

Wprowadzenie do Sztucznej Inteligencji
Laboratorium 2

Algorytmy ewolucyjne

Piotr Niedziątek

304474

1. Wstęp

Algorytm ewolucyjny jest to algorytm metaheurystyczny inspirowany biologiczną ewolucją. Stosuje się go do zadań optymalizacyjnych. Ewolucji podlegają osobniki, czyli na początku losowe punkty z przestrzeni stanów. Osobniki składają się na populację, mogą podlegać mutacji, czyli niewielkim zmianom oraz krzyżowaniu – mieszaniu między sobą. Ważne jest, aby w algorytmie promowane były osobniki najlepsze, dlatego do ich oceny stosuje się funkcję dopasowania – im większa jej wartość tym lepszy osobnik. Reprodukacja ma za zadanie wybrać lepsze osobniki z większym prawdopodobieństwem, natomiast sukcesja decyduje, które osobniki przeżywają do następnej populacji. Poniżej przedstawiono pseudokod podstawowego algorytmu ewolucyjnego:

```
Data:  $q(x), P_0, \mu, \sigma, p_c, t_{max}$ 
Result:  $\hat{x}^*, \hat{o}^*$ 
1 begin
2    $t \leftarrow 0$ 
3    $o \leftarrow \text{ocena}(q, P_0)$ 
4    $\hat{x}^*, \hat{o}^* \leftarrow \text{znajdź\_najlepszego}(P_0, o)$ 
5   while nie spełnione kryterium stopu(  $t, t_{max}, P_t, o$  ) do
6      $R \leftarrow \text{reprodukcja}(P_t, o, \mu)$ 
7      $M \leftarrow \text{operacje\_genetyczne}(R, \sigma, p_c)$ 
8      $o_m \leftarrow \text{ocena}(q, M)$ 
9      $x_t^*, o_t^* \leftarrow \text{znajdź\_najlepszego}(M, o_m)$ 
10    if  $o_t^* < \hat{o}^*$  then
11       $\hat{o}^* \leftarrow o_t^*$ 
12       $\hat{x}^* \leftarrow x_t^*$ 
13    end
14     $P_{t+1}, o \leftarrow \text{sukcesja}(P_t, M, o, o_m)$ 
15     $t \leftarrow t + 1$ 
16  end
17 end
```

2. Zadanie

Celem laboratorium była implementacja algorytmu ewolucyjnego oraz przetestowanie go na tych samych funkcjach co w laboratorium 1, czyli funkcji Rastrigina i Griewanka:

- Funkcja Rastrigina:

$$f(x) = 20 + x_1^2 + x_2^2 - 10 \cdot (\cos(2 \cdot \pi \cdot x_1) + \cos(2 \cdot \pi \cdot x_2))$$

Zakres:

$$x \in [-5.12, 5.12], x \in R^2$$

- Funkcja Griewanka:

$$f(x) = \frac{x_1^2}{4000} + \frac{x_2^2}{4000} - \cos(x_1) \cdot \cos\left(\frac{x_2}{\sqrt{2}}\right) + 1$$

Zakres:

$$x \in [-50, 50], x \in R^2$$

3. Rozwiązanie

W implementacji zastosowano reprodukcję, mutację, krzyżowanie oraz sukcesję.

- Funkcja dopasowania została zaimplementowana jako odwrotność badanej funkcji, czyli:

$$f_d(x) = \frac{1}{f(x_1, x_2)}$$

- Wybrana została reprodukcja turniejowa. Z jednakowym prawdopodobieństwem, z powtórzeniami wybierane są 2 osobniki, następnie najlepszy z nich jest zwracany jako zwycięzca turnieju. Turnieje powtarzane są tyle razy, ile jest osobników w populacji.
- Mutacja danego osobnika odbywa się z prawdopodobieństwem p_m . Polega ona na dodaniu losowego szumu do wartości współrzędnych z rozkładu normalnego:

$$x^{t+1} = x^t + N(0, \sigma^2)$$

- Zastosowano krzyżowanie ważone uśrednione z losowymi wagami. Krzyżowanie odbywa się z prawdopodobieństwem p_k . Jeśli osobnik został wybrany do krzyżowania, drugi osobnik jest losowany z całej populacji.

$$k[i] = w_i \cdot r_1[i] + (1 - w_i) \cdot r_2[i]$$

Gdzie:

w_i – losowana waga z rozkładu jednorodnego od 0 do 1

i – indeks cechy

- Zaimplementowano sukcesję elitarną z elitą o rozmiarze k . W tej sukcesji do następnego pokolenia trafia na pewno k najlepszych osobników z bieżącej populacji. Od sumy tych osobników z osobnikami po mutacji i krzyżowaniu odejmuje się k najgorszych osobników ze zbioru połączonego:

$$P_{t+1} = \{k \text{ najlepszych z } P_t\} \cup M\{k \text{ najgorszych z połączonego zbioru}\}$$

