# Estrategias Evolutivas

# Resumen de Estrategias Evolutivas

- Desarrolladas: Alemania en los 70s
- Pioneros: I. Rechenberg, H.-P. Schwefel
- Generalmente applicadas a :
  - optimisacion numerica
- Caracteristicas principales:
  - rapidas
  - buenas para optimisacion de valores reales
  - en comparacion con otros AE hay (alguna) teoria
- Caracteristica especial:
  - autoadaptacion de parametros

# Tableau tecnico

| | |
|---|---|
| Representacion | Vectores reales |
| Recombinacion | Discreta o intermedia |
| Mutacion | Perturbacion Gausiana |
| Seleccion de padres | Uniforme al azar |
| Seleccion de sobrevivientes | $(\mu,\lambda)$ o $(\mu+\lambda)$ |
| Especialidad | auto-adaptacion de los pasos de mutacion |

# Ejemplo Introductorio

- Tarea: minimisar f : $R^n \rightarrow R$
- Algoritmo: “EE de dos miembros” usando
  - Vectoress en $R^n$ como cromosomas
  - |poblacion| = 1
  - solo mutacion, crea un hijo
  - Seleccion “golosa”

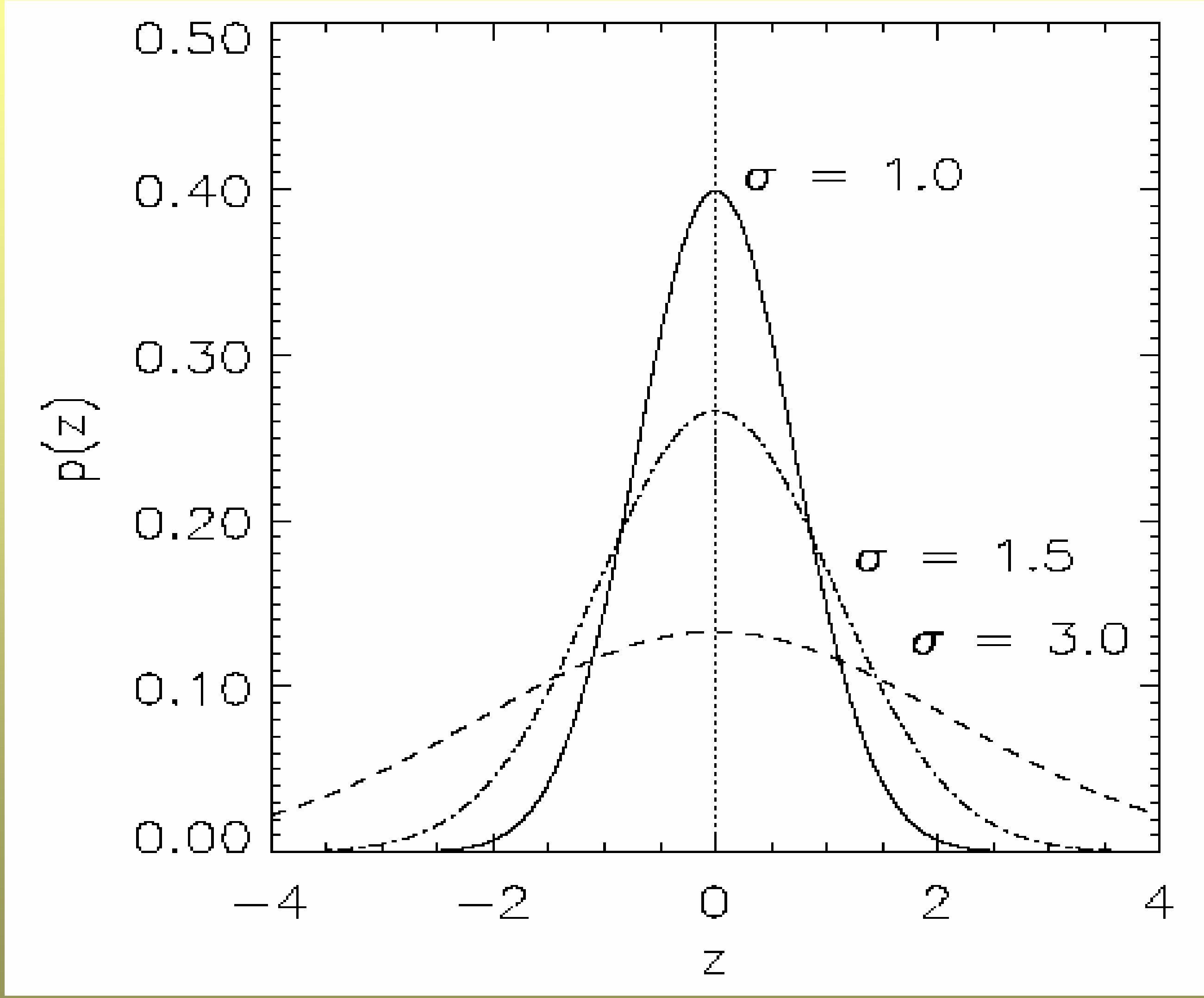
# pseudocodigo

- Setear $t = 0$
