

**Assunto: Atividades Acadêmicas – 2023-02**

**Belo Horizonte, 07 de Setembro de 2023.**

**PLANEJAMENTO: Roteiro e Código para o Primeiro Trabalho**

**DISCIPLINA - DIG ELE037 – OTIMIZACAO**

|                    | Atividade   |
|--------------------|---|
| <b>Avaliação 1</b> | - Primeiro Trabalho Computacional. Entrega - Valor 20 pontos<br><a href="#">Entrega : 29/09</a> |

**Repositório Otimização - Graduação 2023.1:**

<https://drive.google.com/drive/folders/1Asd2jIJ7WcgWds3VjJ4S-dEWInj2IPYN?usp=sharing>

**Roteiro e código para o primeiro trabalho**

**Material do trabalho**

**Pasta - Exercicio-Computacional-I.zip**

**Data de Entrega da Avaliação I - [29/09/2023](#)**

**O trabalho impresso deverá ser entregue na data de 29/09/2023, de forma a garantir a nota total.**

Prof. Rodney Rezende Saldanha

Departamento de Engenharia Elétrica

Universidade Federal de Minas Gerais

# Estudo de uma ferramenta da inteligência computacional - Computação Evolutiva - Algoritmo Genético

## 1- Introdução

O objetivo geral do trabalho é fazer uso da Inteligência Computacional, por meio do estudo de um Algoritmo Genético (GA).

## 2- Metodologia

A *IEEE Computational Intelligence Society* define seus temas de interesse como redes neurais artificiais, sistemas nebulosos (*fuzzy*) e computação evolutiva, incluindo inteligência coletiva (*swarm intelligence*). A Figura 1 ilustra como estes temas podem estar relacionados de forma a conceber-se sistemas híbridos.

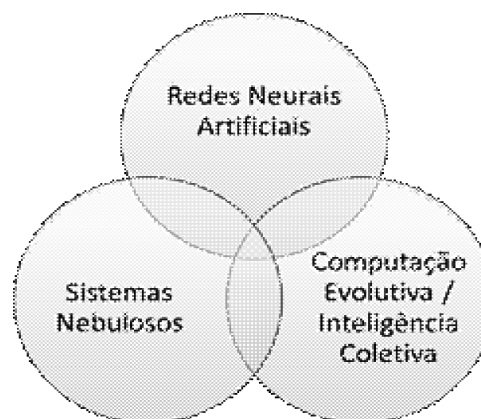


Figura 1 - Temas da Inteligência Computacional relacionados no diagrama de Venn

Para alcançar os objetivos específicos do trabalho, algumas atividades deverão ser realizadas durante o estudo, como:

- Estudar os parâmetros de um Algoritmo Genético aplicado à solução de problemas de otimização;

- Analizar as respostas obtidas através do uso do GA escolhido;
- Ajustar um Algoritmo Genético para solução de um problema proposto.

### 3 -Algoritmos Genéticos

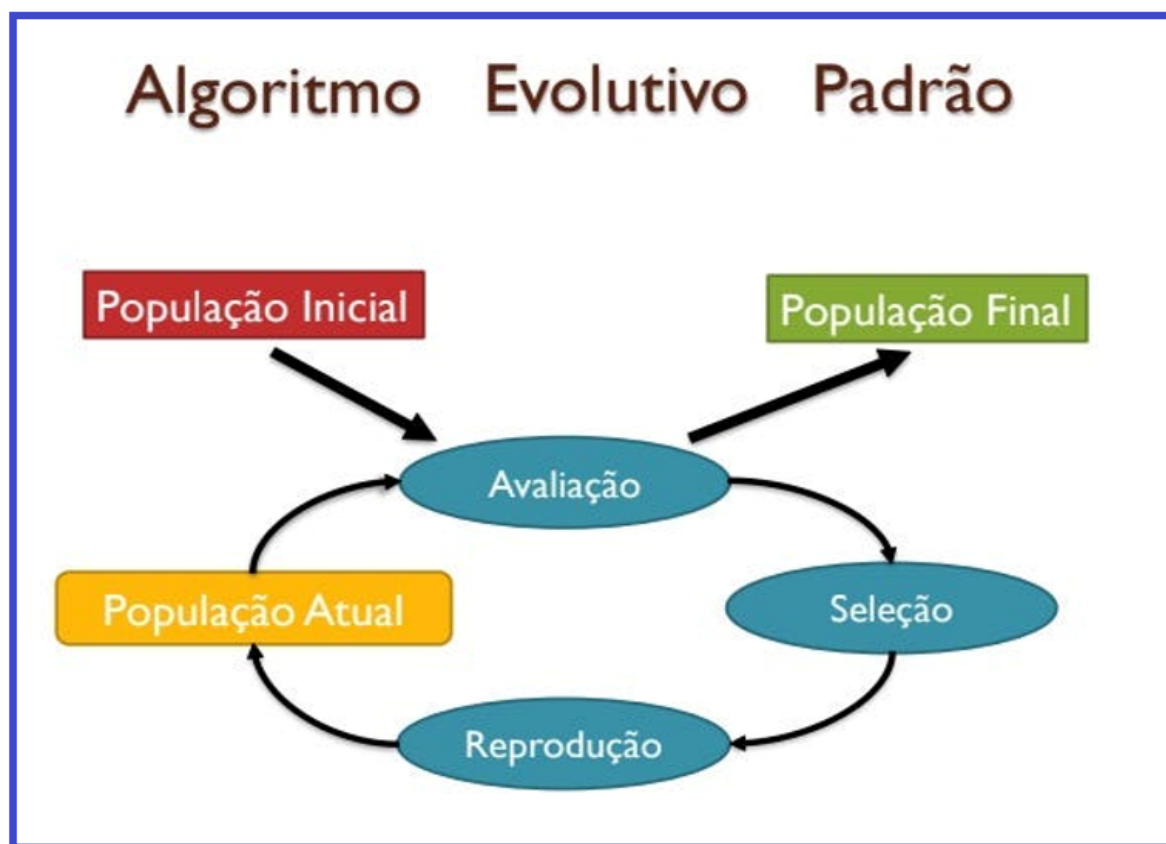
Os Algoritmos Genéticos (GA) são uma importante técnica heurística, baseada na teoria evolutiva, para a otimização de problemas em geral. O primeiro GA foi desenvolvido por John Holland na Universidade de Michigan (1975) e, envolve uma população inicial de soluções possíveis para um problema que valendo-se de operadores genéticos e gerações sucessivas, produz indivíduos otimizados. Por ser um método que espelha-se na Genética Natural, prevalecendo os indivíduos (soluções) mais adaptados ao ambiente (função *fitness*), o GA herdou alguns termos da biologia. A Tabela 1 lista a analogia do vocabulário desses termos.

