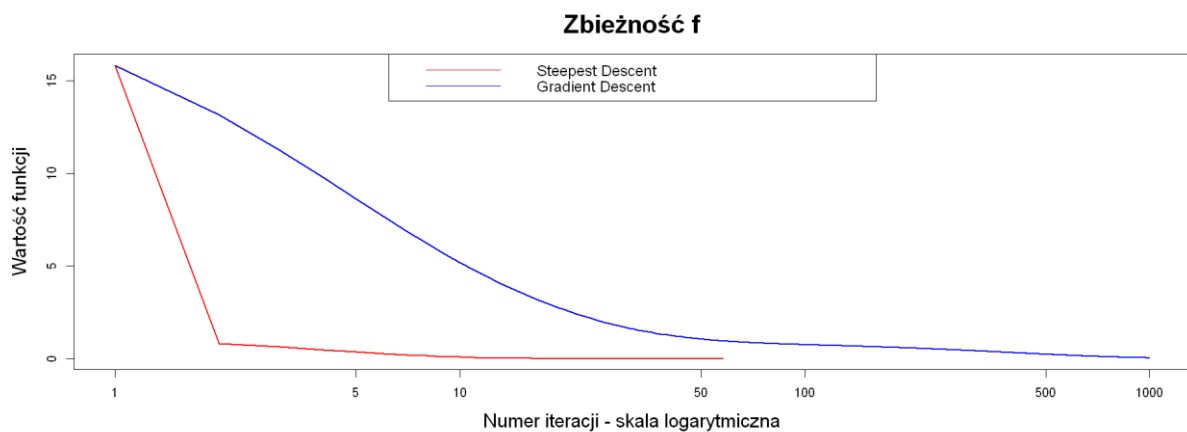


Proszę o wysłanie rozwiązania na lopins@sgh.waw.pl , podając numer albumu.

1. Opisz kluczowe Twoim zdaniem różnice pomiędzy metodami Gradient Descent a Steepest Descent.

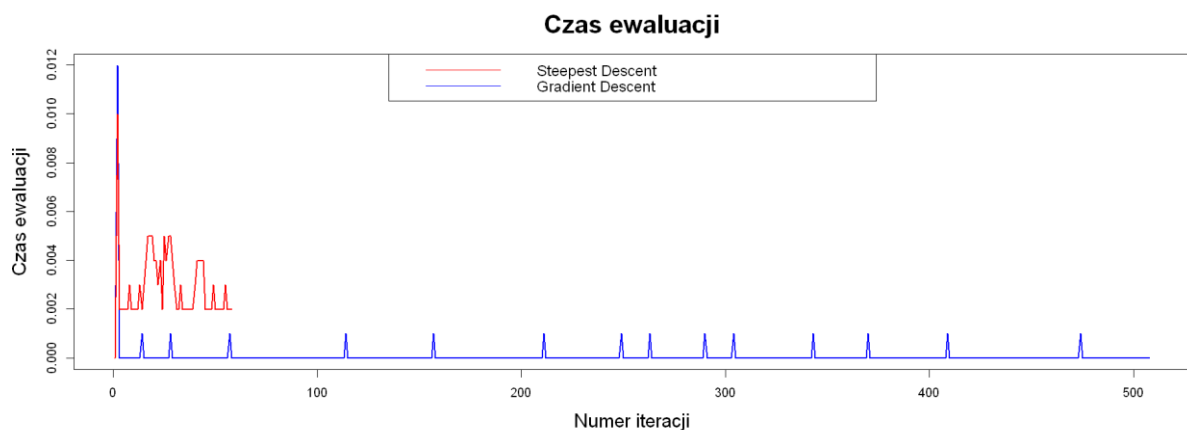
2. Czym jest tzw. learning rate/alpha w ww. metodach, i jakie mogą być konsekwencje zbyt niskiej lub zbyt wysokiej wartości tego parametru?

3. Czym wytłumaczysz różnicę w zbieżności metod Gradient Descent i Steepest Descent?
Przykładowy wykres poniżej.



4. Czym wytłumaczysz różnice w czasie wykonania?

Poniżej przykładowe dane.



Gradient Descent - czas ewaluacji: 0.02592325, optim x: $-1.517637e-09$, optim y: 0.0001560221

Steepest Descent - czas ewaluacji: 0.1685507, optim x: $-5.732683e-06$, optim y: $5.124782e-05$