

Sprawozdanie

Ćw. 9 – Algorytm ewolucyjny (AE)

Filip Horst 311257

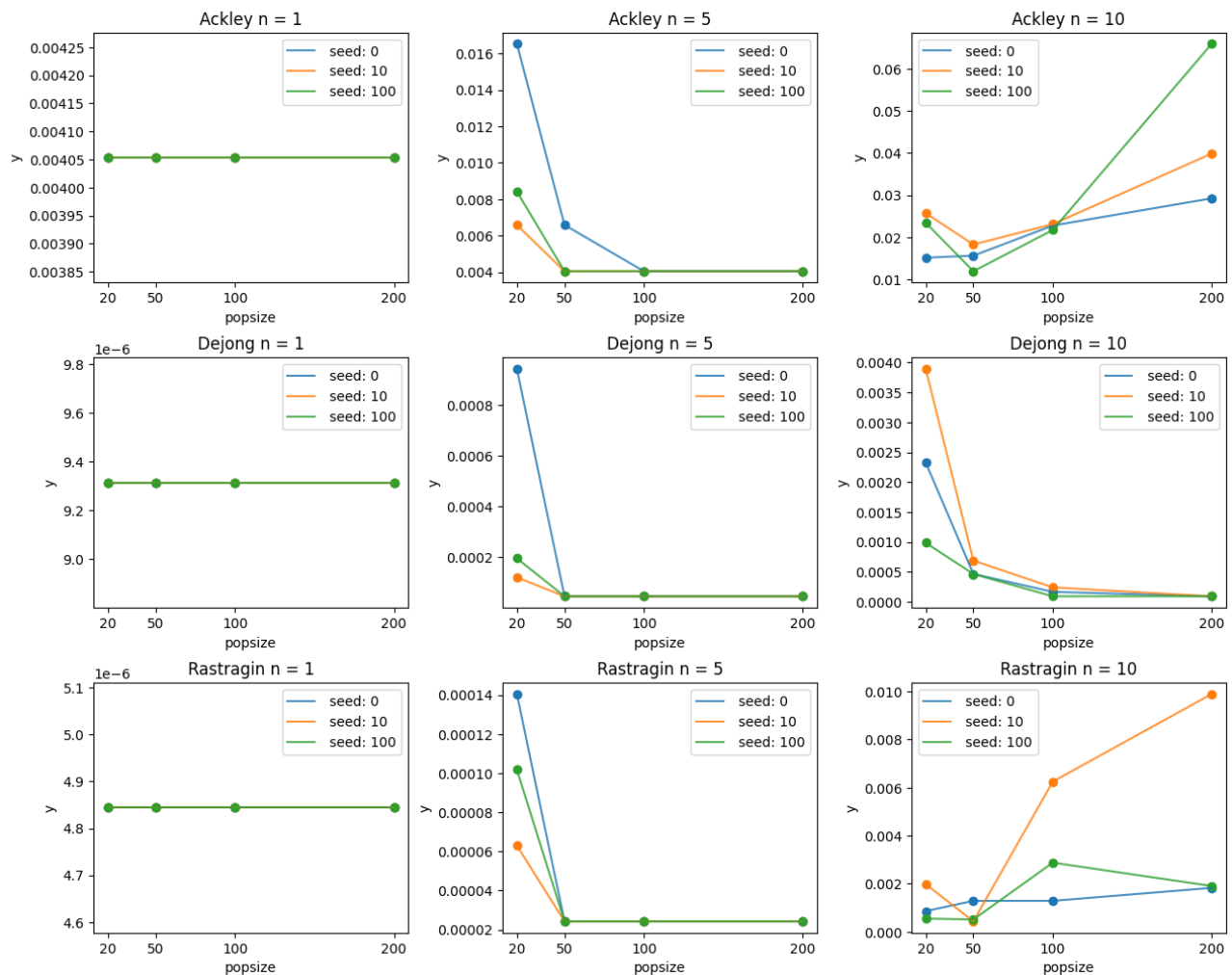
1 Badanie wpływu parametrów na wyniki

Badania polegały na wykonaniu automatycznego przesiewu ustawień, gdzie zmieniany był tylko jeden parametr jednocześnie. Domyślnymi ustawieniami były:

- Popsize 50
- Maxgens 20000
- Pxover 0.5
- Pmutation 0.07
- Krzyz_arytm 0
- Skal 1
- Elit 1

1.1 Popsze

Parametr popsize



Zależność wyniku od ustawienia parametru dla różnych funkcji, wymiarowości i ziaren losowych

Na wykresach można zaobserwować kilka zależności. Ustawienie optymalne zależy od złożoności problemu. Dla jednego wymiaru niezależnie od rozmiaru populacji AE osiąga bardzo dobry wynik. W przypadku 5 wymiarów widać, że zwiększenie rozmiaru z 20 do 50 poprawia skuteczność AE, ale dalsze zwiększanie już nic nie daje. Najciekawsze są funkcje 10-wymiarowe. W przypadku funkcji DeJonga zachowanie jest analogiczne do 5-wymiarowych, ale dla funkcji Ackley'a oraz Rastragina zauważalne jest pogorszenie wyników dla zbyt wysokich ustawień. Dla tych dwóch funkcji wyniki poprawiają się przy zwiększeniu populacji z 20 do 50, ale przy powiększeniu do 100 wartość się pogarsza (poza jednym losowym przypadkiem seed 0 dla f. Rastragina). Potencjalną hipotezą jest to, że duża populacja zwiększa element eksploracji algorytmu, a w przypadku funkcji Rastragina oraz Ackleya, które posiadają ogromną liczbę minimów lokalnych sprawia to, że populacja jest rozrzucona po wielu z nich i ma problemy ze zbiegiem do najlepszego. W przypadku f. DeJonga (nr 1) nie występuje takie zjawisko, ponieważ nie ma ona minimów lokalnych.

