Dokumentacja zadania laboratoryjnego - badanie algorytmu gradientu prostego

Krzysztof Wyrzykowski, nr indeksu 331455

23 października 2024

0.1 Opis działania algorytmu

Algorytm gradientu prostego pozwala na wyznaczenie przybliżonych współrzędnych minimum lokalnego zadanej funkcji f. Obliczenia rozpoczynamy w punkcie początkowym x_k. Następnie wyliczamy współrzędne nowego punktu x_{k+1} korzystając ze wzoru:

$$x_{k+1} = x_k - \alpha \Delta f(x_k)$$

gdzie α jest parametrem kroku, a $\Delta f(x_k)$ jest gradientem badanej funkcji w punkcie x_k. Konieczne jest dobranie odpowiedniej wartości parametru α, aby algorytm był w stanie znaleźć rozwiązanie, a w szczególności znajdował rozwiązanie w możliwie małej liczbie kroków. Obliczanie współrzędnych punktu x_1 kontynujemy, aż do spełnienia minimum jednego warunku stopu. W tej implemntacji algorytmu gradientu prostego zostały użyte trzy warunki stopu, dla zadanej precyzji ϵ:

1) Norma odległości między kolejnymi punktami:

$$||x_{k+1} - x_k|| < \epsilon$$

2) Norma gradientu:

$$||\Delta f(x_k)|| < \epsilon$$

3) Maksymalna liczba iteracji:

$$k > K_{max}$$

0.2 Opis planowanych eksperymentów numerycznych

0.2.1 Cel przeprowadzenia eksperymentów

Celem eksperymentów było zbadanie zachowania algorytmu dla różnych wartości parametru α przy zadanej funkcji f i punkcie początkowym x_0 oraz ustalenie od czego zależna jest optymalna wartość parametru α.

0.2.2 Parametry eksperymentów

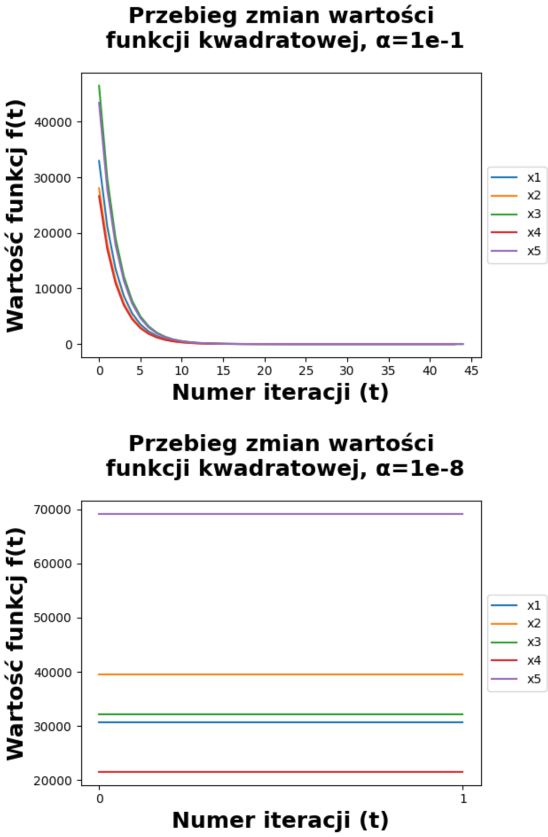
Do testów zostały wykorzystałem trzy funkcje - funkcję kwadratową oraz funkcje f3 i f12 z benchmarku CEC2017. Jako dziedziny funkcji przyjęty został zbiór $[-100, 100]^{10} \in R^{10}$. Wartość parametru precyzji ϵ została ustawiona na 10^{-5}. Przyjąłem maskymalną liczbę iteracji równą 1e5.

0.2.3 Przebieg badań

Dla każdej funkcji sprawdziłem działanie algorytmu dla parametrów α równych 1e-1, 1e-4 i 1e-8, powtarzając symulację dla 5 różnych losowo wybranych punktów należacych do dziedziny. Na tej podstawie przygotowałem wykresy zbieżności, które pokazują jaka była wartość badanej funkcji w punkcie x_k dla k-tej iteracji.

0.3 Wyniki badań

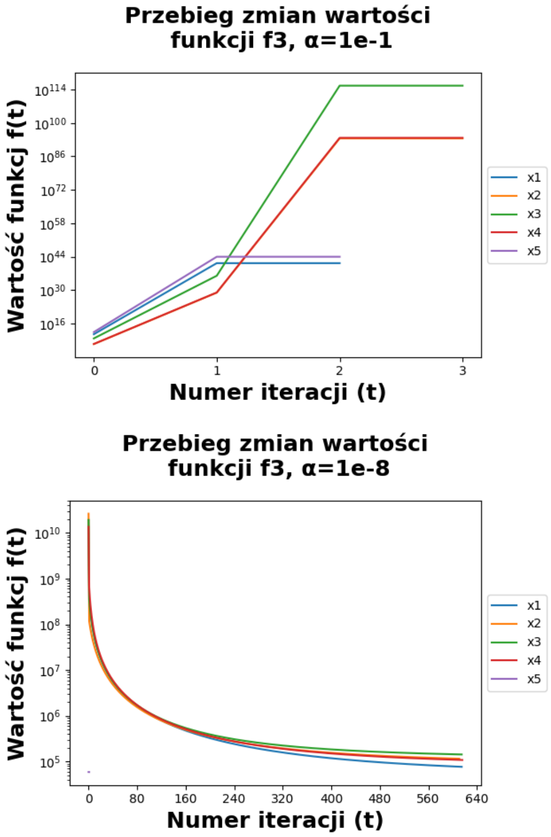
0.3.1 Badanie zachowania algorytmu dla funkcji kwadrtowej