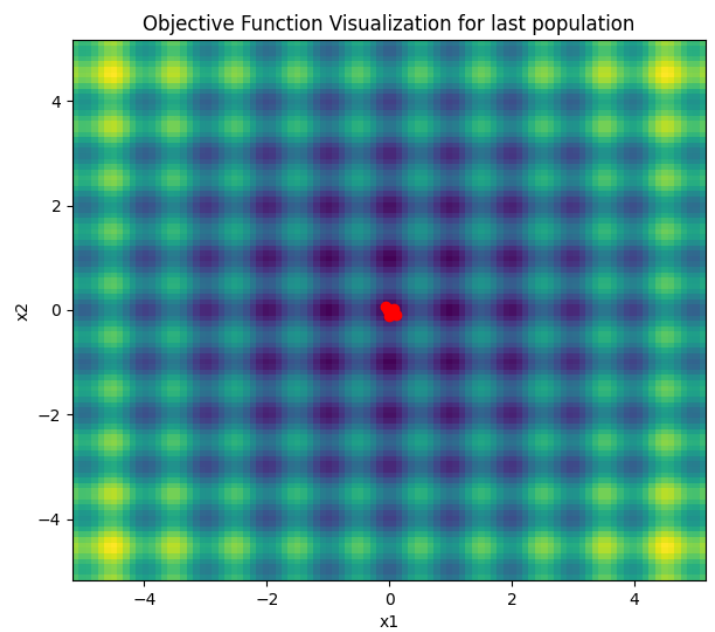
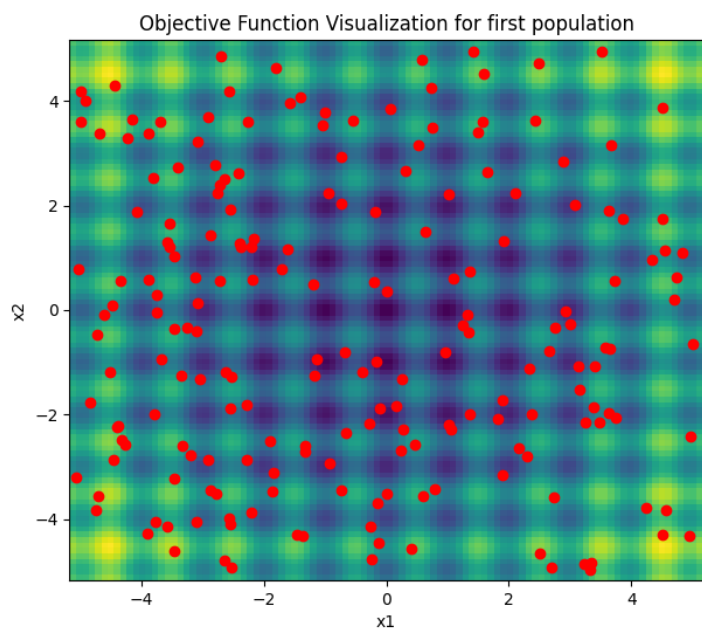
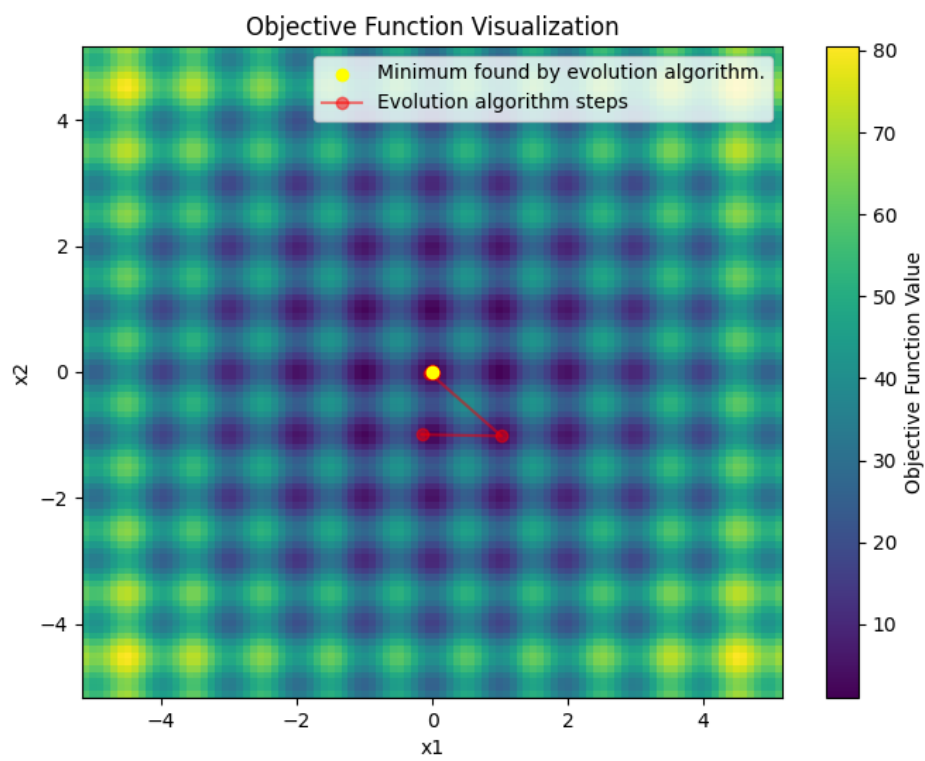
4. Wyniki

4.1. Wizualizacja kolejnych znalezionych najlepszych rozwiązań oraz porównanie pierwszej i ostatniej populacji.

- Funkcja Rastrigina

Dobre zostały następujące parametry algorytmu:

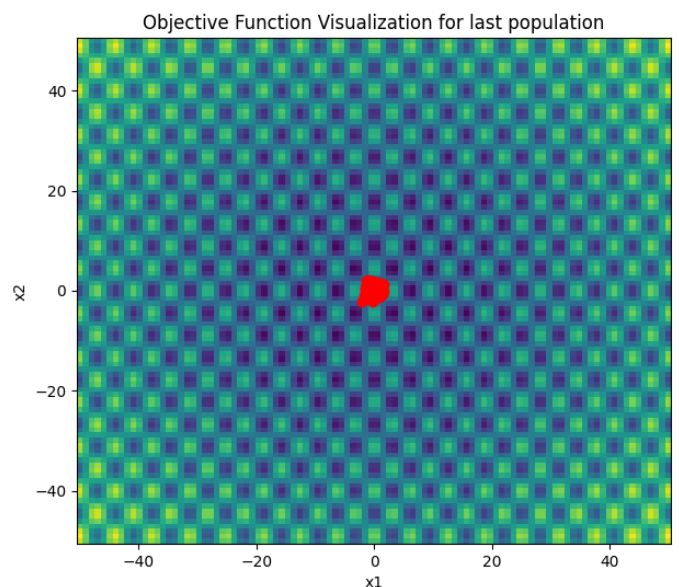
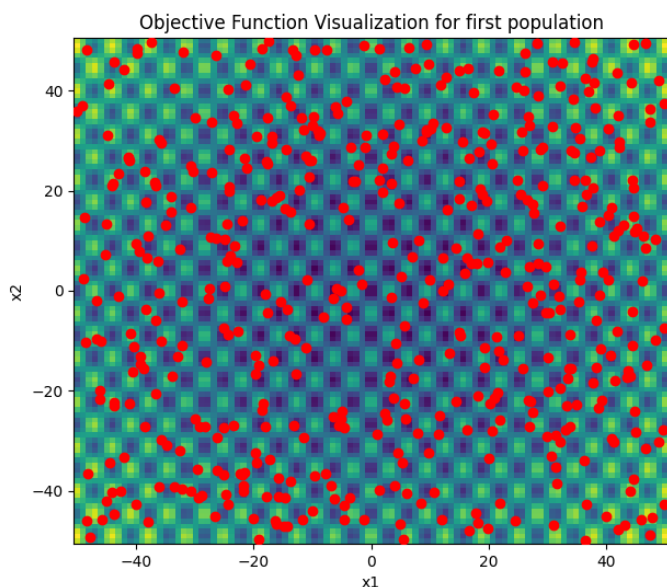
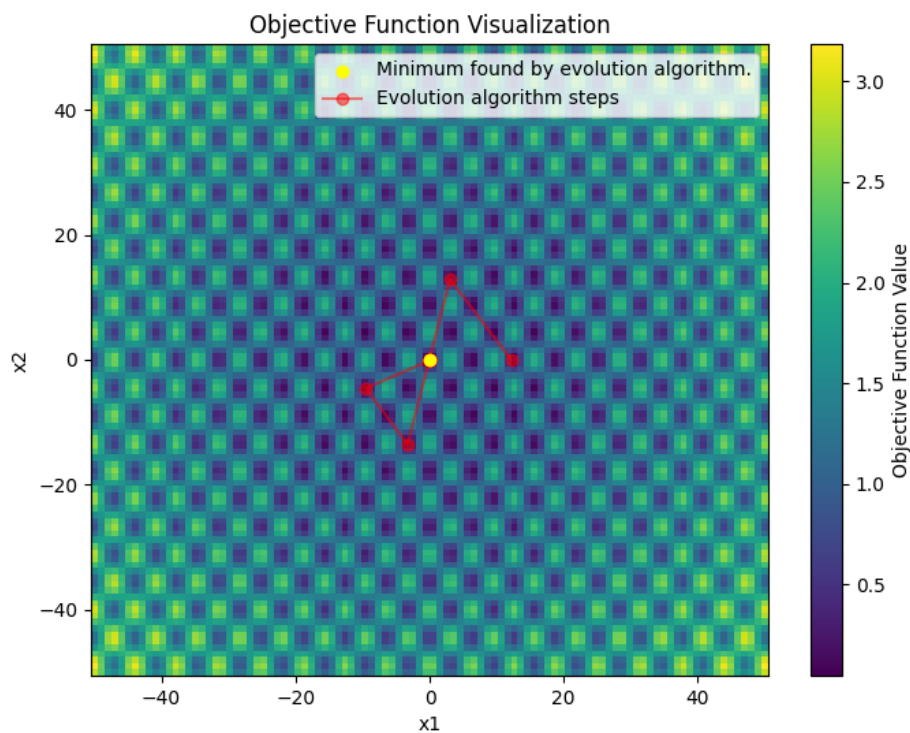
<i>Liczba osobników</i>	200
<i>p_{mutacji}</i>	0.2
<i>p_{krzyżowania}</i>	0.8
<i>σ</i>	0.2
<i>k</i>	30
<i>Liczba iteracji</i>	10000