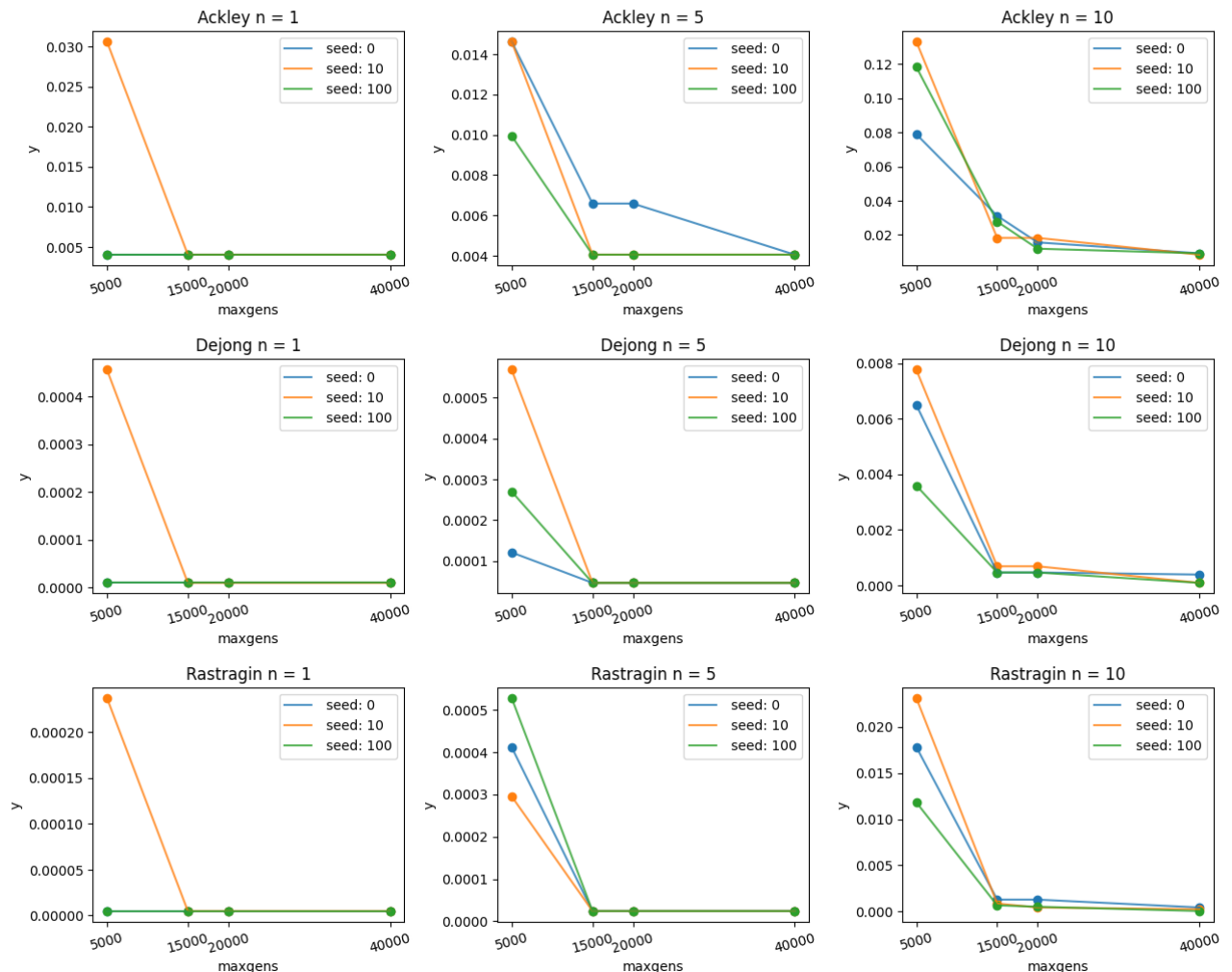
Podsumowując:

- Liczebność populacji to narzędzie manipulacji elementem eksploracji AE
- Istnieje możliwość nasycenia się liczby osobników, gdzie dalsze ich zwiększanie nic nie zmienia

- Dla funkcji z wieloma minimami lokalnymi może wystąpić zjawisko pogorszenia się wyników wraz ze zwiększaniem populacji – prawdopodobnie spowodowane słabą zbieżnością wynikającą z połączenia dużej eksploracji z dużą liczbą minimów lokalnych

1.2 Maxgens

Parametr maxgens

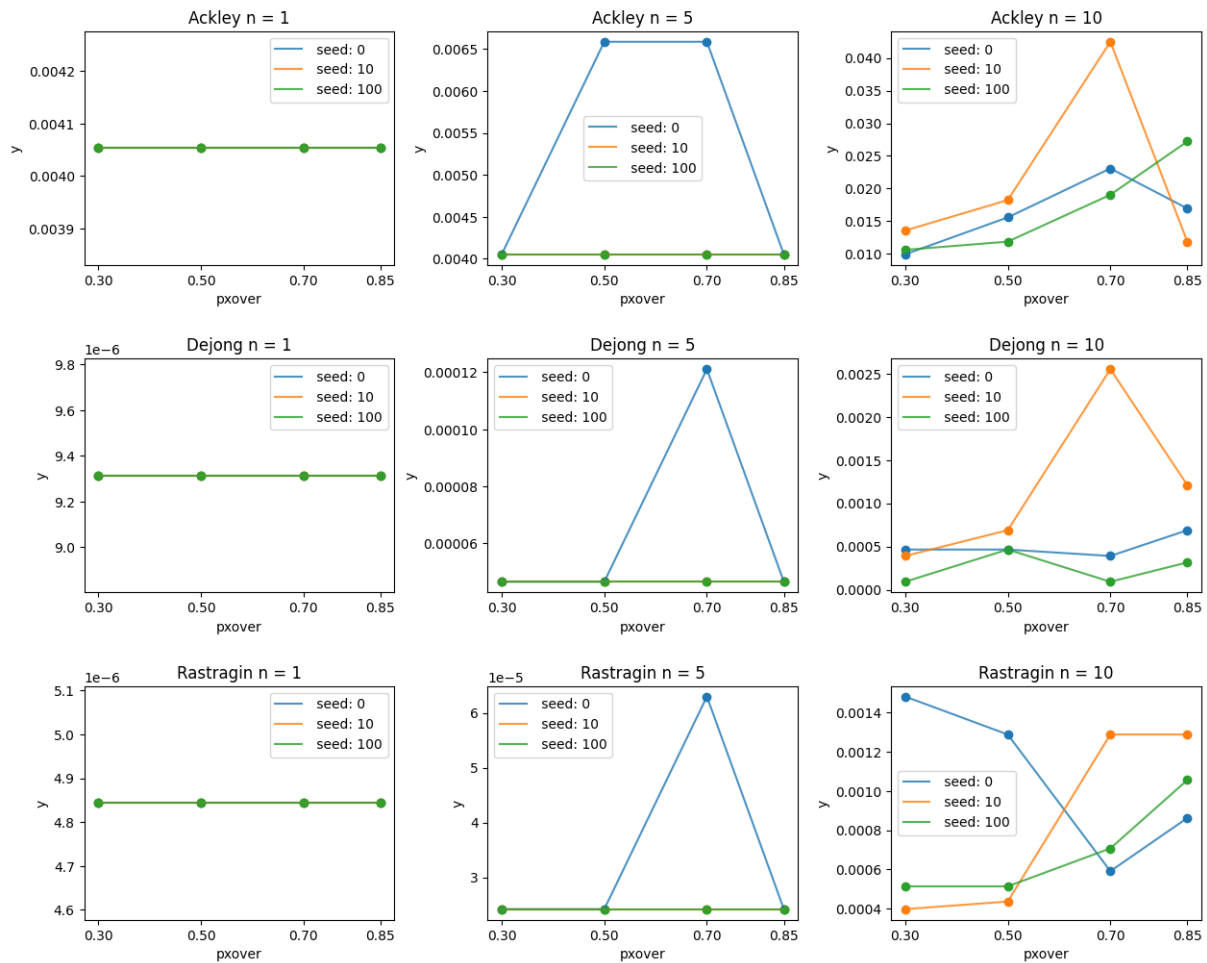


Zależność wyniku od ustawienia parametru dla różnych funkcji, wymiarowości i ziaren losowych

W przypadku maxgens zależność jest oczywista – im więcej tym lepiej. Istnieje jednak pewien próg powyżej którego zysk z dodatkowych iteracji jest bardzo mały, bądź dosłownie zerowy. Wynika to z tego, że po znalezieniu rozwiązania bardzo bliskiego optimum dojście jeszcze bliżej jest blokowane przez inne parametry, niedoskonałości reprezentacji liczb oraz czynniki losowe związane z krzyżowaniem i mutacją – czasem populacja po prostu nie może osiągnąć dokładnego optimum.

