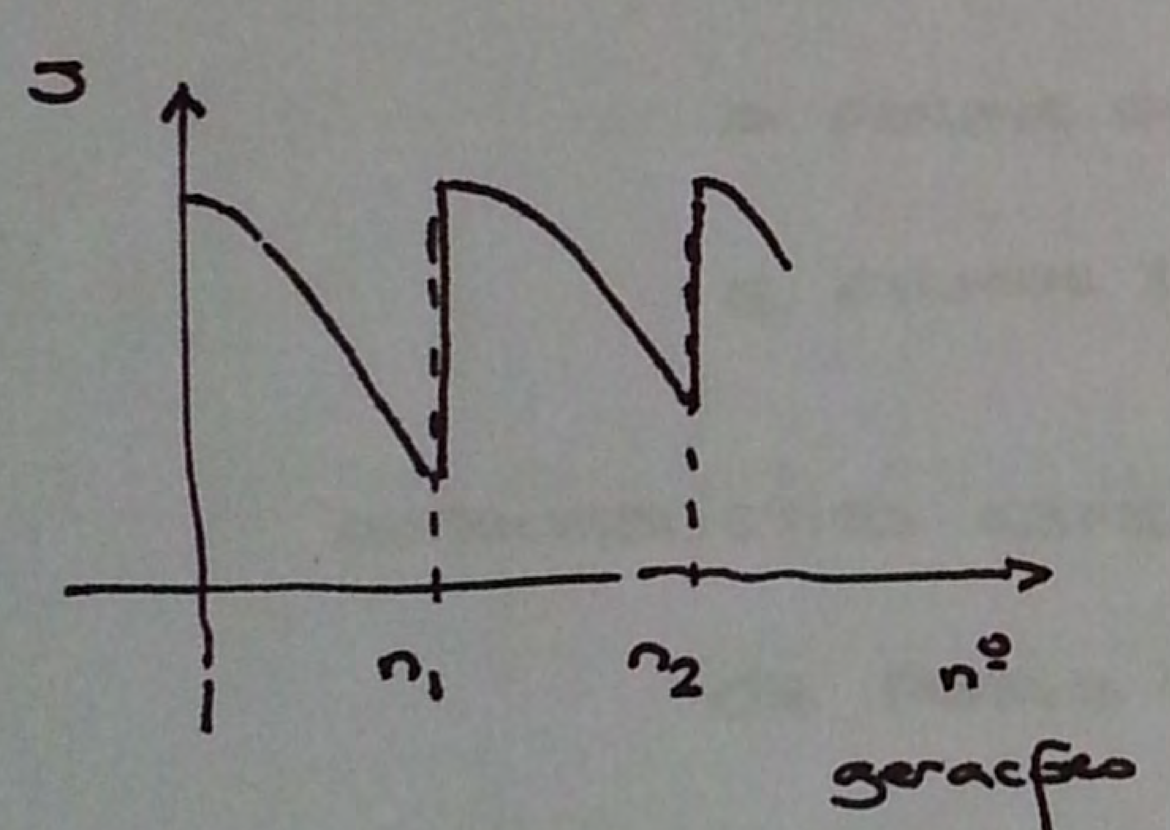
4 ESTRATÉGIAS DE EVOLUÇÃO (ES)

