

1001000101001010010101011000111010 **ISSN Nº 1390 - 3802**

1010011001000101010100010100101010100101

0010010101000101010100010100101010100101

1010010101000101010100010100101010100101

0100101000101010100010100101010100101001

0100101000101010100010100101010100101001

0100101000101010100010100101010100101001

0010010101000101010100010100101010100101

1010010101000101010100010100101010100101

0100101000101010100010100101010100101001

1000101000101010100010100101010100101001

1100101000101010100010100101010100101001

01100101000101010100010100101010100101001

1001001000101010100010100101010100101001

0101001000101010100010100101010100101001

01001001000101010100010100101010100101001

10010001000101010100010100101010100101001

10110001000101010100010100101010100101001

010010001000101010100010100101010100101001

100100001000101010100010100101010100101001

101100001000101010100010100101010100101001

0100100001000101010100010100101010100101001

1001000001000101010100010100101010100101001

1011000001000101010100010100101010100101001

01001000001000101010100010100101010100101001

10010000001000101010100010100101010100101001

10110000001000101010100010100101010100101001

010010000001000101010100010100101010100101001

100100000001000101010100010100101010100101001

101100000001000101010100010100101010100101001

0100100000001000101010100010100101010100101001

1001000000001000101010100010100101010100101001

1011000000001000101010100010100101010100101001

01001000000001000101010100010100101010100101001

10010000000001000101010100010100101010100101001

10110000000001000101010100010100101010100101001

010010000000001000101010100010100101010100101001

100100000000001000101010100010100101010100101001

101100000000001000101010100010100101010100101001

0100100000000001000101010100010100101010100101001

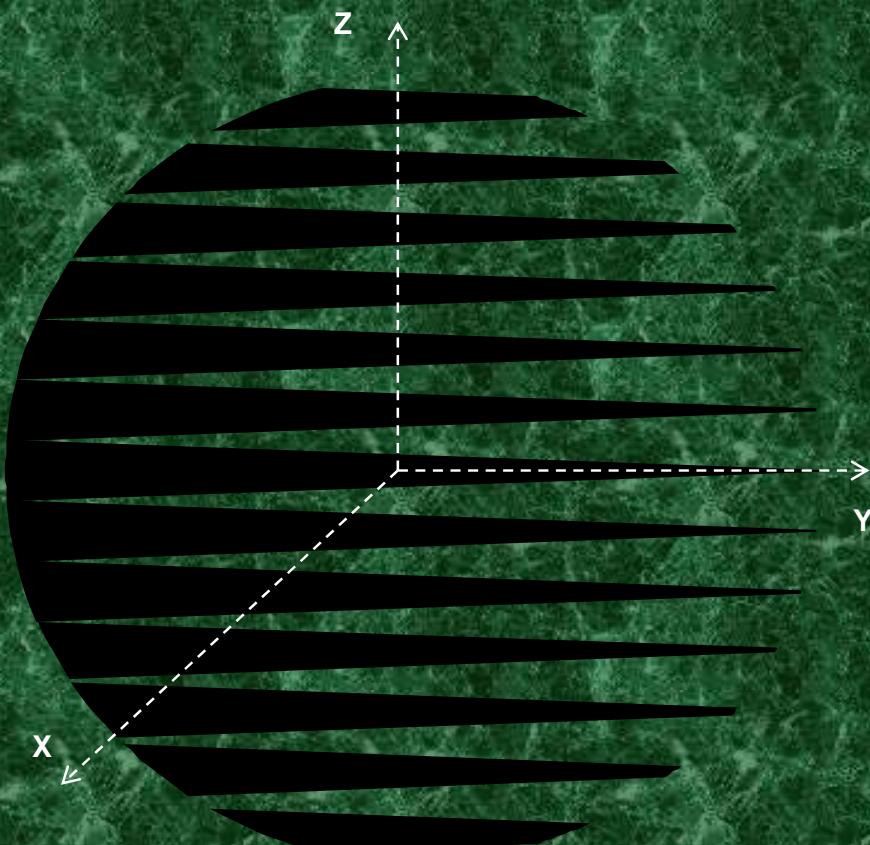
1001000000000001000101010100010100101010100101001

1011000000000001000101010100010100101010100101001

# matemática

UNA PUBLICACIÓN DE FCNM - ESPOL

Volumen 11 Número 1 Abril 2013



Escuela Superior Politécnica del Litoral - ESPOL  
Facultad de Ciencias Naturales y Matemáticas - FCNM

00101010110010001010010100101011000111010100010100

10100110010001010101000101001000101001000101010110100101

## **FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS**

### **DEPARTAMENTO DE MATEMÁTICAS**

El Departamento de Matemáticas (DM) es una unidad académica de la ESPOL. Desde el inicio la función del DM ha sido la docencia en Matemáticas, Ciencias Gráficas e Informática, para la formación de profesionales en ingeniería, tecnología y ciencias, habiendo tenido a su cargo en los albores de la ESPOL, el dictado de 10 materias. Con el transcurso del tiempo y acorde con la era de la información, el Departamento de Matemáticas creó en mayo de 1995 la carrera de “Ingeniería en Estadística Informática”, como alternativa en ingeniería en información y servicios. Posteriormente, con el fin de garantizar la eficiencia en el control y gestión empresarial con profesionales capacitados y de excelencia se creó la carrera de “Auditoría y Control de Gestión” en mayo de 2000. También el Instituto ha incursionado en una de las más importantes ramas de la matemática aplicada que tiene grandes aplicaciones en el mundo moderno, esto es la Investigación de Operaciones, la Teoría de Optimización, y particularmente las aplicaciones logísticas, a través del ofrecimiento de programas de pre-grado y post-grado en estas áreas. Así es como desde el año 2005 se viene ofreciendo la maestría en Control de Operaciones y Gestión Logística y desde el año 2006 la carrera de Ingeniería en Logística y Transporte.

El DM también cuenta con el CENTRO DE INVESTIGACIONES ESTADÍSTICAS, a través del cual, se realizan: estudios de predicción, estudios actuariales, estudios de mercado, diseños de experimentos, planificación y dirección de censos, análisis financieros, bases de datos estadísticos, formulación de proyectos, ingeniería de la calidad, etc.

Entre otras actividades que desarrolla el DM anualmente están: las JORNADAS EN ESTADÍSTICA E INFORMÁTICA que actualmente está en su decimoséptima versión, el CONCURSO INTERCOLEGIAL DE MATEMÁTICAS que se viene realizando en forma continua desde 1988.

.



Más información: [www.icm.espol.edu.ec](http://www.icm.espol.edu.ec) o escribirnos al e-mail: [icm@espol.edu.ec](mailto:icm@espol.edu.ec), [jvaldi@espol.edu.ec](mailto:jvaldi@espol.edu.ec), [erivaden@espol.edu.ec](mailto:erivaden@espol.edu.ec), 30 ½ vía Perimetral: Edificios 25 – B Planta alta (Área de Institutos) Telfs.: (593-4) 2269525 – 2269526, fax: (593-4) 853138.  
Guayaquil – Ecuador



**Volumen 11**

**Número 1**

**Abril 2013**

**Rector ESPOL:**

**M.Sc. Sergio Flores**

**Vicerrectora General ESPOL:**

**Ph.D. Cecilia Paredes**

**Decano FCNM:**

**M.Sc. Gaudencio Zurita Herrera**

**Subdecano FCNM:**

**M.Sc. Oswaldo Valle Sánchez**

**Directora Departamento de**

**M.Sc. Janet Valdiviezo**

**Matemáticas:**

**M.Sc. Hernando Sánchez**

**Director Departamento de Física:**

**Ph.D. David Matamoros**

**Director Departamento de**

**M.Sc. Eduardo Rivadeneira Molina**

**Química:**

**Editor de publicaciones:**

**Comité Editorial:**

**M.Sc. Efrén Jaramillo Carrión**

**Ph.D. David Matamoros**

**M.Sc. Luis Rodríguez Ojeda**

**Ph.D. Francisco Vera**

**Asesores Editoriales:**

**Ph.D. Joseph Páez Chávez**

**Mae. Sandra García Bustos**

**M.Sc. Olga González Sánchez**

**M.Sc. Justo Huayamave Navarrete**

**Mg. Eva María Mera Intriago**

**Mg. María Nela Pastuizaca**

**M.Sc. Fernando Sandoya Sánchez**

**M.Sc. Francisco Torres Andrade**

**Ing. Antonio Chong Escobar**

**Edición:**

**Srta. Carolina Carrasco Salas**



**matemática** es una publicación del Departamento de Matemáticas de la Facultad de Ciencias Naturales y Matemáticas de la Escuela Superior Politécnica del Litoral, y pretende constituirse en un órgano de difusión científico – tecnológico, con el fin de incentivar y motivar el desarrollo y avance de la matemática y sus aplicaciones.

**matemática** publica artículos teóricos y de tipo experimental tales como ensayos, resúmenes de tesis de grado y trabajos de investigación relacionados con la aplicación de la matemática en los diferentes ámbitos de la realidad.

# CONTENIDO

<b>EDITORIAL.....</b>	<b>5</b>
<b>MODELANDO EL TRANSPORTE DE IONES DE POTASIO A TRAVÉS DE UN CANAL IONICO DE UNA MEMBRANA CELULAR</b>	
Iza Peter.....	7
<b>HOMOTOPÍA Y CONTINUACIÓN NUMÉRICA EN SISTEMAS NO LINEALES</b>	
Martín Carlos.....	13
<b>SISTEMAS NO LINEALES Y LA METAHEURÍSTICA SCATTER SEARCH</b>	
Martín Carlos.....	24
<b>INCORPORANDO BROSSEAU EN DOS TEOREMAS FUNDAMENTALES: PROBABILIDAD TOTAL Y BAYES</b>	
Petakos Kyriakos.....	34
<b>DISEÑO DE MUESTRA PARA CONTEO RÁPIDO</b>	
Vera Francisco.....	39

## EDITORIAL

Los profesores de las llamadas Ciencias Básicas tienen un nuevo desafío. La creación de la Facultad de Ciencias Naturales y Matemáticas (FCNM) es un paso adelante en la consolidación de la Investigación Sistemática en la ESPOL. Ahora los profesores de Química, Física y Matemáticas deben afrontar el reto de la investigación tanto de manera individual como colectiva.

Hemos visto con optimismo como se dan los primeros pasos para conseguir este objetivo. La meta de la FCNM es la de liderar la Investigación Científica en la ESPOL. Para ello, deseamos todo éxito a los Colegas que han recibido la responsabilidad inmensa de crear las bases materiales y espirituales de esta nueva Facultad en el seno de la ESPOL.

Una vez más, la revista **Matemática** es el recipiente de los esfuerzos de los Investigadores de toda la comunidad científica de la ESPOL.

Invitamos pues, a todos los Investigadores y Científicos de la ESPOL se sumen a este esfuerzo para fortalecer las Ciencias Básicas, que representan el fundamento de toda actividad académica en la ESPOL.

## MODELANDO EL TRANSPORTE DE IONES DE POTASIO A TRAVÉS DE UN CANAL IÓNICO DE UNA MEMBRANA CELULAR

Iza Peter<sup>1</sup>

**Resumen.** Se desarrolla dos modelos de conducción de iones a través de una membrana celular en condiciones simétricas de concentración considerando canales iónicos activados por voltaje. Estos canales son altamente selectivos a iones K o Na en presencia de una diferencia de potencial. Primero se plantea un modelo de conducción similar al de la conducción de electrones libres en metales donde se modela una distribución aleatoria de obstáculos que impiden el movimiento libre del ion dentro del canal. Esta distribución de espaciamiento entre obstáculos resultó ser muy angosta y su valor medio muy pequeño, implicando un casi continuo choque del ion con los obstáculos presentes en el canal. Adicionalmente, se propone un modelo considerando al canal como una estructura periódica. Para que estos modelos describan los valores experimentales de conducción, los parámetros asociados a cada modelo tienen valores que no son fáciles de comprobar experimentalmente los cuales tienen relación con la estructura del canal.

**Palabras clave:** membrana celular, canal iónico, modelo de conducción.

**Abstract.** Conduction models of ion through voltage-gated ion channels, at symmetrical concentration in cellular membrane, have been developed. These channels are highly selective to K or Na ions due to a potential difference. First, it is considered the theory of metallic conduction, where, it is assumed a random distribution of obstacles, which oppose to free motion of ions inside the channel. This spacing distribution among obstacles turned out to be very narrow and its mean value is very small, that implies an almost continuous collision of ion with the channel obstacles. In addition, a model based on the crystal periodic structure is proposed. These models would describe the experimental values of conduction, but the parameters associated to each model would take values that are not easy to verify experimentally, and they are related to the channel structure.

**Key words:** cellular membrane, ion channel, conduction model.

Recibido: Febrero 2013

Aceptado: Marzo 2013

### 1. INTRODUCCIÓN

La membrana celular tiene una estructura laminar formada por una doble capa de lípidos, que representa un 35%, proteínas con un 62% y un 3% de polisacáridos; el espesor de la membrana tiene un valor no mayor a 100 Å. La membrana es una barrera permeable, la cual contiene sistemas de transporte como: bombas iónicas, compuertas moleculares específicas; que permiten el paso de moléculas o iones de cierta clase. Los procesos de transporte en la membrana tienen papeles importantes: regulan el volumen celular y mantienen el pH y la composición iónica intracelular, con el fin, de suministrar un ambiente favorable a la actividad enzimática. Extraen y concentran combustibles metabólicos, precursores del entorno y eliminan sustancias tóxicas [13, 14, 15, 16, 17]. Generan gradientes iónicos esenciales para la excitabilidad del nervio y del músculo. Existen dos procesos de transporte a través de la membrana, el *Transporte Activo*: consiste en el transporte de sustancias en contra de un gradiente de concentración, lo que requiere un gasto energético; se realiza a expensas de un gradiente de H<sup>+</sup> (potencial electroquímico de protones) y por hidrólisis de adenosin trifosfato (ATP del inglés Adenosine TriPhosphate); y el *Transporte Pasivo*: un ion que atraviesa la membrana celular

debido a una diferencia de concentración, o por un gradiente electroquímico y sin el desgaste de energía metabólica (ATP), se dice que es transportado pasivamente [7, 8, 9].

La descripción del movimiento de iones a través de una membrana biológica es un tema de mucho interés, por su papel en el metabolismo de la célula. Este movimiento de iones a través de membranas ha dependido principalmente de la estructura de la membrana y de la concentración de los iones. Los biólogos han estado siempre interesados en el movimiento de los iones a través de la membrana celular; el intercambio de sustancias en la membrana puede ser modificado por la variación de concentración, por el potencial eléctrico o inclusive por sistemas biológicos como: canales iónicos, bombas Na<sup>+</sup>/K<sup>+</sup>, etc. Este interés ha producido que se desarrollen aparatos electrónicos para realizar mediciones de flujo de iones en la membrana, bajo ciertas condiciones, inclusive se han mejorado los aparatos de rayos X para observar de mejor manera la estructura de la membrana; y además se ha desarrollado modelos teóricos, con diferentes propósitos.

Algunos investigadores por ejemplo han utilizado la técnica de Patch-Clamp para estudiar los tres tipos principales de canales iónicos: los activados por voltaje, los activados por neurotransmisores, y los activados mecánicamente, lo que ha permitido comprender la importancia de la composición de los lípidos en la membrana y sobre todo el funcionamiento de los canales. A partir de los registros que se

<sup>1</sup>Iza Peter, Ph.D., Profesor de la Escuela Superior Politécnica del Litoral (ESPOL).  
(e\_mail: piza@espol.edu.ec).

obtienen con esta técnica se ha podido determinar la probabilidad de apertura de un canal a ciertos valores de voltaje.

Los modelos teóricos están basados en simulaciones, las cuales por ejemplo analizan la energía que necesitan los iones para atravesar el canal, las fuerzas de repulsión que sufre el ion debido a la presencia de dipolos en la pared del canal. En este tipo de simulaciones surgen dificultades principalmente por el vacío entre las escalas de tiempo a nivel atómico y las escalas de tiempo macroscópico [2, 3, 4, 10, 11, 12, 18].

## 2. MODELOS

Existen diferentes tipos de canales iónicos en la membrana celular, en este trabajo se considerará a los canales iónicos activados por voltaje, los cuales son selectivos a cierto tipo de iones [20]. Se conoce que los canales iónicos tienen un medio acuoso y ciertos grupos carboxilos ubicados en lugares específicos de la pared del canal, esto para cada tipo de canal.

**Modelo 1:** Debido a la falta de conocimiento estructural de un canal iónico se procede a idealizar un canal iónico, como si éste tuviera obstáculos distribuidos aleatoriamente a lo largo del canal, estos obstáculos representarán la estructura y los grupos carboxilos, ver figura 1 [20]. Dos reservorios de forma cilíndrica, de área y longitud fija, están en contacto con el canal iónico, con igual concentración de iones potasio, denominando a esto: condiciones simétricas de concentración. La sección transversal del canal tiene un radio promedio de 3 Å; el radio cristalino del ion potasio es de 1.3 Å, el radio hidratado del ion es de 1.83 Å y la longitud del canal es 12 Å. Debido a una diferencia de potencial (V) entre los dos reservorios, los iones estarán pasando del reservorio 1 al 2 y por tanto habrá un flujo de iones a través del canal.

Se puede considerar que el paso de iones por un canal iónico es similar al flujo de electrones por un metal. Por tanto, la densidad de corriente ( $j$ ) de un canal iónico se puede determinar utilizando la ecuación:

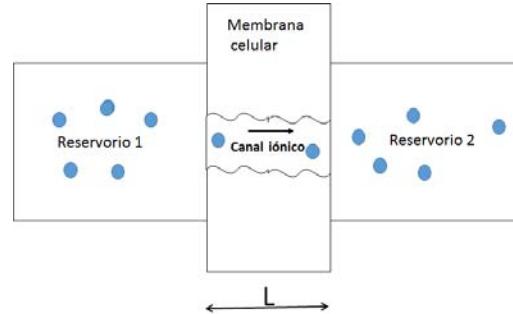
$$j = \sigma E \quad (1)$$

donde  $\sigma$  representa la conductividad, que caracteriza la facilidad con que un campo eléctrico ( $E$ ) induce una corriente en un medio, y es igual a:

$$\sigma = \frac{n e^2 \tau}{m} \quad (2)$$

$n$  representa la densidad de los iones,  $e$  la carga eléctrica del ion,  $\tau$  tiempo de relajación y  $m$  la masa del ion.

**FIGURA 1**  
*Paso de iones en presencia de un campo eléctrico*  
**Modelando el transporte de iones de potasio a través de un canal iónico de una membrana celular**



**Movimiento de iones.** Considerando que los iones tienen un movimiento rectilíneo acelerado dentro del canal iónico, se tiene que el desplazamiento de un ion es igual a:

$$d = v_0 t + \frac{1}{2} a t^2 \quad (3)$$

la aceleración que experimenta el ion, es debido a la presencia de un campo eléctrico  $E$  y por tanto, la ecuación (3), se sustituye por:

$$d = v_0 t + \frac{e E}{2 m} t^2 \quad (4)$$

El término  $v_0$  es igual a cero al tiempo  $t=0$  ya que no existe campo eléctrico, y por tanto el desplazamiento queda expresado como:

$$d = \frac{e E}{2 m} t^2 \quad (5)$$

Despejando  $t$  y promediando cada lado de la ecuación (5), se obtiene:

$$\bar{t} = \sqrt{\frac{2 m}{e E} d^{1/2}} \quad (6)$$

Denominando  $\tau$  al tiempo promedio, que representará el tiempo promedio entre colisiones.

$$\tau = \sqrt{\frac{2 m}{e E} d^{1/2}} \quad (7)$$

Utilizando las ecuaciones de conductividad (2) y densidad de corriente (1) se puede obtener una expresión para la densidad de corriente de iones dentro del canal:

$$j = n e^{3/2} \sqrt{\frac{2 E}{m}} \bar{d}^{1/2} \quad (8)$$

**Distribución de espaciamiento medio entre obstáculos dentro del canal.** Para calcular las propiedades promedio de un conjunto de partículas, por lo general se necesita conocer la forma en que dichas partículas están distribuidas. En varias situaciones físicas reales existe una gran cantidad de eventos aleatorios que están

gobernados por la distribución normal. Se dice que una variable  $x$  tiene una distribución normal con media  $u$  ( $-\infty < u < \infty$ ) y varianza  $\sigma_1^2 > 0$ , si tiene la siguiente función de densidad:

$$FD = \frac{1}{\sigma_1 \sqrt{2\pi}} \text{Exp} \left( -\left( \frac{x-u}{\sqrt{2\sigma_1^2}} \right)^2 \right) \quad (9)$$

Partiendo de la expresión de distribución normal se puede encontrar la raíz cuadrática media de  $x$ , que representará la “distribución de espaciamiento entre obstáculos dentro del canal”:

$$\overline{x^{1/2}} = \int_0^{+\infty} \frac{1}{\sigma_1 \sqrt{2\pi}} \sqrt{x} \text{Exp} \left( -\left( \frac{x-u}{\sqrt{2\sigma_1^2}} \right)^2 \right) dx \quad (10)$$

Hay que tener en cuenta que los términos:  $\sigma_1$  y  $u$  son mayores a cero y que se ha considerado  $u > \sigma_1$  para tener los límites de integración.

