
Aula 04:

Computação Evolutiva e Conexão – Estratégias Evolutivas

Prof. Hugo Puertas de Araújo
hugo.puertas@ufabc.edu.br
Sala: 509.2 (5º andar / Torre 2)

■ Agenda

■ Estratégias Evolutivas

- ❖ Visão Geral
- ❖ Definições dos parâmetros da Estratégia Evolutiva (EE)

Computação Evolutiva e Conexionista

■ Visão geral

■ Desenvolvida: Na Alemanha na década de 60

- ❖ Pioneiros: I. Recheiberg, H.-P. Schwefel na Technical University of Berlin para otimização em mecânica de fluidos
- ❖ Aplicações típicas:
 - Otimização Numérica

■ Características:

- ❖ Velocidade de convergência
- ❖ Bom otimizador para problemas de valores reais

■ Especialidade:

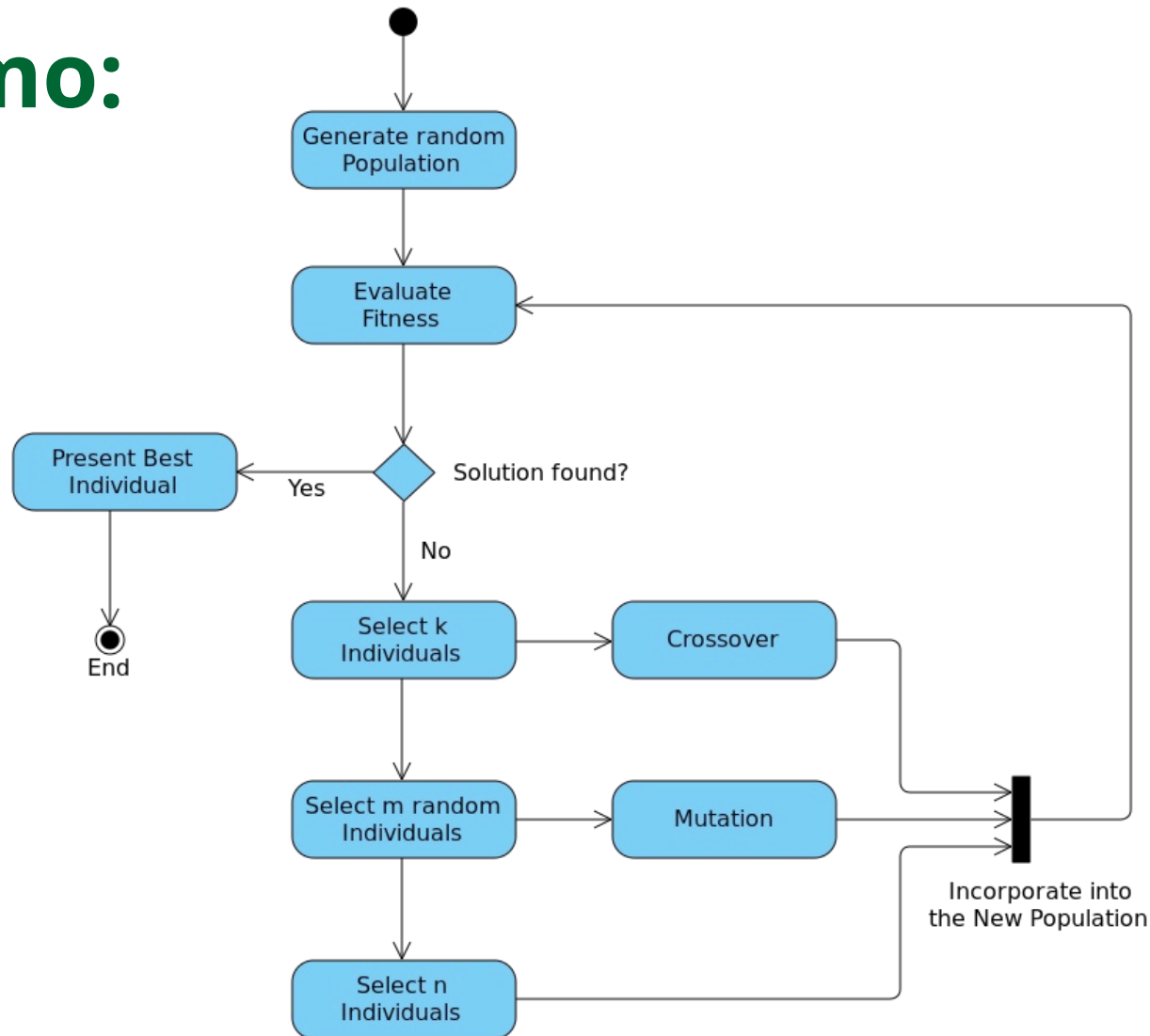
- ❖ Autoadaptação de parâmetros padrões (mutação)

