T)$ is homeomorphic to the space pictured in the next figure ([8, Figure 17]). In the hyperspace $C(T)$, there are three triangles that intersect with the cube in three different sides. The set $cl_{C(T)}(\Omega(T))$ is the union of the sides of the triangles mentioned, except for the side in which it intersects with the cube. Observes that $cl_{C(T)}(\Omega(T) \approx T$.

The following theorem is our principal result.

: Theorem 3.3. *Let Γ be a class of continua such that for each $X \in \Gamma$, we have $cl(\Omega(X)) \approx X$. Then the members of the class Γ are C-determined.*

Proof: Let $X, Y \in \Gamma$ such that $C(X) \approx C(Y)$. Let $h : C(X) \to C(Y)$ be a homeomorphism. Then $h(\Omega(X)) = \Omega(Y)$. Hence, $h(cl(\Omega(X))) = cl(\Omega(Y))$. Therefore, as $X \approx cl(\Omega(X))$ and $cl(\Omega(Y)) \approx Y$, we obtain the $X \approx Y$. □

4. CONVENTIONS AND APPLICATIONS

Let (Y, τ) be a topological space, let $A \subset Y$, and let β be a cardinal number. We say that A is of *order less than or equal to* β in Y, written $ord(A, Y) \leq \beta$, provided that for each $U \in \tau$ such that $A \subset U$, there exists $V \in \tau$ such that $A \subset V \subset U$ and $|Bd_Y(V)| \leq \beta$. We say that A is of *order* β in Y, written $ord(A, Y) = \beta$, provided that $ord(A, Y) \leq \beta$ and $ord(A, Y) \not\leq \alpha$ for any cardinal number $\alpha < \beta$.

If $A = \{p\}$, then we write $ord(p, Y)$ instead of $ord(\{p\}, Y)$ and write p is of order n instead of writing $\{p\}$ is of order n.

Let G be a finite graph. The points of order 1 in G are called *end points* of G (this is obviously a generalization of the notion of an end point of an arc defined in the introduction); the set of all end points of G is denoted by $E(G)$. Points of order 2 are called *ordinary points* of G; the set of all ordinary point of G is denoted by $O(G)$. Points of order at least 3 are called *ramification points* of G; the set of all ramification points of G is denoted by $R(G)$. Therefore, if G is a finite graph such that G is not an arc, then $G = O(G) \cup R(G) \cup E(G)$.

If G is a finite graph, in G are defined *edges* (arcs or simple closed curves of G) and *vertices*. The vertices of G are the end points of

the edges of G. We are interested in distinguishing the ramifications points of the graph G from the rest of the points, so we assume the each vertex of a graph G, different of a simple closed curve, is either an end point of G or a ramification point of G. With this restriction the two end points of an edge of G may coincide and such an edge is a simple closed curve. This kind of edges will be called *loops*. Thus the edges of G are arcs or simple closed curves, and in G there are only three kind of edges namely: loops, edges that contain some end points of G and edges joined ramification points.

Let G be a finite graph such that G is not an arc. An arc $[p, q]$ (edge of G) is an

(a) *internal arc* of G, if $[p, q] \cap R(X) = \{p, q\}$,

(b) *external arc* of G if $[p, q] \cap R(X) = \{p\}$ and $[p, q] \cap E(X) = \{q\}$.

We assume that the metric d in G is the metric of arc length and each edge of G has length equal to one.

For $n \in \mathbb{N}$ an *n-od* in a continuum X is an element $B \in C(X)$ for which there exists $A \in C(B)$ (called *heart*) such that $B - A$ has at least n components. Note that if Y contains an n-od, then contains an m-od, for each $m \leq n$.

Let Y be a metric space with metric d. If $a \in Y$ and $\varepsilon > 0$, then $B_Y(a, \varepsilon) = \{x \in Y : d(a, x) < \varepsilon\}$ is the open *d-ball* in Y with radius ε and center a.

Lemma 4.1. [1, Lemma 8] *Let $K \in C(X)$ and let T be an n-od in X such that, for some $\varepsilon > 0$, $T \in B_{C(X)}(K, \frac{\varepsilon}{2})$. Then there is an n-cell Γ in $C(X)$ such that $T \in \Gamma \subset B_{C(X)}(K, \varepsilon)$.*

Now, we will to establish some lemmas to prove the Theorem 4.5.

Lemma 4.2. *Let G be a finite graph such that G is not an arc and $A \in C(G)$. If $A \in \Omega(G)$, then $A \cap R(G) = \emptyset$.*

Proof: Let $A \in \Omega(G)$ such that $A \cap R(G) \neq \emptyset$. Let V be a neighborhood of A in $C(G)$ such that V is a 2-cell. Then there exist $\varepsilon > 0$ such that $B_{C(G)}(A, \varepsilon) \subset V$. As $A \cap R(G) \neq \emptyset$ we can construct an n-od T in G such that $T \in B_{C(G)}(A, \frac{\varepsilon}{2})$ (T is constructed from A and its heart is a ramification point). By the Lemma 4.1, there is a n-cell, Γ, such that $T \in \Gamma \subset B_{C(G)}(A, \varepsilon) \subset V$, and it is a contradiction. Therefore, $A \cap R(G) = \emptyset$. $\square$

Lemma 4.3. *Let G be a finite graph such that G is not an arc. Then $\Omega(G) = \{\{p\} : p \in G - R(G)\} \cup \{[t, e] \subset G - R(G) : t \in O(G)$ and $e \in E(G)\}$.*

Proof: Suppose that $A \in \Omega(G)$. By the Lemma 4.2, $A \cap R(G) = \emptyset$. Hence A is a singleton or A is an arc. If A is a singleton, for example, $A = \{p\}$, then $p \in G - R(G)$. If A is an arc, then one of the following cases holds.
(1) $A = [a, b] \subset [p, q]$, where $[p, q]$ is an internal arc.

(2) $A = [a, b] \subset [p, q]$, where $[p, q]$ is an external arc and $a, b \notin E(G)$.
(3) $A = [a, b] \subset [p, q]$, where $[p, q]$ is a loop (here $p = q$).
(4) $A = [a, b] \subset [p, q]$, where $[p, q]$ is an external arc, $q \in E(G)$, $b = q$ and $a \in O(G)$.

We now consider the cases (1)-(3). Let V be a neighborhood of A in $C(G)$ such that V is an 2-cell. Therefore, there is $\varepsilon > 0$ such that $B_{C(G)}(A, \varepsilon) \subset V$ and $B_{C(G)}(A, \varepsilon) \subsetneq C([p, q])$. Thus there is an arc $J = [c, d] \in B_{C(G)}(A, \varepsilon)$ such that $A \subsetneq J$ where a and b are neither of p nor q. There is $\varepsilon_1 > 0$ such that $B_{C(G)}(A, \varepsilon_1) \subsetneq C([c, d])$, $B_{C(G)}(A, \varepsilon_1) \cap C(\{c\}, J) = \emptyset$, $B_{C(G)}(A, \varepsilon_1) \cap C(\{d\}, J) = \emptyset$ and $B_{C(G)}(A, \varepsilon_1) \cap F_1(J) = \emptyset$. Thus, by (a) of Example 3.2, we have that $B_{C(G)}(A, \varepsilon_1) \cap Contour(C([c, d])) = \emptyset$. As $C([c, d])$ is an 2-cell and $C([c, d]) \subset V$, by (b) of Lemma 2.1, we obtain that $C([c, d]) - Contour(C([c, d]))$ is an open set in V. Note that $A \in C([c, d]) - Contour(C([c, d]))$. Thus $A \in int(V)$. Therefore we obtain $A \notin Contour(V)$. By the definition of $\Omega(G)$, we have that $A \notin \Omega(G)$, a contradiction to supposition.