Realizando la integración respectiva y con el

cambio de  $x^{1/2}$  por  $d^{1/2}$  se obtiene:

$$\overline{d^{1/2}} = \frac{1}{4\sigma_1} \text{Exp} \left( -\frac{u^2}{4\sigma_1^2} \right) \sqrt{\frac{\pi}{2}} u^{3/2} \\ \left( \text{Bessel } I \left[ -\frac{3}{4}, \frac{u^2}{4\sigma_1^2} \right] + \text{Bessel } I \left[ -\frac{1}{4}, \frac{u^2}{4\sigma_1^2} \right] + \right. \\ \left. \text{Bessel } I \left[ \frac{1}{4}, \frac{u^2}{4\sigma_1^2} \right] + \text{Bessel } I \left[ \frac{3}{4}, \frac{u^2}{4\sigma_1^2} \right] \right) \quad (11)$$

**Conductancia del canal iónico.** La conductancia es una propiedad de la membrana celular y específicamente del canal iónico, y representa una medida de la facilidad con que los iones atraviesan un canal iónico de la membrana celular. Como los iones tienen carga eléctrica, la conductancia se manifiesta en forma de corrientes eléctricas que atraviesan a la membrana celular. La conductancia se mide en unidades llamadas Siemens. Utilizando la ecuación (8) y (11) se obtiene la expresión de densidad de corriente del ion:

$$j = n e^{3/2} \sqrt{\frac{2E}{m}} \frac{1}{4\sigma_1} \text{Exp} \left( -\frac{u^2}{4\sigma_1^2} \right) \sqrt{\frac{\pi}{2}} u^{3/2} \\ \left( \text{Bessel } I \left[ -\frac{3}{4}, \frac{u^2}{4\sigma_1^2} \right] + \text{Bessel } I \left[ -\frac{1}{4}, \frac{u^2}{4\sigma_1^2} \right] + \right. \\ \left. \text{Bessel } I \left[ \frac{1}{4}, \frac{u^2}{4\sigma_1^2} \right] + \text{Bessel } I \left[ \frac{3}{4}, \frac{u^2}{4\sigma_1^2} \right] \right) \quad (12)$$

Sabiendo que el campo eléctrico se puede expresar en función de la diferencia de voltaje de la siguiente manera:

$$E = \frac{V}{d} \quad (13)$$

y considerando las siguientes condiciones:

$$u = \alpha V$$

$$\sigma_1 = \frac{u}{3} = \frac{\alpha V}{3} \quad (14)$$

Se encuentra una expresión de corriente iónica para un canal iónico provocado por el paso de iones, si multiplicamos la densidad de corriente por el área del canal ( $S$ ):

$$I = \frac{3nS}{4 \text{Exp}^{9/4}} \sqrt{\frac{\pi e^3 \alpha}{md}} \\ \left( \text{Bessel } I \left[ -\frac{3}{4}, \frac{9}{4} \right] + \text{Bessel } I \left[ -\frac{1}{4}, \frac{9}{4} \right] + \right. \\ \left. \text{Bessel } I \left[ \frac{1}{4}, \frac{9}{4} \right] + \text{Bessel } I \left[ \frac{3}{4}, \frac{9}{4} \right] \right) V \quad (15)$$

Se puede observar otra variante de la Ley de Ohm:

$$I = \sigma V \quad (16)$$

con

$$\sigma = \frac{3nS}{4 \text{Exp}^{9/4}} \sqrt{\frac{\pi e^3 \alpha}{md}} \\ \left( \text{Bessel } I \left[ -\frac{3}{4}, \frac{9}{4} \right] + \text{Bessel } I \left[ -\frac{1}{4}, \frac{9}{4} \right] + \right. \\ \left. \text{Bessel } I \left[ \frac{1}{4}, \frac{9}{4} \right] + \text{Bessel } I \left[ \frac{3}{4}, \frac{9}{4} \right] \right) \quad (17)$$

que representa la conductancia del canal iónico. Note que esta expresión de conductancia está en función de magnitudes del ion y del canal, así como de la distribución de obstáculos representada por  $\alpha$ .

La concentración ( $n$ ) de iones potasio en los reservorios es 0.302 Molar (0.302 moles/litro), lo que implica la presencia de  $\theta=13$  iones potasio. La forma de los contenedores es cilíndrica y sus dimensiones son: longitud ( $l_1$ ) de  $27 \times 10^{-10}$  m, y radio ( $r_1$ ) igual a  $29 \times 10^{-10}$  m,  $S$  área del canal iónico, el radio ( $r$ ) es  $3 \times 10^{-10}$  m,  $m$  la masa del ion potasio  $6.5 \times 10^{-26}$  Kg,  $d$  la longitud del canal  $12 \times 10^{-12}$  m, e la carga del electrón  $1.602 \times 10^{-19}$  C y  $\alpha$  una constante que debería tener unidades de  $m/V$ . Además, la función BesselI [ $n,z$ ] representa las funciones de modificadas de Bessel de primera clase.

**Modelo 2:** Desde el punto de vista cuántico cuando un electrón se encuentra en un cristal, es como si el electrón se encontrara en un potencial perfectamente periódico, que tiene la periodicidad de la red cristalina [1, 19, 20]. El potencial periódico se lo considera como resultado de la distribución de carga asociada con los núcleos

iónicos situados en los puntos reticulares, más el potencial “extendido” promedio aportado por todos los demás electrones libres que pertenecen al cristal, de tal manera que se tiene en cuenta la interacción promedio de un solo electrón con todos los demás. Cuando el electrón es acelerado dentro de un cristal, la masa del electrón aparece diferente a la masa del electrón libre, y se le denomina *masa efectiva*. Este modelo considera al canal iónico como una red cristalina con cierta periodicidad; esta periodicidad está representada por la estructura molecular uniforme del canal, es decir, los dipolos que forman parte del canal que hacen la oposición al movimiento de los iones y representarían los pozos de potencial. Por tanto, la velocidad del ion en el canal es representada por su velocidad de grupo, asociada con el “paquete de ondas” del ion:

$$v_g = \frac{1}{\hbar} \frac{\partial \mathcal{E}}{\partial k} \quad (18)$$

donde  $k$  es el número de onda asociado al electrón,  $\mathcal{E}$  es la energía del ion en la red y  $\hbar$  es la constante de Planck dividido por 2. Según la física clásica “ $d\mathcal{E}$ ” representa el trabajo realizado por una partícula viajando una distancia  $v_g dt$  bajo la influencia de una fuerza  $qE$ , por tanto:

$$d\mathcal{E} = q E v_g dt = q E \frac{1}{\hbar} \frac{\partial \mathcal{E}}{\partial k} dt \quad (19)$$

Diferenciando la ecuación (18):

$$\frac{dv_g}{dt} = \frac{1}{\hbar} \frac{d}{dt} \left( \frac{\partial \mathcal{E}}{\partial k} \right) = \frac{1}{\hbar} \frac{\partial^2 \mathcal{E}}{\partial k^2} \frac{dk}{dt} \quad (20)$$

de (19) despejando  $dk/dt$  y sustituyendo en (20):

$$\frac{dv_g}{dt} = \frac{1}{\hbar^2} \frac{\partial^2 E}{\partial k^2} qE \quad (21)$$

Comparando con la segunda ley de Newton para la partícula clásica:

$$m \frac{dv}{dt} = qE \quad (22)$$

se puede definir:

$$m^* = \hbar^2 \left( \frac{\partial^2 E}{\partial k^2} \right)^{-1} \quad (23)$$

como la masa efectiva del ion. Note que la fuerza debido a la estructura del cristal está contenido en esta masa efectiva. Si el ion en reposo ( $k=0$ ) es acelerado por un campo eléctrico, el ion se moverá a valores más altos de  $k$  y el ion llegará a ser más pesado. Considerando el efecto de la red periódica en el valor de la masa efectiva del ion, y suponiendo que éste se mueve en un medio viscoso, utilizamos la ecuación de Langevin:

$$m^* v' = -m^* \gamma v + e E \quad (24)$$

resolviendo la ecuación diferencial y calculando la velocidad media, se puede hallar la expresión de corriente iónica:

$$I = \frac{n S \left( \exp(-\gamma t) - 1 + \gamma t \right)}{m^* d t \gamma^2} e^2 V \quad (25)$$

y finalmente la expresión de conductancia:

$$\sigma = \frac{n S \left( \exp(-\gamma t) - 1 + \gamma t \right)}{m^* d t \gamma^2} e^2 \quad (26)$$

### 3. DISCUSIÓN

La estructura, forma y dimensiones de un canal iónico no es conocida, por lo que, varios autores la modelan, por ejemplo: ubican anillos dipolares a la entrada y salida del canal o crean una barrera de potencial energética [2, 5, 6]. Los obstáculos propuestos en estos trabajos harían las veces de grupos carboxilos o amino ubicados en la pared celular del canal iónico, lo que provocaría los choques aleatorios del ion. La ecuación de Langevin utilizada para describir el movimiento Browniano de los iones, caracteriza a todas las interacciones de las moléculas como resultado de fuerzas conservativas y disipativas, al atravesar el canal. Además, se asume que el campo eléctrico y el coeficiente de fricción son independientes del tiempo, lo cual podría ser verdadero o falso en una situación real. Se consideró que la función de distribución de espaciamiento medio entre los obstáculos dentro del canal es una función Normal. Como los canales iónicos activados por voltaje tienen un segmento especial ( $S_4$ ) el cual tiende a cambiar la estructura conformacional del canal; los términos  $u$  y  $\sigma_1$  dependerán del voltaje, ecuación (14), y tendrían alguna relación con el segmento ( $S_4$ ), lo cual indica que para la distribución normal existe una gran concentración de valores alrededor de la media. Por tanto, al término  $\alpha$  se le puede considerar como una constante cinética y en este modelo tendría relación directa con la estructura del canal iónico representando una fuerza de oposición al movimiento de los iones. Tomando en cuenta el valor de conductancia citado en [6] y la expresión de conductancia (17) se encuentra un valor para  $\alpha$  igual a  $1.116 \times 10^{-7} \text{ Å/mV}$ . Si a este valor lo multiplicamos por 100 mV (voltaje típico externo aplicado a la membrana) implicaría que el valor de  $u$  (distancia media entre obstáculos) es  $1.16 \times 10^{-5} \text{ Å}$  lo cual indica que el ion está en un continuo choque con las paredes u obstáculos todo el tiempo.

Si analizamos el valor medio de la raíz cuadrada del espaciamiento entre obstáculos dentro del canal, ecuación (11). Se puede observar que mientras se incrementa el voltaje este parámetro aumenta de una manera proporcional a la raíz cuadrada del voltaje; aunque estos valores sean muy pequeños, esto implica que la estructura interna del canal está cambiando y que los iones

pueden pasar por el canal con cierta facilidad para mayores voltajes ya que la distancia entre los obstáculos va incrementándose. Según Stryer [16] “el canal iónico se hace más corto y ancho cuando se enlazan los cationes: litio, sodio, potasio, rubidio y cesio; lo que pone de manifiesto la naturaleza dinámica de las moléculas transportadoras. Los grupos carboxilo que rodean el poro acuoso forman transitoriamente enlaces coordinados temporales con el cation durante su paso por el conducto”. Por tanto, la distribución de espaciamiento medio entre obstáculos que sé consideró justificaría esta tendencia.

El modelo 2 propuesto considera al canal iónico como una estructura periódica unidimensional, donde los dipolos eléctricos (representados por los radicales amino o carboxilo de la estructura del canal), forman dicha estructura. En este caso se asume que el ion se moverá en un potencial periódico dentro del canal. Este potencial se ve reflejado en la masa efectiva de acuerdo con el orbital donde se encuentre. Suponiendo que dicho ion está moviéndose en un medio acuoso, se plantea usar la ecuación de Langevin; con la diferencia de que ahora la masa considerada es la masa efectiva. Tomando el valor conductancia igual a 55pS obtenemos una masa efectiva ( $m^*$ ) de  $6.166 \times 10^{-25}$  Kg, 10 veces mayor que la masa real del ion ( $6.5 \times 10^{-26}$  Kg). La relación de dispersión  $\epsilon(k)$  del ion en la red periódica, permitiría obtener información acerca de la estructura del canal en base a este modelo.

Los potenciales de interacción no son conocidos y por lo tanto, este modelo está fuera del alcance del presente trabajo, pero podría considerarse en un futuro.

#### 4. CONCLUSIONES

En el estudio de los procesos de transporte en una membrana celular es necesario un conocimiento más detallado de la estructura y las funciones de los diferentes tipos de canales iónicos que se encuentran en la membrana; estos conocimientos ayudarían a comparar propiedades físicas como: conductancia, permeabilidad, probabilidades de apertura y cierre de canales; con modelos teóricos planteados.

Los modelos desarrollados muestran de forma sencilla como modelar la conducción iónica a través de la membrana celular. En el presente trabajo se ha considerado el caso más simple donde tenemos condiciones simétricas de concentración (iones potasio), y debido a la escasez de datos experimentales, no se puede verificar el éxito de estas modelizaciones.

En el modelo que se considera una estructura periódica de átomos, presenta una posible solución al problema de la conducción de iones en la membrana; la interpretación de la masa efectiva y su relación con su estructura periódica podrían simular el ordenamiento de dipolos eléctricos dentro del canal. El modelo propuesto ha servido para observar el comportamiento de la conductancia y además se ha podido comparar, obteniéndose resultados casi similares, pero como se trata de un modelo básico tiene ciertas limitaciones.

REFERENCIAS BIBLIOGRAFICAS Y ELECTRÓNICAS

- [1]. **ASHCROFT, N.W. MERMIN, N.D.** (1976). *“Solid State Physics. Saunders College Publishing”*.
- [2]. **ALLEN, T.W. HOYLES, M. KUYUCAK, S.** (1999). *“Molecular and Brownian dynamics study of ion selectivity and conductivity in the potassium channel”*. Chemical Physics Lectures 313, 358-365.
- [3]. **BEZANILLA, F.** *“Canales Iónicos dependientes del potencial eléctrico”*. Capítulo 10.
- [4]. **BLATT, M. THIEL, G.** (1993). *“Hormonal control of ion channel gating”*. Annu. Rev. Physiol. Plant. Biol. 44:543-67.
- [5]. **CHUNG, S.H. HOYLES, M. ALLEN, T.** *“Study of Ionic Currents across a Model Membrane Channel Using Brownian Dynamics”*. Biophysical Journal 75:793-809.
- [6]. **CHUNG, S.H. ALLEN, T. HOYLES, M. KUYUCAK, S.** *“Permeation of Ions Across the Potassium Channel: Brownian Dynamics Studies”*. Biophysical Journal. 77:2517-2533.
- [7]. **GARRAHAN, P.J Y REGA, A.F.** (1977) *“Transporte a través de la membrana celular”*. Programa Regional del Desarrollo Científico y Tecnológico. Secretaría General de la O.E.A. Washington D.C.
- [8]. **GONZALES IBEAS, J.** (1974) *“Introducción a la Física y Biofísica”*. Editorial Alhambra.
- [9]. **KANDEL, E. SCHWARTZ, J. JESSELL, T.** *“Neurociencia y Conducta”*. Prentice Hall.
- [10]. **POMÉS, R. ROUX, B.** *“Structure and Dynamics of a Proton Wire”*, Biophysical Journal 71:19-39.
- [11]. **ROUX, B.** *“Statistical Mechanical Equilibrium Theory of Selective Ion Channels”*. Biophysical Journal 77: 139-153.
- [12]. **ROUX, B.** *“Statistical Mechanical Equilibrium Theory of Selective Ion Channels”*. Biophysical Journal 68: 876-892.
- [13]. **SHEELER, P. BIANCHI, D.** *“Biología Celular”*. Editorial Limusa, S.A.
- [14]. **VILLE CLAUDE A.** *“Biología”*, Segunda Edición, Interamericana. McGraw-Hill.
- [15]. **VOLKENSHTEIN, M.V.** (1985) *“Biofísica”*, Editorial MIR, Moscú.
- [16]. **STRYER, L.** (1985) *“Bioquímica”*, Editorial Reverté, s.a.
- [17]. **CÓRDOBA, C. LEGAZ, M.** *“Biofísica”*. Editorial Síntesis.
- [18]. **HOYLES, M. KUYUCAK, S. CHUNG, SH.** (1998). *“Computer simulation of ion conductance in membrane channels. Physical Review E”*. Volume 58, Number 3.
- [19]. **MCKELVEY, J.** (1991). *“Física del Estado Sólido y de Semiconductores”*. Editorial Limusa, S.A.
- [20]. **IZA, P.** (2002). *“Modelización del transporte de iones potasio a través de un canal iónico en condiciones simétricas de concentración”*, Tesis previa a la obtención del título de Físico, mención aplicaciones a la biología.

## **HOMOTOPÍA Y CONTINUACIÓN NUMÉRICA EN SISTEMAS NO LINEALES**

**Martín Carlos<sup>1</sup>**

**Resumen.** Este artículo presenta la idea general de un método numérico para resolver sistemas no lineales que es distinto a los métodos numéricos iterativos tradicionales y que está basado en el concepto de homotopía y en la continuación numérica. Se demuestra que resolver un sistema no lineal de ecuaciones es equivalente a resolver un problema de valor inicial de primer orden. Se utiliza el método de Runge-Kutta clásico de cuarto orden para encontrar numéricamente la solución del problema de valor inicial. Finalmente, se incluyen algunos resultados sobre la ejecución del método numérico propuesto.

**Palabras Clave:** Sistema de ecuaciones no lineales, homotopía, continuación numérica, problema de valor inicial, métodos de Runge-Kutta.

**Abstract.** This article presents the general idea of a numerical method for solving nonlinear systems which is different from the traditional iterative numerical methods and is based on the concept of homotopy and numerical continuation. We prove that solving a nonlinear system of equations is equivalent to solve an initial value problem of the first order. We use the classic fourth order Runge-Kutta method to numerically solve the initial value problem. Finally, we include some results on the implementation of the proposed numerical method.

**Key Words:** System of nonlinear equations, homotopy, numerical continuation, initial value problem, Runge-Kutta methods.

Recibido: Febrero 2013

Aceptado: Marzo 2013

### **1. INTRODUCCIÓN**

Existen muchos métodos numéricos iterativos convencionales para resolver sistemas de ecuaciones no lineales, entre los que destacan los métodos de punto fijo y los métodos de Newton y sus variantes. Usualmente los métodos iterativos clásicos utilizan fórmulas recursivas que producen una sucesión vectorial que, dadas ciertas condiciones, converge a una solución del sistema no lineal. Para garantizar la convergencia de la sucesión construida, normalmente se necesita, entre otras cosas, que el dato inicial se encuentre “cerca” de una solución del sistema no lineal, además de que las funciones de varias variables que componen el sistema cumplan ciertas condiciones de continuidad y diferenciabilidad. Otra técnica consiste en trabajar sobre un problema equivalente de optimización. Sobre este nuevo problema se pueden aplicar métodos de optimización numérica conocidos, como el método del gradiente o el método BFGS, sin embargo, desde hace menos de una década, se están usando algoritmos metaheurísticos que hacen uso de la computación evolutiva, como los “algoritmos genéticos” y la “búsqueda dispersa”, para resolver el problema equivalente de optimización, con bastante éxito en cuanto al consumo de recursos computacionales y a los resultados obtenidos. En [4] se propone un algoritmo evolutivo para resolver sistemas no lineales.