- Crear designoid inicial $x^t = \langle x_1^t,\ldots,x_n^t \rangle$
- Repetir haste (Cond *TERMIN*. satisfecha) hacer
- Tomar $z_i$ de una distribucion normal para todo $i = 1,\ldots,n$
- $y_i^t = x_i^t + z_i$
- Si $f(x^t) < f(y^t)$ Entonces $x^{t+1} = x^t$
  - Sino $x^{t+1} = y^t$
- Is
- Setear $t = t+1$
- recah

# Mecanismo de Mutacion

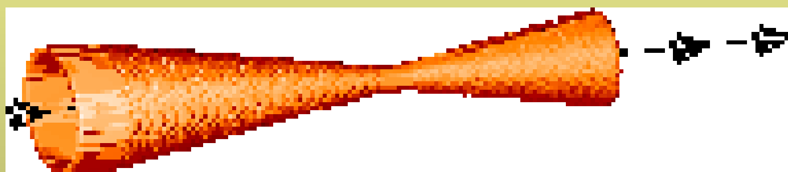
- valores z tomados de una distribucion normal $N(\xi,\sigma)$
  - esperanza $\xi$ se pone a 0
  - varianza $\sigma$ se llama el paso de mutacion
  - $\sigma$ varia on-line usando la regla “1/5 success rule”:
- Esta regla re-setea $\sigma$ despues de k iteraciones asi:
  - $\sigma = \sigma / c$ si $p_s > 1/5$
  - $\sigma = \sigma \cdot c$ si $p_s < 1/5$
  - $\sigma = \sigma$ if $p_s = 1/5$
- donde $p_s$ es el % de mutaciones exitosas, $0.8 \leq c \leq 1$

# La distribucion normal

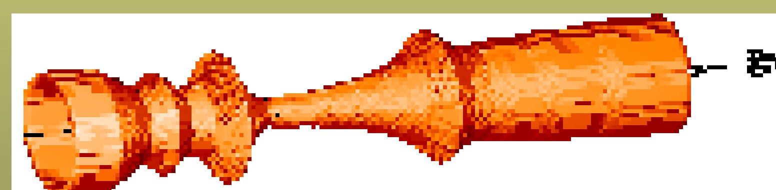


# Un Ejemplo Historico
# Jet Nozzle

Tarea: optimisar la forma del jet nozzle
Enfoque: mutaciones randomicas de la forma + seleccion