1.3 Pxover

Parametr pxover



Zależność wyniku od ustawienia parametru dla różnych funkcji, wymiarowości i ziaren losowych

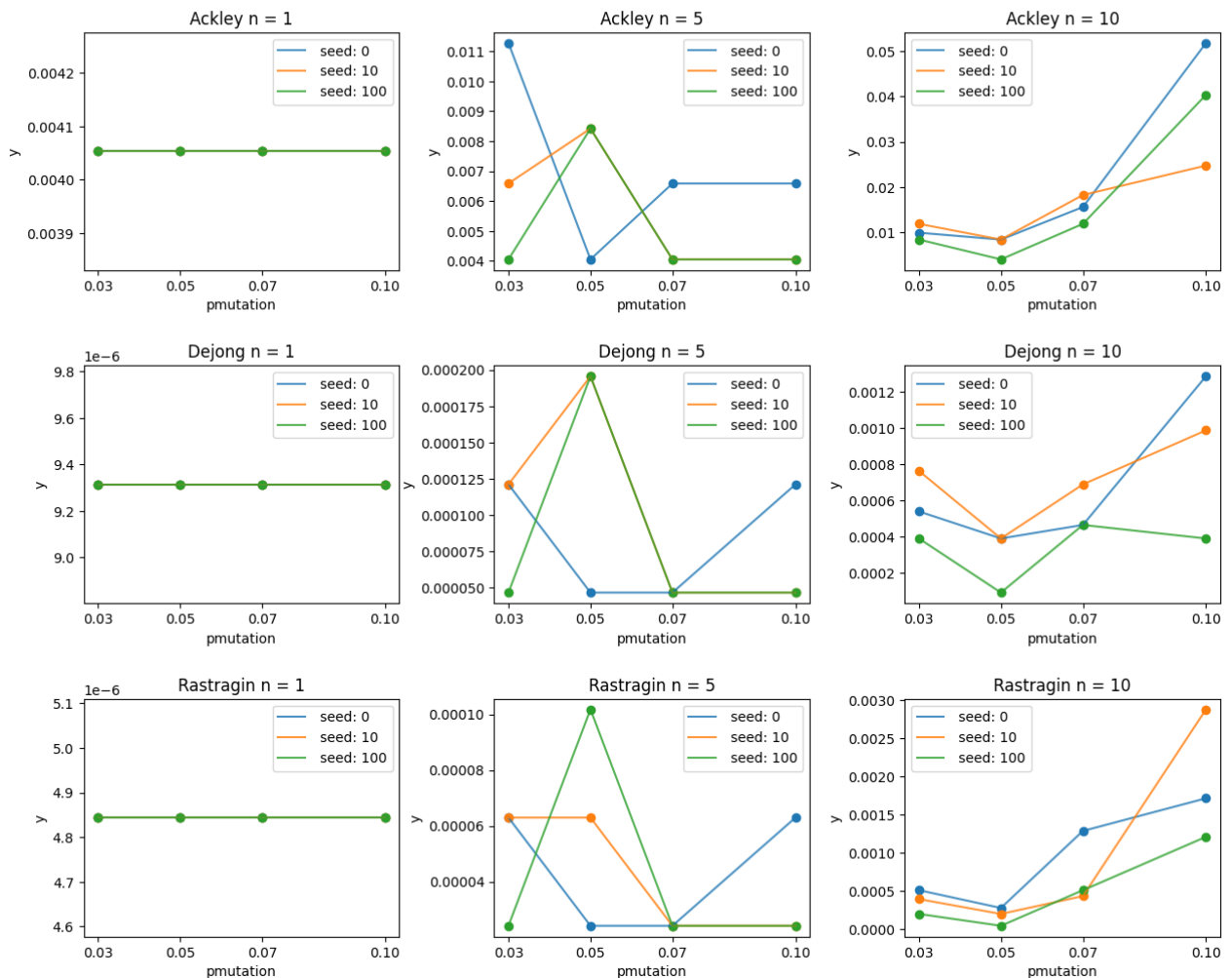
Na podstawie zebranych wyników ciężko ocenić działanie p. krzyżowania w algorytmie. Dla funkcji 1 oraz 5 wymiarowych ustawienie nie ma większego znaczenia poza przypadkiem ziarna = 0 dla f. 5-wymiarowych gdzie dochodziło do dziwnych zjawisk losowych (wada małej próbki badawczej).

Najwięcej mówią wykresy dla f. 10-wymiarowych. Dla trudnych Ackley'a oraz Rastragin'a doszło do niespodziewanego zjawiska – zwiększanie prawdopodobieństwa z wartości 0.3 powodowało pogorszenie wyników. Jednak dla f. Ackley'a z ziarnami 10 oraz 0 zwiększenie do 0.85 ponownie je poprawiało, a w szczególności dla ziarna 10 doprowadziło do najlepszych wyników.

Już sama obecność dużego wpływu ziarna generatora na zależności sprawia, że badanie wpływu p. krzyżowania zaproponowaną metodą jest mało wiarygodne.

1.4 Pmutation

Parametr pmutation



Zależność wyniku od ustawienia parametru dla różnych funkcji, wymiarowości i ziaren losowych

