

“Modelo poblacional con algoritmos genéticos”

Eduardo Véliz Quintero¹, Luis Rodríguez Ojeda².

RESUMEN

Para el desarrollo de este trabajo, “MODELO POBLACIONAL CON ALGORITMOS GENÉTICOS”, he investigado la rama de la inteligencia artificial, como son los algoritmos genéticos. Primero presento en forma general los aspectos que envuelven los algoritmos genéticos, parto de la necesidad de optimizar, así como su historia y posibles aplicaciones y luego he cubierto detalladamente todo lo que pude investigar sobre la teoría de los algoritmos genéticos, sus fundamentos matemáticos, tipos de algoritmos genéticos, representación del mismo, partes y operadores más importantes de los algoritmos genéticos. Además se presenta una pequeñísima comparación con otros métodos de programación evolutiva, tales como los algoritmos genéticos paralelos, también llamados algoritmos meméticos, las redes neuronales, y los autómatas celulares.

Además hago un estudio de los principales modelos de crecimiento poblacional, aplicados tanto a humanos como a animales y luego propongo un modelo basado en el paradigma de los algoritmos genéticos llevando a cabo una simulación poblacional utilizando una aplicación que he desarrollado completamente en el lenguaje de cuarta generación, Visual C++ 6.0. Finalmente, comparo entre mi modelo de crecimiento poblacional y los modelos convencionales, basándome en un subconjunto de datos de los que son generados por mi modelo, y realizar una predicción de los restantes.

INTRODUCCION

En la naturaleza, los individuos compiten entre sí por recursos, tales como comida, agua y refugio. Adicionalmente, entre los animales de una misma especie, aquellos que no obtienen acierto tienden a tener un número reducido de descendientes, teniendo por tanto, menor probabilidad de que sus genes sean propagados a lo largo de sucesivas generaciones.

Los algoritmos genéticos es una técnica de búsqueda basada en la teoría de la evolución de Darwin, que ha cobrado tremenda popularidad en todo el mundo durante los últimos años. Aquí presento los conceptos básicos que se requieren para comprenderla, así como unos ejemplos que permitan a los lectores comprender cómo aplicarla al problema de su elección.

Un aspecto por demás importante de ellos es su aplicación como técnica de optimización que se basa en el azar, pero que aprovecha criterios que la naturaleza ha desarrollado, tales como la selección de los cromosomas más aptos, el cruce de genes en los cromosomas y la mutación. Por esto, no es de extrañarse que en algoritmos genéticos se utilicen términos tomados de la genética natural.

Un algoritmo genético puede converger en una búsqueda de azar, pero su utilización asegura que ningún punto del espacio de búsqueda tiene probabilidad cero de ser examinado.

En el desarrollo de esta tesis utilizo la técnica de los algoritmos genéticos para implantar un modelo de crecimiento poblacional algorítmico de gran robustez que permite la simulación, así como la optimización de sus parámetros.

Para mayor comprensión de la presente tesis de parte de cualquier lector, presento una explicación modesta de los modelos de crecimiento poblacional, aunque es también bastante buena y suficiente para el aprendizaje de los mismos, y la posterior aplicación de los algoritmos genéticos a los modelos poblacionales.

CONTENIDO

- **El teorema de los esquemas**

Veamos algunos conceptos que son necesarios para comprender el Teorema:

