

[PSZT-P] Kiedy się poddać?

Piotr Frątczak

Bartosz Świtalski

25 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 25 wywołań algorytmu.

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$
```

```
python3 when-to-surrender/main.py
```

```
<funkcja> <kryterium> <p1> <p2> <p3> <p4>
```

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.

Użycie skryptu

```
/when-to-surrender$
```

```
./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 (8 różnych). Pojedyncze uruchomienie skryptu jest dość kosztowne czasowo (>2h), 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`. Do czyszczenia służy skrypt `xclean.sh`.

5 Kryteria przerywań

- **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 $\epsilon \geq 0$.

¹Shifted Schwefel's Problem 1.2 with Noise in Fitness

²Schwefel's Problem 2.6 with Global Optimum on Bounds

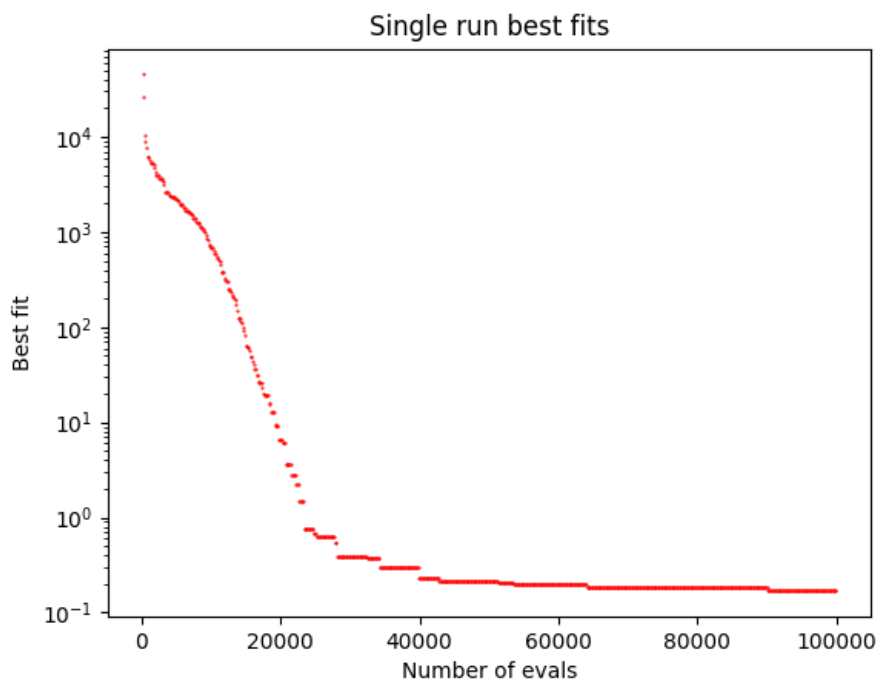
³Shifted Rosenbrock's Function

- **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 $1 \gg \epsilon \geq 0$.

