WSI – ćwiczenie 1.
Algorytm gradientu prostego

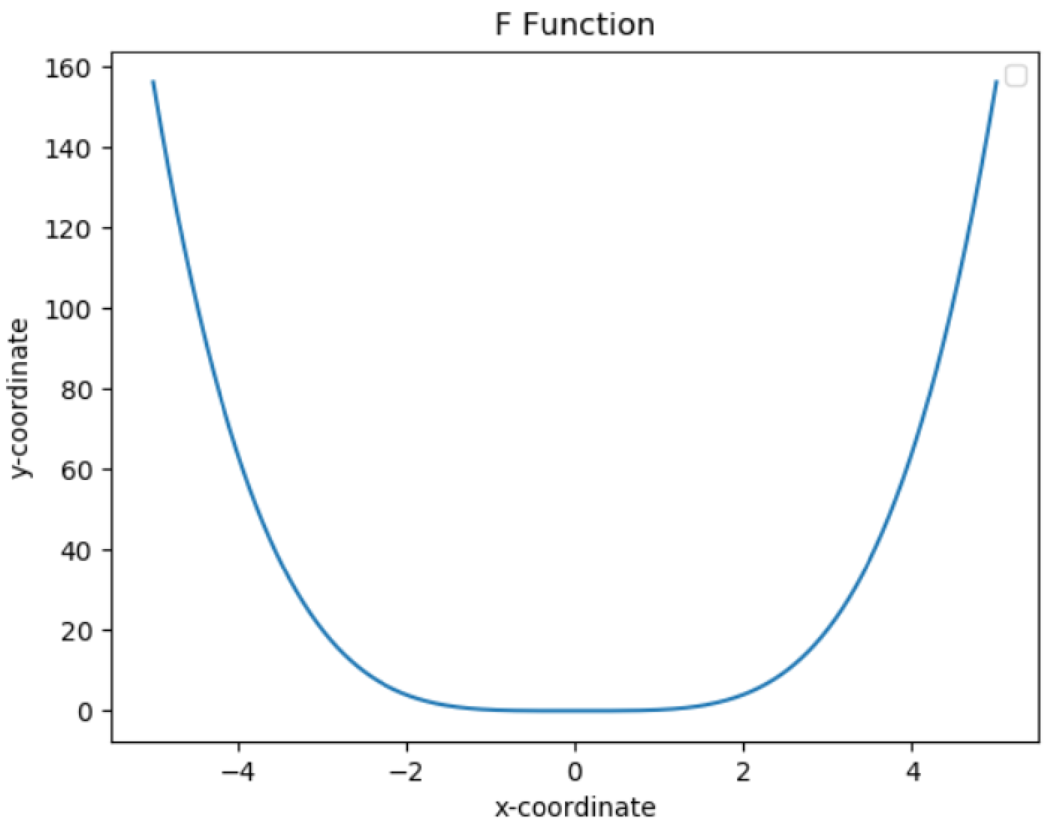
1. Przebieg ćwiczenia

Ćwiczenie będzie polegać na badaniu wpływu współczynnika β na szybkość odnajdywania minimum funkcji f i g. Do pomocy w obliczeniach została użyta biblioteka numpy. Do wyświetlania efektów pracy algorytmu biblioteka matplotlib.

2. Funkcja f

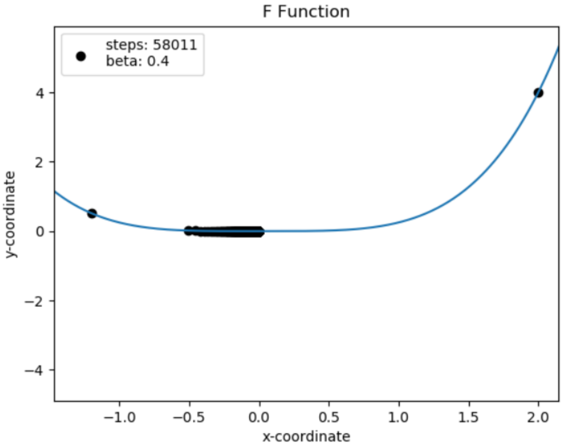
$$f(x) = \frac{1}{4}x^4$$

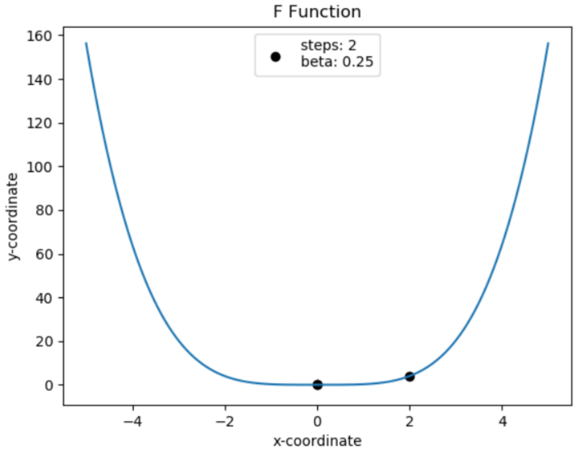
$$\nabla f(x) = x^3$$

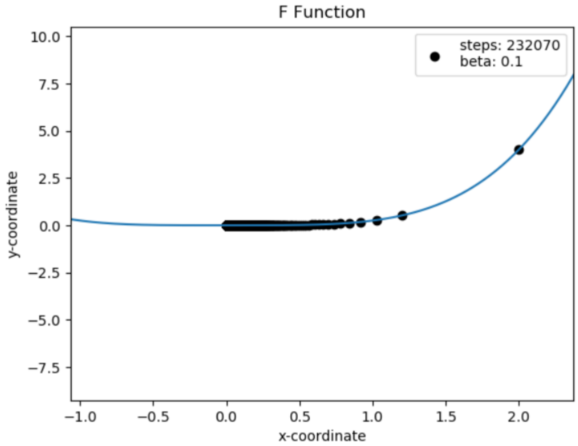


Funkcja f posiada minimum globalne w punkcie (0, 0). Dla dwóch wybranych punktów będziemy zmniejszać współczynnik β i obserwować zmianę w ilościach iteracji.

