

Algoritmos Genéticos: a otimização aplicando a teoria da evolução

Sezimária F. Pereira Saramago

Faculdade de Matemática
Universidade Federal de Uberlândia
saramago@ufu.br

Resumo. Este artigo apresenta um texto introdutório sobre algoritmos genéticos. O objetivo é dar a fundamentação teórica necessária para que um estudante de graduação possa compreender o método e utilizá-lo como uma ferramenta eficaz de otimização. Os principais operadores são introduzidos através do desenvolvimento de um exemplo matemático, permitindo que, de uma forma simples e didática, o estudante possa acompanhar o desenvolvimento dos mesmos.

Palavras-Chave. otimização, métodos evolutivos, algoritmo genético.

1. Introdução

A otimização é aplicada em situações em que se deseja maximizar ou minimizar uma função numérica de várias variáveis, num contexto em que podem existir restrições. Tanto as funções acima mencionadas como as restrições dependem dos valores assumidos pelas variáveis de projeto ao longo do procedimento de otimização, geralmente são funções não-lineares e acopladas entre si.

Pode-se utilizar otimização em várias áreas, como por exemplo no projeto de sistemas ou componentes, planejamento e análise de operações, problemas de otimização de estruturas, otimização de forma, controle de sistemas dinâmicos. A grande vantagem é determinar a melhor configuração de projeto sem ter que testar todas as possibilidades envolvidas.

A otimização tem como vantagens diminuir o tempo dedicado ao projeto, possibilitar o tratamento simultâneo de uma grande quantidade de variáveis e restrições de difícil visualização gráfica e/ou tabular, possibilitar a obtenção de algo melhor, obtenção de soluções não tradicionais, menor custo.

Como limitações tem-se o aumento do tempo computacional quando aumenta-se o número de variáveis de projeto, pode-se surgir funções descontínuas que apresentam lenta convergência, funções com presença de muitos mínimos locais onde o mínimo global raramente é obtido.

As técnicas clássicas de otimização são conhecidas à bem mais de um século, sendo utilizadas na física e na geometria, servindo-se de ferramentas associadas às equações diferenciais ao Cálculo Variacional. A sofisticação dos recursos computacionais, desenvolvidos nos últimos anos, tem motivado um grande avanço nas técnicas de otimização. Aliado ao fato de que os problemas tornam-se cada vez mais complexos.

Técnicas clássicas de otimização são confiáveis e possuem aplicações nos mais diferentes campos de engenharia e de outras ciências. Porém, estas técnicas podem apresentar algumas dificuldades numéricas e problemas de robustez relacionados com: a falta de continuidade das funções a serem otimizadas ou de suas restrições, funções não convexas, multimodalidade, existência de ruídos nas funções, necessidade de se trabalhar com valores discretos para as variáveis, existência de mínimos ou máximos locais, etc. Assim, os estudos de métodos heurísticos, com busca randômica controlada por critérios probabilísticos, reaparecem como uma forte tendência nos últimos anos, principalmente devido ao avanço dos recursos computacionais, pois um fator limitante destes métodos é a necessidade de um número elevado de avaliações da função objetivo (Schwefel e Taylor, 1994).

Os algoritmos genéticos são mecanismos de busca baseados nos processos de seleção natural da luta pela vida e da genética de populações. Trata-se de um método pseudo-aleatório, portanto pode-se dizer que é um método, um procedimento de exploração inteligente, no espaço de parâmetros codificados (Braga, 1998).

O surgimento dos algoritmos genéticos deu-se por volta de 1950 quando vários biólogos usavam técnicas computacionais para a simulação de sistemas biológicos. Entre 1960 e 1970, na Universidade de Michigan, sob a direção de John Holland (1975), iniciou-se o estudo de algoritmos genéticos como os conhecidos atualmente. David Goldberg (1989) apresentou a solução de complexos problemas de engenharia usando algoritmos genéticos, o que ajudou o método a se tornar popular entre os pesquisadores.

2. Problema Geral de Otimização

O problema geral de otimização consiste em minimizar uma função objetivo, sujeita, ou não, a restrições de igualdade, desigualdade e restrições laterais. A função objetivo e as funções de restrições podem ser funções lineares ou não lineares em relação às variáveis de projeto, implícitas ou explícitas, calculadas por técnicas analíticas ou numéricas. Seja o problema geral de otimização dado por:

Minimizar :

$$F(X), \quad X = [X_1, X_2, \dots, X_n]^T, \quad X \in \mathfrak{R}^n \quad (1)$$

Sujeito a:

$$\begin{aligned} g_j(X) &\geq 0, \quad j=1,2,\dots,J \\ h_k(X) &= 0, \quad k=1,2,\dots,K \\ X_i^{(L)} &\leq X_i \leq X_i^{(U)}, \quad i=1,2,\dots,n \end{aligned} \quad (2)$$

onde, $F(X)$ representa a função objetivo, g_j e h_k as restrições de desigualdade e de igualdade, $X_i^{(L)}$ e $X_i^{(U)}$ as restrições laterais. Todas estas funções assumem valores em \mathfrak{R}^n e são, na maioria dos casos, não-lineares.