Odnalezienie minimum: $x1 = -1.407977225952007e-09$, $x2 = 1.0489653238574209e-10$, $f(x1, x2) = 0.0$

○ Funkcja Griewanka

<i>Liczba osobników</i>	500
<i>p_{mutacji}</i>	0.5
<i>p_{krzyżowania}</i>	0.8
σ	0.8
k	10
<i>Liczba iteracji</i>	10000



Odnalezione minimum: $x1 = -2.066279716546198e-09$, $x2 = 6.5776610658269295e-09$, $f(x1, x2) = 0.0$

4.2. Badanie wpływu zmiany parametrów na odnalezione minima

Przeprowadzono badania odnalezionego rozwiązania w zależności od prawdopodobieństwa mutacji, prawdopodobieństwa krzyżowania, liczby iteracji oraz rozmiaru populacji. Parametry te były zmieniane w podanych zakresach:

iteracje = [10, 20, 50, 100, 200]

prawdopodobieństwa mutacji = [0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]

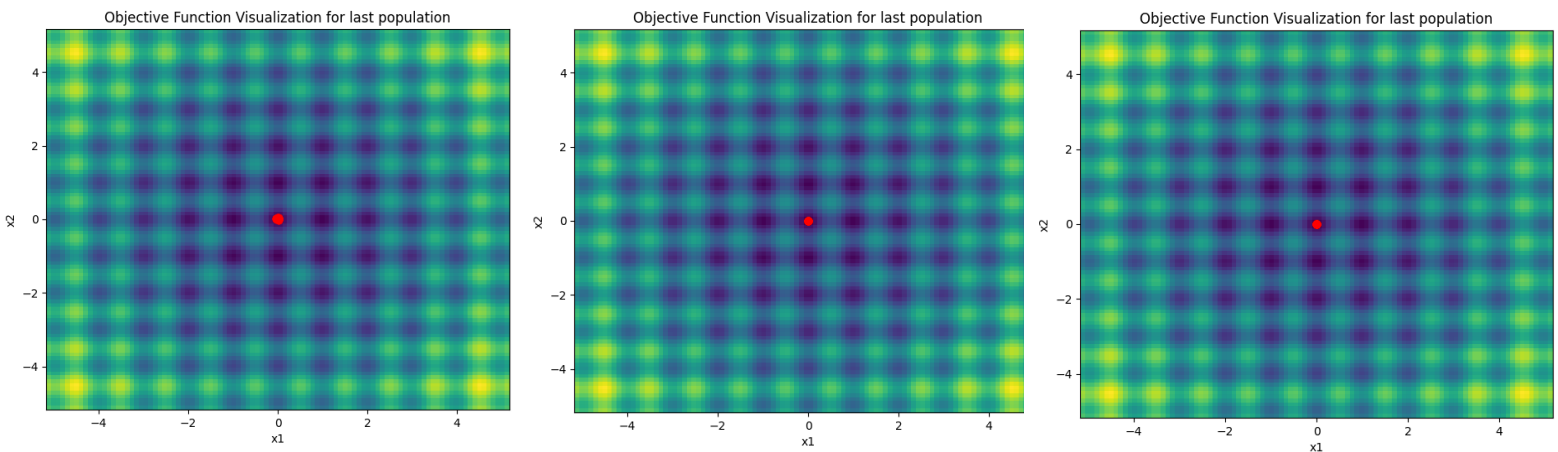
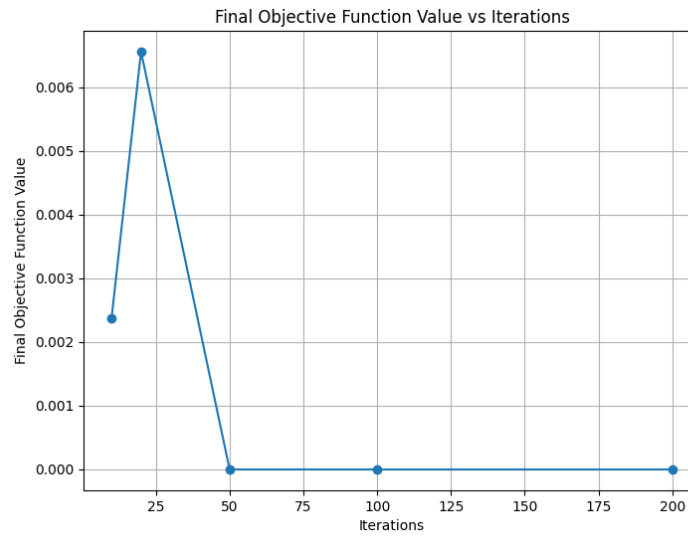
prawdopodobieństwa krzyżowania = [0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]

rozmiary populacji = [2, 10, 50, 100, 200, 400]

Wartości pozostałych parametrów były przyjmowane zgodnie z poprzednimi tabelkami, za wyjątkiem iteracji, które zostały ustawione na 100. Poniżej przedstawiono otrzymane wykresy oraz tabelki.

