

[PSZT-P] Kiedy się poddać?

Piotr Frątczak
(300207)

Bartosz Świtalski
(300279)

29 listopada 2020

1 Opis problemu

Algorytmy ewolucyjne w klasycznym wydaniu nie mogą same zdecydować, kiedy zakończyć swoje działanie. W związku z tym należy rozważyć implementację określonego z góry kryterium. Nie istnieje uniwersalne kryterium wykrywania bezcelowości dalszej pracy algorytmu ewolucyjnego. W naszej pracy proponujemy przykładowe rozwiązania, zaimplementujemy je oraz dokonamy analizy ich działania.

2 Decyzje projektowe

- Optymalizacja zostanie przeprowadzona na funkcjach z [cec2005](#).
- Zaimplementowana została strategia ewolucyjna $(\mu + \lambda)$.
- Przyjęty budżet możliwych ewaluacji funkcji celu dla pojedynczej próby optymalizacji wynosi $10000 \cdot \text{wymiarowość zadania}$.
- Skupiamy się na wymiarowości $D = 10$.
- Podczas jednego uruchomienia programu dokonujemy uśrednienia wyników z 51 wywołań algorytmu.
- W przypadku zajścia ustalonego kryterium stopu obliczenia w danym uruchomieniu są przerywane.

3 Cele eksperymentu

Implementacja kryteriów przerwań optymalizacji. Dobranie przykładowych parametrów dla odpowiednich kryteriów. Zbadanie wpływu tychże kryteriów na ogólny czas optymalizacji oraz dokładność uzyskanego wyniku (optimum).

4 Użycie

```
/when-to-surrender$  
  
./xsetup.sh  
python3 when-to-surrender/main.py  
<funkcja> <kryterium> <p1> <p2> <p3> <p4> <plik_wykres>  
# po wykonaniu  
./xclean.sh
```

Oznaczenia argumentów

<funkcja> - optymalizowana funkcja (dozwolone wartości: F_4 ¹, F_5 ², F_6 ³).
<kryterium> - kryterium przerywania (dozwolone wartości: *k-iter*, *sd*, *best-worst*, *variance*).
<p_i> - kolejne wartości parametru do wcześniej sprecyzowanego kryterium.
<plik_wykres> - nazwa pliku do którego zostanie zapisany wykres wynikowy.

Użycie skryptu

```
/when-to-surrender$  
  
./xsetup.sh  
./xscript.sh  
# po wykonaniu  
./xclean.sh
```

Komentarz do użycia

Skrypt umożliwia uruchomienie optymalizacji wszystkich dostępnych funkcji (3) według wszystkich zaimplementowanych kryteriów (4) z wcześniej określonymi parametrami. Pojedyncze uruchomienie skryptu jest dość kosztowne czasowo, więc w celu skrócenia czasu wykonania zalecany jest np. przydział zadań do różnych rdzeni (komenda `taskset`), ale optymalizacja czasu wykonania wielu uruchomień na raz nie jest tematem projektu.

Wygenerowane dane zostaną zapisane do plików w folderze `output`. Wygenerowane wykresy zostaną zapisane do plików w folderze `graphs`. Do czyszczenia służy skrypt `xclean.sh`.

¹Shifted Schwefel's Problem 1.2 with Noise in Fitness

²Schwefel's Problem 2.6 with Global Optimum on Bounds

³Shifted Rosenbrock's Function

5 Kryteria przerwań

- ***K-iterations* (*K-iteracji*).** Wykorzystane jako kryterium bazowe. Kryterium jest spełnione, jeśli nie ma poprawy wartości funkcji celu przez K kolejnych iteracji. Należy wybrać odpowiednią wartość K przy założeniu, że niemożliwe jest uzyskanie lepszego wyniku po K kolejnych iteracjach.
- **Standard Deviation (Odchylenie Standardowe).** Kryterium jest spełnione, jeżeli odchylenie standardowe wartości każdej z cech osobników obecnej generacji jest mniejsze lub równe niż dane próg $\epsilon \geq 0$.
- **Best-worst (Najlepszy-Najgorszy).** Kryterium jest spełnione, gdy różnica funkcji celu między najlepszym i najgorszym osobnikiem jest mniejsza lub równa niż dany próg $\epsilon \geq 0$.
- **Fit-Variance (Wariancja Dopasowania).** Kryterium uwzględniające zróżnicowanie funkcji celu wszystkich osobników populacji. Jest spełnione, jeśli wariancja funkcji celu dla wszystkich osobników w pokoleniu jest mniejsza lub równa niż dany próg ϵ przy czym $1 \gg \epsilon \geq 0$.

6 Testowanie

Testowaliśmy na wybranych ustawieniach początkowych:

- $\mu = 20$
- $\lambda = 7 * \mu$
- $\sigma_0 \in (0.9; 1.1)$

Zbadaliśmy różne parametry dla każdej funkcji wg każdego kryterium. Następnie na podstawie wyników wybraliśmy cztery odpowiednio dobrane parametry dla każdej pary (funkcja, kryterium).

Uzyskane wyniki zostały omówione w dalszej części niniejszego sprawozdania.

7 Wyniki

Objaśnienia

`best fit` - najlepsze znalezione optimum

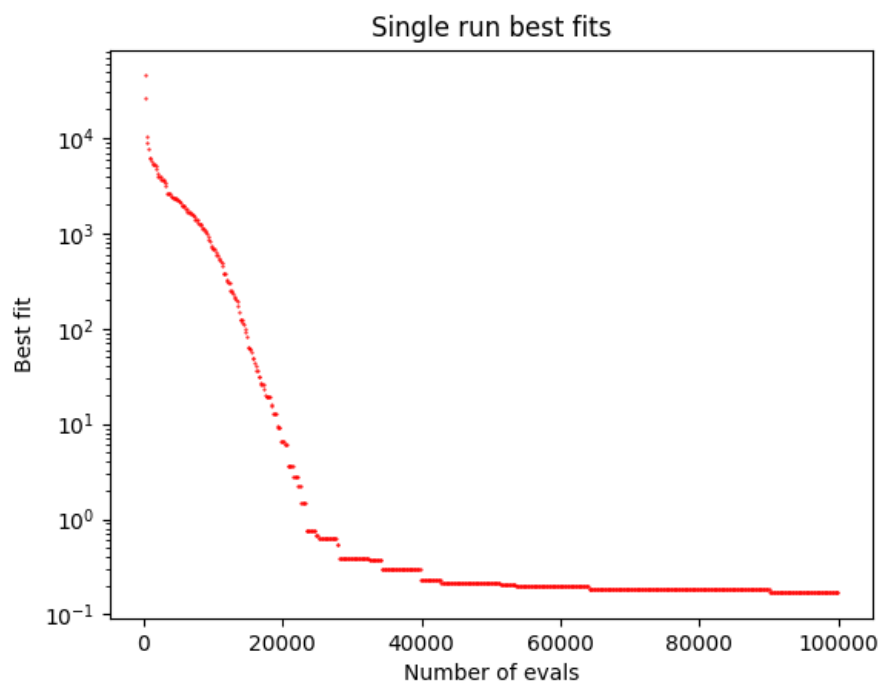
`best fit mean` - średnia najlepszych znalezionych optimów

