Algorytm gradientu prostego

Sprawozdanie, laboratorium numer 1 WSI

Rafał Kuśmierz 305858

Analiza dla funkcji celu:

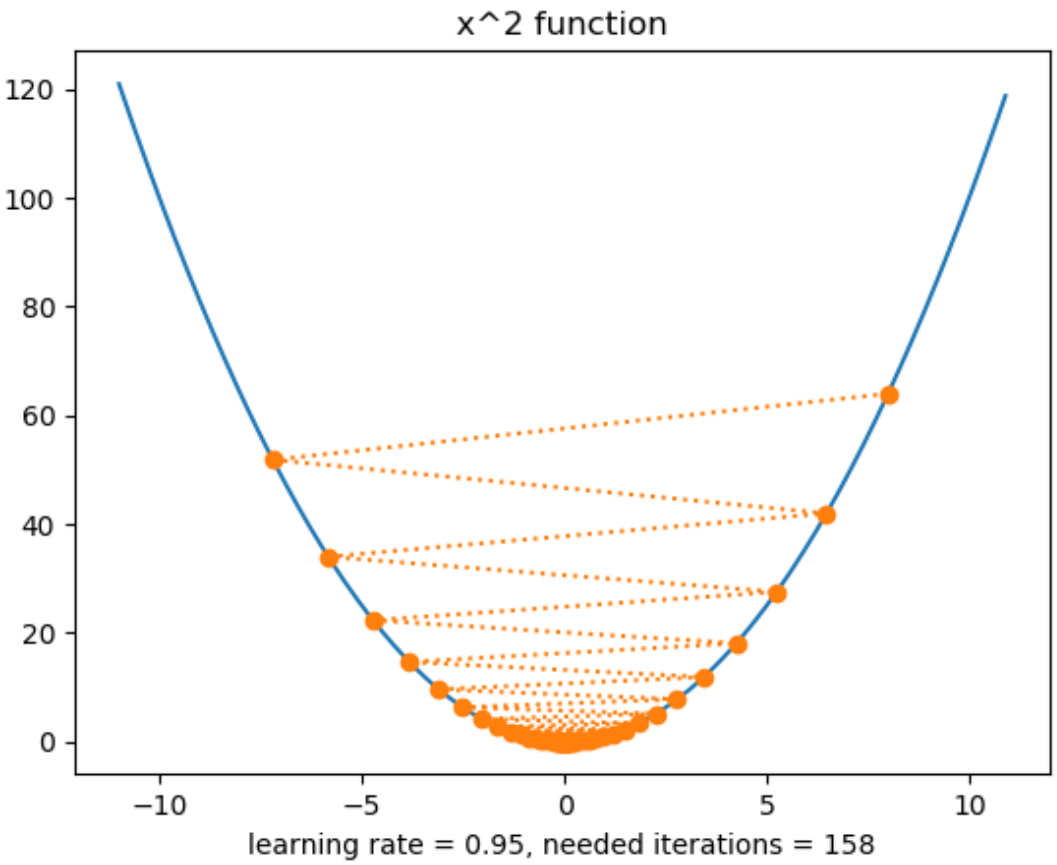
x^2

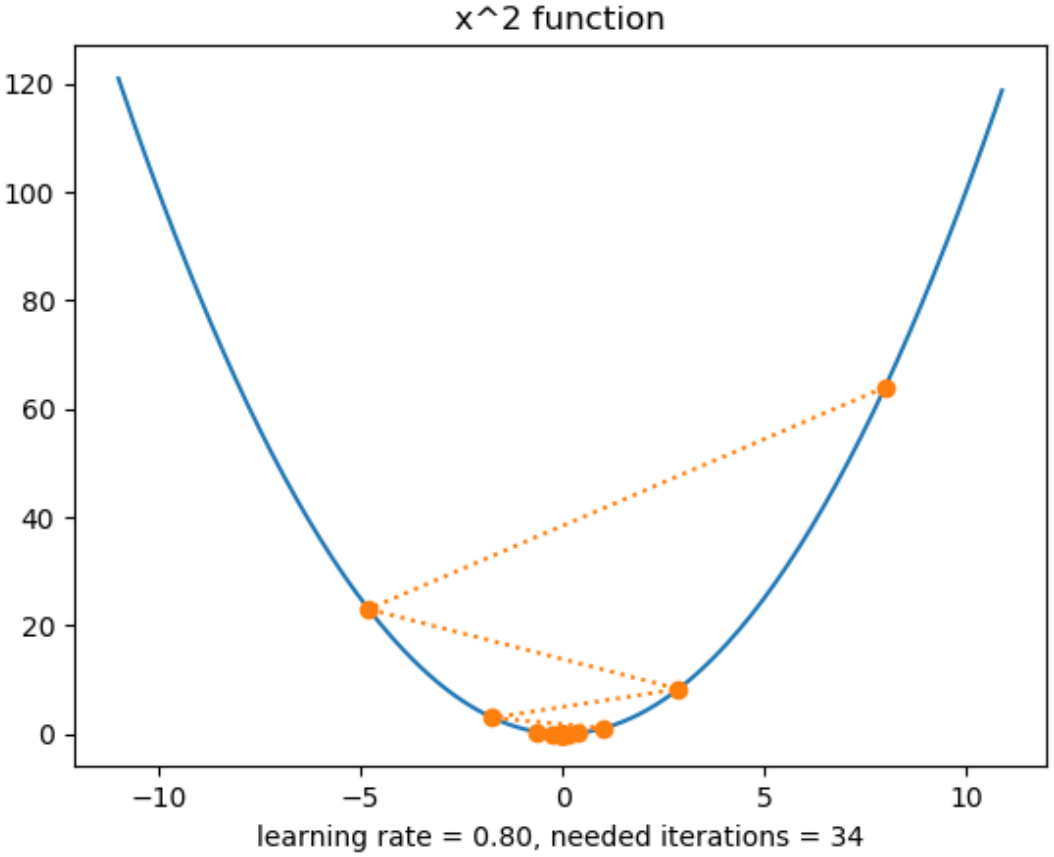
Badamy działanie algorytmu gradientu prostego. W przypadku problemu jednowymiarowego wydaje się on działać prosto i przejrzyście. Dla rozmiaru kroku o wartości większej niż 1 algorytm gubi się. Pierwsze obserwacje przeprowadziłem dla punktu startowego x = 8. Dla rozmiaru kroku = 0.95 algorytm potrzebował aż 158 iteracji, aby dotrzeć do oczekiwanego rezultatu. Można również zaobserwować pewne 'przestrzelanie' kolejnych punktów – droga do minimum wielokrotnie zawraca. Zmniejszenie wartości rozmiaru kroku do 0.45 zapobiegło zawracaniu, a ilość potrzebnych iteracji znacząco zmalała – do jedynie 8 powtórzeń. Kontynuacja zmniejszania rozmiaru kroku poskutkowała wzrostem potrzebnych iteracji. Dla rozmiaru kroku = 0.001 potrzebne było ich aż 4837.

