

# Algoritmos Genéticos

Módulo 4 : Técnicas computacionales avanzadas para modelar fenómenos sociales  
Concentración en Economía Aplicada y Ciencia de Datos  
ITESM

6 de octubre de 2022



# Algoritmos genéticos

- Un **Algoritmo Genético** es una meta-heurística de optimización basada en los principios de la genética y la selección natural ([La Rocca, 2022](#), p. 628).
- Incorporan procesos de adaptación basados en mecanismos de la evolución natural:
  - ▶ **Teoría de la herencia**,
  - ▶ **Selección natural**.
- Explotan los principios de la selección natural, esto es, de qué tan bien un organismo se adapta a su ambiente, sobrevive y genera descendencia.

# Algoritmos genéticos

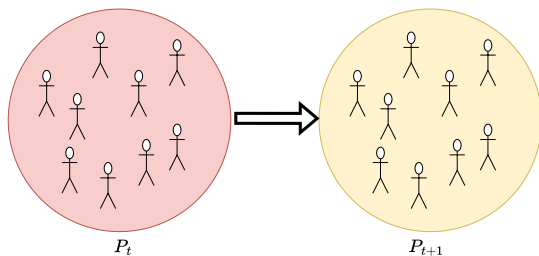
Los algoritmos genéticos presentan un balance adecuado entre

- **Explotación**, ¿cómo utilizamos la solución ya conocida para construir una mejor solución?
- **Exploración** del espacio de búsqueda de posibles soluciones.

Este balance lo logran por los mecanismos de **selección** de los individuos de la población en función a su **aptitud** y del **operador genetico** de **cruza**, para el caso de la **explotación**, mientras que para la **exploración** la base es el operador de **mutación**.

# Algoritmos genéticos

- Un algoritmo genético simula la dinámica de una población genética.
- En este proceso se genera una base de conocimientos en función de la forma cómo los individuos se adaptan al medio al cual se enfrentan (el problema a resolver).
- La población se va modificando de acuerdo a la **probabilidad de selección** para contribuir en la generación de una nueva generación de la población.



**En este contexto ¿Qué significa la adaptación?**

## En este contexto ¿Qué significa la adaptación?

La adaptación involucra una **transformación progresiva** de los individuos.

# Elementos generales de un algoritmo genético

- **Codificación de la población.**
- **Decodificación de la población.**
- **Selección de individuos.**
- **Reproducción de individuos.**

# Algoritmos Genéticos vs Algoritmos tradicionales de optimización

Genéticos	Tradicionales
<ul style="list-style-type: none"><li>• Tratan el problema como una caja negra (no requieren información a priori del problema).</li><li>• Sólo requieren evaluar la función.</li><li>• La búsqueda se realiza en una población de soluciones posibles, logrando con ello reducir la probabilidad de caer en óptimos locales.</li><li>• Emplea reglas de transición probabilísticas.</li></ul>	<ul style="list-style-type: none"><li>• Necesitan información del problema, como derivadas, gradientes, hessianos, etc.</li><li>• Trabajan <i>punto a punto</i> para determinar el siguiente punto a analizar, ocasionando llegar a óptimos locales.</li><li>• Emplea reglas determinísticas.</li><li>• Necesita partir de un punto factible.</li></ul>



# Factores que afectan el desempeño del algoritmo

- **Convergencia:**

- ▶ Presión selectiva: buscar en regiones del espacio de búsqueda que son prometedoras. Esto se logra mediante el cruzamiento de los individuos con mejor aptitud.
- ▶ Esto es necesario para que la búsqueda no sea aleatorio.

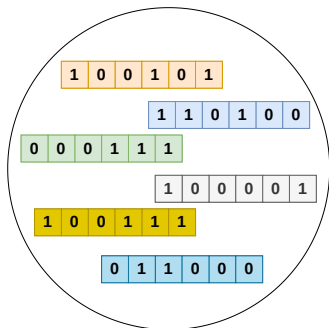
- **Diversidad de la población:**

- ▶ Evita una convergencia prematura.
- ▶ La carencia de la diversidad se traduce en que todos los individuos son iguales, llevando a un proceso de explotación no adecuado, llevando así a óptimos locales.