Os métodos para a solução de problemas de otimização podem ser divididos em dois grandes grupos, os métodos baseados no cálculo (Deterministic Optimization) e os métodos aleatórios (Random Strategies). Quanto à presença de limitantes ao problema, tem-se a otimização irrestrita e a otimização restrita. No grupo dos métodos aleatórios encontra-se os métodos de ordem zero (métodos tradicionais), Algoritmos Genéticos, Simulated Annealing, Evolução Diferencial, Tabu Search, etc.

2.1. Revisão sobre Técnicas Seqüenciais

A maioria dos algoritmos seqüenciais de otimização requer um conjunto inicial de variáveis de projeto X^0 (Vanderplaats, 1984). A partir daí, o projeto é atualizado iterativamente:

$$X^q = X^{q-1} + \mathbf{a}^* S^q \quad (3)$$

onde q representa o número da iteração, S o vetor direção de busca no espaço de projeto, \mathbf{a}^* é o escalar que define o passo que se deseja dar na direção de S . Os algoritmos de otimização não-linear, baseados no cálculo, necessitam da determinação da direção de busca S e do parâmetro escalar \mathbf{a}^* .

Raramente pode-se garantir a existência e unicidade de um projeto ótimo, isto ocorre devido a existência de várias soluções, mal condicionamento numérico ou lenta convergência. A estratégia prática utilizada é inicializar o processo de otimização à partir de diferentes configurações de X^0 , caso se encontre o mesmo valor para F_{\min} , tem-se alguma garantia de mínimo global.

Considerando problemas sem restrição, para que $F(x)$ seja mínimo, uma condição necessária, mas não suficiente, é que $\nabla F(x) = 0$. Para $F(x)$, função de uma variável, no ponto de mínimo a 2ª derivada deve ser positiva.

Para o caso de uma função de várias variáveis, a matriz Hessiana H deve ser positiva definida, o que implica dizer que todos os autovalores de H devem ser positivos.

2.2. Métodos de Primeira Ordem (Método da Descida Máxima)

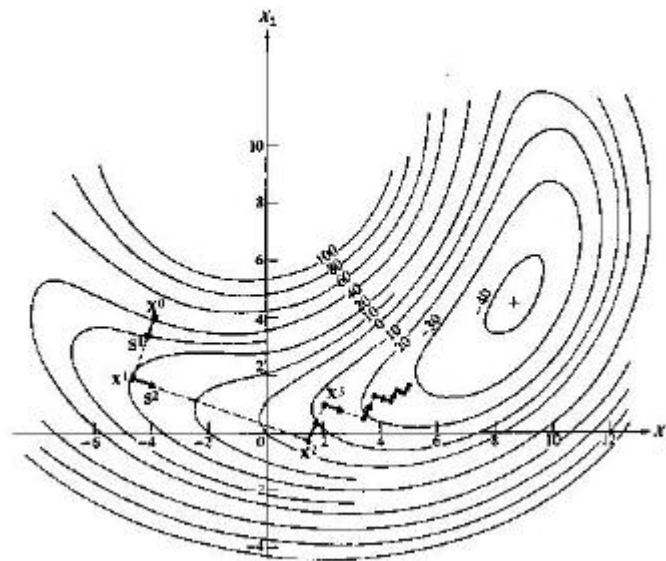


Figura 1. Representação do Método da Descida Máxima

A direção de busca é dada pelo gradiente $\nabla F(X)$. Usando $\nabla F(X)$ limita-se a direção de busca, evitando a busca em todo espaço de projeto (Saramago e Steffen, 1996). O gradiente deve ser recalculado a cada nova direção, conforme representado na Fig. 1. Sua importância é permitir estabelecer o ponto inicial para métodos mais sofisticados. Toma-se como direção de busca aquela oposta ao gradiente:

$$S^q = -\nabla F(X^q) \text{ , onde } X^{q+1} = X^q + a_q^* S^q \quad (4)$$

A cada passo, determina-se a^* que ao mínimo nesta direção, busca unidimensional. Nestes métodos, a convergência é boa no início, mas muito lenta ao se aproximar do mínimo

3. Métodos de Ordem Zero

São métodos simples, de fácil implementação, confiáveis e capazes de trabalhar com valores discretos. Requerem apenas o cálculo de $F(X)$, não dependem do gradiente e da continuidade da função. Necessitam de um grande número de avaliações da função objetivo, o que aumenta o custo computacional.

A idéia básica é selecionar um grande número de vetores de projeto X e calcular $F(X)$ correspondente a cada um. O vetor correspondente ao menor valor de $F(X)$ será adotado como o valor ótimo X^* . O vetor X é selecionado de forma randômica no espaço de projeto. Para limitar a busca, utiliza-se as restrições laterais. Um número randômico r é gerado, $r \in [0, 1]$, e as variáveis de projeto da q -ésima iteração atualizadas:

$$X_i^q = X_i^l + r(X_i^u - X_i^l) \quad (5)$$

O processo do método de Ordem Zero está apresentado no fluxograma da Fig. 2. Neste caso, o critério de parada adotado é o número máximo de iterações. Porém, outros critérios podem ser incorporados ao programa.

