

Busca por Extremos de Funções Multidimensionais Utilizando Algoritmo Genético

Davi Valadares Rodrigues Feliciano

Universidade de Brasília - Instituto de Física

4 de maio de 2022

Motivação

Problemas de Otimização Numérica

- Ajuste de Funções → Método dos Mínimos Quadrados
- *Machine Learning* → *Fitting* de parâmetros em redes neurais, sistemas de classificação
- Indústria → Otimização de recursos e maximização de lucros
- Física → *Fitting* de curvas de potencial de moléculas diatômicas [3]



Figura: A antena evoluída, assim chamada por ter seu formato calculado por meio de um AG.

Motivação

Métodos para Otimização Numérica

- Baseados em Cálculo
 - Indiretos → Solução de $\nabla f = 0$
 - Diretos → Procuram extremos escalando na direção de ∇f . Conhecidos por escalada de morro, ou *hill climbing*
- Enumerativos → Iteração sobre os membros de um espaço de busca finito, ou um espaço discreto infinito
- Procura Aleatória → *Random walks*

Um Exemplo

Otimização numérica da função

$$f(x, y) = \cos(n\pi r) \exp\left(-\frac{r^2}{\sigma^2}\right)$$

- Sabemos que existem máximos em $2k/n$ com $k \in \mathbb{N}$
- O máximo global ocorre quando $k = 0$

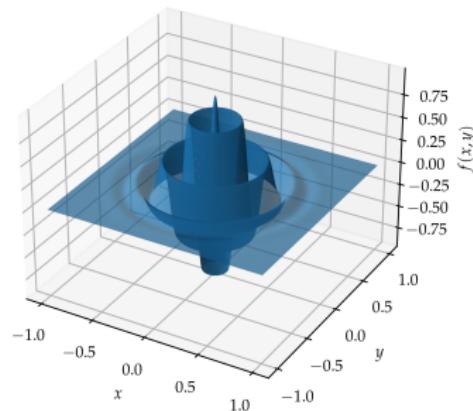


Figura: Gráfico de $f(x, y)$ com $n = 9$ e $\sigma = 0,4$

Um Exemplo

- Algoritmos do tipo *hill climbing* só teriam sucesso em $r < 1/n$, em $\approx 1\%$ das vezes, no caso ilustrado
- Os máximos locais atuariam como barreiras protegendo o máximo local

Como criar um algoritmo para otimização numérica que contorna essas dificuldades?

O Algoritmo Genético

Vantagens

- Não é limitado por restrições no espaço de busca — não depende da existência de derivada ou continuidade.
- A função objetivo pode ser aleatória ou mesmo depender do tempo
- É facilmente paralelizável
- São simples, do ponto de vista computacional
- Permite o mapeamento também dos máximos locais

O Algoritmo Genético

Em que consiste um AG?

- Codificação da informação de um elemento do espaço de busca em um número de cromossomos
- Um conjunto de cromossomos qualifica um indivíduo
- Esses cromossomos são compostos por um número de genes, que podem ter como valores os alelos 0 ou 1
- Um conjunto de indivíduos similares qualifica uma população

O Algoritmo Genético

Em que consiste um AG?

- Um mecanismo de seleção probabilístico escolhe um grupo de indivíduos
- A população evolui com a recombinação desses indivíduos
- Nesse processo, existe uma probabilidade de mutação nos genes
- A tendência será de que, após um longo período, os melhores indivíduos da população representem a solução do problema

O Algoritmo Genético

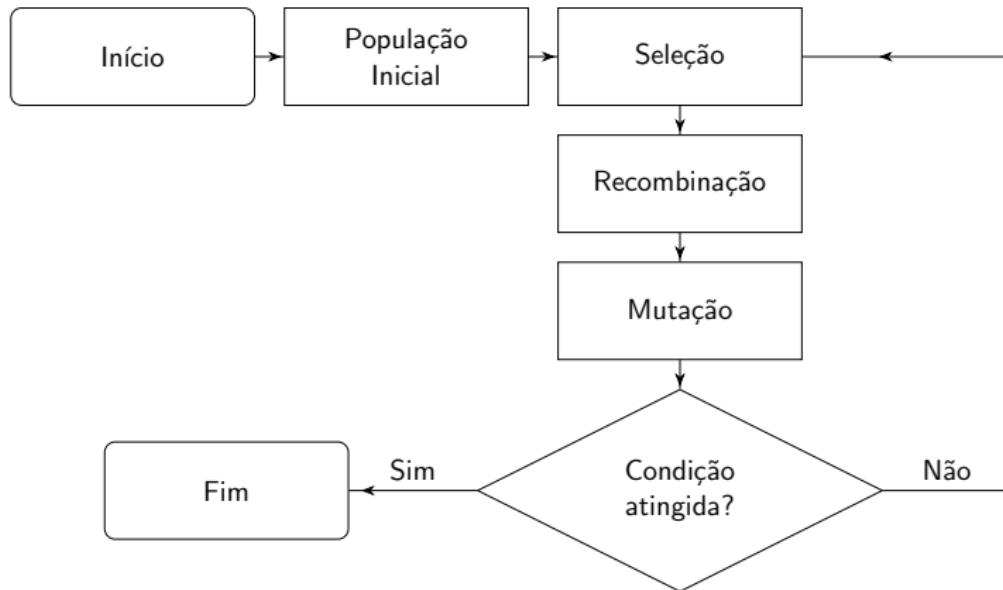


Figura: *Flowchart* geral de um algoritmo genético.