Lo que se propone en este documento es resolver un sistema no lineal planteando otro tipo de problema equivalente: un problema de valor inicial de primer orden. Esta idea la propuso D. F. Davidenko alrededor de 1960 y actualmente tiene un sinnúmero de aplicaciones. Para esta forma diferente de resolver un sistema no lineal hacemos uso del concepto de homotopía y de la continuación numérica. Se demostrará más adelante que resolver el sistema de ecuaciones diferenciales de primer orden con el dato inicial es equivalente a resolver el sistema no lineal de ecuaciones. Se hará uso del método de cuarto orden de Runge-Kutta clásico para resolver numéricamente el problema de valor inicial. Por lo tanto, se tomará como una solución del sistema no lineal la solución obtenida en el problema de valor inicial. Se concluye este artículo compartiendo algunos experimentos computacionales.

### **2. CONCEPTOS Y FUNDAMENTOS**

**Definición 1.-** Un sistema no lineal de ecuaciones con incógnitas es de la forma:

$$\begin{cases} f_1(x_1, x_2, x_3, \dots, x_n) = 0 \\ f_2(x_1, x_2, x_3, \dots, x_n) = 0 \\ f_3(x_1, x_2, x_3, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, x_3, \dots, x_n) = 0 \end{cases}$$

O en presentación vectorial  $F(\vec{X}) = \vec{0}$ , donde:

<sup>1</sup> Martín Carlos, M.Sc., Profesor de la Escuela Superior Politécnica del Litoral (ESPOL).  
(e\_mail: cmmartin@espol.edu.ec).

$$X = (x_1, x_2, x_3, \dots, x_n),$$

$$F(X) = \begin{pmatrix} f_1(x_1, x_2, x_3, \dots, x_n) \\ f_2(x_1, x_2, x_3, \dots, x_n) \\ f_3(x_1, x_2, x_3, \dots, x_n) \\ \vdots \\ f_n(x_1, x_2, x_3, \dots, x_n) \end{pmatrix} \text{ y } \vec{0} \text{ es el}$$

vector nulo de  $R^n$ .

Las  $n$  funciones  $f_1, f_2, f_3, \dots, f_n$  son, en general, funciones no lineales de  $n$  variables reales independientes y que tienen un dominio en común el cual se va a denotar como  $\Omega \subseteq R^n$ . Es decir, para cada  $i = 1, 2, 3, \dots, n$ :

$$f_i : \Omega \mapsto R$$

$$X = (x_1, x_2, x_3, \dots, x_n) \rightarrow f_i(X)$$

**Definición 2.-** Suponga que las funciones  $F, G : \Omega \subseteq R^n \rightarrow R^n$  son continuas. Se dice que  $F$  y  $G$  son homotópicas, si existe una función continua  $H : \Omega \times [0,1] \rightarrow R^n$  tal que para toda  $X \in \Omega$  ocurre que  $H(X, 0) = G(X)$  y  $H(X, 1) = F(X)$ . Si tal es el caso, llamaremos a la función  $H$  una homotopía de  $G$  hacia  $F$ .

La idea del concepto de homotopía es que una función continua  $G$  se pueda “deformar continuamente” hasta convertirse en la función continua  $F$ . La función que permite esta transformación es la función continua  $H$  mediante el parámetro  $\lambda \in [0,1]$ . Se podría definir  $H$  de muchas maneras, pero es muy común trabajar con homotopías convexas y definir  $H$  mediante la expresión:

$$H(X, \lambda) = \lambda F(X) + (1 - \lambda)G(X)$$

De esta manera, conforme  $\lambda$  varía continuamente de 0 hasta 1, la función  $H$  permite que  $G$  se deforme de manera continua hasta convertirse en  $F$ .

¿Cómo este concepto nos puede ayudar para resolver el sistema no lineal  $F(X) = \vec{0}$ ? La idea es la siguiente: debemos partir de un sistema no lineal  $H(X, 0) = G(X) = \vec{0}$  del cual se conozca una solución  $\tilde{X}$ . A medida que “movemos continuamente”  $\lambda$  desde el valor 0 hasta el valor 1, y tomando como base de partida

la solución  $\tilde{X}$ , vamos resolviendo toda la familia de sistemas no lineales  $H(X, \lambda) = \vec{0}$  hasta encontrar finalmente la solución deseada  $X^*$  del sistema no lineal  $H(X, 1) = F(X) = \vec{0}$ .

¿Qué sistema  $G(X) = \vec{0}$  utilizar? Tenemos también muchas respuestas a esta pregunta, pero es típico definir para algún  $\tilde{X} \in \Omega$ :

$$G(X) = F(X) - F(\tilde{X})$$

Obviamente, debido a la manera en que se construyó la función  $G$ , es de observar que  $\tilde{X}$  es una solución del sistema  $G(X) = \vec{0}$ . Note que el sistema  $H(X, \lambda) = \vec{0}$  define, bajo ciertas condiciones, una curva continua y diferenciable  $X = \psi(\lambda)$  en  $R^n$  con punto inicial en  $\tilde{X}$  cuando  $\lambda = 0$  y punto final en  $X^*$  cuando  $\lambda = 1$ . Es decir, se tiene la función:

$$\psi : [0, 1] \rightarrow R^n$$

$$\lambda \rightarrow X = \psi(\lambda)$$

La continuidad y la diferenciabilidad son condiciones deseadas sobre la función  $X = \psi(\lambda)$  en todo su dominio. Consideremos las siguientes preguntas: ¿Cómo se puede garantizar que la curva  $X = \psi(\lambda)$  existe, es única y es de clase  $C^1$ ? ¿Cómo obtener la curva  $X = \psi(\lambda)$  numéricamente? ¿Qué garantía existe para que podamos afirmar que  $X^* = \psi(1)$  es una solución del sistema no lineal  $H(X, 1) = F(X) = \vec{0}$ ? Para responder a estas preguntas necesitamos algunas definiciones y teoremas que se presentan a continuación.

**Definición 3.-** Sea  $H : U \subseteq R^n \times R \rightarrow R^n$  una función con  $U$  conjunto abierto. Suponga que existe un punto  $(X_0, \lambda_0) \in U$  tal que  $H(X_0, \lambda_0) = \vec{0}$ . Se dice que  $H(X, \lambda) = \vec{0}$  define a  $X$  como una función implícita de  $\lambda$  en un entorno de  $(X_0, \lambda_0)$  si existe un conjunto abierto  $V \subset R^n$  tal que  $X_0 \in V$ , un intervalo

abierto  $(\lambda_0 - \rho, \lambda_0 + \rho)$  con  $\rho > 0$  como radio tal que  $V \times (\lambda_0 - \rho, \lambda_0 + \rho) \subset U$ , y una única función  $\psi : (\lambda_0 - \rho, \lambda_0 + \rho) \rightarrow V$  tal que para todo  $\lambda \in (\lambda_0 - \rho, \lambda_0 + \rho)$  la función  $X = \psi(\lambda)$  es solución del problema

$$H(X, \lambda) = \vec{0} \quad (\text{Es decir que})$$

$$H(\psi(\lambda), \lambda) = \vec{0})$$

Según el concepto anterior, si se cumplen ciertas condiciones, se puede garantizar la existencia de la función  $X = \psi(\lambda)$  que define una curva en

$R^n$  con extremos en  $\tilde{X}$  con  $\lambda = 0$  y  $X^*$  con  $\lambda = 1$ . Si ahora enunciamos el conocido teorema de la función implícita aplicado a nuestra función  $H$ , dicho enunciado quedaría como se detalla a continuación.

**Teorema 1.-** Sea  $H : U \subset R^n \times R \rightarrow R^n$  una función de clase  $C^1(U)$  con  $U$  conjunto abierto. Si se tiene el punto  $(X_0, \lambda_0) \in U$  tal que  $H(X_0, \lambda_0) = \vec{0}$ , y además  $H_x(X_0, \lambda_0)$  es una matriz inversa, entonces la ecuación  $H(X, \lambda) = \vec{0}$  define implícitamente la función  $X = \psi(\lambda)$  en un entorno abierto del punto  $(X_0, \lambda_0)$ . Además,  $\psi$  es también una función de clase  $C^1$  en algún intervalo alrededor de  $\lambda_0$  y es única.

Este teorema es muy importante porque proporciona condiciones suficientes para garantizar la existencia y unicidad de la curva  $X = \psi(\lambda)$  en el intervalo  $\lambda \in [0,1]$ . No sólo eso, también garantiza que  $X = \psi(\lambda)$  es de clase  $C^1[0,1]$ . Dependiendo del problema a resolver, puede ocurrir que  $U = R^n \times R$ . Lo importante es que el dominio de la función  $H$  de donde se considera la existencia, unicidad, continuidad y diferenciabilidad continua de  $X = \psi(\lambda)$  contenga al conjunto  $\Omega \times [0,1]$  de nuestro interés. Dependiendo del problema, podría no ocurrir esto. Es importante entonces estudiar cómo se comporta la curva para el problema en particular que se desea resolver. La curva podría no llegar a alcanzar el punto para el cual  $\lambda = 1$  o podría recorrer una distancia

“infinita” desde  $\lambda = 0$  hasta  $\lambda = 1$  debido a su longitud no finita.

Procedemos ahora a enunciar el teorema de la función inversa para funciones de un subconjunto de  $R^n$  en  $R^n$ . Este teorema es de gran utilidad para discutir una proposición que se enunciará y demostrará más adelante.

**Teorema 2.-** Sea  $F : U \subseteq R^n \rightarrow R^n$  una función de clase  $C^1$ . Sea  $X_0 \in U$  tal que  $F'(X_0)$  es inversible y  $F(X_0) = Y_0$ . Entonces, existen conjuntos abiertos  $V$  y  $W$  tales que  $X_0 \in V$ ,  $Y_0 \in W$  y  $F : V \rightarrow W$  es una función biyectiva. Por lo tanto la función inversa  $F^{-1} : W \rightarrow V$  existe y es también de clase  $C^1$ . Además, ocurre que:  $(F^{-1})'(Y_0) = [F'(X_0)]^{-1}$

**Definición 4.-** Se dice que la función  $F : R^n \rightarrow R^n$  es un homeomorfismo si se cumple que  $F$  es una función biyectiva continua con función inversa  $F^{-1}$  también continua.

### 3. EL PROBLEMA DE VALOR INICIAL

A continuación se va a enunciar y a demostrar formalmente un teorema que justifica el plantear, como problema equivalente de un sistema no lineal de ecuaciones, un problema de valor inicial. Este teorema nos permite tomar la solución del problema de valor inicial como una solución del sistema de ecuaciones no lineales.

**Teorema 3.-** Sea  $H(X, \lambda)$  una función de clase  $C^1$  y la matriz Jacobiana  $H_x(X, \lambda)$  inversible. Suponga también que  $X = \psi(\lambda)$  es una función de clase  $C^1[0,1]$ . La función  $X = \psi(\lambda)$  es una solución de la ecuación

$H(X, \lambda) = \vec{0}$  para todo  $0 \leq \lambda \leq 1$ , si y sólo si,  $X = \psi(\lambda)$  es una solución del problema de valor inicial

$$X'(\lambda) = -[H_x(X, \lambda)]^{-1} H_\lambda(X, \lambda) \quad \text{con}$$

$$H(X(0), 0) = \vec{0} \quad \text{y además } \lambda \in [0,1]$$

Demo

Primera Parte: Supongamos que la función  $X = \psi(\lambda)$  satisface la ecuación

$H(X, \lambda) = \vec{0}$  para todo  $\lambda \in [0,1]$  y definase la función  $\phi(\lambda) = H(X = \psi(\lambda), \lambda)$ . Podemos derivar entonces aplicando la regla de la cadena, respecto de la variable  $\lambda$ , ambos lados de la ecuación:  $\phi(\lambda) = H(X = \psi(\lambda), \lambda) = \vec{0}$

Entonces:

$$\begin{aligned}\phi'(\lambda) &= H_x(X = \psi(\lambda), \lambda) \psi'(\lambda) \\ &+ H_\lambda(X = \psi(\lambda), \lambda) = \vec{0}\end{aligned}$$

Al despejar  $\psi'(\lambda)$  de la ecuación anterior, suponiendo que  $H_x(X, \lambda)$  tiene inversa, tenemos:

$$\psi'(\lambda) = -[H_x(X = \psi(\lambda), \lambda)]^{-1} H_\lambda(X = \psi(\lambda), \lambda)$$

El  $X$  que satisface la condición inicial  $H(X(0), 0) = \vec{0}$  es la solución del sistema

$$H(X(0), 0) = G(X) = \vec{0}, \quad \text{es decir,} \\ \tilde{X} = \psi(0)$$

Segunda Parte: Lo que debemos probar ahora es que  $X = \psi(\lambda)$  satisface la ecuación

$$H(X, \lambda) = \vec{0} \quad \text{para todo } 0 \leq \lambda \leq 1, \\ \text{suponiendo que se cumple} \\ \psi'(\lambda) = -[H_x(\psi(\lambda), \lambda)]^{-1} H_\lambda(\psi(\lambda), \lambda)$$

con  $\psi(0) = \tilde{X}$  y  $\lambda \in [0,1]$ .

Entonces:

$$[H_x(\psi(\lambda), \lambda)] \psi'(\lambda) = -H_\lambda(\psi(\lambda), \lambda)$$

Luego:

$$[H_x(\psi(\lambda), \lambda)] \psi'(\lambda) + H_\lambda(\psi(\lambda), \lambda) = \vec{0}$$

$$\text{Después: } \frac{d}{d\lambda} [H(X = \psi(\lambda), \lambda)] = \vec{0}$$

Se concluye entonces que  $H(X = \psi(\lambda), \lambda) = \vec{C}$  donde  $\vec{C}$  es un vector constante de  $R^n$  para todo  $0 \leq \lambda \leq 1$ . Pero por la condición inicial  $H(X(0), 0) = \vec{0}$ , el vector constante  $\vec{C}$  es el vector nulo  $\vec{0}$ . Por lo tanto  $H(X = \psi(\lambda), \lambda) = \vec{0}$  para todo  $\lambda \in [0,1]$ .

Otra demostración para esta segunda parte de la prueba es la siguiente: Apliquemos el teorema del valor medio a la función vectorial  $\phi(\lambda) = H(X = \psi(\lambda), \lambda)$ . Debido a que  $\phi$  es una función vectorial de clase  $C^1$  en el intervalo  $[0,1]$ , se procede a fijar un  $\lambda$  en el intervalo  $(0,1]$  y aplicamos el teorema del valor medio para funciones vectoriales en el intervalo  $[0, \lambda]$  donde obviamente la función vectorial  $\phi$  es también de clase  $C^1$ . Entonces, tendríamos  $\|\phi(\lambda) - \phi(0)\| \leq (\lambda - 0) \|\phi'(\lambda)\|$  para algún  $t$  entre 0 y  $\lambda$ . Pero por la hipótesis que  $X = \psi(\lambda)$  satisface la ecuación diferencial para todo  $\lambda$  entre 0 y 1 se tiene que

$$\phi'(t) = H_X(X = \psi(t), t) \psi'(t) + H_\lambda(X = \psi(t), t) = \vec{0}$$

para cualquier  $t$  entre 0 y  $\lambda$ . Además, de la condición inicial tenemos

$$H(X = \psi(0), 0) = \vec{0} = \phi(0). \quad \text{Entonces:}$$

$$\|\phi(\lambda)\| \leq (\lambda - 0) \|\phi'(\lambda)\| = 0$$

Se puede concluir de la última expresión que

$$\phi(\lambda) = \vec{0}, \quad \text{pero } \phi(\lambda) = H(X = \psi(\lambda), \lambda).$$

Por lo tanto  $H(X = \psi(\lambda), \lambda) = \vec{0}$  para  $\lambda \in [0,1]$ .

Con la definición de

$$G(X) = F(X) - F\left(\tilde{X}\right), \quad \text{la ecuación para la}$$

función  $H(X, \lambda)$  es la siguiente:

$$H(X, \lambda) = \lambda F(X) + (1 - \lambda) G(X) = \vec{0}$$

$$H(X, \lambda) = \lambda F(X) + (1 - \lambda) \left[ F(X) - F\left(\tilde{X}\right) \right] = \vec{0}$$

$$H(X, \lambda) = F(X) + (\lambda - 1) F\left(\tilde{X}\right) = \vec{0}$$

Entonces, la ecuación que define implícitamente a la curva  $X = \psi(\lambda)$  es:

$$F(X) = (1 - \lambda) F\left(\tilde{X}\right)$$

Al derivar ambos lados de la ecuación anterior respecto de  $\lambda$ , aplicando regla de la cadena y suponiendo que el Jacobiano de  $F(X)$  es no singular, para obtener el problema de valor inicial equivalente al sistema no lineal, tenemos

$F'(X)X'(\lambda) = -F(\tilde{X})$ , de donde la ecuación diferencial es  $X'(\lambda) = -[F'(X)]^{-1}F(\tilde{X})$  con dato inicial  $X(0) = \tilde{X}$ .

**Teorema 4.-** Sea  $F: R^n \rightarrow R^n$  una función de clase  $C^1$ . Suponga que el Jacobiano  $F'$  es no singular para todo  $X \in R^n$  y que  $F$  es un homeomorfismo. Entonces, para cualquier

$\tilde{X} \in R^n$  existe una única función  $X(\lambda): [0,1] \rightarrow R^n$  de clase  $C^1[0,1]$  tal que  $F(X) = (1-\lambda)F(\tilde{X})$ . Además, ocurre que

$X'(\lambda) = -[F'(X)]^{-1}F(\tilde{X})$  para todo  $0 \leq \lambda \leq 1$  y  $X(0) = \tilde{X}$ .

#### Demostración

Recordar que de la ecuación  $H(X, \lambda) = \lambda F(X) + (1-\lambda) \left[ F(X) - F(\tilde{X}) \right] = 0$  se obtiene que  $F(X) = (1-\lambda)F(\tilde{X})$ . Pero

por hipótesis  $F$  es un homeomorfismo y por ende  $F^{-1}$  existe y es continua. Además, por el teorema de la función inversa  $F^{-1}$  es también de clase  $C^1$  ya que  $F$  es de clase  $C^1$  y  $F'$  es no singular para todo  $X \in R^n$ . Entonces, para cada  $\lambda \in [0,1]$  se tiene la solución única:

$$X(\lambda) = F^{-1} \left[ (1-\lambda)F(\tilde{X}) \right]$$

$$\text{Note que } X(0) = F^{-1} \left[ F(\tilde{X}) \right] = \tilde{X}$$

Finalmente, derivando ambos lados de la ecuación anterior respecto de  $\lambda$ , aplicando la regla de la cadena y haciendo uso también del teorema de la función inversa, tenemos:

$$\begin{aligned} X'(\lambda) &= D_\lambda F^{-1} \left[ (1-\lambda)F(\tilde{X}) \right] \\ &= \left( F^{-1} \right)' \left[ (1-\lambda)F(\tilde{X}) \right] D_\lambda \left[ (1-\lambda)F(\tilde{X}) \right] \end{aligned}$$

$$X'(\lambda) = [F'(X)]^{-1} \left[ -F(\tilde{X}) \right]$$

$$X'(\lambda) = -[F'(X)]^{-1} F(\tilde{X})$$

Por tanto se tiene la ecuación diferencial

$$X'(\lambda) = -[F'(X)]^{-1} F(\tilde{X}) \text{ para todo}$$

$0 \leq \lambda \leq 1$  con dato inicial  $X(0) = \tilde{X}$  con solución única de clase  $C^1[0,1]$

## 4. EL MÉTODO NUMÉRICO

Existen muchos métodos numéricos para resolver problemas de valor inicial, pero los métodos de Runge-Kutta son muy populares por su facilidad de implementación y por el orden de error local y global que presentan. Se usará entonces el método clásico de cuarto orden de Runge-Kutta para resolver el problema de valor inicial antes mencionado.

