



Seminario de Solución de Problemas de Inteligencia Artificial I: Estrategias Evolutivas

M.C. Jesús Hernández Barragán

UDG - CUCEI

Ciclo: 2018-B

Contenido

- Introducción
- (1+1)-ES
 - (1+1)-ES Adaptativo
- $(\mu + 1)$ -ES
- Recombinación
 - Recombinación sexual discreta
 - Recombinación sexual intermedia
 - Recombinación global discreta
- $(\mu + \lambda)$ -ES y (μ, λ) -ES

Introducción

Las Estrategias Evolutivas (ES) están basadas en el principio de evolución de las teorías de Darwin. Las principales operaciones de las ES son las siguientes:

- Mutación
- Recombinación
- Selección

La Mutación y la Selección se implementan diferente a los Algoritmos Genéticos. Además, en algunos casos la selección de individuos es elitista.

Generalmente, las ES están conformados de una población de padres e hijos donde las mejores soluciones se consideran para la siguiente generación.

Introducción (continuación)

Existen diferentes versiones de los ES, como por ejemplo:

- (1+1)-ES: Un padre genera un hijo, solo uno es elegido para la siguiente generación
- \bullet $(\mu+1)\text{-ES}:$ Una población de μ individuos genera un hijo, las mejores soluciones pasan a la siguiente generación
- $(\mu + \lambda)$ -ES: μ individuos generan λ hijos, las mejores soluciones pasan a la siguiente generación
- (μ, λ) -ES: μ individuos generan λ hijos, los padres no sobreviven para la siguiente generación
- CMA-ES y CMSA-ES: Utilizan una matriz de covarianza para mejorar las soluciones en cada iteración

(1+1)-ES

Es la ES mas simple de todas. Consiste en una población de un padre \mathbf{x}^p y un hijo \mathbf{x}^h , donde solo el mejor entre estos es seleccionado para la siguiente generación. El algoritmo (1+1)-ES es el siguiente:

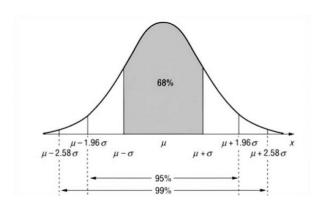
Algorithm 1 (1+1)-ES para resolver problemas de minimización.

- 1: $\sigma^2 \in \mathbb{R}^+ \leftarrow \text{inicializar varianza positiva}$
- 2: $\mathbf{x}^p \leftarrow \text{inicializar padre aleatoriamente}$
- 3: Hacer
- 4: $\mathbf{r} \leftarrow \text{generar vector aleatorio con } r_j \sim N\left(0, \sigma^2\right) \text{ para } j \in [1, D]$
- 5: $\mathbf{x}^h \leftarrow \mathbf{x}^p + \mathbf{r}$ (Mutación)
- 6: Si $f(\mathbf{x}^h) < f(\mathbf{x}^p)$ Entonces
- 7: $\mathbf{x}^p \leftarrow \mathbf{x}^h$
- 8: **Fin Si**
- 9: Mientras que se cumpla el total de generaciones

D indica el tamaño de la dimensión, f es la Función Objetivo y $N\left(0,\sigma^2\right)$ es una distribución normal. Además, $\mathbf{x}^p, \mathbf{x}^h \in \mathbb{R}^D$.

(1+1)-ES (continuación)

 $N\left(\mu,\sigma^2\right)$ es una distribución normal con media μ y desviación estándar σ . En las ES, es común utilizar la media seleccionada como $\mu=0$.



(1+1)-ES (continuación)

En esta ES, el padre genera al hijo mediante una mutación aleatoria que depende de la desviación estándar σ .

La varianza σ^2 es un parámetro de ajuste, donde:

- \bullet σ debería ser grande para generar mutaciones que exploren el espacio de búsqueda.
- \bullet σ debería ser pequeña para para explotar la solución del problema.

(1+1)-ES Adaptativo

Existe una versión (1+1)-ES Adaptativa que se encarga de adaptar σ en cada generación. Esto se logra utilizando una regla llamada 1/5.

La regla 1/5 establece lo siguiente:

En el (1+1)-ES, si el porcentaje de éxitos es menor que 1/5, entonces la desviación estándar σ debería disminuir. Si el cantidad de éxitos es mayor que 1/5, entonces σ debería incrementar.

8 / 20

$\overline{(1+1)\text{-ES Adapta}}$ tivo (continuación)

El algoritmo de (1+1)-ES Adaptativo esta descrito a continuación:

Algorithm 2 (1+1)-ES Adaptativo para resolver problemas de minimización.