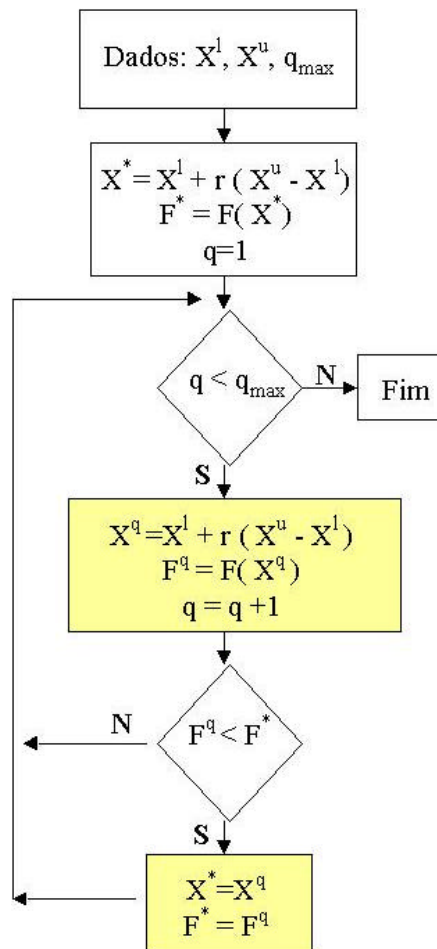


Figura 2. Esquema do Método de Ordem Zero

4. Algoritmos Genéticos

Os algoritmos genéticos usam um vocabulário emprestado da genética natural. Fala-se sobre indivíduos (genótipos) de uma população. Estes indivíduos também são chamados de cromossomos (Michalewicz, 1998).

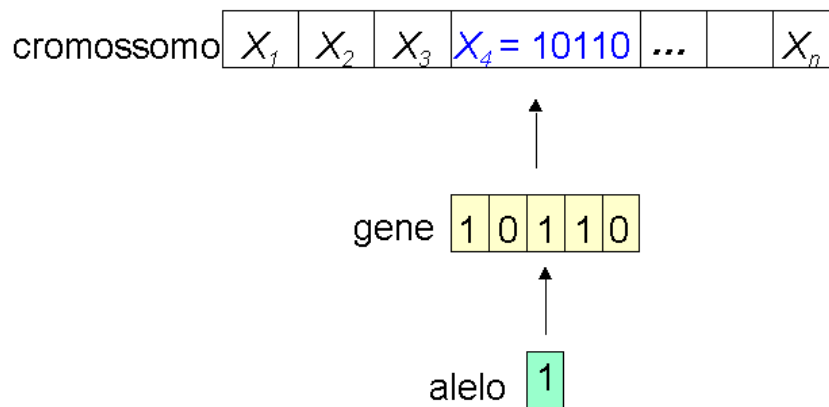


Figura 3. Esquemática de um cromossomo.

Cromossomos são compostos de unidades ou elementos, cada elemento equivale a um gene, dispostos em uma sequência linear, conforme exemplificado na Fig. 3.

Sendo n o número de variáveis (genes) e a cada gene tem o comprimento m . Assim, uma função de duas variáveis $f(x, y)$, $n = 2$, será representada através de um cromossomo com 2 genes. Seja $m = 7$, o número de alelos de cada gene. Neste caso, tem-se:

$$\text{Cromossomo: } \left[\underbrace{1100001}_x \quad \underbrace{0011101}_y \right] \quad (6)$$

Algoritmos genéticos são algoritmos iterativos, e a cada iteração a população é modificada, usando as melhores características dos elementos da geração anterior e submetendo-as a três tipos básicos de operadores, para produzir melhores resultados. São eles:

Reprodução: é um processo no qual cada cadeia é copiada levando em conta os valores da função de adaptação f .

Cruzamento: é um processo no qual a combinação em partes de cada um de dois cromossomos gera um novo descendente.

Mutação: é a modificação aleatória ocasional (de baixa probabilidade) do valor de um alelo da cadeia.

O primeiro passo para a aplicação de algoritmos genéticos a um problema qualquer é encontrar alguma representação cromossômica conveniente, cujo gene represente o espaço de busca do problema, com vetores binários de zeros e um (0,1), os quais são gerados aleatoriamente. O comprimento m do gene depende da precisão requerida para o problema. Temos na Fig. 4 um fluxograma do algoritmo genético (Haupt, 1998).

Com finalidade de ilustrar a aplicação dos operadores, vamos considerar o problema de minimização de uma função dada por:

$$g(x, y) = x \sin(4x) + 1,1 y \sin(2y) \quad (7)$$

no intervalo, $8 < x < 10$, $8 < y < 10$, que define a região viável do problema.

A maioria dos códigos computacionais para algoritmos genéticos costuma maximizar a função, portanto, a função objetiva em estudo será reescrita como:

$$\max g(x, y) = - [x \sin(4x) + 1,1 y \sin(2y)]. \quad (8)$$

Para este exemplo, será adotada a precisão de duas casas decimais. Como o espaço de busca, ou seja, o domínio da função tem amplitude $I = 10 - 8 = 2$ e considerando uma precisão de duas casas

decimais, este intervalo deve ser dividido em $l \times 10^m$ subintervalos iguais, neste caso, $2 \times 10^2 = 200$ pontos. Portanto a sequência binária (cada gene) deverá ter pelo menos 8 bits, pois $128 = 2^7 < 200 < 2^8 = 256$.