- n : Numero total de individuos de una población.
- $f(h)$: Valor de la función de adaptación para un individuo concreto denominado h .
- $f(t)$: Valor medio de la función de adaptación de todos los individuos de una población en una generación t .
- $p(t)$: Representa la población en una generación t .
- $h \in p(t)$: Expresión que indica que un individuo h pertenece tanto a la población en un instante de tiempo como al conjunto de individuos representados por el esquema s .
- $\hat{u}(s,t)$: Valor medio de la función de adaptación de los individuos del esquema s en la generación t .
- $m(s,t)$: Numero de individuos del esquema s en la generación t .
- $m(s,t+1)$: Numero de individuos del esquema s en la siguiente generación.
- $E[m(s,t+1)]$: Significa lo mismo que la expresión anterior.
- $Pr(h) = f(h)/nf(t)$: Probabilidad de que un individuo sea seleccionado en la generación t
- $Pr(h \in s) = \hat{u}(s,t)m(s,t)/nf(t)$: Probabilidad de que un individuo del esquema s sea seleccionado
- $E[m(s,t+1)] = \hat{u}(s,t)m(s,t)/f(t)$: Considerando solamente el efecto de la selección sobre la población, esta expresión nos da el valor del numero de individuos del esquema s en la siguiente generación, que será la probabilidad de que seleccionemos un individuo por el número de individuos totales de la población.

¹ Ingeniero en Estadística Informática 2001

² Director de Tesis, Ingeniero Eléctrico, Escuela Superior Politécnica del Litoral, Maestría en Ciencias de Computación, en Estados Unidos, Universidad Northwester, 1978, Profesor de ESPOL desde 1978.

El análisis hecho arriba considera sólo el paso de selección del algoritmo genético, la combinación y la mutación deben ser considerados también. El Teorema fundamental considera únicamente la posibilidad de influencia negativa de estos operadores genéticos. El Teorema completo se enuncia a continuación, tras haber definido antes algunos conceptos:

- L: Longitud de un individuo.
- pc: Probabilidad de que el operador de combinación se aplique a un individuo.
- pm: Probabilidad de que un bit de un individuo sea mutado.
- o(s): Numero de bits definidos en el esquema s.
- d(s): Distancia máxima entre dos bits definidos en el esquema s.

$$E[m(s, t+1)] \geq \left[\frac{\hat{u}(s,t)m(s,t)}{f(t)} \right] \cdot \left[1 - \frac{pc \cdot d(s)}{(L-1)} \right] \cdot \left[(1 - pm)^{o(s)} \right]$$

El primer factor de la expresión representa, como ya hemos visto, el efecto del operador de selección en la siguiente generación. El factor medio y último describen el efecto de los operadores de combinación (de un solo bit) y mutación, en particular describen la probabilidad de que un individuo concreto que es representado por el esquema s pueda seguir siendo representado por dicho esquema tras haber sido aplicado sobre el operador de combinación y mutación respectivamente.

- **El algoritmo genético básico**

<p><i>Inicio</i></p> <p><i>Genere la población inicial</i></p> <p><i>Halle su genotipo</i></p> <p><i>Halle la función de adaptación</i></p> <p><i>Mientras < no converge ></i></p> <p style="padding-left: 2em;"><i>Seleccione los de padres</i></p> <p style="padding-left: 2em;"><i>Realice el crossover</i></p> <p style="padding-left: 2em;"><i>Realice la mutación</i></p> <p style="padding-left: 2em;"><i>Actualice la población</i></p> <p style="padding-left: 2em;"><i>Halle su genotipo</i></p> <p style="padding-left: 2em;"><i>Halle la función de adaptación</i></p> <p><i>Fin_Mientras</i></p> <p><i>Muestre la solución optima</i></p> <p><i>Fin</i></p>

El algoritmo genético básico

Cuadro 2.1

Al igual que todo algoritmo, el Algoritmo genético básico también tiene una estructura principal que consta de operadores genéticos tales como la selección natural, el crossover, la mutación y la adaptación. Además de estos puede agregarse otros según las necesidades de la programación.

Además de los algoritmos genéticos existen otras técnicas de programación evolutiva como las redes neuronales artificiales y los autómatas celulares, ya que no en todos los problemas se puede aplicar los algoritmos genéticos.

- **Modelos de crecimiento poblacional**

Las poblaciones experimentan procesos que hacen variar el número de individuos y que cuando se compensan entre sí el número el tamaño se mantiene constante.