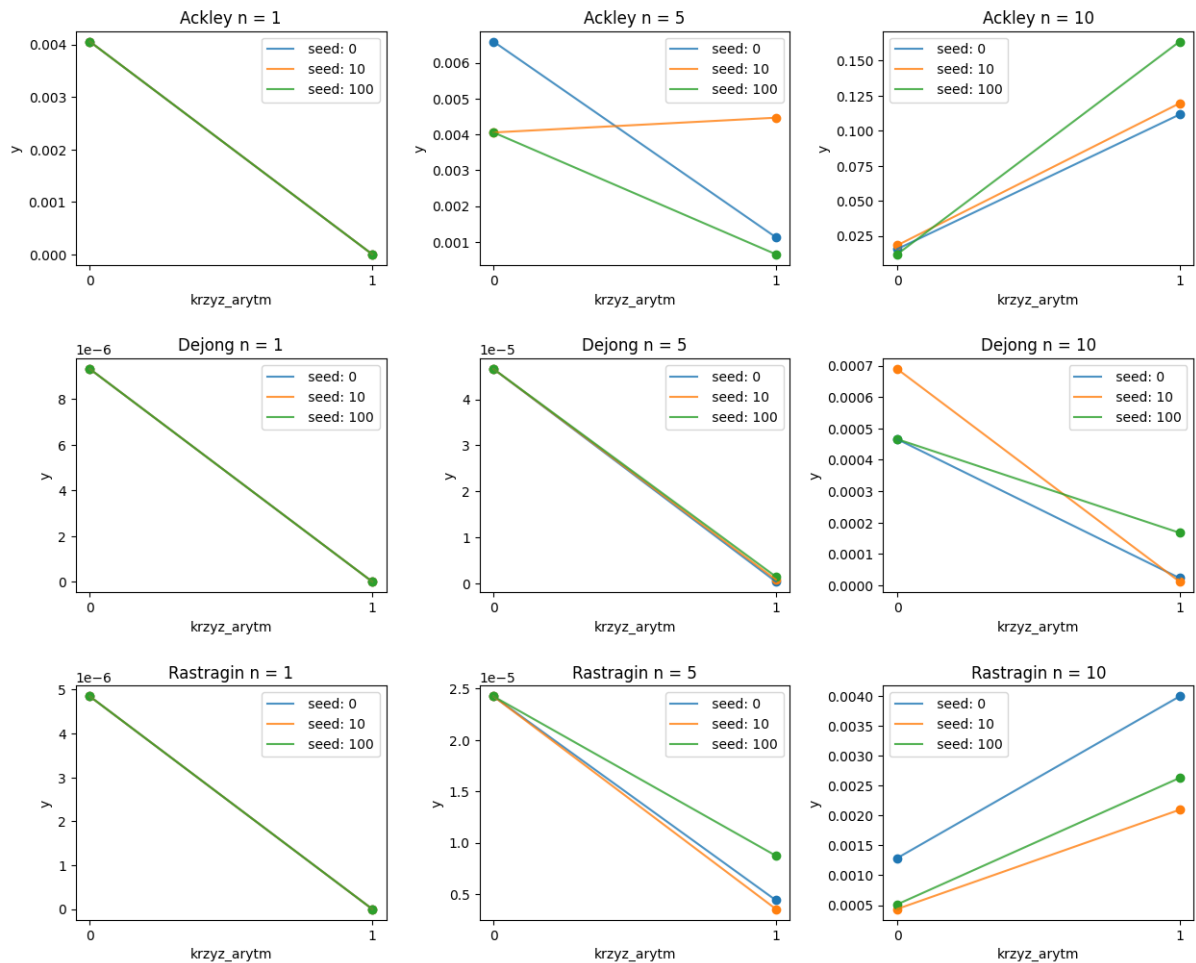
Dla funkcji 5-wymiarowych doszło do bardzo dziwnego zjawiska, gdzie charakterystyka danego ziarna losowego jest podobna niezależnie od funkcji. Powodem istnienia takiego zjawiska w wizualizacji jest zbyt mała próbka badawcza.

W zależnościach funkcji 10-wymiarowych można zaobserwować, że zarówno zbyt mała, jak i zbyt duża wartość pogarsza działanie. Istnieje pewne optimum, gdzie mutacje są wystarczająco duże, by pomóc z minimami lokalnymi, ale wciąż nie na tyle duże by powodować losowe przeszukiwanie i słabą zbieżność.

Zarówno PXOVER, jak i PMUTATION zostały dokładniej zbadane w dalszych etapach sprawozdania, aby upewnić się że działają zgodnie z oczekiwaniami związanymi z wiedzą teoretyczną.

1.5 Krzyz_arytm

Parametr krzyz_arytm

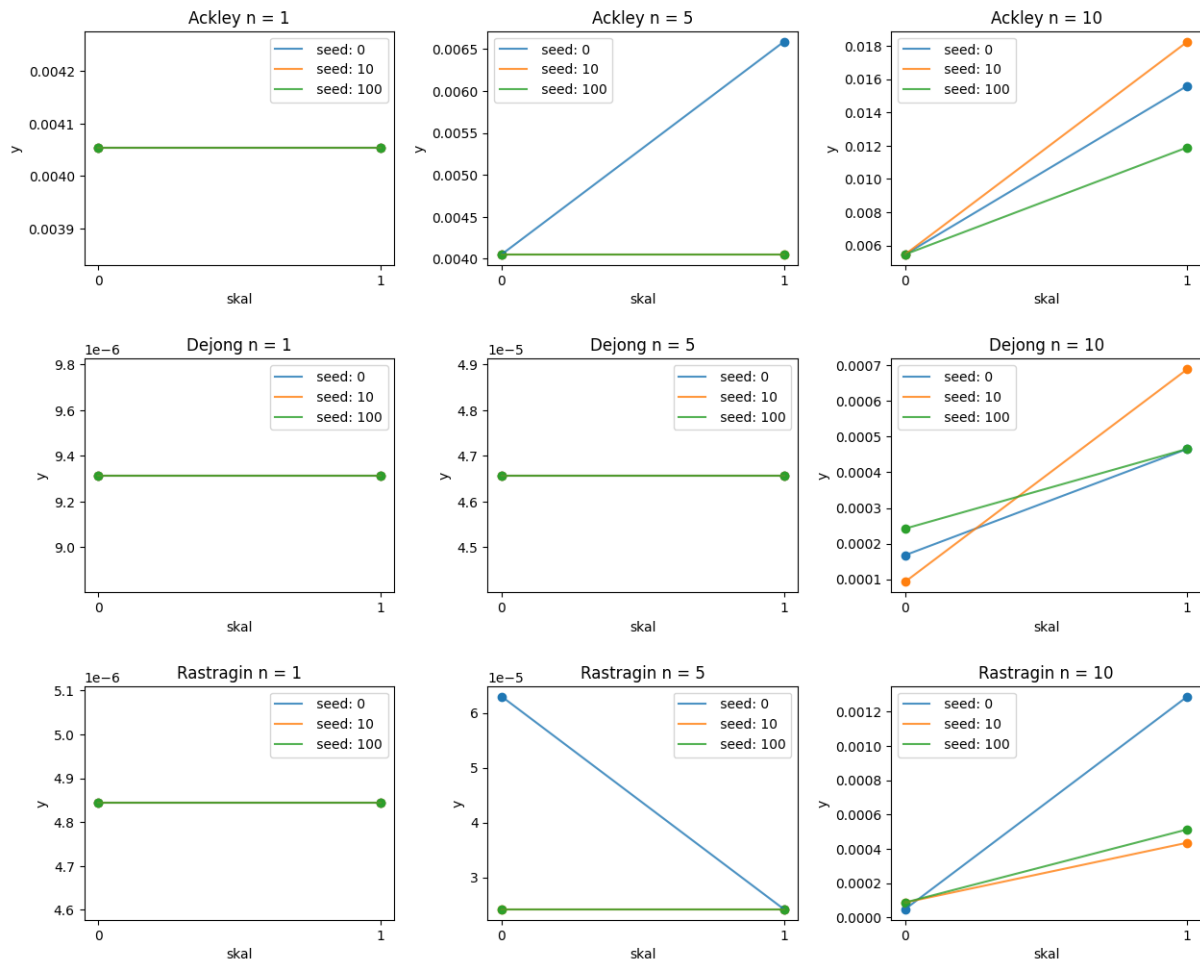


Zależność wyniku od ustawienia parametru dla różnych funkcji, wymiarowości i ziaren losowych