`best fit std. deviation` - odchylenie standardowe najlepszych znalezionych optimów

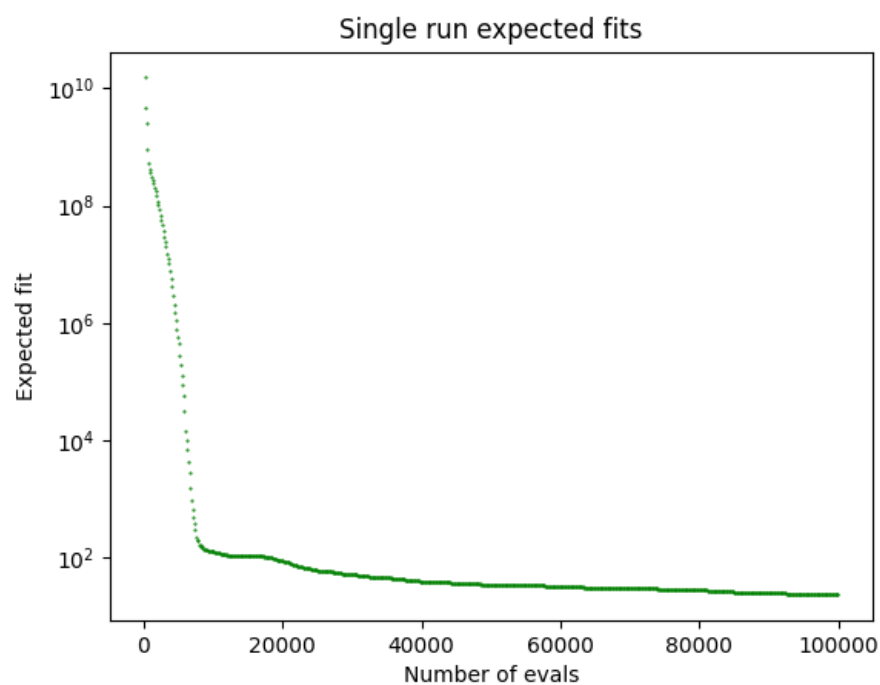
`budget use` - średnie wykorzystanie budżetu

`number of evals` - średnia liczba ewaluacji funkcji celu

Przykłady działania strategii ewolucyjnej ($\mu + \lambda$)



Rysunek 1: Dopasowanie najlepszego osobnika populacji w zależności od liczby ewaluacji funkcji celu dla F_4 .

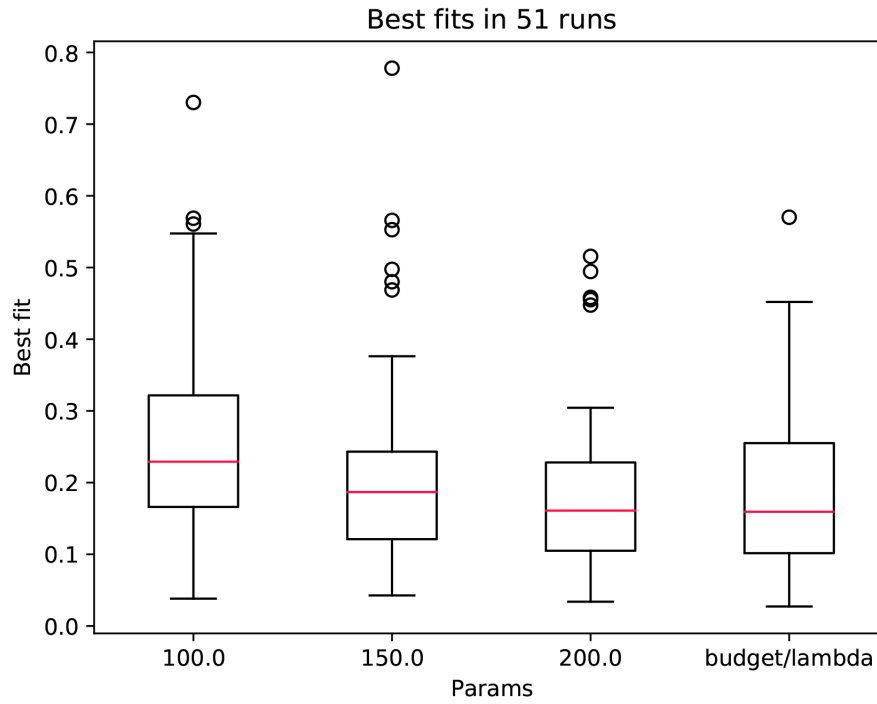


Rysunek 2: Dopasowanie średniego osobnika populacji w zależności od liczby ewaluacji funkcji celu dla F_6 .