- 1: $\sigma^2 \in \mathbb{R}^+ \leftarrow$ inicializar varianza positiva
- 2: $\mathbf{x}^p \leftarrow \text{inicializar padre aleatoriamente}$
- 3: $n_e \leftarrow 0$, para calcular porcentaje de éxitos
- 4: Hacer
- 5: $\mathbf{r} \leftarrow \text{generar vector aleatorio con } r_j \sim N\left(0, \sigma^2\right) \text{ para } j \in [1, D]$
- 6: $\mathbf{x}^h \leftarrow \mathbf{x}^p + \mathbf{r}$ (Mutación)
- 7: Si $f(\mathbf{x}^h) < f(\mathbf{x}^p)$ Entonces
- 8: $\mathbf{x}^p \leftarrow \mathbf{x}^h$
- 9: $n_e \leftarrow n_e + 1$
- 10: Fin Si
- 11: Etapa de ajuste de σ
- 12: Mientras que se cumpla el total de generaciones

D indica el tamaño de la dimensión, f es la Función Objetivo y $N\left(0,\sigma^2\right)$ es una distribución normal. Además, $\mathbf{x}^p, \mathbf{x}^h \in \mathbb{R}^D$.

(1+1)-ES Adaptativo (continuación)

Descripción de la etapa de ajuste de σ :

Algorithm 3 Etapa de ajuste de σ .

- 1: $\phi \leftarrow \frac{n_e}{t_i}$, para establecer el porcentaje de éxitos
- 2: Si $\phi < 1/5$ Entonces
- 3: $\sigma \leftarrow c^2 \sigma$
- 4: O si $\phi > 1/5$ Entonces
- 5: $\sigma \leftarrow \frac{\sigma}{c^2}$
- 6: **Fin Si**

donde c modifica la contribución de incrementos y decrementos de σ , y el valor t_i indica la generación actual. Se recomienda utilizar c=0.817.

(1+1)-ES Adaptativo (continuación)

Se puede concluir que:

- \bullet Si ϕ es más grande que 1/5, entonces se obtienen mutaciones pequeñas con pequeñas mejoras.
- \bullet Si ϕ es más chico que 1/5, entonces se generan grandes mutaciones pero con raras mejoras.

$$(\mu + 1)$$
-ES

En la estrategia ($\mu+1$)-ES, μ indica los padres utilizados para cada generación. Algunas características de esta estrategia se muestran a continuación:

- Cada padre es asociado con un vector σ que controla la magnitud de las mutaciones.
- Los padres se recombinan con otros padres para crear un hijo, el cual es mutado.
- \bullet Las mejores soluciones son seleccionadas entre los padres y el hijo, de tal forma que solo μ individuos permanecen para las siguientes generaciones.

$(\mu + 1)$ -ES (continuación)

A continuación se muestra el algoritmo ($\mu + 1$)-ES:

Algorithm 4 $(\mu + 1)$ -ES.

- 1: $\{(\mathbf{x}_i^p, \sigma_i^p)\}$ \leftarrow generar aleatoriamente $i \in [1, \mu]$ individuos tal que $\mathbf{x}_i^p \in \mathbb{R}^D$ y $\sigma_i^p \in \mathbb{R}^D$ con elementos positivos
- 2: Hacer
- Seleccionar aleatoriamente dos padres $\{(\mathbf{x}_{r_1}^p, \sigma_{r_1}^p)\}\ y\ \{(\mathbf{x}_{r_2}^p, \sigma_{r_2}^p)\}\$ Recombinar los padres para crear un hijo $\{(\mathbf{x}^h, \sigma^h)\}$ 3:
- 4:
- $\mathbf{r} \leftarrow \text{generar vector aleatorio con } r_j \sim N\left(0, \left(\sigma_j^h\right)^2\right) \text{ para } j \in [1, D]$ 5:
- $\mathbf{x}^h \leftarrow \mathbf{x}^h + \mathbf{r}$ 6:
- Eliminar al peor individuo de $\{(\mathbf{x}_1^p, \sigma_1^p), \cdots, (\mathbf{x}_{\mu}^p, \sigma_{\mu}^p), (\mathbf{x}^h, \sigma^h)\}$ 7:
- 8: Mientras que se cumpla el total de generaciones G

los padres r_1 y r_2 no necesariamente tienen que ser diferentes, aunque lo mejor es seleccionarlos tal que $r_1 \neq r_2$. D indica la dimensión del problema.

Recombinación (continuación)

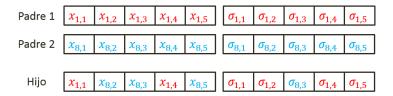
Existen diversión métodos de recombinación, como por ejemplo:

- Recombinación sexual discreta
- Recombinación sexual intermedia
- Recombinación global discreta

Es necesario notar que la recombinación se aplica a las soluciones de los padres, tanto como a la solución \mathbf{x}_i^p como a la desviación estándar σ_i^p .

Recombinación sexual discreta

Esta recombinación consiste en generar un hijo a partir de dos padres. Ejemplo:



Se muestra un problema de dimensión D=5, donde cada solución y desviación estándar del hijo es seleccionada aleatoriamente de los padres.