# Algoritmo Genético Binario (simple)

Un **Algoritmo Genético Binario** se caracteriza por contar con un mecanismo de selección de los individuos más aptos de la población, los cuales contribuirán en la generación de la siguiente generación, y del uso de los operadores de cruce y mutación.

- **Generación de la población inicial.**

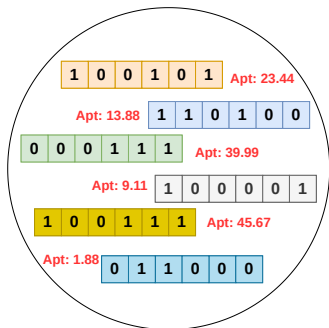


**Figura:** Generación de la población: se parte de individuos representados como cadenas de  $m$  bits, los cuales pueden ser generados de forma aleatoria.

# Algoritmo Genético Binario (simple)

Un **Algoritmo Genético Binario** se caracteriza por contar con un mecanismo de selección de los individuos más aptos de la población, los cuales contribuirán en la generación de la siguiente generación, y del uso de los operadores de cruce y mutación.

- Generación de la población inicial.
- Evaluación de Fitness.



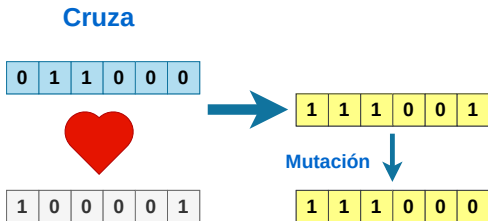
**Figura:** Se evalúa cada individuo con respecto al problema a resolver y se le asigna una medida llamada **fitness** o **aptitud** que es un valor numérico de qué tan bueno es el individuo para resolver el problema



# Algoritmo Genético Binario (simple)

Un **Algoritmo Genético Binario** se caracteriza por contar con un mecanismo de selección de los individuos más aptos de la población, los cuales contribuirán en la generación de la siguiente generación, y del uso de los operadores de cruce y mutación.

- Generación de la población inicial.
- Evaluación de Fitness.
- Selección
- Cruza y Mutación



**Figura:** Se aplican los operadores genéticos de **cruza** y **mutación** a cada pareja seleccionada de la generación presente hasta que se crean  $n$  nuevos individuos. La generación actual es reemplazada por completo por la generación nueva

# Algoritmo Genético Binario (simple)

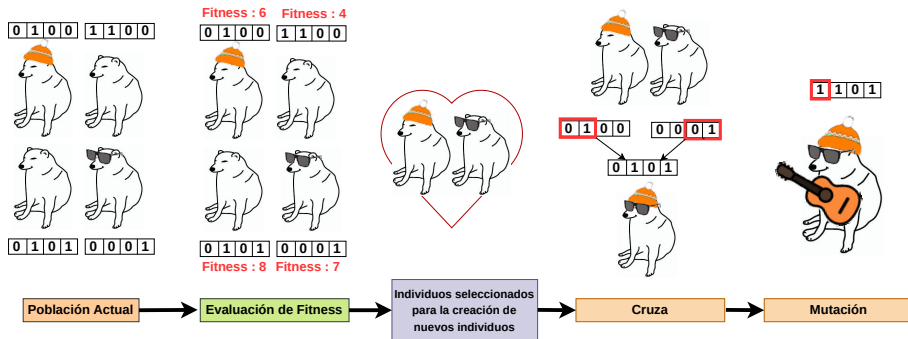


Figura: Elementos de un Algoritmo Genético Binario (simple)

# Algoritmo Genético Binario (simple)



Figura: Elementos de un Algoritmo Genético Binario (simple)

En la selección se escogen a los individuos que cuentan con las mejores condiciones para contribuir en la formación de la siguiente generación.

Algunos métodos de selección son:

- **Reproducción proporcional o Método de la ruleta:**

- ▶ Muestreo determinístico.
- ▶ Muestreo aleatorio.

- **Torneo:**

- ▶ Con substitución.
- ▶ Sin substitución.