This proves that if $A \in \Omega(G)$ and A is an arc, then $A = [a, b] \subset G - R(G)$ with $a \in O(G)$ and $b \in E(G)$ (case (4)).

Let
$\mathfrak{L} = \{\{p\} : p \in G - R(G)\} \cup \{[t, e] \subset G - R(G) : t \in O(G)$ and $e \in E(G)\}$. We next prove that $\Omega(G) \subset \mathfrak{L}$. Suppose first that $A \in \mathfrak{L}$. If $A = \{p\}$, where $p \in G - R(G)$. We obtain three cases: $\{p\} \in [r, q]$, where $[r, q]$ is an internal arc; or $\{p\} \in [t, e]$, where $[t, e]$ is an external arc; and if p is in a loop S of the G. In the first two cases, either $C([r, q])$ or $C([t, e])$ is a neighborhood of $\{p\}$ and by (a) of Example 3.2, we have that $\{p\} \in Contour(C([r, q]))$ or $\{p\} \in Contour(C([t, e]))$, hence $\{p\} \in \Omega(G)$. In the third case let an arc J contained in S such that $p \in J$ (p is not end point of J) and $J \cap R(G) = \emptyset$. Then $C(J)$ is a neighborhood of p such that is an 2-cell and $\{p\} \in F_1(J)$. Therefore $\{p\} \in Contour(C(J))$.

Now suppose that $A = [t, e]$ with $t \in O(G)$, $e \in E_2(G)$ and $[t, e] \subset [r, e] = J$ where J is an external arc. By the (a) of Example 3.2, the contour of the 2-cell $C(J)$, is the union of the sets $C(\{r\}, J)$, $C(\{e\}, J)$ and $F_1(J)$. In this cases $A \in C(\{e\}, J)$ and $C(J)$ is a neighborhood of A. By this reason $A = [t, e] \in \Omega(G)$. □

Lemma 4.4. *Let G be a finite graph such that G is not an arc. Then*

$cl_{C(G)}(\Omega(G)) = F_1(G) \bigcup \{[r, e] : e \in E(G)$ *and* $[r, e] \cap R(G) = \{r\}\}$
$\bigcup \{[r, e] : e \in E(G)$ *and* $r \in O(G)\}$.

Proof: If $p \in G$, then $p = \lim_{n\to\infty} p_n$, where $p_n \in O(G)$. Hence $\{\{p_n\} : n \in \mathbb{N}\} \subset \Omega(G)$, and therefore $\{p\} \in cl_{C(G)}(\Omega(G))$, i.e., $F_1(G) \subset cl_{C(G)}(\Omega(G))$.

Let $\mathfrak{M} = \{[r, e] : e \in E(G)$ *and* $[r, e] \cap R(G) = \{r\}\}$
$\bigcup \{[r, e] : e \in E(G)$ *and* $r \in O(G)\}$. Let $[r, e] \in \mathfrak{M}$ with $r \in R(G)$ and $e \in E(G)$, then there is $\{o_n\}_{n=1}^{\infty}$ in $O(G) \cap [r, e]$ such that $\lim_{n\to\infty} o_n = r$. Therefore $\lim_{n\to\infty} [o_n, e] = [r, e]$. Since $[o_n, e] \in \Omega(G)$, we have that $[r, e] \in cl_{C(G)}(\Omega(G))$. This proves that $\mathfrak{M} \subset cl_{C(G)}(\Omega(G))$. Therefore $F_1(G) \cup \mathfrak{M} \subset cl_{C(G)}(\Omega(G))$.

To prove the other inclusion let now $A \in cl_{C(G)}(\Omega(G))$ and $A \notin F_1(G)$. Then $A = \lim_{n\to\infty} [t_n, e_n]$, where $t_n \in O(G)$ and $e_n \in E(G)$. Since the number of arcs is finite, there are $N \in \mathbb{N}$, $e \in E(G)$ such that if $n > N$, then $e_n = e$. Therefore $A = \lim_{n\to\infty} [t_n, e]$. But $[t_n, e] \cap R(G) = \emptyset$, for each $n \in \mathbb{N}$, then $A = [r, e]$, where $\lim_{n\to\infty} t_n = r$ with $r \in R(G) \cup O(G)$. Hence $A = [r, e] \in \mathfrak{M}$. This completes the proof of the Lemma. □

Theorem 4.5. *If G is a finite graph such that G is not an arc, then $cl_{C(G)}(\Omega(G)) \approx G$.*

Proof: Let us first find an homeomorphism from $cl_{C(G)}(\Omega(G))$ onto G. To do that, let $A \in \Omega(G)$. Since $A \cap R(G) = \emptyset$, then A is contained in an internal arc or A is contained in an external arc or a A is contained in loop:

(1) If A is contained in an internal arc $[a, b]$, it follows from Lemma 4.3 that $A = \{p\}$ with $p \in (a, b)$.

(2) If A is contained in a loop $[a, b]$, it follows from Lemma 4.3 that $A = \{p\}$ with $p \in (a, b)$.

(3) If A is contained in an external arc $[r, e]$, it follows from Lemma 4.3 that $A = \{q\}$ with $q \in [r, e] - R(G)$ or $A = [q, e]$, with $q \in (r, e)$.

Let $[r,e]$ be an external arc such that $[r,e] \cap R(G) = \{r\}$ and $[r,e] \cap E(G) = \{e\}$, where $s \in (r,e)$. For each external arc, there are two homeomorphisms, $f_1 : [r,e] \to [r,s]$ such that $f_1(r) = r$ and $f_1(e) = s$ and $f_2 : [r,e] \to [s,e]$ such that $f_2(r) = e$ and $f_2(e) = s$.
Now we define the function $g : \Omega(G) \to G$ by the following rule:
If A satisfies condition (1) and (2), we have that $g(A) = g(\{p\}) = p$.
If A satisfies condition (3), $g(A) = f_1(q)$ if $A = \{q\}$ and $g(A) = f_2(q)$ if $A = [q,e]$, where $[q,e] \subset [r,e]$. To see that g is a continuous function, we consider two cases (see below Figure). Suppose that $p \in (a,b)$, where $[a,b]$ is an internal arc or a loop. Then $p = g(\{p\})$. Let $\varepsilon > 0$ and $\delta < \min\{d(p,a), d(p,b), \varepsilon\}$ such that $B_{C(G)}(\{p\}, \delta) \subset C([a,b])$. We obtain that $g(B_{C(G)}(\{p\}, \delta) \cap \Omega(G)) \subset B_G(p,\varepsilon)$. Therefore g is a continuous function in $\{p\}$.
Now suppose that $A = \{q\}$, with $q \in [r,e] - R(G)$, where $[r,e]$ is an external arc. In this case $g(\{q\}) = f_1(q)$. We will prove that g is a continuous function in $\{q\}$. Let $\varepsilon > 0$, take the ball $B_G(f_1(q),\varepsilon)$. By continuity of f_1, there is $\delta > 0$ such that $f_1(B_G(q,\delta) \cap [r,e]) \subset B_G(f_1(q),\varepsilon)$. For this reason, $g(B_{C(G)}(\{q\}, \delta) \cap \Omega(G)) \subset B_G(f_1(q),\varepsilon)$. Therefore g is a continuous function in $\{q\}$.