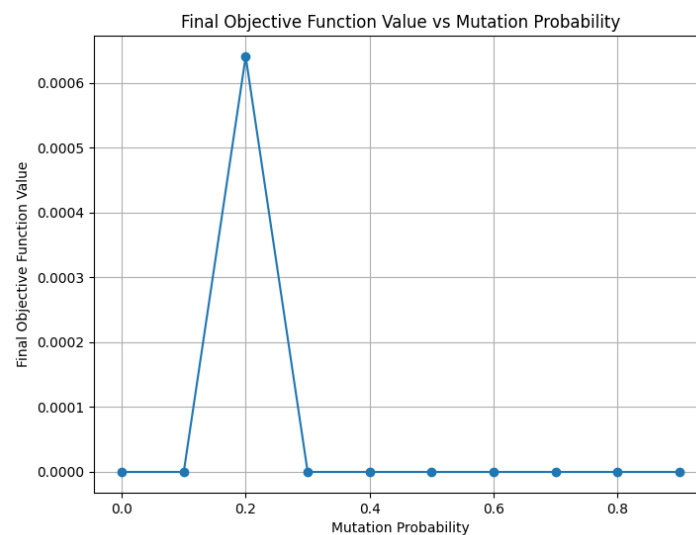
- Funkcja Rastrigina

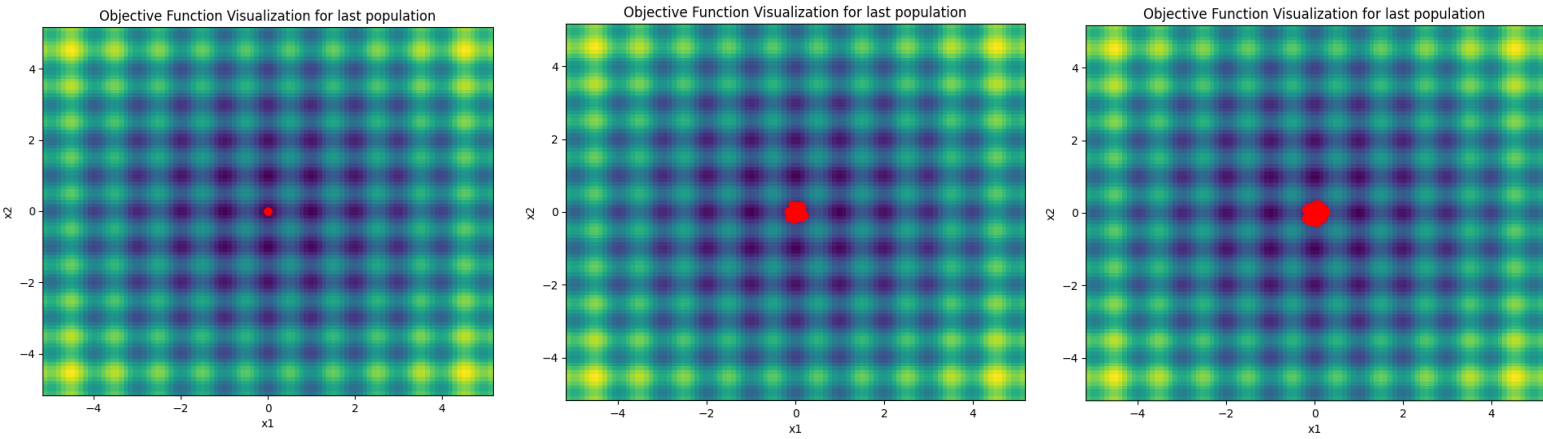
- Iteracje – wykres zależności i ostatnie populacje dla 10, 50 i 200



<i>ilość iteracji</i>	$x_1, x_2, f(x_1, x_2)$	Błąd średniokwadratowy
10	0.0032, 0.0012, 0.0024	0.00000576
20	-0.00029, -0.0057, 0.0066	0.00004356
50	-1.59e-09, -1.22e-09, 0.0	0.0
100	-1.37e-09, 2.26e-09, 0.0	0.0
200	1.49e-09, -2.58e-09, 0.0	0.0