Seja a seguinte população inicial, obtida aleatoriamente:

C₁ - 10000101 00100111
 C₂ - 00001110 00001001
 C₃ - 10010001 00000001
 C₄ - 11000101 00101001
 C₅ - 01111100 10101100
 C₆ - 11100010 01001010

Tem-se assim, a população inicial de cromossomos definida, onde cada gene é um vetor binário de m bits, sendo m função da precisão exigida (10^{-2}) e da amplitude do intervalo definido pelas restrições laterais ($l = 2$). A seguir, todos esses indivíduos serão modificados, submetendo-os aos operadores genéticos.

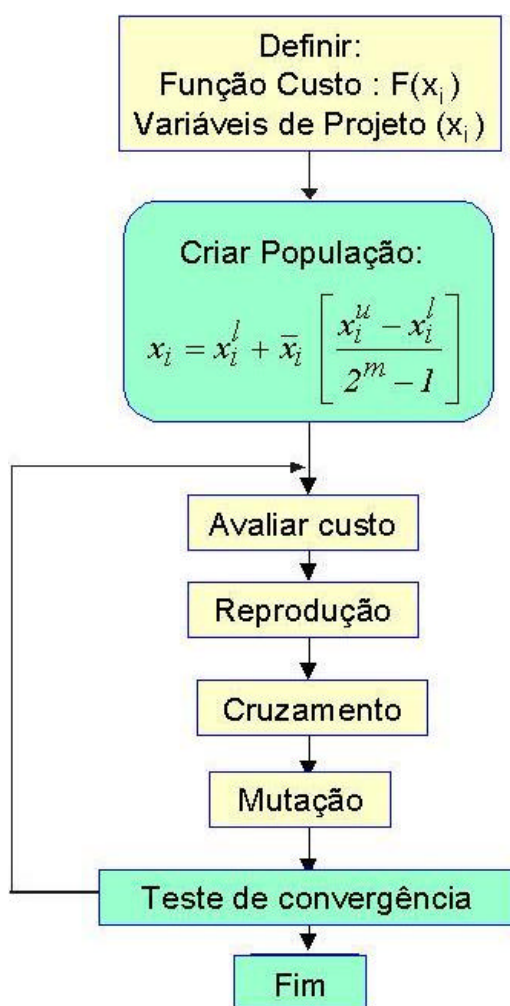


Figura 4. Fluxograma do algoritmo genético contínuo.

4.1 Reprodução

Reprodução é um processo que será atribuído às cadeias que possuem o maior valor objetivo e, portanto uma probabilidade mais elevada de contribuir à geração seguinte, criando pelo menos um descendente. Quanto maior o valor da função objetivo, maiores são as chances do indivíduo sobreviver no ambiente e reproduzir-se passando parte de seu material genético a gerações posteriores (Braga, 1998).

Usando a probabilidade, expressa pela Eq.(9), tem-se que se o indivíduo for de baixa adequabilidade, tem alta probabilidade de desaparecer da população, caso contrário, os indivíduos terão grandes chances de permanecer na população.

$$P_i = \frac{f(x)}{F(x)}, \text{ sendo } F(x) = \sum f(x) \quad (9)$$

Para se calcular o valor da função de adaptação f , deve-se converter a seqüência binária (base 2) para base 10, ou seja, deve-se decodificar um cromossomo, conforme Eq. (10).

$$C = [b_7 b_6 \dots b_2 b_1 b_0 \ a_7 a_6 \dots a_2 a_1 a_0]$$

$$\bar{x} = \sum_{i=0}^{m-1} b_i \times 2^i \quad e \quad \bar{y} = \sum_{i=0}^{m-1} a_i \times 2^i \quad (10)$$

Feito isso, deve-se calcular o valor de x e y reais, dentro da região viável, através da seguinte equação:

$$\begin{cases} x = a_i + decimal(100 \dots 010)_2 \times \frac{b_i - a_i}{2^m - 1} \\ y = a_i + decimal(010 \dots 100)_2 \times \frac{b_i - a_i}{2^m - 1} \end{cases} \quad (11)$$

sendo:

a_i e b_i - domínio das variáveis x e y .
 m - comprimento total do gene.

$decimal(100 \dots 010)_2 = \bar{x}$

$decimal(010 \dots 100)_2 = \bar{y}$

Como a população inicial já está definida, o próximo passo será o cálculo da função objetivo (adaptação).

A título de ilustração, será mostrado o cálculo da função objetivo do primeiro cromossomo da população criada para o Exemplo (7) em estudo.