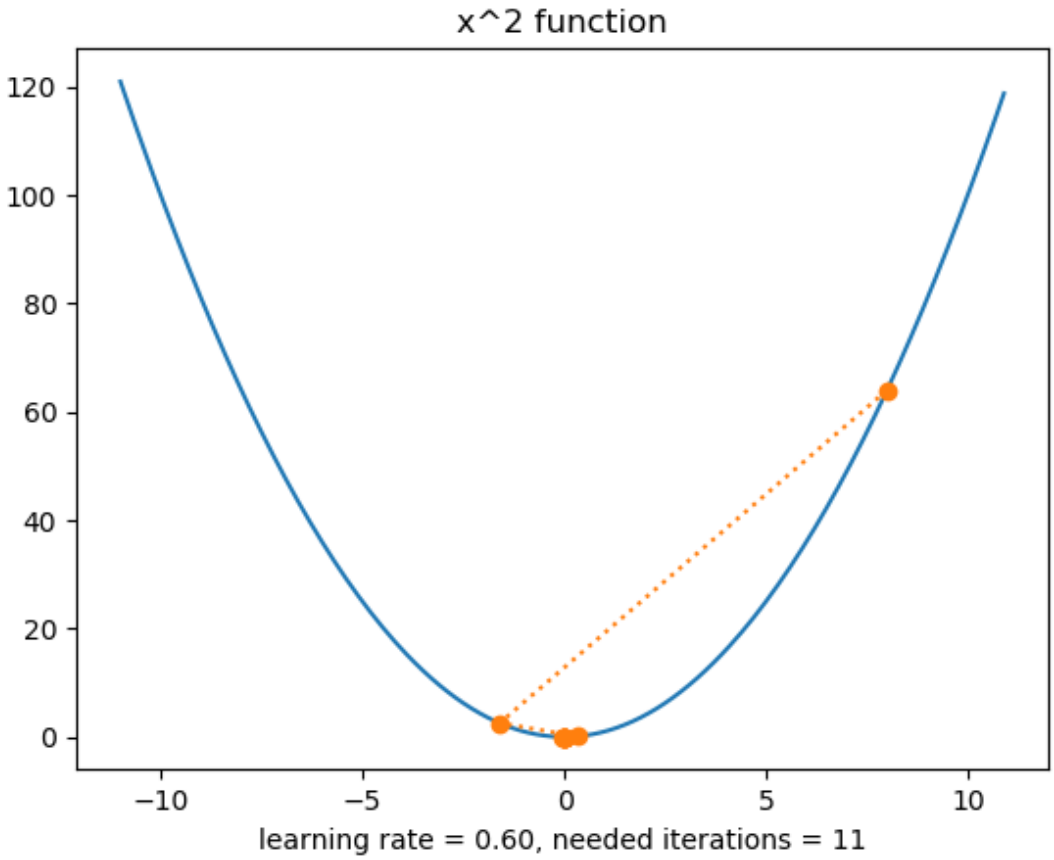
Powtórzyłem symulacje dla kolejnych różnych wartości punktu startowego. Umieściłem wykresy dla x= 40 oraz x = -18.5. Ich analiza pozwala potwierdzić przeprowadzone obserwacje.

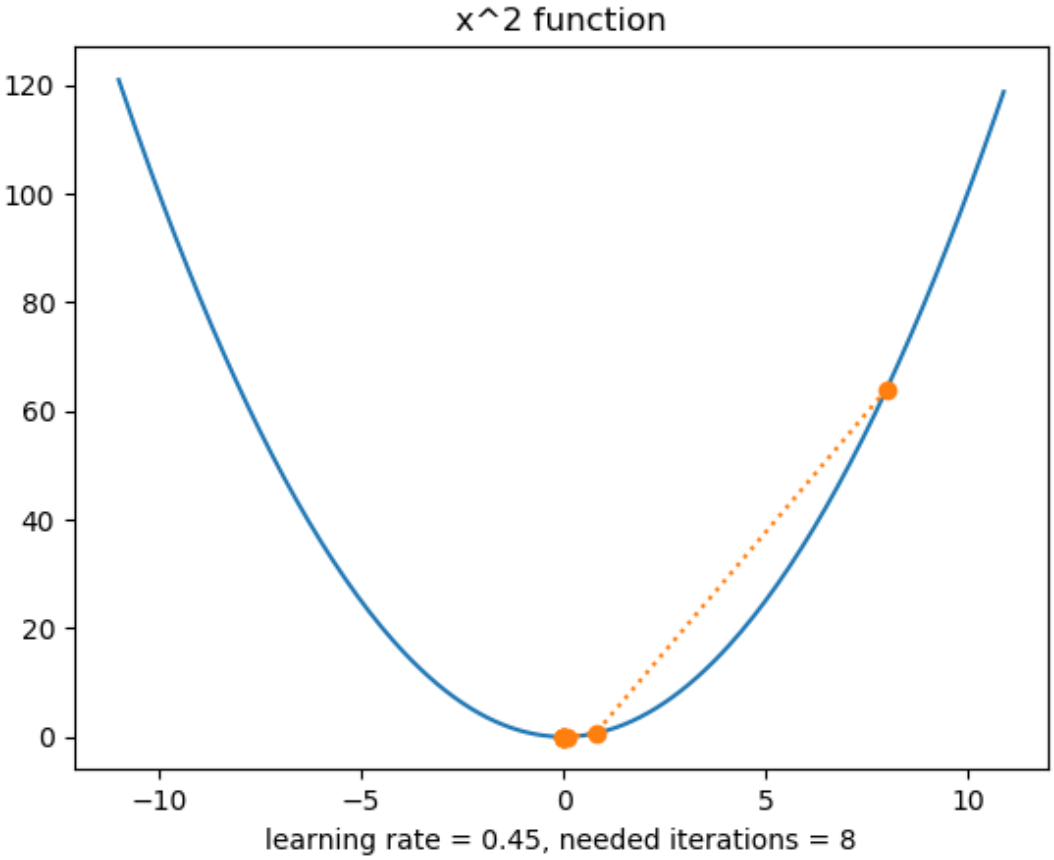
Dla każdego rozmiaru kroku poniżej 1 udawało się uzyskiwać oczekiwane rezultaty, jedynie przy bardzo małych jego wartościach ilość potrzebnych iteracji znacząco rosła (była to jednak akceptowalna liczba – wymagała kilku sekund odczekania aż algorytm poradzi sobie z zadaniem).