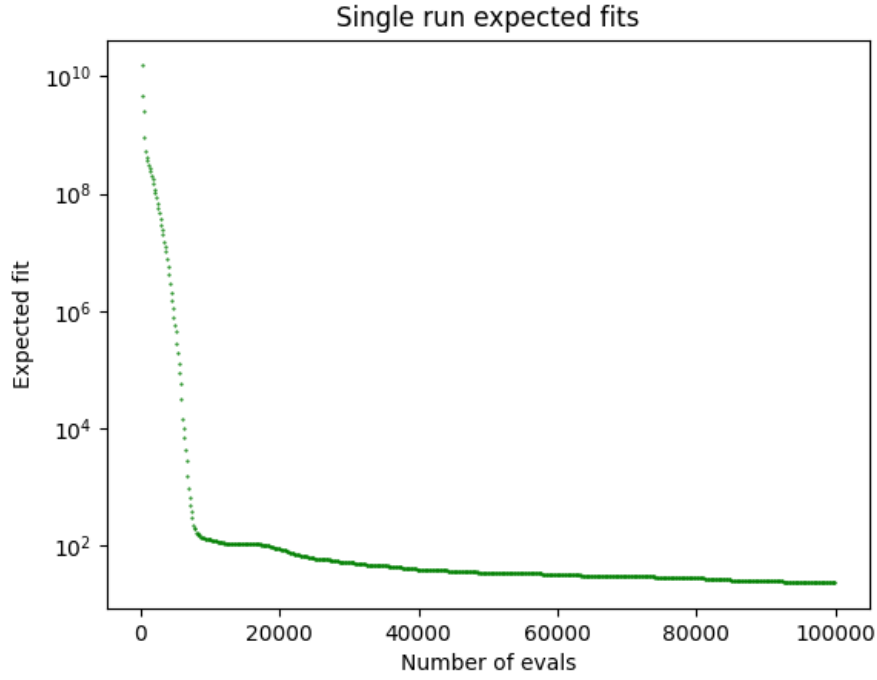
6 Uzyskane wyniki

Uzyskane wyniki dla uruchomień funkcji według każdego kryterium z różnymi parametrami zostały zebrane w pojedynczy plik `output.pdf` i zostaną omówione w dalszej części tego sprawozdania.

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 .

7 Wnioski

Objaśnienia

k value - liczba K kolejnych iteracji w kryterium K -iteracji

best fit - najlepsze znalezione optimum

best fit mean - średnia najlepszych znalezionych optimów

best fit std. deviation - odchylenie standardowe najlepszych znalezionych optimów

number of evals mean - średnia liczba ewaluacji funkcji celu

Wnioski do kryterium K -iteracji

k value	best fit	best fit mean	best fit std. deviation	number of evals mean
≈ 100	0.5	0.24	0.14	$\approx 4/5$ of budget

Tablica 1: Najlepsza wartość parametru k dla F_4 - Shifted Schwefel's Problem 1.2 with Noise in Fitness

k value	best fit	best fit mean	best fit std. deviation	number of evals mean
≈ 150	1.75	3.48	1.31	$\approx 9/10$ of budget

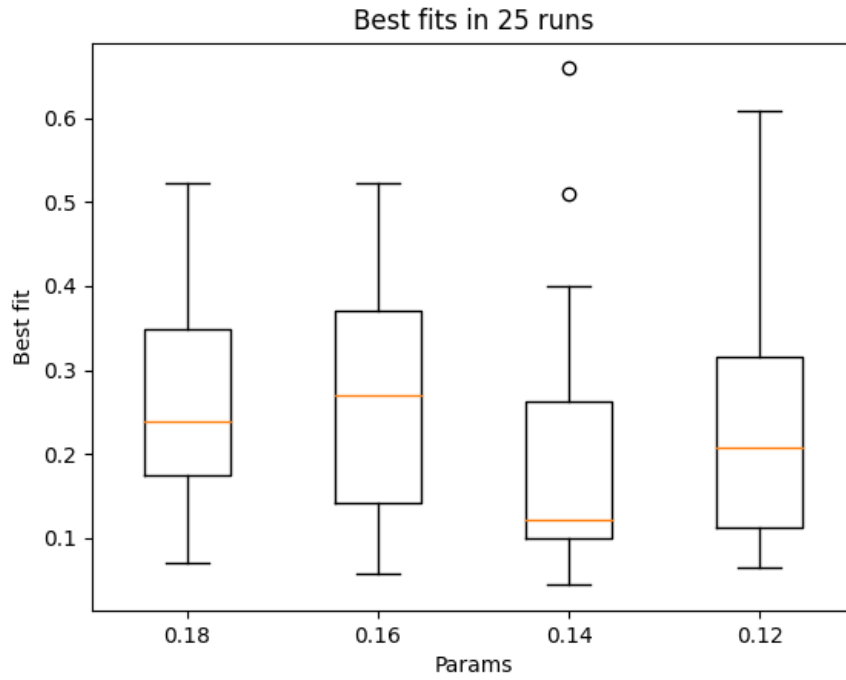
Tablica 2: Najlepsza wartość parametru k dla F_5 - Schwefel's Problem 2.6 with Global Optimum on Bounds

k value	best fit	best fit mean	best fit std. deviation	number of evals mean
≈ 75	7.77	445.8	1355.54	$\approx 1/2$ of budget