We will prove that g is an injective function. Suppose that $g(A) = g(D)$. If $A = \{p\}$ with $p \in (a,b)$, where $[a,b]$ is an internal arc, then $D = \{s\}$ and $s \in (a,b)$. In fact, if $D = \{r\}$, with r a point that is not in (a,b), then $g(A) \neq g(D)$ (because these imagines will be in different arcs). And if $D = [r,e]$, we have that $g(D)$ is in an external arc. Hence $g(D) \notin (a,b)$. Since $f_1(\{p\}) = g(A) = g(D) = f_1(\{s\})$ and f_1 is injective, we have that $\{p\} = \{s\}$. The other cases are proved similarly.

Now we extend g to a continuous function $\overline{g}$ from $cl_{C(G)}(\Omega(G))$ to G. We use the relation in Lemma 4.4 to define the function $\overline{g}$ at the points of $cl_{C(G)}(\Omega(G)) - \Omega(G)$: For $A \in cl_{C(G)}(\Omega(G)) - \Omega(G)$ we consider three cases.
(j) If $A = \{r\}$ and $r \in R(G)$, we define $\overline{g}(A) = \overline{g}(\{r\}) = r$;
(jj) If $A = [r,e]$ and $[r,e]$ is an external arc, $e \in E(G)$, we define $\overline{g}(A) = \overline{g}([r,e]) = e$.

Notice that the function $\overline{g}$ is bijective and continuous. Therefore $\overline{g}$ is an homeomorphism. □

If $\mathfrak{G}$ is the class of graphs G such that G is not an arc, by an application of Theorem 4.5, we obtain the following.

Theorem 4.6. *The members of the class $\mathfrak{G}$, are C-determined.*

Proof: Let $G \in \mathfrak{G}$. By the Theorem 4.5, we obtain $cl_{C(G)}(\Omega(G)) \approx G$. Next, by Theorem 3.3, we obtain the result. □

D. Herrera-Carrasco considered the class $\mathfrak{D}$ of *dendrites* X such that $cl_X(E(X)) = E(X)$ and X is not an arc. He has the next result.

Theorem 4.7 [4, Theorem] *If $X \in \mathfrak{D}$, then $cl_{C(X)}(\Omega(X)) \approx X$.*

Corollary 4.8. *The members of the class $\mathfrak{D}$ are C-determined.*

Corollary 4.9. *If $\Gamma = \mathfrak{D} \cup \mathfrak{G}$, then the members of the class Γ are C-determined.*

In fact, the previous result can be generalized:

Theorem 4.10. *If Γ is the union of classes of continua such that for each $X \in \Gamma$, we have $cl(\Omega(X)) \approx X$. Then the members of the class Γ are C-determined.*

References

[1] G. Acosta, *Continua with unique hyperspace*, Lecture notes in pure and applied mathematics 230, 33-49, Marcel Dekker, Inc., 2002.

[2] R. Duda, *On the hyperspace of subcontinua of a finite graph, I*, Fund. Math. 62 (1968), 265-286.

[3] C. Eberhart and S. B. Nadler, Jr., *Hyperspaces of cones and fans,* Proc. Amer. Math. Soc. 77 (1979), 279-288.

[4] D. Herrera-Carrasco, *Dendrite whose hyperspace of subcontinua is unique*, preprint.

[5] A. Illanes, *Chainable continua are not C-determined*, Topology Appl. 98 (1999), 211-216.

[6] A. Illanes, *Fans are not C-determined*, Colloq. Math. 81 (2) (1999), 299-308.

[7] A. Illanes, S. B. Nadler, Jr., *Hyperspaces: Fundamentals and Recent Advances*, Monographs and Textbooks in Pure and Applied Math., 216, Marcel Dekker, Inc., New York, 1999.

[8] A. Illanes, *Hiperespacios de Continuos,* Aportaciones Matemáticas, Serie Textos 28, Sociedad Matemática Mexicana, 2004.

[9] S. Macías, *On C-determined continua*, Glasnik Mat., 32(52) (1997), 259-262.

[10] S. B. Nadler, Jr., *Hyperspaces Of Sets*, Marcel Dekker, 1978.

[11] S. B. Nadler, Jr., *Dimension Theory: An introduction with exercises*, Aportaciones Matemáticas, Serie Textos 18, Sociedad Matemática Mexicana, 2002.

Facultad de Ciencias Físico Matemáticas, BUAP.
Av. San Claudio y Río Verde, Ciudad Universitaria.
San Manuel. C.P. 72570, Puebla, Pue. México.
dherrera@fcfm.buap.mx
fmacias@fcfm.buap.mx

DESIGUALDEDES

ARMANDO MARTÍNEZ GARCÍA
BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

Resumen. La idea es trabajar, ciertas desigualdades que nos llevan directamente a la definición de límite de una función en un punto.

1. Introducción

El objetivo de estas notas es presentarle al alumno de nuevo ingreso a una Facultad de Ciencias una introducción al concepto de **limite de una función**, por medio de la solución de algunas desigualdades que nos permita alcanzar este fin.

Se ha de mencionar que el concepto de **limite de una función**, se presenta en el primer curso de Cálculo que se lleva en dichas facultades, por lo que debemos de tomar en cuenta que este no es un material nuevo, mas si su forma de presentarlo.

Esta forma de presentar este material no pretende resolver todos los problemas concernientes a este tema, sino darle al alumno una idea geometrica de este, para poder entender mejor lo que es presisamente el **limite de una función**.

2. Desigualdades

Al conjunto de los números reales, lo denotaremos como $\mathbb{R}$.

Sean $a, b, c \in \mathbb{R}$ recordemos que una ecuación de la forma

$$ax + b = c \text{ con } a \neq 0$$

tiene solución si existe $x_0 \in \mathbb{R}$ tal que,

$$ax_0 + b = c.$$

En este caso

$$x_0 = c - b/a.$$

En forma analoga dados $a, b, c \in \mathbb{R}$ una ecuación de la forma

$$ax^2 + bx + c = 0 \text{ con } b^2 - 4ac \geq 0$$

tiene solución si existe $x_0 \in \mathbb{R}$ tal que,

$$ax_0^2 + bx_0 + c = 0.$$

En este caso tenemos dos soluciones diferentes o iguales las cuales seran;

$$x_0 = (-b + (b^2 - 4ac)^{1/2})/2a \text{ y } x_0 = (-b - (b^2 - 4ac)^{1/2})/2a$$
$$\text{si } b^2 - 4ac > 0$$

y

$$x_0 = -b/a \text{ si } b^2 - 4ac = 0.$$

En forma analoga dados $a, b, c \in \mathbb{R}$ una desigualdad de la forma

$$ax + b < c \text{ con } a \neq 0$$

tiene solución si existe $x_0 \in \mathbb{R}$ tal que,

$$ax_0 + b < c$$

En este caso

$$x_0 < (c-b)/a \text{ si } a > 0 \text{ y } x_0 > (c-b)/a \text{ si } a < 0$$

es decir la solución es el conjunto $S \subset \mathbb{R}$ con

$$S = \{x \in \mathbb{R} : x < (c-b)/a\} \text{ si } a > 0 \text{ y } S = \{x \in \mathbb{R} : x > (c-b)/a\} \text{ si } a < 0.$$

Lo cual significa que

$$\text{si } x \in S \text{ entonces } ax + b < c.$$

Los siguientes resultados nos permitiran resolver otro tipo de desigualdades.