Seja $C_1 = 1000010100100111$

Passando para a base 10, utilizando as Eqs. (10) e (11) tem-se:

$$\bar{x} = \sum_{i=0}^{8-1} b_i \times 2^i = 1 \times 2^0 + 0 \times 2^1 + 1 \times 2^2 + 0 \times 2^3 + 0 \times 2^4 + 0 \times 2^5 + 0 \times 2^6 + 1 \times 2^7 = 133$$

$$\bar{y} = \sum_{i=0}^7 a_i \times 2^i = 1 \times 2^0 + 1 \times 2^1 + 1 \times 2^2 + 0 \times 2^3 + 0 \times 2^4 + 1 \times 2^5 + 0 \times 2^6 + 0 \times 2^7 = 39$$

Os valores reais x e y , dentro da região viável, são dados por:

$$x = 8 + \frac{133 \times (10 - 8)}{2^8 - 1} \Rightarrow x = 9,04$$

$$y = 8 + \frac{39 \times (10 - 8)}{2^8 - 1} \Rightarrow y = 8,31$$

e o valor da função de adaptação é:

$$g(x, y) = -[x \sin(4x) + 1,1y \sin(2y)]$$

$$g(x, y) = -[9,04 \sin(4 \cdot 9,04) + 1,1 \cdot 8,31 \sin(2 \cdot 8,31)]$$

$$g(x, y) = + 16,26$$

De forma análoga, obteve-se os resultados mostrados na Tab. 1, para cada cromossomo da população inicial.

Tabela 1. Cromossomos da população inicial.

Cromossomos	x	y	$g(x, y)$.
1000010100100111	9,04	8,31	16,26
0000111000001001	8,11	8,07	-3,21
1001000100000001	9,14	8,01	11,01
1100010100101001	9,55	8,32	2,76
0111110010101100	8,98	9,35	10,32
1110001001001010	9,77	8,58	- 0,22
$\sum g(x,y)$			36,92

Como citado anteriormente, a função de adaptação $g(x, y)$ é o árbitro final que decide sobre a vida ou a morte de cada cromossomo. O mecanismo para seleção das melhores cadeias, ou seja, as mais adaptadas, são definidas pelo uso das probabilidades proporcionais, dadas pela Eq. (9), resultando os seguintes valores:

$$p_1 = \frac{16,26}{36,92} = 0,44$$

$$p_2 = \frac{-3,21}{36,92} = -0,09$$

$$p_3 = \frac{11,01}{36,92} = 0,30$$

$$p_4 = \frac{2,76}{36,92} = 0,07$$

$$p_5 = \frac{10,32}{36,92} = 0,28$$

$$p_6 = \frac{-0,22}{36,92} = -0,00$$

sendo $p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1,00$

Considerando que as probabilidades acumulativas q_i cada cromossomo são dadas por:

$$q_i = \sum_{j=1}^i p_j \quad (12)$$

obtem-se aos seguintes valores acumulativos:

$$q_1 = p_1 = 0,44$$

$$q_2 = p_1 + p_2 = 0,44 - 0,09 = 0,35$$

$$q_3 = p_1 + p_2 + p_3 = 0,44 - 0,09 + 0,30 = 0,65$$

$$q_4 = p_1 + p_2 + p_3 + p_4 = 0,44 - 0,09 + 0,30 + 0,07 = 0,72$$

$$q_5 = p_1 + p_2 + p_3 + p_4 + p_5 = 0,44 - 0,09 + 0,30 + 0,07 + 0,28 = 1,00$$

$$q_6 = p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 0,44 - 0,09 + 0,30 + 0,07 + 0,28 + 0,00 = 1,00$$

A seguir deve-se selecionar as cadeias que irão contribuir para a geração seguinte. Esta seleção considera um conjunto de números r , escolhidos randomicamente entre $[0,1]$, em quantidade igual ao número de cadeias.

A análise é feita através das seguintes opções:

Se $r < q_1$, então se seleciona o 1º cromossomo C_1 .

Se $r > q_1$, então se passa para o subsequente e faz a análise novamente.

Vale ressaltar que alguns cromossomos poderão ser selecionados mais de uma vez, ou seja, os melhores serão copiados mais vezes, enquanto que outros irão morrer.

Seja o exemplo em estudo. Considere que foram gerados os seguintes números randômicos:

$r_1 = 0,64$; $r_2 = 0,08$; $r_3 = 0,47$
 $r_4 = 0,88$; $r_5 = 0,93$; $r_6 = 0,70$

A seleção dos cromossomos é dada por:

$r_1 = 0,64 > q_1 = 0,44 \Rightarrow r_1 = 0,64 > q_2 = 0,35 \Rightarrow r_1 = 0,64 < q_3 = 0,65 \quad \therefore \text{seleciona-se } q_3 \Rightarrow C_3$

$r_2 = 0,08 < q_1 = 0,44 \quad \therefore \text{seleciona-se } q_1 \Rightarrow C_1$

$r_3 = 0,47 > q_1 = 0,44 \Rightarrow r_3 = 0,47 > q_2 = 0,35 \Rightarrow r_3 = 0,47 < q_3 = 0,65 \quad \therefore \text{seleciona-se } q_3 \Rightarrow C_3$

$r_4 = 0,88 > q_1 = 0,44 \Rightarrow r_4 = 0,88 > q_2 = 0,35 \Rightarrow r_4 = 0,88 > q_3 = 0,65 \Rightarrow$
 $r_4 = 0,88 > q_4 = 0,72 \Rightarrow r_4 = 0,88 < q_5 = 1,00 \quad \therefore \text{seleciona-se } q_5 \Rightarrow C_5$