Codificação

Vamos considerar uma função objetivo $f : \mathcal{C} \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$

- A população é representada por n matrizes

$$A_k = \begin{bmatrix} a_{11}^{(k)} & \cdots & a_{1m}^{(k)} \\ \vdots & \ddots & \vdots \\ a_{l1}^{(k)} & \cdots & a_{lm}^{(k)} \end{bmatrix}$$

- As respectivas n posições são

$$X_k = (x_1^{(k)}, \dots, x_j^{(k)}, \dots, x_m^{(k)})$$

Codificação

- O espaço de busca é $S \subseteq \mathcal{C}$ tal que

$$S = \bigtimes_{j=1}^m [x_j^{(min)}, x_j^{(max)}]$$

- l bits representam inteiros em $[0, 2^l - 1]$
- As posições são obtidas dos cromossomos por

$$x_j^{(k)} = x_j^{(min)} + \frac{x_j^{(max)} - x_j^{(min)}}{2^l - 1} \sum_{i=1}^l a_{ij}^{(k)} 2^{i-1}$$

Codificação

- A relação inversa é obtida por

$$\sum_{i=1}^I a_{ij}^{(k)} 2^{i-1} = \left\lfloor \frac{(x_j^{(k)} - x_j^{(min)})(2^I - 1)}{x_j^{(max)} - x_j^{(min)}} \right\rfloor$$

Seleção

- Função Desempenho $g : \mathbb{R} \rightarrow \mathbb{R}$ tal que

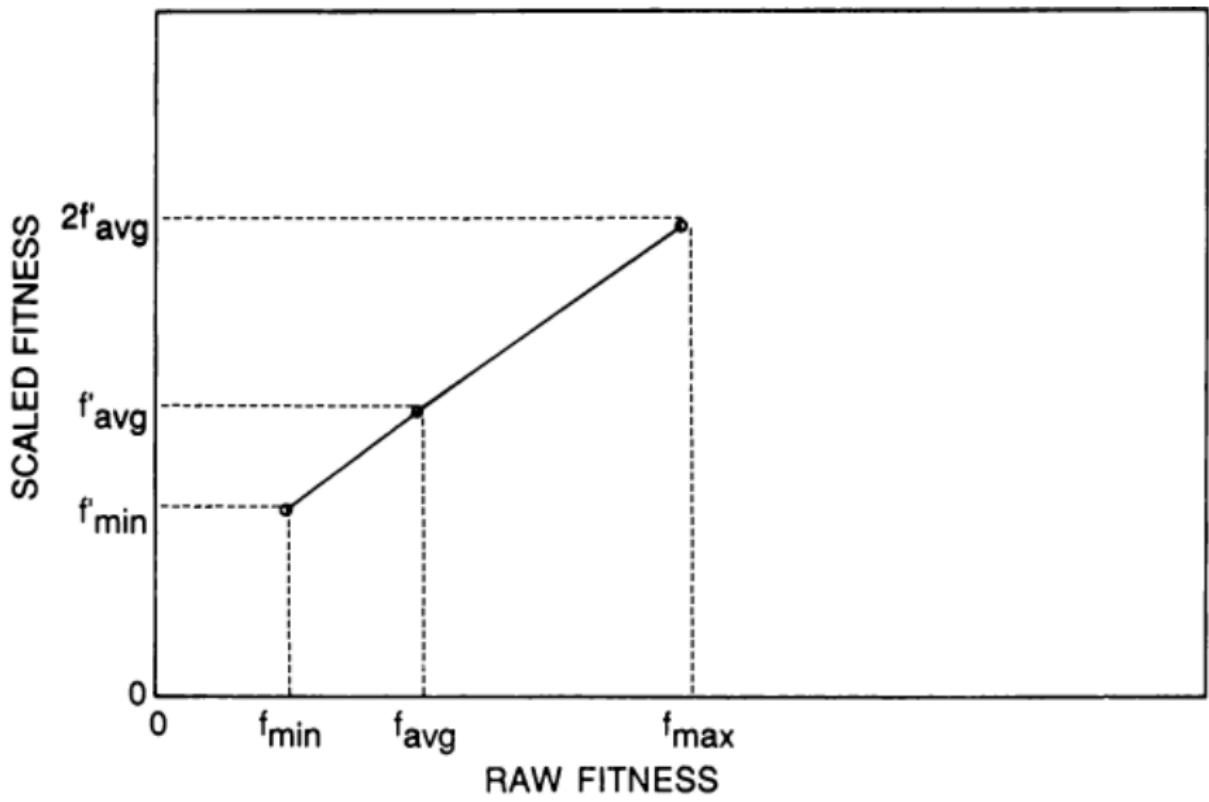
$$P_k = \frac{g(f(X_k))}{\sum_{j=1}^n g(f(X_j))}$$

- A função usada foi um escalamento linear

$$g(x) = ax + b$$

$$a = \begin{cases} \frac{\mu(h-1)}{f_{max}-\mu} , & \text{se } f_{min} > \frac{h\mu-f_{max}}{h-1} \\ \frac{\mu}{\mu-f_{min}} , & \text{caso contrário} \end{cases}$$

$$b = \begin{cases} \frac{\mu(f_{max}-h\mu)}{f_{max}-\mu} , & \text{se } f_{min} > \frac{h\mu-f_{max}}{h-1} \\ -\frac{\mu f_{min}}{\mu-f_{min}} , & \text{caso contrário} \end{cases}$$