Visão geral

Sumário: características geralmente encontradas

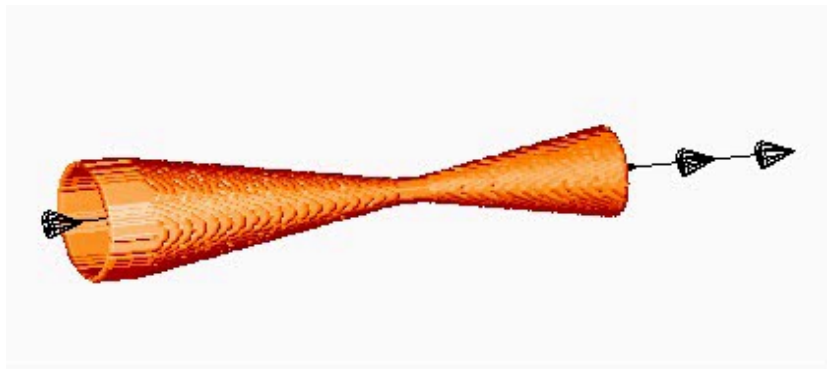
Representação	Vetores com valores reais
Recombinação	Discreta ou intermediária
Mutação	Gaussiana com possível auto-adaptação dos parâmetros de mutação
Seleção dos pais	Aleatória com distribuição uniforme
Seleção para próxima população	(μ, λ) ou $(\mu + \lambda)$

O algoritmo:



Exemplo de utilização

- Tarefa: otimizar o formato de um bocal de 2 fases
- Método: mutações aleatórias, seleção ($\mu+\lambda$)



<https://ls11-www.cs.tu-dortmund.de/people/schwefel/EADemos/>

Exemplo de utilização

– Tarefa:

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \in \mathbb{R} \\ \text{sujeito a} & \mathbf{x} \in \mathbb{R}^n \end{array}$$

- Algoritmo: ES com um único indivíduo
- Cromossomos: vetores reais n -dimensionais
 - Tamanho da população: 1
 - Mutação cria um filho a partir do único pai
 - Seleção elitista
 - Observação: no trabalho original, algumas restrições tiveram que ser feitas

Exemplo de utilização

Algoritmo ES com um único indivíduo

Início

$t \leftarrow 0$

Crie ponto inicial $\mathbf{x}_t = [x_t(1), \dots, x_t(n)]^T$

avale indivíduo \mathbf{x}_t

enquanto (critérioConvergencia == 0)