$r_5 = 0,93 > q_1 = 0,44 \Rightarrow r_5 = 0,93 > q_2 = 0,35 \Rightarrow r_5 = 0,93 > q_3 = 0,65 \Rightarrow$
 $r_5 = 0,93 > q_4 = 0,72 \Rightarrow r_5 = 0,93 < q_5 = 1,00 \quad \therefore \text{seleciona-se } q_5 \Rightarrow C_5$

$r_6 = 0,70 > q_1 = 0,44 \Rightarrow r_6 = 0,70 > q_2 = 0,35 \Rightarrow r_6 = 0,70 > q_3 = 0,65 \Rightarrow$
 $r_6 = 0,70 < q_4 = 0,72 \quad \therefore \text{seleciona-se } q_4 \Rightarrow C_4$

Depois de selecionados, os cromossomos dão origem a uma nova população representada como:

C_1' - 1001000100000001 \Rightarrow gerados de C_3 ; $g(x, y) = 11,01$.
 C_2' - 1000010100100111 \Rightarrow gerados de C_1 ; $g(x, y) = 16,26$.
 C_3' - 1001000100000001 \Rightarrow gerados de C_3 ; $g(x, y) = 11,01$.
 C_4' - 0111110010101100 \Rightarrow gerados de C_6 ; $g(x, y) = 10,32$.
 C_5' - 0111110010101100 \Rightarrow gerados de C_5 ; $g(x, y) = 10,32$.
 C_4' - 1100010100101001 \Rightarrow gerados de C_4 ; $g(x, y) = 2,76$.

4.2. Cruzamento

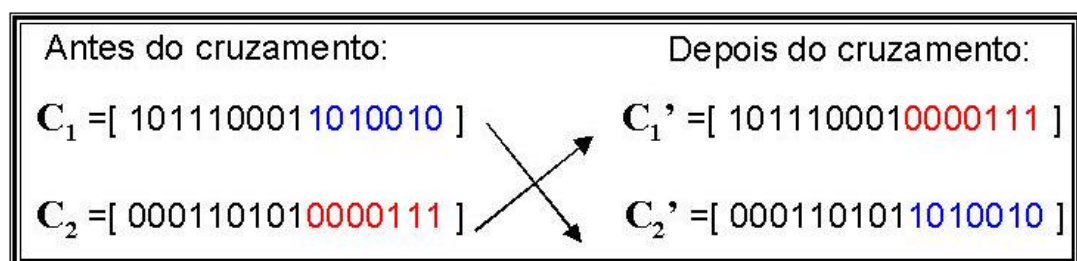


Figura 5. Representação do Operador Cruzamento

Tem-se várias formas para se obter o cruzamento, neste trabalho será utilizada a seguinte técnica para se fazer o cruzamento:

Seja o ponto k que define a posição de cruzamento na cadeia de bits de cada cromossomo escolhido aleatoriamente. A quantidade de cromossomos a ser submetida ao processo de cruzamento é definida através da probabilidade de cruzamento p_c , especificada pelo usuário. Cada cadeia é partida neste ponto k

e todas as informações do cromossomo A, a partir do ponto escolhido, são copiadas para o cromossomo B e vice-versa, conforme esquematizada na Fig. 5.

O processo de escolha de quem será cruzado deve ser feito em pares, sorteando números randômicos (r_i). Quando não for possível formar os pares um novo sorteio deverá ser feito até obter os pares necessários para o cruzamento. Por exemplo, se r_1 for menor que a probabilidade p_c , então o cromossomo C_1' será selecionado.

Na maioria das literaturas especializadas, a probabilidade de cruzamento é de $p_c = 25\%$, a qual será adotada neste trabalho.

Após de ter feito isso, temos que gerar um novo número randômico para determinar a posição k , onde duas novas cadeias são formadas pela troca de todos os caracteres compreendidos entre as posições $k + 1$ e o comprimento total do cromossomo ($n \times m$). Esta posição k é determinada pela seguinte fórmula:

$$k = 1 + \text{rand} [(n \times m - 1) - 1] \quad (13)$$

Dando continuidade ao Exemplo (7) em estudo, submetem-se as populações (C_i') ao cruzamento. Sejam os seguintes números randômicos, r_i :

$$\begin{array}{ll} r_1 - 0,50 & \Rightarrow C_1' > p_c \\ r_2 - 0,17 & \Rightarrow C_2' < p_c \\ r_3 - 0,40 & \Rightarrow C_3' > p_c \\ r_4 - 0,15 & \Rightarrow C_4' < p_c \\ r_5 - 0,20 & \Rightarrow C_5' < p_c \\ r_6 - 0,23 & \Rightarrow C_6' < p_c \end{array}$$

sendo selecionados para o cruzamento, os cromossomos C_2' e C_4' ; C_5' e C_6' .