Seleção

- Caso $g(f(X_k)) < 0$ para algum k , fazemos $g = 0$ para os indivíduos correspondentes
- É feito o sorteio de $n/2$ indivíduos por meio de uma roleta simples
- Os indivíduos selecionados são

$$S = \{S_1, \dots, S_k, \dots, S_{n/2}\}$$

Motivos para essa escolha de g

- Evita a convergência prematura da população
- Aumenta a variabilidade genética
- Aumenta a probabilidade de que vários máximos locais coexistam na população com o máximo global

Recombinação e Mutação

- Cada cromossomo terá dois pontos de recombinação
- Sorteio de $\alpha_j^{(k)}$ e $\beta_j^{(k)}$ tais que $0 \leq \alpha_j^{(k)} < \beta_j^{(k)}$ e $\alpha_j^{(k)} \leq \beta_j^{(k)} < l$ onde $j = 0, 1, \dots, m$ e $k = 0, 2, 4, \dots, n/2 - 2$
- Com esses valores, criamos a máscara de combinação

$$r_{ij}^{(k)} = \begin{cases} 0, & \text{se } i \leq \alpha_j^{(k)} \text{ ou } i > \beta_j^{(k)} \\ 1, & \text{caso contrário} \end{cases}$$

- Criamos uma máscara de mutação

$$m_{ij}^{(k)} = \begin{cases} 1, & \text{se houver mutação} \\ 0, & \text{caso contrário} \end{cases}$$

Recombinação e Mutação

- O conjunto dos indivíduos filhos será

$$S' = \{S'_1, \dots, S'_k, \dots, S'_{n/2}\}$$

$$s'_{ij}^{(k)} = (s_{ij}^{(k)} \wedge r_{ij}^{(k)}) \vee (s_{ij}^{(k+1)} \wedge \neg r_{ij}^{(k)}) \oplus m_{ij}^{(k)}$$

$$s'_{ij}^{(k+1)} = (s_{ij}^{(k)} \wedge \neg r_{ij}^{(k)}) \vee (s_{ij}^{(k+1)} \wedge r_{ij}^{(k)}) \oplus m_{ij}^{(k)}$$

- A população da geração seguinte é $S \cup S'$

Recombinação e Mutação

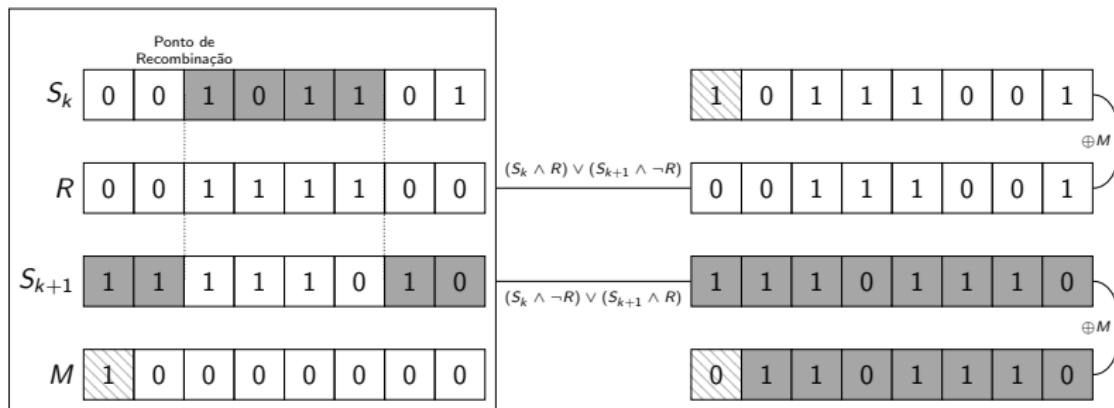


Figura: Diagrama ilustrando—em um caso com $m = 1$ e $l = 8$ —a recombinação entre os indivíduos S_k e S_{k+1} e mutação nos filhos gerados segundo as respectivas máscaras de recombinação e mutação R e M .

Estratégia Elitista

$$S = \left\{ \underbrace{S_\epsilon, \dots, S_\epsilon}_{e_1 \text{ cópias}}, \underbrace{S_\epsilon, \dots, S_\epsilon}_{e_2 \text{ cópias}}, \underbrace{S_{e_1+e_2+1}, \dots, S_{n/2}}_{e_3 \text{ cópias dentre os restantes}} \right\}$$

- S_ϵ é o melhor indivíduo
- As primeiras e_1 cópias de S_ϵ terão probabilidade de mutação $p_1 = 0$
- As e_2 cópias de S_ϵ terão probabilidade de mutação p_2
- As e_3 cópias de S_ϵ terão probabilidade de mutação p_3

Estratégia Elitista

Motivos para o uso de elitismo

- Garante o S_ϵ na população seguinte
- Mutações de S_ϵ na geração seguinte
- Filhos de S_ϵ com outros membros da população
- Ganho de performance em funções comportadas

Observação

- Essa é apenas uma dentre muitas estratégias [1]
- Os parâmetros devem ser ajustados caso a caso

Funções de Teste

$$f_1(x, y) = \cos(9\pi r) \exp\left\{-\frac{r^2}{(0, 4)^2}\right\}$$

$$r = \sqrt{(x - 0, 5)^2 + (y - 0, 5)^2}$$

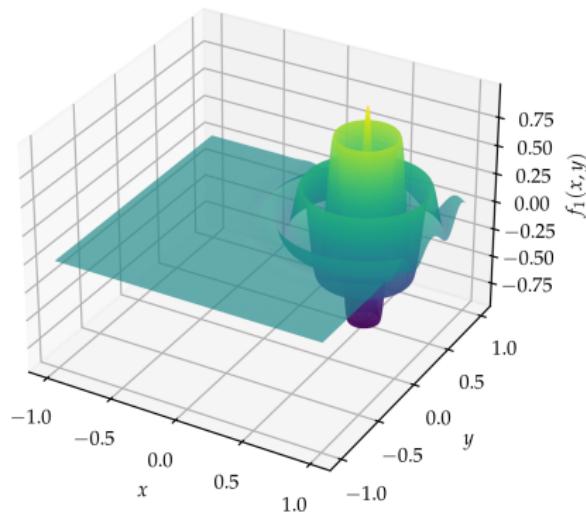


Figura: Gráfico de $f_1(x, y)$.