Recordemos que dado $a \in \mathbb{R}$ al producto de $a.a$ lo denotaremos como a^2, es decir $a^2 = a.a$.

2.1. Proposición. Sean $a > 0$ y $b > 0$. Entonces

$$a < b \text{ si y sólo si } a^2 < b^2.$$

Demostración. Si $a < b$ como $a > 0$ y $b > 0$ se sigue que, $a^2 < ab$ y $ab < b^2$ por lo tanto, $a^2 < b^2$.

Ahora si $a = b$ entonces, $a^2 = b^2$ lo cual no puede ser.

Analogamente si $a > b$ por la primera parte tendriamos que, $a^2 > b^2$ lo cual tampoco puede ser por lo tanto $a < b$. □

2.2. Definición. Sean $a \geq 0$ y $b \geq 0$. a es la raiz cuadrada de b si, $a^2 = b$.

En caso de que a sea la raiz cuadrada de b lo escribiremos como

$$a = b^{1/2}.$$

Es claro que si $a^2 = b$ es decir $a = b^{1/2}$ entonces

$$(b^{1/2})^2 = b.$$

2.3. Proposición. Sea $b > 0$. Entonces

1) $a^2 < b$ si y sólo si $-(b^{1/2}) < a < b^{1/2}$.

2) $b < a^2$ si y sólo si $a < -(b^{1/2})$ o $b^{1/2} < a$.

Demostración. Es claro que, si $a = 0$, (1) se satisface.

Ahora si $a > 0$ y $a^2 < b$. Como $b > 0$, $b = (b^{1/2})^2$ se sigue que, $a^2 < (b^{1/2})^2$ de donde aplicando la Proposición anterior tenemos que, $a < b^{1/2}$ por lo tanto

$$-(b^{1/2}) < a < b^{1/2}.$$

Analogamente si $a < 0$ y $a^2 < b$. Como $b > 0$, $b = (b^{1/2})^2$ y $(-a)^2 = a^2$ se sigue que, $(-a)^2 < (b^{1/2})^2$ de donde aplicando la Proposición anterior tenemos que $-a < b^{1/2}$ de donde se sigue que,

$$-(b^{1/2}) < a < b^{1/2}.$$

Ahora si $-(b^{1/2}) < a < b^{1/2}$ y $a > 0$ se sigue que, $0 < a < b^{1/2}$ de donde

$$a^2 < b.$$

Analogamente si $-(b^{1/2}) < a < b^{1/2}$ y $a < 0$ se sigue que, $0 < -a < b^{1/2}$ de donde

$$a^2 < b.$$

En forma similar se tiene el inciso (2). □

2.4. Corolario. Sean $a > 0$ y $b > 0$. Entonces

$$a < b \text{ si y sólo si } a^{1/2} < b^{1/2}.$$

Aplicando la Proposición anterior resolver las siguientes desigualdades.

Ejemplo 1. Encontrar la solución de la desigualdad $4x^2 + 6 < 22$.
Es claro que

$$4x^2 + 6 < 22 \iff 4x^2 < 22 - 6 \iff 4x^2 < 16 \iff 4x^2 < 16 \iff 4x^2 < 4 \iff x^2 < 4$$

de donde se sigue que,

$$-2 < x < 2.$$

Ejemplo 2. Encontrar la solución de la desigualdad $x^2 - 4x < 12$.
Es claro que $x^2 - 4x = (x-2)^2 - 4$ por lo tanto

$$x^2 - 4x < 12 \iff (x-2)^2 - 4 < 12 \iff (x-2)^2 < 16 \iff -4 < x - 2 < 4$$

de donde se sigue que,

$$-2 < x < 6.$$

3. Intervalos

3.1. Definición. Dados $a, b \in \mathbb{R}$ con $a < b$ el intervalo abierto a, b el cual denotaremos como (a, b) es el conjunto:

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}.$$

3.2. Definición. Dados $a, b \in \mathbb{R}$ con $a < b$ el intervalo cerrado a, b el cual denotaremos como $[a, b]$ es el conjunto:

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}.$$

3.3. Definición. Dado $a, \in \mathbb{R}$ el intervalo semi abierto a, el cual denotaremos como (a, ∞) es el conjunto:

$$(a, \infty) = \{x \in \mathbb{R} : a < x\}.$$

3.4. Definición. Dado $b \in \mathbb{R}$ el intervalo semi abierto b el cual denotaremos como (∞, b) es el conjunto:

$$(\infty, b) = \{x \in \mathbb{R} : x < b\}.$$

Es claro que si $a < b$ y $c = (a + b)/2$ entonces, $b - c = c - a$ es decir la distancia de b a c es igual a la distancia de c a a, por tal motivo en los intervalos (a, b) y $[a, b]$ a c se le llama el centro del intervalo y a $b - c$ se le llama el radio del intervalo.

3.5. Definición. Sean $x_0 \in \mathbb{R}$ y $r > 0$. El intervalo abierto con centro en x_0 y radio r el cual denotaremos como $(x_0 - r, x_0 + r)$ es el conjunto:

$$(x_0 - r, x_0 + r) = \{x \in \mathbb{R} : x_0 - r < x < x_0 + r\}$$

Es claro que si $x \in (x_0 - r, x_0 + r)$ entonces, la distancia de x a x_0 es menor que r.

Observación. Para conciderar un intervalo abierto con centro en un punto dado $x_0 \in \mathbb{R}$, es suficiente dar su radio el cual es un número real $r > 0$.

Por lo que en el Ejemplo (1) la solución de la desigualdad $4x^2 + 6 < 22$ son las $x \in (-2, 2)$ es decir, son las x que estan en el intervalo abierto con centro en 0 y radio 2.

Analogamente en Ejemplo (2) la solución de la desigualdad $x^2 - 4x < 12$ son las $x \in (-2, 6)$ es decir son todas las x que estan en el intervalo abierto con centro en 2 y radio 4.

4. Valor Absoluto

4.1. Definición. Sea $x \in \mathbb{R}$ el valor absoluto de x el cual denotamos como $|x|$ es

Observemos que

$$\text{para todo } x \in \mathbb{R} \text{ se tiene que } |x| \geq 0$$

y que

$$|x| = 0 \Longleftrightarrow x = 0$$

lo cual nos permite pensar a

$$|x|$$

como la distancia de x a 0 independientemente si x es positivo o negativo.

4.2. Proposición. Sea $x \in \mathbb{R}$. Entonces

$$|x| = (x^2)^{1/2}.$$

Demostración. Para esto sera suficiente ver que $|x|^2 = x^2$.

Como $|x|^2 = |x||x| = |x^2| = x^2$. □

4.3. Teorema. Sea $x \in \mathbb{R}$ y $r > 0$. Entoces

$$|x| < r \text{ si y sólo si } -r < x < r.$$

Demostración. Es claro que si $x = 0$ la afirmación se satisface.

