Introduction to Evolution Strategies

1. Evolution Strategies

Evolution Strategies (ESs) [1, 2] were initially developed by Rechenberg and Schwefel in 1964 with a strong focus on building systems capable of solving difficult real-valued parameter optimization problems. The natural representation was a vector or real-valued genes that were manipulated primarily by mutation operators designed to perturb the real-valued parameters in useful ways.

There are different kinds of ESs. Next we shall introduce the two of them most known.

1.1. The $(1+1)$-Evolution Strategy

The first ES algorithm, the so-called *(1+1)-ES*, was based on only two individuals per generation, one parent and one descendent. The parent string is evolved by applying a mutation operator to each one of its components. The mutation strength is determined by a value σ, a standard deviation of a normally distributed random variable. This parameter is associated to the parent and it is evolved in each process step as well. If the evolution has been performed successfully, then the descendent substitutes the parent in the next generation. The process is iterated until a certain finishing condition is satisfied.

The mutation operator **mut** has two components. The first one, $\mathbf{mu}_\sigma$, evolves the value of the standard deviation σ using Rechenberg's 1/5-success rule:

$$\sigma' = \mathbf{mu}_\sigma(\sigma) = \begin{cases} \frac{\sigma}{\sqrt[n]{c}}, & \text{if } p > \frac{1}{5} \\ \sigma \cdot \sqrt[n]{c}, & \text{if } p < \frac{1}{5} \\ \sigma, & \text{if } p = \frac{1}{5} \end{cases}$$

where p is the relative frequency of succesful mutations and c is a constant determining the updating amount of σ.

The second one, $\mathbf{mu}_x$, mutates each component in the real coded string by adding normally distributed variations with standard deviation σ' ($z_i \sim N_i(0, \sigma'^2)$) to it:

$$x' = \mathbf{mu}_x(x) = (x_1 + z_1, \ldots, x_n + z_n)$$

2. The (μ, λ)-Evolution Strategy

This second kind of ES is based on performing evolution on a population of μ possible n-dimensional solutions, obtaining λ offspring and selecting the best μ from them to form the new population.The offspring are obtained by first recombining a single or some parents in a single n-dimensional vector of object variables, and then creating a new one from this by applying mutations with identical or different standard deviations to each object variable. The main quality of the algorithm is its ability to incorporate the most important parameters from the strategy (standard deviations and correlation coefficients of normally distributed mutations) into the search process, such that adaption also takes place in the strategy parameters according to the current local topology of the search space. This property is called *self-adaptation* [1].

Therefore, each population individual consists of three vectors, $\vec{a} = (\vec{x}, \vec{\sigma}, \vec{\alpha})$, representing, respectively, the object variable, the standard deviation and the rotation angle values. The vector $\vec{x}$ has n dimensions, equal to the number of problem variables. The n_σ dimensions of a vector $\vec{\sigma}$ can be up to n (in this case, each object variable x_i, $i = 1, \ldots, n$, has associated a different step size σ_i), and n_α can be up to $\frac{(2 \cdot n - n_\sigma) \cdot (n_\sigma - 1)}{2}$. The set of strategy parameters consisting of standard deviations and rotation angles provides a complete description of the generalized n-dimensional distribution with an expectation value vector $\vec{0}$. Anyway, n_α may be set to zero, indicating that the rotation angles are not considered. The more usual values for n_σ and n_α are the following [1]: $(n_\sigma, n_\alpha) = \{(1, 0), (n, 0), (n, \frac{n \cdot (n-1)}{2}), (2, n-1)\}$

The following algorithm generically describes the behavior of the (μ, λ)-ES. The parameter t stands for the number of the current generation and $P(t)$ for the population in it:

1. Initialize and evaluate $P(0)$. Initialize $t \leftarrow 0$
2. Recombine ζ of the μ individuals of $P(t)$ λ times, by using one of the following gene recombination mechanisms, $r \in \{0, 1, 2, 3\}$, $i = 1, \ldots, n + n_\sigma + n_\alpha$:

$$a'_i = \begin{cases} a_{S,i};\, S \sim U(\{1, \ldots, \zeta\}) \text{ equal } \forall i & r = 0, \text{ no recombination} \\ \frac{\sum_{j=1}^{\zeta} a_{j,i}}{\zeta} \quad r = 1, \text{ global intermediary} \\ u \cdot a_{S,i} + (1 - u) \cdot a_{T,i}; \quad S, T \sim U(\{1, \ldots, \zeta\}) & r = 2, \text{ local intermediary} \\ a_{S,i};\; S \sim U(\{1, \ldots, \zeta\}) \quad r = 3, \text{ discrete} \end{cases}$$

   This operation generates λ individuals forming $P'(t)$.
3. Mutate $P'(t)$ by adapting the λ individuals to obtain λ offspring forming $P''(t)$ in the way:

   3.1. Mutate the values of $\vec{\sigma}'$ to obtain the array $\vec{\sigma}''$:

$$\vec{\sigma}'' = (\sigma'_1 \cdot \exp(z_1 + z_0), \ldots, \sigma'_{n_\sigma} \cdot \exp(z_{n_\sigma} + z_0)$$

   where $z_i \sim N(0, \frac{1}{\sqrt{2 \cdot \sqrt{n}}}^2)$, $i = 1, \ldots, n_\sigma$ and $z_0 \sim N(0, \frac{1}{\sqrt{2 \cdot n}}^2)$.

   3.2. Mutate the values of $\vec{\alpha}'$ to obtain the vector $\vec{\alpha}''$:

$$\vec{\alpha}'' = (\alpha'_1 + z_1, \ldots, \alpha'_{n_\alpha} + z_{n_\alpha})$$

   where $z_i \sim N(0, 0.0873^2)$, $i = 1, \ldots, n_\alpha$.

   3.3. Mutate the values of $\vec{x}'$ to obtain the vector $\vec{x}''$:

$$\vec{x}'' = (x'_1 + cor_1(\vec{\sigma}'', \vec{\alpha}''), \ldots, x'_n + cor_n(\vec{\sigma}'', \vec{\alpha}''))$$

   where $cor(\vec{\sigma}'', \vec{\alpha}'')$ is a normally distributed random vector of correlated values.
4. Evaluate $P''(t)$ and select the best μ individuals to form $P(t + 1)$.
5. Set the counter of generations $t \leftarrow t + 1$
6. If not (termination condition), then go to 2, else Stop.

For more information about (μ, λ)-ES refer to [1].

References

[1] T. Bäck, *Evolutionary alghorithms in theory and practice*, Oxford University Press, 1996.

[2] H. P. Schwefel, *Evolution and optimum seeking.Sixth-Generation Computer Technology Series*, John Wiley and Sons, 1995.