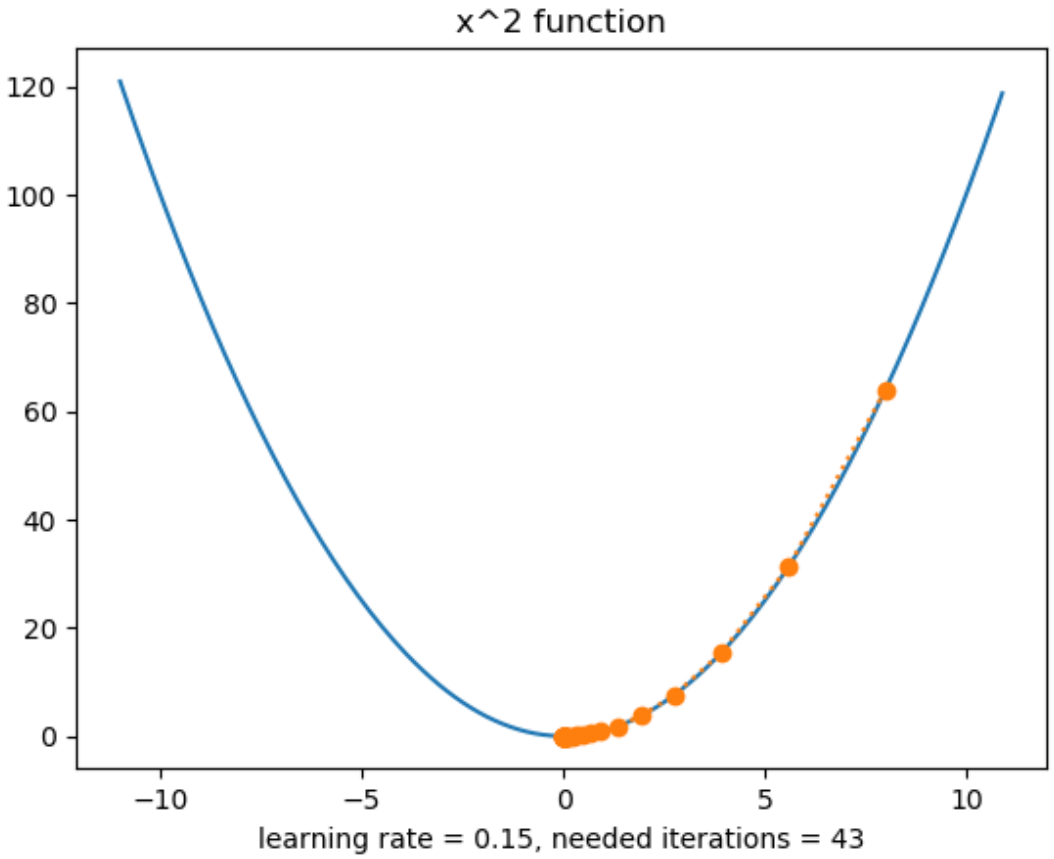
Dla punktu startowego x = 8

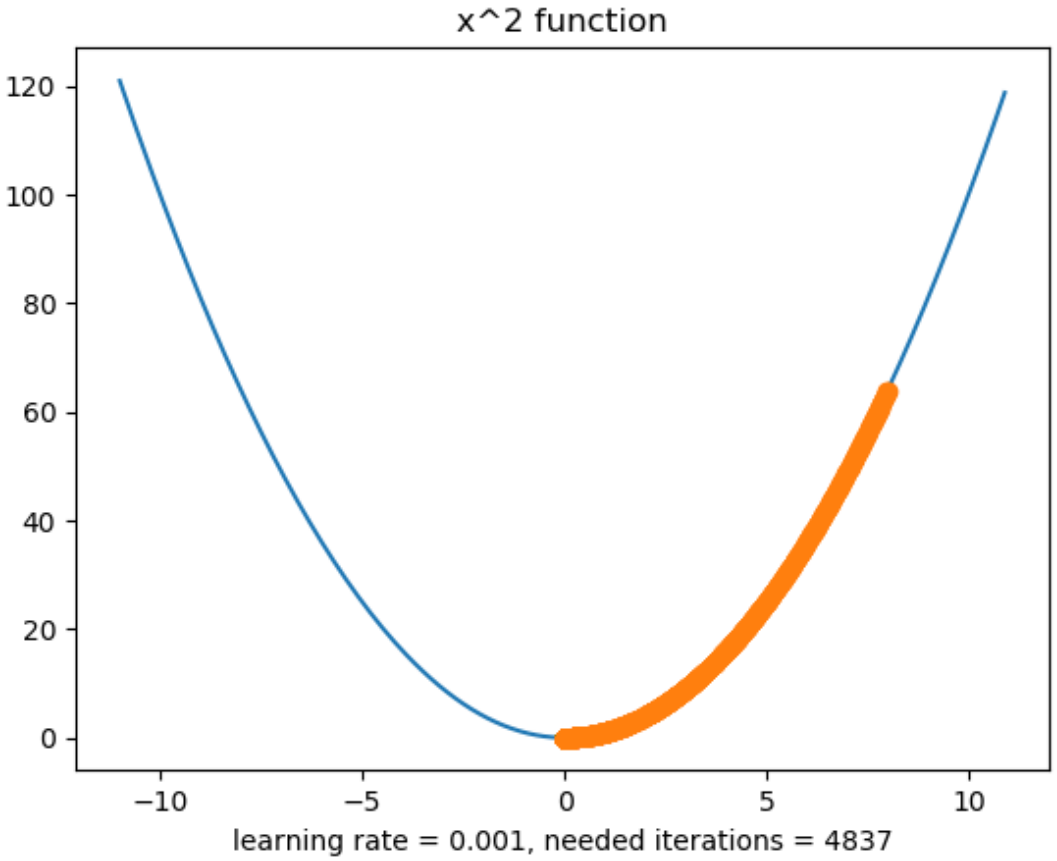




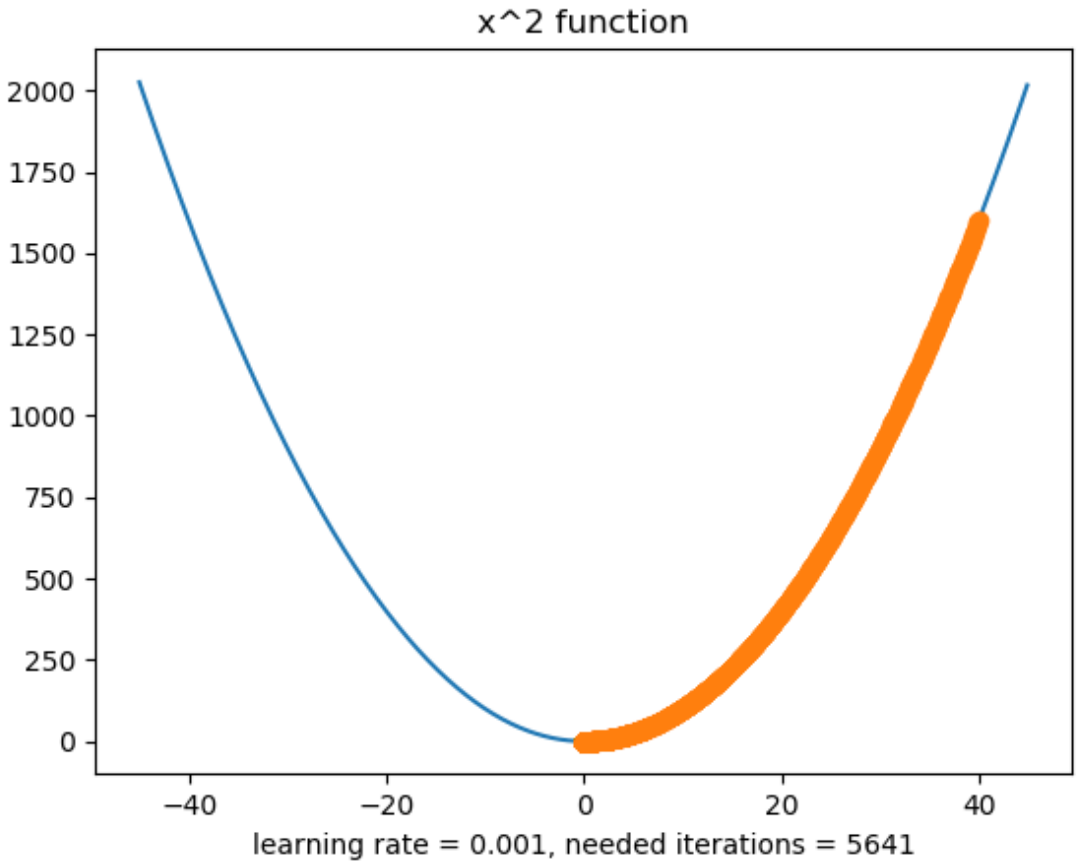

x^2 function
120
100
80
60
40
20
0
−10
−5
0
5
10
learning rate = 0.60, needed iterations = 11

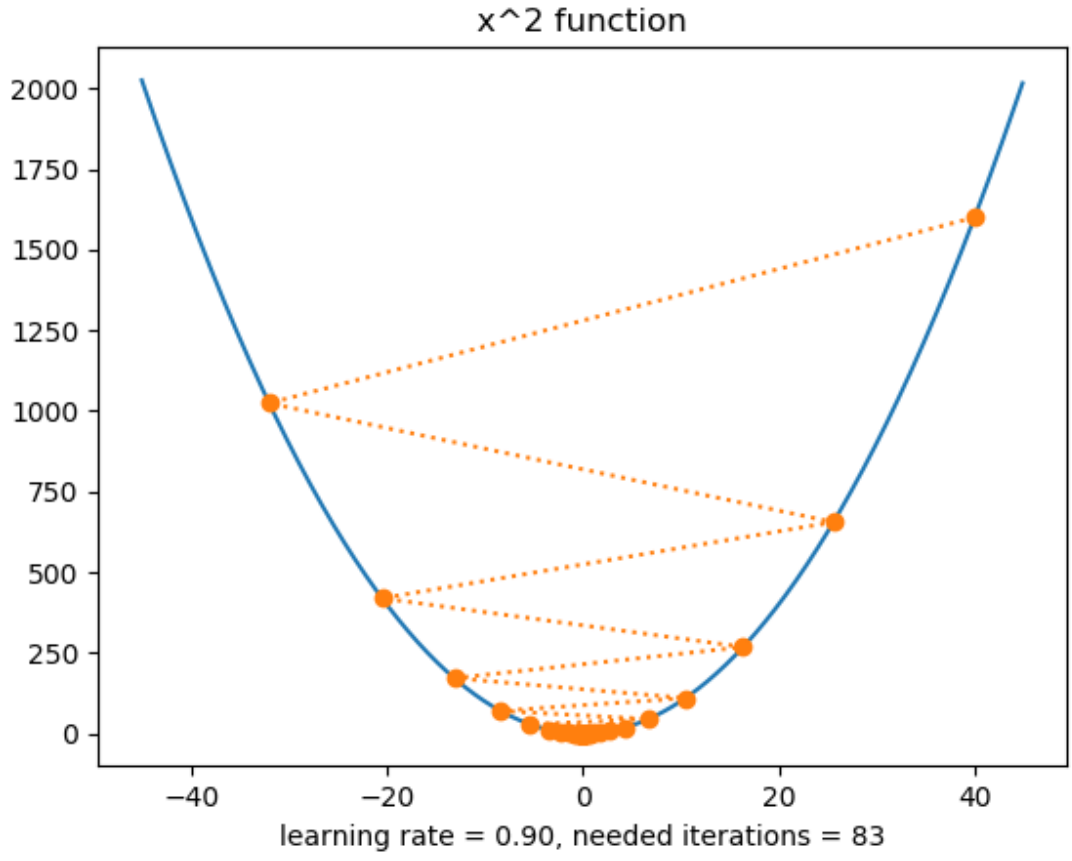

x^2 function
120