La Mortalidad y la natalidad se pueden expresar en tasas porcentuales y diferenciales. Los procesos que hacen cambiar a las poblaciones son: emigración, inmigración, natalidad y mortalidad.

- Mortalidad: Número de individuos muertos.
- Natalidad: Número de individuos nacidos.

En un periodo de tiempo, con relación al número de individuos al inicio del periodo de tiempo. La natalidad y la mortalidad se dan en forma simultánea y su diferencia mostrará que la población crezca o disminuya.

Los modelos de crecimiento poblacional más conocidos son el exponencial y el logístico, aunque también existen una gran cantidad de modelos estocásticos.

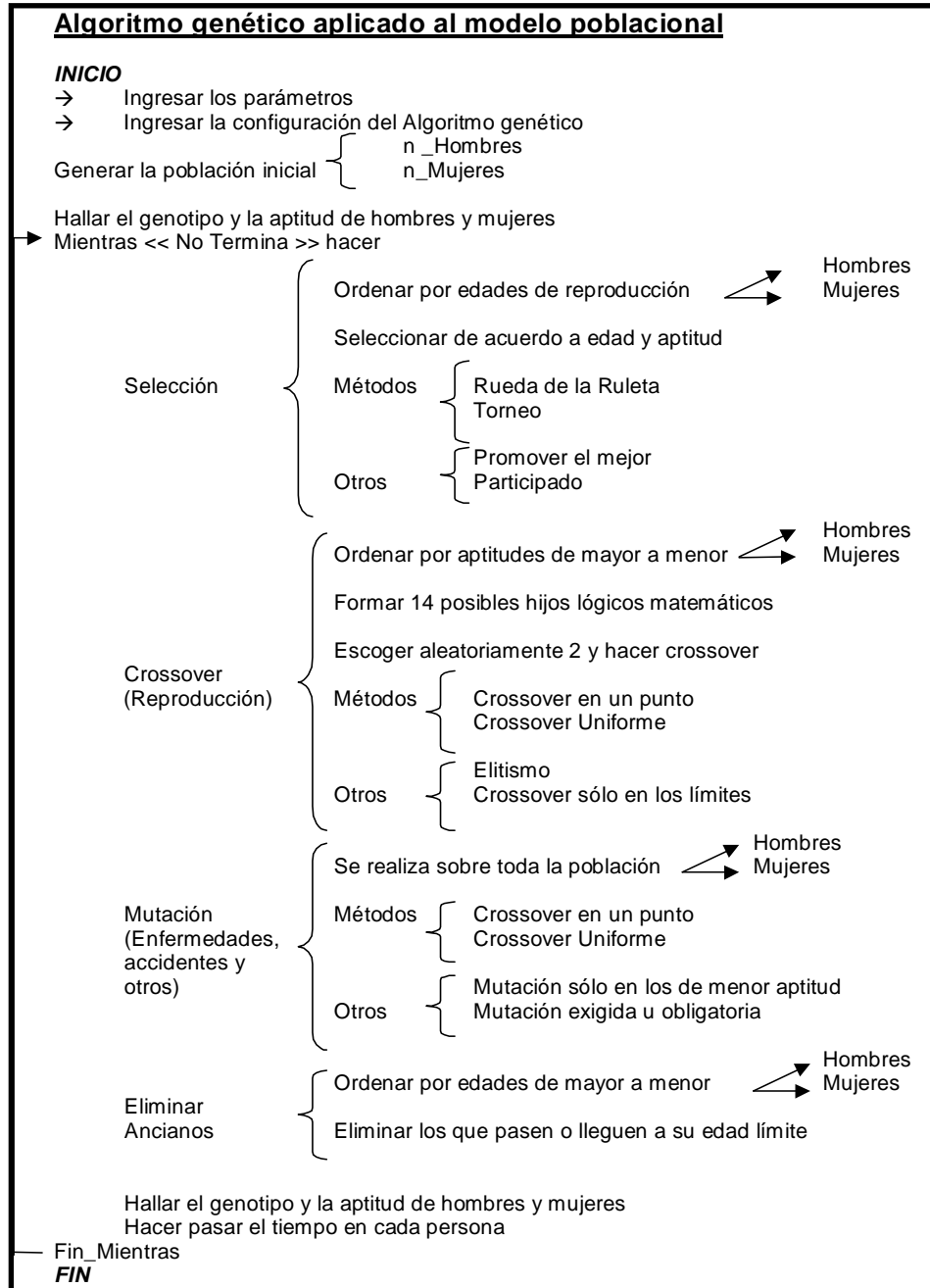
Modelo de crecimiento exponencial

$$Y = \beta_0 e^{\beta_1 x}$$

Modelo de crecimiento logístico

$$Y = \frac{1}{\frac{1}{U} + \left(\beta_0 \beta_1^x \right)}$$

Yo desarrollé una simulación de la población usando el esquema de los algoritmos genéticos y obtuve los resultados necesarios para mostrar el crecimiento de la población mediante el siguiente algoritmo, el cual fue desarrollado completamente en Visual C++ 6.0 como aplicación de tipo MFC.



Algoritmo genético aplicado a los modelos poblacionales