Kryterium K -iteracji

k value	best fit	best fit mean	best fit std. deviation	budget use mean
100	0.04	0.26	0.14	84.74%
150	0.04	0.22	0.15	95.41%
200	0.03	0.19	0.11	96.63%
budget/lambda	0.03	0.19	0.12	99.98%

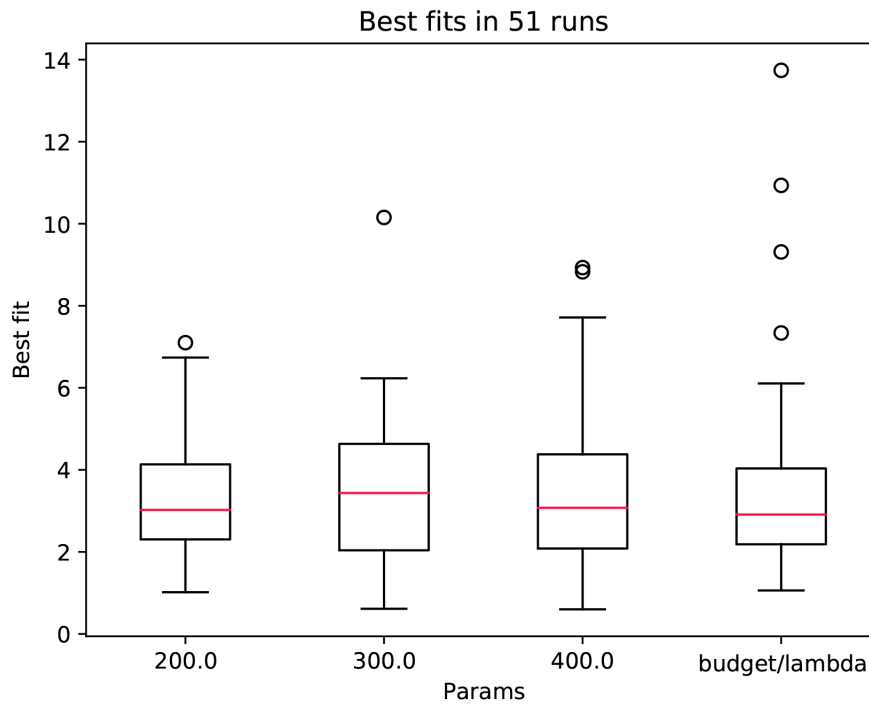
Tablica 1: Porównanie najlepszych wartości parametru k dla F_4



Rysunek 3: Wartości najlepszego dopasowania funkcji F_4 wg kryterium K -iteracji. Najlepszy parametr $k \approx 200$, wykorzystanie $\approx 96.63\%$ budżetu.

k value	best fit	best fit mean	best fit std. deviation	budget use mean
200	1.02	3.32	1.33	94.65%
300	0.61	3.5	1.74	99.52%
400	0.6	3.49	1.87	99.87%
budget/lambda	1.06	3.56	2.38	99.98%

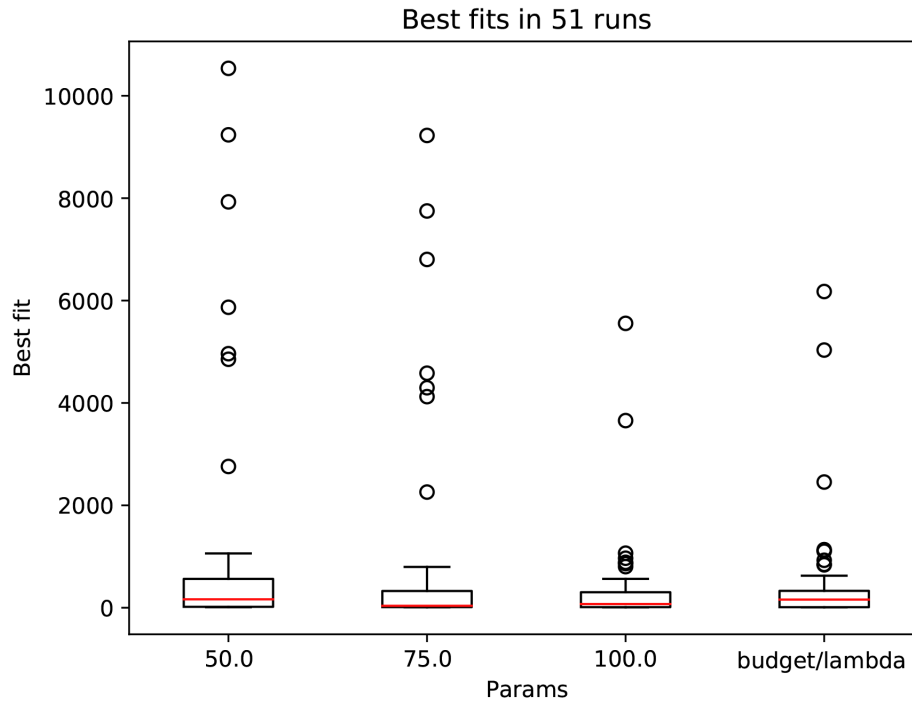
Tablica 2: Porównanie najlepszych wartości parametru k dla F_5



Rysunek 4: Wartości najlepszego dopasowania funkcji F_5 wg kryterium K -iteracji. Najlepszy parametr $k \approx 200$, wykorzystanie $\approx 94.65\%$ budżetu.

k value	best fit	best fit mean	best fit std. deviation	budget use mean
50	9.88	1090.58	2400.09	42.34%
75	8.06	886.17	2057.39	56.76%
100	7.69	364.44	918.76	65.22%
budget/lambda	7.37	493.17	1126.43	99.98%