Agora, é só gerar um número randômico para determinar k , a posição de cruzamento usando a Eq.(13). Seja $\text{rand} = 0,7$; segue-se que:

$$\begin{aligned} k &= 1 + 0,7 [(2 \times 8 - 1) - 1] = 1 + 0,7 (15 - 1) \\ k &= 1 + 0,7 \cdot 14 \quad \Rightarrow k = 10,8 \end{aligned}$$

Como k é um número inteiro, então $k = 11$. Daí,

C_2' - 1000010100100111
 C_4' - 0111110010101100

Trocando os caracteres, tem-se:

C_2'' - 1000010100101100
 C_4'' - 0111110010100111
e
 C_5' - 0111110010101100
 C_6' - 1100010100101001

Trocando os caracteres, tem-se:

C_5'' - 0111110010101001
 C_6'' - 1100010100101100

Assim, após a aplicação do operador cruzamento, a população é dada por:

C_1' - 1001000100000001; $g(x, y) = 11, 01$.
 C_2'' - 1000010100101100; $g(x, y) = 16, 72$.
 C_3' - 1001000100000001; $g(x, y) = 11, 01$.
 C_4'' - 0111110010100111; $g(x, y) = 11, 02$.
 C_5'' - 0111110010101001; $g(x, y) = 10, 67$.
 C_6'' - 1100010100101100; $g(x, y) = 3, 10$.

4.3 Mutação

A mutação é uma modificação aleatória do valor de um alelo da cadeia. Caso o alelo escolhido seja zero passa a ser um e vice-versa, conforme esquematizado na Fig. 6.

Segundo Haupt (1998), uma técnica de se fazer à mutação é gerar pares (a, b) randômicos onde a representa a linha e b representa a coluna da mudança do bit. Nesta forma de aplicar o operador mutação exclui-se da seleção o melhor cromossomo. No exemplo em estudo, sejam os pares (1, 10) e (5, 3), logo tem-se (o cromossomo C_2' não será objeto de mutação por apresentar o maior valor para a função objetiva):

(1, 10) \Rightarrow C_1' e posição 10
 1001000100000001 \Rightarrow 1001000101000001

(5, 3) \Rightarrow C_5'' e posição 3
 0111110010101001 \Rightarrow 0101110010101001

Neste artigo, será utilizada outra metodologia onde se seleciona randomicamente uma posição em um cromossomo, obedecendo a probabilidade de mutação p_m , e muda o valor deste bit.

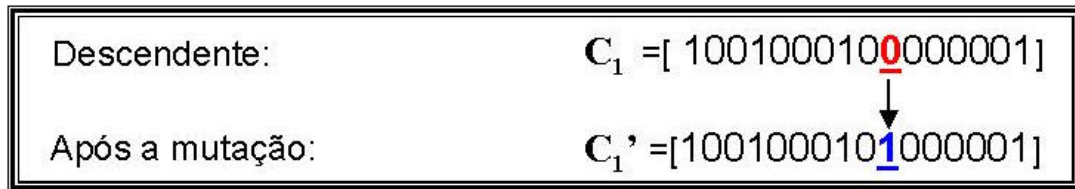


Figura 6. Representação do Operador Mutação.

O processo de mutação é controlado por um parâmetro fixo p_m , probabilidade de mutação, que é geralmente recomendado como de 1%. Este operador tem um papel importante e necessário, porque a reprodução e o cruzamento podem perder material genético potencialmente útil. O operador de mutação protege os algoritmos genéticos contra perdas irreparáveis. Tomada isoladamente, a mutação se constituiria na exploração aleatória do espaço das cadeias. Utilizada com cuidado, juntamente com os outros dois operadores, protege-se o procedimento da perda prematura de informações importantes (Braga, 1998).

Aplicando o operador mutação ao Exemplo (7) em estudo, torna-se necessário gerar $(n \times m \times N)$ números randômicos r entre $[0,1]$, onde N é o número total de indivíduos da população. Se r for menor que a probabilidade $p_m = 0,01$ será feita a mutação no bit correspondente. Considere que foram gerados 96 ($2 \times 8 \times 6$) números r entre 0 e 1 e que três tiveram probabilidades menores que p_m , como mostrdo a seguir:

$r_{13} = 0,009 < p_m = 0,01$
 $r_{39} = 0,0025 < p_m = 0,01$
 $r_{83} = 0,0004 < p_m = 0,01$

Considerando a população atual,

C_1' – 1001000100000001
 C_2'' – 1000010100101100
 C_3' – 1001000100000001
 C_4'' – 0111110010100111
 C_5'' – 0111110010101001
 C_6'' – 1100010100101100

torna-se possível selecionar a posição onde deve-se ocorrer a mutação, conforme Tab. 2.

Tabela 2. Seleção da posição para aplicação do operador mutação.

Posição	Cromossomo	Probabilidade (p_m)
13	C_1'	0,009
39	C_3'	0,0025
83	C_6''	0,0004

Submetendo os bits 13, 39 e 83 ao processo de mutação têm-se:

C_1' – 1001000100001001
 C_2'' – 1000010100101100
 C_3' – 1001001100000001
 C_4'' – 0111110010100111
 C_5'' – 0111110010101001
 C_6'' – 1110010100101100

Após a aplicação dos três operadores, tem-se encerrado o ciclo da 1ª geração. Assim, torna-se interessante observar os valores das funções de adaptação para se ter uma idéia de como está ocorrendo a evolução dos cromossomos da população inicial. Estes dados podem ser acompanhados na Tab. 3.

Tabela 3. Cromossomos da população após a 1ª geração.