100
80
60
40
20
0
−10
−5
0
5
10
learning rate = 0.45, needed iterations = 8


x^2 function
120
100
80
60
40
20
0
−10
−5
0
5
10
learning rate = 0.15, needed iterations = 43

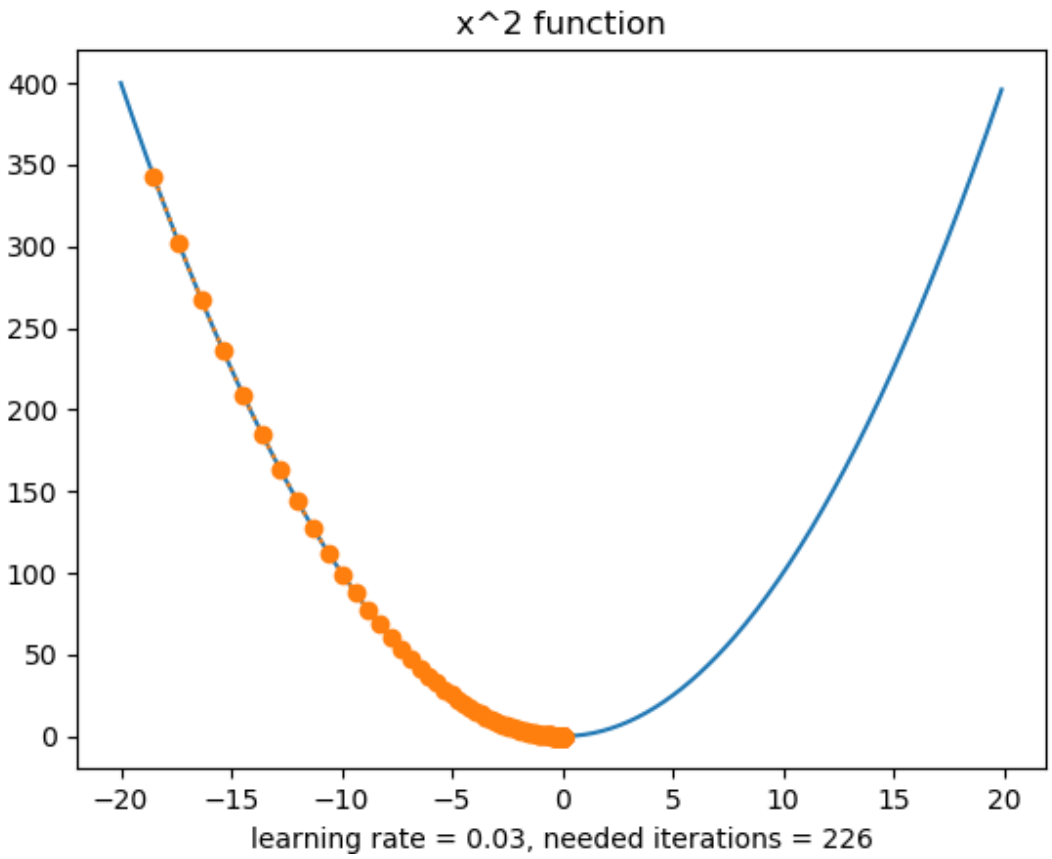

x^2 function
120
100
80
60
40
20
0
−10
−5
0
5
10
learning rate = 0.001, needed iterations = 4837