Sin embargo, suponga por un momento que se desea implementar el método de Euler que presenta un error local  $O(h^2)$  y un error global  $O(h)$ . La ecuación recursiva que resuelve el problema de valor inicial es:

$$X_{i+1} = X_i - h [F'(X_i)]^{-1} F(X_0); \\ i = 0, 1, 2, \dots, n-1$$

Recuerde que el método numérico discretiza el problema y la variable  $h = 1/n$  representa el tamaño de paso. El número de subintervalos en el cual es particionado el intervalo  $[0,1]$  es representado por la variable  $n$ . Además

$X_0 = X(0) = \tilde{X}$  y el valor de  $X_n$  es la estimación numérica de  $X(1) = X^*$ .

Si ahora introducimos una nueva variable independiente  $t$  y la relacionamos con la variable independiente  $\lambda$  de la ecuación diferencial

$$X'(\lambda) = -[F'(X)]^{-1} F(\tilde{X}) \text{ mediante la}$$

ecuación  $t = -\ln(1-\lambda)$ , al hacer el cambio de variable tendríamos:

$$X'(\lambda) = X'(t) \cdot \frac{dt}{d\lambda} = X'(t) \cdot \frac{1}{1-\lambda} = -[F'(X)]^{-1} F(\tilde{X})$$

Finalmente:

$$X'(t) = -(1-\lambda)[F'(X)]^{-1} F(\tilde{X})$$

$$\text{Pero antes vimos que: } F(X) = (1-\lambda)F(\tilde{X})$$

Por lo tanto, la nueva ecuación diferencial ahora con  $t$  como variable independiente es:

$$X'(t) = -[F'(X)]^{-1} F(X)$$

Observe que si  $\lambda = 0$ , entonces  $t = 0$ . Además, si  $\lambda \rightarrow 1^-$ , entonces  $t \rightarrow +\infty$ . Por lo tanto, el nuevo problema de valor inicial equivalente a nuestro problema de valor inicial original es

$$X'(t) = -[F'(X)]^{-1} F(X) \quad \text{para}$$

$t \in [0, +\infty]$  con  $X(0) = \tilde{X}$ . Si ahora aplicamos el método de Euler con tamaño de paso  $h = 1$  tenemos la ecuación recursiva:

$$X_{i+1} = X_i - [F'(X_i)]^{-1} F(X_i); \\ i = 0, 1, 2, \dots$$

Note que lo que se ha obtenido por tanto es el método iterativo convencional de Newton para resolver sistemas de ecuaciones no lineales que, bajo ciertas condiciones, converge a la solución deseada  $X^*$ . Es de observar que si vemos al método de Newton desde esta última perspectiva, se ha tomado un tamaño de paso que no es bueno, junto con el método de Euler cuya aproximación es pobre en comparación con la diversidad de métodos numéricos existentes para resolver ecuaciones diferenciales que presentan mejores estimaciones de los valores exactos.

El método de Runge-Kutta que usaremos presenta un error local  $O(h^5)$  y un error global  $O(h^4)$ . Al discretizar el problema con un tamaño de paso  $h = 1/n$ , donde  $n$  es el número de subintervalos, tenemos la ecuación recursiva:

$$X_{i+1} = X_i + \frac{h}{6} (K_1 + 2K_2 + 2K_3 + K_4); \\ i = 0, 1, 2, \dots, n-1$$

Las funciones  $K_i$  son:

$$K_1 = -[F'(X_i)]^{-1} F(\tilde{X})$$

$$K_2 = -\left[ F'\left(X_i + \frac{h}{2} K_1\right) \right]^{-1} F(\tilde{X})$$

$$K_3 = -\left[ F'\left(X_i + \frac{h}{2} K_2\right) \right]^{-1} F(\tilde{X})$$

$$K_4 = -[F'(X_i + h K_3)]^{-1} F(\tilde{X})$$

La estimación de  $X(1) = X^*$  es el valor de  $X_n$ . Se hicieron pruebas con sistemas no lineales de hasta 5 ecuaciones con 5 incógnitas con excelentes resultados que se muestran en la siguiente sección de este documento. Cabe indicar que la precisión obtenida además del tamaño de paso  $h$ , depende del valor inicial  $X(0) = \tilde{X}$ .

## 5. SOLUCIÓN DE ALGUNOS SISTEMAS NO LINEALES

Como un primer ejemplo tomemos la función de variable real  $f(x) = x^3 - e^x \sin(x)$  con derivada

$$f'(x) = 3x^2 - e^x [\sin(x) + \cos(x)].$$

Luego tenemos que  $H(x, \lambda) = x^3 - e^x \sin(x) + (\lambda - 1)(x_0^3 - e^{x_0} \sin(x_0))$  con  $x_0 \in R$ . Finalmente, el problema de valor inicial correspondiente es:

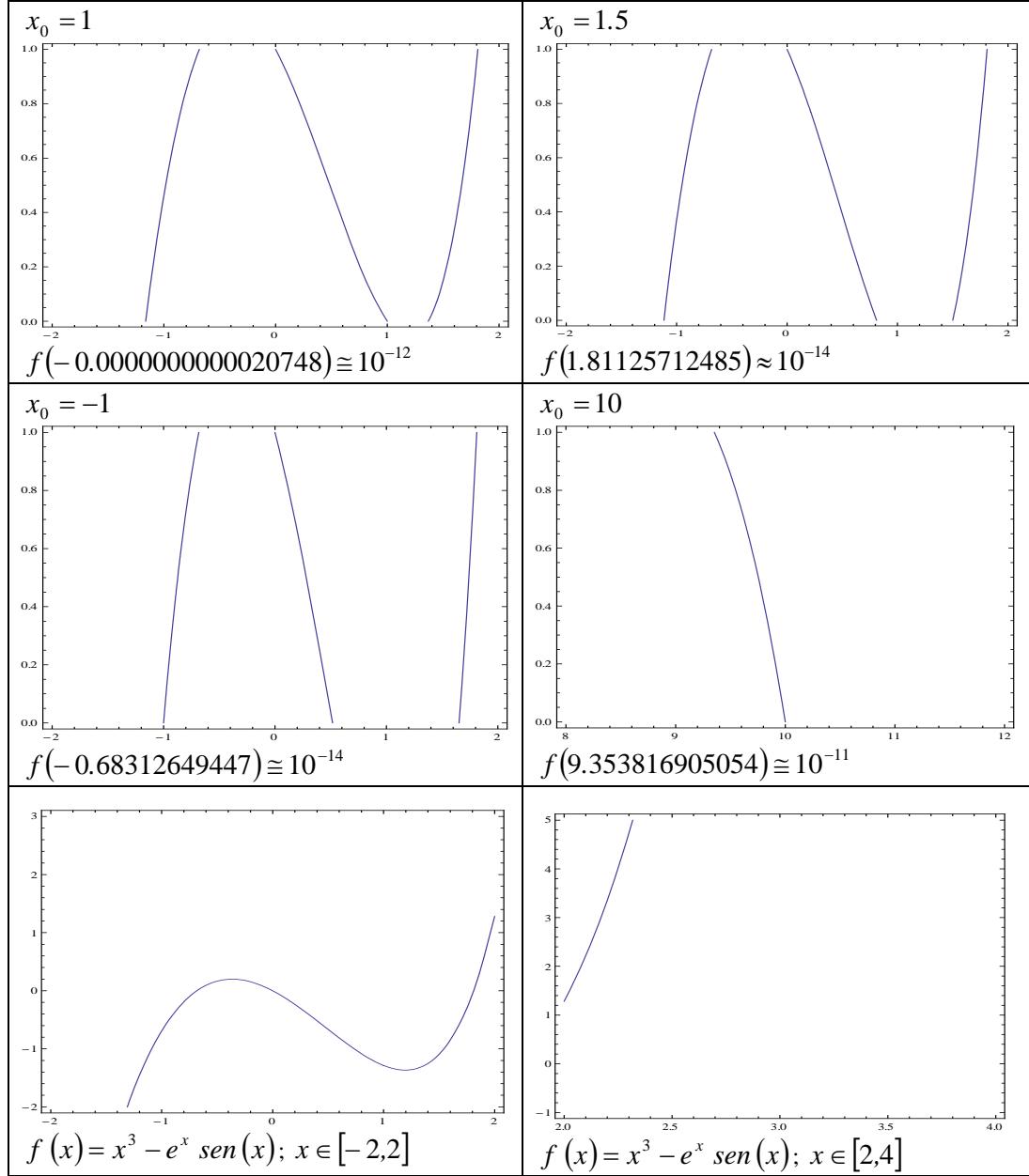
$$x'(\lambda) = \frac{e^{x_0} \sin(x_0) - x_0^3}{3x^2 - e^x [\sin(x) + \cos(x)]} \quad \text{para}$$

todo  $0 \leq \lambda \leq 1$  y  $x(0) = x_0$ .

A continuación, en la figura # 1, se muestra la curva de nivel  $H(x, \lambda) = 0$  en  $R^2$  con distintos valores para  $x_0$ . Se incluyen los resultados obtenidos al ejecutar el método numérico de Runge-Kutta de cuarto orden en cada caso. También se incluye la gráfica de la función  $f(x) = x^3 - e^x \sin(x)$  en distintos intervalos.

El tamaño de paso  $h = 0.001$  se utilizó en todas las ejecuciones. En el eje horizontal hemos ubicado los valores que toma la variable  $x$ , mientras que los valores de la variable  $\lambda \in [0, 1]$  se encuentran en el eje vertical. Note que  $f$  tiene muchos "ceros".

**FIGURA 1**  
*Curvas de nivel de  $H$  y Gráfico de  $f$*   
**Homotopía y continuación numérica en sistemas no lineales**



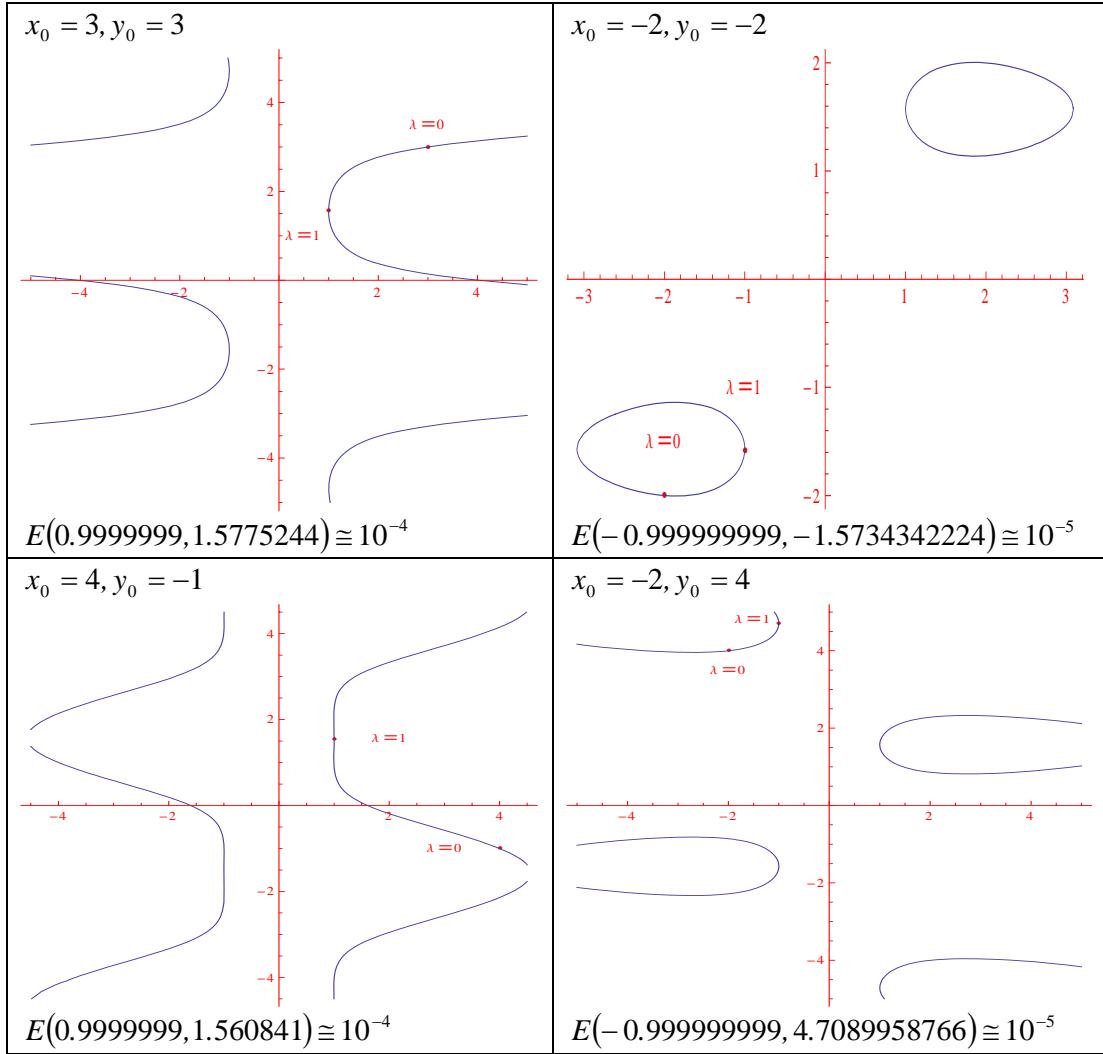
Como un segundo ejemplo, veamos lo que ocurre con el sistema no lineal:

$$\begin{cases} x \sin(y) - 1 = 0 \\ x^2 + \cos(2y) = 0 \end{cases}$$

Recuerde que el valor de  $X_n$  que entrega el método de Runge-Kutta es la estimación del valor exacto  $X^*$ . De aquí en adelante usaremos la siguiente expresión para calcular el error:

$E(X_n) = \sum_{i=1}^n |f_i(X_n)|$  donde las  $f_i$  son las funciones componentes del sistema no lineal. Los valores de la variable  $x$  se encuentran en el eje horizontal, los valores de la variable  $y$  en el eje vertical. Se utilizó  $h = 0.001$  como tamaño de paso. En la figura # 2 se muestra, para cada corrida, la trayectoria homotópica de las soluciones.

**FIGURA 2**  
 Trayectoria de las Soluciones Segundo Ejemplo  
 Homotopía y continuación numérica en sistemas no lineales



Otro ejemplo es el siguiente sistema no lineal:

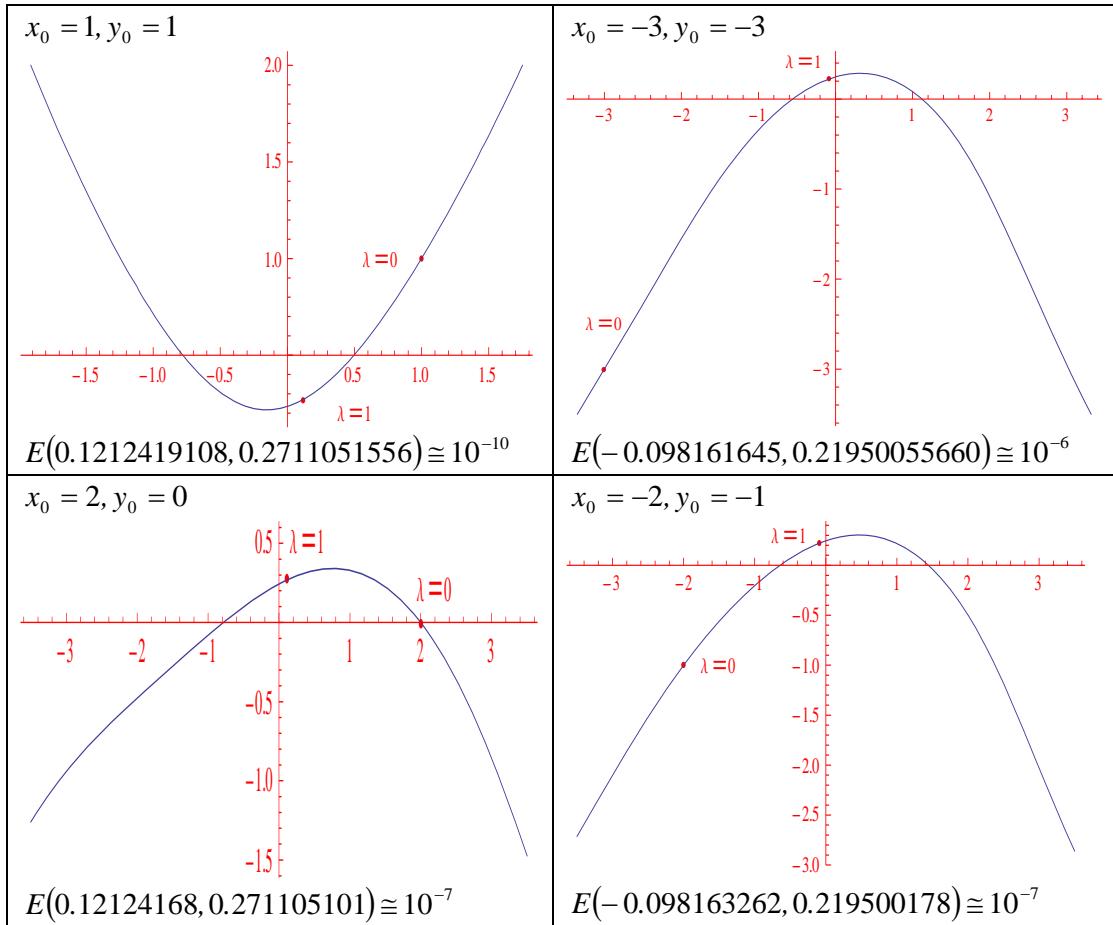
$$\begin{cases} 5x^2 - y^2 = 0 \\ 4y - \sin(x) - \cos(y) = 0 \end{cases}$$

También con  $h = 0.001$  como tamaño de paso, se ejecutó el método numérico con distintos

valores iniciales  $(x_0, y_0)$ . En la figura # 3 se muestra, para cada ejecución, la trayectoria homotópica de las soluciones de cada uno de los sistemas no lineales que se van resolviendo desde

$\lambda = 0$  hasta  $\lambda = 1$ .

**FIGURA 3**  
*Trayectoria de las Soluciones Tercer Ejemplo*  
**Homotopía y continuación numérica en sistemas no lineales**



Consideremos ahora el sistema no lineal:

$$\begin{cases} (x-2)^2 + (y-1)^2 + x y - 3 = 0 \\ x e^{x+y} + y z - 3 = 0 \\ \sin(x+z) + x - y - z + 1 = 0 \end{cases}$$

Con el mismo tamaño de paso  $h = 0.001$  usado en los ejemplos anteriores, se muestran los resultados de dos ejecuciones del método numérico para este cuarto ejemplo.

**TABLA I**  
*Resultado de dos ejecuciones del método numérico con  $h = 0.001$*   
**Homotopía y continuación numérica en sistemas no lineales**

$x_0 = 1, y_0 = 1, z_0 = 1$ $x_n = 0.3885222610700$ $y_n = 1.0345501950139$ $z_n = 1.341346675089$ $E(x_n, y_n, z_n) \approx 10^{-13}$	$x_0 = 4, y_0 = -2, z_0 = 7$ $x_n = 3.039247388237$ $y_n = -1.610482827797$ $z_n = 6.013364389595$ $E(x_n, y_n, z_n) \approx 10^{-13}$
--	--

Finalmente, tenemos el sistema no lineal:

$$\begin{cases} x^2 + 2y^2 + \cos(z) - w^2 = 0 \\ 3x^2 + y^2 + \operatorname{sen}^2(z) - w^2 = 0 \\ -2x^2 - y^2 - \cos(z) + w^2 = 0 \\ -x^2 - y^2 - 2\cos^2(z) + w^2 = 0 \end{cases}$$

Se muestran también un par de ejecuciones de este sistema con  $h = 0.00001$  como tamaño de paso.

**TABLA II**  
*Resultado de dos ejecuciones del método numérico con  $h = 0.00001$*   
**Homotopía y continuación numérica en sistemas no lineales**

$x_0 = 3, y_0 = 3, z_0 = 3, w_0 = 3$	$x_0 = -2, y_0 = -2, z_0 = -2, w_0 = -2$
$x_n = 0.998525959$	$x_n = -0.9989511610$
$y_n = 0.998525959$	$y_n = -0.9989511610$
$z_n = 0.0456785223$	$z_n = -0.0385347375$
$w_n = 1.9975435860$	$w_n = -1.9982520966$
$E(x_n, y_n, z_n) \approx 10^{-4}$	$E(x_n, y_n, z_n) \approx 10^{-4}$

Cabe mencionar que todos los ejemplos que se mencionan en este documento fueron tomados de [4]. Se realizaron pruebas con sistemas de hasta cinco ecuaciones con cinco incógnitas y con otros sistemas no lineales de menor tamaño, que no se muestran en este documento. En todos ellos se trabajó con  $h = 10^{-3}, 10^{-4}, 10^{-5}$  como tamaño de paso.