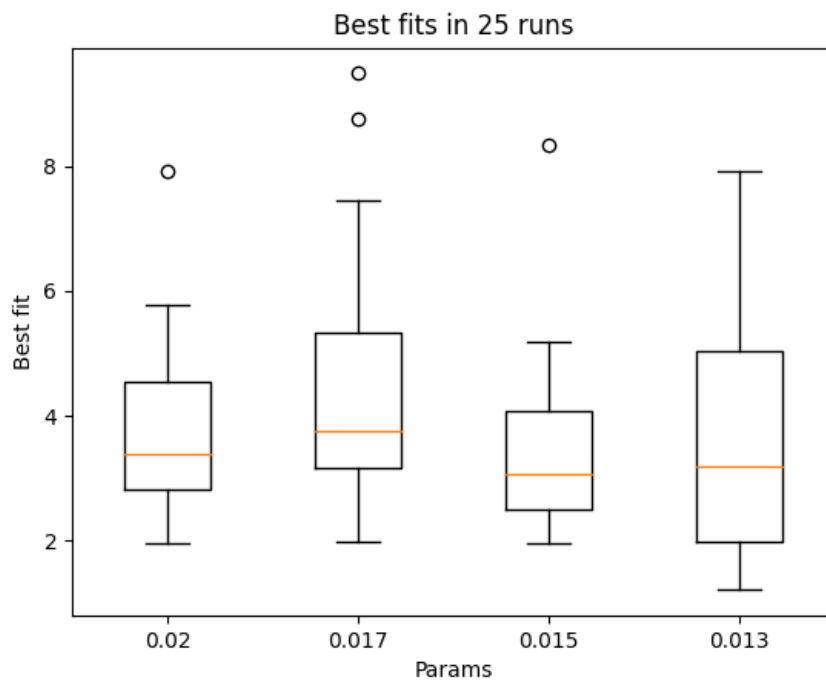
Tablica 3: Najlepsza wartość parametru k dla F_6 - Shifted Rosenbrock's Function

Restarty wg kryterium K -iteracji mają sens. Nie wykorzystując budżetu w pojedynczym uruchomieniu możemy wykorzystać niewykorzystane ewaluacje funkcji w nowym uruchomieniu, co może skutkować uzyskaniem lepszego wyniku (dodatkowe uruchomienie jest możliwe: po 4 uruchomieniach dla F_4 , po 9 uruchomieniach dla F_5 , po 1 uruchomieniu dla F_6).

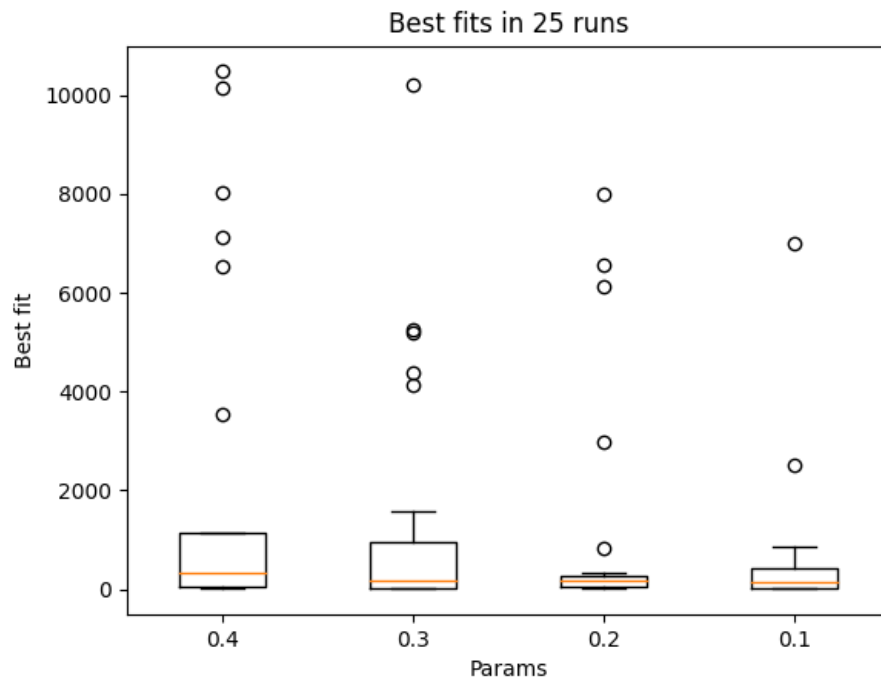
Wnioski do kryterium Odchylenia Standardowego