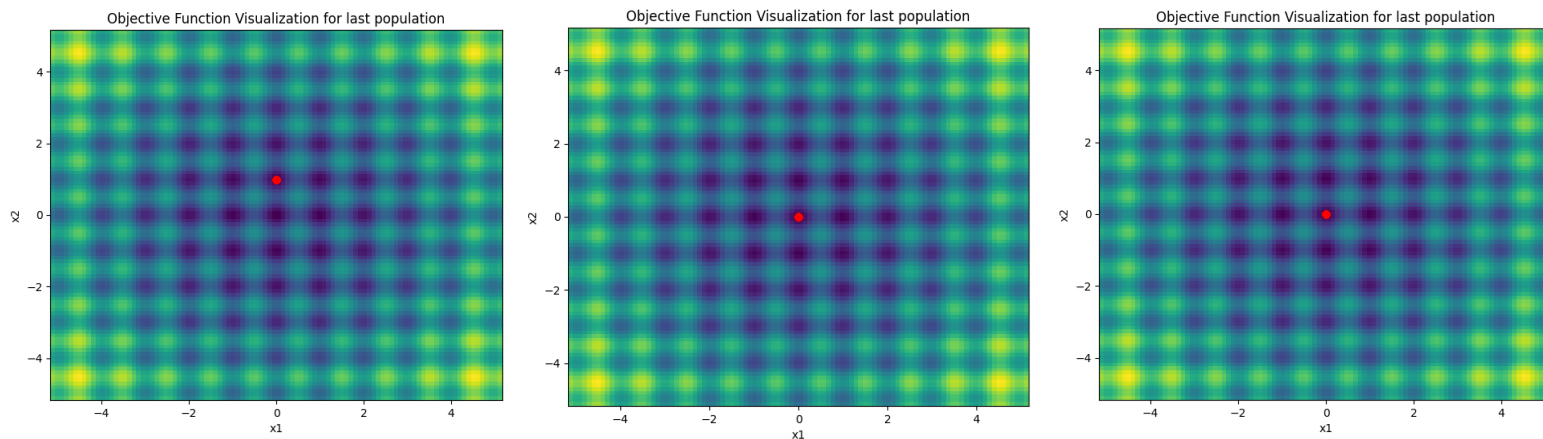
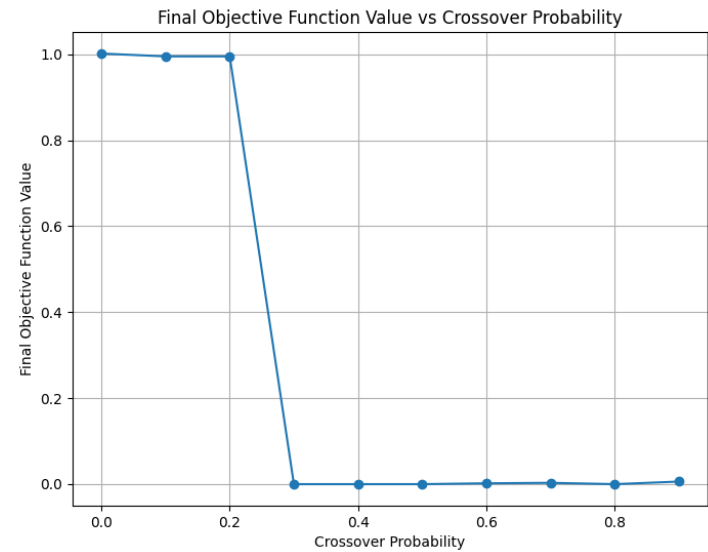
- Prawdopodobieństwo mutacji – wykres zależności i ostatnie populacje dla 0.1, 0.5, 0.9





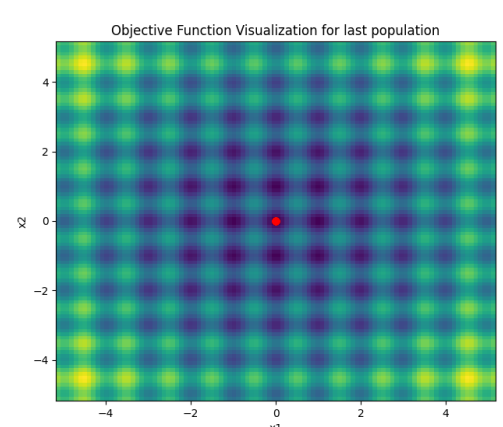
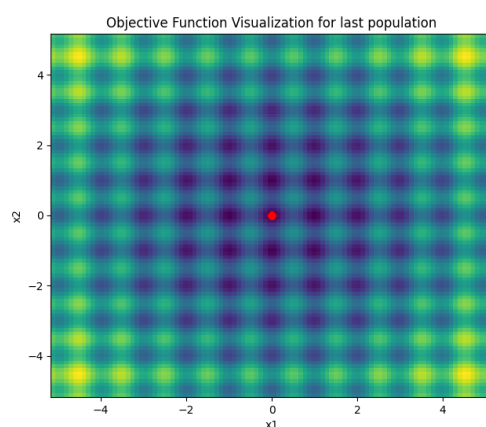
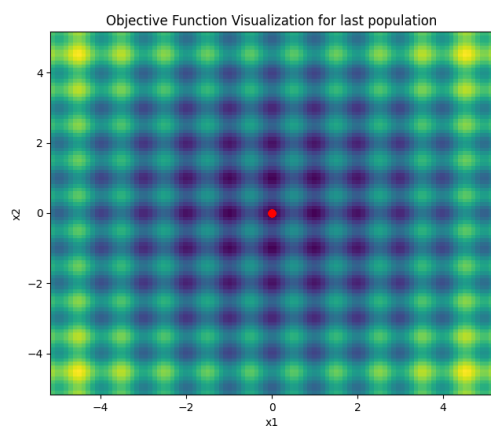
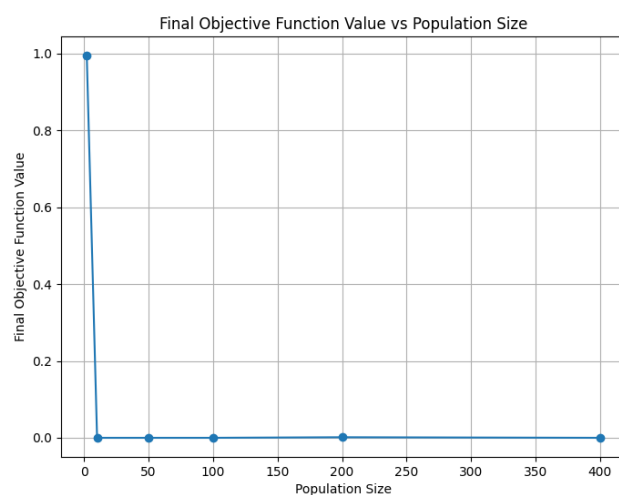
<i>prawdopodob.mutacji</i>	$x_1, x_2, f(x_1, x_2)$	błąd
0.0	1.23e-09, -8.24e-09, 1.42e-14,	2.016e-28
0.1	1.98e-09, 3.74e-10, 0.0	0.0
0.2	0.0018, -1.44e-06, 0.00064	4.09e-07
0.3	9.73e-08, 7.13e-07, 1.028e-10	1.057e-20
0.4	2.39e-09, -3.24e-10, 0.0	0.0
0.5	6.8e-10, -2.13e-09, 0.0	0.0
0.6	1.027e-09, 2.23e-09, 0.0	0.0
0.7	-1.13e-10, -3.42e-07, 2.32e-11	5.38e-22
0.8	-1.32e-10, 7.36e-10, 0.0	0.0
0.9	1.83e-08, -1.873e-09, 6.75e-14	4.56e-27

- Prawdopodobieństwo krzyżowania wykres zależności i ostatnie populacje dla 0.1, 0.5, 0.9