gere \mathbf{z}_t a partir de distribuição normal

$\mathbf{y}_t \leftarrow \mathbf{x}_t + \mathbf{z}_t$

avale indivíduo \mathbf{y}_t

se $f(\mathbf{x}_t) < f(\mathbf{y}_t)$

$\mathbf{x}_{t+1} \leftarrow \mathbf{x}_t$

senão

$\mathbf{x}_{t+1} \leftarrow \mathbf{y}_t$

fim se

$t \leftarrow t + 1$

fim enquanto

Fim

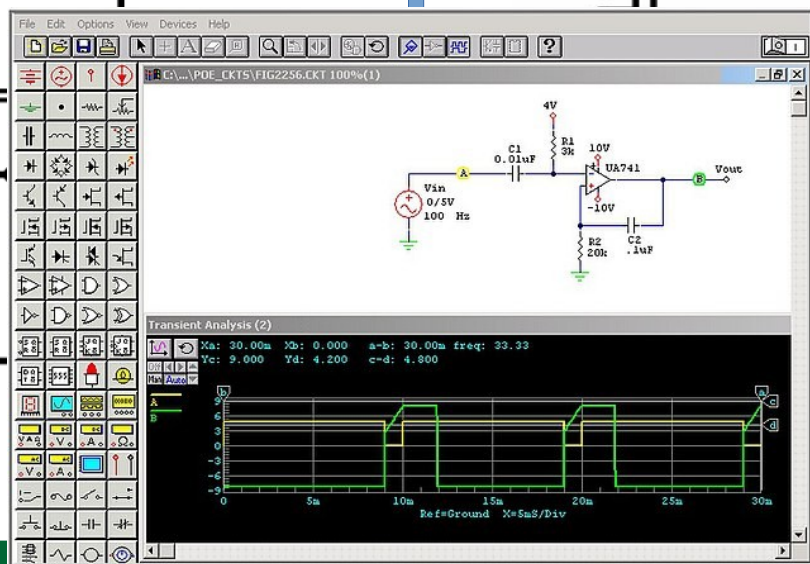
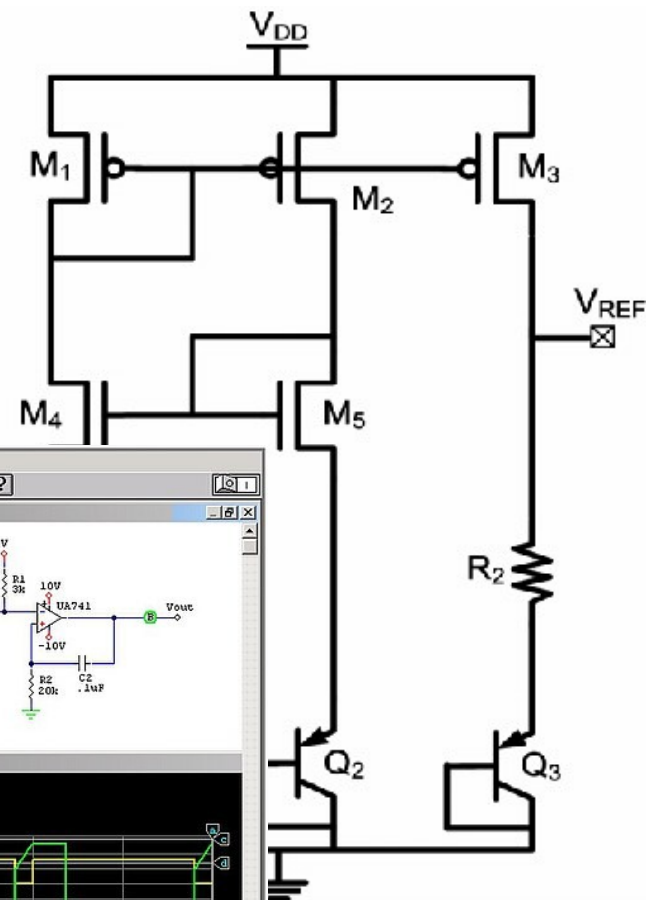
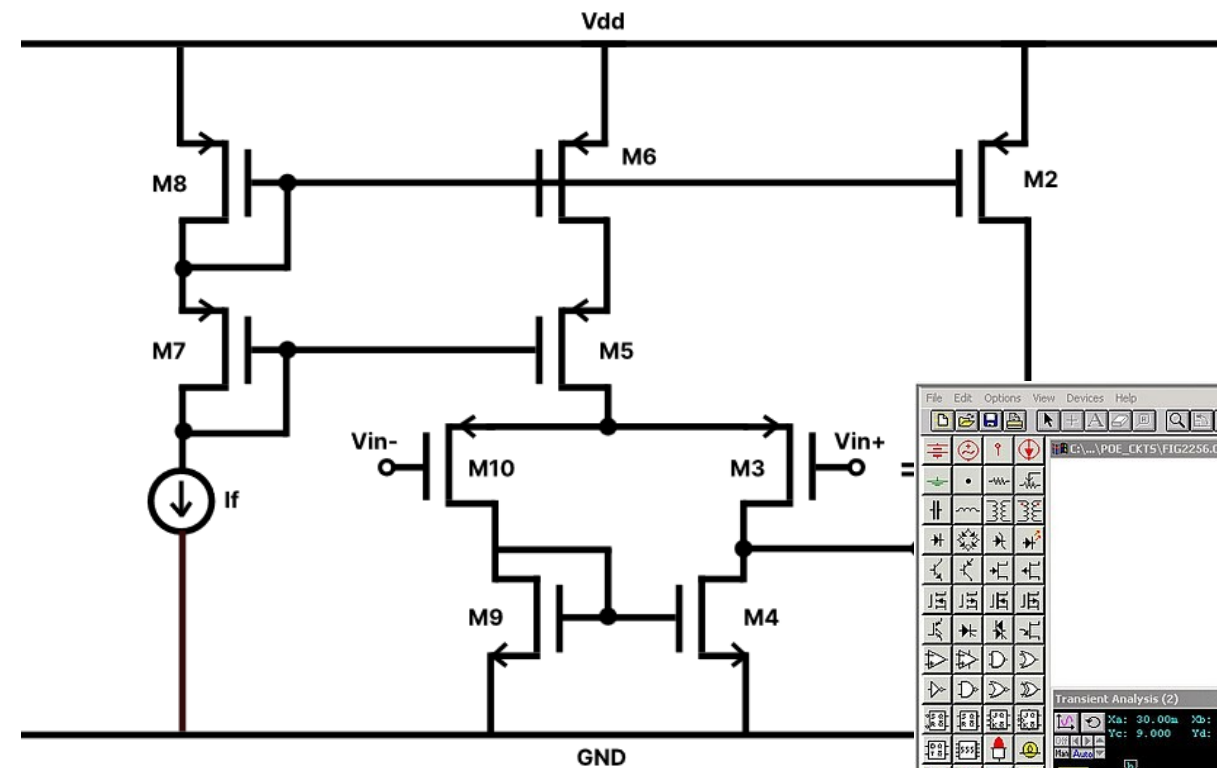