Funções de Teste

$$f_2(x, y) =$$

$$+ 0,8 \exp \left\{ -\frac{r_1^2}{(0,3)^2} \right\}$$

$$+ 0,88 \exp \left\{ -\frac{r_2^2}{(0,03)^2} \right\}$$

$$r_1 = \sqrt{(x - 0,5)^2 + (y - 0,5)^2}$$

$$r_2 = \sqrt{(x - 0,6)^2 + (y - 0,1)^2}$$

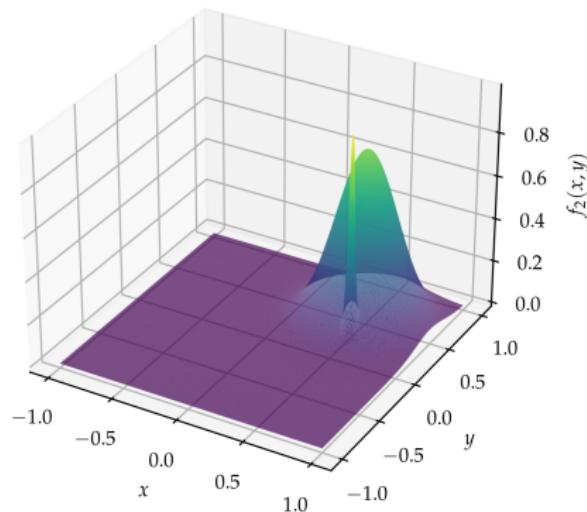
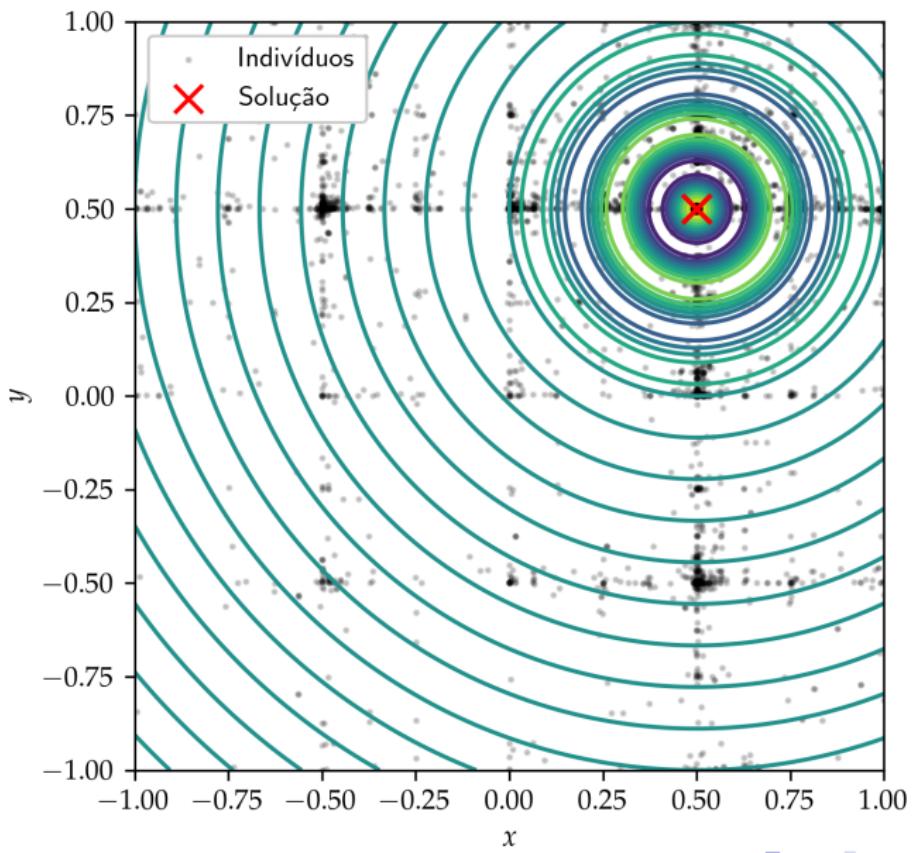
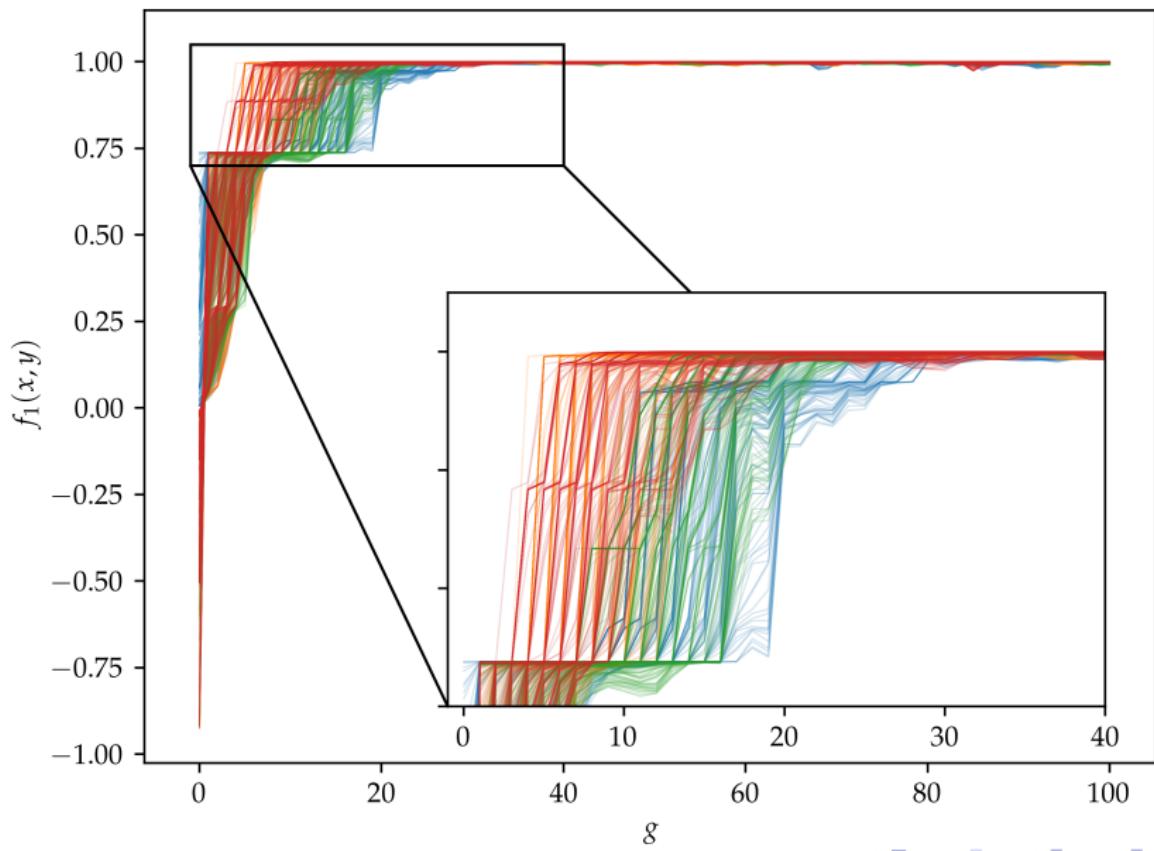