Rysunek 1: Wykresy zbieżności dla funkcji kwadratowej

Z wykresów wynika, że funkcja jest zbieżna dla α równego 1e-1 i 1e-4. Dla wartości 1e-8 algorytm zatrzymuje się po wykonaniu jednej iteracji. Wynika to z tego, że przy tak małej wartości kroku odległość między kolejnymi punktami jest na tyle mała, że niespełniony jest warunek normy odległości między punktami. Widać również, że dla wartośći 1e-1 algorytm potrzebował w każdym przypadku nie więcej niż 45 iteracji, natomiast dla wartości 1e-4 liczba iteracji przekraczała 10000

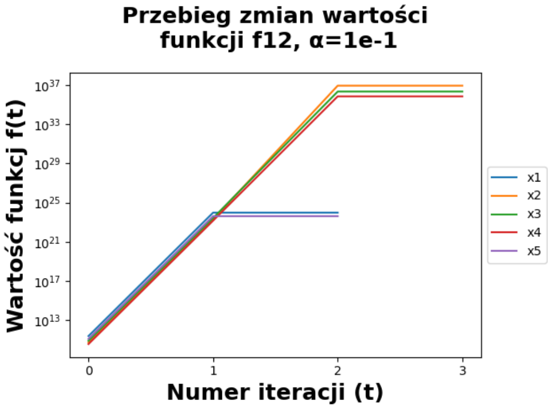
0.3.2 Badanie zachowania algorytmu dla funkcji f3

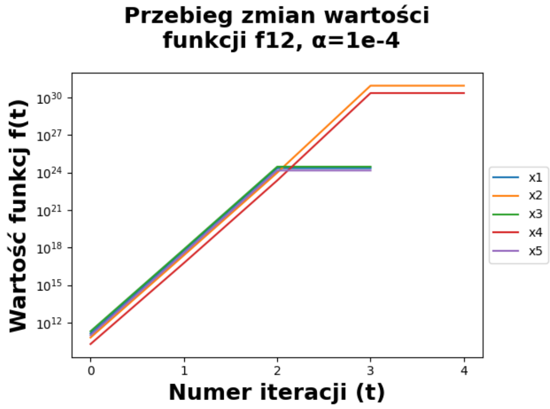


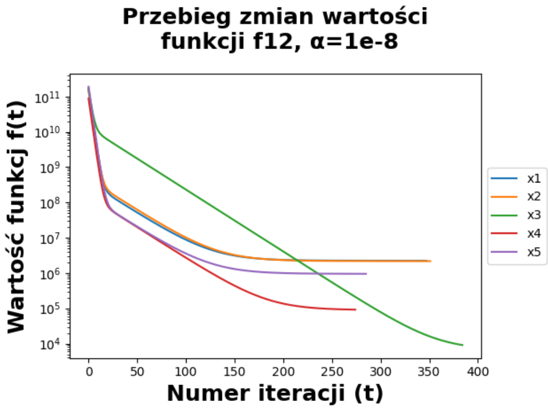
Rysunek 2: Wykresy zbieżności dla funkcji f3

Wykresy dla funkcji f3 pokazują, że w tym przypadku potrzebne są mneijsze wartości parametru rzędu 1e-8, bo tylko dla tego badanego parametru algorytm jest zbieżny. Dla pozostałych badancych parametrów wartość funkcji szybko rozbiego do $+\inf$, a punkt x_k juz w drugiej lub trzeciej iteracji znajduje sie poza dziedziną.

0.3.3 Badanie zachowania algorytmu dla funkcji f12







Zachowanie algorytmu dla funkcji f12 jest podobne do zachowania algorytmu dla funkcji f3, również wymaga ona do osiągenięcia zbieżności wartości parametru rzędu 1e-8

0.4 Wnioski

Z obserwacji wynika, że nie istnieje jedna uniwersalna wartość parametru α, która gwarantuje zbieżność algorytmu niezależnie od badanej funkcji. Z tego powodu nie możemy korzystać z tej metody w sytuacjach, gdy nie znamy charakterystki funkcji.

Na przykładzie funkcji kwadratowej widać, że nawet jeśli algorytm jest zbieżny dla róznych wartości α, to liczba potrzebnych kroków może się różnić o kilka rzędów wielkości. Zatem algorytm działa wydajnie dla bardzo wąskiego zakresu wartości parametru.

Jeśli algorytm gradientu prostego mógłby stać się metodą uniwersalną, ponieważ do każdej funkcji da się dobrać odpowiednią wartośc parametru. Natomiast konieczne jest wprowadzenie parametru który będzie zmienny w czasie i będzie się dostosowywał do funkcji celu.