## 6. CONCLUSIONES Y RECOMENDACIONES

Los métodos numéricos iterativos convencionales, como los métodos de Newton, para resolver sistemas no lineales, siempre deben ser la primera opción cuando se tiene algún tipo de sospecha de dónde se encuentra alguna solución, por su precisión y rápida convergencia, siempre que las funciones que componen el sistema cumplan con las hipótesis de algún teorema de convergencia. Caso contrario, el método de la homotopía y la continuación numérica es una opción válida para resolver sistemas no lineales. Algo positivo de este método es que permite ampliar el intervalo de convergencia, tal como lo hacen los algoritmos metaheurísticos para resolver sistemas no lineales. Si la solución encontrada al resolver la ecuación

diferencial no cumple con algún criterio de precisión requerido, podría tomarse la salida del método homotópico como el “dato inicial” que se necesita en algún método iterativo convencional para tener una mejor estimación de la solución del sistema no lineal. Es decir, ambos tipos de métodos pueden trabajar de forma conjunta para finalmente resolver el sistema no lineal. Cuando se plantea el problema de valor inicial equivalente al sistema no lineal se debe tener cuidado con el

valor  $\tilde{X}$  con el que se desea iniciar el trabajo iterativo porque no necesariamente se va a tener convergencia. Es importante, dependiendo del problema, analizar la curva  $X = \psi(\lambda)$  para, según el dato inicial escogido, estudiar si existe  $\psi(1)$  y si efectivamente este valor está estimando bien alguna solución del sistema no lineal. En [5] se dan sugerencias de cómo escoger el dato inicial analizando regiones de convergencia. El método numérico usado para resolver el problema de valor inicial es de un paso, podrían implementarse métodos mutipaso y hacer un estudio comparativo de los resultados.

## REFERENCIAS BIBLIOGRÁFICAS Y ELECTRÓNICAS

- [1]. **ORTEGA Y RHEINBOLDT** (2001). “*Iterative Solution of Nonlinear Equations in Several Variables*”, SIAM’s Classics in Applied Mathematics, Chapters 5, 7, 10
- [2]. **ALLGOWER Y GEORG** (1990). “*Introduction to Numerical Continuation Methods*”, Colorado State University
- [3]. **HAIRER, NORSETT Y WANNER** (2008). “*Solving Ordinary Differential Equations I*”, Springer Series in Computational Mathematics Chapters 1, 2.
- [4]. **MARTÍN** (2013). “*Sistemas No Lineales y la Metaheurística Scatter Search*”, Revista Matemática Volumen 11, Número 1 Páginas 24 – 33. ESPOL, Guayaquil – Ecuador.
- [5]. **LEE Y CHIANG** (2001). “*Convergent Regions of the Newton Homotopy Method for Nonlinear Systems: Theory and Computational Applications*”, IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications, Vol. 48, No. 1.

## SISTEMAS NO LINEALES Y LA METAHEURÍSTICA SCATTER SEARCH

**Martín Carlos<sup>1</sup>**

**Resumen.** Este artículo presenta un algoritmo metaheurístico que permite encontrar soluciones reales a sistemas de ecuaciones no lineales, en contraste a los métodos numéricos iterativos convencionales. Se demuestra que resolver un sistema no lineal de ecuaciones es equivalente a resolver un problema de optimización, particularmente un problema de minimización. Se utiliza la metaheurística “scatter search” para el problema de optimización. Finalmente, se incluyen algunos resultados sobre la ejecución del algoritmo en sistemas no lineales.

**Palabras Clave:** Sistema de ecuaciones no lineales, scatter search, metaheurística, problema de optimización.

**Abstract.** This article presents a metaheuristic algorithm that allows finding real solutions to systems of nonlinear equations, in contrast to the traditional iterative numerical methods. We prove that solving a nonlinear system of equations is equivalent to solving an optimization problem, particularly a minimization problem. We use the metaheuristic “scatter search” for the optimization problem. Finally, include some results of execution of the algorithm used in nonlinear systems as examples.

**Key Words:** System of nonlinear equations, scatter search, metaheuristic, optimization problem.

Recibido: Febrero 2013

Aceptado: Marzo 2013

### 1. INTRODUCCIÓN

Existen muchos métodos numéricos iterativos para resolver sistemas de ecuaciones no lineales. Los algoritmos que implementan estos métodos normalmente trabajan con matrices, y necesitan además las primeras y segundas derivadas parciales, o sus aproximaciones, de algunas funciones de varias variables. Es necesario también efectuar operaciones matriciales como sumas, productos, incluso en algunos métodos calcular la inversa de matrices, todo esto por cada iteración, lo que hace que el costo computacional sea alto. Son muy conocidos los métodos de Newton y sus variaciones como Newton Amortiguado, Quasi-Newton o Newton Mejorado. Existen también un sinnúmero de métodos de punto fijo. Otra característica de estos métodos iterativos clásicos es que son localmente convergentes, es decir, el punto inicial, con el cual se comienza a iterar, debe estar “algo cerca” de la solución para que el algoritmo converja en un número aceptable de iteraciones. Lo que se propone en este documento es un algoritmo distinto. Lo primero que haremos es expresar el sistema de ecuaciones no lineales en términos de un problema de optimización y, se va a demostrar formalmente, que resolver el problema de optimización es equivalente a resolver el sistema no lineal. Una vez hecho esto, nos enfocaremos en resolver el problema de optimización, que consiste en minimizar una función objetivo que se construye usando cada una de las ecuaciones que componen el sistema no lineal.

Se usará la metaheurística de “scatter search” para resolver el problema de optimización. Se dará una explicación de cada uno de los pasos del algoritmo. Un objetivo secundario es mostrar la metodología y la filosofía de trabajo del algoritmo scatter search cuando se resuelven problemas de optimización. Se podrá apreciar que la mezcla de calidad y diversidad que utiliza el algoritmo es fascinante. Luego, la solución real encontrada para el problema de optimización es también la solución real del sistema no lineal de ecuaciones. Finalmente, se procede a compartir los resultados obtenidos al ejecutar el algoritmo que proponemos sobre algunos sistemas de ecuaciones no lineales que usaremos como ejemplos.

### 2. SISTEMAS DE ECUACIONES NO LINEALES

Un sistema no lineal de ecuaciones con incógnitas es de la forma:

$$\begin{cases} f_1(x_1, x_2, x_3, \dots, x_n) = 0 \\ f_2(x_1, x_2, x_3, \dots, x_n) = 0 \\ f_3(x_1, x_2, x_3, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, x_3, \dots, x_n) = 0 \end{cases}$$

O en presentación vectorial  $F(X) = \vec{0}$ , donde:  
 $X = (x_1, x_2, x_3, \dots, x_n)$ ,

<sup>1</sup> Martín Carlos, M.Sc., Profesor de la Escuela Superior Politécnica del Litoral (ESPOL).  
(e-mail: cmmartin@espol.edu.ec).

$$F(X) = \begin{pmatrix} f_1(x_1, x_2, x_3, \dots, x_n) \\ f_2(x_1, x_2, x_3, \dots, x_n) \\ f_3(x_1, x_2, x_3, \dots, x_n) \\ \vdots \\ f_n(x_1, x_2, x_3, \dots, x_n) \end{pmatrix} \quad \text{y} \quad \vec{0} \text{ es el}$$

vector nulo de  $R^n$

Se va a suponer que las  $n$  funciones  $f_1, f_2, f_3, \dots, f_n$  son, en general, funciones no lineales de  $n$  variables reales independientes y que tienen un dominio en común el cual se va a denotar como  $\Omega \subseteq R^n$ . Es decir, para cada  $i = 1, 2, 3, \dots, n$ :

$$\begin{aligned} f_i : \Omega &\rightarrow R \\ X = (x_1, x_2, x_3, \dots, x_n) &\rightarrow f_i(X) \end{aligned}$$

Se procede ahora a definir la función objetivo  $g$  la cual tiene también obviamente su dominio en el conjunto  $\Omega$ :

$$\begin{aligned} g(X) &= \sum_{i=1}^n |f_i(X)| = \\ &= |f_1(X)| + |f_2(X)| + |f_3(X)| + \dots + |f_n(X)| \end{aligned}$$

### 3. EL PROBLEMA DE OPTIMIZACIÓN

A continuación se va a enunciar y a demostrar matemáticamente una proposición que justifica el plantear un sistema de ecuaciones no lineales en términos de un problema de minimización.

Proposición.- El vector  $X^* = (x_1^*, x_2^*, x_3^*, \dots, x_n^*) \in \Omega$  es una solución de  $F(X) = \vec{0}$  (es decir:  $F(X^*) = \vec{0}$ ), si y sólo si, la función  $g$  (definida anteriormente) toma como valor mínimo en  $X^*$  el valor 0

#### Demostración

Primera Parte: Supongamos que  $F(X^*) = \vec{0}$ . Entonces para cada  $i = 1, 2, 3, \dots, n$  ocurre que  $f_i(X^*) = 0$ . Luego:

$$g(X^*) = \sum_{i=1}^n |f_i(X^*)| = 0$$

Además para toda  $i = 1, 2, 3, \dots, n$  y para toda  $X \in \Omega$  se tiene que  $|f_i(X)| \geq 0$ , por lo que  $g(X) \geq 0$  para toda  $X \in \Omega$ .

Entonces la función  $g$  toma el valor mínimo cero en  $X^* \in \Omega$

Segunda Parte: Ahora supongamos que la función  $g$  toma el valor mínimo cero en  $X^* \in \Omega$ . Se debe entonces demostrar que  $F(X^*) = \vec{0}$ .

Por hipótesis  $g(X^*) = 0$ . Entonces

$$g(X^*) = \sum_{i=1}^n |f_i(X^*)| = 0$$

Es decir

$$|f_1(X^*)| + |f_2(X^*)| + |f_3(X^*)| + \dots + |f_n(X^*)| = 0$$

Se debe probar que para toda  $i = 1, 2, 3, \dots, n$  ocurre que  $f_i(X^*) = 0$ . Hagamos la prueba por contradicción. Supongamos que lo anterior es falso, es decir, que existe una  $1 \leq i \leq n$  tal que  $f_i(X^*) \neq 0$ . Si esto es así, sucede que

$$\begin{aligned} g(X^*) &= \sum_{i=1}^n |f_i(X^*)| \\ &= |f_1(X^*)| + |f_2(X^*)| + |f_3(X^*)| \\ &\quad + \dots + |f_n(X^*)| > 0 \end{aligned}$$

lo que obviamente contradice la hipótesis que dice que  $g(X^*) = 0$ , lo cual no puede ser. Por lo tanto, para toda  $1 \leq i \leq n$  se da que  $f_i(X^*) = 0$  y  $F(X^*) = \vec{0}$ . Es decir,  $X^*$  es una solución de  $F(X) = \vec{0}$ .

De aquí en más, para resolver un sistema no lineal en particular, lo primero que haremos es construir la función  $g$  utilizando las funciones  $f_1, f_2, f_3, \dots, f_n$  que componen el sistema. Luego, el conjunto  $\Omega$ , para la ejecución del algoritmo, lo vamos a suponer de la forma:

$$\Omega = \left\{ (x_1, x_2, x_3, \dots, x_n) \in R^n / \alpha_i \leq x_i \leq \beta_i \quad \begin{array}{l} \\ \text{para cada } i = 1, 2, 3, \dots, n \end{array} \right\}$$

Es decir, para resolver el sistema de ecuaciones no lineales, se debe minimizar la función no lineal  $g$  y se deben fijar cotas inferiores y superiores para cada una de las variables del sistema. Se espera que exista una solución para el sistema no lineal en el conjunto  $\Omega$ , para que la solución encontrada en el problema de optimización la podamos tomar también como solución del

sistema no lineal. Debido a que el sistema no lineal es el modelo matemático de alguna aplicación o problema del mundo real, se puede construir un conjunto  $\Omega$  donde se sospecha o se considera se encuentra la solución del sistema no lineal. La experiencia del modelador, o de las personas involucradas en el problema que es representado por el sistema no lineal, es vital y sumamente importante para definir dicho conjunto  $\Omega$ .

#### 4. EL ALGORITMO METAHEURÍSTICO

El problema de optimización que se pretende resolver es un problema de minimización en el cual la función objetivo  $g(x_1, x_2, x_3, \dots, x_n)$  es una función no lineal y, cada una de las variables independientes de  $g$ , está acotada tanto inferior como superiormente. Es decir, para cada  $i = 1, 2, 3, \dots, n$  ocurre que  $x_i \in [\alpha_i, \beta_i]$ . En otras palabras:

$$\Omega = \prod_{i=1}^n [\alpha_i, \beta_i] = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] \times [\alpha_3, \beta_3] \times \dots \times [\alpha_n, \beta_n] \subseteq \mathbb{R}^n$$

Para ejecutar el algoritmo necesitamos ingresar entonces la función  $g$  y, los  $n$  intervalos para cada una de las  $n$  variables donde se cree que se encuentra la solución del sistema no lineal.

Algo importante de mencionar es que al usar cualquier algoritmo metaheurístico estamos conscientes que no existe garantía de encontrar el óptimo del problema, pero podemos encontrar una “buena solución” en un tiempo prudencial y con un costo computacional aceptable, en comparación con cualquier método exacto que se pueda usar para el problema de optimización. Ahora, ¿por qué el algoritmo “scatter search”? Porque ha sido muy ampliamente usado y con mucho éxito para la optimización de funciones lineales y no lineales de varias variables sobre un espacio de soluciones que puede ser continuo y también discreto. A continuación se explica cada etapa del algoritmo.

##### *Generación del Conjunto $P$*

Lo primero que debe hacerse al resolver un problema de optimización con “scatter search” es construir un conjunto al que llamaremos  $P$ . Una entrada del algoritmo entonces es el tamaño del conjunto  $P$ , definida con la variable  $PSize$ . Los elementos del conjunto  $P$  serán vectores de  $\mathbb{R}^n$  que pertenecen al conjunto  $\Omega$ . Es decir, cada elemento del conjunto  $P$  es una solución factible del problema de minimización. Estos elementos se van a generar aleatoriamente pero de forma tal que podamos cubrir o abarcar todo el conjunto

$\Omega$ , para así tener diversidad en esta primera generación de soluciones factibles. Sin embargo, el tema de la calidad es también importante en esta primera etapa del algoritmo. A continuación se explica cómo se van a generar los vectores.

Cada variable independiente  $x_i$  de la función  $g$  “vive” en un intervalo  $[\alpha_i, \beta_i]$ . Se va a dividir o a particionar cada intervalo  $[\alpha_i, \beta_i]$  en  $m$  subintervalos. El número de subintervalos de igual tamaño en el que se va a dividir el intervalo  $[\alpha_i, \beta_i]$  se va a representar entonces con la variable  $m$ , el cual es también un parámetro de entrada para el algoritmo. De esta manera, la longitud de cada subintervalo (todos los subintervalos para la variable  $x_i$  tienen la misma longitud) es:  $\Delta x_i = (\beta_i - \alpha_i)/m$

Ahora se van a generar  $PSize$  elementos con una mezcla de diversidad y calidad. Primero hacemos uso de la diversidad, ¿cómo trabaja? Así: Para cada variable  $x_i$  se selecciona aleatoriamente un subintervalo (se genera un número entero aleatorio entre 1 y  $m$ ), luego se genera un número aleatorio de punto flotante dentro del subintervalo seleccionado anteriormente, lo que va a representar el valor de la  $i$ -ésima componente del vector. Esto se debe hacer entonces por cada una de las  $n$  componentes del vector de  $\mathbb{R}^n$ . Finalmente, se aplica un método de mejora al vector que se acaba de construir, en otras palabras, a través de una “local search” se busca si existe “cerca” de la solución generada otro vector que mejora el comportamiento de la función objetivo (es decir que minimiza la función  $g$ ). Posteriormente se explicará la forma de trabajar de la búsqueda local. Pero, para asegurar la diversificación de los elementos en el conjunto  $P$ , la probabilidad de elegir un subintervalo, para una variable cualquiera  $x_i$ , es inversamente proporcional a la cantidad de números de punto flotante generados en dicho subintervalo de esa variable  $x_i$ . De esta manera se trata de “balancear” las soluciones a lo largo de todo el conjunto  $\Omega$ . Como podemos apreciar, se tiene ahora un conjunto  $P$  con una mezcla muy buena de calidad y diversidad.

##### *Búsqueda Local*

La “local search” o búsqueda local implementada recibe como entrada la solución a mejorar y un parámetro adicional que le dice al método el número de intentos que debe hacer para buscar un “mejor” vector en una vecindad del vector inicial, este número de intentos está en función del número de veces que se va a evaluar la función

$g$ . ¿Cómo trabaja? A la  $i$ -ésima componente del vector inicial se le genera un número aleatorio que se le suma o se le resta a esa  $i$ -ésima componente (lo cual también es aleatorio) de manera tal que se obtiene un nuevo valor para esa  $i$ -ésima componente y, de esta forma, al efectuar esto  $n$  veces, se obtiene un nuevo vector. Aquí se debe tener todo el cuidado de que el valor de la  $i$ -ésima componente del nuevo vector respete los límites definidos del subintervalo al cual pertenece el valor de la  $i$ -ésima componente del vector original (esto es lo que hace que la búsqueda sea local, el nuevo valor de cada variable no se debe salir del subintervalo en el que se encuentra el valor original de esa variable). Cada vez que se construye un vector en la vecindad del vector original se ha realizado “un intento”. Como resultado del intento se tiene un nuevo vector que es descartado si no mejora la función  $g$  o tomado como el nuevo vector original si la función  $g$  mejora con este nuevo vector. Después de una cantidad de intentos (lo que es un argumento de este procedimiento) se detiene la búsqueda. Si se ha llegado al número de intentos tope de búsqueda definido a través del argumento “número de evaluaciones” y no se ha encontrado una mejor solución, el método devuelve la primera solución que se pasó como argumento, es decir, se indica así que “nos quedamos igualitos”, no se encontró nada mejor. Cabe indicar que otro método que se implementó para la búsqueda local es el conocido algoritmo heurístico “Nelder-Mead” propuesto por Jhon Nelder y Roger Mead en 1965 para la optimización continua de una función objetivo sin restricciones. Este algoritmo sólo utiliza la regla de correspondencia de la función a optimizar y no utiliza las derivadas parciales de ningún orden para dicha función, motivo por el cual fue escogido también. Además se trata de un algoritmo que ha tenido mucho éxito en diversos problemas presentando buenos resultados, pero hay que tener presente que se trata de un algoritmo local y no global. De esta manera, el algoritmo scatter search que se ha construido tiene la capacidad de ejecutar con cualquiera de los dos algoritmos de búsqueda local antes indicados.

#### **Construcción del Conjunto $R$**

En el método de “scatter search” existe un segundo conjunto importante, el conjunto  $R$ , el cual es construido a partir del conjunto  $P$ . La cardinalidad de este conjunto  $R$  es también un parámetro del algoritmo. Para construir el conjunto  $R$  nos preocupamos que la variable  $RSize$ , que define el tamaño de  $R$ , sea par, ya que la mitad de los elementos de  $R$  serán escogidos por calidad mientras que la otra mitad por diversificación, como ya se dijo, haciendo uso

del conjunto  $P$ . Aquí es importante indicar que se pudo haber parametrizado esto, es decir, que  $b_1$  elementos sean escogidos por calidad y  $b_2$  por diversidad teniendo presente que  $b_1 + b_2 = RSize$ . Retomando, se toman entonces los  $RSize/2$  “primeros en calidad” elementos del conjunto  $P$  (es decir, los mejores elementos de dicho conjunto) y se incorporan al conjunto  $R$ . ¿Cómo escoger los elementos restantes?

