Understanding working of Differential Evolution using step by step procedure

The working procedure of differential evolution (DE) can be explained in 4 parts, namely problem definition, DE parameters, initialization and DE position update. DE is explained here by considering a simple optimization problem (sphere function) with two decision variables. There are many variants of DE which are called DE strategies. In this example, only DE/1/rand/bin strategy is considered for explanation purpose.

1) Problem Definition

Consider an optimization problem

$$\text{Minimize } f(x) = x_1^2 + x_2^2 \qquad -5 \leq x_1, x_2 \leq 5$$

2) Differential Evolution and Problem Parameters

Population size $N = 5$
Dimension of the problem = 2
Stopping Criteria = Maximum Number of iterations = 2
Scaling Factor $F = 0.5$
Crossover Probability $P_{cr} = 0.7$

3) Initialization

Position	**Function Value**
x_1 = (1.7667, -4.1337)	20.2089
x_2 = (4.8071, 0.6642)	23.5502
x_3 = (-2.9232, -4.0439)	24.8983
x_4 = (-4.3747, -4.7421)	41.6270
x_5 = (-1.6587, 0.5680)	3.07408

4) Differential Evolution Position Update

Strategy: DE/rand/1/bin

In this notation, '*DE*' stands for differential evolution, '*rand'* makes the random selection of the target vector, *'1'* is for the number of differentials while *'bin'* says about the type of crossover which is binary crossover here.

Generation – 1

In DE position update is carried out by two operators, mutation and crossover. It is different from the evolutionary algorithms where the solution is mutated after the crossover. But here mutation is first operator and then crossover is applied.

Mutation

In differential evolution, the mutation operator is used to create a ***Trial Vector*** (say u) for each solution (***Parent Vector***). It is done by mutating a ***Target Vector*** (selection of ***Target Vector*** depends upon different strategies of DE).

$$\boldsymbol{Trial\,Vector} = \boldsymbol{Target\,Vector} + \boldsymbol{Scale\,Factor} \times (Randomly\ selected\ solution_1 - Randomly\ selected\ solution_2)$$

Here it should be noted that $Randomly\ selected\ solution_1$, $Randomly\ selected\ solution_2$, $Target\ Vector$ and *Parent Vector* should be different be different with each other . $Scale\ Factor$ is a user defined parameter $\in (0, \infty)$. The recommended value of this parameter is 0.5.

Now to begin with mutation, select the first solution as the *Parent Vector*.
Say $x_1 = (1.7667, \ -4.1337)$ is the Parent Vector.
Corresponding function value is $f(x_1) = 20.2089$

For mutation, let *Target Vector* (selected randomly from the current population) is
$x_4 = (-4.3747, -4.7421)$

$Randomly\ selected\ solution_1 = x_5$
$Randomly\ selected\ solution_2 = x_3$

Let $Scale\ Factor = 0.5$

Now the Trial Vector u_1 is calculated as below:

$$\begin{aligned}
u_{11} &= x_{41} + 0.5 \times (x_{51} - x_{31}) \\
&= -4.3747 + 0.5 \times (-1.6587 - (-2.9232)) \\
&= -3.7425 \\
u_{12} &= x_{42} + 0.5 \times (x_{52} - x_{32}) \\
&= -4.7421 + 0.5 \times (0.5680 - (-4.0439)) \\
&= -2.4362
\end{aligned}$$

Trial Vector $u_1 = (-3.7425, \ -2.4362)$

At this point, we will check whether the *Trial Vector* u_1 is within the search space or not. As we see, $-5 \le x'_{11}, x'_{12} \le 5$, we will accept it.

This completes the mutation process.

Crossover

In the DE crossover an *Offspring* is generated using the discrete recombination of *Parent Vector* and *Trial Vector*. Consider the crossover probability $P_{cr} = 0.7$. Since we are going to calculate *Offspring* corresponding to the *Parent Vector* x_1, so let us denote the *Offspring* by x'_1.

Offspring $x'_{1j} = \begin{cases} u_{1j}, & if\ j \in I \\ x_{1j}, & otherwise \end{cases}$

Here I is the set of crossover points which depends upon the choice of crossover type and the crossover probability. Binomial crossover is quite popular in DE which is explained below:

In Binomial crossover, crossover points are selected from the set {1, 2, ……, problem dimension} with probability P_{cr}. This may lead to the situation when no point is selected for I (more probable for the small dimension problems, e.g. this problem as it is only 2 dimensional problem). If this happens, i.e. $I = \phi$ then there will not be any change in the *Offspring* and it will be same as parent. To avoid this situation, I is always considered as a non-empty set by including a random point (say j_0) from the set {1, 2, ……, problem dimension} initially.

