Zadanie 1 - Sprawozdanie

Jan Stusio

Marzec 2024

1 Wstep

Metoda gradientu prostego (ang. Gradient descent) - iteracyjny algorytm służacy wyznaczeniu minimum lokalnego funkcji celu. Gradient prosty przyjmuje wartości wektorowe i określa, w którym kierunku funkcja, dla której został wyznaczony rośnie najszybciej.

Etapy algorytmu w pseudokodzie

1. $x \leftarrow x_0$.	Wybór punktu startowego x_0
2. while! $\|\| \bigtriangledown g(x) \|\| \leq \epsilon$	Sprawdzenie kryterium stopu
3. $d \leftarrow - \bigtriangledown g(x)$.	Obliczenie kierunku poszukiwań
4. $x \leftarrow x + \beta * d$	β - krok

Analizowane ta metoda zostana funkcje:
1. Rastrigina(https://www.sfu.ca/ ssurjano/rastr.html),
dla zakresu $x \in [-5,12; 5,12], x \in R^2$

$$f(x) = 10d + \sum_{i=1}^{d} \left(x_i^2 - 10\cos(2\pi x_i) \right)$$

2. Griewanka(https://www.sfu.ca/ ssurjano/griewank.html),
dla zakresu $x \in [-5,5], x \in R^2$

$$f(x) = \sum_{i=1}^{d} \frac{x_i^2}{4000} - \prod_{i=1}^{d} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

2 Implementacja

W implementacji gradientu prostego nie uwzgledniam kryterium stopu, ponieważ zakładam stała liczbe iteracji (równa 100). Zatem pseudokod dla zaimplementowanej funkcji wyglada nastepujaco:

1. $x \leftarrow x_0$.	Wybór punktu startowego x_0
2. for i in max	max - maksymalna liczba iteracji
3. $d \leftarrow - \bigtriangledown g(x)$.	Obliczenie kierunku poszukiwań
4. $x \leftarrow x + \beta * d$	β - krok

Rozważam tylko wymiar $d = 2$, zatem analizowane funkcje można uprościć:
Rastrigin

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

Griewank

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

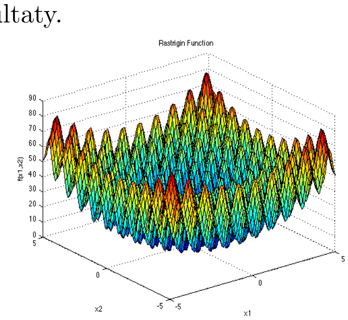
3 Badane parametry

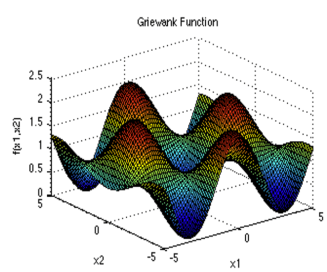
Jednymi z najistotniejszych parametrów, które przebadam sa β i x_0, który jest punktem inicjalizacji.

Parametr β przyjmuje wartości $0.66, 0.3, 0.1, 0.01, 0.001, 0.0001$

Parametr $x_0 = [a, 2]$ przyjmuje wartości $a = -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5$

Na podstawie wykresów obu analizowanych funkcji zakładam, że mniejszy krok β dla funkcji rastrigina w dziedzinie określonej w zadaniu bedzie dawał lepsze rezultaty.





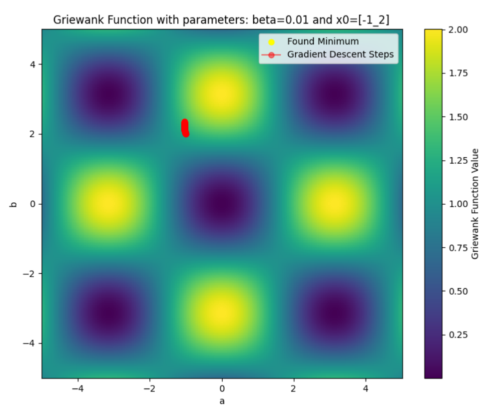
4 Testy

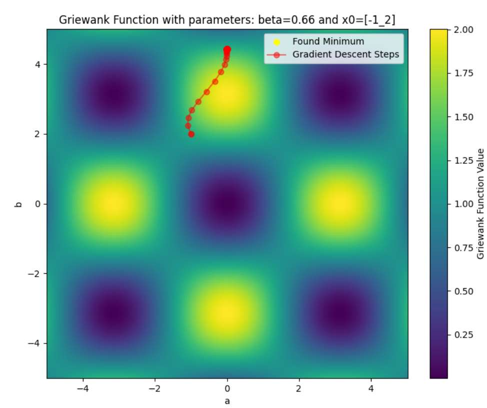
Testowana była poprawność implementacji poprzez plotowanie samych wyników w inny sposób. Oraz analizowałem wyniki w poszukiwaniu zgodności z założeniami przy różnych parametrach β oraz x_0.