De los elementos que quedan de  $P$  se deben tomar  $RSize/2$  elementos más para completar el conjunto  $R$ , ¿cómo se lo hace para garantizar diversidad? Se calcula la distancia de cada elemento de  $P$  (ya no se consideran aquellos que se llevan a  $R$ ) al conjunto  $R$  con la fórmula: Sea  $v \in P$ , luego  $d(v, R) = \min_{w \in R} \|v - w\|$  y

el elemento que se almacena en  $R$  es aquel elemento de  $P$  que dista más de  $R$ , es decir, el más lejano al conjunto  $R$ . La norma utilizada para calcular la distancia mencionada es la norma euclíadiana.

En este momento se tiene ya construido totalmente el conjunto  $R$  con una mezcla muy buena y adecuada de calidad y diversidad.

#### **Actualización y Mantenimiento del Conjunto $R$**

Durante la búsqueda del óptimo el conjunto  $R$  debe actualizarse muchas veces y eso es lo que se va a explicar. Cabe indicar aquí que el número de actualizaciones del conjunto  $R$  es también un parámetro del algoritmo. Una vez construido el conjunto  $R$  se procede a crear subconjuntos de  $R$ . En esta implementación el tamaño de cada subconjunto es de dos elementos de  $R$  (pero en otras implementaciones con scatter search la cardinalidad de estos subconjuntos puede ser de más elementos), es decir, usando  $R$  se generan todas las parejas posibles de elementos sin repetir, porque se les va a aplicar un método de combinación. ¿Cómo trabaja el método de combinación? Eso se explica unas líneas más adelante. Al aplicar un método de combinación a cada subconjunto, se generan cuatro elementos y, a cada uno de estos nuevos elementos (llamados vectores prueba) se les aplica un método de mejora, es decir, se emplea la operación de búsqueda local explicada anteriormente para estudiar la vecindad de estas “soluciones prueba” en busca de unas mejores soluciones. Entonces, por cada subconjunto se tienen cuatro “soluciones mejoradas” que se añaden a un conjunto denominado *Pool*. Es decir que finalmente en *Pool* se tiene el resultado de aplicar métodos de “combinación” y “mejora” a cada uno de los subconjuntos construidos a partir del gran

conjunto  $R$ . Una vez que el conjunto  $Pool$  está lleno se construye un conjunto de “vida corta”  $Temp$  de forma tal que ahí se almacenen todos los elementos de  $R$  y de  $Pool$ , es decir  $Temp = R \cup Pool$ . Este conjunto temporal, tiene una cardinalidad igual a la cardinalidad de  $R$  más la cardinalidad de  $Pool$ . Los  $RSize$  mejores elementos son los que pasan finalmente al conjunto  $R$ . ¿Cuántas veces ocurre esto? Hay un parámetro llamado “Número de Iteraciones” que se debe ingresar antes de correr el algoritmo de scatter search, pues este parámetro es el que sirve para configurar el número de veces que el conjunto  $R$  se va a actualizar. La opción de reconstruir el conjunto  $R$ , aclaramos, no se ha incorporado en la implementación. Es decir, si el conjunto  $R$  permanece sin cambios después de una actualización (debido a que el “combinar y mejorar” no produce nada nuevo) se recomienda una reconstrucción donde participa nuevamente el conjunto  $P$  y el conjunto  $R$  se actualiza nuevamente por calidad y diversidad. No se implementó esta etapa debido a que no se lo consideró necesario ya que en las pruebas que se realizaron el algoritmo siempre encontraba soluciones muy cercanas al óptimo, es decir, el software encontraba correctamente valores muy próximos al óptimo del problema de minimización en todos los ejemplos de sistemas no lineales en los que se lo puso a prueba. Cabe indicar que se probó el algoritmo en sistemas no lineales de hasta cinco ecuaciones y cinco incógnitas, y con conjuntos  $\Omega$  “pequeños” donde se sospechaba se encontraba la o las soluciones. Pero sería importante considerar la reconstrucción del conjunto  $R$  si se tienen sistemas no lineales de mayor tamaño.

#### ***Método de Combinación***

El método de combinación implementado es muy sencillo y opera así: se generan cuatro números aleatorios de punto flotante en el intervalo  $(0,1)$  con la idea de aplicar combinaciones lineales “convexas”. Es decir, el método de combinación lo único que recibe como parámetro es el subconjunto de  $R$  que tiene dos soluciones  $SolA$  y  $SolB$ , y devuelve como salida cuatro combinaciones lineales convexas distintas. Es importante decir que esta salida no pasa directamente a  $R$  ya que antes se aplica el método de mejora a cada una de las cuatro soluciones obtenidas por combinación, y ellas sí pasan ya al conjunto  $R$ .

#### ***Método Principal del Algoritmo***

A continuación se va a mostrar el método principal llamado “EjecutarSS” que gobierna el comportamiento y la operación general del algoritmo de scatter search que se implementó.

Todas las etapas o fases anteriores coordinan y se enlazan aquí. A continuación el método:

```

public double[] EjecutarSS(out double
MejorValor)
{
    ArrayList Pool = new ArrayList();
    ArrayList EvalPool = new ArrayList();
    Stopwatch Tiempo;
    double[,] MejoresValores =
    new double[this.NumIteraciones, 2];
    Tiempo = Stopwatch.StartNew();
    this.GenerarP();
    this.ConstruirR(this.P, this.EvalP, this.R,
    this.EvalR);

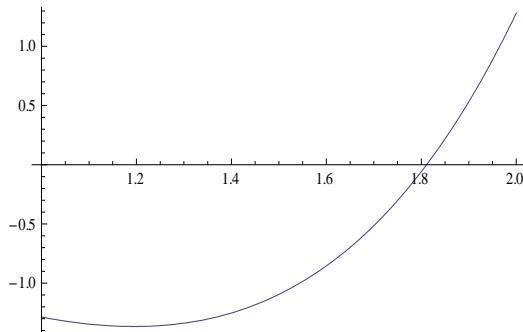
    ArrayList SubSets = new ArrayList();
    int Iteraciones = 0;
    while (true)
    {
        this.GenerarSubSets(SubSets, R);
        while (SubSets.Count != 0)
        {
            double[][] SubSet =
            (double[][])SubSets[0];
            double[][] TrialSolutions =
            this.Combinar(SubSet);
            double Valor1 =
            this.LS(TrialSolutions[0], 100);
            Pool.Add(TrialSolutions[0]);
            EvalPool.Add(ValorUno);
            double Valor2 =
            this.LS(TrialSolutions[1], 100);
            Pool.Add(TrialSolutions[1]);
            EvalPool.Add(ValorDos);
            double Valor3 =
            this.LS(TrialSolutions[2], 100);
            Pool.Add(TrialSolutions[2]);
            EvalPool.Add(ValorTres);
            double Valor4 =
            this.LS(TrialSolutions[3], 100);
            Pool.Add(TrialSolutions[3]);
            EvalPool.Add(ValorCuatro);
            SubSets.RemoveAt(0);
        }
        this.ActualizarR(R, EvalR, Pool,
        EvalPool);
        MejoresValores[Iteraciones, 0] =
        EvalR[0];
        MejoresValores[Iteraciones, 1] =
        Tiempo.ElapsedMilliseconds;
        Iteraciones++;
        if (Iteraciones == this.NumIteraciones)
        break;
    }
    MejorValor = EvalR[0];
    return R[0];
}

```

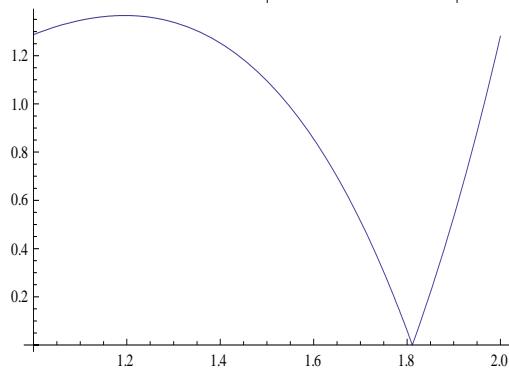
## 5. SOLUCIÓN DE ALGUNOS SISTEMAS NO LINEALES

A continuación se presentan algunas corridas del algoritmo scatter search. Para empezar, y para poder dar una interpretación geométrica de los resultados, vamos a mostrar lo que ocurre con la ecuación  $f(x) = x^3 - e^x \operatorname{sen}(x)$ . Haciendo uso del teorema de Bolzano se puede probar que  $f$  tiene un CERO en el intervalo  $\Omega = [1,2]$ . Por la proposición enunciada y demostrada anteriormente se puede afirmar que  $x^*$  satisface  $f(x) = x^3 - e^x \operatorname{sen}(x) = 0$ , si y sólo si,  $x^*$  minimiza la función  $g(x) = |x^3 - e^x \operatorname{sen}(x)|$  en el intervalo  $[1,2]$  con el valor cero. La figura # 1 muestra la gráfica de la función  $f$  en el intervalo  $[1,2]$ , mientras que la figura # 2 nos muestra la gráfica de la función a minimizar  $g$  en el mismo intervalo  $[1,2]$ .

**FIGURA 1**  
*Sistemas no lineales y la metaheurística scatter search*  
**Gráfico de**  $f(x) = x^3 - e^x \operatorname{sen}(x)$



**FIGURA 2**  
*Sistemas no lineales y la metaheurística scatter search*  
**Gráfico de**  $g(x) = |x^3 - e^x \operatorname{sen}(x)|$



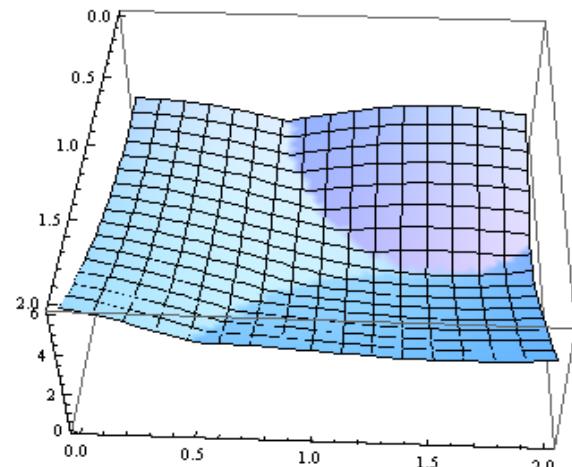
Se puede observar que mayor a 1.8 es el valor de  $x^*$ . En una de las corridas del algoritmo metaheurístico de “scatter search” encontramos que  $x^* = 1.8112693646$  y que  $g(x^*) = 0.0000655681$ . Para esta ejecución se dividió el intervalo  $\Omega = [1,2]$  sólo en 10 subintervalos y se solicitaron 5 actualizaciones del conjunto  $R$ .

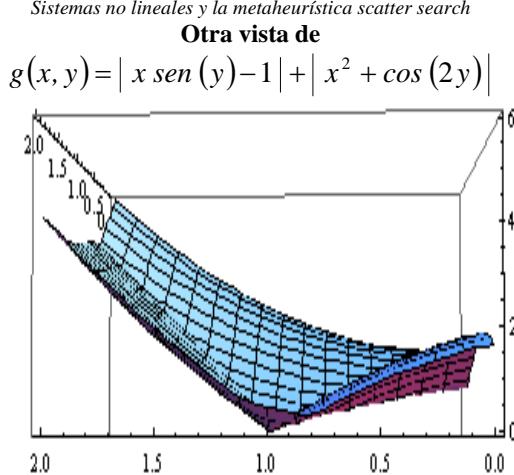
Un segundo ejemplo es el sistema:

$$\begin{cases} x \operatorname{sen}(y) - 1 = 0 \\ x^2 + \cos(2y) = 0 \end{cases}$$

Se puede apreciar que el par ordenado  $(1, \pi/2)$  es una solución del sistema en la región  $\Omega = \{(x, y) \in R^2 / 0 \leq x \leq 2, 0 \leq y \leq 2\}$ . La gráfica de la función  $g(x, y) = |x \operatorname{sen}(y) - 1| + |x^2 + \cos(2y)|$  sobre la región  $\Omega$  del plano  $xy$  se muestra en la figura 3 y en la figura 4.

**FIGURA 3**  
*Sistemas no lineales y la metaheurística scatter search*  
**Gráfico de**  $g(x, y) = |x \operatorname{sen}(y) - 1| + |x^2 + \cos(2y)|$



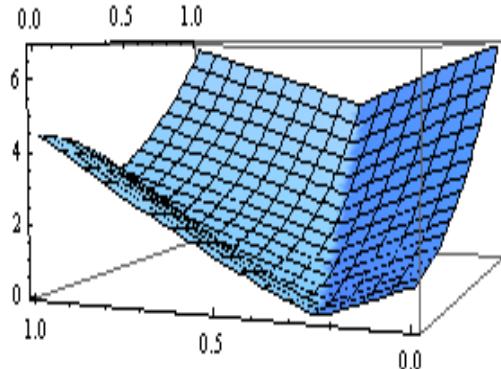
**FIGURA 4**  
Sistemas no lineales y la metaheurística scatter search

En una de las ejecuciones del algoritmo “scatter search”, para este segundo ejemplo, se obtuvo que  $(x^*, y^*) = (1.0, 1.5707963290621922)$  y que  $g(x^*, y^*)$  es cero. Para esta corrida se solicitaron 5 actualizaciones del conjunto  $R$  y, para cada variable, se ingresó el intervalo  $[-10, 10]$ , el cual se partió en 10 subintervalos.

Otro ejemplo con interpretación gráfica es el siguiente:

$$\begin{cases} 5x^2 - y^2 = 0 \\ 4y - \operatorname{sen}(x) - \cos(y) = 0 \end{cases}$$

Este sistema no lineal tiene solución en  $\Omega = [0, 1] \times [0, 1]$ . La gráfica de la función  $g(x, y) = |5x^2 - y^2| + |4y - \operatorname{sen}(x) - \cos(y)|$  se muestra en la figura 5.

**FIGURA 5**  
Sistemas no lineales y la metaheurística scatter search  
Gráfico de

Con los mismos parámetros de entrada de la ejecución del segundo ejemplo se obtuvo que

$$(x^*, y^*) = (0.12124191148054805, 0.27110515579243255)$$

y que  $g(x^*, y^*)$  es del orden de  $10^{-14}$ . Es importante indicar que se pueden solicitar más actualizaciones del conjunto  $R$ , aumentar el valor del parámetro  $m$ , aumentar la cardinalidad de los conjuntos  $P$  y  $R$ , todo esto con el objetivo de mejorar la precisión. Sin embargo, los costos computacionales serán mayores.

En otro ejemplo, considere el sistema no lineal que se muestra a continuación:

$$\begin{cases} (x-2)^2 + (y-1)^2 + xy - 3 = 0 \\ x e^{x+y} + y z - 3 = 0 \\ \operatorname{sen}(x+z) + x - y - z + 1 = 0 \end{cases}$$

Se ejecutó el algoritmo “scatter search” y se obtuvo como resultado la solución

$$(x^*, y^*, z^*)$$

$$= (3.03924738054514, -1.6104828215273, 6.01336437928736)$$

y además que  $g(x^*, y^*, z^*) = 0.0000000345641$  en la región sólida

$\Omega = [-10, 10] \times [-10, 10] \times [-10, 10]$ . Se solicitaron nuevamente 5 actualizaciones del conjunto  $R$  y además  $m = 10$ .

Finalmente, como un último ejemplo, se tiene el sistema no lineal:

$$\begin{cases} x^2 + 2y^2 + \cos(z) - w^2 = 0 \\ 3x^2 + y^2 + \operatorname{sen}^2(z) - w^2 = 0 \\ -2x^2 - y^2 - \cos(z) + w^2 = 0 \\ -x^2 - y^2 - 2\cos^2(z) + w^2 = 0 \end{cases}$$

Este sistema no lineal presenta muchas soluciones. Dos de ellas son:  $(1, -1, 0, 2)$  y  $(-1, 1, 0, -2)$ . Se ingresaron como parámetros  $m = 10$ , un total de 5 actualizaciones para el conjunto  $R$  y el espacio de búsqueda  $\Omega = [-10, 10] \times [-10, 10] \times [-10, 10] \times [-10, 10]$ . El valor de la función  $g$  en el punto encontrado

$$(x^*, y^*, z^*, w^*)$$

$$= (0.99896239, -0.99896202, 0.03718569, 1.99827054)$$

es igual a  $0.000002593939$ , es decir, del orden de  $10^{-6}$ .

Se realizaron pruebas con sistemas de hasta cinco ecuaciones con cinco incógnitas y con otros sistemas no lineales de menor tamaño, que no se muestran en este documento, con excelentes resultados.



## 6. TABLA COMPARATIVA

A continuación se muestra una tabla que compara los algoritmos metaheurísticos con los algoritmos iterativos convencionales para la solución de sistemas no lineales. Se presentan ventajas y desventajas de cada uno.

Algoritmos Metaheurísticos		Algoritmos Iterativos Convencionales	
Ventajas	Desventajas	Ventajas	Desventajas
Dominio de convergencia ampliado debido a las estrategias de “diversificación”	La implementación del algoritmo es más complicada, existen muchas fases o etapas involucradas	La implementación del algoritmo es más sencilla, usualmente mediante fórmulas recursivas	Normalmente un “pequeño” dominio de convergencia
Se los puede utilizar para determinar el “dato inicial” que necesitan los algoritmos iterativos convencionales o para futuras corridas del mismo algoritmo en un espacio de búsqueda más reducido	No existe garantía de encontrar el óptimo del problema, son algoritmos que utilizan en su trabajo la aleatoriedad. Además, pueden quedarse atrapados en óptimos locales	Si el “dato inicial” se encuentra dentro del intervalo de convergencia se garantiza que el método converge a la solución exacta	Pueden tener dependencia de otros métodos que proporcionen el “dato inicial” para poder comenzar con las iteraciones
Las funciones de varias variables que se utilizan no necesitan cumplir con condiciones de continuidad y diferenciabilidad de ningún orden	Espacio de búsqueda considerable, en general convergencia lenta. La velocidad de convergencia depende del tamaño del espacio de búsqueda	En general se tiene convergencia al menos cuadrática, es decir, son métodos rápidamente convergentes	Dependiendo del método que se utilice, las funciones de varias variables empleadas deben ser de clase $C^1$ o de clase $C^2$ o deben satisfacer o cumplir la condición de Lipschitz
No se necesitan cálculos matriciales tediosos ni calcular derivadas parciales de ningún orden	No se puede estimar el número de iteraciones necesarias para estar tan cerca de la solución exacta como se deseé.	Se puede tener control sobre el error. Se puede estimar el número de iteraciones para que la distancia entre la solución exacta y el valor estimado no supere la tolerancia deseada	Por cada iteración muchos cálculos matriciales (en algunos casos el cálculo de matrices inversas y producto de matrices) y cálculos de derivadas parciales
En una misma ejecución se pueden encontrar muchas soluciones, si es que existen	La técnica principal en que basan su trabajo es una computación evolutiva. Existe muy poca teoría matemática detrás del método que estudie o analice convergencia y error	Muchos teoremas y proposiciones que estudian convergencia y error de los métodos	Si el sistema no lineal tiene muchas soluciones, sólo una se puede encontrar por cada corrida

## 7. CONCLUSIONES Y RECOMENDACIONES

Los algoritmos metaheurísticos usados para resolver sistemas de ecuaciones no lineales tienen la ventaja de que las funciones de varias variables que componen el sistema no necesitan ser de clase  $C^1$ , por el uso del Jacobiano, ni de clase  $C^2$ , por

el empleo de la matriz Hessiana, ni ser diferenciables, ni satisfacer condiciones de Lipschitz como en los métodos de punto fijo. Los métodos numéricos iterativos convencionales exigen estas condiciones en las funciones del sistema no lineal  $f_1, f_2, f_3, \dots, f_n$ . En los algoritmos metaheurísticos podrían las funciones

anteriores incluso no ser funciones continuas sobre su dominio  $\Omega \subseteq R^n$  y sin embargo algoritmos como scatter search podrían ejecutar sin inconvenientes y resolver los sistemas no lineales. Sólo necesitamos que las funciones se encuentren definidas sobre el conjunto  $\Omega$ . Además, estos algoritmos no necesitan efectuar cálculos matriciales que puedan complicar la implementación, consumir mucha memoria y tiempo excesivo de procesador. A pesar de que en general los algoritmos metaheurísticos no garantizan entregar la respuesta exacta, para sistemas no lineales de pequeño o mediano tamaño han probado tener mucho éxito. La recomendación principal para futuros trabajos es la de implementar al menos un par de métodos iterativos clásicos y de comparar los resultados de ejecución, en eficacia y eficiencia, con el algoritmo metaheurístico que se presenta en este artículo. Finalmente se podría intentar construir un “algoritmo híbrido”, es decir, un algoritmo donde algunas partes provengan por ejemplo del método de Newton o de una de sus variantes y, por otro lado, también tenga componentes que utilicen “inteligentemente” la aleatoriedad como lo hacen los algoritmos metaheurísticos.