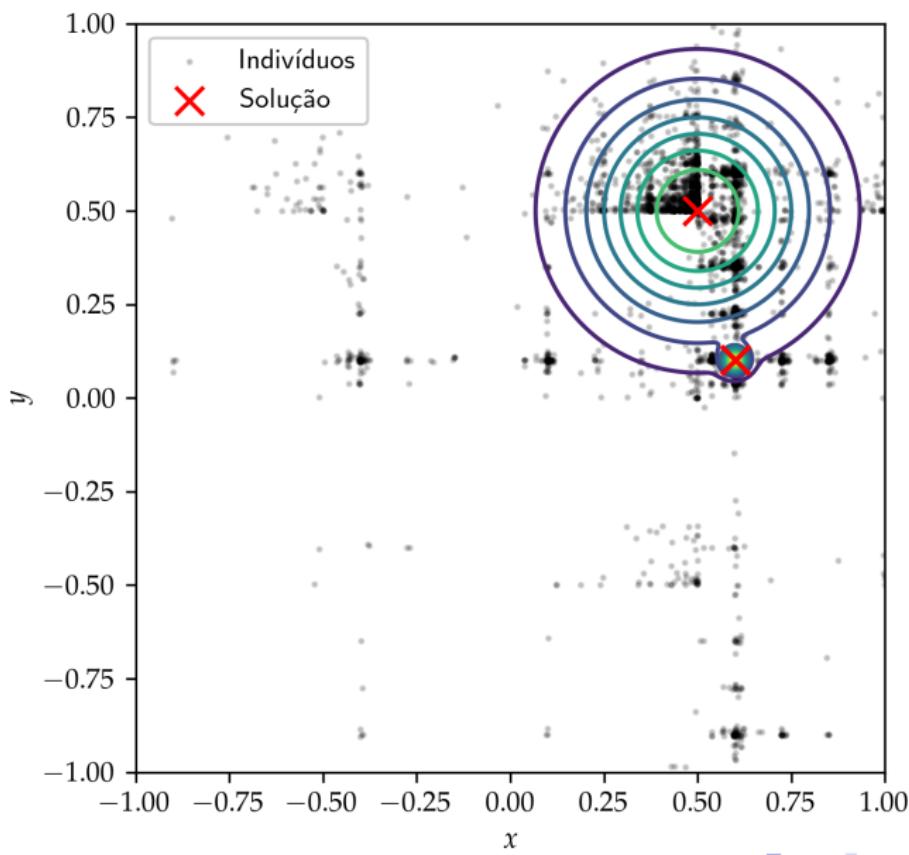
Figura: Gráfico de $f_2(x, y)$.