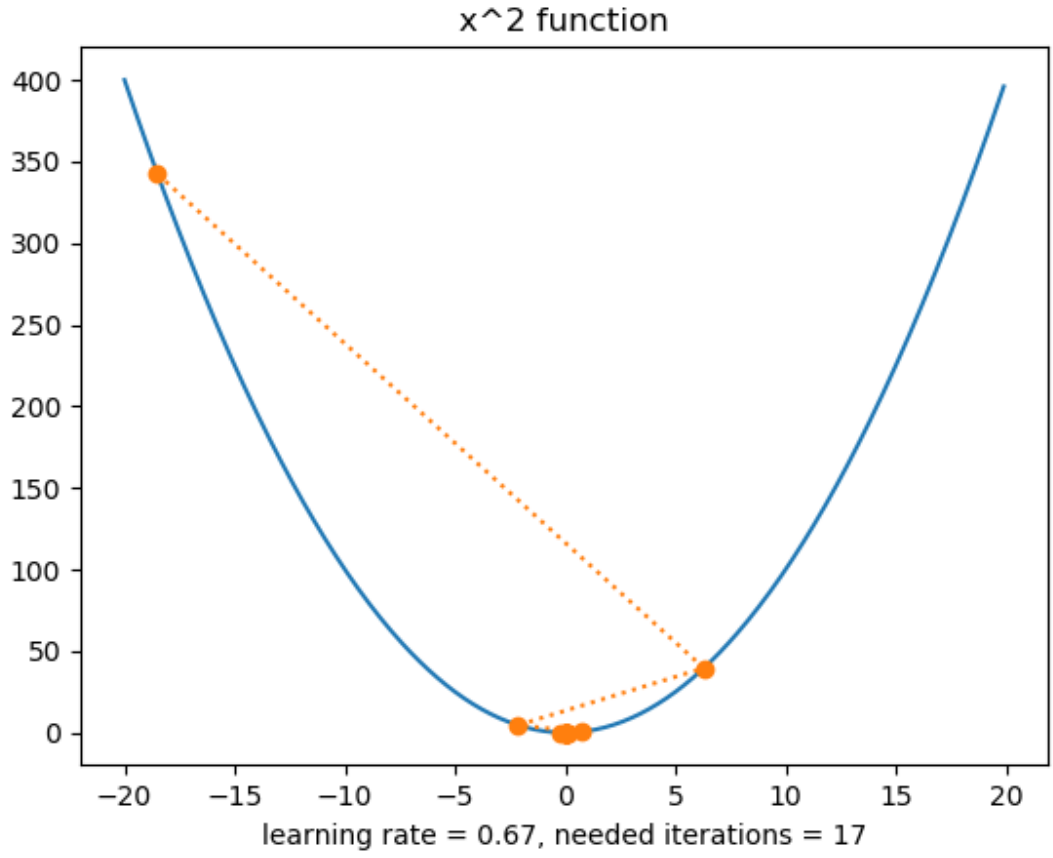
Dla punktu startowego x = 40





Dla punktu startowego x = -18.5





Analiza dla funkcji celu:

$$(x_1+a)^2+(x_2-a)^2-5\cos\left(10\sqrt{(x_1+a)^2+(x_2-a)^2}\right)$$ dla a = 8

Problem dwuwymiarowy na pierwszy rzut oka wydaje się być bardziej skomplikowany. Po wygenerowaniu odpowiednich wykresów można przejść do sensownej analizy działania algorytmu.

Patrząc na pierwsze wygenerowane wykresy, można zaobserwować wpływ cosinusa na kształt naszej funkcji, która wydaje się mieć ‘falistą’ naturę.

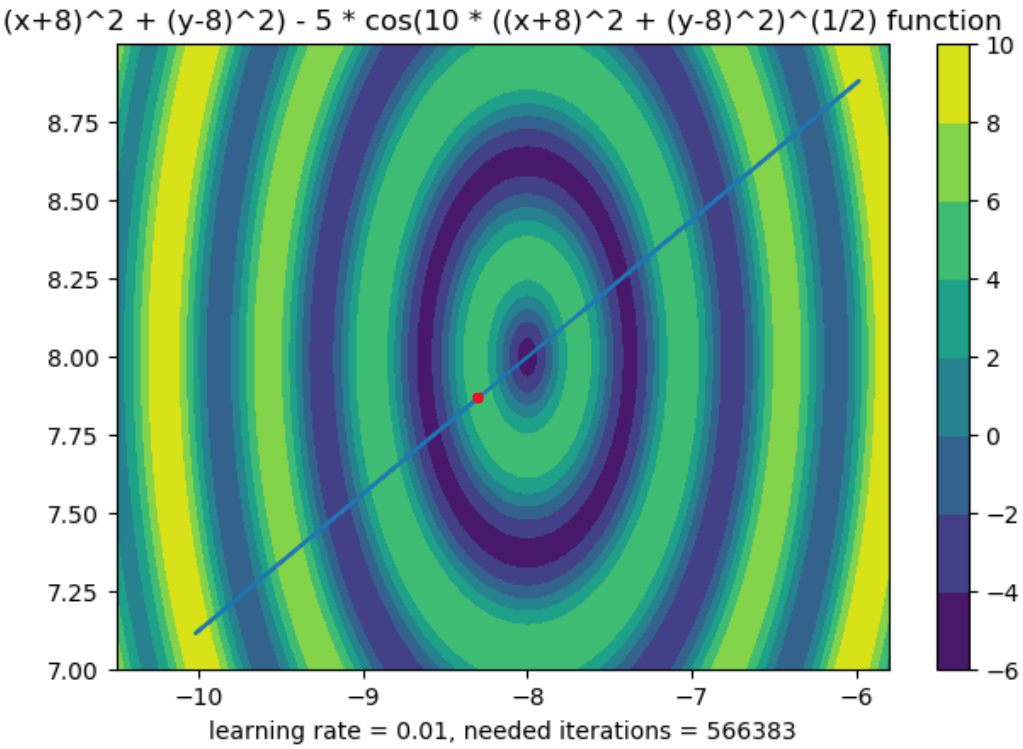
Dla wartości rozmiaru kroku większej od 0.01 ciężko ocenić otrzymywane rezultaty jako zadowalające – pomimo dużej liczby iteracji.