Initial shape



Final shape

Exemplo de utilização



■ Operadores genéticos

- Seleção
- Mutação
- Recombinação

■ Seleção

Algoritmo Evolutivo Estacionário

Início

$t \leftarrow 1$

inicializePopulacao(\mathbf{P}_t)

avaliePopulacao(\mathbf{P}_t)

enquanto (criterioConvergencia == 0)

$\mathbf{Q}_t \leftarrow \text{transformePopulacao}(\mathbf{P}_t)$

 avaliePopulacao(\mathbf{Q}_t)

$\mathbf{P}_{t+1} \leftarrow \text{selecao}(\mathbf{P}_t \cup \mathbf{Q}_t)$

$t \leftarrow t + 1$

fim enquanto

Fim

■ Seleção

■ $(\mu + \lambda)$ -ES

- ❖ Utilizando mutação e recombinação
- ❖ uma população com μ pais cria λ descendentes
- ❖ Os melhores μ indivíduos entre os μ pais E os λ descendentes formam a nova população:
 - Elitista

■ Algoritmo Evolutivo Estacionário

■ Seleção

■ (μ, λ) -ES

- ❖ Uma população com μ pais cria $\lambda \geq \mu$ descendentes utilizando mutação e recombinação
- ❖ Os melhores μ descendentes formam a nova população
 - Não é elitista
- ❖ Variação:
 - $P_{t+1} \leftarrow \text{selecao}(Q_t)$

■ Seleção

■ (μ, λ) -ES ou $(\mu + \lambda)$ -ES ?

- ❖ Muitas vezes (μ, λ) -ES é utilizada pois, geralmente,
 - É melhor em escapar de ótimos locais
 - É melhor para seguir ótimos que se movem continuamente
 - Não conserva (como a $(\mu + \lambda)$ -ES) por muito tempo valores ruins dos parâmetros que geram o passo de mutação (quando estes são evoluídos) em indivíduos promissores
- ❖ Pressão seletiva geralmente é mantida alta em ESs
 - $\lambda \approx 7 \cdot \mu$ é muitas vezes utilizada

■ Seleção

■ Seleção dos indivíduos a serem reproduzidos

- ❖ Pais são selecionados aleatoriamente com distribuição uniforme
 - para qualquer tipo de operador de reprodução
 - assim, a seleção é sem bias
- ❖ Todo indivíduo tem a mesma chance de ser selecionado para ser reproduzido

■ Seleção

Exercício

■ Seleção dos indivíduos a serem reproduzidos

- ❖ Pais são selecionados aleatoriamente com distribuição uniforme
 - para qualquer tipo de operador de reprodução
 - assim, a seleção é sem bias
- ❖ Todo indivíduo tem a mesma chance de ser selecionado para ser reproduzido

E.: Esse tipo de seleção é do tipo 'roleta' ou 'torneio'?