**A Tabela 1 - Analogia do vocabulário do GA.**

| Genética Natural  | Algoritmo Genético   |
|---|--|
| Cromossomos   | <i>Strings</i>   |
| Genótipo (conjunto de cromossomos de um indivíduo)                              | Estruturas (conjuntos de <i>Strings</i> )  |
| Fenótipo (corpo formado pela interação de carga genética para o meio ambiente)  | Descodificar a estrutura de um conjunto de parâmetros que compõem uma solução particular |
| Genes (elementos que compõem os cromossomos são compostos de sequências de DNA) | Parte de uma sequência codificando um determinado parâmetro                              |
| Alelos (proteína que compreende o DNA)  | Cada um dos bits que constituem a estrutura de cadeia                                    |
| Ambiente  | Função objetivo ou aptidão ( <i>Fitness</i> )  |

|            |         |
|------------|---------|
| Indivíduos | Solução |
|------------|---------|

Os princípios básicos do GA estão mostrados na Figura 2, onde os detalhes das etapas do GA são descritos a seguir.



**Figura 2 – Fluxograma Algoritmo Genético Padrão.**

### 3.1 - População inicial

Os indivíduos que representam as soluções para um problema específico, no GA, podem ser descritos por meio de cadeias binárias que consistem de zeros e uns.

O GA é principiado com a criação de uma população inicial de

'*tam\_pop*' indivíduos, cada um representado por uma *string* (cadeia binária) contendo uma sequência aleatória de '*tam\_crom*' de zeros e uns (Figura 3).

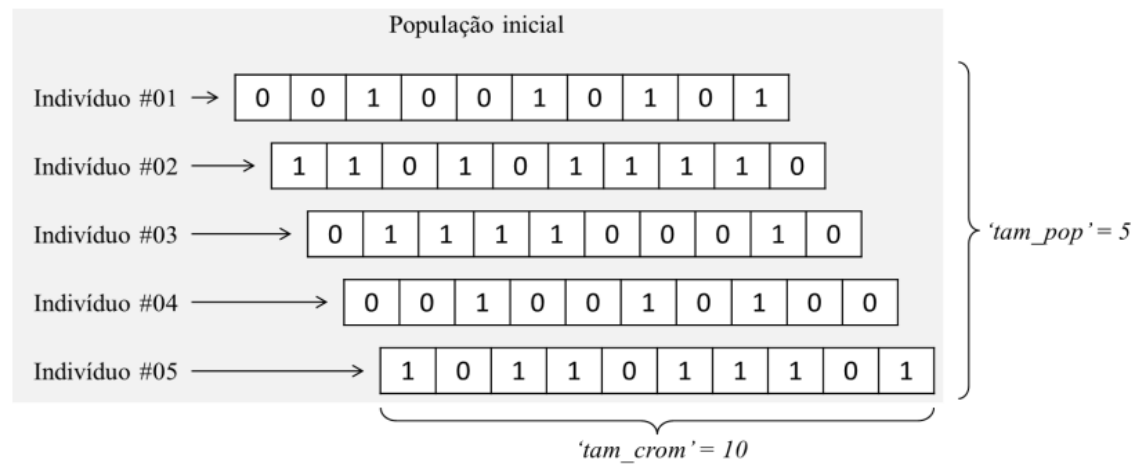


Figura 3 - Representação da população binária

Na Figura 3, os 05 (cinco) indivíduos da população inicial estão representados por uma *string* de 10 (dez) dígitos binários (*bits*) de comprimento.

### 3.2-Decodificação

Aplicando os pesos ilustrados na Figura 3 em cada um dos bits, o valor decimal inteiro que o Indivíduo #01 representa é 149.

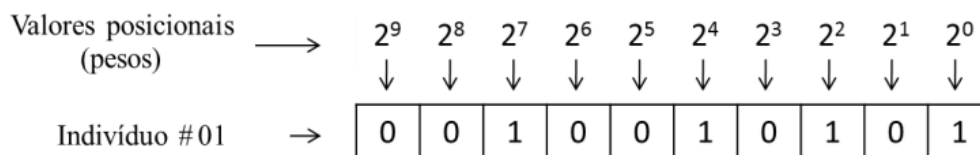


Figura 4 - Exemplo dos valores posicionais para cálculo do valor de  $X_b$ .

A *string* #01 representa  $X_b = 149$ ,  $X_b$  é o inteiro que corresponde ao binário,  $n$  é o número de bits do cromossomo e  $X_R$  é o real no intervalo  $[X_{min}, X_{max}]$ .  
 $X_{min} = 0$  e  $X_{max} = 10$ ,

$$X_R = X_b \cdot \frac{|X_{min} - X_{max}|}{2^n - 1} + X_{min}$$

$$X_R = 149 \cdot \frac{|0 - 10|}{2^{10} - 1} + 0 = 1,4565$$

Na Figura 3, os 05 (cinco) indivíduos da população inicial estão representados por *strings* de 10 (dez) dígitos binários (*bits*) de comprimento. Desta forma, os indivíduos podem ser igualmente decodificados e seus valores são expressos na Tabela 2.