Tablica 3: Porównanie najlepszych wartości parametru k dla F_6

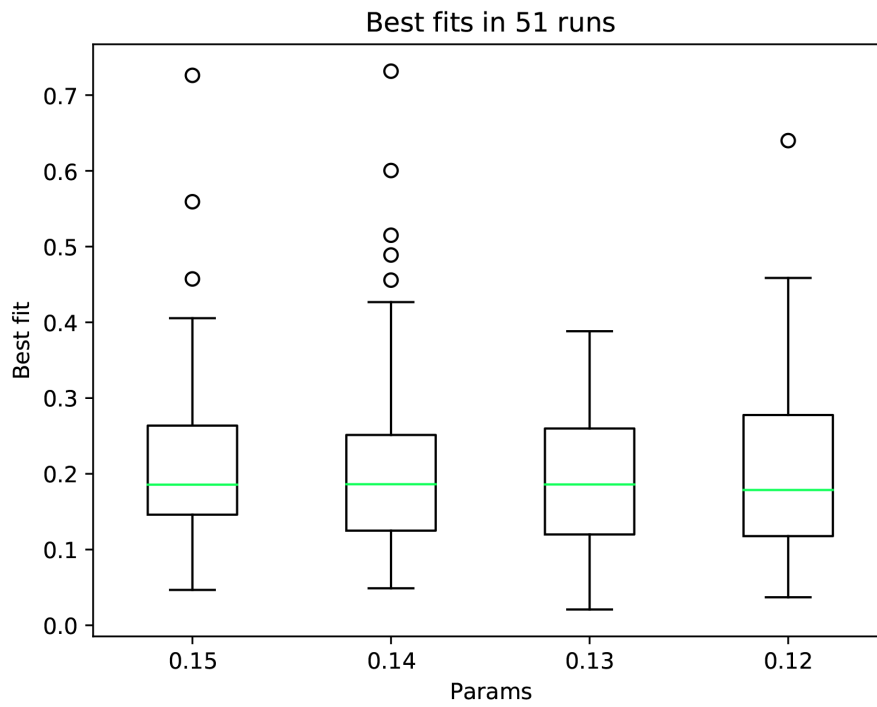


Rysunek 5: Wartości najlepszego dopasowania funkcji F_6 wg kryterium K -iteracji. Najlepszy parametr $k \approx 100$, wykorzystanie $\approx 65.22\%$ budżetu.

Kryterium Odchylenia Standardowego

ϵ	best fit	best fit mean	best fit std. deviation	budget use mean
0.15	0.05	0.22	0.12	89.3%
0.14	0.05	0.22	0.14	95.3%
0.13	0.02	0.2	0.09	95.27%
0.12	0.04	0.21	0.12	97.63%

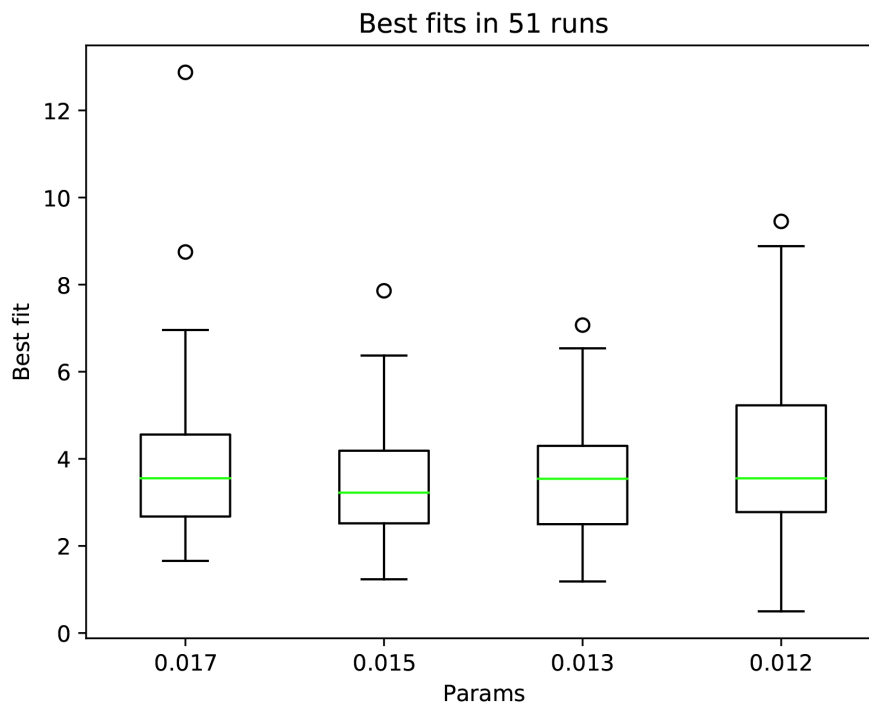
Tablica 4: Porównanie najlepszych wartości parametru ϵ dla F_4



Rysunek 6: Wartości najlepszego dopasowania funkcji F_4 wg kryterium Odchylenia Standardowego. Najlepszy parametr $\epsilon \approx 0.13$, wykorzystanie $\approx 95.27\%$ budżetu.

ϵ	best fit	best fit mean	best fit std. deviation	budget use mean
0.017	1.66	3.89	1.94	83.7%
0.015	1.23	3.48	1.38	93.03%
0.013	1.18	3.5	1.34	94.6%
0.012	0.5	4.02	1.87	97.92%

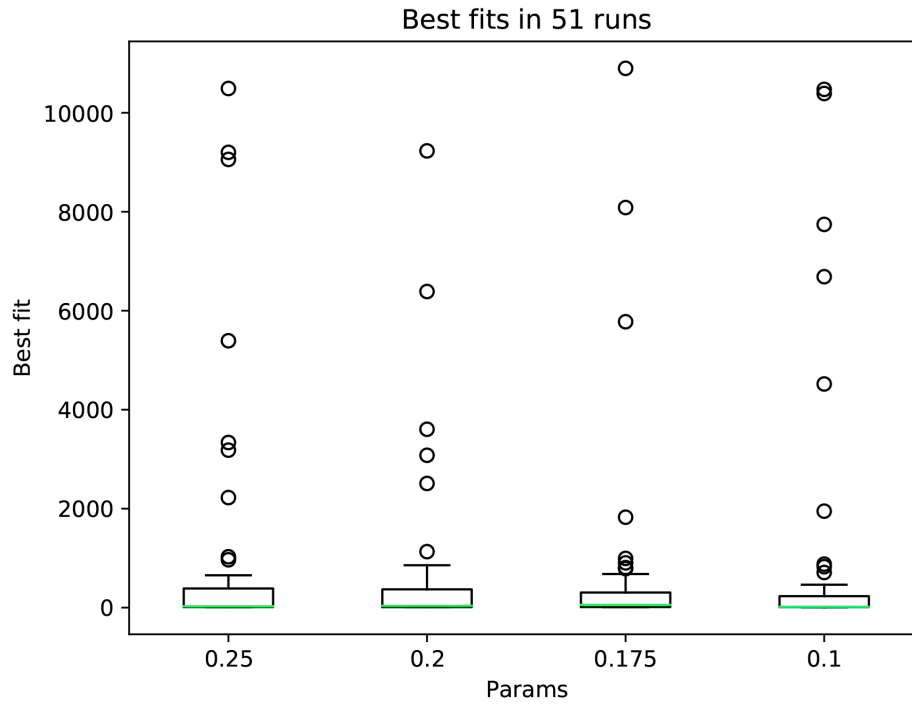
Tablica 5: Porównanie najlepszych wartości parametru ϵ dla F_5



Rysunek 7: Wartości najlepszego dopasowania funkcji F_5 wg kryterium Odchylenia Standardowego. Najlepszy parametr $\epsilon \approx 0.015$, wykorzystanie $\approx 93.03\%$ budżetu.