■ Mutação

- Geralmente é utilizada a mutação Gaussiana

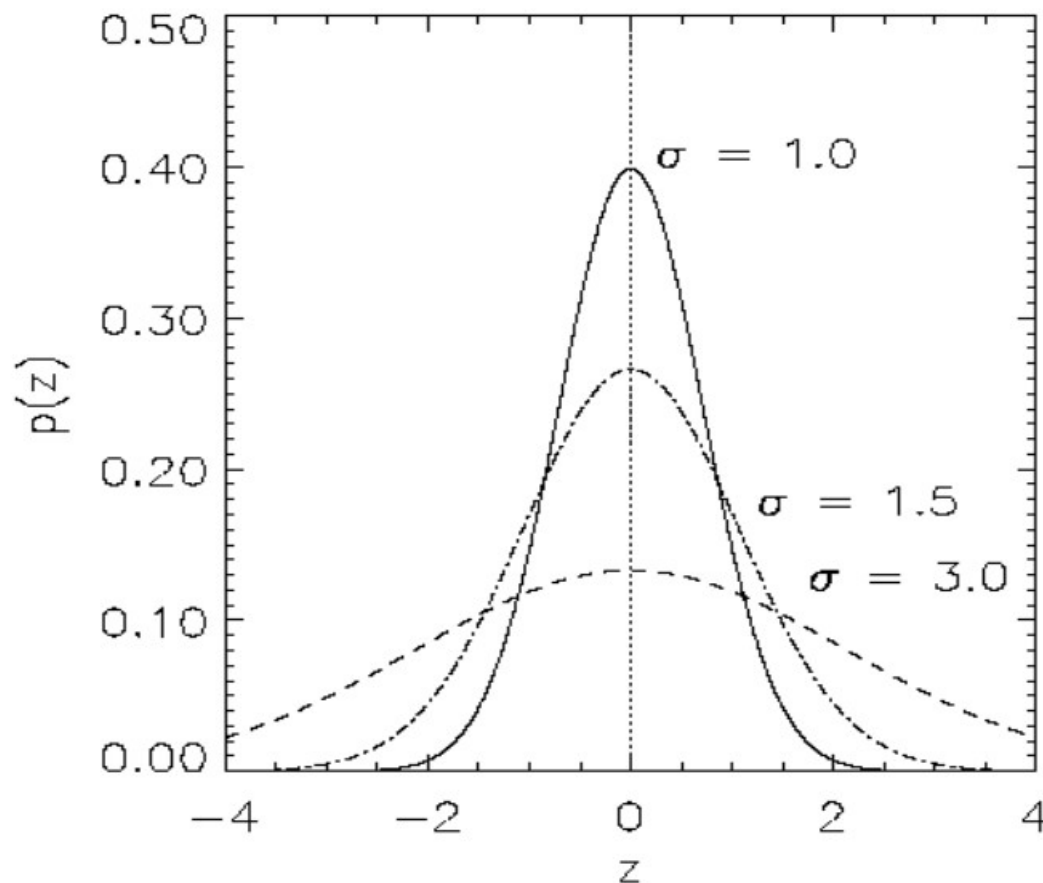
$$\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{C}_{t+1} \mathbf{N}(0,1)$$

- $\mathbf{N}(0,1)$ é um vetor aleatório gerado a partir de uma distribuição normal com média 0 e variância 1
- A matriz \mathbf{C} define o espalhamento da distribuição