Cromossomos	x	y	$g(x, y)$
C ₁ '- 1001000100001001	9,14	8,04	11,52
C ₂ "-1000010100101100	9,04	8,35	16,72
C ₃ '- 1001001100000001	9,15	8,00	10,68
C ₄ "- 0111110010100111	8,97	9,31	11,06
C ₅ "- 0111110010101001	8,97	9,33	10,63
C ₆ "- 1110010100101100	9,80	8,35	-2,09
$\sum g(x,y)$			58,52

Observando as Tabs. 2. e 4, nota-se que a população inicial melhorou no sentido de caminhar na direção da maximização da função objetiva, após aplicar os três operadores. Observa-se que o valor de $\sum g(x,y)$ passou de 36,92 para 58,52. Nesta primeira iteração, o ponto ótimo obtido corresponde a: $x= 9,04$ $y= 8,35$ e $f(x,y)= - 16,72$. Obviamente, executando outras iterações espera-se uma adaptação muito melhor da população.

5. Conclusões

Cada iteração do algoritmo genético, denominada geração, é composta pela execução dos três operadores descritos anteriormente. Após algumas gerações sobrevivem os indivíduos melhores adaptados, ou seja, são escolhidos os pontos do espaço de trabalho que correspondem ao maior valor da função objetivo. Desta forma, obtém-se a maximização, ou minimização, de uma função de várias variáveis, sem a exigência de continuidade da mesma ou de conhecimento de suas derivadas. Devido ao seu caráter puramente aleatório, esta metodologia possui a característica de não ficar "presa" em mínimos locais, sendo esta sua grande vantagem em relação aos métodos sequenciais de otimização. Sua desvantagem é, obviamente, o tempo computacional. Várias modificações já foram acrescentadas ao método, tornando-o mais eficiente e rápido. Por questões didáticas, este artigo apresentou apenas a formulação clássica de cada operador, visando dar um conhecimento básico ao leitor. Na bibliografia existem vários livros e artigos sobre o assunto, onde o tema é desenvolvido com maior profundidade.

6. Bibliografia

- Bazarra, M.S., Shetali, H.D. e Shetty, C.M., 1993, "Nonlinear Programming: Theory and Algorithms", John Wiley & Sons, second edition, New York.
- Braga, C.G., 1998, "O Uso de Algoritmos Genéticos para Aplicação em Problemas de Otimização de Sistemas Mecânicos", Dissertação de Mestrado, Universidade Federal de Uberlândia, Uberlândia, MG.
- Corana, A., Marchesi, M., Martini, C. e Ridella, S., 1987, "Minimizing multimodal functions of continuous variables with the simulated annealing algorithm", *ACM Transactions on Mathematical Software*, 13, n. 3, 262-280.
- Goldberg, D.E., 1989, "Genetic Algorithms in Search, Optimization and Machine Learning", New York, Addison- Wesley.
- Grace, A., 1992, "Optimization Toolbox- For Use with Matlab", The Math Works Inc., Natick.
- Haupt, R.L. e Haupt, S.E., 1998, "Practical Genetic Algorithm", John Wiley & Sons Inc; New York, pp.25-48.
- Holland, J.H., 1975, "Adaptation in Natural and Artificial Systems", Ann Arbor, The University of Michigan Press, United States of America.
- Lopes, L.C.G., Santos, E.A.A. e Karlec, V., 1999, "Determinação de raízes de equações não lineares utilizando algoritmos genéticos", *Computational Methods in Engineering (Cd room)*, pp. 141.1-141.11, XX CILAMCE, São Paulo.
- Masters, T., 1993, "Practical Neural Network Recipes In C++", Academic Press, pp. 117-134.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M.N. e Teller, A.H., 1953, "Equations of state calculations by fast computing machines", *Journal of Chemical Physics*, 21, 1087-1092.
- Michalewicz, Z., 1998, "Evolutionary Computation for NonLinear Programming Problems", [ftp.uncc.edu](http://ftp.uncc.edu/directory/coe/evol), directory coe/evol.

- Reklaitis, G.V., Ravindran, A. e K.M. Ragsdell, 1983, "Engineering Optimization - Methods and Applications", John Wiley and Sons, United States of America.
- Saramago, S.F.P. e Steffen Jr., V., 1996, "Aspectos fundamentais ao usar técnicas de otimização no projeto de sistemas mecânicos", Proc. IV Congr. de Eng. Mecânica- NNE, V.1, pp.421-426, Recife, Brasil, Junho.
- Schwefel, H.P. e Taylor, L., 1994, "Evolution and Optimum Seeking", John Wiley & Sons Inc, United States of America, pp. 87-88.
- Vanderplaats, G., 1984, "Numerical Optimization Techniques for Engineering Design", McGraw-Hill, United States of America, 70-97.

7. Direitos Autorais

Os autores são os únicos responsáveis pelo conteúdo do material impresso incluído no seu trabalho.

Genetic algorithms: the optimization applying the evolution theory

Sezimária F. Pereira Saramago

Faculdade de Matemática
Universidade Federal de Uberlândia
saramago@ufu.br

Abstract. *This article presents an introductory text on genetic algorithms. The objective is to give the necessary theoretical concepts so that a graduation student can understand the method and to use it as an effective tool of optimization. The main operators are introduced by using a mathematical example, allowing that, in a simple and didactic way, the student can accompany the development of the same ones.*

Keywords. *optimization, evolutionary methods, genetic algorithm.*