ϵ	best fit	best fit mean	best fit std. deviation	budget use mean
0.25	7.6	972.81	2374.26	56.62%
0.2	7.7	631.97	1646.65	76.59%
0.175	7.9	686.5	1987.23	81.41%
0.1	5.14	929.13	2446.75	99.3%

Tablica 6: Porównanie najlepszych wartości parametru ϵ dla F_6

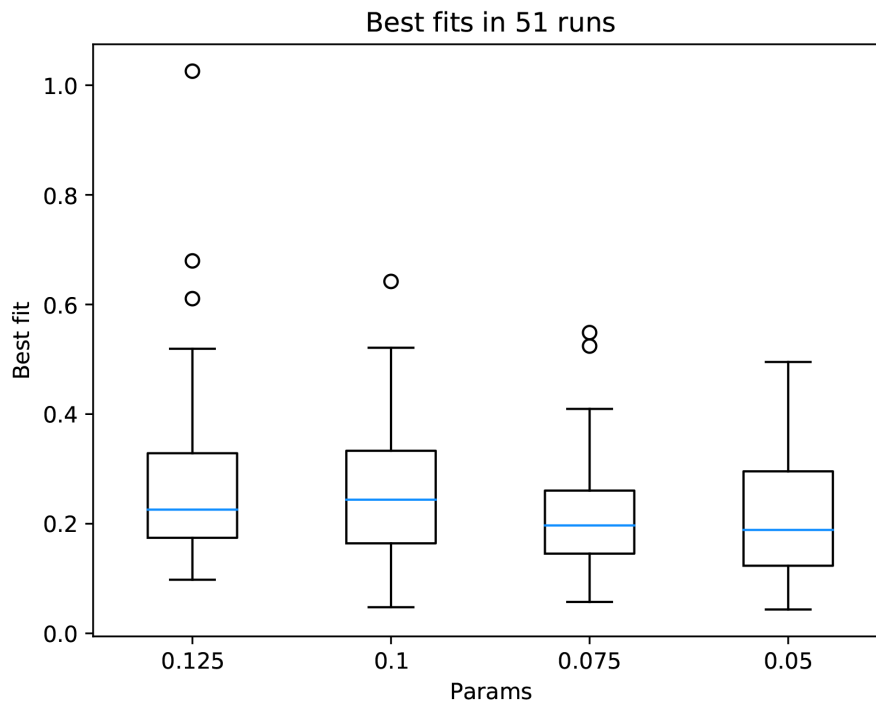


Rysunek 8: Wartości najlepszego dopasowania funkcji F_6 wg kryterium Odchylenia Standardowego. Najlepszy parametr $\epsilon \approx 0.2$, wykorzystanie $\approx 76.59\%$ budżetu.

Kryterium Najlepszy-Najgorszy

ϵ	best fit	best fit mean	best fit std. deviation	budget use mean
0.125	0.1	0.28	0.17	73.41%
0.1	0.05	0.26	0.12	85.76%
0.075	0.06	0.21	0.1	93.59%
0.05	0.04	0.21	0.12	99.3%

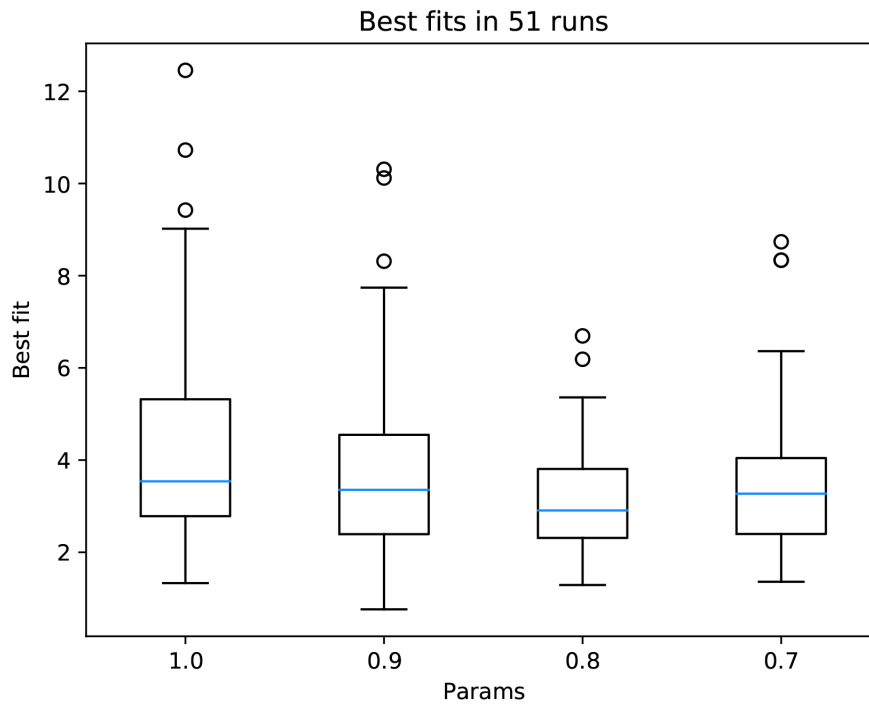
Tablica 7: Porównanie najlepszych wartości parametru ϵ dla F_4



Rysunek 9: Wartości najlepszego dopasowania funkcji F_4 wg kryterium Najlepszy-Najgorszy. Najlepszy parametr $\epsilon \approx 0.075$, wykorzystanie $\approx 93.59\%$ budżetu.