Mutação

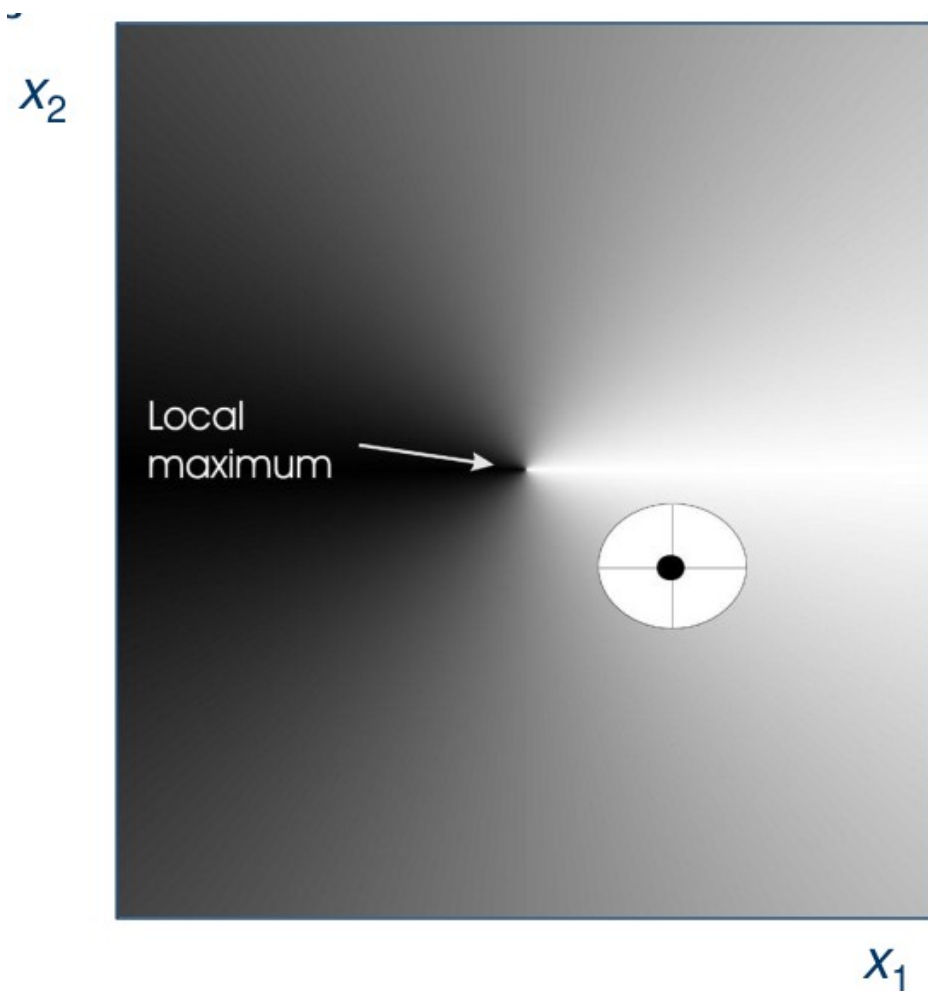
Mutação Gaussiana

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{C}_{t+1} \mathbf{N}(0,1)$$



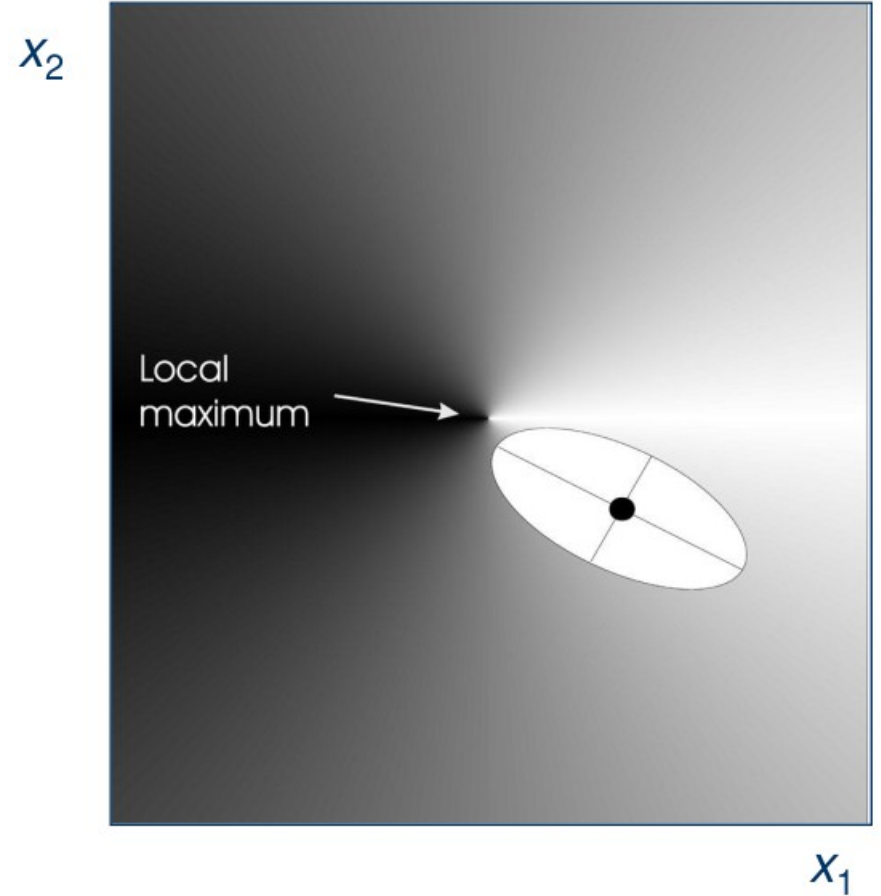
Mutação

- Mutação Gaussiana não-correlacionada com um parâmetro de mutação:



Mutação

■ Mutação Gaussiana correlacionada:



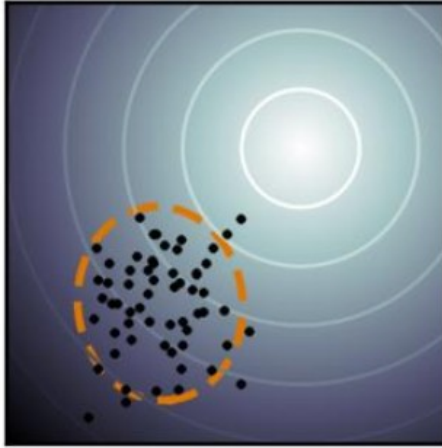
Mutação

■ Covariance Matrix Adaptation Evolution Strategy (CMA-ES)