Ahora si $|x| < r$ y $x > 0$ entonces $|x| = x$ de donde, se sigue que, $x < r$ y como $-r < 0$ y $0 < x$ se tiene que,

$$-r < x < r.$$

Analogamente si $|x| < r$ y $x < 0$ entonces $|x| = -x$ de donde, se sigue que, $-x < r$ es decir $-r < x$ y como $x < 0$ y $0 < r$ tenemos que,

$$-r < x < r.$$

Ahora si $-r < x < r$ y $x > 0$ entonces $|x| = x$ y $x < r$ de donde, se sigue que,

$$|x| < r.$$

Analogamente si $-r < x < r$ y $x < 0$ entonces $|x| = -x$ y $-r < x$ de donde, se sigue que,

$$|x| < r.$$

□

Es decir

$$|x| < r \Longleftrightarrow -r < x < r \Longleftrightarrow x \in (-r, r).$$

Por lo tanto dado $x_0 \in \mathbb{R}$ y $r > 0$ la desigualdad

$$|x - x_0| < r$$

puede ser leida de las siguientes formas siendo todas ellas equivalentes entre si:

i) Valor absoluto de x menos x_0 menor que r.

ii) La distancia de x a x_0 menor que r.

iii) x esta en el intervalo abierto con centro en x_0 y radio r.

Observación. Por lo que cuando se pida resolver una desigualdad de la forma $|x - x_0| < r$ esto es equivalente segun el inciso

(ii) a encontrar $S \subset \mathbb{R}$ tal que si $x \in S$ entonces la distancia de x a x_0 sea menor que r, o equivalentemente segun el inciso

(iii) a encontrar $S \subset \mathbb{R}$ tal que si $x \in S$ entonces x este en el intervalo abierto con centro en x_0 y radio r.

Ejemplo 3. Encontrar todas las $x \in \mathbb{R}$ tales que la distancia de $4x$ a 2 sea menor que $1/10$.

Por la observación anterior esto es equivalente a encontrar $S \subset \mathbb{R}$ tal que

$$\text{si } x \in S \text{ entonces } |4x - 2| < 1/10.$$

Aplicando el Teorema anterior tenemos que

$|4x - 2| < 1/10 \iff -1/10 < 4x - 2 < 1/10 \iff 2 - 1/10 < 4x < 2 + 1/10 \iff 19/10 < 4x < 21/10 \iff 19/40 < x < 21/40$.

Po lo tanto

$$\text{si } x \in (19/40, 21/40) \Rightarrow |4x - 2| < 1/10.$$

Ejemplo 4. Encontrar todas las $x \in \mathbb{R}$ tales que x^2 este en el intervalo abierto con centro en 9 y radio $1/20$.

Por la observación anterior esto es equivalente a encontrar $S \subset \mathbb{R}$ tal que

$$\text{si } x \in S \text{ entonces } |x^2 - 9| < 1/20.$$

Aplicando el Teorema anterior tenemos que

$|x^2 - 9| < 1/20 \iff -1/20 < x^2 - 9 < 1/20 \iff 9 - 1/20 < x^2 < 9 + 1/20 \iff 179/20 < x^2 < 181/20$.

Ahora aplicando el Corolario 2.4 y la Proposición 4.2 tenemos que

$$179/20 < x^2 < 181/20 \iff (179/20)^{1/2} < |x| < (181/20)^{1/2}.$$

Ahora si $x > 0$, $|x| = x$ tenemos que

$$(179/20)^{1/2} < x < (181/20)^{1/2}.$$

Y si $x < 0$, $|x| = -x$ tenemos que

$$-(181/20)^{1/2} < x < -(179/20)^{1/2}.$$

Por lo tanto

$$\text{si } x \in (-(181/20)^{1/2}, -(179/20)^{1/2}) \cup ((179/20)^{1/2}, (181/20)^{1/2}) \Rightarrow |x^2 - 9| < 1/20.$$

Ejemplo 5. Encontrar todas las $x \in \mathbb{R}$ tales que la distancia de $(x^2-16)/(x-4)$ a 8 sea menor que $1/100$.

Por la observación anterior esto es equivalente a encontrar $S \subset \mathbb{R}$ tal que

$$\text{si } x \in S \text{ entonces } |(x^2 - 16)/(x - 4) - 8| < 1/100.$$

Observemos que

$$x \neq 4 \text{ y que } (x^2 - 16)/(x - 4) = x + 4$$

por lo tanto $|(x^2 - 16)/(x - 4) - 8| = |x - 4|$ de donde la desigualdad que tenemos que resolver es

$$|x - 4| < 1/100 \text{ con } x \neq 4.$$

Aplicando el Teorema anterior tenemos que

$|x - 4| < 1/100 \iff -1/100 < (x - 4) < 1/100 \iff 4 - 1/100 < x < 4 + 1/100 \iff 399/100 < x < 401/100$, y $x \neq 4$.

Por lo tanto

$$\text{si } x \in (399/400, 4) \cup (4, 401/100) \Rightarrow |(x^2 - 16)/(x - 4) - 8| < 1/100.$$

Ejemplo 6. Encontrar todas las $x \in \mathbb{R}$ tales que $1/x$ este en el intervalo abierto con centro en 2 y radio $1/1000$.

Por la observación anteriro esto es aquivalente a encontrar $s \subset \mathbb{R}$ tal que

$$\text{si } x \in S \text{ entonces } |1/x - 2| < 1/1000.$$

Observemos que $x \neq 0$.

Ahora $|1/x - 2| < 1/1000 \iff -1/1000 < 1/x - 2 < 1/1000 \iff 2 - 1/1000 < 1/x < 2 + 1/1000 \iff 1999/1000 < 1/x < 2001/1000$.

Como $1999/1000 < 1/x \Rightarrow x > 0$ de donde se sigue que $1000/2001 < x < 1000/1999$.

Por lo tanto

$$\text{si } x \in (1000/2001, 1000/1999) \Rightarrow |1/x - 2| < 1/1000.$$

Observemos que en cada uno de estos ejemplos no se impone ninguna condición sobre el conjunto solución de la desigualdad a resolver.

Un problema un poco distinto a este es resolver una desigualdad imponiendo una cierta condición sobre el conjunto solución como nos lo muestran los siguientes ejemplos.

Recordemos que dado $x_0 \in \mathbb{R}$ para dar un intervalo con centro en x_0 lo unico que tenemos que dar es su radio el cual es un número positivo r.

Ejemplo 7. Encontrar el radio r del intervalo abierto con centro en $1/2$ tal que: si x esta en este intervalo entonces la ditancia de $4x$ a 2 sea menor que $1/10$.

Problema que queda escrito de la siguiente forma:

$$\text{si } |x - 1/2| < r \Rightarrow |4x - 2| < 1/10$$

o equivalentemente a

$$|4x - 2| < 1/10 \text{ si } |x - 1/2| < r.$$

Ahora del Ejemplo (3) tenemos que si $x \in (19/40, 21/40) \Rightarrow |4x - 2| < 1/10$, en este caso $1/2$ es el centro del intevalo abierto $(19/40, 21/40)$ por lo tanto tomando $r = 1/40$ podemos afirmar que

$$\text{si } |x - 1/2| < 1/40 \Rightarrow |4x - 2| < 1/10.$$

Ejemplo 8. Encontrar el radio r del intervalo abierto con centro en -3 tal que: x^2 este en el intervalo abierto con centro en 9 y radio $1/20$ si x esta en el intervalo abierto abierto con centro en -3 y radio r.