Cuadro 3.1

Además mi aplicación, en otro módulo, tratará de estimar los valores de los beta en la regresión de tipo exponencial o logística, y finalmente en un tercer módulo, seleccionará el mejor modelo posible cuando se explica el crecimiento mediante varias variables.

Resultados de la Simulación							
Año	Hombres	Mujeres	Total	Crecimiento	Año	M. Aptitud	Aptitud P.
00	184	187	371	+0.000000	00	0.041208	0.027689
01	220	227	447	+0.204852	01	0.037542	0.025094
02	253	263	516	+0.154362	02	0.034942	0.023368
03	248	256	504	-0.023256	03	0.035355	0.023628
04	246	249	495	-0.017857	04	0.035675	0.023821
05	239	244	483	-0.024242	05	0.036116	0.024226
06	234	239	473	-0.020704	06	0.036496	0.024379
07	230	231	461	-0.025370	07	0.036967	0.024631
08	226	226	452	-0.019523	08	0.037334	0.024853
09	222	217	439	-0.028761	09	0.037882	0.025280
10	217	212	429	-0.022779	10	0.038321	0.025680
11	213	207	420	-0.020979	11	0.038730	0.025964
12	210	202	412	-0.019048	12	0.039104	0.026224
13	205	195	400	-0.029126	13	0.039686	0.026703
14	201	192	393	-0.017500	14	0.040038	0.026936
15	194	184	378	-0.038168	15	0.040825	0.027398
16	202	182	384	+0.015873	16	0.040505	0.027304
17	215	184	399	+0.039063	17	0.039736	0.026744