So let I ={2}

Now other crossover points are selected using following algorithm:

$$for\ \ j\ \in\ \{1,2,\ldots,problem\ dimension\}$$

$$if\ rand(0,1) < P_{cr}\ and\ j \neq j_0$$
$$I = I\ \cup\ \{j\};$$
$$end$$

For this example, let $rand\ (0, 1)\ =\ 0.67$. Then I = *{1, 2}*.

Now the *Offspring* x'_1 will be formed from *Parent Vector* $x_1\ =\ (1.7667,\ \ -4.1337)$ and *Trial Vector* $u_1 = (-3.7425,\ -2.4362)$. Obviously, since the set of crossover points I has both the dimensions, both the variables of the *Offspring* x'_1 is are selected from the *Trial Vector* $u_1 = (-3.7425,\ -2.4362)$.

Thus the *Offspring* $x'_1 = (-3.7425,\ -2.4362)$.

At this point of time, we will decide whether in the population *Parent Vector* x_1 or *Offspring* x'_1 will survive. For that we will compare the function values of x_1 and x'_1.

Since $f(x'_1) = 19.9417$ which is better than $f(x_1)\ =\ 20.2089$, therefore for next generation, in the population *Offspring* x'_1 will survive and *Parent Vector* x_1 will die out or in other words, x'_1 will replace x_1. Thus **new** $\boldsymbol{x_1} = (\mathbf{-3.7425},\ \mathbf{-2.4362})$.

The same procedure will be applied to next solution $x_2 = (4.8071, 0.6642)$with corresponding function value 23.5503.

Mutation

Target Vector = x_5

$Randomly\ selected\ solution_1 = x_3$

$Randomly\ selected\ solution_2 = x_1$

After applying the same procedure, the *Trial Vector* $u_2 = (-1.2491, -0.2358)$ which will be accepted as both the variables are within the search space.

Crossover
Let I ={1, 2} using the binomial crossover.

Therefore, *Offspring* $x_2' = (-1.2491, -0.2358)$. $f(x_2') = 1.6158$.

Clearly, $f(x_2') < f(x_2)$ so the **new** $\boldsymbol{x_2} = (\mathbf{-1.2491, -0.2358})$

For solution $\boldsymbol{x_3}$

Mutation
Target Vector = x_4
$Randomly\ selected\ solution_1 = x_1$
$Randomly\ selected\ solution_2 = x_5$

After applying the same procedure, the *Trial Vector* $u_3 = (-5.4167,\ -6.2443)$. Now since u_{31} and u_{32} both are beyond the search space therefore we will pull the solution to the corresponding boundary of the search space, i.e.

Trial Vector $u_3 = (-5,\ -5)$.

Crossover
Let I ={1, 2} using the binomial crossover.
-2.9232, -4.0439
Therefore, *Offspring* $x_3' = (-5, -5)$. $f(x_3') = 50$.

Clearly, $f(x_3') > f(x_3)$ so the **new** $\boldsymbol{x_3} = (\mathbf{-2.9232, -4.0439})$ which is the *Parent Vector*.

Applying the same procedure to remaining two solutions, we get

new $\boldsymbol{x_4} = (\mathbf{-4.3748,\ -4.3402})$ **with** $\boldsymbol{f(x_4)} = \mathbf{37.9765}$

new $\boldsymbol{x_5} = (\mathbf{-1.6587, 0.5681})$ **with** $\boldsymbol{f(x_5)} = \mathbf{3.0741}$

After first generation, the updated population is

Position	Function Value
x_1 = (-3.7425,-2.4362)	19.9417
x_2 = (-1.2491,-0.2358)	1.6158
$x_3 = (-2.9232, -4.0439)$	24.8982
x_4 = (-4.3748,-4.3402)	37.9765
x_5 = (-1.6587,0.5681)	3.0741

Best solution after first generation is x_2 = (-1.2491,-0.2358) with $f(x_2) = 1.6158$.

Generation -2

Position	**Function Value**
$x_1 = (-1.3603\ -1.9917)$	5.8173
$x_2 = (-1.2491, -0.2358)$	1.6158
$x_3 = (-2.9232, -4.0439)$	24.8982
$x_4 = (0.5233, -0.0876)$	0.2815
$x_5 = (-0.4676, 0.7903)$	0.8433

It can be observed that solutions x_1, x_4 and x_5 are modified. Also after second generation the best solution is x_4=(0.5233,-0.0876) with $f(x_4)$= 0.2815 which is better than that of first generation.

Reference:

Storn, R. and Price, K., 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization, 11(4), pp.341-359.