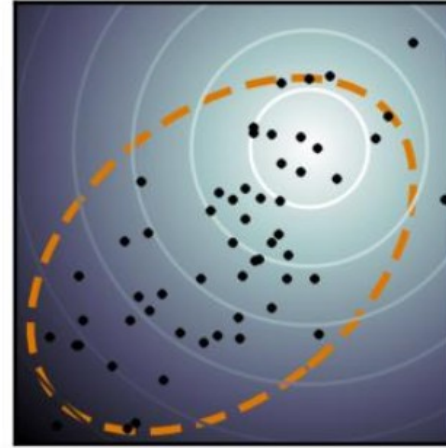
- ❖ A média da distribuição é atualizada de modo que a probabilidade de soluções candidatas que previamente apresentaram sucesso é maximizada
- ❖ Os elementos da matriz de covariância C são incrementadas de tal forma que a probabilidade de repetir mudanças nos passos de busca que apresentaram sucesso é aumentada
 - Para isso, a evolução da distribuição no tempo é armazenada (evolution paths)
- ❖ Assim, de maneira indireta, estima-se o gradiente da superfície de fitness

Mutação

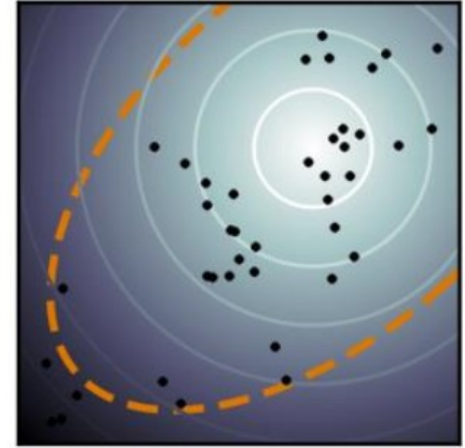
Generation 1



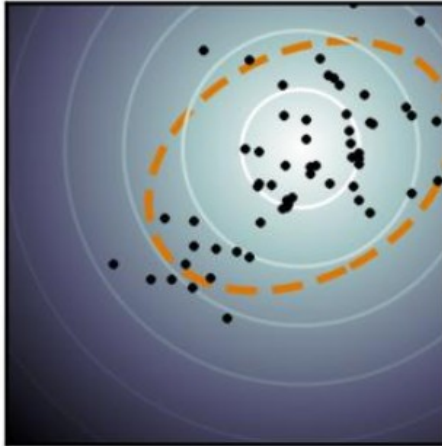
Generation 2



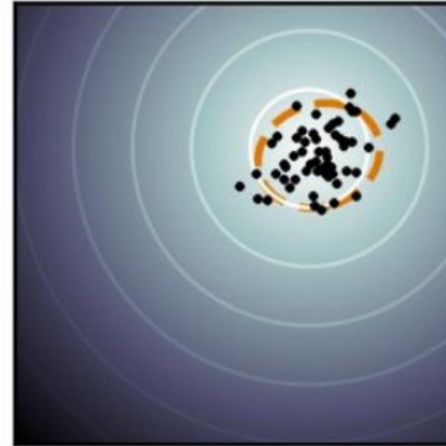
Generation 3



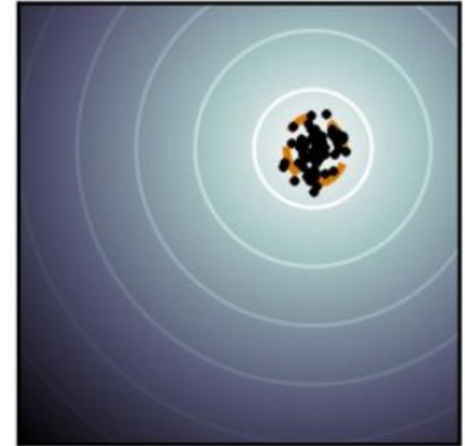
Generation 4



Generation 5



Generation 6



■ Representação cromossomial

■ Cromossomos consistem de até três partes:

❖ Variáveis do objeto: x_1, \dots, x_n

● Vetor x (sempre presente)

❖ 2 vetores de parâmetros de mutação (strategy parameters)

● Tamanhos dos passos de mutação: $\sigma_1, \dots, \sigma_n$

– Pode usar menos do que n parâmetros

● Ângulos de rotação: $\alpha_1, \dots, \alpha_k$

– Sendo $k = n(n-1)/2$

❖ Pode usar menos do que k ângulos

■ Vetor completo: $[x_1, \dots, x_n, \sigma_1, \dots, \sigma_n, \alpha_1, \dots, \alpha_k]$

■ Representação cromossomial

Exercício

- Cromossomos consistem de até três partes:
 - ❖ Variáveis do objeto: x_1, \dots, x_n
 - Vetor x (sempre presente)
 - ❖ 2 vetores de parâmetros de mutação (strategy parameters)
 - Tamanhos dos passos de mutação: $\sigma_1, \dots, \sigma_n$
 - Pode usar menos do que n parâmetros
 - Ângulos de rotação: $\alpha_1, \dots, \alpha_k$
 - Sendo $k = n(n-1)/2$
 - ❖ Pode usar menos do que k ângulos
- Vetor completo: $[x_1, \dots, x_n, \sigma_1, \dots, \sigma_n, \alpha_1, \dots, \alpha_k]$

E.: Como vocês representariam o cromossomo p/ os exemplos dados?