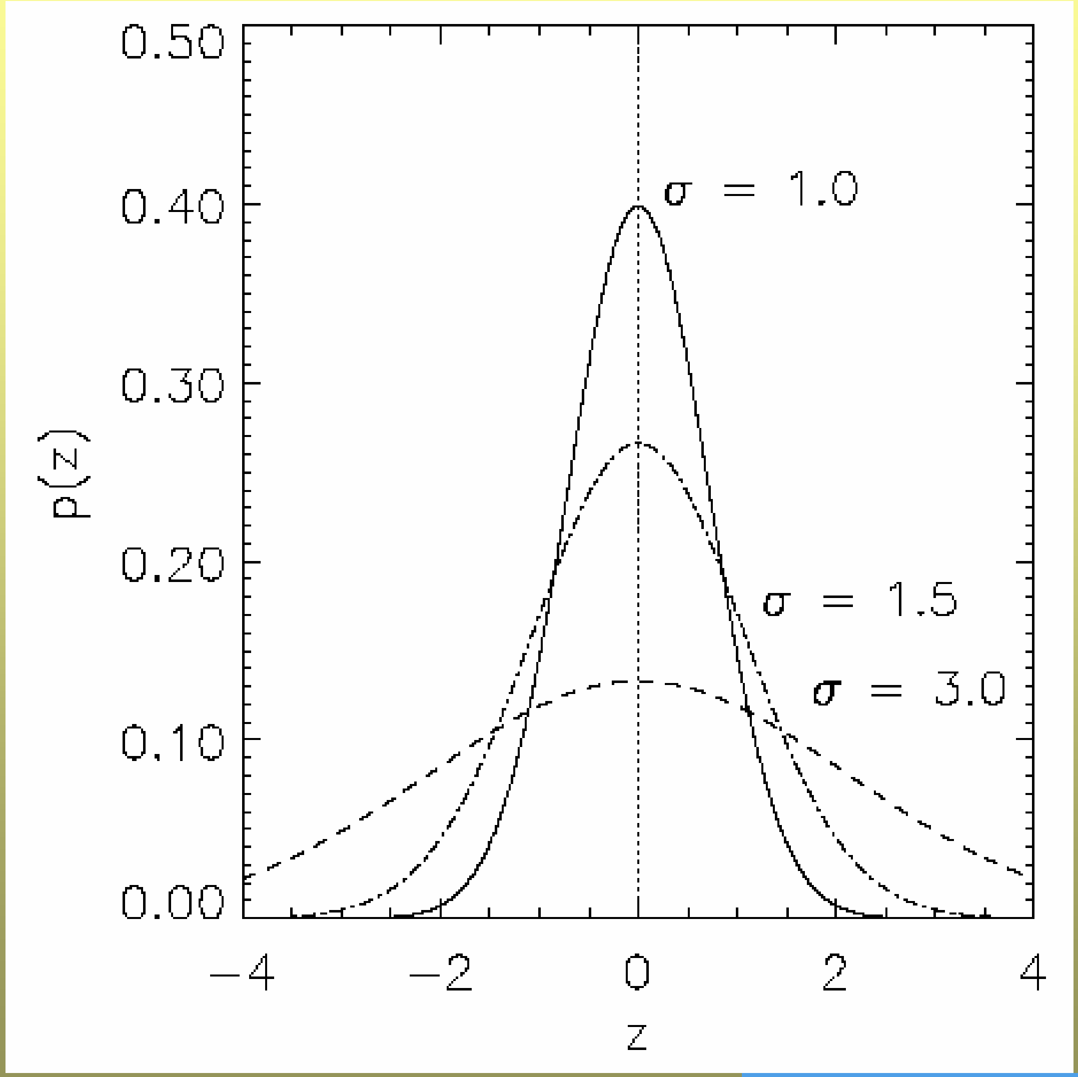
Forma inicial



Forma optimisada

# Operadores Geneticos:  Mutaciones (2)

El caso unidimensional



# Representacion

- El cromosoma consiste de tres partes:
  - variables del designoide: $x_1,\ldots,x_n$
  - parametros de la estrategia:
    - Tamanio de los pasos de mutacion: $\sigma_1,\ldots,\sigma_{n_\sigma}$
    - Direcciones de mutacion: $\alpha_1,\ldots, \alpha_{n_\alpha}$
- No siempre se usan todos los componentes
- Cromosoma completo: $\langle\, x_1,\ldots,x_n, \sigma_1,\ldots,\sigma_n ,\alpha_1,\ldots, \alpha_k \,\rangle$
- donde k = n(n-1)/2 (no. de pares i,j)

# Mutacion

- Mecanismo principal: cambiar valores al agregar ruido de una distribucion normal
- $x'_i = x_i + N(0,\sigma)$
- Idea clave:
  - $\sigma$ es parte del cromosoma $\langle x_1,\ldots,x_n, \sigma \rangle$
  - $\sigma$ tambien se muta a $\sigma'$ (ver mas abajo)
- Entonces: el tamanio de la mutacion co-evoluciona con la solucion