Otro trabajo a futuro que podría efectuarse sería el de preparar el algoritmo de forma tal que no sólo se puedan encontrar soluciones reales sino también soluciones complejas para los sistemas no lineales que las posean. Se puede apreciar también que “scatter search”, el algoritmo metaheurístico utilizado, no hace un uso intensivo e indiscriminado de la aleatoriedad, es decir, no abusa de ella como sí lo hacen, por ejemplo, en general los algoritmos genéticos. La mezcla de diversidad y calidad del algoritmo “scatter search” muestra lo inteligente de su forma de trabajar. Podríamos también considerar el empleo de otros algoritmos metaheurísticos para resolver sistemas no lineales. Finalmente, cabe mencionar que existe otro método iterativo no tradicional que permite resolver sistemas no lineales de ecuaciones y que usa “homotopía y continuación numérica”. En este método se expresa el sistema no lineal en términos de un sistema de ecuaciones diferenciales ordinarias de primer orden donde se podría, por ejemplo, usar algún método de Runge-Kutta. Se trata de un problema equivalente, de forma tal que la solución del sistema de ecuaciones diferenciales es también la solución del sistema no lineal de ecuaciones.

**REFERENCIAS BIBLIOGRÁFICAS Y ELECTRÓNICAS**

- [1]. **GROSAN Y ABRAHAM** (2008). “*A New Approach for Solving Nonlinear Equations Systems*”, IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, Vol. 38, No. 3 *Implementations in C*”, Kluwer Academic Publishers, Norwell Massachusetts
- [2]. **LAGUNA Y MARTÍ** (2003). “*Scatter Search: Methodology and*
- [3]. **MARTÍ Y LAGUNA** (2003). “*Scatter Search: Diseño Básico y Estrategias Avanzadas*”, Universidad de Valencia, España

## INCORPORANDO BROUSSEAU EN DOS TEOREMAS FUNDAMENTALES: PROBABILIDAD TOTAL Y BAYES

**Petakos Kyriakos<sup>1</sup>**

**Resumen.** El presente artículo intenta presentar la popular teoría de las situaciones didácticas TDS (theory of didactical situations), establecida por Brousseau (Brousseau, 1997) y refinada por Chavarría (Chavarría, 2006), usando la llamada voz del aula en dos momentos críticos para el aprendizaje de la probabilidad del nivel universitario. Vemos la formación de un milieu, sobre el entendimiento de dos teoremas de probabilidad, la Probabilidad Total y Bayes.

**Palabras Clave:** Probabilidad total, Bayes, Teoría de las Situaciones Didácticas TDS.

**Abstract.** The present article presents the popular theory of the didactical situations (TDS), as established by Brousseau (Brousseau, 1997) and refined by Chavarría (Chavarría, 2006), employing the so-called voices of the classroom in two critical points for the learning of probability on the university level. We will see the formation of teaching environment that concerns understanding two probability theorems, Total probability law and Bay's theorem

**Key Words:** Total probability, Bayes, Theory of Didactical Situations TDS.

Recibido: Octubre 2012

Aceptado: Enero 2013

### 1. INTRODUCCIÓN

En los primeros años universitarios la Teoría de la Probabilidad, habitualmente como seguidora de un curso de análisis combinatorio (Sanabria, 2006), se considera fundamental en el desarrollo del pensamiento matemático en muchas disciplinas diferentes. Forma un paso imprescindible para organizar un programa de estudios, que se puede justificar como terciario y para preparar a los estudiantes para que piensen de una manera más eficaz en sus carreras de futuro. Un curso como el descrito arriba contiene necesariamente las reglas y las leyes básicas de la probabilidad con el objetivo de familiarizar a los estudiantes con una forma de pensar tanto práctica como teórica, la que les pueda parecer útil en su futuro próximo y lejano. Aquí vamos a destacar dos, el teorema de la Probabilidad Total y la ley de Bayes, (Feller, 1968).

*Teorema de Probabilidad Total.* Si  $\Omega$  simboliza el espacio total de un experimento y  $A_i$  es una secuencia de eventos mutuamente excluyentes, cuya unión forma el conjunto  $\Omega$ , entonces por cada posible evento B

$$P(B) = \sum_i P(B/A_i)P(A_i)$$

*Teorema de Bayes* Bajo las mismas condiciones de arriba, se concluye

$$P(A_i/B) = \frac{P(B/A_i)P(A_i)}{P(B)} = \frac{P(B/A_i)P(A_i)}{\sum_i P(B/A_i)P(A_i)}$$

### 2. METODOLOGÍA

Nuestra muestra consiste de 16 estudiantes de la Universidad de Pireas, divididos por 4 grupos. La división de los grupos se basa principalmente a las relaciones ya desarrolladas entre ellos en un sentido assertivo (Cadoche, 2009, Pérez, 2006). Los encuentros realizados con ellos son 2 veces a la semana que duran 2-3 horas cada vez. Se presenta por parte del profesor el material didáctico con sus pruebas apropiadas, sus consecuencias y un número predefinido de ejemplos adecuadamente elegidos. Al final se pone para resolución un problema, lo diríamos especie de instrumento para poder formar una imagen de lo entendido, cuyo desarrollo tendrá lugar en el segundo encuentro.

#### Ejercicio

En una carrera de caballos la probabilidad de victoria de un caballo es 0,3 si el tiempo está bueno y 0,5 si el tiempo está malo. El evento de llover pasa con probabilidad 0,4.

- 1) ¿Cuál es la probabilidad de victoria?
- 2) En el caso de que el caballo perdiera la carrera, ¿cuál es la probabilidad que el tiempo estaba bueno el día de la carrera?

#### Solución

(según el modo de Sanabria, (Sanabria, 2010))

#### Etapa 1

Definición de los eventos

$V$  victoria del caballo,  $D$  derrota del caballo,  $M$  el tiempo está malo,  $B$  el tiempo está bueno. Entonces

<sup>1</sup>Petakos Kyriakos, Dr., Academia Turística Rodos Grecia.  
(e\_mail: kyriakospetakos66@gmail.com).

$$P(V/B) = 0,3, \quad P(V/M) = 0,5, \\ P(M) = 0,4$$

**Etapa 2**

Aplicación de Probabilidad total

$$P(V) = P(V \cap B) + P(V \cap M)$$

y más detalladamente

$$P(V) = P(V/B)P(B) + P(V/M)P(M).$$

Sustituyendo los números

$$P(V) = 0,3 * 0,6 + 0,5 * 0,4 = 0,38$$

**Etapa 3**

Aplicación de Bayes

$$P(B/D) = \frac{P(B \cap D)}{P(D)} = \frac{P(D/B)P(B)}{P(D)}$$

Pero

$$P(D) = 1 - P(V) = 1 - 0,38 = 0,62$$

Por otro lado

$$P(D/B) = 1 - P(V/B) = 1 - 0,3 = 0,7$$

Combinando las tres últimas relaciones,

$$P(B/D) = \frac{0,7 * 0,6}{0,62} = 0,68.$$

**3. PRESENTACIÓN**

Los cuatro grupos han escrito correctamente y detalladamente la primera etapa del ejercicio, es decir, la definición de los eventos en la manera apropiada, pero los dos tienen dudas sobre las probabilidades condicionales involucradas en el problema

$$P(V/B) = 0,3 \\ P(V/M) = 0,5$$

Podemos ver unas ideas que tuvieron los estudiantes:

$$P(V/M) = 1 - P(V/B)$$

$$P(V/B) = 1 = P(V/M)$$

Cuando son entrevistados, obtenemos el diálogo siguiente. P simboliza desde ahora en adelante el profesor y A, B, C, D los grupos designados con

índices diferentes para cada estudiante -miembro del grupo-separadamente.

*A<sub>1</sub>* Así que el tiempo malo es el complemento del tiempo bueno, ¿no se concluye consecuentemente a la relación de abajo?

$$P(V/M) = 1 - P(V/B)$$

*B<sub>1</sub>* Naturalmente en el contenido del ejercicio vemos dos números que no suman 1. Pero quizás esconde alguna interpretación difícil de los teoremas.

P Si tenemos en cuenta que la ley probabilística del complemento se aplica cuando la negación está en el numerador, i.e.

$$P(A'/B) = 1 - P(A/B)$$

entonces se puede evadir un error así.

*A<sub>1</sub>* al menos de ahora en adelante nos queda en la memoria. Los números verdaderamente no sostienen algo así.

El último comentario pone mucha importancia en el concepto del ejemplo, de paradigma y del papel que desempeña entre las dos partes involucradas, el profesor y el estudiante. Diríamos que ataña a la tercera parte de Brousseau (Chavarría, 2006), al medio didáctico. Elegir un ejemplo como nuestro ejercicio inicial sirve a un doble objetivo: ayudar a aclarar la relación de arriba de forma práctica, sólo a través de una comparación de números, por parte del estudiante y facilitar la tarea del profesor para diagnosticar errores eventuales.

*Segunda etapa*

Aplicación del Teorema de Probabilidad total. Una vez subsanada la situación de la probabilidad condicional, escribir la fórmula del teorema probabilidad total viene casi naturalmente

$$P(V) = P(V/B)P(B) + P(V/M)P(M)$$

El adverbio *casi* refiere a la realidad de que los libros están abiertos y así mientras un solo grupo escribe la relación con velocidad, los otros la reproducen paso a paso consultando el texto.

*P* ¿Era verdaderamente fácil entender que el Teorema de la Probabilidad total estaba escondido?

*A<sub>1</sub>* Seguro que no. Nos ha ocupado muchas veces desde el principio de semestre. Y al menos, en ese concepto los ejemplos ya resueltos estaban escritos muy detalladamente. No quedó espacio de duda.

*B<sub>1</sub>* (y *C<sub>1</sub>* a la vez) Una vez definidas las probabilidades condicionales, los componentes de este teorema y su aplicación parece como la resolución de una fórmula sencilla, a la que estamos acostumbrados desde hace muchos años.

### Tercera etapa

La aplicación de Bayes, un teorema cuya comprensión a lo mejor formó para nosotros el obstáculo más grande durante el desarrollo de la resolución. Verdaderamente lo que pasa, lo que se observa es una incapacidad por parte de los grupos para representar lo que se pregunta. La frase ¿cuál es la probabilidad de un tiempo bueno, dada la derrota del caballo? no se puede interpretar probabilísticamente. En un esfuerzo de evadir el efecto Topaze (Chavarría, 2006) como lo caracteriza la teoría de Brousseau, tenemos que profundizar en las razones, para las que nuestra muestra encuentra tanta dificultad de adelantarse. A la vez combinamos de grado permitido lo que sostiene la zona de desarrollo próximo (Radford, 2008). Un proceso de equilibrar dos especies de capacidades, la capacidad personal independiente y la que se cultiva junto a una persona con mayor experiencia.

*C*<sub>1</sub> Lo que personalmente me causa no sólo dificultad, sino también una falta del contacto, es este tipo de inversión de la realidad. Digamos mejor la inversión de la lógica. La lógica implica preguntar ¿cuál es la probabilidad de la derrota, así que el tiempo está bueno?

*D*<sub>1</sub> Para mí, usando la primera etapa, creo que se escribe algo probabilístico. Pero me da miedo expresarlo. Me da la impresión de que todo viene al revés.

*P* (viendo que el resto de los grupos comparte estas preocupaciones) Entonces ¿por qué no emplear esta violación de la lógica, como ustedes la caracterizan para asociarla con la aplicación de Bayes?

*A*<sub>2</sub> Si ocurre nuevamente otra situación para nosotros paradoja, que no está resuelta por Bayes, ¿entonces qué haremos?

El comentario de arriba nos ofrece la oportunidad de redefinir nuestro papel en la enseñanza de hoy. Somos partidarios de la doctrina que no estamos en el aula simplemente para transmitir el conocimiento, sino para poder contribuir a su construcción (Vrasidas, 2000). Así que no otorgamos una receta para cada instante difícil en el proceso enseñanza-aprendizaje. Provocamos la resolución de un problema, como el anterior, en etapas. En cada etapa observamos los varios tipos de malentendidos y malas interpretaciones que surgen y basándonos a ellos, nos esforzamos en preparar un espacio en su sentido amplio, donde la participación del estudiante sea la mayor posible.

*P* Cada caso en la matemática tiene su propio tratamiento. La satisfacción de nuestro trabajo viene muchas veces de la diversidad de las maneras, con las que podemos afrontar un problema específico. El refinamiento de las exigencias del problema, nos conduce a lo que se

debe hacer. En este caso, escriben sencillamente lo que demanda la lógica, según su propio vocabulario.

$$P(D/B)$$

Y luego inviértanlo como si fuera una fracción sencilla

$$P(B/D)$$

Creo que después, siguiendo la fórmula del libro, pueden desarrollarlo.

Observando lo que pasa en sus cuadernos, veremos que el Teorema de Bayes se aplica correctamente, un poco de manera dudosa y sólo en uno de los grupos, aparece la fórmula siguiente:

$$= \frac{P(D/B)P(B)}{P(D)} \quad (1)$$

Sin otro lado en la igualdad. Tenemos que comunicarnos más intensamente con ese grupo para que ponga al lado izquierdo en (1), empleando una cadena de comentarios y propuestas, sostenidos también por la ayuda de los otros.

Al haber conseguido un cierto grado de movilidad hacia la compleción del ejercicio, generamos un diálogo entre los participantes según el que:

*P* ¿Y ahora hemos completado la resolución?

*B*<sub>2</sub> Otra aplicación de la probabilidad total en el horizonte, calcular *P(D)*

*C*<sub>3</sub> Para mí, sería más fácil interpolar una pregunta, que requería el cálculo de *P(D)*, es decir una segunda aplicación de la probabilidad total. Y después pasar a Bayes.

*A*<sub>4</sub> No hace gran diferencia, si se introduce una pregunta semejante. Lo que a mí me alerta adelantar es escribir lo que pienses, analizarlo según alguna norma establecida y finalmente si era o no una violación de la lógica.

*B*<sub>2</sub> En general, escribe lo que pasa por la mente, aunque parezca extraordinario y haz todos los comentarios al final. Al menos, obtienes una solución, algo concreto, un número, que pueda alejar de algún modo el miedo de moverse.

## 4. CONCLUSIONES

Lo que nos esforzamos presentar en este artículo era la resolución de un ejercicio, que ataña dos teoremas fundamentales en cada asignatura introductiva a la Probabilidad: El teorema de la probabilidad total y el Teorema de Bayes. Estamos bajo la influencia didáctica de Brousseau y su tipología de situaciones didácticas, especialmente la situación de la formulación y la de validación. La primera se expresa con la división de nuestra muestra en grupos, donde el desarrollo de un diálogo productivo, en el sentido

platónico (Theodorakopulos, 2002) es una tarea fácil a causa de las buenas relaciones ya establecidas entre ellos. Seguro en este punto se podría sostener que quizás esa interacción de los grupos pudiera limitar la autonomía, que forma un principio fundamental tanto de los constructivistas como los de Brousseau. Por eso diríamos que el abordar sociocultural, cuyo resultado se considera la autonomía estudiantil, se puede combinar (Radford, 2008) para que se lleve a cabo nuestro objetivo.

La segunda se realiza con la intervención del profesor en un momento oportuno para arreglar una mal interpretación o para alentar una actividad, que genera dudas entre los estudiantes. ¿Y que pasa con la situación adidáctica definida por Brouseeau? Para conseguir esta situación según el fundador del concepto, tenemos que asemejar una situación real, donde los estudiantes utilizarán su propia experiencia y conocimiento previo y serán capaces de articular conjeturas destinadas al revolucionar la pregunta puesta.

El concepto moderno de la enseñanza requiere que el docente provoque la adaptación expectada eligiendo adecuadamente los problemas puestos. Dichos problemas, elegidos de modo que facilite su recibo por parte de los estudiantes, tiene que inducirlos a actuar, articular, pensar, y evolucionarse a través de su propia motivación. (Brouseeau, 1997, Radford, 2008).

En nuestro caso esa semejanza a la realidad no es algo fácil, ni sabemos si sea posible ejecutarla.

En el ejercicio dado, hemos incluido bastante vocabulario de la vida cotidiana, como el estado de tiempo, bueno, malo, un juego de azar etc. Pero todavía esta bastante distanciado que bosqueje una situación verdadera reproducida.

Pero si interpretamos el espacio de las situaciones reales para que se incluya todo lo que pasa en el aula, lo que llaman los norteamericanos en sus revistas voces del aula, entonces todo lo que los estudiantes escriben y dicen, cada detalle de su comportamiento es una componente de la realidad didáctica. En ese sentido, enfatizando en las dificultades afrontadas, en sus comentarios, insistimos en profundizar en su realidad de recibir el nuevo conocimiento. Además el proceso seguido está inscrito tanto en las situaciones de formulación como en las situaciones de validación.

Generalmente, el estudio probabilístico, basado y desarrollado a juegos de azar, esconde una especie de actualidad, que surge en muchas versiones durante su presentación en el aula. Depende de la disposición del profesor, de su esfuerzo para poder ampliar el medio didáctico Brousseano, emplear la realidad estudiantil. Una realidad, que contiene sus miedos, sus preocupaciones, sus contribuciones positivas y negativas hacia el aprendizaje de la matemática.

## REFERENCIAS BIBLIOGRÁFICAS Y ELECTRÓNICAS

- [1]. **BROUSSEAU, G.** (1997). “*Theory of Didactical situations in Mathematics*”. Kluwer Academic publishers.
- [2]. **FELLER, W.** (1968). “*An Introduction to Probability Theory and its Application*”, Wiley.
- [3]. **CADOCHE, L.** (2009). “*Asertividad en alumnos universitarios*”. Universidad Nacional de Litoral, Santa Fe, Argentina. Preprint.
- [4]. **CHAVARRÍA, J.** (2006). “*Teoría de las Situaciones Didácticas*”. Cuadernos de Investigación y formación en educación matemática, 1(2).
- [5]. **PÉREZ, P.R.** (2006). “*Asertividad Laboral*”. Ediciones de ASIMET. Chile.
- [6]. **RADFORD, L.** (2008). “*Theories in Mathematics Education: A Brief Inquiry into their Conceptual Differences*”. Proposed for the ICMI Survey Team 7, The Notion and role of theory in mathematics education research.
- [7]. **SANABRIA, G.** (2006). “*Tópicos precedentes al estudio de las probabilidades*”. Publicaciones Institutito Tecnológico de Costa Rica.
- [8]. **SANABRIA, G.** (2010). “*Una propuesta para la enseñanza de los Elementos de Análisis Combinatorio*”. Revista digital Matemática, Educación e Internet, 10(2).
- [9]. **THEODORAKOPULOS, I.** (2002) “*Introducción a Plato*”. Estia, Atenas.
- [10]. **VRASIDAS, C.** (2000). “*Constructivism vs. Objectivism: Implications of Interaction, Course design and Evaluation in Distance Education*”. International Journal of educational Technologies, 6(4), 339.

## DISEÑO DE MUESTRA PARA CONTEO RÁPIDO

Vera Francisco<sup>1</sup>

**Resumen.** El presente artículo discute la estimación del margen de error en la proyección de proporciones a partir de una muestra, utilizando una combinación de muestreo y censo. El censo observa todas las unidades de una cierta categoría, y la muestra escoge estratificadamente por conglomerados los de otra categoría. La estimación de los márgenes de error se realiza con la técnica del bootstrap.

**Palabras Clave:** muestreo estratificado, muestreo por conglomerados, bootstrap, proyección electoral.

**Abstract.** This article discusses the estimation of the margin of error in the estimation of proportions from a sample, using a combination of a census and a sample. The census observes all units from certain category, while the sample chooses clusters in a stratified way from another category. The estimation of the margins of error is done using bootstrap.