Krzyżowanie arytmetyczne wyraźnie poprawia wyniki dla funkcji 1 i 5 wymiarowych. Jednak dla trudniejszych funkcji 10 wymiarowych wyniki są pogarszane! Można więc powiedzieć, że zastosowanie tego krzyżowania może zarówno poprawić, jak i pogorszyć wyniki i zależy od konkretnego problemu. Ważne jest to, że poprawa/pogorszenie jest zauważalna niezależnie od czynnika losowego, więc można wykonać prostsze wstępne badania, by sprawdzić użyteczność tego podejścia.

1.6 Skal

Parametr skal

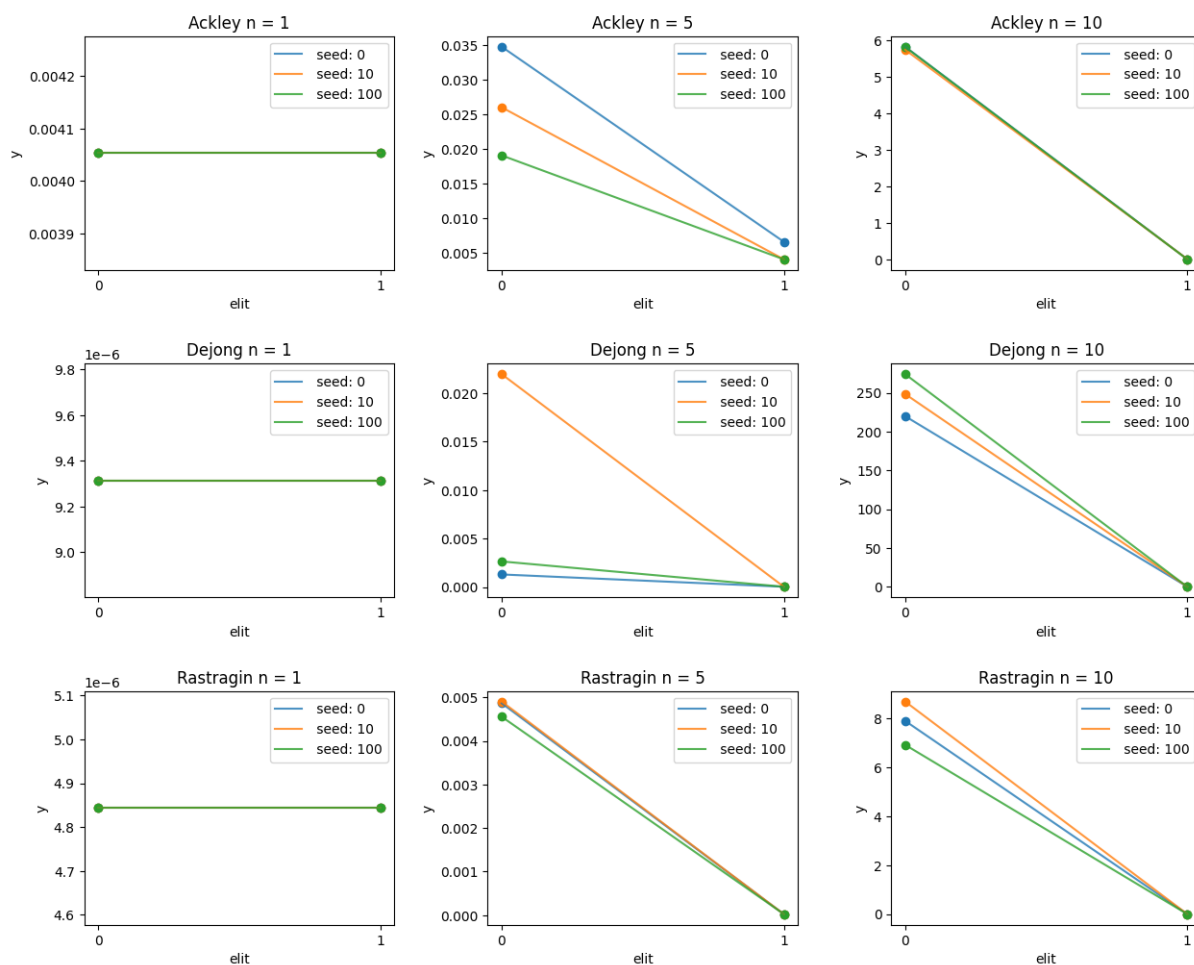


Zależność wyniku od ustawienia parametru dla różnych funkcji, wymiarowości i ziaren losowych

W tym badaniu wyłączenie skalowania na ogół nic nie zmienia lub wręcz pogarsza wyniki, co jest niespodziewanym zjawiskiem. Szczególnie zauważalne jest pogorszenie dla funkcji 10 wymiarowych włącznie z łatwiejszą f. DeJonga. W programie użyte jest skalowanie liniowe razem z selekcją kołem rulety. Jedną z hipotez nietypowego zjawiska jest to, że w tym problemie przy badanych ustawieniach AE więcej zyskuje z dominacji najlepszych osobników (podobnie do opisanego dalej elitaryzmu), niż z ich różnorodności.