Recombinación sexual intermedia

Esta recombinación consiste en generar un hijo a partir de dos padres. Ejemplo:

Padre 1
$$x_{1,1}$$
 $x_{1,2}$ $x_{1,3}$ $\sigma_{1,1}$ $\sigma_{1,2}$ $\sigma_{1,3}$

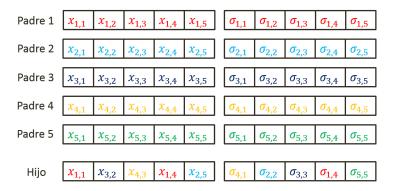
Padre 2 $x_{8,1}$ $x_{8,2}$ $x_{8,3}$ $\sigma_{8,1}$ $\sigma_{8,2}$ $\sigma_{8,3}$

Hijo $\frac{x_{1,1} + x_{8,1}}{2} \frac{x_{1,2} + x_{8,2}}{2} \frac{x_{1,3} + x_{8,3}}{2}$ $\frac{\sigma_{1,1} + \sigma_{8,1}}{2} \frac{\sigma_{1,2} + \sigma_{8,2}}{2} \frac{\sigma_{1,3} + \sigma_{8,3}}{2}$

Se muestra un problema de dimensión D=3, donde cada solución y desviación estándar del hijo es promediada respecto a los padres.

Recombinación global discreta

Esta recombinación consiste en generar un hijo a partir de todos los padres. Ejemplo:



Se muestra un problema de dimensión D=5 y población $\mu=5$, donde cada solución y desviación estándar del hijo es seleccionada aleatoriamente entre la población.

$(\mu + \lambda)$ -ES y (μ, λ) -ES

En la estrategia ($\mu + \lambda$)-ES, de una población de μ padres, se generan μ hijos. Se tiene en total de $\mu + \lambda$ individuos de los cuales solo los mejores μ individuos pasan a la siguiente generación.

En la estrategia (μ, λ) -ES, los padres μ para la siguiente generación son seleccionados entre los hijos λ actuales, es decir, ningún padre sobrevive. Por lo tanto, es necesario asegurarse que $\lambda > \mu$.

En ambas estrategias, se seleccionan los individuos con los mejores atributos, mientras que los demás mueren. Por esta razón, la selección de individuos es elitista.

$(\mu + \lambda)$ -ES y (μ, λ) -ES (continuación)

Algorithm 5 $(\mu + \lambda)$ -ES y (μ, λ) -ES.

```
1: \{(\mathbf{x}_i^p, \sigma_i^p)\} \leftarrow \text{generar aleatoriamente } i \in [1, \mu] \text{ individuos tal que } \mathbf{x}_i^p \in \mathbb{R}^D \text{ y}
\sigma_i^p \in \mathbb{R}^D \text{ con elementos positivos}
```

2: Hacer

6:

- 3: Desde i = 1 Hasta λ
- 4: Seleccionar aleatoriamente dos padres $\{(\mathbf{x}_{r_1}^p, \sigma_{r_1}^p)\}$ y $\{(\mathbf{x}_{r_2}^p, \sigma_{r_2}^p)\}$ 5: Recombinar los padres para crear un hijo $\{(\mathbf{x}_i^h, \sigma_i^h)\}$
 - $\mathbf{r}_i \leftarrow \text{generar vector aleatorio con } r_{ij} \sim N\left(0, \left(\sigma_{ij}^h\right)^2\right) \text{ para } j \in [1, D]$
- 7: $\mathbf{x}_i^h \leftarrow \mathbf{x}_i^h + \mathbf{r}_i$
- 8: Fin Desde
- 9: **En caso de** $(\mu + \lambda)$ -ES
- 10: Seleccionar los mejores μ individuos de:

11:
$$\left\{ \left(\mathbf{x}_{1}^{p}, \sigma_{1}^{p}\right), \cdots, \left(\mathbf{x}_{\mu}^{p}, \sigma_{\mu}^{p}\right), \left(\mathbf{x}_{1}^{h}, \sigma_{1}^{h}\right), \cdots, \left(\mathbf{x}_{\lambda}^{h}, \sigma_{\lambda}^{h}\right) \right\}$$

- 12: **En caso de** (μ, λ) -ES
- 13: Seleccionar los mejores μ individuos de:
- 14: $\left\{ \left(\mathbf{x}_{1}^{h}, \sigma_{1}^{h} \right), \cdots, \left(\mathbf{x}_{\lambda}^{h}, \sigma_{\lambda}^{h} \right) \right\}$
- 15: **Mientras** que se cumpla el total de generaciones G

Gracias por tu atención!

Información de contacto:

M.C. Jesús Hernández Barragán

 $\hbox{E-mail: jesus.hdez.barragan@gmail.com.}\\$