**Key Words:** stratified sampling, cluster sampling, bootstrap, electoral projection.

Recibido: Febrero 2013

Aceptado: Marzo 2013

### 1. INTRODUCCIÓN

El propósito de la muestra para conteo rápido es poder estimar el porcentaje de votos válidos que tendrá cada candidato presidencial en la primera vuelta de las elecciones ecuatorianas, llevada a cabo el 17 del presente mes.

Este documento explica la técnica de selección de la muestra para el conteo rápido, presenta las expresiones matemáticas que se deben usar para la estimación de las proporciones, y un análisis del margen de error de muestreo de la estimación.

### 2. POBLACIÓN

La población consiste en los electores ecuatorianos a nivel nacional y no incluye a los electores del extranjero. Estos están agrupados en juntas electorales, la mayoría de las cuales tiene 300 electores, las otras tienen menos de 300 electores. De esta manera los 11,000,000 de electores están agrupados en 39700 juntas.

Las juntas a su vez están agrupadas por recintos electorales, el más grande de los cuales tiene 123 juntas, mientras que los más pequeños tienen 2 juntas. El mayor costo de tomar una observación es el de trasladarse al recinto electoral. Una vez en el recinto, el costo de observar una junta o todas las juntas es similar, por lo que se puede observar todas las juntas en un recinto.

Hemos dividido los recintos electorales en dos grupos. Los recintos grandes y los recintos pequeños. En el grupo de los recintos grandes se realizará un censo, es decir, no habrá error de muestreo en este grupo. En el grupo de recintos pequeños se tomará una muestra, la cual tendrá un margen de error de muestreo.

Al final se combina la información de ambos grupos para obtener un solo estimador, y se calcula el error de este estimador.

### 3. SELECCIÓN DE LA MUESTRA

La muestra se seleccionó usando muestreo por conglomerados estratificado (Thompson, 2012). En cada estrato (provincias y circunscripciones) se toma una muestra aleatoria de los recintos electorales. En cada conglomerado (recinto), se realiza un censo (no una muestra).

La muestra es tomada solo del grupo de recintos pequeños. Los recintos grandes no son considerados al seleccionar la muestra.

#### Nomenclatura

A continuación describimos los símbolos empleados en las expresiones matemáticas y algunas de las relaciones entre los mismo.

$i$ : Grupo (1: recintos grandes, 2: recintos pequeños).

$N_i$ : Total de electores en grupo  $i$

$M_i$ : Total de recintos en grupo  $i$

$j$ : Provincia o circunscripción (1 a 31)

$N_{ij}$ : Total de electores en grupo  $i$  provincia o circunscripción  $j$ .

$M_{ij}$ : Total de recintos en grupo  $i$  provincia o circunscripción  $j$ .

$$N_i = \sum_{j=1}^{31} N_{ij}, i = 1, 2$$

$$M_i = \sum_{j=1}^{31} M_{ij}, i = 1, 2$$

<sup>1</sup> Vera Francisco, Ph. D., Profesor de la Escuela Superior Politécnica del Litoral (ESPOL). (e\_mail: fvera@espol.edu.ec).

$N_{ijk}$ : Total de electores en grupo  $i$  provincia o circunscripción  $j$  recinto  $k$  ( $k = 1, \dots, M_{ij}$ )

$$N_{ij} = \sum_{k=1}^{M_{ij}} N_{ijk}$$

$N_{ijc}$ : Total de electores en grupo  $i$  provincia o circunscripción  $j$  recinto  $k$  que votan por candidato  $c$  ( $c = 1, \dots, 11$ , incluyendo 8 candidatos, votos nulos, blancos y ausentes)

$N_{ij\cdot c}$ : Total de electores en grupo  $i$  provincia o circunscripción  $j$  que votan por candidato  $c$

$$N_{ij\cdot c} = \sum_{k=1}^{M_{ij}} N_{ijkc}$$

$N_{i\cdot c}$ : Total de electores en grupo  $i$  que votan por candidato  $c$

$$N_{i\cdot c} = \sum_{j=1}^{31} N_{ij\cdot c}$$

$V_{ijk}$ : Total de votos válidos en grupo  $i$  provincia o circunscripción  $j$  recinto  $k$

$$V_{ijk} = \sum_{c=1}^8 N_{ijkc}$$

$V_{ij}$ : Total de votos válidos en grupo  $i$  provincia o circunscripción  $j$

$$V_{ij} = \sum_{c=1}^8 N_{ij\cdot c}$$

$V_i$ : Total de votos válidos en grupo  $i$

$$V_i = \sum_{c=1}^8 N_{i\cdot c}$$

$V$ : Total de votos válidos

$$V = V_1 + V_2$$

$T_c$ : Total de electores que votan por candidato  $c$

$$T_c = N_{1\cdot c} + N_{2\cdot c}$$

$p_c$ : Proporción de electores (de los votos válidos) que votan por candidato  $c$

$$p_c = \frac{T_c}{V}$$

$m$ : Total de recintos en muestra de grupo 2 (los recintos pequeños)

$n$ : Total de electores en muestra de grupo 2

$m_j$ : Total de recintos en muestra de grupo 2 provincia o circunscripción  $j$

$n_j$ : Total de electores en muestra de grupo 2 provincia o circunscripción  $j$

$n_{jc}$ : Total de electores en muestra de grupo 2 provincia o circunscripción  $j$  que votan por candidato  $c$

$v_j$ : Total de votos válidos en muestra de grupo 2 provincia o circunscripción  $j$

$$v_j = \sum_{c=1}^8 n_{jc}$$

#### 4. ESTIMACIÓN DE TOTALES Y PROPORCIONES POR CANDIDATO

Para el cálculo de  $T_c$  necesitamos  $N_{1\cdot c}$  y  $N_{2\cdot c}$ . El primero se obtendrá su valor exacto por censo, mientras que el segundo se estimará a partir de una muestra. El estimador del total sería  $\hat{T}_c = N_{1\cdot c} + \hat{N}_{2\cdot c}$

El sombrero denota que no es el valor exacto sino el estimado a partir de la muestra. Por esa razón el primer término del lado derecho de la igualdad no tiene sombrero.

Para estimar el segundo total se utiliza la

$$\text{expresión matemática } \hat{N}_{2\cdot c} = \sum_{j=1}^{31} N_{2\cdot j} \left( \frac{n_{jc}}{n_j} \right)$$

Para estimar  $V$  necesitamos  $V_1$  y  $V_2$ . El primero se obtendrá su valor exacto por censo, mientras que el segundo se estimará a partir de una muestra. El estimador del total de votos válidos sería  $\hat{V} = V_1 + \hat{V}_2$ . Para estimar el segundo total se utiliza

$$\hat{V}_2 = \sum_{j=1}^{31} N_{2\cdot j} \left( \frac{v_j}{n_j} \right)$$

La estimación de la proporción de electores que votan por el candidato  $c$  es

$$\hat{p}_c = \frac{\hat{T}_c}{\hat{V}}$$

#### 5. ANÁLISIS DEL MARGEN DE ERROR DE LA PROPORCIÓN POR CANDIDATO

En el muestreo por conglomerados es difícil obtener el margen de error del estimador seleccionado cuando los conglomerados (los recintos) tienen distintos tamaños. Así que procedimos a estimar el margen de error usando la técnica de simulación de Monte Carlo. En esta

estimación suponemos un desconocimiento del comportamiento del electorado.

Una vez que se tenga tanto la información del censo de los recintos grandes, como de la muestra de los recintos pequeños, se procede a estimar el margen de error utilizando la técnica de bootstrap (Efron, 1979).

Para la estimación previa a la elección, se supuso que todas las categorías, los 8 candidatos más los votos nulos y blancos, tenían igual probabilidad en cada voto. La muestra fue estratificada por provincias, y en cada provincia se procedió a escoger un número de conglomerados. Se repitió la simulación mil veces, y se obtuvieron mil posibles errores de muestreo. El error estándar de los errores fue de 0.006, mientras que el máximo fue 0.025. El percentil 95 fue de 0.012, indicando que la probabilidad que el error de muestreo sea menor a 1.2 puntos porcentuales es 0.95. Este error es para la estimación a nivel nacional.

La noche de las elecciones, al tener los datos reales que iban llegando del conteo rápido, se procedió a realizar las simulaciones basadas en bootstrap. Por cuestiones logísticas no se pudo obtener todos los recintos en la muestra seleccionada, pero se obtuvo un censo para los recintos con más de 53 mesas.

En la muestra final, el margen de error a nivel nacional fue ahora mucho más bajo, 0.0038, es decir, 0.38 puntos porcentuales. Se proyectó con este margen de error que la proporción del candidato ganador era 56.47% de los votos.

Se obtuvo proyecciones a nivel provincial, obteniendo los mayores márgenes de error en las provincias de Bolívar, Morona Santiago, Napo, Orellana, Sucumbíos y Zamora Chinchipe, las cuales tuvieron márgenes de error superiores a 4 puntos porcentuales, pero menores a 10.

Las provincias de Guayas, Manabí, Pichincha y Azuay, tuvieron márgenes de error de 0.62, 1.04, 0.82 y 1.56 puntos porcentuales, respectivamente.

## 6. CONCLUSIONES

La estimación del margen de error exacta en muestreo por conglomerados, cuando estos son de distintos tamaños, es un problema *np-hard*, que resulta casi imposible calcularlos. La técnica del bootstrap es una buena alternativa y da estimaciones razonables.

## ANEXO: CÓDIGO EN R

```

#Parámetros
corte=53 #Mayores a este número son grandes
basefile="Recintos.xlsx"
muestrafile="CalculoError.xlsx"
varnames=c("EMITIDOS","VALIDOS","NULOS","BLANCOS",
          "LUCIO.GUTIERREZ","ALVARO.NOBOA",
          "NELSON.ZAVALA","ALBERTO.ACOSTA",
          "GUILLERMO.LASSO","MAURICIO.RODAS",
          "NORMAN.WRAPY","RAFAEL.CORREA")

#Trabajo
library(xlsx)

remuestreo=function(pob,mues,index,pmuestra)
{
  res=matrix(0,nrow=nrow(pob),ncol=ncol(pmuestra))
  colnames(res)=varnames
  res=as.data.frame(res)
  if(nrow(pmuestra)>0)
  {
    res[index,]=pmuestra
    res[!index,]=pmuestra[sample.int(nrow(pmuestra),nrow(pob)-nrow(mues),TRUE),]
  }
  res
}

muestrear=function(res,grandes,ng,np)
{
  i=1:nrow(res)
  ig=i[grandes]
  mg=res[ig[sample.int(length(ig),ng)],]
  ip=i[!grandes]
  mp=res[ip[sample.int(length(ip),np)],]
  rbind(mg,mp)
}

estimar=function(pob,mues,grandes)
{
  ig=mues$COD_RECINTO %in% pob[grandes,$COD_RECINTO
  rg=apply(mues[ig,varnames],2,sum)
  sg=sum(mues[ig,"TOT_INSCRITOS"])
  if(sg>0)
    rg=(rg/sg)*sum(pob[grandes,"TOT_INSCRITOS"])
  rp=apply(mues[!ig,varnames],2,sum)
  sp=sum(mues[!ig,"TOT_INSCRITOS"])
  if(sp>0)
    rp=(rp/sp)*sum(pob[!grandes,"TOT_INSCRITOS"])
  rp+rg
}

estprop=function()
{
  rescir=t(mapply(estimar,basecir,muestracir,grandescir))
  pestcir=cbind(rescir[,1:4]/Ncir,rescir[,-(1:4)]/rescir[,,"VALIDOS"])
  totest=colSums(rescir)
  N=sum(Ncir)
  pest=c(totest[1:4]/N,totest[,-(1:4)]/totest["VALIDOS"])
}

```

```

rbind(Nacional=pest,pestcir)
}

error=function(i)
{
  remuestra=mapply(remuestreo,basecir,muestracir,indexcir,pmuestracir,SIMPLIFY=FALSE)
  rescir=mapply(function(pob,remues)
  {
    r=pob$TOT_INSCRITOS*remues
    colnames(r)=varnames
    r
  },basecir,remuestra,SIMPLIFY=FALSE)
  rescir=mapply(function(pob,res)cbind(pob,res),basecir,recir,SIMPLIFY=FALSE)
  muescir=mapply(muestrear,recir,grandescir,ngrandescir,npequeñoscir,SIMPLIFY=FALSE)
  recir=t(mapply(estimar,basecir,muescir,grandescir))
  colnames(recir)=varnames
  pestcir=cbind(recir[,1:4]/Ncir,recir[,-(1:4)]/recir[,"VALIDOS"])
  rcir=t(mapply(function(pob,remues)colSums(pob$TOT_INSCRITOS*remues),basecir,remuestra))
  pcir=cbind(rcir[,1:4]/Ncir,rcir[,-(1:4)]/rcir[, "pVALIDOS"])
  ecir=pcir-pestcir
  totest=colSums(recir,na.rm=TRUE)
  tot=colSums(rcir,na.rm=TRUE)
  N=sum(Ncir)
  pest=c(totest[1:4]/N,totest[-(1:4)]/totest["VALIDOS"])
  p=c(tot[1:4]/N,tot[-(1:4)]/tot["pVALIDOS"])
  e=pest-p
  rbind(Nacional=e,ecir)
}

transf=function(cir)
{
  cir=as.character(cir)
  i1=cir=="Guayas Circ. 1"
  i2=cir=="Guayas Circ. 2"
  i3=cir=="Guayas Circ. 3"
  i4=cir=="Guayas Circ. 4"
  i=i1|i2|i3|i4
  cir[i]="Guayas"
  i1=cir=="Manabi Circ. 1"
  i2=cir=="Manabi Circ. 2"
  i=i1|i2
  cir[i]="Manabí"
  i1=cir=="Pichincha Circ. 1"
  i2=cir=="Pichincha Circ. 2"
  i3=cir=="Pichincha Circ. 3"
  i4=cir=="Pichincha Resto"
  i=i1|i2|i3|i4
  cir[i]="Pichincha"
  cir=as.factor(cir)
}

base=read.xlsx(basefile,1,encoding="UTF-8")
base$CIRCUNSCRIPCION=transf(base$CIRCUNSCRIPCION)

#Correr desde aquí cada vez que se actualiza la muestra

muestra=read.xlsx(muestrafile,1,encoding="UTF-8")
muestra$CIRCUNSCRIPCION=transf(muestra$CIRCUNSCRIPCION)

basecir=split(base,base$CIRCUNSCRIPCION)

```

```

muestracir=split(muestra,muestra$CIRCUNSCRIPCION)
indexcir=mapply(function(pob,mues)pob$COD_RECINTO %in%
mues$COD_RECINTO,basecir,muestracir)
pmuestracir=lapply(muestracir,function(x)x[,varnames]/x$TOT_INSCRITOS)
pvarnames=paste("p",varnames,sep="")
grandescir=lapply(basecir,function(x)x$JUN_JUNR>corte)
ngrandescir=mapply(function(i,j) sum(i&j),indexcir,grandescir)
npequeñoscir=mapply(function(i,j) sum(i&!j),indexcir,grandescir)
Ncir=sapply(basecir,function(x)sum(x$TOT_INSCRITOS))

ep=estprop()
es<-lapply(1:200,error)
efin=array(dim=c(dim(es[[1]]),length(es)))
for(i in 1:length(es)) efin[,i]=es[[i]]
dimnames(efin)=c(dimnames(es[[1]]),list(1:length(es)))
marg=apply(efin,c(1,2),quantile,probs=0.95,na.rm=TRUE)

#wb=createWorkbook()
wb=loadWorkbook("Análisis de Error.xlsx")
wbSheets=getSheets(wb)
#sheetEst=createSheet(wb,sheetName="Proyección")
cs1=CellStyle(wb,dataFormat=DataFormat("0,0%"))
l1=NULL
for(i in 1:ncol(ep))l1=c(l1,list(cs1))
#addDataFrame(ep,sheetEst,colStyle=l1)
addDataFrame(ep,wbSheets[[1]])
#sheetMar=createSheet(wb,sheetName="Margen de Error")
cs2=CellStyle(wb,dataFormat=DataFormat("0,00%"))
l2=NULL
for(i in 1:ncol(marg))l2=c(l2,list(cs2))
#addDataFrame(marg,sheetMar,colStyle=l2)
addDataFrame(marg,wbSheets[[2]])
saveWorkbook(wb,"Análisis de Error.xlsx")

```

### REFERENCIAS BIBLIOGRÁFICAS Y ELECTRÓNICAS

- [1]. **THOMPSON, S.** (2012). Sampling, 3rd edition. Wiley series in probability and statistics, Wiley, New Jersey.
- [2]. **EFRON, B.** (1979). “*Bootstrap methods: another look at the jackknife*”. The annals of Statistics, 7, 1-26.

## 1. ACERCA DEL FORMATO DE LOS ARTÍCULOS

Quienes presenten trabajos para publicación en la revista deberán regirse por las siguientes disposiciones:

- El trabajo será escrito en castellano
- Contendrá un **RESUMEN** y **ABSTRACT** que se presentará al comienzo del mismo, con no más de 100 palabras. Este resumen será conciso, impersonal e incluirá de manera sucinta los resultados y conclusiones de la investigación, luego de lo cual se presentará una lista de no más de cinco **PALABRAS CLAVES** utilizadas en el artículo, facilitando de esta manera la indexación, las mismas deben presentarse en inglés (**KEYWORD**).
- A continuación, irá la primera sección del trabajo denominada **INTRODUCCIÓN**, las restantes secciones las titulará el autor de acuerdo a las características del trabajo.
- La sección final se denominará **CONCLUSIONES** y en la misma se discutirán los resultados a los que haya llegado el autor.
- La sección **REFERENCIAS BIBLIOGRÁFICAS** será numerada y contendrá a más del nombre del autor(es) del artículo o texto, la fecha y lugar de publicación. Se exhorta incluir como referencias las publicaciones esenciales, esto es, las que realmente sirven de sustento al investigador en su trabajo.

## 2. GENERALIDADES

Los manuscritos para publicación en la revista serán enviados al editor de la revista en un **CD o vía electrónica**, escritos en el procesador de palabras **WORD O LATEX**. Se entregarán además cuatro copias en papel A4; la longitud del artículo no excederá las diez páginas, será escrito a espacio simple, en doble columna, letra Times New Roman, tamaño 10 para el texto principal y tamaño 8 para el resumen, las palabras claves y las referencias bibliográficas. Los gráficos, tablas y fotos serán numerados de manera consecutiva en su parte superior, utilizando números arábigos y respectivamente rotulados. Los márgenes superiores, inferiores, derecho e izquierdo serán de 2.5 cm, 2.5 cm, 2 y 3.5 cm respectivamente.

Las secciones y subsecciones serán debidamente numeradas en forma consecutiva. Las secciones se rotularán centradas en mayúsculas, mientras que las subsecciones estarán alineadas a la derecha.

No habrá espacios entre párrafos, cada uno iniciará con una sangría de 2 espacios.

Las copias enviadas al editor no serán devueltas pero sí tres separatas luego de publicado el artículo. Si el trabajo es publicado en la revista, el mismo no podrá ser reproducido total o parcialmente sin el consentimiento del Departamento de Matemáticas de la Facultad de Ciencias Naturales y Matemáticas.

### Recepción de manuscritos e Información:

FCNM-ESPOL, Campus Gustavo Galindo  
Teléfonos: 042 269525 – 042 269542  
Mat. Eduardo Rivadeneira Molina,  
Director de Publicaciones del DM  
(e-mail: erivaden@espol.edu.ec )  
Correo de la Revista Matemática  
(revismat@espol.edu.ec)

## CONTENIDO

<b>EDITORIAL.....</b>	<b>5</b>
<b>MODELANDO EL TRANSPORTE DE IONES DE POTASIO A TRAVÉS DE UN CANAL IONICO DE UNA MEMBRANA CELULAR</b>	
Iza Peter.....	7
<b>HOMOTOPÍA Y CONTINUACIÓN NUMÉRICA EN SISTEMAS NO LINEALES</b>	
Martín Carlos.....	13
<b>SISTEMAS NO LINEALES Y LA METAHEURÍSTICA SCATTER SEARCH</b>	
Martín Carlos.....	24
<b>INCORPORANDO BROSSEAU EN DOS TEOREMAS FUNDAMENTALES: PROBABILIDAD TOTAL Y BAYES</b>	
Petakos Kyriakos.....	34
<b>DISEÑO DE MUESTRA PARA CONTEO RÁPIDO</b>	
Vera Francisco.....	39