Rysunek 3: Wartości najlepszego dopasowania funkcji F_4 wg kryterium Odchylenia Standardowego. Najlepszy parametr $\epsilon \approx 0.14$, wykorzystanie $\approx 85\%$ budżetu.



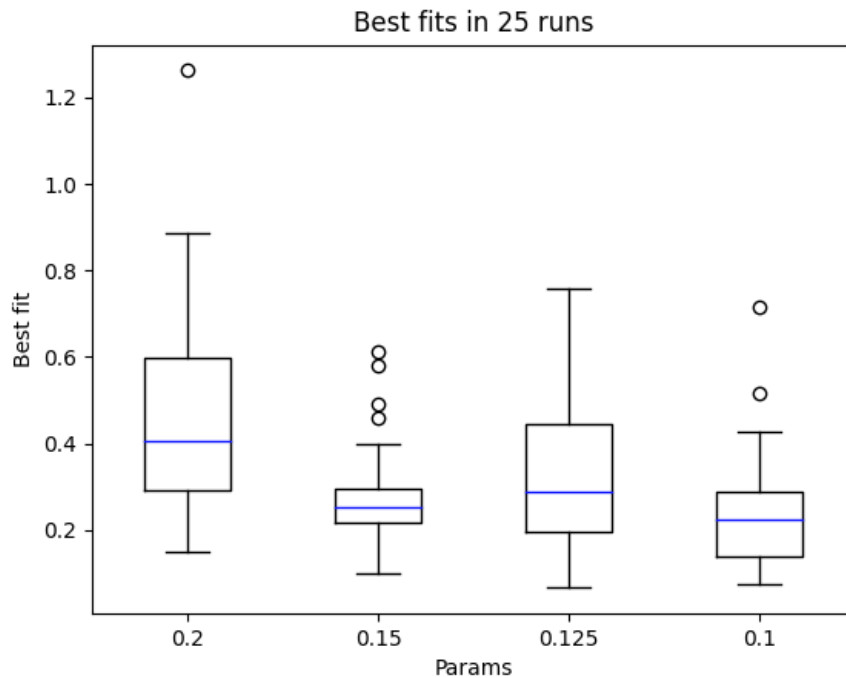
Rysunek 4: Wartości najlepszego dopasowania funkcji F_5 wg kryterium Odchylenia Standardowego. Najlepszy parametr $\epsilon \approx 0.015$, wykorzystanie $\approx 90\%$ budżetu.



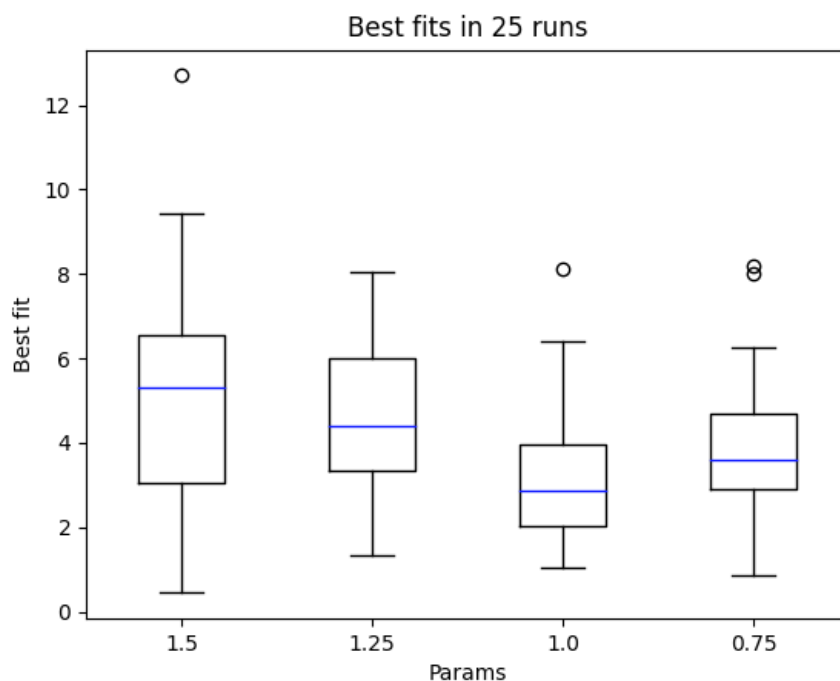
Rysunek 5: Wartości najlepszego dopasowania funkcji F_6 wg kryterium Odchylenia Standardowego. Najlepszy parametr $\epsilon \approx 0.2$, wykorzystanie $\approx 70\%$ budżetu.