Una vez mas el problema que tenemos que resolver es:

$$\text{si } |x - (-3)| < r \Rightarrow |x^2 - 9| < 1/20$$

o equivalentemente a

$$|x^2 - 9| < 1/20 \text{ si } |x - (-3)| < r.$$

Ahora del Ejenplo (4) tenemos que

si $x \in (-(181/20)^{1/2}, -(179/20)^{1/2}) \cup ((179/20)^{1/2}, (181/20)^{1/2}) \Rightarrow |x^2 - 9| < 1/20$.

Es claro que

$-3 \in (-(181/20)^{1/2}, -(179/20)^{1/2})$ por lo tanto tomando $r = min\{r_1, r_2\}$

con

$r_1 = -(179/20)^{1/2}) - (-3)$ y $r_2 = -3 - (-(181/20)^{1/2})$, podemos afirmar que:

$$\text{si } |x - (-3)| < r \Rightarrow |x^2 - 9| < 1/20.$$

Ejemplo 9. Encontrar el radio r del intervalo abierto con centro en 4 tal que si x esta en este intervalo entonces la distancia de $(x^2 - 16)/(x - 4)$ a 8 es menor que $1/100$.

Una vez mas el problema a resolver es:

$$|(x^2 - 16)/(x - 4) - 8| < 1/100 \text{ si } |x - 4| < r \text{ con } x \neq 0.$$

Del Ejemplo (5) tenemos que si

$x \in (399/400, 4) \cup (4, 401/100) \Rightarrow |(x^2 - 16)/(x - 4) - 8| < 1/100$, lo cual lo podemos dar como:

$$\text{si } 0 < |x - 4| < 1/100 \Rightarrow |(x^2 - 16)/(x - 4) - 8| < 1/100.$$

Ejemplo 10. Encontrar el radio r del intervalo abierto con centro en $1/2$ tal que $1/x$ este en el intervalo abierto con centro en 2 y radio $1/1000$ si x esta en el intervalo abierto con centro en $1/2$ y radio r.

Una vez mas el problema a resolver es:

$$|1/x - 2| < 1/1000 \text{ si } |x - 1/2| < r.$$

Del Ejemplo (6) tenemos que si

$x \in (1000/2001, 1000/1999) \Rightarrow |1/x - 2| < 1/1000$, como $1/2$ es el centro del intervalo abierto $(1000/2001, 1000/1999)$ por lo tanto tomando $r = min\{r_1, r_2\}$ donde $r_1 = 1000/1999 - 1/2$ y $r_2 = 1/2 - 1000/2001$ podemos afirmar que:

$$\text{si } |x - 1/2| < r \Rightarrow |1/x - 2| < 1/1000.$$

Es claro que en cada uno de los ejemplos dados lo que estamos encontrando es el limite de una función en punto.

maga@fcfm.buap.mx

TOPOLOGÍA DE $\mathbb{R}$

ARMANDO MARTÍNEZ GARCÍA
BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

Resumen. El objetivo es ver algunos resultados topológicos que se estudian desde los primeros semestres en una facultad de ciencias.

1. Introducción

El objetivo de esta nota es hacerle ver al estudiante que desde el inicio de la carrera de Matemáticas comienza a utilizar conceptos topológicos que son basicos en el estudio de la topología.

Algunos de estos conceptos los estudia desde los primeros semestres de su carrera como son; intervalo abierto, intervalo cerrado, valor absoluto, función continua, densidad de un conjunto en otro y algunos resultados que se desprenden de estos conceptos.

Que posteriormente se generalizan primero en los cursos de Analisis y posteriormente en los cursos de Topología.

2. Intervalos

2.1. Definición. Dados $a, b \in \mathbb{R}$ con $a < b$ el intervalo abierto a, b el cual denotaremos como (a, b) es el conjunto:

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}.$$

2.2. Definición. Dados $a, b \in \mathbb{R}$ con $a < b$ el intervalo cerrado a, b el cual denotaremos como $[a, b]$ es el conjunto:

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}.$$

Es claro que si $a < b$ y $c = (a + b)/2$ entonces, $b - c = c - a$ es decir la distancia de b a c es igual a la distancia de c a a, por tal motivo en los intervalos (a, b) y $[a, b]$ a c se le llama el centro del intervalo y a $b - c$ se le llama el radio del intervalo.

2.3. Definición. Sean $x_0 \in \mathbb{R}$ y $r > 0$. El intervalo abierto con centro en x_0 y radio r el cual denotaremos como $(x_0 - r, x_0 + r)$ es el conjunto:

$$(x_0 - r, x_0 + r) = \{x \in \mathbb{R} : x_0 - r < x < x_0 + r\}$$

Es claro que si $x \in (x_0 - r, x_0 + r)$ entonces, la distancia de x a x_0 es menor que r.

Observación. Para conciderar un intervalo abierto con centro en un punto dado $x_0 \in \mathbb{R}$, es suficiente dar su radio el cual es un número real $r > 0$.

Que junto con la definición de Valor absoluto y algunas de sus propiedades nos permiten dar alguas generalizaciones de intervalo abierto.

3. Valor Absoluto

3.1. Definición. Sea $x \in \mathbb{R}$ el valor absoluto de x el cual denotamos como $|x|$ es

Observemos que

$$\text{para todo } x \in \mathbb{R} \text{ se tiene que } |x| \geq 0$$

y que

$$|x| = 0 \Longleftrightarrow x = 0$$

lo cual nos permite pensar a

$$|x|$$

como la distancia de x a 0 independientemente si x es positivo o negativo.

3.2. Teorema. Sea $x \in \mathbb{R}$ y $r > 0$. Entoces

$$|x| < r \text{ si y sólo si } -r < x < r.$$

Por lo tanto dado $x_0 \in \mathbb{R}$ y $r > 0$ la desigualdad

$$|x - x_0| < r$$

puede ser leida de las siguientes formas siendo todas ellas equivalentes entre si:

i) Valor absoluto de x menos x_0 menor que r.

ii) La distancia de x a x_0 menor que r.

iii) x esta en el intervalo abierto con centro en x_0 y radio r.

Esta desigualdad nos lleva directo a la definición de función continua.

Que es otro de los conceptos fundamentales del estudio de la Topología.

4. Función Continua

4.1. Definición. Sean $x_0 \in \mathbb{R}$ y $f : \mathbb{R} \to \mathbb{R}$. f es continua en x_0 si para todo $\epsilon > 0$ existe $\delta > 0$ tal que

$$|f(x) - f(x_0)| < \epsilon \text{ si } |x - x_0| < \delta.$$

Lo cual lo podemos leer de la siguiente forma: f es continua en x_0 si para todo intervalo abierto de radio ϵ y centro en $f(x_0)$ existe un intervalo abierto de radio δ y centro en x_0 tal que si x esta en este intervalo entonces $f(x)$ esta en el intervalo con centro en $f(x_0)$ y radio ϵ.