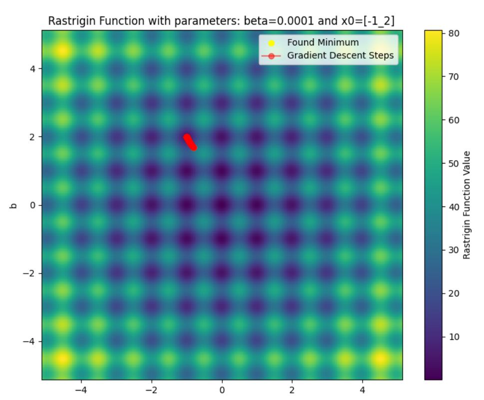
5 Wizualizacje parametrów

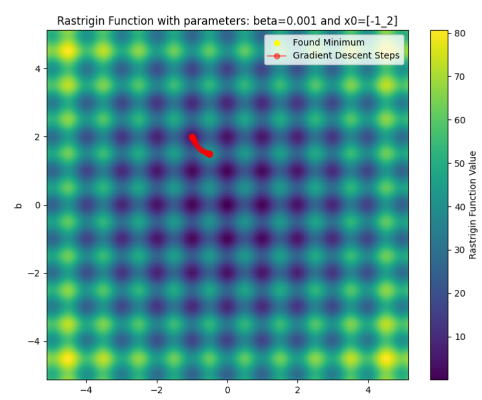
Wszystkie wyniki dla badanych parametrów znajduja sie w repozytorium Zadanie1/results. Ich nazwy zawieraja wartość parametru β oraz punktu inicjalizacji.

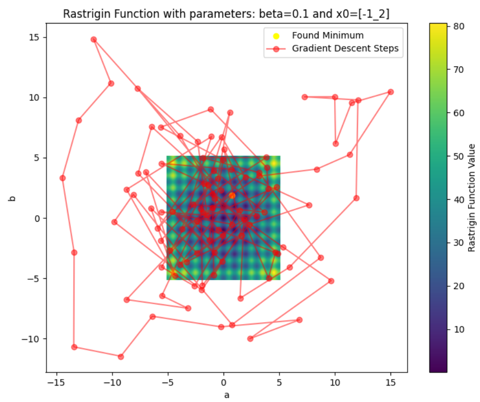
Poniżej zamieszczam tylko wybrane wizualizacje wyników, które dobrze ilustruja moje wnioski z eksperymentu opisane w nastepnej sekcji.

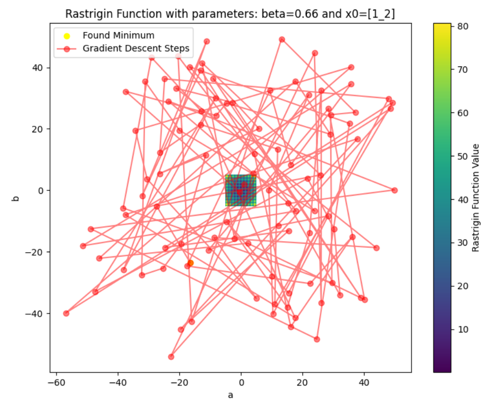












6 Wnioski

100 iteracji przy badanych parametrach zwykle nie pozwala odnaleźć żadnego minimum.
Szukane jest minimum lokalne tak jak zakładano.
Czasami gradient zmierza poza badana dziedzine funkcji (szczególnie dla wysokich wartości β)

Wśród badanych długości kroku dla funkcji Griewanka najlepsze rezultaty otrzymałem dla wartości 0,3 oraz 0,6, natomiast dla funkcji Rastrigina najdłuższy krok "skakał" i najlepsza długościa było 0,0001

Analizowane funkcje znaczaco różnia sie zbiorem wartości i okresami monotoniczności, stad różna jakość wyników dla podobnych parametrów. Dla Griewanka, który wolniej rośnie i ma mniejsza amplitude wartości w zbiorze wyników im wiekszy krok z puli badanych, tym lepszy rezultat. W przypadku funkcji Rastrigina, która szybciej zmienia wartości i ich amplituda jest wieksza, najlepsze rezultaty obserwowałem dla najmniejszych wartości z puli badanych.

Zbierzność funkcji da sie tylko zaobserwować na wykresach z dobrze dobranymi parametrami, które rzeczywiście daża lub osiagaja minimum lokalne. Przy dłuższym kroku widać, że punkty pośrednie oznaczajace kolejne iteracje metody gradientu prostego sie zageszczaja na wizualizacjach.