Przerwania wg kryterium Odchylenia Standardowego również mają sens.

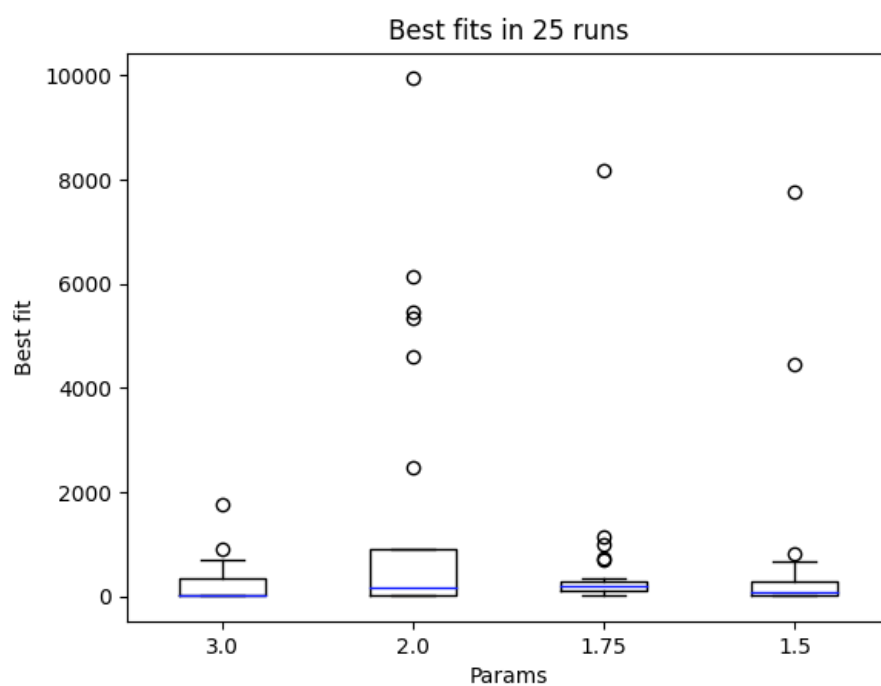
Wnioski do kryterium Najlepszy-Najgorszy



Rysunek 6: Wartości najlepszego dopasowania funkcji F_4 wg kryterium Najlepszy-Najgorszy. Najlepszy parametr $\epsilon \approx 0.15$, wykorzystanie $\approx 80\%$ budżetu.