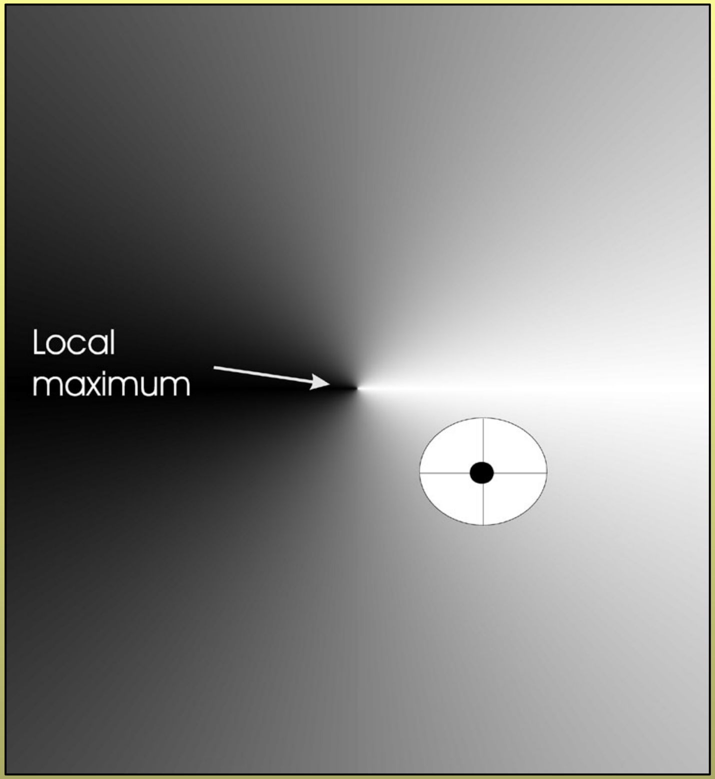
# Mutar σ primero

- Efecto neto de mutacion: $\langle x, \sigma \rangle \rightarrow \langle x', \sigma' \rangle$
- Orden es importante:
  - primero $\sigma \rightarrow \sigma'$ (ver abajo)
  - despues $x \rightarrow x' = x + N(0,\sigma')$
- Racional: nuevo $\langle x', \sigma' \rangle$ es evaluatdo dos veces
  - Primero: x' es bueno si f(x') es bueno
  - Segundo: σ' es bueno si el x' que creo es bueno
- Cambiando el orden no funcionaria

# Mutacion sin correlacion con un $\sigma$

- Cromosomas: $\langle\, x_1,\ldots,x_n,\ \sigma\, \rangle$
- $\sigma' = \sigma \cdot \exp(\tau \cdot N(0,1))$
- $x'_i = x_i + \sigma' \cdot N(0,1)$
- Tipicamente "learning rate" $\tau \propto 1/\, n^{½}$
- Con la condicion de borde $\sigma' < \varepsilon_0 \Rightarrow \sigma' = \varepsilon_0$