ϵ	best fit	best fit mean	best fit std. deviation	budget use mean
1.0	1.33	4.37	2.48	88.9%
0.9	0.76	3.78	2.1	89.1%
0.8	1.29	3.19	1.22	98.47%
0.7	1.36	3.61	1.71	98.25%

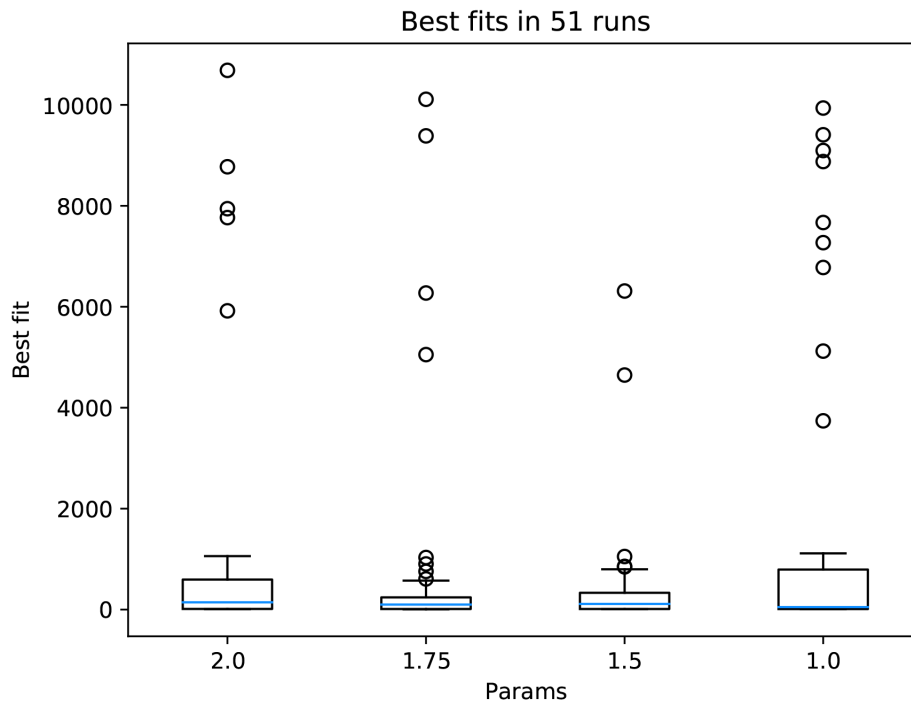
Tablica 8: Porównanie najlepszych wartości parametru ϵ dla F_5



Rysunek 10: Wartości najlepszego dopasowania funkcji F_5 wg kryterium Najlepsy-Najgroszy. Najlepszy parametr $\epsilon \approx 0.8$, wykorzystanie $\approx 98.47\%$ budżetu.

ϵ	best fit	best fit mean	best fit std. deviation	budget use mean
2.0	8.07	1030.86	2437.83	65.55%
1.75	4.67	759.88	2122.63	81%
1.5	8.33	420.16	1069.83	86.58%
1.0	7.68	1489.94	2933.88	99.69%

Tablica 9: Porównanie najlepszych wartości parametru ϵ dla F_6

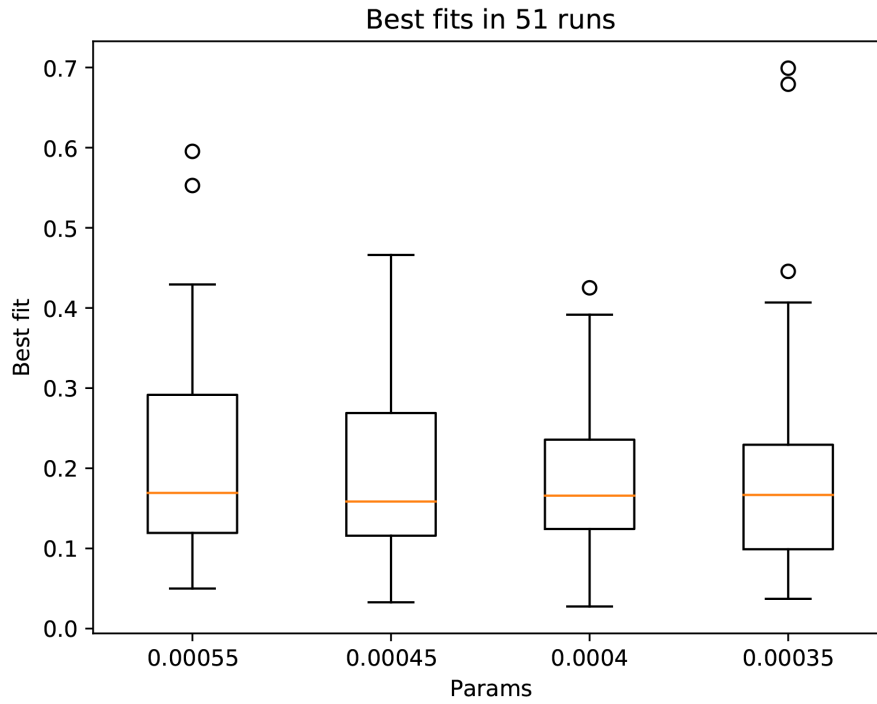


Rysunek 11: Wartości najlepszego dopasowania funkcji F_6 wg kryterium Najlepszy-Najgroszy. Najlepszy parametr $\epsilon \approx 1.5$, wykorzystanie $\approx 86.58\%$ budżetu.

Kryterium Wariancji Dopasowania

ϵ	best fit	best fit mean	best fit std. deviation	budget use mean
0.00055	0.05	0.21	0.12	90.22%
0.00045	0.03	0.2	0.11	91.68%
0.0004	0.03	0.18	0.1	96.89%
0.00035	0.04	0.19	0.14	97.67%

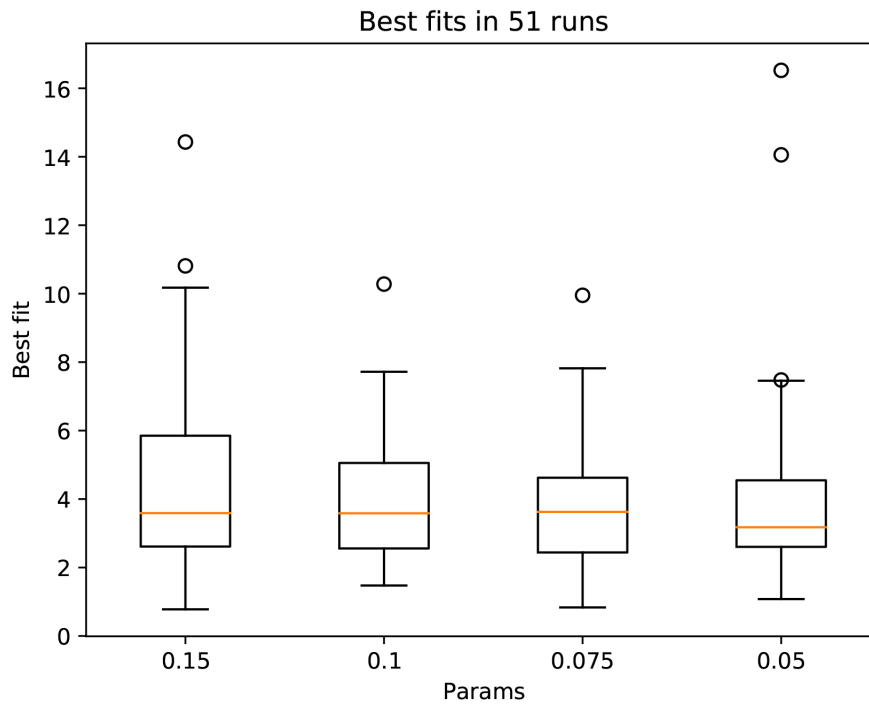
Tablica 10: Porównanie najlepszych wartości parametru ϵ dla F_4