Tabela 2 - Valores inteiro e decodificado de cada indivíduo (cromossomo binário)

| Indivíduo | Cromossomo binário | Valor inteiro ( $X_b$ ) | Valor de codificado ( $X_R$ ) |
|-----------|--------------------|-------------------------|-------------------------------|
| #01       | 0010010101         | 149                     | 1,4565                        |
| #02       | 1101011110         | 862                     | 8,4261                        |
| #03       | 0111100010         | 482                     | 4,7116                        |
| #04       | 0010010100         | 148                     | 1,4467                        |

|     |            |     |        |
|-----|------------|-----|--------|
| #05 | 1011011101 | 733 | 7,1652 |
|-----|------------|-----|--------|

O número mínimo de bits  $k$  para uma determinada precisão de  $p$  casas decimais pode ser obtido a partir da equação:

$$k \geq \log_2 \left( (X_{\max} - X_{\min}) \cdot 10^p \right) .$$

O valor inteiro de  $k$  deve ser tomado como o tamanho do cromossomo ‘tam\_crom’, de forma que o GA opere com  $p$  casas decimais de precisão. Por exemplo, para uma precisão de 3 casas decimais:

$$k \geq \log_2 \left( (10 - 0) \cdot 10^3 \right) = \log_2 (10^4) = 13,2877$$

$$n = \text{ceil} ( k ) = \text{ceil} ( 13,2877 ) = 14$$

**Obs:** (ceil( $k$ ) rounds the elements of  $k$  to the nearest integers towards infinity) . É necessário um cromossomo com 14 *bits* para que, no intervalo  $[0,10]$ , os valores decodificados apresentem precisão de 3 casas decimais.


### 3.3-Avaliação

Funções de avaliação ou aptidão são específicas para cada problema. A avaliação pode ser tratada de forma análoga ao meio ambiente onde os seres vivos são postos a provar sua existência e perpetuar suas respectivas espécies. No GA, ela determina a probabilidade de cada solução (indivíduo)

ser selecionada para reproduzir e perpetuar suas informações para as geração (iterações) futuras.

### 3.3.1 - Seleção (Aptidão = *fitness* )

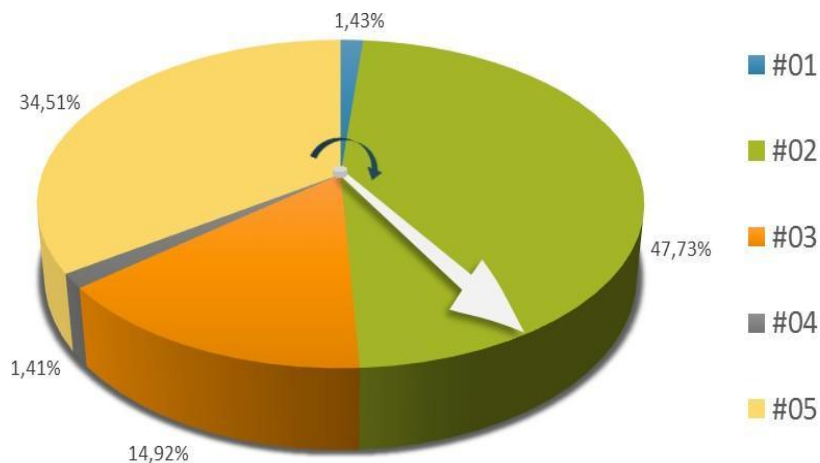
O processo de seleção do algoritmo genético seleciona indivíduos para reprodução de acordo com a aptidão. Indivíduos mais aptos (com maior *fitness*) têm maior probabilidade de serem selecionados para a reprodução. Existem vários métodos de seleção nos GAs, destacam-se os seguintes:

 Seleção por roleta proporcional (*Roulette wheel selection*):

Método mais tradicional, fundamentada em uma busca linear através de uma “roleta” com as fatias da roda ponderados na proporção de valores de aptidão relativa de cada indivíduo. De tal modo que, se  $f_i$  é o valor da *fitness* do indivíduo  $i$  na população corrente, a probabilidade do indivíduo ser selecionado  $p_i$  é dada por:

$$p_i = \frac{f_i}{\sum_{i=1}^N f_i} \quad (N \text{ número de indivíduos na população})$$





**Figura 5 - Representação da roleta**

Onde  $N$  é o número de indivíduos na população . A Figura 5 representa a roleta para uma fitness  $f(x) = x^2$  (quadrado do valor decodificado), conforme dados da Tabela 2.

**Tabela 3 - Valores de *fitness* e probabilidade de seleção de cada indivíduo (cromossomo binário) para seleção do tipo roleta**

| Indivíduo | Cromossomo binário | Valor decodificado ( $XR$ ) | Valor da fitness ( $f_i = XR^2$ ) | Valor da probabilidade ( $p_i$ ) |
|-----------|--------------------|-----------------------------|-----------------------------------|----------------------------------|
| #01       | 0010010101         | 1,4565                      | 2,12                              | 1,43 %                           |
| #02       | 1101011110         | 8,4261                      | 70,99                             | 47,73 %                          |
| #03       | 0111100010         | 4,7116                      | 22,19                             | 14,92 %                          |
| #04       | 0010010100         | 1,4467                      | 2,09                              | 1,41 %                           |
| #05       | 1011011101         | 7,1652                      | 51,34                             | 34,51 %                          |
|           |                    |                             | 148,75                            | 100,00 %                         |

### Seleção randômica (*Random selection*):

Neste tipo de seleção os indivíduos são selecionados aleatoriamente, desprezando o valor da *fitness* dos mesmo, tornando o GA mais rápido, porém, na maioria das vezes ineficaz.