# Mutantes con Chances Iguales

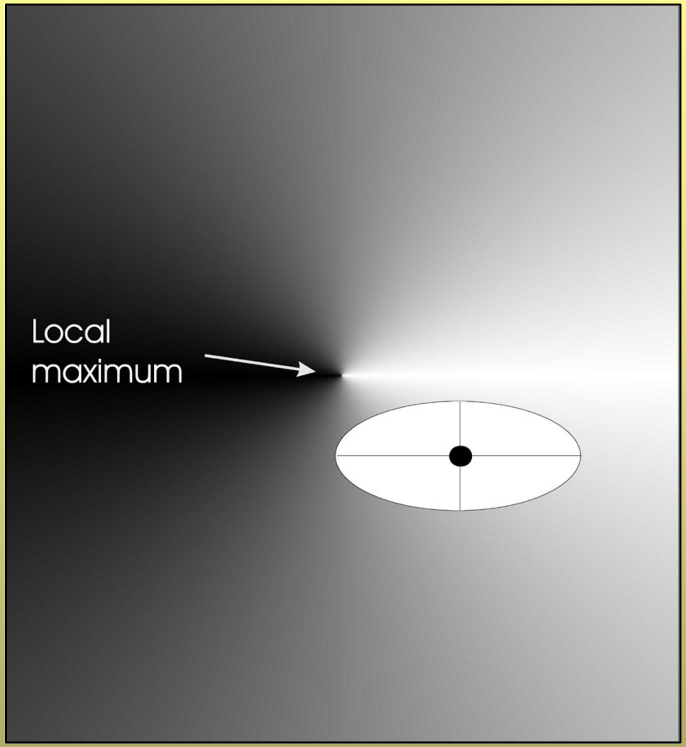



Circulo: los mutantes tienen las mismas chances de ser creados

# Mutacion sin correlacion con n $\sigma$'s

- Cromosomas: $\langle x_1,\ldots,x_n, \sigma_1,\ldots, \sigma_n \rangle$
- $\sigma'_i = \sigma_i \cdot \exp(\tau' \cdot N(0,1) + \tau \cdot N_i\,(0,1))$
- $x'_i = x_i + \sigma'_i \cdot N_i\,(0,1)$
- Dos learning rates (parametros):
  - $\tau'$ learning rate global
  - $\tau$ learning rate por coordenada
- $\tau \propto 1/(2\,n)^{½}$ y $\tau \propto 1/(2\,n^{½})^{½}$
- y $\sigma_i' < \varepsilon_0 \Rightarrow \sigma_i' = \varepsilon_0$

# Mutantes con Chances Iguales




Elipse:mutantes tienen las mismas chances de ser creados

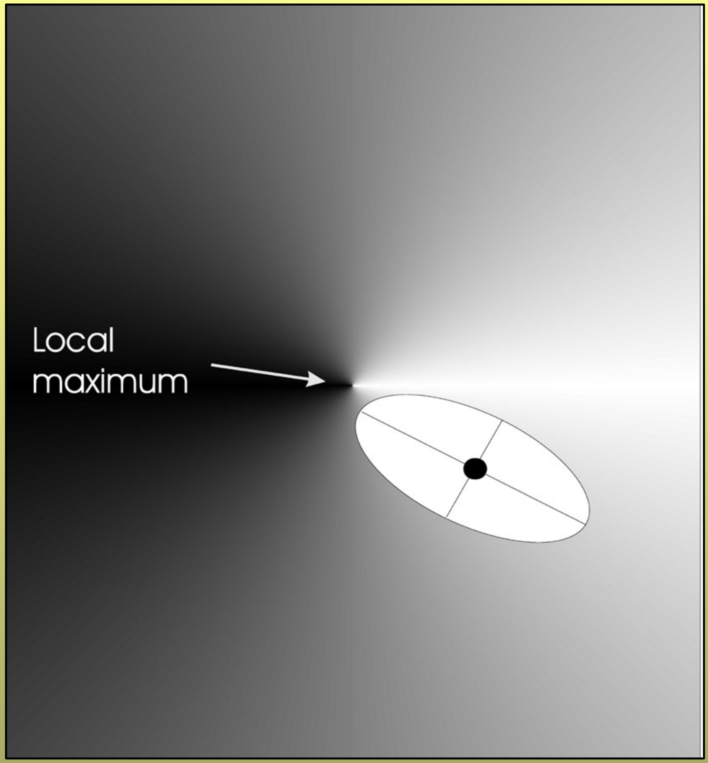
# Mutaciones Correlacionadas

- Cromosomas: $\langle x_1,\ldots,x_n, \sigma_1,\ldots, \sigma_n ,\alpha_1,\ldots, \alpha_k \rangle$
- donde $k = n \cdot (n-1)/2$
- y la matriz de co-varianza C se define como:
  - $c_{ii} = \sigma_i^2$
  - $c_{ij} = 0$ si i y j no estan correlacionadas
  - $c_{ij} = ½ \cdot ( \sigma_i^2 - \sigma_j^2 ) \cdot \tan(2\, \alpha_{ij})$ si i y j estan correlacionads
- Notar los indices de los α's