# Reproducción proporcional o Método de la ruleta

Los individuos son seleccionados con una probabilidad proporcional a su valor relativo de fitness con respecto al total de la población.

La probabilidad de selección  $p_i$  de un individuo en la generación  $t$  está dada por la siguiente expresión:

$$p_i(t) = \frac{f_i(t)}{\sum_{j=1}^n f_j(t)} \quad (1)$$

Donde  $f_i(t)$  es el valor de fitness del individuo  $i$  en el tiempo  $t$ .

# Ejemplo

Individuo	Fitness	Proporción relativa
A	8	7 %
B	12	11 %
C	27	24 %
D	4	3 %
E	45	40 %
F	17	15 %
Total	113	100 %

Cuadro: Valores de fitness individuales

# Ejemplo

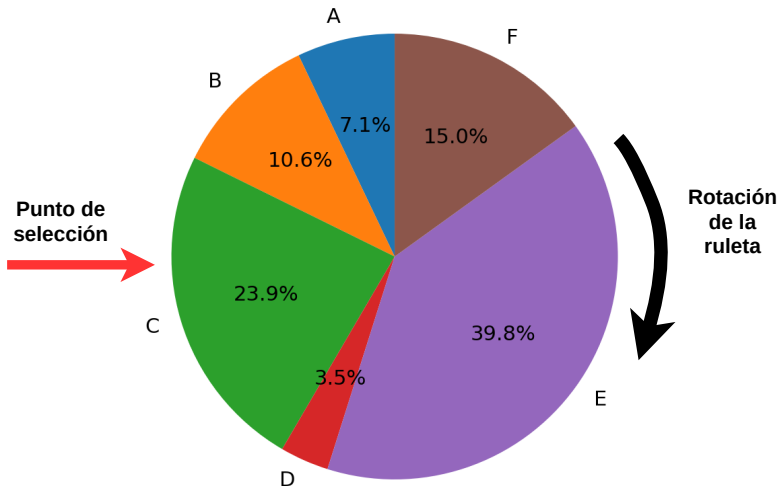


Figura: Método de selección de ruleta

# Muestreo estocástico universal

- Se utiliza la misma ruleta así como las mismas proporciones de los individuos, pero en vez de usar un sólo punto de selección, la ruleta se gira una sola vez y se realizan múltiples puntos de selección que se encuentran igualmente espaciados en la ruleta.
- Los individuos son escogidos al mismo tiempo.
- Este método previene que individuos con valores altos de fitness acaparen la selección para la reproducción de la siguiente generación al ser escogidos una y otra vez.

# Muestreo estocástico universal

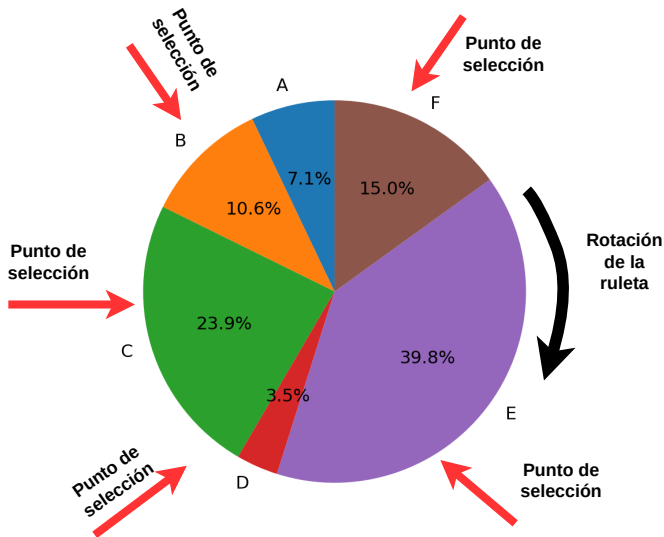


Figura: Muestreo estocástico universal

# Fitness scaling

Fitness scaling aplica una transformación de escalamiento a los valores originales de fitness y los reemplaza con la transformación resultante.