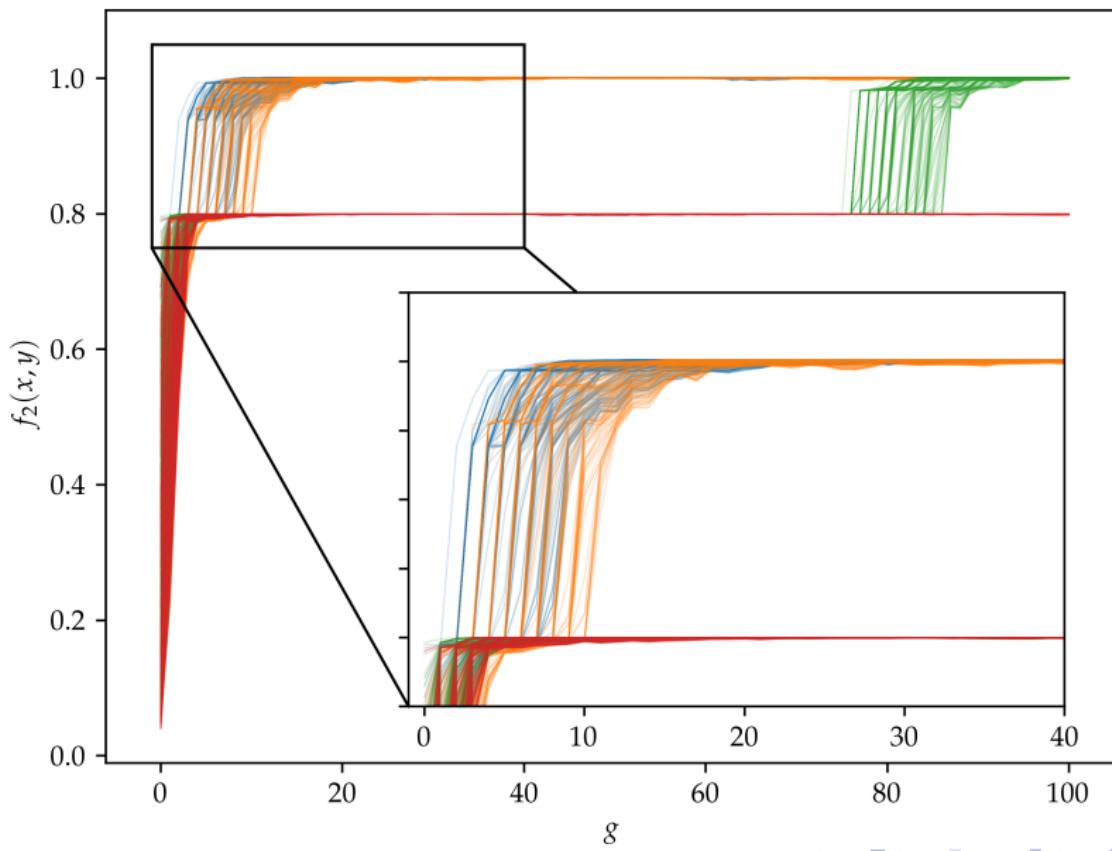
Configurações Utilizadas

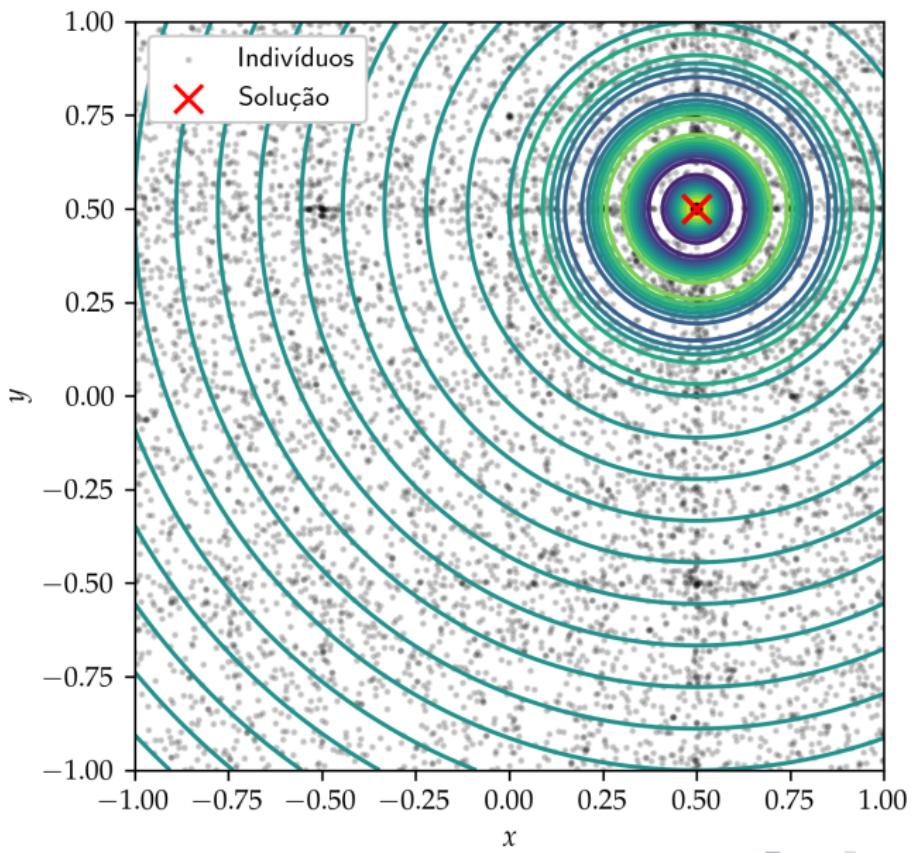
- 8 populações de 1000 indivíduos
- Cromossomos de 32 bits
- Espaço de busca $\mathcal{S} = [-1, 1] \times [-1, 1]$
- Parâmetro da função desempenho $h = 2$
- $e_1 = 4, e_2 = 6, e_3 = 10$
- $p_2 = p_3 = 5\%$ e $p_2 = p_3 = 20\%$