# La mutacion entonces es mas complicada:

- $\sigma'_i = \sigma_i \cdot \exp(\tau' \cdot N(0,1) + \tau \cdot N_i\,(0,1))$
- $\alpha'_j = \alpha_j + \beta \cdot N\,(0,1)$
- $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{N}(\boldsymbol{0},\boldsymbol{C}')$
  - $\boldsymbol{x}$ es el vector $\langle x_1,\ldots,x_n \rangle$
  - $\boldsymbol{C}'$ es la matriz de co-varianza $\boldsymbol{C}$ despues de la mutacion de los valores $\alpha$
- $\tau \propto 1/(2\,n)^{½}$ y $\tau \propto 1/(2\,n^{½})^{½}$ y $\beta \approx 5°$
- $\sigma_i' < \varepsilon_0 \Rightarrow \sigma_i' = \varepsilon_0$ y
- $|\,\alpha'_j\,| > \pi \Rightarrow \alpha'_j = \alpha'_j - 2\,\pi\,\mathrm{sign}(\alpha'_j)$

# Mutantes con Chances Iguales




Elipse: mutantes con chances iguales de creacion

# Recombinacion

- Crea un hijo
- Actua por variable/posicion:
  - Promediando el valor de los padres, o
  - Seleccciondo el valor de uno de los padres
- De dos o mas padres:
  - Usando dos padres para hacer un hijo
  - Seleccionando dos padres nuevos por cada variable

# Nombres de Recombinaciones

| | Dos padres fijos | Dos padres nuevos por cada i |
|---|---|---|
| $z_i = (x_i + y_i)/2$ | Intermediario local | Intermediario global |
| $z_i$ es $x_i$ o $y_i$ al azar | Local discreto | Global discreto |

# Seleccion de Padres

- Padres son seleccionados por medio de una distribucion uniforme

- En EE la seleccion de padres no es “biased”: todo individuo tiene las mismas probabilidades de ser seleccionado

# Seleccion de Sobrevivientes

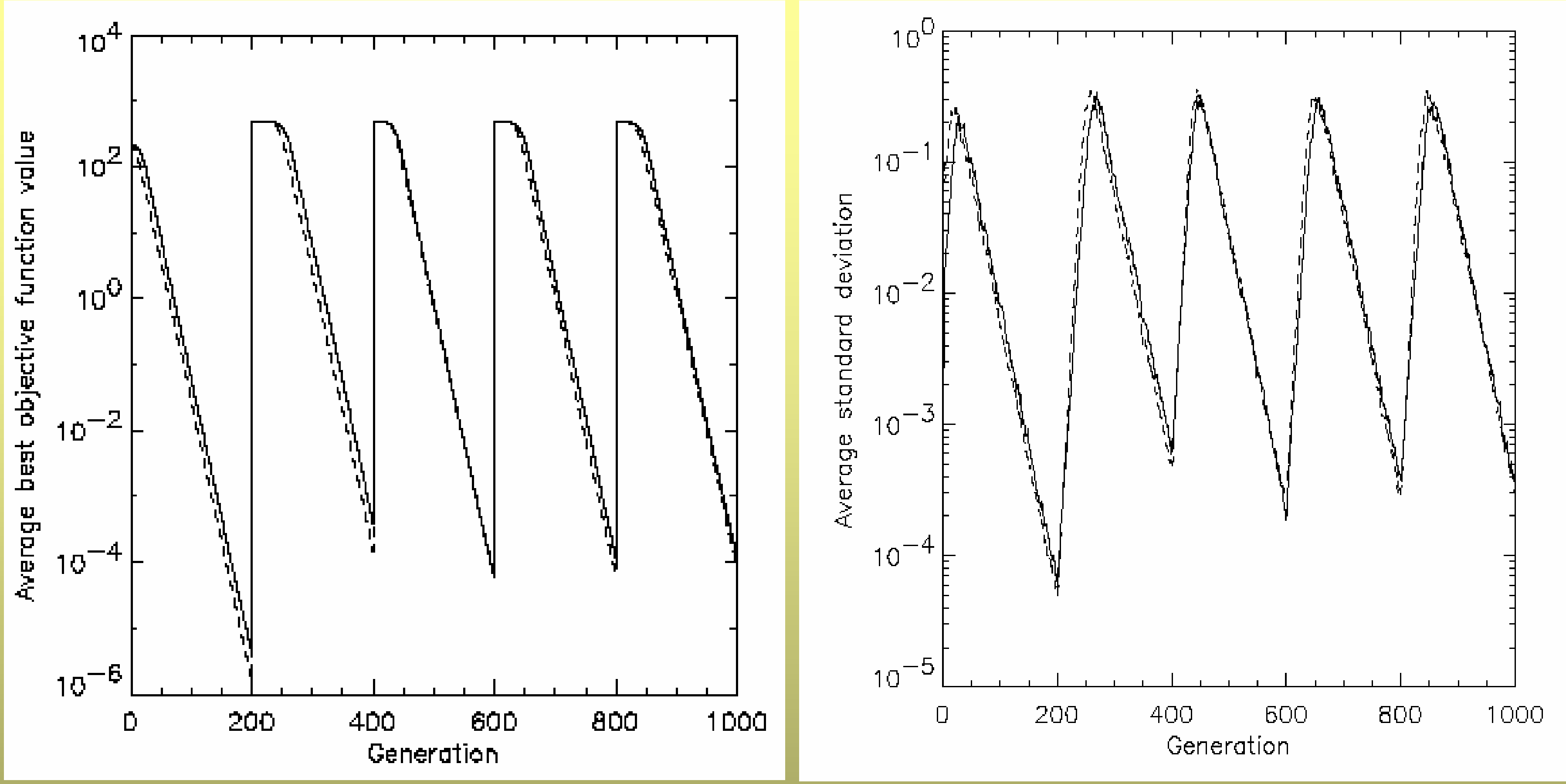
- Se aplica despues de crear $\lambda$ hijos a partir de $\mu$ padres por mutacion/recombinacion
- Deterministicamente elimina el material malo
- La base de la seleccion es:
  - El conjunto de los hijos solamente, seleccion $(\mu,\lambda)$
  - El conjunto de padres e hijos, seleccion $(\mu+\lambda)$