### Seleção Classificatória (*Rank Selection*):

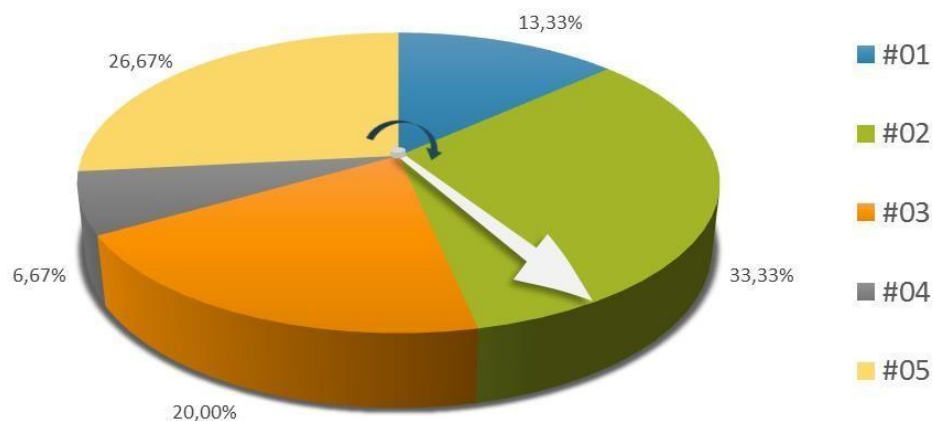
Este método é utilizado quando há grandes diferenças entre os valores de *fitness*. Por exemplo, se a maior probabilidade  $p_i$  entre os indivíduos é 90% da soma de todas as adequações, então haverá soluções com chances muito baixas de serem selecionados.

Por exemplo, o pior terá *fitness* igual a 1, o segundo pior 2 e assim por diante, de forma que o melhor terá *fitness* igual a N (número de indivíduos na população). A Tabela 4 apresenta os novos valores de *fitness*  $f_i'$  para cada indivíduo ranqueados conforme método da seleção classificatória.

**Tabela 4 - Novos valores de *fitness* e probabilidade de seleção de cada indivíduo (cromossomo binário) para seleção do tipo classificatória**

| Indivíduo | Cromossomo binário | Valor da <i>fitness</i><br>( $f_i = X_R^2$ ) | Novo valor da <i>fitness</i><br>( $f_i'$ ) | Novo valor da probabilidade<br>( $p_i'$ ) |
|-----------|--------------------|--|--|---|
| #01       | 0010010101         | 2,12 %                                       | 2  | 13,33 %                                   |
| #02       | 1101011110         | 70,99 %                                      | 5  | 33,33 %                                   |
| #03       | 0111100010         | 22,19 %                                      | 3  | 20,00 %                                   |
| #04       | 0010010100         | 2,09 %                                       | 1  | 6,67 %                                    |
| #05       | 1011011101         | 51,34 %                                      | 4  | 26,67 %                                   |
|           |                    |  | 15   | 100,00 %                                  |

A partir desses novos valores de probabilidade  $pi'$  é possível montar uma nova distribuição de probabilidade e verificar a diferença entre esta e a definida pelo método da roleta. A Figura 7 apresenta como a “nova roleta” estaria distribuída utilizando o método da seleção classificatória.



**Figura 6 - Situação da “nova roleta” depois da Classificação**

#### **Seleção por Torneio (ou Campeonato):**

Neste tipo de seleção os indivíduos são escolhidos, dois a dois (ou em pequenos grupos), aleatoriamente, e os valores de *fitness* são comparados. O indivíduo com maior *fitness* entre os dois (ou no grupo) vence o “torneio” e é assumido como cromossomo “pai” que irá gerar novos indivíduos nos passos seguintes.

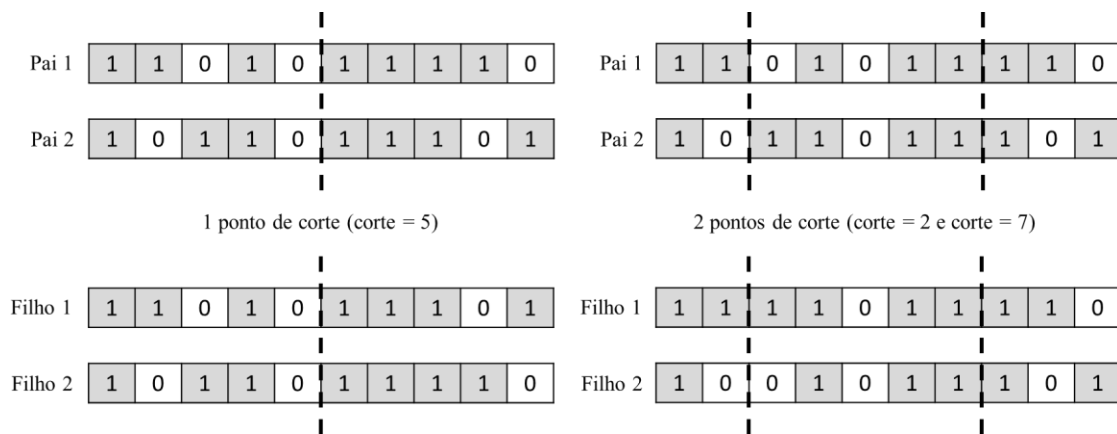
Após o processo de seleção, por qualquer um dos métodos supracitados, os indivíduos escolhidos são enviados para o **operador de cruzamento** e para o **operador de mutação** genética para obter nova geração.