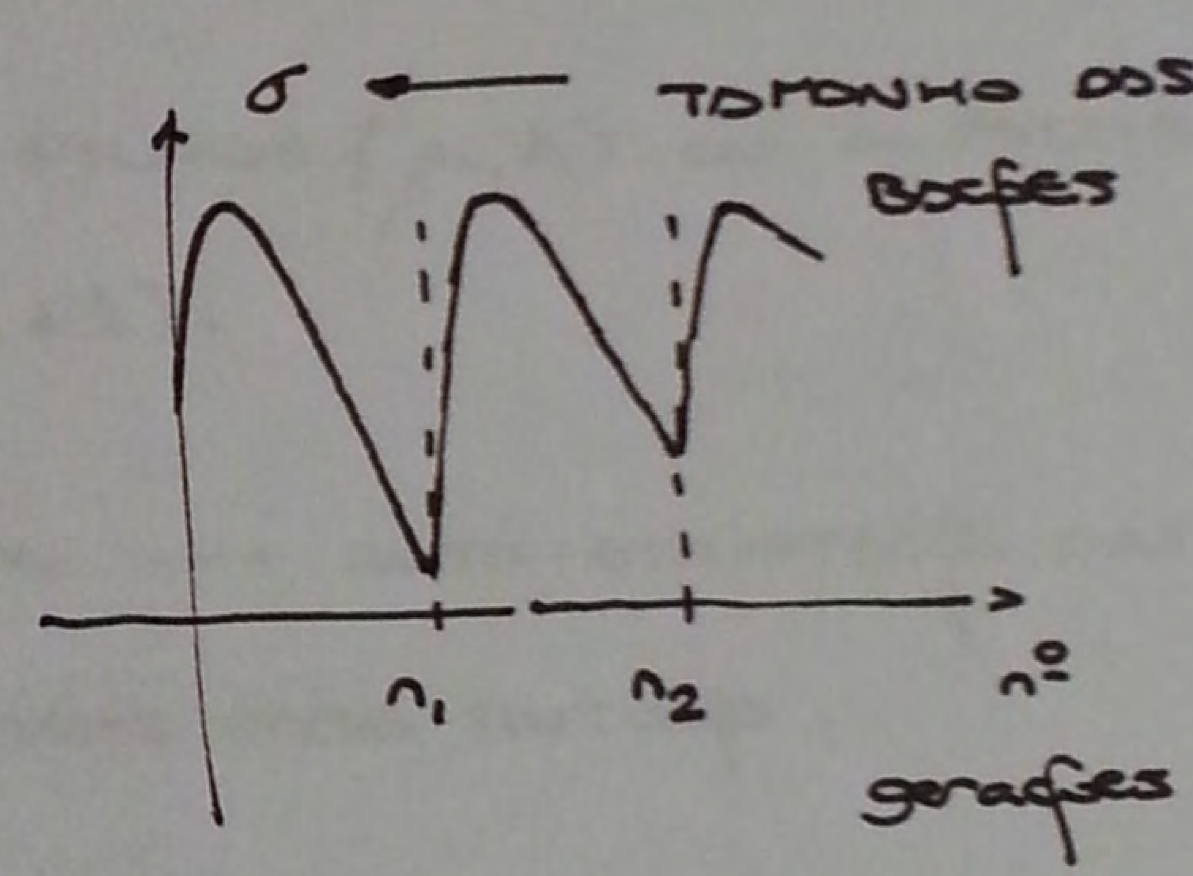
HANS-PAUL SCHWEFEL, 1977*

—> AUTO-ADAPTAÇÃO, NO SENTIDO DE ENCONTRAR TAMANHOS DE PERTUBAÇÕES ADEQUADOS PARA A OBTENÇÃO DE BONS RESULTADOS.

— REDUÇÃO DO TAMANHO DOS PASSOS AO LONGO DO TEMPO

— COMPORTAMENTO ADEQUADO QUANDO HÁ VARIAÇÃO TEMPORAL DA FUNÇÃO CUSTO: $J(x,t)$





— PERTURBAÇÕES INTEIRAMENTE ALEATÓRIAS, SEM CRITÉRIO DE ACEITAÇÃO (OU SEJA, ~~100%~~ p_s DE ACEITAÇÃO):

$$\hat{x}_{k+1} = x_k + \varepsilon . R$$

VETOR COM VARIÁVEIS ALEATÓRIAS GAUSSIANAS DE MÉDIA ZERO E VARIÂNCIA UM.

$$x_{k+1} = \begin{cases} \hat{x}_{k+1}, \text{ SE } J(\hat{x}_{k+1}) < J(x_k) \\ 0, \text{ CASO CONTRÁRIO} \end{cases}$$

* DESCRIÇÃO DETALHADA. TRABALHO ORIGINAL COM INGO RECHENBERG EM 1960 APROXIMADAMENTE.

REGRA DE 20% DE SUCESSO (RECHENBERG):

$$\varepsilon = \begin{cases} \varepsilon/c, & \text{SE } p_s > 0.2 \\ \varepsilon c, & \text{SE } p_s < 0.2 \\ \varepsilon, & \text{SE } p_s = 0.2 \end{cases}, \quad \text{ONDE } 0.817 \leq c \leq 1$$

TÍPICO ES:

REPRESENTAÇÃO CONTÍNUA ("PONTO FLUTUANTE")

SELEÇÃO DE PAIS — ALEATÓRIA E UNIFORME

RECOMBINAÇÃO — DISCRETA OU INTERMEDIÁRIA

MUTAÇÃO — PERTURBAÇÃO GAUSSIANA

SELEÇÃO DE SOBREVIVENTES — ESCOLHA DETERMINÍSTICA A PARTIR DOS FILHOS (μ, λ) OU A PARTIR DOS PAIS E FILHOS $(\mu + \lambda)$.