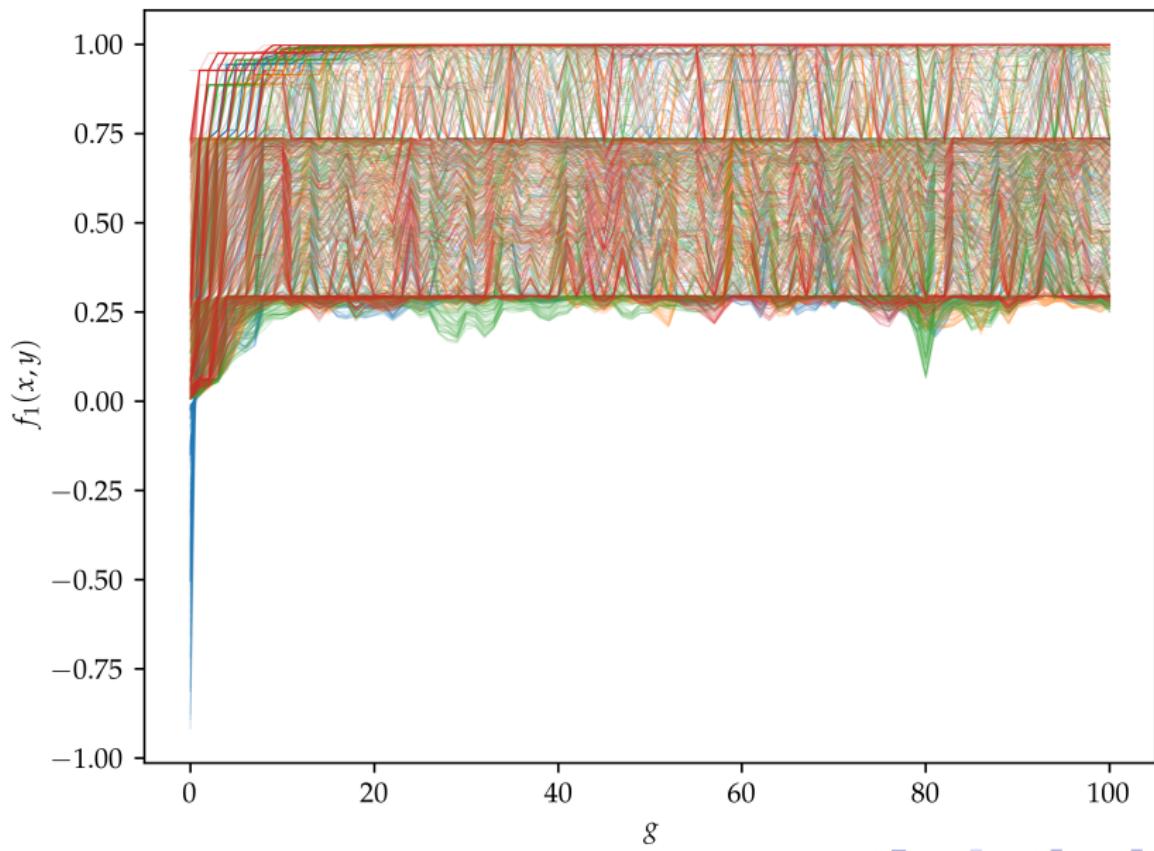


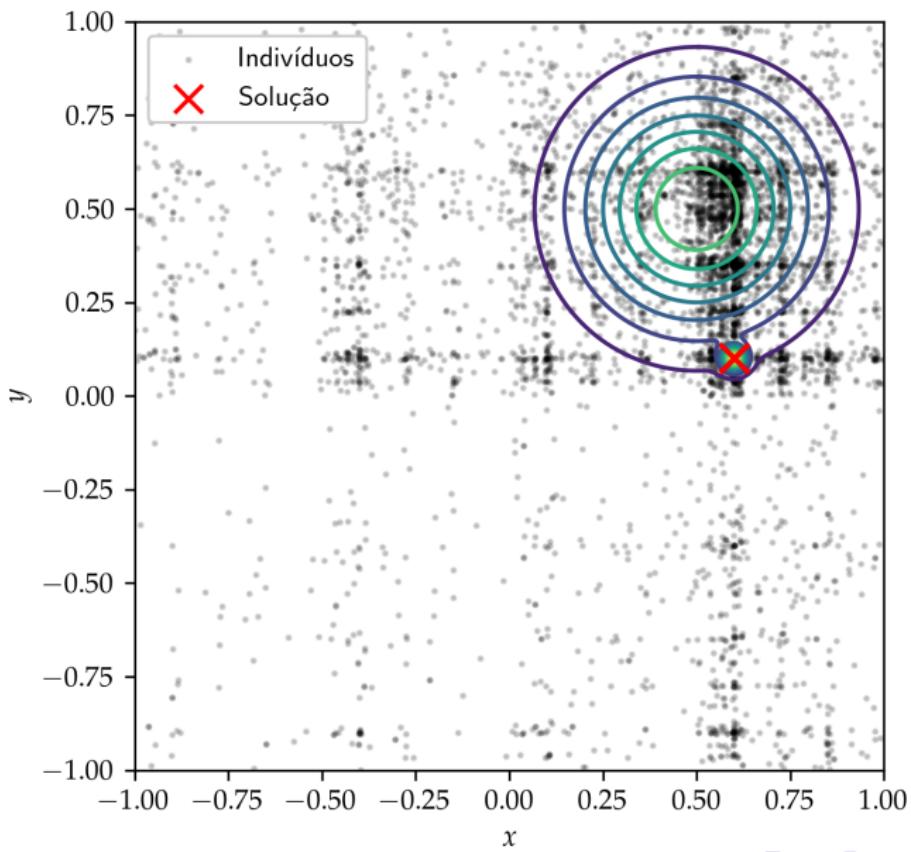


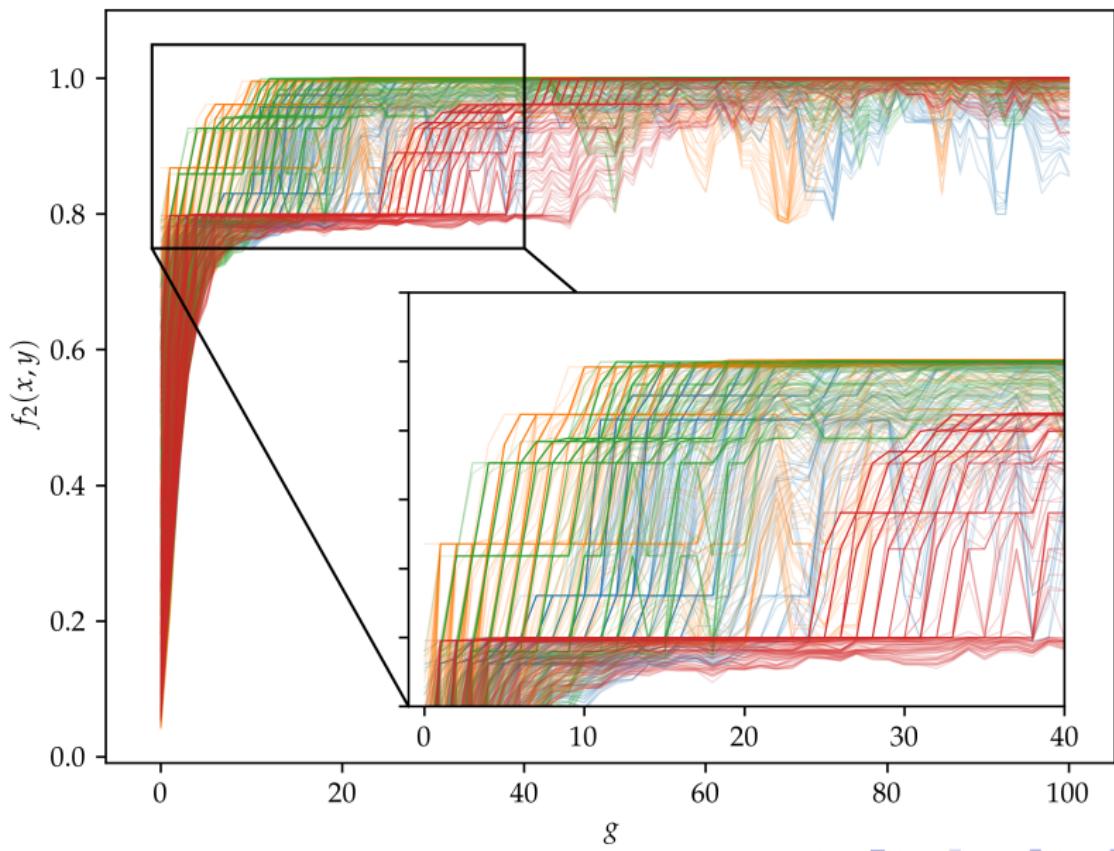












Conclusões

- Em todos os casos o máximo global foi localizado
- Os máximos locais também foram localizados
- Uma variabilidade genética foi mantida
- A evolução por 100 gerações de 8 populações em paralelo levou ≈ 15 segundos
- Se mostrou escalável: 100 gerações com 10^5 indivíduos em ≈ 1 hora

O que não foi testado

- Problemas com dimensão superior, como ajuste de curvas com 4 ou mais parâmetros
- Outros valores de h , e_1 , e_2 e e_3 e outras funções desempenho

Cronograma

- Período 2021/02 — Nesse período foi feita a leitura da bibliografia citada neste texto, e então foi implementada a primeira versão do algoritmo. Posteriormente, foi feita uma série de testes com diferentes funções, dentre as quais estão as citadas nesse trabalho. Finalmente, foi confeccionado este texto.
- Período 2022/01 — No período subsequente, planeja-se escolher um problema físico que envolva a otimização de funções, realizar uma consulta sobre referências bibliográficas relacionadas, fazer a leitura dos textos encontrados e aplicar o programa na solução do problema. Por fim, o Trabalho de Conclusão de Curso será elaborado acerca dos resultados obtidos e a defesa dessa tese será apresentada.



D. E. Goldberg.

Genetic Algorithms in Search, Optimization and Machine Learning.

Addison-Wesley Longman Publishing Co., Inc., USA, 1st edition, 1989.



J. H. Holland.

Genetic algorithms.

Scientific American, 267(1):66–73, 1992.



L. F. Roncaratti, R. Gargano, and G. M. e Silva.

A genetic algorithm to build diatomic potentials.

Journal of Molecular Structure: THEOCHEM, 769(1-3):47–51, 2006.