<i>prawdopodob. krzyżowania</i>	$x_1, x_2, f(x_1, x_2)$	błąd
0.0	-1, -0.0028, 1.0017	1.0034
0.1	-0.00035, 0.99, 0.99	0.9801
0.2	-3.089e-08, 0.99, 0.99	0.9801
0.3	-2.15e-09, -7.7e-10, 0.0	0.0
0.4	-9.68e-05, 1.97e-06, 1.86e-06	3.46e-12
0.5	0.00013, 0.001, 5.44e-06	2.96e-11
0.6	-2.08e-06, 0.0032, 0.002	4e-06
0.7	0.00022, -0.004, 0.003	9e-06
0.8	8.54e-10, 2.19e-09, 0.0	0.0
0.9	0.00016, 0.0055, 0.006	3.6e-05

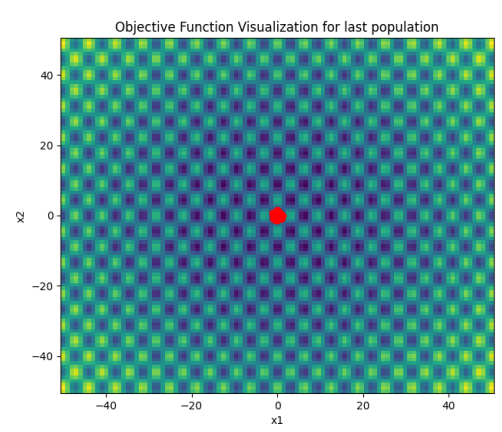
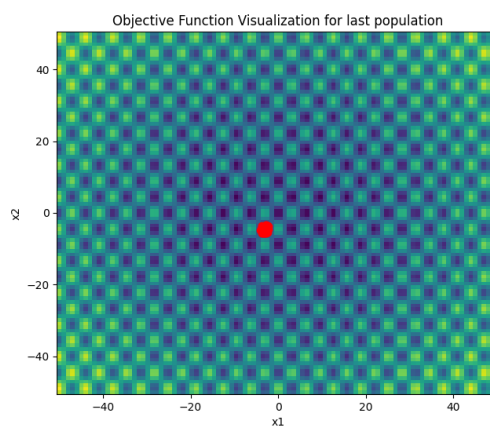
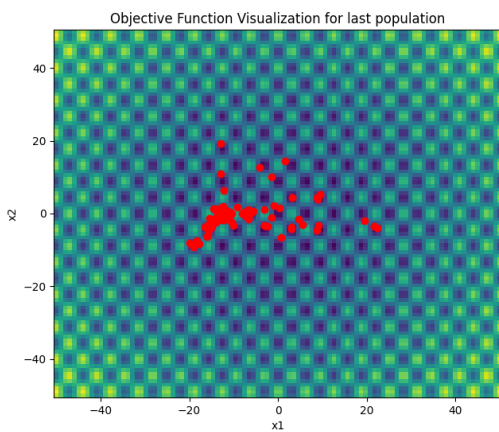
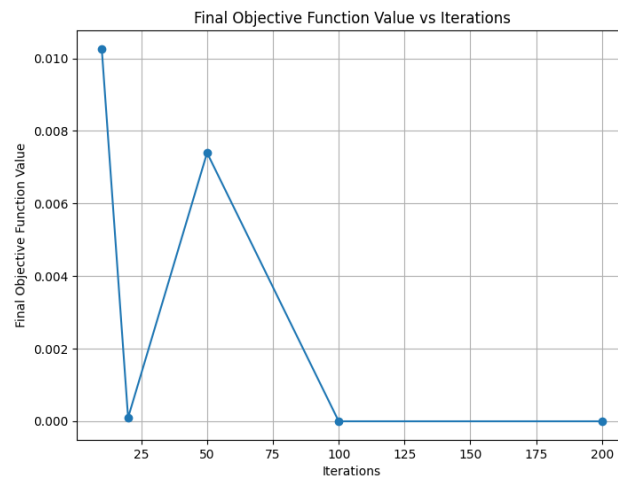
- Rozmiary populacji i ostatnie populacje dla 10, 50 i 200 osobników



<i>ilość osobników</i>	$x_1, x_2, f(x_1, x_2)$	błąd
2	2.16e-09, 0.99, 0.99	0.98
10	-2.88e-10, -6.51e-10, 0.0	0.0
50	1.61e-09, -7.8e-10, 0.0	0.0
100	9.43e-11, 1.88e-10, 0.0	0.0
200	-2.48e-05, 0.0027, 0.0015	2.25e-06
400	1.68e-10, -7.65e-09, 0.0	0.0