Dla rozmiaru kroku = 0.005 można zaobserwować, że rezultat to pewne minimum lokalne naszej funkcji – wydaje się być to najbliższe minimum lokalne od wektora startowego. Ilość potrzebnych iteracji do osiągnięcia i tak niezbyt dokładnego rezultatu wyniosła ustalone przeze mnie maksimum wynoszące 10 000 000. Zwiększenie liczby iteracji *10 nie wpłynęło na otrzymany rezultat – okazał się być identyczny przy przyjętym rozmiarze kroku.

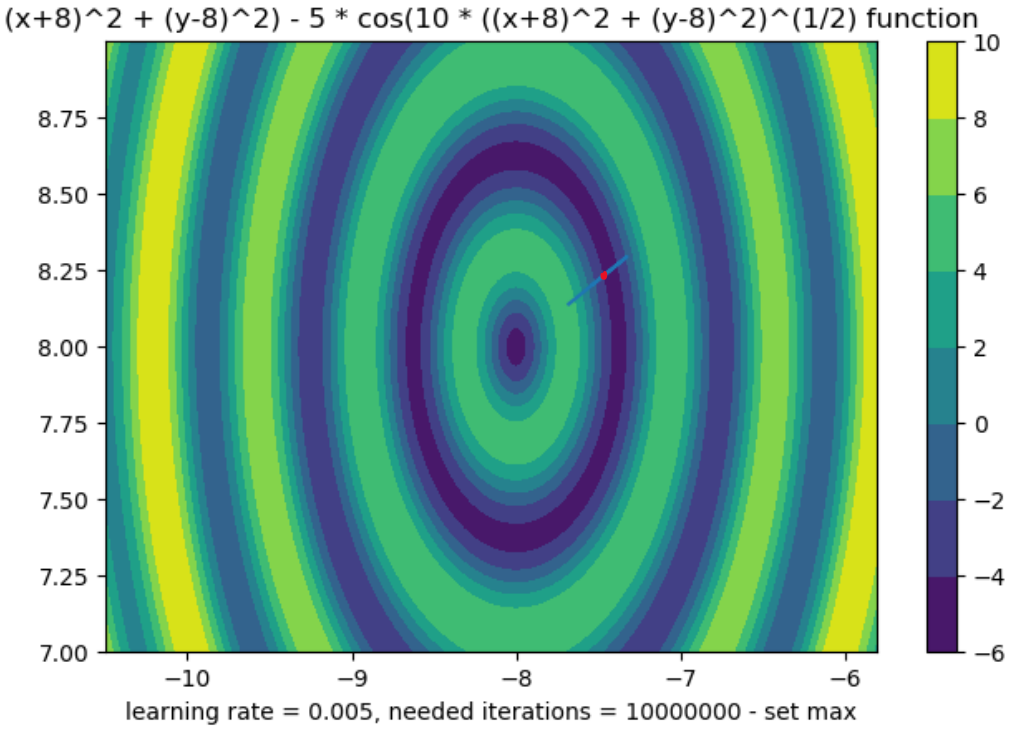
Dalsze zmniejszanie wartości rozmiaru kroku poskutkowało wzrostem dokładności rezultatu oraz zredukowaniem potrzebnej do tego liczby iteracji. Można wywnioskować, że najoptymalniejsze działanie algorytmu gwarantuje nam w tym przypadku rozmiar kroku = 0.002 – przejście od wektora startowego = [-7.68, 8.14] do rezultatu = [-7.42, 8.25] pochłonęło jedynie 7 iteracji.

Analiza kolejnych różnych wektorów startowych pozwala potwierdzić wysunięte wnioski. Algorytm gradientu prostego jako rezultat pozwala uzyskać minimum lokalne. Minimum globalne otrzymamy jedynie w sytuacji, gdy jest ono jednocześnie najbliższym minimum lokalnym danego wektora startowego – sytuacja przedstawiona na wykresach **6** oraz **7**.

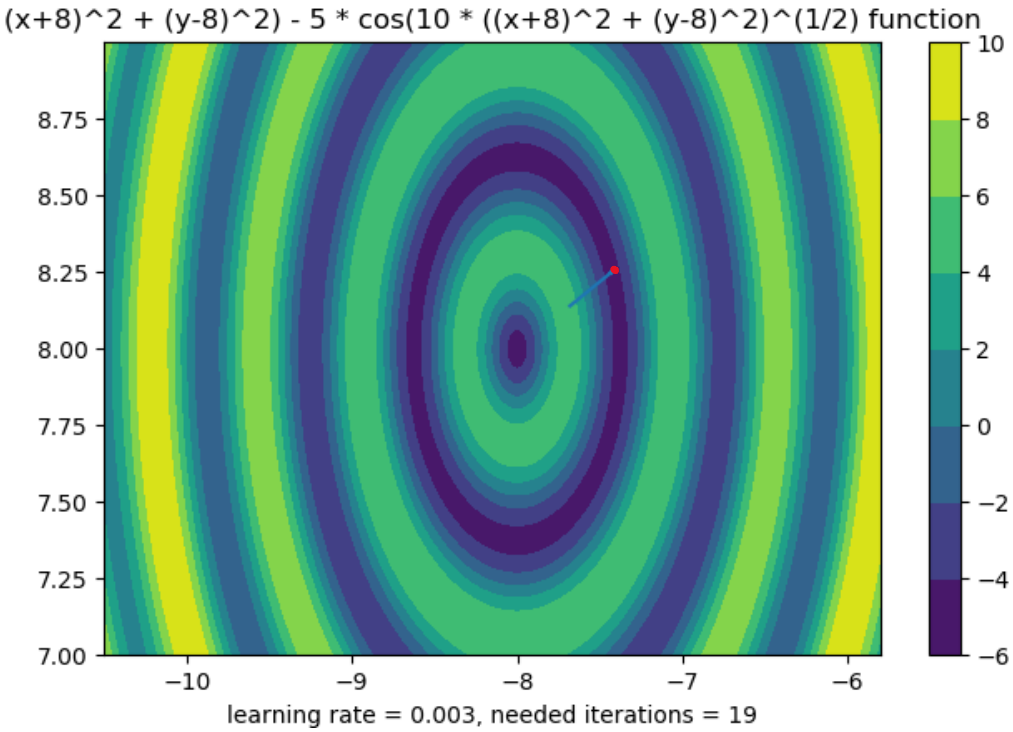
1.Start: [-7.68, 8.14] Rezultat: [-8.29, 7.87]•