1.7 Elit

Parametr elit

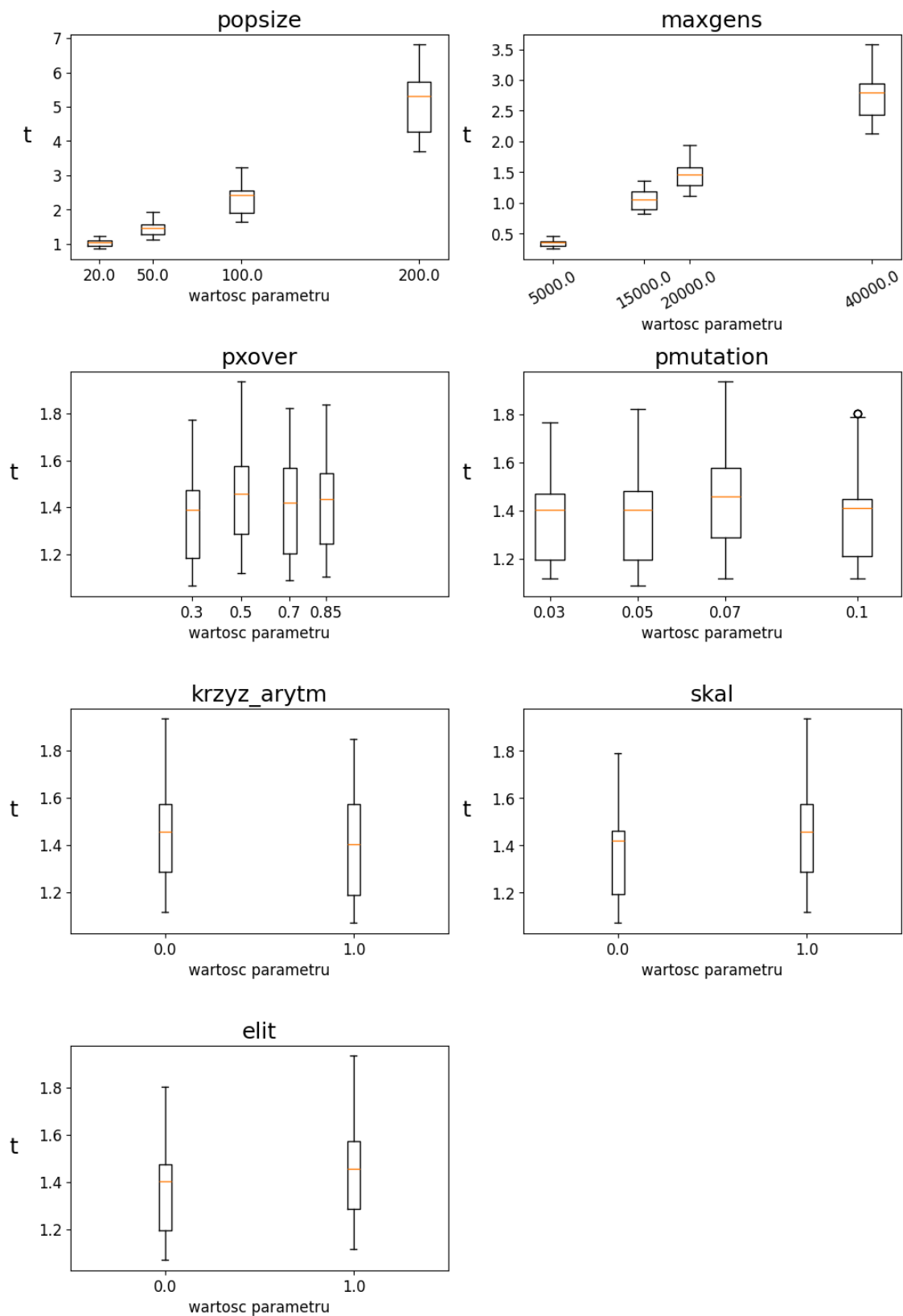


Zależność wyniku od ustawienia parametru dla różnych funkcji, wymiarowości i ziaren losowych

Strategia elitarystyczna zgodnie z oczekiwaniami po prostu poprawia wyniki. Ponadto, im większa złożoność funkcji tym większe są zyski z jej zastosowania – nic dziwnego ponieważ dla wielu wymiarów wartości y dla kolejnych potomków są bardziej różnorodne między iteracjami, dlatego ponowne trafienie w okolice najlepszego jest mniej prawdopodobne niż dla np. jednego wymiaru.

2 Wpływ parametrów na szybkość wykonania

Wpływ ustawień na czas wykonywania



Zależność czasu wykonania od wartości badanego parametru

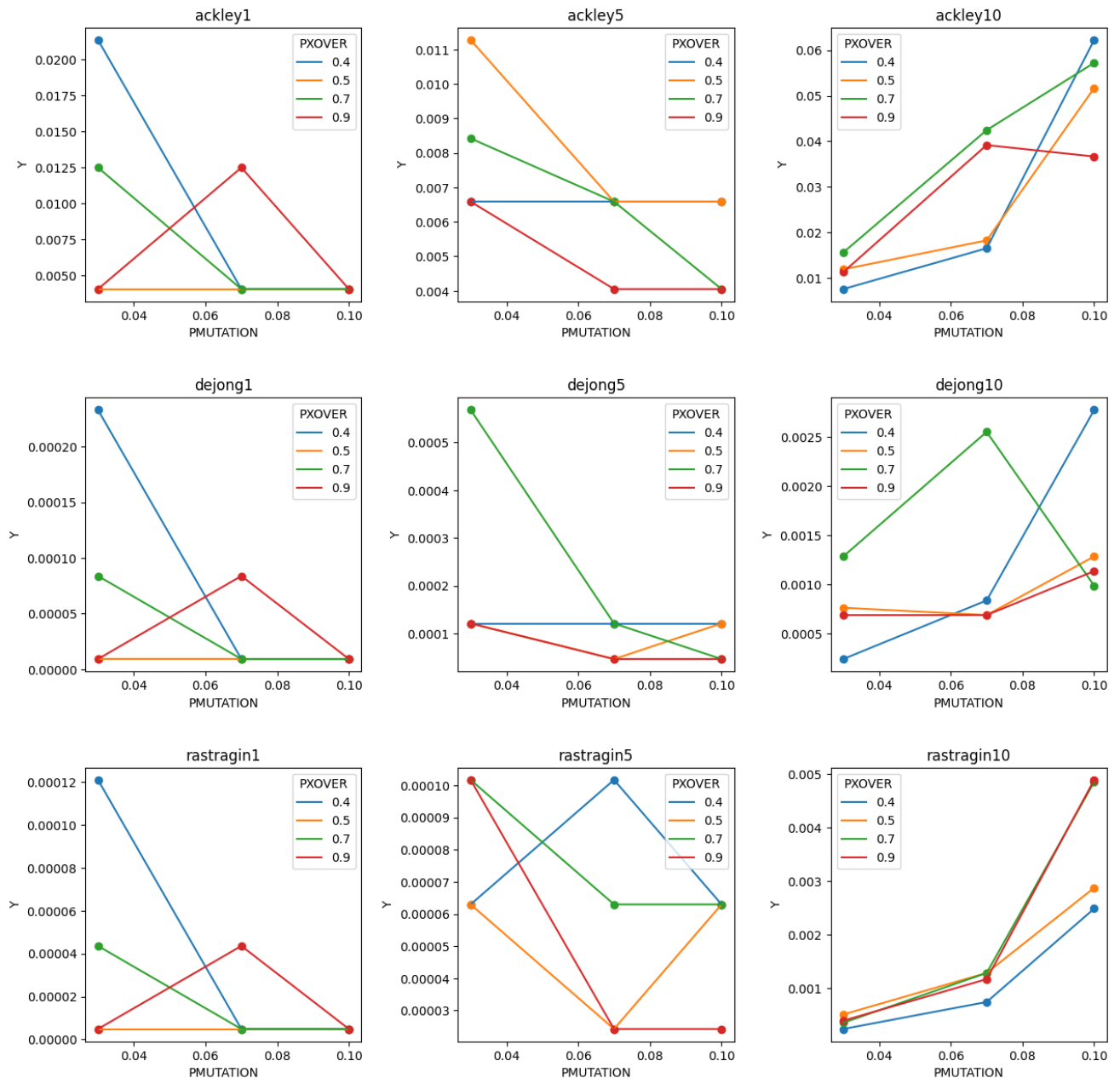
Wyświetlenie czasów wykonania w postaci wykresów pudełkowych zależnych od ustawionej wartości parametru pozwala określić wpływ ustawień na szybkość wykonania. Wnioski najlepiej przedstawić jako tabelkę:

Parametr	Wpływ	Opis
Popsize	Duży	Wyższy popsize = dłuższe działanie Zwiększanie liczebności wymaga większej liczby ewaluacji, krzyżowania i mutacji.
Maxgens	Duży	Więcej iteracji = dłuższe działanie
Pxover	Brak	Brak zauważalnych zależności
Pmutation	Brak	Podobnie jak Pxover
Krzyz_arytm	Niski, pomijalny	Włączenie tego mechanizmu nieznacznie zwolniło działanie – zmiana na tyle mała, że może być skutkiem losowym
Skal	Niski, pomijalny	Włączenie nieznacznie wydłużyło działanie – naturalnie dodanie etapu przetwarzania spowalnia AE, jednak różnica w czasie była b. mała
Elit	Niski, pomijalny	Włączenie nieznacznie wydłużyło działanie Podobnie jak przy skalowaniu, jednak wpływ jeszcze mniejszy

3 Dodatkowe badanie PXOVER i PMUTATION

Badanie polegało na jednoczesnej manipulacji obu tych parametrów w celu dokładniejszego zbadania ich wpływu na wynik oraz siebie nawzajem.

Wpływ kombinacji PMUTATION i PXOVER na wynik y



Wykres przedstawia wyniki y (oś y) w zależności od PMUTATION (oś x) oraz PXOVER (kolor) dla różnych funkcji o 1, 5 lub 10 wymiarach. W badaniu pozostałe parametry były ustawione następująco:
l.iteracji 20 000, krzyż_arytm 0, skalowanie 1, elitaryzm 1, popsize 50.

Dla funkcji 1 i 5 wymiarowych zauważalne są negatywne skutki małej próby badawczej. Mimo różnych dokładnych wartości wyniku y , ogólne charakterystyki są takie same dla danych ziaren niezależnie od funkcji.

Na podstawie funkcji jednowymiarowych można zaobserwować, że niskie p . krzyżowania w połączeniu z niskim p . mutacji nie prowadzi do dobrych wyników – prawdopodobnie z powodu zbyt słabej zbieżności. Po prostu nie ma mechanizmu wprowadzającego znaczące poprawy wraz z iteracjami. Po zwiększeniu p . krzyżowania, bądź p . mutacji wyniki ulegały poprawie (pomijając zauważalne losowe wpadki spowodowane małą l. eksperymentów).

Dla funkcji 5-wymiarowych zależność ulega zmianie. Wynik w zależności od PXOVER jest bardziej losowy – dla f. Ackley’a najgorsze jest ustawienie $p_x = 0.5$, ale dla f. Rastragina jest ono jednym z najlepszych. Zauważalne jest to, że w większości przypadków najlepszym ustawieniem było bardzo duże p. krzyżowania równe 0.9. W przypadku bardziej zbalansowanych ustawień 0.5 oraz 0.7 raz jedno, a raz drugie było w stanie dorównać wyższemu 0.9, ale były one bardziej zależne od p. mutacji z zasadą, że zwiększanie mutacji poprawiało wyniki.

Dla funkcji 10-wymiarowych najłatwiej zauważyć wpływ PMUTATION, ponieważ jej zwiększanie po prostu pogarszało wyniki niezależnie od p. krzyżowania. Jak chodzi o p. krzyżowania najlepsze w większości scenariuszy było niskie 0.3 oznaczone kolorem niebieskim. Preferencja niskich wartości obu prawdopodobieństw, czyli spowolnienie zmian między kolejnymi populacjami prawdopodobnie wynika z wysokiego skomplikowania funkcji. Z wykresu można wywnioskować, że im mniej mutacji, tym lepiej jednak nie ma pewności (a teorie wręcz informuje, że na pewno tak nie jest) i by się upewnić należało by powtórzyć przesiew automatyczny dla innego zakresu danych.

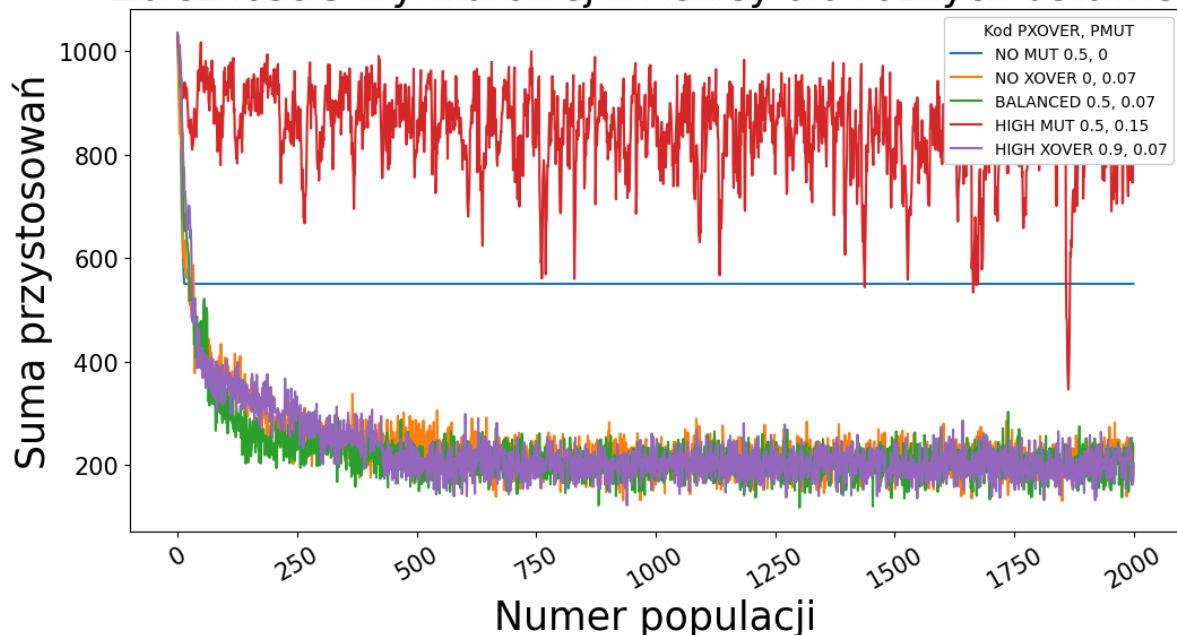
Podsumowując, ciężko jednoznacznie określić wpływ tych dwóch parametrów na działanie algorytmu, ponieważ jest silnie zależny od trudności problemu oraz prawdopodobnie również innych parametrów tj. popsize. W pewien sposób również liczba iteracji jest problemem, ponieważ mimo że pozwala na osiągnięcie lepszych wyników to częściowo ukrywa wpływ pozostałych ustawień na zbieżność populacji.