### 3.4 - Cruzamento

De acordo com o fluxograma da Figura 2, o cruzamento será ou não realizado dependendo do valor de sua taxa. Quando o valor sorteado aleatoriamente for inferior à taxa probabilística de cruzamento indica que os cromossomos dos pais são divididos em duas partes. O filho (nova sequência) recebe a combinação das partes dos dois pais.

Os valores mais usuais na literatura ficam em torno de 70% a 90%.

A Figura 7 ilustra como pode acontecer o cruzamento de 02 (dois) pais selecionados (indivíduos #02 e #05) de forma a gerar descendentes (filhos).



**Figura 7 - Cruzamento com (a) um ponto de corte e (b) dois pontos de corte**

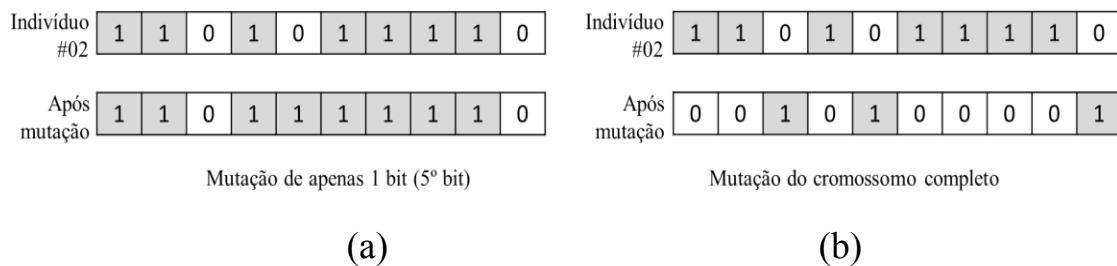
Para apenas um ponto de corte e com a posição definida de forma aleatória, é sorteado um valor entre 0 e '*tam\_crom*' (o tamanho do cromossomo).

### 3.5 - Mutação

Após o cruzamento, pode-se também fazer a mutação dos cromossomos de maneira a alterar completamente as características, inserindo diversidade à população. Geralmente, na mutação cada '0' do cromossomo é substituído por '1' e vice-versa.

Este operador genético é necessário porque insere diversidade à população e reduz as possibilidades de criar uma população de indivíduos iguais, em outras palavras, evita mínimos locais.

Os valores mais usuais na literatura ficam em torno de 1 a 2%. A Figura 8 ilustra como pode acontecer a mutação dos cromossomos.



**Figura 8 - Mutação do Indivíduo #02 (a) em apenas um bit e (b) no cromossomo completo**

Conforme Figura 8, a mutação pode acontecer em 01 (um) ou mais bits (no 5º bit, por exemplo) ou pode acontecer no cromossomo completo.

### **3.6 – Elitismo**

Um grupo com os melhores indivíduos (elite) são selecionados de acordo com a taxa percentual de elitismo, levando-os para a próxima geração juntamente aos novos indivíduos que complementam a população. Por exemplo, uma taxa de elitismo de 75% em uma população com 100 indivíduos indica que serão selecionados os 75 melhores e levados para a geração seguinte juntamente a outros novos 25 criados randomicamente.

Portanto, é muito importante a correta definição de parâmetros como: função de aptidão (*fitness*); população; mecanismo de seleção dos pais; operadores de variação e reprodução; as taxas de recombinação e mutação. Assim, é possível ter um GA suficientemente capaz de desenvolver um adequado mecanismo de busca.

## **4 – Parte Experimental**

### **4.1- Estudo e o Desenvolvimento de uma Ferramenta Computacional Baseada no GA.**

A proposta do trabalho é o estudo e o desenvolvimento de uma ferramenta computacional baseada no GA. A ferramenta deverá permitir ao usuário criar um problema de otimização (ou utilizar um existente) e conceber a sua solução com o uso de um Algoritmo Genético.

## 4.2 – Exemplos de Funções Objetivo

As funções podem ser escolhida pelo usuário, de forma a adaptarem-se aos problemas de maximização ou de minimização .

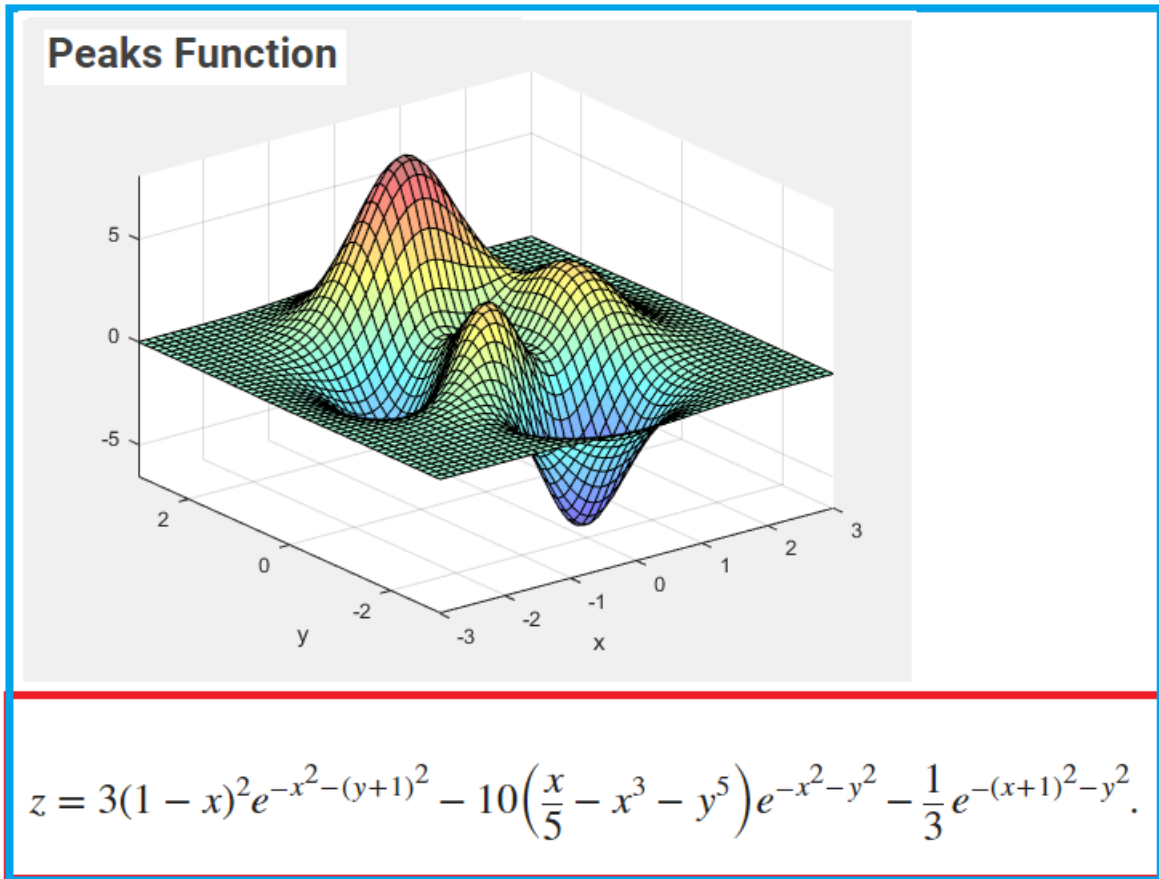
### 4.2.1 – Exemplo I : Função Picos – (Peaks Function)