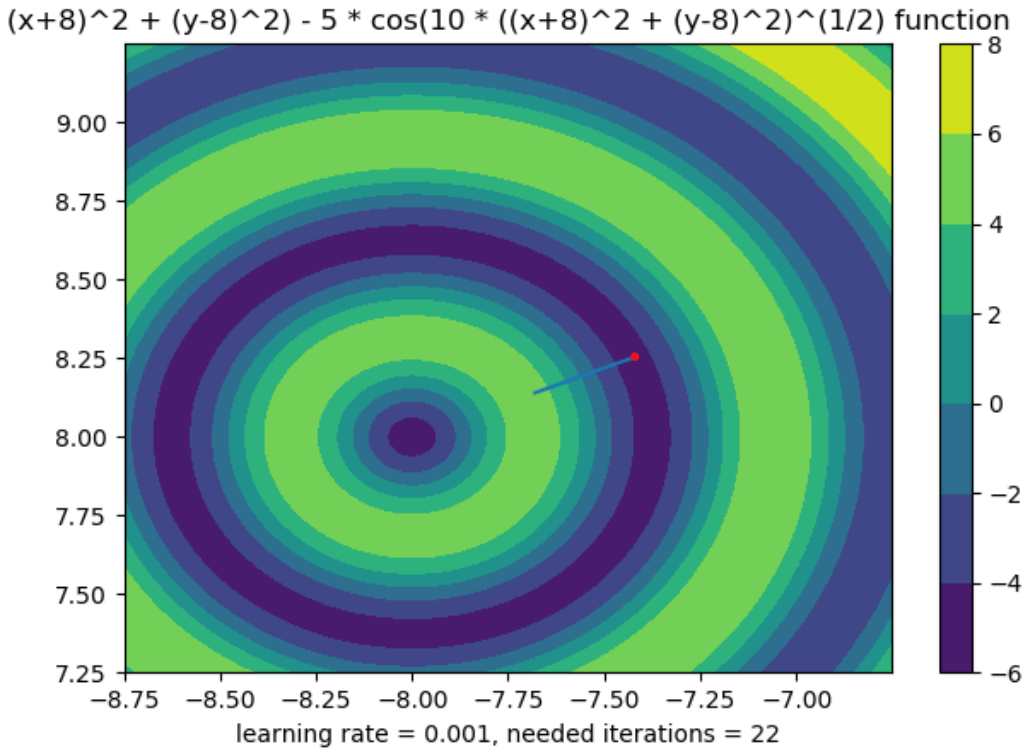
2.Start: [-7.68, 8.14] Rezultat: [-7.53, 8.20]•



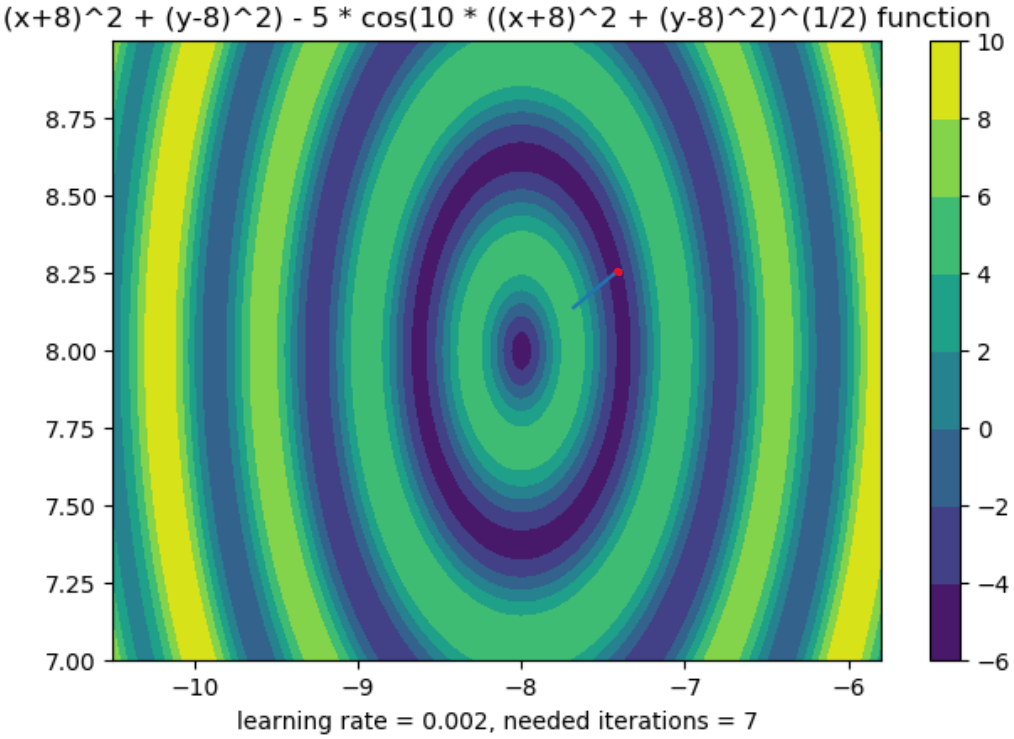
3.Start: [-7.68, 8.14] Rezultat: [-7.42, 8.25]



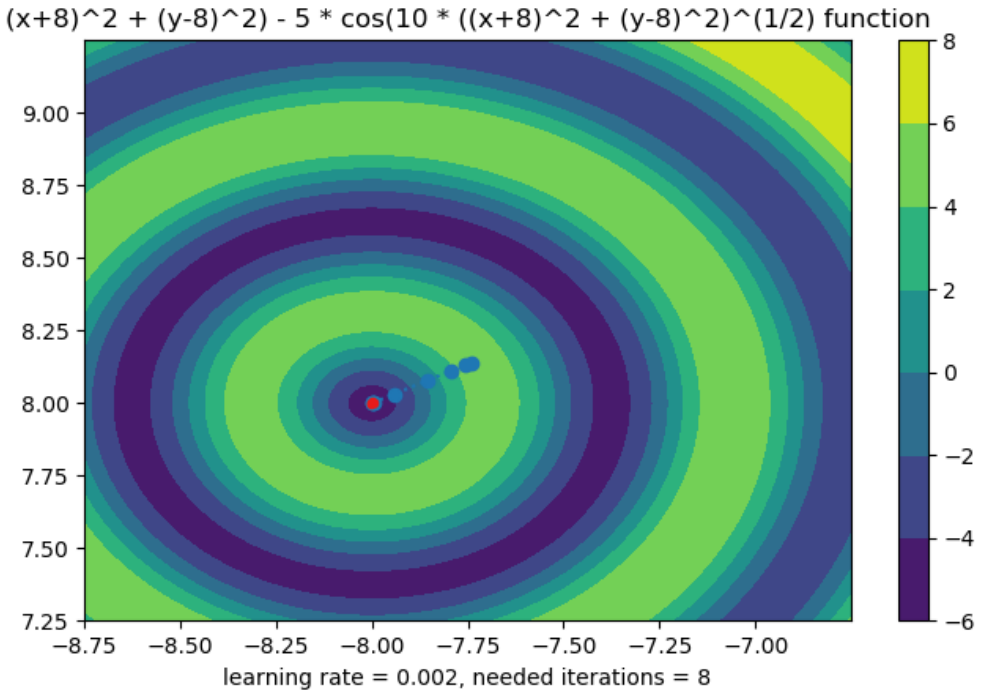
4.Start: [-7.68, 8.14] Rezultat: [-7.42, 8.25]