Ogólnie zauważalne jest, że p. krzyżowania określające element eksploatacji najbardziej konsekwentnie osiąga dobre wyniki dla dużych wartości, ale optymalna wartość dla pojedynczego problemu może być zupełnie inna – stąd wniosek, że duże PXOVER jest dobrym punktem startowym. P. mutacji silnie zależy od złożoności problemów – dla łatwych jest mile widziane, natomiast dla trudniejszych lepiej jest je ograniczyć (ale nie wyłączać).

Co do samego badania, obszerność zadania i wymóg analizy trzech funkcji sprawił, że dokładniejsze zbadanie większej liczby kombinacji było by zbyt czasochłonne. Wpływ małej liczby próbek był wyraźnie zauważalny szczególnie przy tych dwóch parametrach, co utrudniało skuteczną analizę ich nietrywialnych zależności.

4 Analiza zbieżności w zależności od PXOVER, PMUTATION

Zbieżność 5-wymiarowej f. Ackley dla różnych ustawień



Zbieżność populacji w wybranych zestawach ustawień p. mutacji i p. krzyżowania
Oś y jest sumą wartości przystosowania osobników w populacji.

Kolejnym etapem badania zależności p. krzyżowania i mutacji była analiza zbieżności w zależności od ustawień. Aby zapewnić lepszą widoczność wpływu parametrów wykorzystane zostały skrajne wartości. Pozostałe parametry miały wartość: popsize 20, maxgens 2000 (mniej iteracji to gorsze wyniki, ale większa czytelność), krzyz_arytm 0, elitaryzm 1, skalowanie 1.

Mutacja odpowiada za eksplorację – zauważalne jest, że w sytuacji jej braku populacja bardzo szybko zbiega do pewnego zestawu osobników, a następnie nie ulega poprawie. Jest to tożsame z wystąpieniem sytuacji, gdzie krzyżowanie osobników daje takie samo potomstwo. Widać więc, że mutacje zapobiegają takiemu utykaniu populacji. Z drugiej strony, zbyt duże p. mutacji w sytuacji oznaczonej czerwonym kolorem powoduje, że mimo ogólnego trendu dążącego do minimum, spadek jest bardzo wolny. Jest to efekt wielu skoków po różnych wartościach – zbyt duży element eksploracji. Pod kątem mutacji najlepsze jest ustawienie zbalansowane – w tym przykładzie wybrane na 0.07.

Brak krzyżowania nie powoduje zatrzymania działania algorytmu tak jak to miało miejsce przy wyłączeniu mutacji. Wynik wciąż zbiega się do całkiem dobrych wartości, jednak ma to miejsce wolniej niż przy aktywowanym krzyżowaniu. Dla bardzo dużego p.krzyżowania mimo tego, że osiągnany wynik jest najlepszy to po wykresie widać, że zbieżność jest dużo wolniejsza od wersji zbalansowanej (porównywalna z brakiem krzyżowania!), a fakt, że osiąga najlepszy wynik jest efektem czynników losowych.

Podsumowując, p. krzyżowania odpowiada za eksploatację i pozwala na przyspieszenie zbiegu populacji w okolice optimum. Mechanizm mutacji zapobiega utykaniu populacji w minimach lokalnych oraz pomaga szukać nowych rozwiązań zapewniając element eksploracji. Dokładne wartości zależą od konkretnego problemu, ale ogólną zasadą (dobrym startem) jest utrzymanie p. krzyżowania w wartości kilku dziesiątych, a p. mutacji poniżej 0.1.

W badaniu wyniki zbiegają się bardzo szybko do pewnego progu a następnie poprawy są bardzo małe. To zjawisko jest zauważalne już po 1000 iteracji, a biorąc pod uwagę że w poprzednich badaniach było

ich aż 20 000 tłumaczy to obecność zaskakujących obserwacji i dużego wpływu losowości na końcowe y .

5 Optymalizacja i porównanie

Ostatnim etapem było wykorzystanie wniosków i próba znalezienia najlepszych ustawień dla każdego zadania. Ustawienia nie są maksymalnie zoptymalizowane, ponieważ dalsze badania po uzyskaniu tak dobrych wyników jak np. $4e-3$ po prostu byłyby nieopłacalne – rozwiązanie już jest świetne.

Funkcja	Wymiary	Wynik	popsze	maxiter	pxover	pmutation	Krzyz_arytm	skal	Elit	seed
Ackley	1	7.255416e-009	50	20000	0.5	0.07	1	1	1	100
Ackley	5	4.434522e-006	30	20000	0.1	0.07	1	1	1	100
Ackley	10	4.053382e-003	50	20000	0.5	0.05	0	0	1	100
DeJong	1	6.345581e-015	50	20000	0.5	0.07	1	1	1	100
DeJong	5	4.215115e-007	50	20000	0.5	0.07	1	1	1	100
DeJong	10	1.184261e-005	50	20000	0.5	0.07	1	1	1	10
Rastragin	1	1.776357e-015	40	20000	0.5	0.05	1	1	1	10
Rastragin	5	6.766072e-010	20	20000	0.4	0.05	1	1	1	10
Rastargin	10	4.843852e-005	50	20000	0.5	0.07	0	0	1	0

Uzyskane wyniki końcowe są prawie idealne. Dla każdej funkcji dowolnej wymiarowości udało się uzyskać wynik, który jest ekstremalnie bliski optimum. W niektórych przypadkach tylko niedoskonałości reprezentacji nie pozwoliły na osiągnięcie równego 0. Ponadto, należy zaznaczyć, że we wszystkich przypadkach praca algorytmu nie przekraczała czasem trwania 2-3s.

Biorąc pod uwagę podobieństwo parametrów do domyślnych w większości przypadków można powiedzieć, że zadanie jest łatwe dla tego algorytmu. Ponadto, szczegóły ustawień nie miały aż tak dużego znaczenia, ponieważ szybkość ewaluacji pozwalała na zastosowanie po prostu dużej liczby iteracji (a mogła by być jeszcze większa), co sprawiało że potencjalne dostosowywanie prędkości i kształtu zbiegu nie miało większego znaczenia, bo ostatecznie wszystko sprowadzało się do szczęścia przy losowych mutacjach kiedy populacja docierała już w okolice optimum.

Żeby dokładniej sprawdzić zrozumienie ustawiania AE należałoby zastosować go do trudniejszych zadań lub w celach edukacyjnych wymyślić odgórne ograniczenie liczby iteracji do liczby np. 1000, czy 2000 i starać się zminimalizować popełniany wtedy błąd ignorując fakt, że y był by daleki optimum.

Podsumowując, AE jest łatwy w użyciu, posiada zrozumiałe i łatwo edytowalne parametry, a nawet w domyślnej wersji jest o wiele skuteczniejszy od SA. Jedyną potencjalną wadą jest większy koszt obliczeniowy, który w tym prostym badaniu nie był realnym problemem, ale wydłużenie czasu wykonania było zauważalne.