Dicha transformación mapea los valores originales dentro de un rango deseado mediante la siguiente expresión:

$$\text{scaled fitness} = a \times \text{valor original de fitness} + b \quad (2)$$

Donde  $a$  y  $b$  son constantes que seleccionamos para lograr el rango deseado de los valores de fitness escalados.

## Ejemplo

Utilizando los mismos valores de fitness, los mapearemos a un nuevo rango entre 50 y 100.

Calculamos los valores de  $a$  y  $b$  tomando a los individuos con el valor de fitness más bajo y más alto, respectivamente.

- $50 = a \times 4 + b$
- $100 = a \times 45 + b$

Resolviendo el sistema de ecuaciones obtenemos los siguientes valores para  $a = 1,22$  y  $b = 45,12$ .

De manera que los valores escalados se pueden calcular con la siguiente expresión:

$$\text{scaled fitness} = 1,22 \times \text{valor original de fitness} + 45,12$$

# Ejemplo

Individuo	Fitness	Scaled fitness	Proporción relativa
A	8	50	13 %
B	12	60	15 %
C	27	78	19 %
D	4	50	12 %
E	45	100	25 %
F	17	66	16 %
Total	113	404	100 %

Cuadro: Valores de fitness individuales



# Fitness scaling

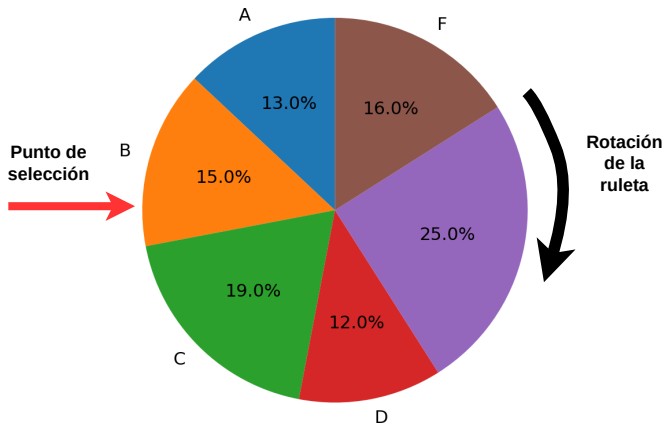


Figura: Fitness scaling

# Selección por torneo

- En cada ronda se toman aleatoriamente dos o más individuos de la población, y aquel individuo con la puntaje de fitness más alto gana la contienda y es seleccionado.
- El número de individuos que participan en cada ronda del torneo se denomina *tamaño del torneo*.
- En la medida que el tamaño del torneo es grande, hay más oportunidades que el mejor individuo participe en las contiendas y las gane, mientras que para un tamaño de torneo pequeño, hay más oportunidades que sean seleccionados los individuos con menor fitness.

# Ejemplo

Individuo	Fitness
A	8
B	12
C	27
D	4
E	45
F	17

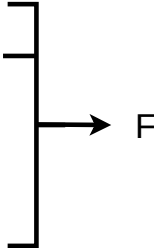


Figura: Selección por torneo con tamaño de torneo igual a 3

# Cruza o Recombinación

- Es un proceso que genera una recombinación de los alelos mediante el intercambio de segmentos entre pares de cromosomas.
- El operador de cruza es el operador principal en algoritmos genéticos.
- Es usado para combinar la información genética de dos individuos para generar (usualmente dos) nuevos individuos.
- Típicamente el operador de cruza se aplica con una probabilidad relativamente alta, de forma análoga a la vida natural ( $\in [0,6,1]$ ).
- Si **no** se aplica el operador de cruza, ambos individuos seleccionados son clonados directamente para la siguiente generación.

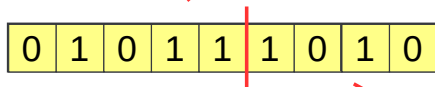
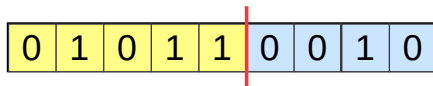
# Cruza o Recombinación

El operador de craza puede ejecutarse en dos pasos:

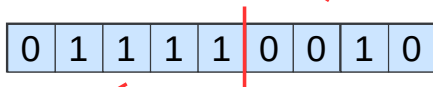
- Se seleccionan los individuos con algún método de selección.
- Los individuos se combinan con alguno de los siguientes métodos;
  - ▶ Cruza en un punto.
  - ▶ Cruza en dos puntos.
  - ▶ Cruza multipunto.
  - ▶ Cruza uniforme.