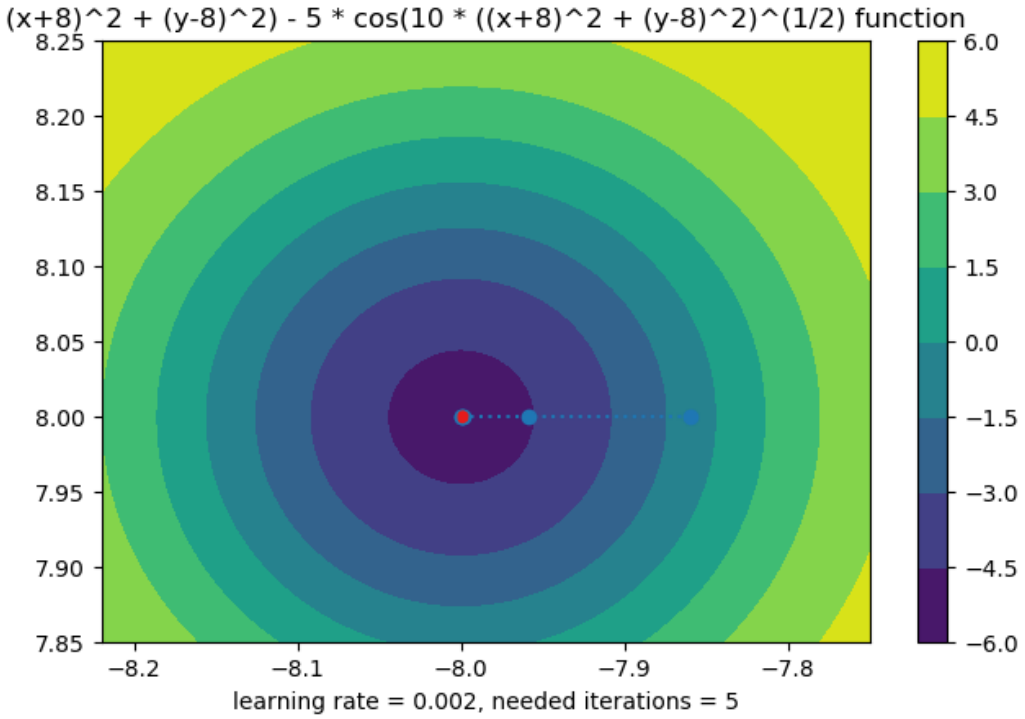
5.Start: [-7.68, 8.14] Rezultat: [-7.42, 8.25]•



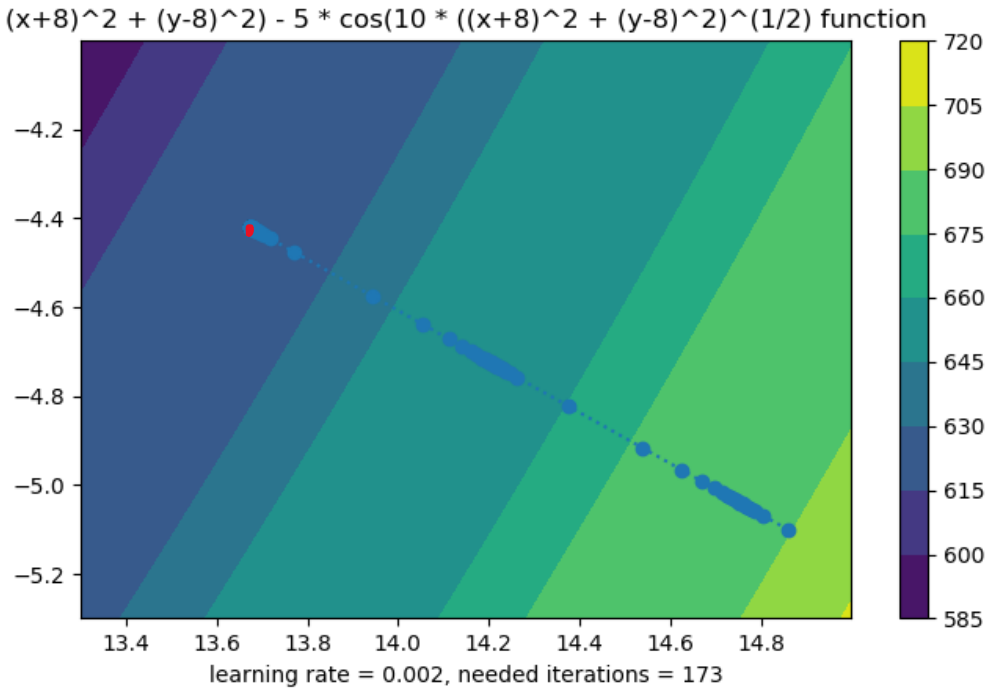
6.Start: [-7.74, 8.14] Rezultat: [-8.00, 8.00]•



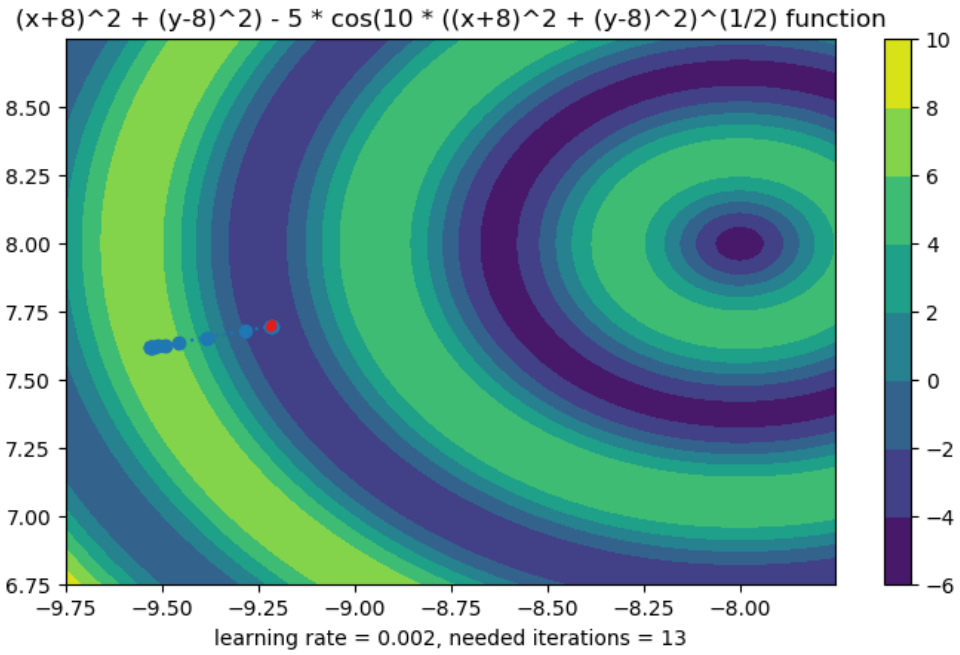
7.Start: [-7.86, 8.00] Rezultat: [-8.00, 8.00]



8.Start: [14.86, -5.10] Rezultat: [13.67, -4.42]



9.Start: [-9.53, 7.62] Rezultat: [-9.21, 7.70]



10. Start: [-9.31, 8.78] Rezultat: [-9.08, 8.64]