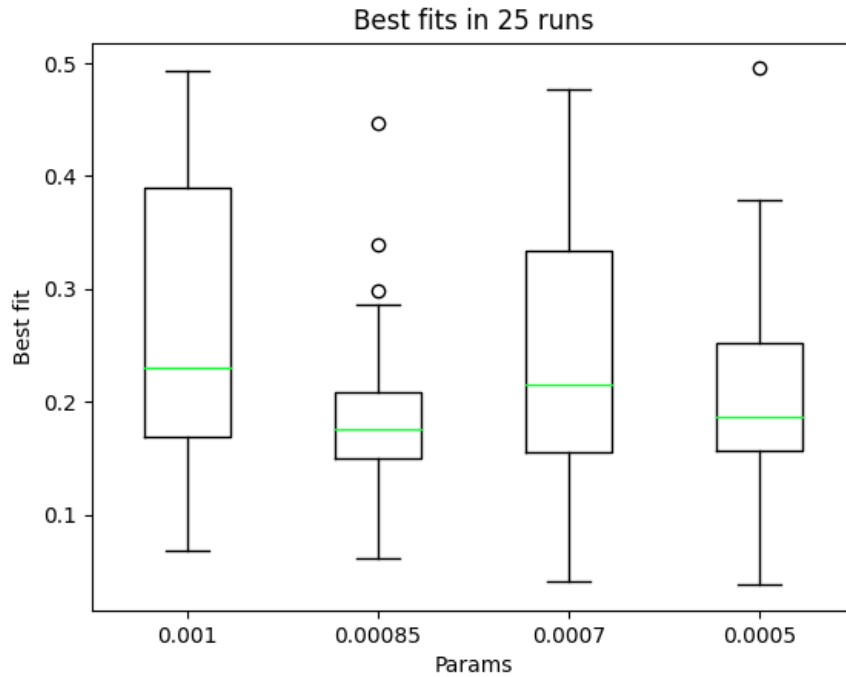
Rysunek 7: Wartości najlepszego dopasowania funkcji F_5 wg kryterium Najlepiej-Najgorszy. Najlepszy parametr $\epsilon \approx 1$, wykorzystanie $\approx 90\%$ budżetu.



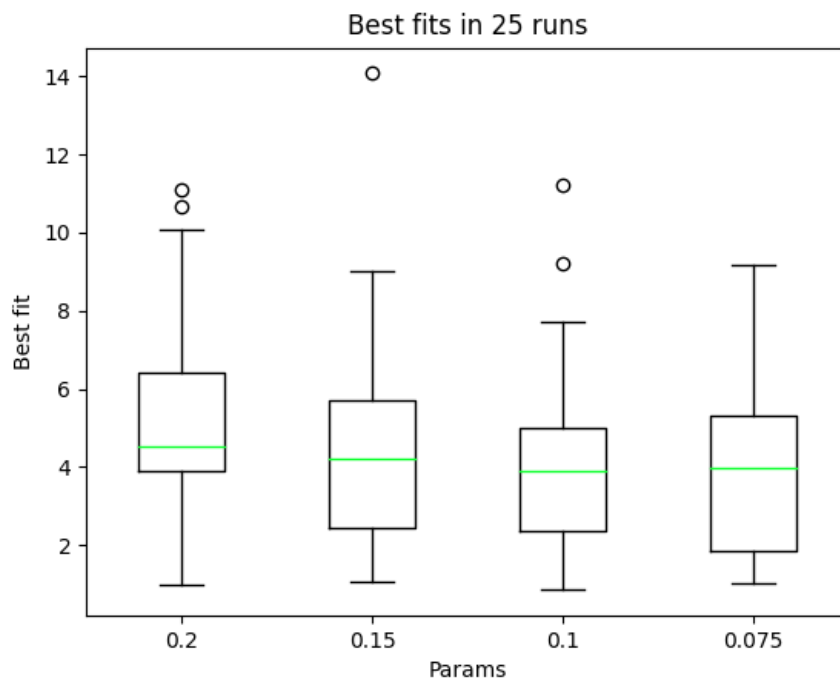
Rysunek 8: Wartości najlepszego dopasowania funkcji F_6 wg kryterium Najlepiej-Najgorszy. Najlepszy parametr $\epsilon \approx 3$, wykorzystanie $\approx 45\%$ budżetu.

Przerwania wg kryterium Najlepiej-Najgorszy również mają sens.