CARACTERÍSTICA ESPECIAL — AUTO-ADAPTAÇÃO DOS TAMANHOS DE PASSO USADOS PARA MUTAÇÃO.

OBSERVAÇÕES:

1. TIPICAMENTE SE USA ES PARA OTIMIZAÇÃO DE PARÂMETROS CONTÍNUOS
2. HÁ UMA GRANDE ÊNFASE NA MUTAÇÃO
3. MUTAÇÃO USA PERTURBAÇÃO GAUSSIANA
4. PARÂMETROS DA MUTAÇÃO SÃO ALTERADOS DURANTE A EXECUÇÃO DO ALGORITMO.
5. EM DIFERENTES CIRCUNSTÂNCIAS*, PASSOS COM TAMANHOS DIFERENTES LEVAM A DESEMPENHOS DIFERENTES (ALGUNS MELHORES QUE OUTROS).

* TEMPORAIS — ESTÁGIOS DO PROCESSO EVOLUCIONÁRIO

ESPACIAIS — CARACTERÍSTICAS LOCAIS DA FUNÇÃO CUSTO

4.1 REPRESENTAÇÃO DE INDIVÍDUOS

VARIÁVEIS DE DESCRIÇÃO DO PROBLEMA + PARÂMETROS DE ESTRATÉGIA

GENÓTIPO : $(x_1, \ldots, x_n, \sigma_1, \ldots, \sigma_{n_\sigma}, \alpha_1, \ldots, \alpha_{n_\alpha})$

n_σ : TIPICAMENTE 1 OU n

n_α : ATÉ $\left(n - \frac{n}{2}\right)(n-1) = \frac{n(n-1)}{2}$

4.2 SELEÇÃO DOS PAIS

SORTEIO ALEATÓRIO, COM DISTRIBUIÇÃO UNIFORME, A PARTIR DA POPULAÇÃO COM μ INDIVÍDUOS.

4.3 RECOMBINAÇÃO (CROSSOVER)

DOIS PAIS $\longrightarrow$ UM FILHO. A RECOMBINAÇÃO É REPETIDA λ VEZES.

OPÇÕES

- RECOMBINAÇÃO LOCAL
  - RECOMBINAÇÃO INTERMEDIÁRIA : $z_i = (x_i + y_i)/2$ → P/ PARÂMETROS DE ESTRATÉGIA
  - RECOMBINAÇÃO DISCRETA : $z_i = x_i$ OU y_i, ALEATORIAMENTE → P/ VARIÁVEIS DO PROBLEMA
- RECOMBINAÇÃO GLOBAL

4.4 MUTAÇÃO

$$x_{k+1} = x_k + \underset{(\sigma)}{\varepsilon} R$$

R : VETOR COM VARIÁVEIS ALEATÓRIAS GAUSSIANAS DE MÉDIA ZERO E VARIÂNCIA UM.

ε : <u>NÃO</u> É ESCOLHIDO PELO USUÁRIO.

INDIVÍDUO $(x_1, \ldots, x_n, \underset{(\sigma)}{\varepsilon})$ É EFETIVAMENTE AVALIADO DUAS VEZES :

1. NA SELEÇÃO DE SOBREVIVENTES, COM BASE EM $J(x)$

2. INDIRETAMENTE, ATRAVÉS DA CAPACIDADE DE GERAR FILHOS: UM INDIVÍDUO (x_{k+1}, σ_{k+1}) CONTÉM (REPRESENTA) UM BOM x_{k+1}, QUE SOBREVIVEU À SELEÇÃO, E TAMBÉM UM BOM σ_{k+1} QUE FOI BEM-SUCEDIDO EM GERAR ESTE x_{k+1} A PARTIR DO x_k.

a) MUTAÇÃO NÃO-CORRELACIONADA, COM SÓ UM TAMANHO DE PASSO:

- $n_\sigma = 1$ (SÓ UM PARÂMETRO DE ESTRATÉGIA EM CADA INDIVÍDUO)
- VALORES "α" NÃO SÃO USADOS
- IMPLEMENTAÇÃO:

$$\sigma_{k+1} = \sigma_k e^{\tau N(0,1)}$$

$$x_{i,k+1} = x_{i,k} + \sigma_{k+1}(N(0,1))_i \quad (i = 1, \ldots, n)$$

SE $\sigma_{k+1} < \varepsilon_0 \Rightarrow \sigma_{k+1} = \varepsilon_0$

$N(0,1)$ = V.A. GAUSSIANA COM MÉDIA 0 E VARIÂNCIA 1, SORTEADA TODA VEZ QUE FOR NECESSÁRIO.