Nesta etapa, o usuário define os parâmetros para o problema estudado. Estes parâmetros consistem nos limites inferior e superior de cada uma das variáveis.

**Tabela 5 - Parâmetros do Problema.**

| Variáveis de Projeto | Xmin | Xmax |
|----------------------|------|------|
| x                    | -3   | 3    |
| y                    | -3   | 3    |

Estes parâmetros escolhidos definem o domínio de busca do GA. O limite inferior e superior representam  $Xmin$  e  $Xmax$ .



**Figura 9 – Função Picos**

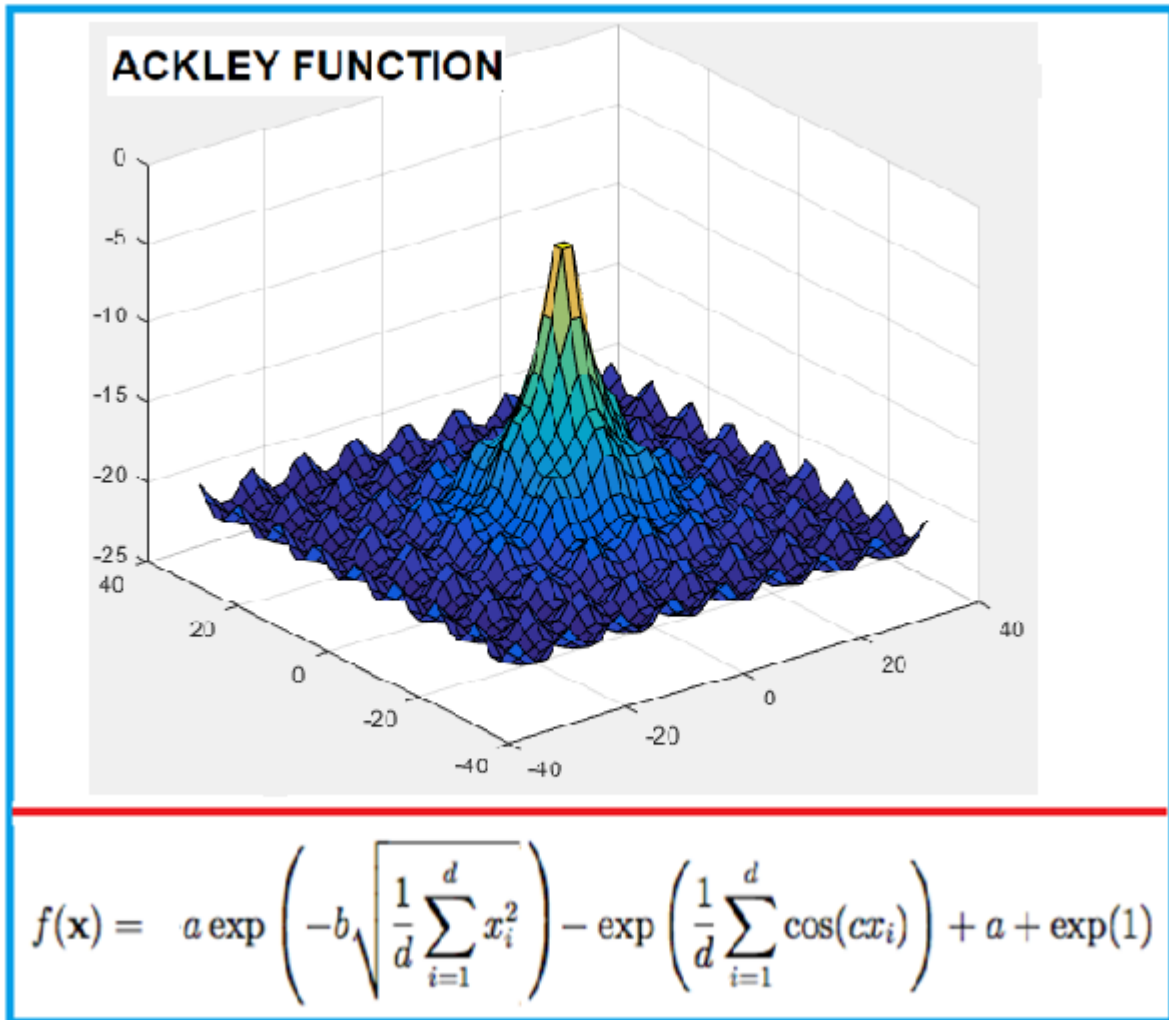
#### 4.2.2 Example II: Função Ackley

Os limites inferior e superior estão representados por Xmin e Xmax. Os parâmetros definem o domínio de busca para o GA.