# Seleccion de Sobrevivientes (2)

- $(\mu+\lambda)$ es una seleccion elitista
- $(\mu,\lambda)$ puede perder los mejores individuos
- $(\mu,\lambda)$ se prefiere aveces por:
  - puede escapar optimos locales mas facilmente
  - puede seguir optimos locales en problemas dinamicos
  - Usando la estrategia “+” valores malos de $\sigma$ pueden sobrevivir mucho tiempo si $\langle x,\sigma \rangle$ es muy bueno, esto es x es fit.
- La presion selectiva en EE es usualmente muy alta ($\lambda \approx 7 \cdot \mu$ en general)

# Ilustracion de auto-adaptacion

- Dado un problema dinamico (el optimo se mueve cada 200 generaciones)
- Las EE pueden:
  - seguir el optimo
  - ajustar el valor de mutacion despues de cada cambio !




Cambios en fitness (izquierda) y en los pasos de mutacion (derecha)
* tratamos de minimizar fitness

# Prerequisitos para auto-adaptacion

- $\mu > 1$ deben tener varias estrategias
- $\lambda > \mu$ generar un “surplus” de hijos
- Presion selectiva alta pero no muy alta, e.g., $\lambda \approx 7 \cdot \mu$
- $(\mu,\lambda)$ preferible para poder eliminar los malos $\sigma$‘s
- Mezcla de los parametros de las estrategias por medio de crossover intermedio

# Ejemplo: El color del "cherry"

- Tarea: crea un color especifico, en este caso, el color del Cherry brandy
- Ingredientes: agua + colorantes rojo, amarillo, azul
- Representacion: $\langle$ w, r, y ,b $\rangle$ sin auto-adaptacion
- Valores scalados para dar un volumen especifico (30 ml)
- Mutacion: valores bajos / medios / altos de $\sigma$ usados con igual chances
- Seleccion: estrategia (1,8)

- Fitness: Estudiantes arman la mezcla y la comparan con el color deseado
- Criterio de finalizacion: Estudiantes satisfechos con el color obtenido
- La solucion se encuentra en mas o menos 20 generaciones
- Buena precision

# Otro Ejemplo: Ackley function

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

- Estrategia Evolutiva
  - Representacion:
    - $-30 < x_i < 30$
    - 30 tamanios de mutacion
  - (30,200) seleccion
  - Terminacion : luego de 200000 evaluaciones de fitness
  - Resultados: promedio mejor solucion es $7.48 \cdot 10^{-8}$ (muy bueno!)