τ: ESCOLHIDO PELO USUÁRIO. NORMALMENTE: $\tau \propto \frac{1}{\sqrt{n}}$, SENDO n O COMPRIMENTO DO VETOR x

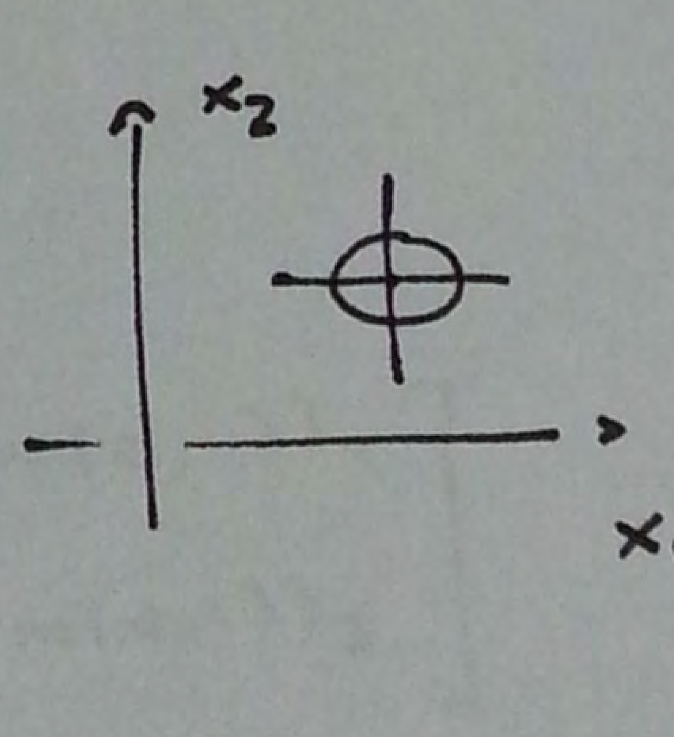
- MOTIVOS PARA USO DO FATOR $e^{\tau N(0,1)}$:

MODIFICAÇÕES MENORES $\longrightarrow$ MAIS FREQUENTES

DESVIOS-PADRÃO SÃO POSITIVOS

EM MÉDIA, A MUTAÇÃO É NEUTRA

- REPRESENTAÇÃO DE UM INDIVÍDUO:



$n = 2$

$n_\sigma = 1$

$n_\alpha = 0$

b) MUTAÇÃO NÃO-CORRELACIONADA, COM n TAMANHOS DE PASSO:

· $n_\sigma = n$ (UM PARÂMETRO DE PASSO PARA CADA INDIVÍDUO)

· VALORES "α" NÃO SÃO USADOS

· IMPLEMENTAÇÃO:

$$\sigma_{i,k+1} = \sigma_{i,k}\, e^{\tau_1 N(0,1)}\, e^{\tau_2 (N(0,1))_i} \qquad (i = 1, \dots, n)$$

$$x_{i,k+1} = x_{i,k} + \sigma_{i,k+1} (N(0,1))_i \qquad (i = 1, \dots, n)$$

SE $\sigma_{i,k+1} < \varepsilon_0 \Rightarrow \sigma_{i,k+1} = \varepsilon_0$

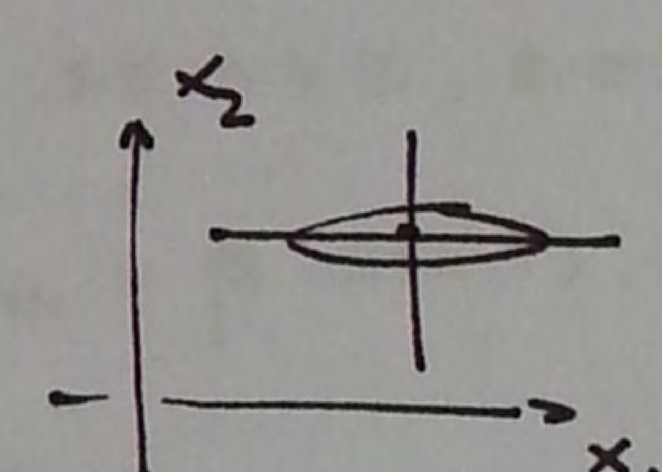
τ_1 E τ_2: ESCOLHIDOS PELO USUÁRIO. NORMALMENTE:

$$\tau_1 \propto \frac{1}{\sqrt{2n}} \quad \text{E} \quad \tau_2 \propto \frac{1}{\sqrt{2\sqrt{n}}}$$

· MOTIVOS PARA USO DO "FATOR DUPLO" $e^{\tau_1 N(0,1)} e^{\tau_2 (N(0,1))_i}$:

O PRIMEIRO FATOR É UMA BASE COMUM A TODAS AS MUTAÇÕES, ENQUANTO QUE O SEGUNDO FATOR PERMITE DIFERENTES ESTRATÉGIAS EM DIREÇÕES DIFERENTES.

· REPRESENTAÇÃO DE UM INDIVÍDUO:



$n = 2$

$n_\sigma = n = 2$

$n_\alpha = 0$

c) MUTAÇÕES CORRELACIONADAS:

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad ; \quad \text{PERTURBAÇÃO: } \begin{bmatrix} \sigma_1 (N(0,1))_1 \\ \sigma_2 (N(0,1))_2 \end{bmatrix}$$

MATRIZ DE COVARIÂNCIA (ROTAÇÃO α):

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

PERTURBAÇÃO "RODADA" POR UM ÂNGULO α:

$$x_{k+1} = x_k + \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \sigma_1 (N(0,1))_1 \\ \sigma_2 (N(0,1))_2 \end{bmatrix}$$

PARA UM NÚMERO DE COMPONENTES n QUALQUER:

$$C = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{bmatrix}$$

E $\alpha_{ij} = \alpha_{ji}$. POR ISSO, O NÚMERO DE COEFICIENTES α É $n_\alpha = \frac{n(n-1)}{2}$.

. IMPLEMENTAÇÃO:

$$\sigma_{i,k+1} = \sigma_{i,k} e^{\tau_1 N(0,1)} e^{\tau_2 (N(0,1))_i} \qquad (i = 1, \dots, n)$$

$$\alpha_{ij,k+1} = \alpha_{ij,k} + \text{"}\beta\text{"}(N(0,1))_{ij} \qquad (i=1,\dots,n ;\ \underline{\underline{j \geq i}})$$

O QUE EQUIVALE A: $\alpha_{j,k+1} = \alpha_{j,k} + \beta(N(0,1))_j$, SENDO α_j O j-ÉSIMO ÂNGULO DE ROTAÇÃO. SE $n=2$, ENTÃO SÓ HÁ UM ÂNGULO DE ROTAÇÃO: $\alpha_{k+1} = \alpha_k + \beta N(0,1)$.

$$x_{k+1} = x_k + \left[\begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{bmatrix} \right] \begin{bmatrix} \sigma_1 (N(0,1))_1 \\ \sigma_2 (N(0,1))_2 \\ \vdots \\ \sigma_n (N(0,1))_n \end{bmatrix}$$