**Tabela 6 - Parâmetros do Problema: Função Ackley**

| Variáveis de Projeto | Xmin | Xmax |
|----------------------|------|------|
| x                    | -35  | 35   |
| y                    | -35  | 35   |





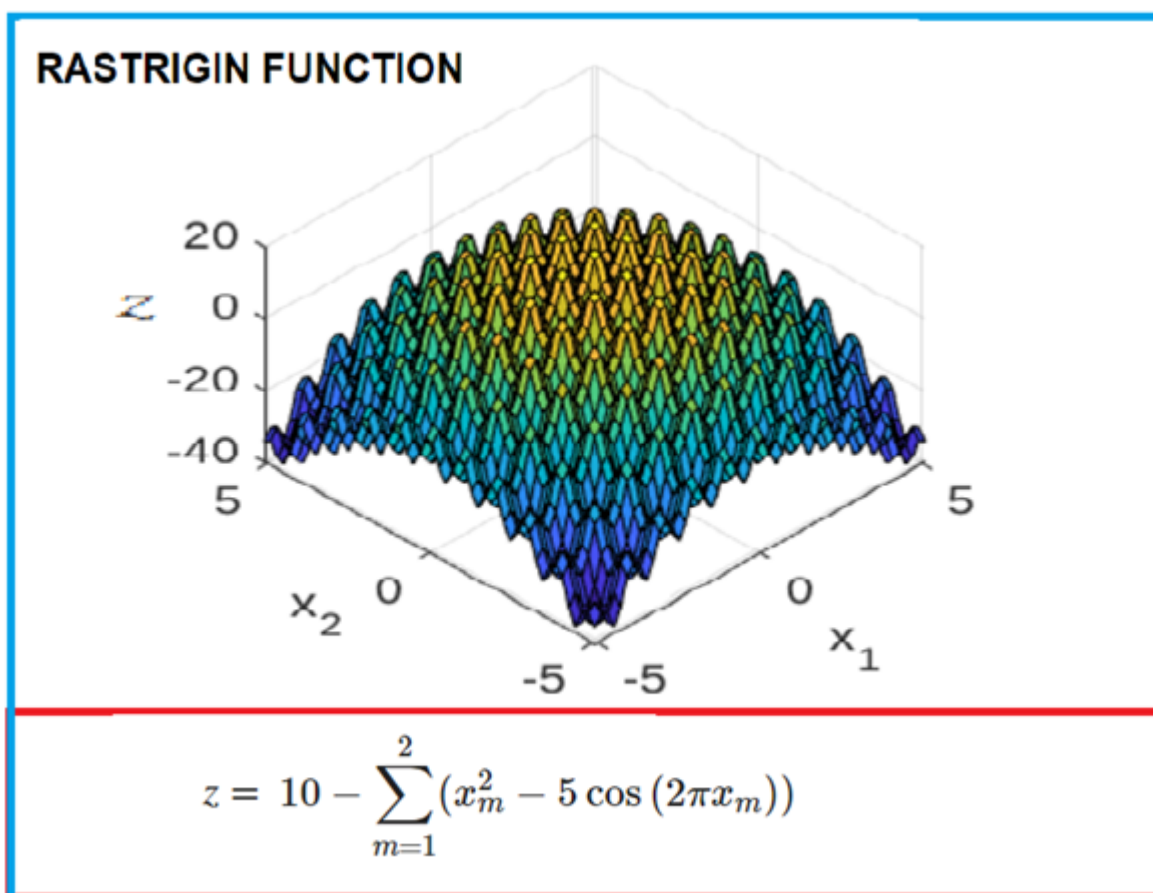
**Figura 10 – Função Ackley**

### **4.2.3 Example III: Função Rastrigin – (Rastrigin Function)**

Os limites inferior e superior estão representados por Xmin e Xmax. Os parâmetros definem o domínio de busca para o GA.

**Tabela 7 - Parâmetros do Problema - Função Rastrigin.**

| Variáveis de Projeto | Xmin  | Xmax |
|----------------------|-------|------|
| x                    | -5.12 | 5.12 |
| y                    | -5.12 | 5.12 |



**Figura 11 – Função Rastrigin**

### 4.3 – Variação dos Parâmetros do Algoritmo Genético

Nesta etapa, o usuário deve definir os parâmetros do GA. Na Tabela 5, ao lado de cada um desses parâmetros, foram inseridos faixas de valores para auxiliar o usuário nos estudos a serem feitos e no preenchimento afim

de levar o algoritmo à convergir para um resultado satisfatório.

Os parâmetros a serem estudados do GA são:

**Tabela 8 - Parâmetros do GA para solucionar o Problema.**

| <b>Parâmetro</b>          | <b>Valor / Opção</b>   |
|---------------------------|------------------------|
| Tamanho da população      | 10 - 100               |
| Tamanho do cromossomo     | 10 - 35                |
| Número máximo de gerações | 10 - 50                |
| Taxa de cruzamento        | 60 - 80 %              |
| Taxa de mutação           | 1 - 5 %                |
| Taxa de elitismo          | 55 - 75 %              |
| Função <i>Fitness</i>     | Minimizar ou Maximizar |

Cada um desses valores, conforme apresentados inicialmente têm um significado específico e influenciam de forma particular nos resultados. Na coluna de valores da Tabela 8 são apresentadas faixas de variação com sugestão de valor conforme a literatura.