Rysunek 12: Wartości najlepszego dopasowania funkcji F_4 wg kryterium Wariancji Dopasowania. Najlepszy parametr $\epsilon \approx 0.0004$, wykorzystanie $\approx 96.89\%$ budżetu.

ϵ	best fit	best fit mean	best fit std. deviation	budget use mean
0.15	0.78	4.47	2.78	72.52%
0.1	1.47	4.0	1.84	85.24%
0.075	0.83	3.84	1.82	92.95%
0.05	1.08	4.07	2.79	96.18%

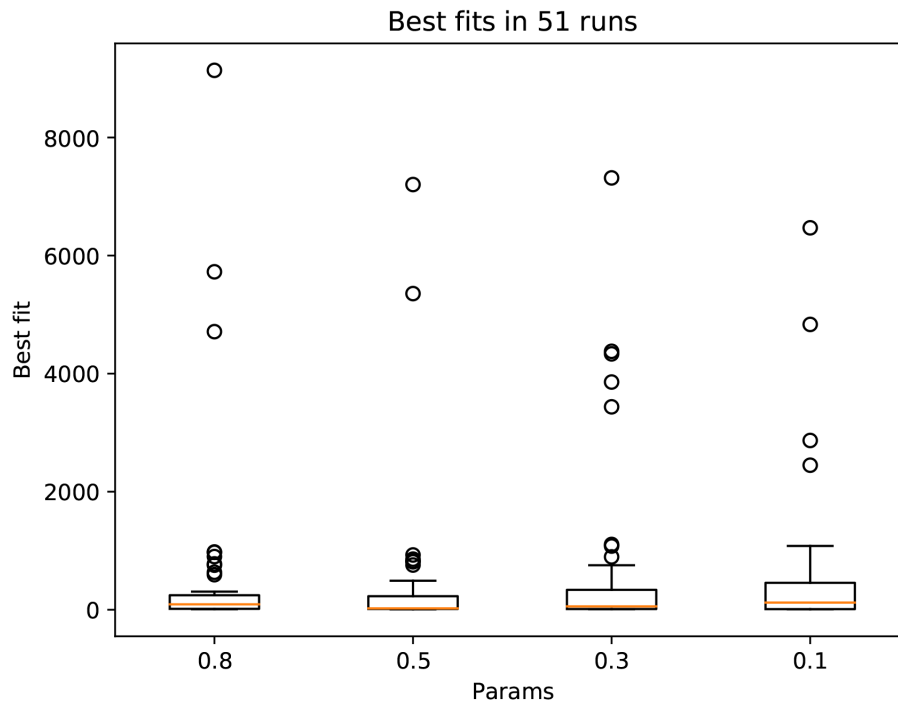
Tablica 11: Porównanie najlepszych wartości parametru ϵ dla F_5



Rysunek 13: Wartości najlepszego dopasowania funkcji F_5 wg kryterium Wariacji Dopasowania. Najlepszy parametr $\epsilon \approx 0.075$, wykorzystanie $\approx 92.95\%$ budżetu.

ϵ	best fit	best fit mean	best fit std. deviation	budget use mean
0.8	8.3	560.79	1582.09	46.6%
0.5	6.82	409.4	1225.4	50.01%
0.3	8.43	615.06	1427.96	75.1%
0.1	7.46	524.33	1189.06	98.09%

Tablica 12: Porównanie najlepszych wartości parametru ϵ dla F_6



Rysunek 14: Wartości najlepszego dopasowania funkcji F_6 wg kryterium Wariacji Dopasowania. Najlepszy parametr $\epsilon \approx 0.5$, wykorzystanie $\approx 50.01\%$ budżetu.

Porównanie najlepszych wyników

kryterium	średnie dopasowanie	odchylenie standardowe	wykorzystanie budżetu
K -iteracji	0.19	0.11	$\approx 96.63\%$
Odchylenie Standardowe	0.2	0.09	$\approx 95.27\%$
Najlepszy-Najgorszy	0.21	0.1	$\approx 93.59\%$
Wariancja Dopasowania	0.18	0.1	$\approx 96.89\%$

Tablica 13: Porównanie kryteriów przerwania dla F_4 .

kryterium	średnie dopasowanie	odchylenie standardowe	wykorzystanie budżetu
K -iteracji	3.32	1.33	$\approx 94.65\%$
Odchylenie Standardowe	3.48	1.38	$\approx 93.03\%$
Najlepszy-Najgorszy	3.19	1.22	$\approx 98.47\%$
Wariancja Dopasowania	3.84	1.82	$\approx 92.95\%$

Tablica 14: Porównanie kryteriów przerwania dla F_5 .

kryterium	średnie dopasowanie	odchylenie standardowe	wykorzystanie budżetu
K -iteracji	364.44	918.76	$\approx 65.22\%$
Odchylenie Standardowe	631.97	1646.65	$\approx 76.59\%$
Najlepszy-Najgorszy	420.16	1069.83	$\approx 86.58\%$
Wariancja Dopasowania	409.4	1225.4	$\approx 50.01\%$

Tablica 15: Porównanie kryteriów przerwania dla F_6 .