Si denotamos con $B(x_0, r) = \{x \in \mathbb{R} : |x - x_0| < r\}$ entonces tenemos que f es continua en x_0 si para todo $\epsilon > 0$ existe $\delta > 0$ tal que

$$f(B(x_0, \delta) \subset B(f(x_0, \epsilon)$$

que es precisamente la definición de función continua en x_0 que se trabaja en los cursos de Analisis.

Que es la definición de función continua que posteriormente se generaliza en espacios topologicos siendo esta uno de los pilares en el estudio de la Topología.

Tambien obtenemos la definición de conjunto abierto que se trbaja en los cursos de Analisis.

4.2. Definición. Sea $U \subset \mathbb{R}$. U es un conjunto abierto de $\mathbb{R}$ si

$$\text{para cada } x \in U \text{ existe } r > 0 \text{ tal que } B(x, r) \subset U.$$

Concepto que que tambien se generaliza y es el concepto basico de la Topología.

Asi mismo se aplican las propiedades de conexidad y compacidad que tiene el intervalo abierto (a, b) y el intervalo cerrado $[a, b]$ sin dar las definiciones correspondientes en forma explicita para dar dos de los teoremas basicos de Calculo Diferencial en una Variable que posteriormente tambien se generalizan.

4.3. Teorema. Sea $f : [a, b] \to \mathbb{R}$ una función continua en $[a, b]$ con $f(a) < f(b)$ (o $f(b) < f(a)$) y $c \in (f(a), f(b)$ entonces

$$\text{existe } x_0 \in (a, b) \text{ tal que } f(x_0) = c.$$

4.4. Teorema. Sea $f : [a, b] \to \mathbb{R}$ una función continua en (a, b) entonces

$$\text{existen } x_0, x_1 \in (a, b) \text{ tal que } f(x_0) \leq f(x) \leq f(x_1) \text{ para todo } x \in [a, b].$$

Es importante observar que la propiedad compacidad y conexidad que tienen los intervalos $[a, b]$ y (a, b) son dos de los conceptos principales que se generalizan y de los cuales se encarga el estudio de la Topología.

maga@fcfm.buap.mx

SOBRE EL CONCEPTO DE FUNCIÓN MEDIBLE

FRANCISCO JAVIER MENDOZA TORRES
VÍCTOR FEDERICO XOCHICALE VÁZQUEZ
FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS
BENÉMERITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

Abstract. Consideremos las siguientes definiciones de una función medible: Definición 1. Sean $(X, \mathcal{M})$ un espacio medible y (Y, τ) un espacio topológico. La función $f : (X, \mathcal{M}) \longmapsto (Y, \tau)$ es medible si $f^{-1}(G) \in \mathcal{M}$ para cualquier $G \in \tau$.

Definición 2. Sean $(X_1, \mathcal{M}_1)$ y $(X_2, \mathcal{M}_2)$ espacios medibles. La función $f : (X_1, \mathcal{M}_1) \longmapsto (X_2, \mathcal{M}_2)$ será medible si $f^{-1}(M) \in \mathcal{M}_1$ para cualquier $M \in \mathcal{M}_2$. Al enfrentar estas dos definiciones es común que se llegue a ciertas confusiones. Por ejemplo, existen funciones que cumplen la definición 1 y no la definición 2, pensando que hay una contradiccion, El trabajo presenta una función con las características mencionadas, y aclara que no existe tal contradicción.

Palabras clave: Espacio medible, $\sigma-$álgebra, función medible.

1. Introducción

En general, en los libros de texto de teoría de la integral que se utilizan en la licenciatura, la definición usual de función medible o una de sus equivalencias es la siguiente: *sean* $(X, \mathcal{M})$ *un espacio medible, y* (Y, τ) *un espacio topológico,.* $f : (X, \mathcal{M}) \longmapsto (Y, \tau)$ *es medible si* $f^{-1}(G) \in \mathcal{M}$ *para cualquier* $G \in \tau$.

Se avanza en su generalización al decir que f será medible si al considerar en Y la $\sigma-$álgebra de los borelianos, se tiene que la preimagen de cualquier boreliano es un conjunto medible en X. Debido a que casi siempre se trabaja con funciones de valor real, no es necesario avanzar más en la comprensión del concepto de función medible, quedandonos con los dos conceptos anteriores. Sin embargo, al considerar el caso en que Y es un espacio medible con una $\sigma-$álgebra distinta a la boreliana podemos pensar que la función será medible si la preimagen de cualquier conjunto medible en Y es medible en X.

Esta generalización no es tan facil de asimilar debido a que, como lo expondremos, existen fuciones que satisfacen las primeras dos definiciones y no la última. En este trabajo nos proponemos aclarar esta situación.

2. Construcción de un conjunto no medible

Repasemos la construcción del conjunto de Cantor. Sea $E = [0, 1]$ y consideremos los siguientes conjuntos:

Body Math

$$\begin{aligned} C_1 &= E - \left(\tfrac{1}{3}, \tfrac{2}{3}\right) \\ C_2 &= C_1 - \left\{\left(\tfrac{1}{9}, \tfrac{2}{9}\right) \cup \left(\tfrac{7}{9}, \tfrac{8}{9}\right)\right\} \\ &\bullet \\ &\bullet \\ &\bullet \\ C_n &= C_{n-1} - \left\{\left(\tfrac{1}{3^n}, \tfrac{2}{3^n}\right) \cup \cdots \cup \left(\tfrac{3^n-2}{3^n}, \tfrac{3^n-1}{3^n}\right)\right\}. \end{aligned}$$

El conjunto de Cantor se define como $C = \overset{\infty}{_{n=1}} C_n$. Debido a que en cada paso, para pasar de C_{n-1} a C_n, se retiran 2^n intervalos abiertos disjuntos de longitud $1/3^{n+1}$, entonces la medida de Lebesgue de C es:

$$m(C) = 1 - \overset{\infty}{_{n=0}} \frac{2^n}{3^{n+1}} = 1 - \frac{1}{3}\left(\frac{1}{1-\frac{2}{3}}\right) = 0.$$

Además C es perfecto y nada denso en E

Basandonos en la construcción del conjunto de Cantor, construyamos un conjunto $D \subset [0,1]$ medible, con medida $m(D) > 0$ y nada denso en E. Para ésto sea $\xi \in (0,1)$ y construimos conjuntos D_n de tal forma que para pasar de D_n a D_{n+1} quitamos 2^n intervalos de longitud $(1/\xi + 2)^{-n-1}$. Con esto tenemos que la suma de las longitudes quitadas a E es $\sum_{n=0}^{\infty} \frac{2^n}{(1/\xi+2)^{n+1}}$. Además, considerando que $D = \overset{\infty}{_{n=0}} D_n$, tendriamos que su medida es

$$m(D) = 1 - \overset{\infty}{_{n=0}} \frac{2^n}{(1/\xi+2)^{n+1}} = 1 - \xi.$$

Como ξ esta en $(0,1)$, entonces la medida de D será positiva. De forma semejante como se demuestra que el conjunto de Cantor es nada denso, se demuestra que D es nada denso en E.

En lo que sigue veremos que existe un conjunto $B \subset D$ no medible.

Definición. Definamos la *suma modulo 1* en E como sigue:

$$x \dot{+} y = \begin{cases} x+y & \text{si } x+y \leq 1 \\ x+y-1 & \text{si } x+y > 1 \end{cases}.$$

Por ejemplo; $1/2 \dot{+} 3/5 = 11/10 - 1 = 1/10$.