■ Recombinação

- Diferentemente dos AGs, gera apenas um filho de cada vez
- Dois tipos principais:
 - ❖ Intermediária
 - ❖ Discreta

Recombinação

Intermediária

$$x_g = \frac{1}{\rho} \sum_{i=1}^{\rho} x_g(i)$$

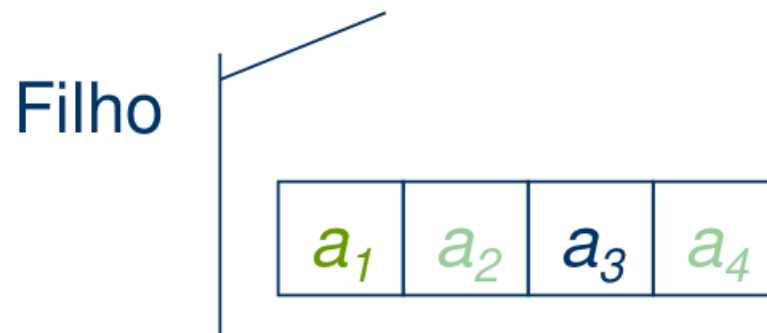
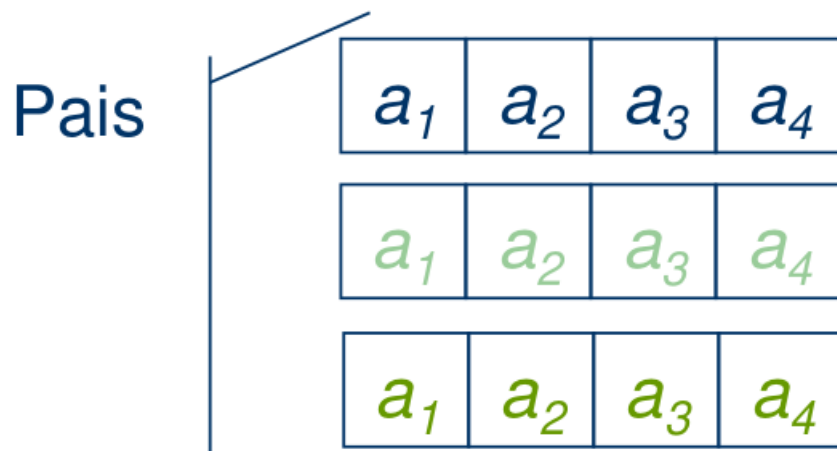
na qual:

- ❖ g indica o índice do elemento do vetor x
- ❖ i indica o índice do pai
- ❖ ρ indica o número de pais utilizado para gerar um filho
 - Geralmente 2 ou μ

Recombinação

Discreta (ou dominante)

- ❖ Os genes são escolhidos aleatoriamente (com distribuição uniforme) entre os genes do p pais
- ❖ Exemplo:



Autoadaptação

Parâmetros de mutação são permitidos evoluir também

- Mutação Gaussiana Correlacionada

$$\sigma'_i = \sigma_i \exp(\tau' N(0,1) + \tau N_i(0,1))$$

$$\alpha'_j = \alpha_j + \beta N(0,1)$$

- Tipicamente, as “taxas de aprendizado” são

$$\tau' \propto 1/(2n)^{1/2}$$

$$\tau \propto 1/(2n^{1/2})^{1/2}$$

$$\beta \approx 5^\circ$$

- Limite inferior para vetor α

$$|\alpha'_j| > \pi \Rightarrow \alpha'_j = \alpha'_j - 2\pi \operatorname{sign}(\alpha'_j)$$

■ Autoadaptação

Pré-requisitos

- $\mu > 1$: para testar diferentes estratégias
- $\lambda > \mu$: mais filhos do que pais
- Pressão seletiva não demasiadamente alta, e.g.,
 $\lambda \approx 7 \mu$
- (μ, λ) -ES : diminuir o problema de parâmetros de mutação mal-adaptados
- Recombinação dos parâmetros de mutação através de recombinação (intermediária)