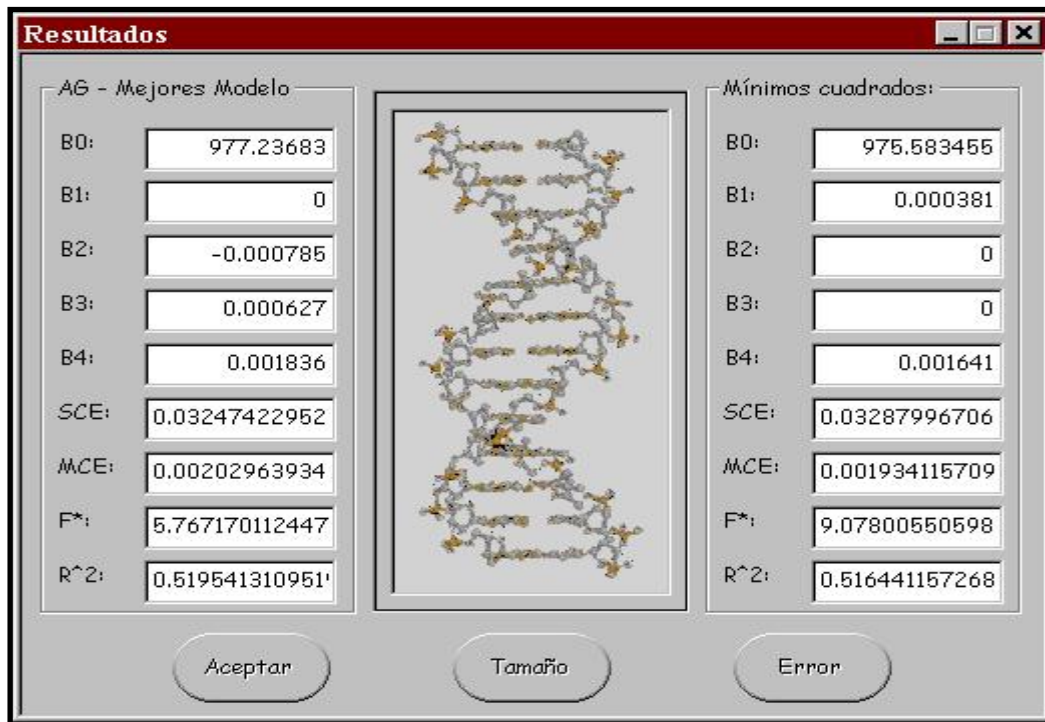
Pantalla de los resultados del Simulador de los Algoritmos Genéticos – Modelo Poblacional

Figura 4.4

Resultados	
AG - Mejores Estimadores:	Mínimos cuadrados:
B0: 9355.46875	B0: 9865.29104554
B1: 0.015625	B1: 0.01893850368
SCE: 0.93632454670	SCE: 0.159818776994
MCE: 0.02035488145	MCE: 0.00347432123
F*: 124.173354768	F*: 950.988540785
R^2: 0.787207817312	R^2: 0.95589596797
Modelo:	
AG - Mejores Estimadores:	$P(t) = 1 / (-1.08e-062 + 9355 * 1.563e-002^t)$
Est. Mínimos Cuadrados:	$P(t) = 1 / (-1.08e-062 + 9865 * 1.894e-002^t)$
<input type="button" value="Aceptar"/> <input type="button" value="Modelo"/> <input type="button" value="Error"/> <input type="button" value="Intervalos"/>	

Pantalla de resultados para los modelos dependientes del tiempo mediante algoritmos genéticos

Figura 4.9



Pantalla de resultados para los modelos dependientes de varias variables y mediante algoritmos genéticos

Figura 4.14

- **Observaciones.**

En este capítulo, he revisado los resultados de realizar una simulación poblacional utilizando los algoritmos genéticos, y luego ajusté los datos a un modelo exponencial, debido a la tendencia de los mismos, tratando inclusive de predecir valores que ya eran conocidos, y finalmente he comparado los resultados de la ejecución de mi aplicación del modelo exponencial con algoritmos genéticos y con mínimos cuadrados.

Este simulador de crecimiento poblacional con algoritmos genéticos funciona con valores de aptitudes aleatorias tomadas del algoritmo genético, y así mismo con cada uno de sus parámetros, por lo que los resultados podrían variar de una corrida a otra. En las aplicaciones de los modelos exponencial y logístico, a mayor cantidad de nodos de búsqueda, darán una mejor aproximación de los resultados.

En el algoritmo genético de optimización para selección de un mejor modelo, el algoritmo dará un buen resultado, que probablemente no sea el mejor modelo que tengamos, pero sí es uno de los que más explican dichos resultados.