Definición. Sean $x, y \in E$, diremos que $x \sim y$ si $x - y \in \mathbb{Q}$.

Esta es una relacion de equivalencia, y la coleccion de las clases de equivalencia definen una particion de E. Denotemos a $\widehat{x}$ como el representante de cada una de ellas y sea $F = \{\widehat{x} : \ x \in E\}$.

Sea $\{q_i\}_0^{\infty}$ una enumeracion de $\mathbb{Q} \cap E$ y sea $F_i = F \dot{+} q_i$, $i \in \mathbb{N} \cup \{0\}$, donde $F_0 = F$. Entonces se tiene que:

i) $F_i \cap F_j = \emptyset$, si $i \neq j$

ii) $\overset{\infty}{_{i=1}} F_i = [0,1] = E$.

Lema 1. *F no es medible*

Demostración. Si suponemos que es medible, como $F_i = F \dot{+} q_i$, entonces cada F_i sería medible y tendríamos que

$$1 =_{i=0}^{\infty} m(F_i) =_{i=0}^{\infty} m(F \dot{+} q_i) =_0^{\infty} m(F),$$

lo cual no puede ser, ya que si $m(F) = 0$ entonces se tendría que $1 = 0$, ó si $m(F) > 0$ tendríamos una suma infinita de términos constantes positivos que es igual a uno, por lo tanto F no puede ser medible. ■

Lema 2. *Si A es medible y $A \subset F$, entonces $m(A) = 0$*

Demostracion. Sean $A_i = A \dot{+} q_i = \{a \dot{+} q_i : a \in A\}$. $\{A_i\}$ es una familia de conjuntos disjuntos, medibles y $m(A_i) = m(A)$. Entonces

$$1 = m(E) \geq m(\,_0^{\infty} A_i) =_0^{\infty} m(A_i) =_0^{\infty} m(A),$$

de donde no puede ser que $m(A) > 0$, por lo tanto $m(A) = 0$.■

Lema 3. *Existe un conjunto $B \subset D$ que no es medible*

Demostración. Para cada $i \in \mathbb{N}$, sea $B_i = B \cap F_i$. Si cada uno de ellos es medible, entonces, por el lema 2, $m(B_i) = 0$ y tendríamos que

$$0 =_0^{\infty} m(B_i) = m(\,_0^{\infty} B_i) = m(D) = 1 - \xi > 0,$$

por lo tanto existe un $i \in \mathbb{N}$ tal que B_i no es medible. Denotemos este conjunto por B.■

3. Una función medible con preimagen no medible

Ahora construyamos una funcion $f : [0,1] \longrightarrow [0,1]$ medible (continua y monotona) tal que $H = f(B)$ sea medible, de tal forma que tengamos que $f^{-1}(H) = B$ no es medible. Para ésto sigamos los pasos siguientes:

a) Sea $g : [0,1] \setminus D \longrightarrow [0,1] \setminus C$ definida de tal forma que hagamos corresponder linealmente los intervalos que se quitan en cada paso para obtener C_n y D_n. Por esta construccion se tiene que g es creciente y biyectiva.

b) Para cada $x_0 \in D$ se tiene que:

i) $\lim_{x \longrightarrow x_0^-} g(x) = \sup \{g(x) \ / \ x \in [0,1] - D, \ x < x_0\}$

ii) $\lim_{x \longrightarrow x_0^+} g(x) = \inf \{g(x) \ / \ x \in D, \ x_0 < x\}$.

c) Sea $x_0 \in D$, como $g(x) \leq g(x')$ para todo $x \in [0,1] - D$, $x < x_0$, y para todo $x' \in [0,1] - D : x' < x_0$. Haciendo $a = \lim_{x \longrightarrow x_0^-} g(x)$ y $b = \lim_{x \longrightarrow x_0^+} g(x)$ y por las igualdades anteriores de **b)**, se tiene que $a \leq b$.

d) Lema 4. $a = b$

Demostración. Supongamos que $a < b$. Por la construcción de g sabemos que $a, b \in C$. Como C es nada denso en E entonces existe $g(x^*) \notin C$ con $x^* \in [0,1] - D$ tal que $a < g(x^*) < b$. Pero por la monotonia de g, x^* deberá estar a la izquierda y a la vez a la derecha de x_0, por lo tanto $a = b$. ■

e) Sea $f : [0,1] \to [0,1]$ definida como

$$f(x) = \begin{cases} g(x) & \text{si } x \in [0,1] \setminus D \\ \lim_{x \longrightarrow x_0} g(x) & \text{si } x \in D \end{cases}.$$

Esta función es continua, creciente y sobreyectiva, por lo tanto es medible..

f) Sea $H = f(B)$. Por la definición de f, y como $B \subset D$, entonces $f(B) \subset C$. Como la medida de Lebesgue es completa, entonces: $H = f(B)$ es medible y $m(H) = 0$.

Hemos construido una función medible $f : [0,1] \longrightarrow [0,1]$, *para la cual existente un conjunto* $H \subset [0,1]$ *medible Lebesgue tal que* $f^{-1}(H) = B$ *no es medible.*

4. Conclusión

La definición general de función medible es la siguiente: *sean* $(X, \mathcal{M}_1)$ *y* $(Y, \mathcal{M}_2)$ *dos espacios medibles, la función* $h : (X, \mathcal{M}_1) \longmapsto (Y, \mathcal{M}_2)$ *será medible si* $h^{-1}(M) \in M_1$ *para cualquier* $M \in M_2$.

Al considerar cualquier espacio medible $(X, \mathcal{M})$ y el espacio medible $(Y, \mathcal{B})$, donde Y es de origen un espacio topológico y $\mathcal{B}$ es la $\sigma-$ álgebra de los borelianos sobre Y, tendremos, según la definición anterior, que la función $g : (X, \mathcal{M}) \to (Y, \mathcal{B})$ es medible si $g^{-1}(B) \in \mathcal{M}$ para cualquier $B \in \mathcal{B}$. Debido a que usualmente se trabaja con funciones que van de un espacio medible a un topológico, como es el caso de las funciones de valor real, las definiciones de medibilidad que se dan para ese tipo de funciones son las equivalentes a la anterior Por ejemplo, $g : (X, \mathcal{M}) \to \mathbb{R}$ será medible si $g^{-1}((a,b)) \in \mathcal{M}$ para cualquier intervalo abierto $(a,b) \subset \mathbb{R}$.

Ahora bien, al considerar una función medible los espacios sobre la que esta definida son básicos. Si los cambiamos, la función puede dejar de ser medible, como es el caso de la función f que construimos en los apartados anteriores. Nuestra función f es medible cuando va de $(E, \mathcal{L})$ a $(E, \mathcal{B})$, siendo $\mathcal{L}$ y $\mathcal{B}$ las σ-álgebras de Lebesgue y Borel, respectivamente. Pero esa función deja de serlo si en lugar de $(E, \mathcal{B})$ tenemos $(E, \mathcal{L})$.

5. Bibliografía

[1] Bartle, R. G., *The Elements of Integration.* John Wiley and Sons, Inc., Nueva York, 1965.

[2] Wilcok H.J. y Myers D.L., *An Introduction to Lebesgue Integration and Fourier Series.* Dover Publications, Inc., 1978.