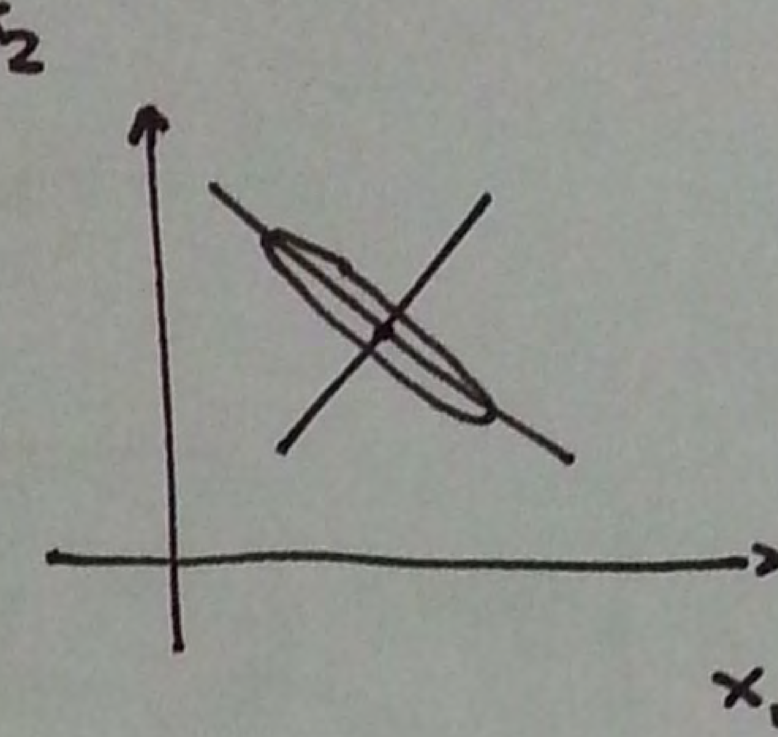
$\alpha_{11,k+1}$, $\sigma_{1,k+1}$

E ASSIM POR DIANTE...

PARÂMETROS ESCOLHIDOS PELO USUÁRIO:

$$\tau_1 = \frac{1}{\sqrt{2n}} \ ; \ \tau_2 = \frac{1}{\sqrt{2\sqrt{n}}} \ : \ \beta \simeq 5^\circ$$

· REPRESENTAÇÃO DE UM INDIVÍDUO:



$n = 2$; $n_\sigma = 2$; $n_\alpha = 1$

· ASSUMINDO QUE TODOS OS ÂNGULOS α_j DEVEM FICAR NO INTERVALO $[-\pi, \pi]$:

SE $|\alpha_{j,k+1}| > \pi \Rightarrow \alpha_{j,k+1} = \alpha_{j,k+1} - 2\pi \operatorname{sign}(\alpha_{j,k+1})$

OBSERVAÇÕES:

6. COMEÇAR COM MUTAÇÕES NÃO-CORRELACIONADAS E σ_1 ATÉ σ_n $\xrightarrow{\text{RESULTADOS BONS}}$ USAR SOMENTE UM σ, PARA ACELERAR ALGORITMO

$\xrightarrow{\text{RESULTADOS RUINS}}$ USAR MUTAÇÕES CORRELACIONADAS

4.5 SELEÇÃO DE SOBREVIVENTES:

POPULAÇÃO INICIAL: μ. SÃO CONHECIDOS OS "$J(\mu)$".

SÃO GERADOS λ FILHOS E AVALIADOS TODOS OS "$J(\lambda)$".

A GERAÇÃO SEGUINTE (PRÓXIMOS μ INDIVÍDUOS) É FORMADA DE MANEIRA <u>DETERMINÍSTICA</u> (POR RANKING, SEM SORTEIO) A PARTIR:

- DOS FILHOS (PREFERENCIAL)* (μ, λ) ($\lambda = 7\mu$: PRESSÃO SELETIVA ALTA)
- DOS PAIS <u>E</u> DOS FILHOS

* ESCAPAR DE MÍNIMOS LOCAIS; DESCARTAR SOLUÇÕES ANTIGAS, SE $J(x,t)$; DESCARTAR RAPIDAMENTE PARÂMETROS DE ESTRATÉGIA RUINS.

4.6 FUNÇÃO DE ACKLEY (EXEMPLO ES):

$$J(x) = 20 + e - 20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right)$$

$$J(x)\Big|_{x=0} = 0 \quad (\text{MÍNIMO GLOBAL})$$

REPRESENTAÇÃO —— DIRETA, EM PONTO FLUTUANTE (VETOR x)

MUTAÇÕES NÃO-CORRELACIONADAS

INDIVÍDUO: $(x_1, \ldots, x_n, \sigma_1, \ldots, \sigma_n)$

NESTE EXEMPLO, $n = 30$

SELEÇÃO DOS PAIS —— ALEATÓRIA, UNIFORME EM "μ"

RECOMBINAÇÃO —— DISCRETA NA PARTE x E INTERMEDIÁRIA GLOBAL NA PARTE σ.

MUTAÇÃO —— JÁ ESPECIFICADA, NA REPRESENTAÇÃO

SELEÇÃO DOS SOBREVIVENTES —— (μ, λ) COM $\mu = 30$ E $\lambda = 200$

INICIALIZAÇÃO ALEATÓRIA E INTERRUPÇÃO APÓS 2×10^5 AVALIAÇÕES DE J.

$x_i \in [-30, +30]$.