CONCLUSIONES

1. Los algoritmos genéticos permiten realizar una excelente simulación del crecimiento de una población, haciendo unas pequeñas modificaciones a los operadores del algoritmo genético básico y aumentando algunos operadores.
2. En los algoritmos genéticos el operador de mayor importancia es el operador de selección, ya que éste permite una mejor convergencia en menor tiempo de ejecución si es bien implantado; pero puede suceder lo contrario si se lo implanta mal.
3. En los algoritmos genéticos la función de adaptación es la misma función que se desea optimizar; pero puede utilizarse algún conocimiento matemático para optimizar el tiempo de ejecución.
4. En los algoritmos genéticos el crossover se encarga de buscar mejores individuos mediante el intercambio de bits y la mutación permite la diversificación de la población para evitar la convergencia prematura; pero el abuso de la mutación provoca una búsqueda aleatoria.
5. Al utilizar un mayor número de individuos se puede abarcar un mayor número de puntos discretos en el intervalo de búsqueda, y por lo tanto se obtendrá estimadores más cercanos a los que se obtienen mediante métodos convencionales o búsquedas heurísticas.
6. Los modelos poblacionales tienen crecimiento que puede modelarse bastante bien mediante el modelo exponencial y el modelo logístico; pero un modelo lineal es muy malo para predecir la tendencia de los mismos.
7. Los estimadores de β_0 y β_1 de los modelos exponenciales y logísticos que se obtienen mediante los algoritmos genéticos son casi los mismos que se obtienen mediante el método convencional de los mínimos cuadrados, siempre que se tome en cuenta el intervalo de búsqueda de estos estimadores.
8. La selección de un modelo que explique la regresión de la mejor manera mediante algoritmos genéticos es una de las principales utilidades que estos algoritmos presentan, ya que cuando existe un gran número de posibles modelos, los algoritmos genéticos pueden encontrar un modelo de gran explicación en poco tiempo.
9. El tiempo de ejecución de los algoritmos genéticos crece exponencialmente y es directamente proporcional al de acuerdo al número de individuos de la población o nodos de la búsqueda, así como al número de iteraciones.

REFERENCIAS

1. Goldberg, D.E. (1989) "Genetic Algorithms in Search, Optimization and Machine Learning", Reading. Mass.: Addison- Wesley.
2. Holland, J. (1975) "Adaptation in Natural and Artificial Systems", The University of Michigan Press, Ann Arbor.
3. Dawkins, R. (1982) "The Extended Phenotype: The Gene as a Unit of Selection", Oxford Univ. Press, Oxford.
4. Obitko, M. (1998) "Introduction to Genetic Algorithms", Alemania, University of Applied Sciences.
<http://cs.felk.cvut.cz/~xobitko/ga/>
5. Herrán , M. (1995) "Vida Artificial", España, El grupo Gaia.
<http://www.geocities.com/SiliconValley/Vista/7491/>
6. Contreras, J. (1996) "Operadores genéticos", Venezuela,
<http://www.iamnet.com/users/jcontre/genetic/operadores.htm>
7. Larrañaga, P. (2000) "Introducción a los algoritmos genéticos",
<http://www.geocities.com/CapeCanaveral/9802/3d5ca000.htm>
8. Ferrer, M. (1997) "Población y recursos naturales", Universidad de navarra, Departamento de humanidades biomedicas
<http://www.eumed.net/cursecon/2/index.htm>
9. "Población, Conceptos Básicos"
<http://www.lafacu.com/>
10. Preciado, J. (1999) "Futuro",
<http://www.pagina12.com.ar/1999/futuro.htm>
11. "La ecología y los insectos", Perú, Universidad de Lamolina
<http://tarwi.lamolina.edu.pe/~acg/ç>
12. Cordon J. (1999) "La ecuación logística, modelo poblacional", Dpto. de Matemáticas, Estadística y Computación (MATESCO)
<http://193.144.183.17/~cordon/doctor98/proyecto/logistic.htm>
13. "Redes neuronales artificiales"
<http://www.gc.ssr.upm.es/inves/neural/ann2/introduc/intro.htm>