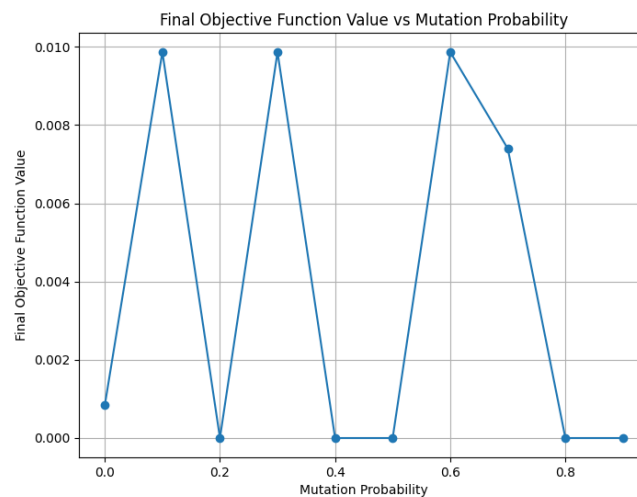
- Funkcja Griewanka

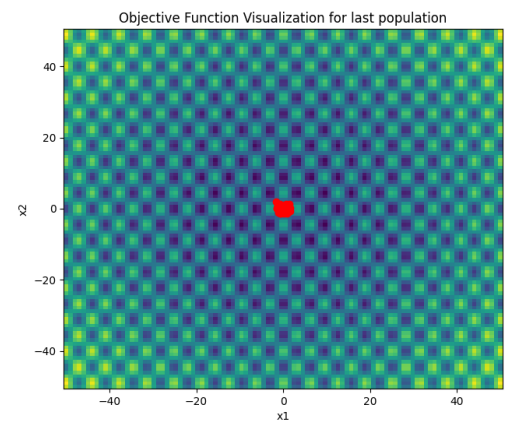
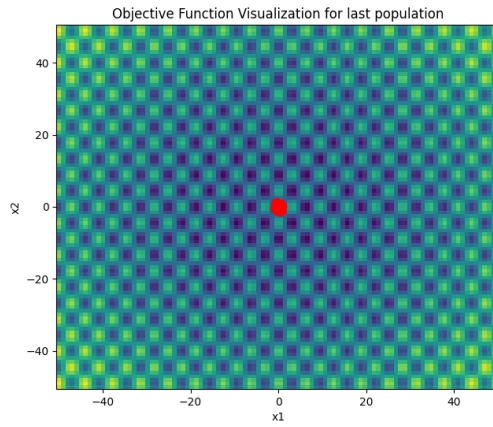
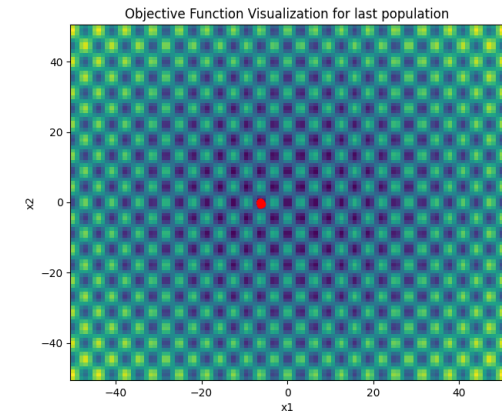
- Iteracje – wykres zależności i ostatnie populacje dla 10, 50 i 200



<i>ilość iteracji</i>	$x_1, x_2, f(x_1, x_2)$	błąd
10	-6.3, 0.03, 0.01	1e-04
20	-0.014, 0.029, 9.96e-05	9.9e-09
50	-3.14, -4.44, 0.0074	5.5e-05
100	7.44e-05, 1.66e-05, 2.84e-09	8.0656e-18
200	-5.5e-09, -9.97e-09, 0.0	0.0

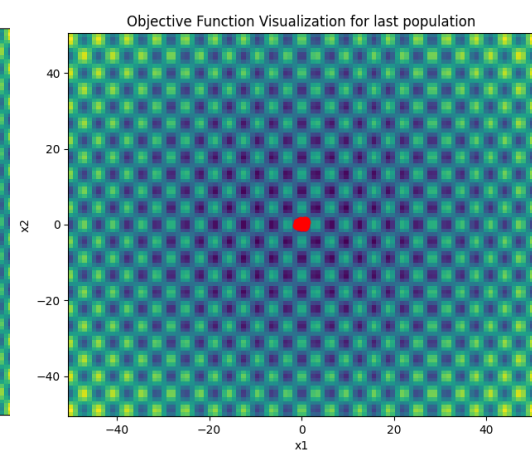
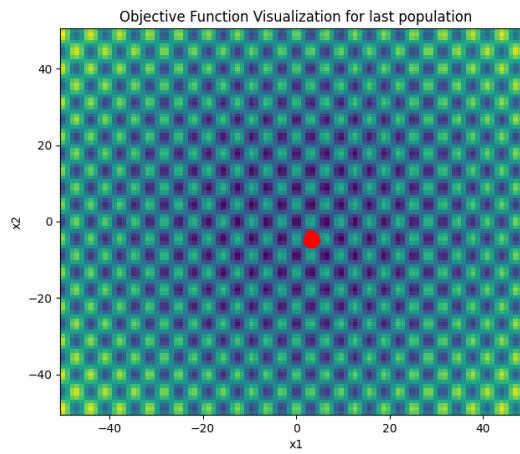
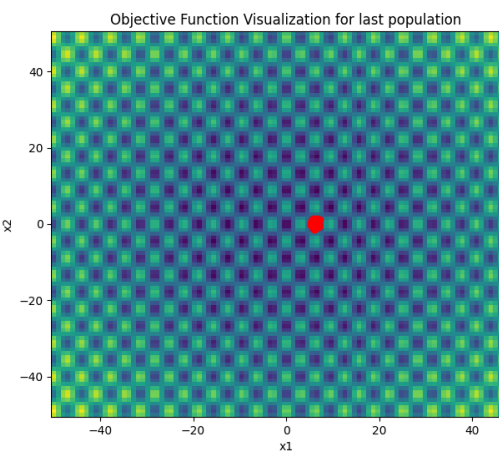
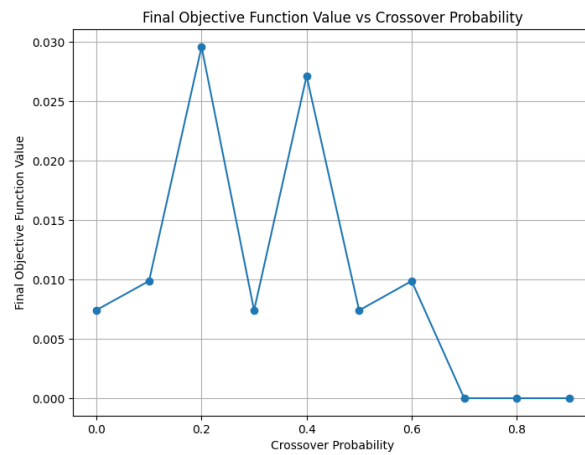
- Prawdopodobieństwo mutacji – wykres zależności i ostatnie populacje dla 0.1, 0.5, 0.9