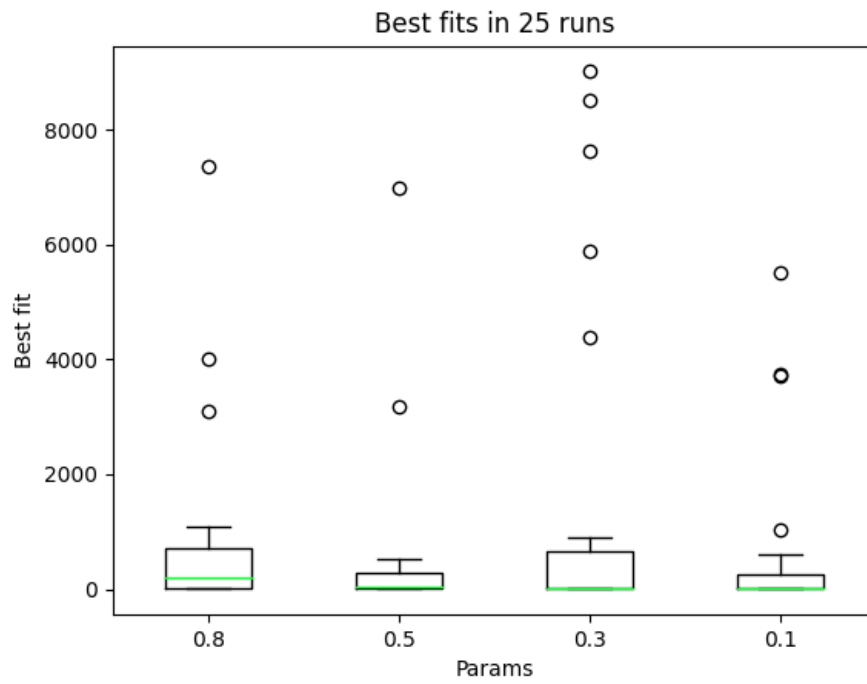
Wnioski do kryterium Wariacji Dopasowania



Rysunek 9: Wartości najlepszego dopasowania funkcji F_4 wg kryterium Wariacji Dopasowania. Najlepszy parametr $\epsilon \approx 0.00085$, wykorzystanie $\approx 80\%$ budżetu.



Rysunek 10: Wartości najlepszego dopasowania funkcji F_5 wg kryterium Wariacji Dopasowania. Najlepszy parametr $\epsilon \approx 0.1$, wykorzystanie $\approx 85\%$ budżetu.



Rysunek 11: Wartości najlepszego dopasowania funkcji F_6 wg kryterium Wariacji Dopasowania. Najlepszy parametr $\epsilon \approx 0.5$, wykorzystanie $\approx 50\%$ budżetu.

Przerwania wg kryterium Wariacji Dopasowania również mają sens.

Dyskusja

kryterium	średnie dopasowanie	wykorzystanie budżetu
K -iteracji	0.24	$\approx 80\%$
Odchylenie Standardowe	0.21	$\approx 85\%$
Najlepszy-Najgorszy	0.25	$\approx 80\%$
Wariancja Dopasowania	0.19	$\approx 80\%$

Tablica 4: Porównanie kryteriów przerwań dla F_4 .

kryterium	średnie dopasowanie	wykorzystanie budżetu
K -iteracji	3.48	$\approx 90\%$
Odchylenie Standardowe	3.21	$\approx 90\%$
Najlepszy-Najgorszy	3.78	$\approx 90\%$
Wariancja Dopasowania	3.9	$\approx 85\%$

Tablica 5: Porównanie kryteriów przerwań dla F_5 .

kryterium	średnie dopasowanie	wykorzystanie budżetu
K -iteracji	445.8	$\approx 50\%$
Odchylenie Standardowe	957.59	$\approx 70\%$
Najlepszy-Najgorszy	1081.83	$\approx 45\%$
Wariancja Dopasowania	380.98	$\approx 50\%$

Tablica 6: Porównanie kryteriów przerwań dla F_6 .


Wyniki są uśrednieniem z czterech niezależnych badań 25 uruchomień z określonym parametrem.

Dla funkcji F_4 najlepiej zastosować kryterium Wariancji Dopasowania. Dla funkcji F_5 najlepiej zastosować kryterium Odchylenia Standardowego. Na pierwszym rzut oka dla funkcji F_6 najlepiej zastosować kryterium Wariancji Dopasowania, jednakże przy większej liczbie niezależnych badaniach 25 uruchomień okazuje się, że funkcja F_6 jest całkiem nieprzewidywalna, dlatego ciężko określić właściwe dla niej kryterium przerwań.

8 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 sensowności stosowania tychże przerwań.

9 Powiązane linki

 [Repozytorium projektowe](#)

10 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.