8 Wnioski

F4

Według zebranych wyników, w celu osiągnięcia najlepszego wyniku dla F_4 najlepiej zastosować kryterium Wariancji Dopasowania z parametrem $\epsilon \approx 0.0004$. Niemniej wykorzystane zostało w takim wypadku $\approx 97\%$ budżetu. Takie zużycie może nie być zadowalające, ponieważ czas wykonania jest długi. W takim wypadku możemy zastosować inne kryteria lub zmienić parametry obecnego i uzyskać nieco gorsze wyniki, przy znacznym ograniczeniu zużycia budżetu. Jak widać poniżej średnie dopasowanie dla kryterium Najlepszy-Najgorszy z parametrem $\epsilon = 0.12$ jest niewiele gorsze, a wykorzystany budżet jest niższy o ponad 20 punktów procentowych.

kryterium	średnie dopasowanie	odchylenie standardowe	wykorzystanie budżetu
Najlepszy-Najgorszy	0.28	0.17	$\approx 73\%$
K -iteracji	0.26	0.14	$\approx 85\%$
Wariancja Dopasowania	0.21	0.12	$\approx 90\%$

Tablica 16: Porównanie wyników według różnego zużycia budżetu dla F_4 .

F5

W celu osiągnięcia najlepszego wyniku dla F_5 najlepiej zastosować kryterium Najlepszy-Najgorszy z parametrem $\epsilon \approx 0.8$. Wykorzystujemy wtedy prawie cały budżet $\approx 98\%$. Porównywalne wyniki możemy także uzyskać zmniejszając nieco zużycie budżetu - stosując inne kryteria bądź inne parametry. Na przykład według Tablicy 11, dla kryterium Wariancji Dopasowania z parametrem $\epsilon = 0.15$ (pierwszy wiersz poniższej tabeli) najlepszy znaleziony wynik jest bardzo bliski optimum przy zdecydowanie mniejszym wykorzystaniu budżetu. Porównywalne wyniki zostały przedstawione w tabeli poniżej.

kryterium	średnie dopasowanie	odchylenie standardowe	wykorzystanie budżetu
Wariancja dopasowania	4.47	2.78	$\approx 73\%$
Odchylenie standardowego	3.89	1.94	$\approx 84\%$
Wariancja Dopasowania	3.84	1.82	$\approx 93\%$

Tablica 17: Porównanie wyników według różnego zużycia budżetu dla F_5 .

F6

Dla funkcji F_6 najlepszy wynik został osiągnięty przy relatywnie niskim wykorzystaniu budżetu (w porównaniu do pozostałych kryteriów), przy zastosowaniu kryterium K -iteracji. Funkcja F_6 jest multimodalna, stąd wysokie odchylenia standardowe oraz różnorodność uzyskiwanych wyników. Warto zauważyć, że dłuższe poszukiwania nie w każdym przypadku oznaczały lepszy wynik, szczególnie widoczne jest to w zestawieniu danych w Tablicy 15. Zbliżony jakościowo wynik uzyskaliśmy także przy zastosowaniu kryterium Wariancji Dopasowania, jednakże przy znacznie mniejszym wykorzystaniu budżetu. Optymalizacja F_6 jest typowym przypadkiem, w którym poszukujemy ogólnie dobrego rozwiązania, niekoniecznie optimum globalnego. W tabeli poniżej zostały przedstawione inne akceptowalne uzyskane średnie wartości dla mniejszego wykorzystania budżetu przeszukiwań.

kryterium	średnie dopasowanie	odchylenie standardowe	wykorzystanie budżetu
K -iteracji	1090.58	2400.09	$\approx 42\%$
Wariancja Dopasowania	560.79	1582.09	$\approx 47\%$
Wariancja Dopasowania	409.4	1225.4	$\approx 50\%$

Tablica 18: Porównanie wyników według różnego zużycia budżetu dla F_6 .

Dyskusja


Z natury algorytmów ewolucyjnych, im dłużej algorytm pracuje, tym lepsze wyniki osiąga. Zgodnie z Rysunkami 1 i 2 algorytmy ewolucyjne na początku uruchomienia w szybkim tempie poprawiają swój wynik i odnajdują rozwiązania akceptowalne. Jednak w późniejszej fazie przeszukiwań coraz trudniej jest znaleźć lepsze rozwiązanie, dlatego tempo poprawy spada znacznie i odnalezienie każdego lepszego rozwiązania trwa coraz dłużej. Ponadto zauważalne są sytuacje, w których po wydłużeniu obliczeń wyniki nie są lepsze, jeżeli algorytm nie trafi na lepsze rozwiązanie i utknie w dobrym optimum lokalnym, może się okazać, że już jego nie opuści. W takim wypadku większe korzyści może przynieść przerwanie pracy algorytmu wcześniej, aby wykorzystać pozostały budżet z danego uruchomienia w innym celu (np. w nowym uruchomieniu z mniejszym budżetem lub w kolejnym uruchomieniu z powiększonym budżetem, jeśli budżet jest np. czasowy). Stąd pomysł na stosowanie kryteriów stopu.

Przedwczesne zakończenie pracy algorytmu według ustalonego kryterium może być bardziej rozsądnym i korzystnym rozwiązaniem. Wszystko zależy od tego, jakie wyniki chcemy uzyskać. Jeśli celem jest otrzymanie dobrego, ale niekoniecznie najlepszego rozwiązania i ważnym czynnikiem jest czas, stosowanie kryteriów przerwania algorytmu jest wskazane, ponieważ umożliwia sprecyzowanie wymagań i optymalizację procesu wyszukiwania rozwiązania.

9 Podsumowanie

Projekt wprowadzający w tematykę algorytmów ewolucyjnych. Dzięki implementacji przykładowej strategii ewolucyjnej poznano istotę algorytmów ewolucyjnych. Implementacja przykładowych kryteriów przerywania pracy algorytmu pozwoliła na analizę procesu optymalizacji funkcji i praktycznego zastosowania tychże przerw.

10 Powiązane linki

 [Repozytorium projektowe](#)

11 Bibliografia

GHOREISHI, Seyyedeh Newsha; CLAUSEN, Anders; JØRGENSEN, Bo Nørregaard. Termination Criteria in Evolutionary Algorithms: A Survey. In: IJCCI. 2017. p. 373-384.

SUGANTHAN, Ponnuthurai N., et al. Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization. KangAL report, 2005, 2005005.2005: 2005.