No trabalho, os usuários devem estudar a variação dos parâmetros abaixo, dentro de suas respectivas faixas:

|                           |          |
|---------------------------|----------|
| Tamanho da população      | 10 - 100 |
| Tamanho do cromossomo     | 10 - 35  |
| Número máximo de gerações | 10 - 50  |

|                    |           |
|--------------------|-----------|
| Taxa de cruzamento | 60 - 80 % |
| Taxa de mutação    | 1 - 5 %   |
| Taxa de elitismo   | 55 - 75 % |

## 4.4 – Estudo dos Critérios de Parada

Existem duas principais formas para o término da execução de um GA. A **primeira** condição é em relação às características dos indivíduos que compõem a solução do problema. Quando é possível identificar um padrão ótimo em relação aos indivíduos da população, não existindo mais a necessidade de se continuar executando o GA. A **segunda** condição de parada de um GA acontece quando não se sabe identificar um padrão ótimo dos indivíduos. Fatores que podem fazer com que um GA termine sua execução:

- Tempo máximo de execução do algoritmo ou número máximo de gerações é excedido;
- Número total de avaliações da função fitness é alcançado;
- Melhorias em indivíduos feitas através de operadores genéticos e seleção já alcançaram um certo limite, não havendo mais mudanças.

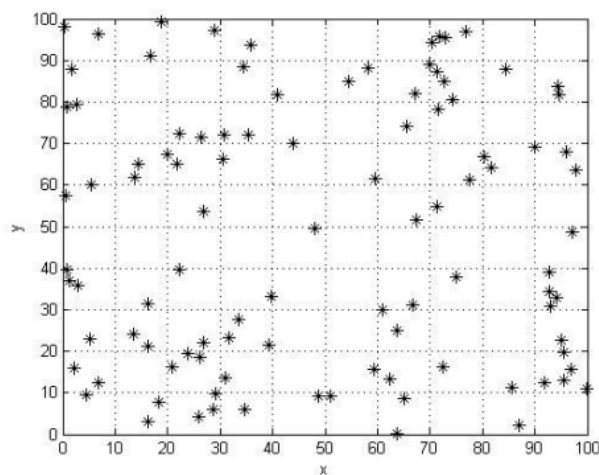
Geralmente, um GA tem sua execução terminada quando se satisfaz uma das formas descritas anteriormente: quando um certo valor ótimo (ou satisfatório) é alcançado pelos indivíduos ou quando uma condição de parada é satisfeita.

## 4.4 – Resultados, Discussões e Conclusões

Os resultados de cada um dos problemas estudados deverão ser relatados e analisados e discutidos pelo usuário. Os experimentos computacionais devem ser obtidos com as variações paramétricas do GA.

## 5 – Desafio I – Problema do Caixeiro Viajante

Considere um conjunto de  $N$  cidades cujas coordenadas  $(x,y)$  estão definidas no arquivo `cidades.txt`.



**Figura 12 – Ploter do Arquivo `cidades.txt`**

- a) Deseja-se encontrar o percurso de menor comprimento total que passe por todas as cidades exatamente uma única vez. Escolha uma representação adequada para as soluções candidatas e uma função de avaliação (fitness) e para a codificação adotada implemente um algoritmo genético (GA) para a resolução do problema. Indique quais as escolhas feitas referentes aos operadores (e.g., seleção, cruzamento e

mutação) e seus respectivos parâmetros (e.g., tamanho da população e critério de parada). No exercício devem ser estudados os parâmetros: Tamanho da população; Tamanho do cromossomo; Número máximo de gerações; Taxa de cruzamento; Taxa de mutação; Taxa de elitismo.

- b) Execute o algoritmo por 10 vezes, gerando, em cada caso, uma população inicial aleatória. Apresente a melhor solução obtida pelo algoritmo em cada caso e compute o valor médio e o desvio padrão do grupo formado pelos fitness das soluções finais. Comente.
- c) Apresente, para uma execução do algoritmo, as curvas de fitness médio, fitness do melhor e fitness do pior indivíduo em função do número de gerações. Faça uma breve descrição/análise do algoritmo em termos da eficiência de busca e da manutenção de diversidade.

## **7 - Referências Bibliográficas**

Goldberg, D.; David Edward, G.; Goldberg, D. e Goldberg, V. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning. Artificial Intelligence. Addison-Wesley Publishing Company.

Eiben, A. E. e Smith, J. E. (2007). Introduction to Evolutionary Computing. Springer Verlag.

Diaz-Gomez, P. A. e Hougen, D. F. (2007). Initial population for genetic algorithms: A metric approach. In Arabnia, H. R.; Yang, J. Y. e Yang, M. Q., editores, Proceedings of the 2007 International Conference on Genetic and Evolutionary Methods, GEM 2007, June 25-28, 2007, Las Vegas, Nevada, USA, pp. 43–49. CSREA Press.

Meadows, B.; Riddle, P.; Skinner, C. e Barley, M. M. (2013). Evaluating the seeding genetic algorithm. In Cranefield, S. e Nayak, A., editores, AI 2013: Advances in Artificial Intelligence, pp. 221–227, Cham. Springer International Publishing.

Emmanuel Ramon Marques Dantas; Dissertação Mestrado; “Desenvolvimento de uma toolbox para aplicação de inteligência computacional em sistemas de controle clássico”; Universidade Federal Rural do Semi-Árido, PPGSCA, Mossoró, 27/12/2013.

GABRIEL, P. H. R.; DELBEM, A. C. B. Fundamentos de Algoritmos Evolutivos. Notas Didáticas do ICMC-USP. São Carlos/SP: ICMC-USP, v. 75, 2008. 35 p. ISBN 0103-2585.