# Cruza en un punto

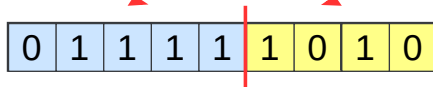
Seleccionamos un punto de corte aleatorio para los cromosomas de ambos individuos



Punto de corte

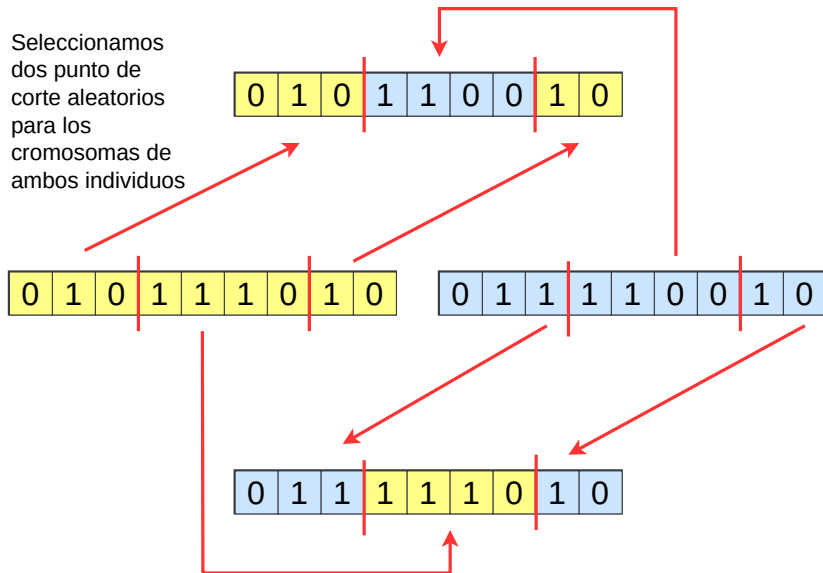


Punto de corte



# Cruza en dos punto

Seleccionamos dos punto de corte aleatorios para los cromosomas de ambos individuos



# Mutación

- La mutación es el último operador genético en el proceso de creación de la nueva población.
- La mutación puede interpretarse de varias maneras:
  - ▶ Un mecanismo para asegurar diversidad en la población.
  - ▶ Una protección contra pérdidas prematuras de material genético generado durante la cruce.
  - ▶ Permite explorar el espacio de búsqueda, escapar de óptimos locales y evitar la convergencia prematura.
- La mutación es un proceso donde un alelo de un gen es reemplazado aleatoriamente para generar un nuevo cromosoma.
- Al igual que en la vida natural, la mutación aparece con una probabilidad baja.



# Mutación

Se ha propuesto que la probabilidad de mutación esté en función de la longitud del cromosoma,  $L$ , de acuerdo a la siguiente expresión:

$$\text{Probabilidad de mutación} = 0,1 \times \left(\frac{1}{L}\right)$$

Es decir, en promedio, se estará mutando 1 % de la población en un bit.

# Flip bit mutation

Se selecciona un gen de forma aleatoria y se reemplaza su valor.

0	1	0	1	1	1	0	1	0
---	---	---	---	---	---	---	---	---



0	1	1	1	1	1	0	1	0
---	---	---	---	---	---	---	---	---

Podemos extenderlo a seleccionar y reemplazar más de un gen.

# Swap mutation

Se seleccionan dos genes de forma aleatoria y se intercambian sus valores.

0	1	0	1	1	1	0	1	0
---	---	---	---	---	---	---	---	---



0	1	1	1	1	1	0	0	0
---	---	---	---	---	---	---	---	---

# References I

La Rocca, M. (2022). *Advanced Algorithms and Data Structures*.  
Manning, first edition.