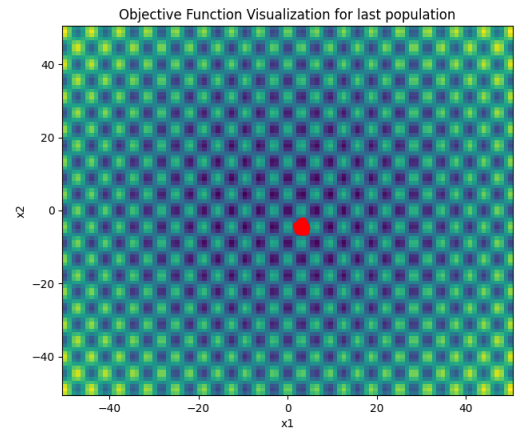
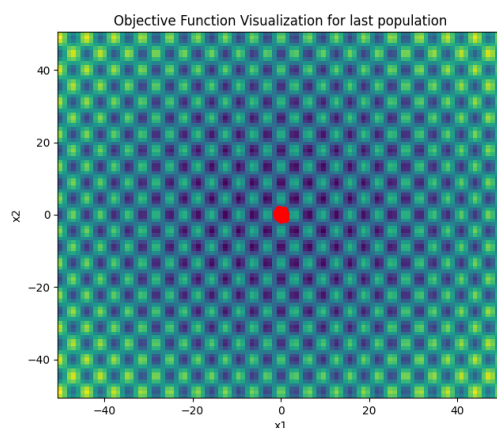
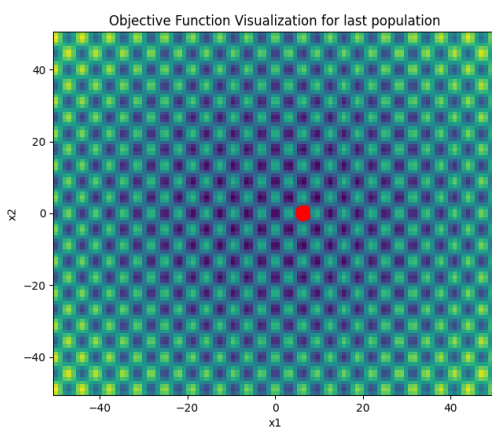
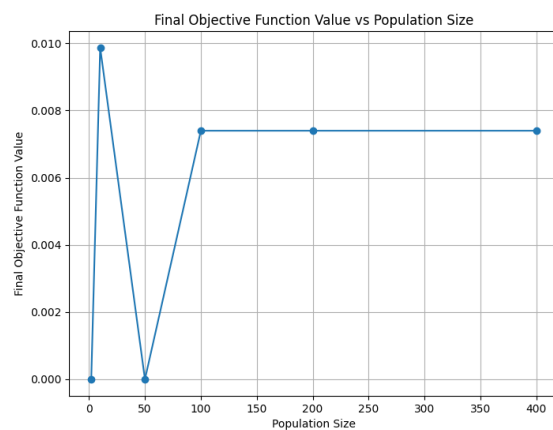
<i>prawdopodob.mutacji</i>	$x_1, x_2, f(x_1, x_2)$	błąd
0.0	-0.24, 0.047, 0.00084	7.056e-07
0.1	-6.28, 3.22e-09, 0.0098	9.6e-05
0.2	-0.00044, -5.89e-05, 9.84e-08	9.68e-15
0.3	6.28, -1.31e-08, 0.0099	9.8e-5
0.4	1.016e-08, 1.086e-08, 0.0	0.0
0.5	-5.022e-09, -3.16e-09, 0.0	0.0
0.6	-6.28, -2.48e-05, 0.001	1e-03
0.7	3.14, -4.44, 0.0074	5.48e-05
0.8	-2.53e-09, 1.26e-08, 0.0	0.0
0.9	5.68e-09, 8.24e-07, 1.7e-13	2.89e-26

- Prawdopodobieństwo krzyżowania wykres zależności i ostatnie populacje dla 0.1, 0.5, 0.9



<i>prawdopodob.krzyżowania</i>	$x_1, x_2, f(x_1, x_2)$	błąd
0.0	3.15, -4.44, 0.0074	5.48e-05
0.1	6.28, 0.00073, 0.0099	9.8e-05
0.2	-6.28, -8.88, 0.03	9e-04
0.3	3.14, -4.44, 0.0074	5.48e-05
0.4	9.42, -4.44, 0.027	7.29e-04
0.5	3.14, -4.44, 0.0074	5.48e-05
0.6	6.28, 8.24e-09, 0.0099	9.8e-05
0.7	4.83e-08, -3.32e-07, 4.83e-08	2.33e-15
0.8	-8.88e-05, -7.19e-06, 3.96e-09	1.57e-17
0.9	2.92e-05, -9.26e-09, 4.28e-10	1.83e-19

- Rozmiary populacji i ostatnie populacje dla 10, 50 i 200 osobników



<i>ilość osobników</i>	$x_1, x_2, f(x_1, x_2)$	błąd
2	7.91e-09, -4.32e-09, 0.0	0.0
10	6.28, -1.14e-08, 0.0099	9.8e-05
50	-1.0028e-08, -7.35e-09, 0.0	0.0
100	-3.14, -4.44, 0.0074	5.48e-05
200	3.14, -4.44, 0.0074	5.48e-05
400	-3.14, -4.44, 0.0074	5.48e-05

5. Wnioski

Dobór parametrów algorytmu ewolucyjnego bardzo zależy od badanej funkcji. Najważniejsze jest, aby algorytm miał odpowiednio wysokie prawdopodobieństwo krzyżowania – bez mutacji algorytm może funkcjonować, bez krzyżowania już nie można go nazwać algorytmem ewolucyjnym. Wyższe prawdopodobieństwo mutacji daje większe rozprzestrzenienie osobników w przeszukiwanej przestrzeni. Pozwala to wyjść z minimum lokalnego. Dla funkcji Griewanka ustawienie wysokiego prawdopodobieństwa mutacji było kluczowe do odnalezienia globalnego minimum – algorytm często utykał w lokalnych minimach, których dla tej funkcji na danym przedziale było bardzo wiele i ich wartość była bliska zeru. Większa liczba populacji daje zazwyczaj lepsze wyniki, jednakże jest kosztowniejsza obliczeniowo. Kluczowa jest duża ilość iteracji, w przeciwnym razie algorytm często utyka w minimach lokalnych. Przedstawione wykresy, ze względu na prawdopodobieństwa, za każdym razem wychodziły różne, jednakże przy wielokrotnym uruchomieniu można zaobserwować opisane prawidłowości.

Dla funkcji Rastrigina wystarczyło 50 iteracji do osiągnięcia minimum globalnego.

Prawdopodobieństwo mutacji w przypadku tej funkcji nie było aż tak istotne, już przy bardzo niewielkich wartościach osiągnięto cel. Prawdopodobieństwo krzyżowania wynoszące 0.3 pozwoliło osiągnąć cel, algorytm bardzo dobrze radził sobie z tą funkcją. Dla prawdopodobieństwa krzyżowania wynoszącego 0.9 wyniki pogorszyły się. Może to wynikać z faktu, że z dużym prawdopodobieństwem krzyżowane będzie wiele osobników, co też może nie być korzystne, gdyż wprowadza zbyt dużo losowości. Już 10 osobników było w stanie osiągnąć minimum globalne.

Funkcja Griewanka była dużo bardziej wymagająca i zadawalające wyniki otrzymano dla 100 iteracji. Prawdopodobieństwo mutacji było istotne i minimum globalne otrzymano dopiero dla wartości 0.4. Najlepsze wyniki otrzymano dla wysokich wartości prawdopodobieństwa krzyżowania. Dla badań populacji problematyczne jest wyciągnięcie wniosków, układ wyników był niefortunny, z jednej strony dla 2 osobników osiągnięto minimum globalne a dla 400 już nie.