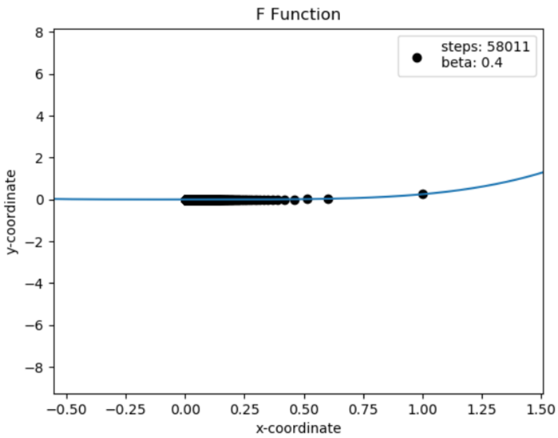
Pierwszym badanym punktem początkowym będzie (2, 4).

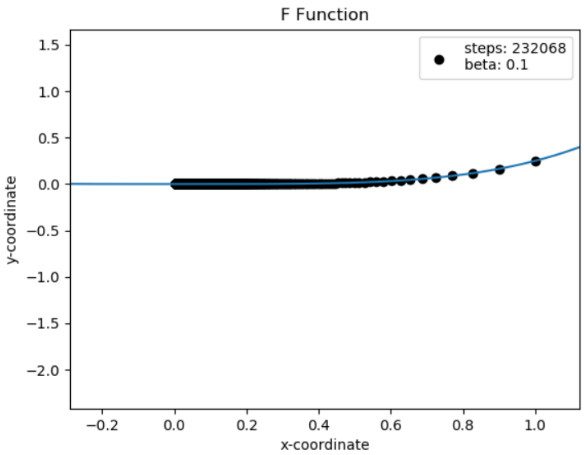


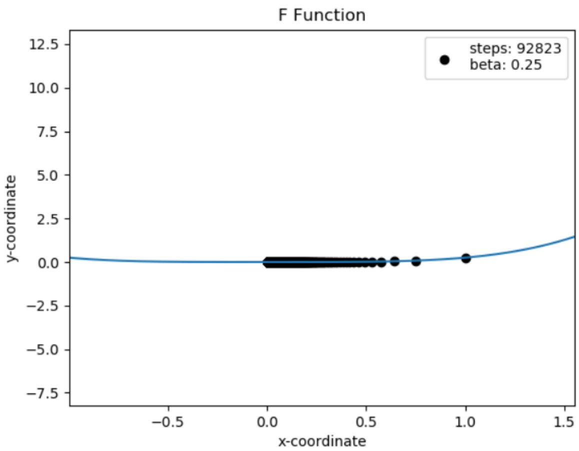




Drugim badanym punktem będzie (1, 0.25)





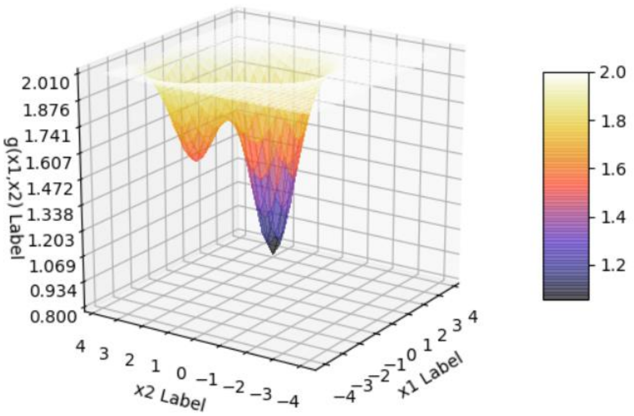


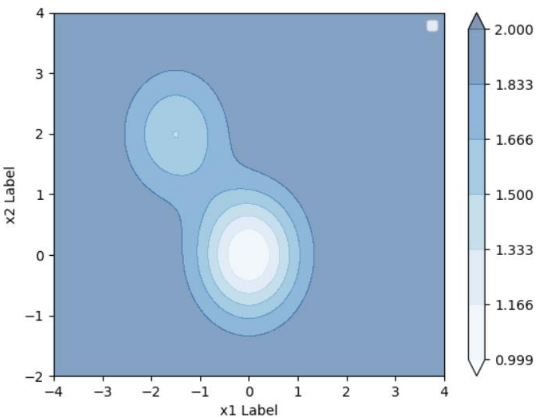
Współczynnik β wyraźnie wpływa na odległość między następnymi badami punktami. Można również zauważyć, że wraz z zmniejszaniem współczynnika β istotnie rośnie ilość iteracji, co czyni algorytm mało optymalnym. Wyjątkiem jest sytucja dla punktu (2, 4) i β=0.25, gdzie minimum jest znalezione po pierwszej iteracji. Warto również zauwżyć sytuacje dla punktu (2, 4) i β=0.4, gdzie pierwszy „krok” jest na tyle duży, że przeskakuje przez minimum globalne.

3. Funkcja g

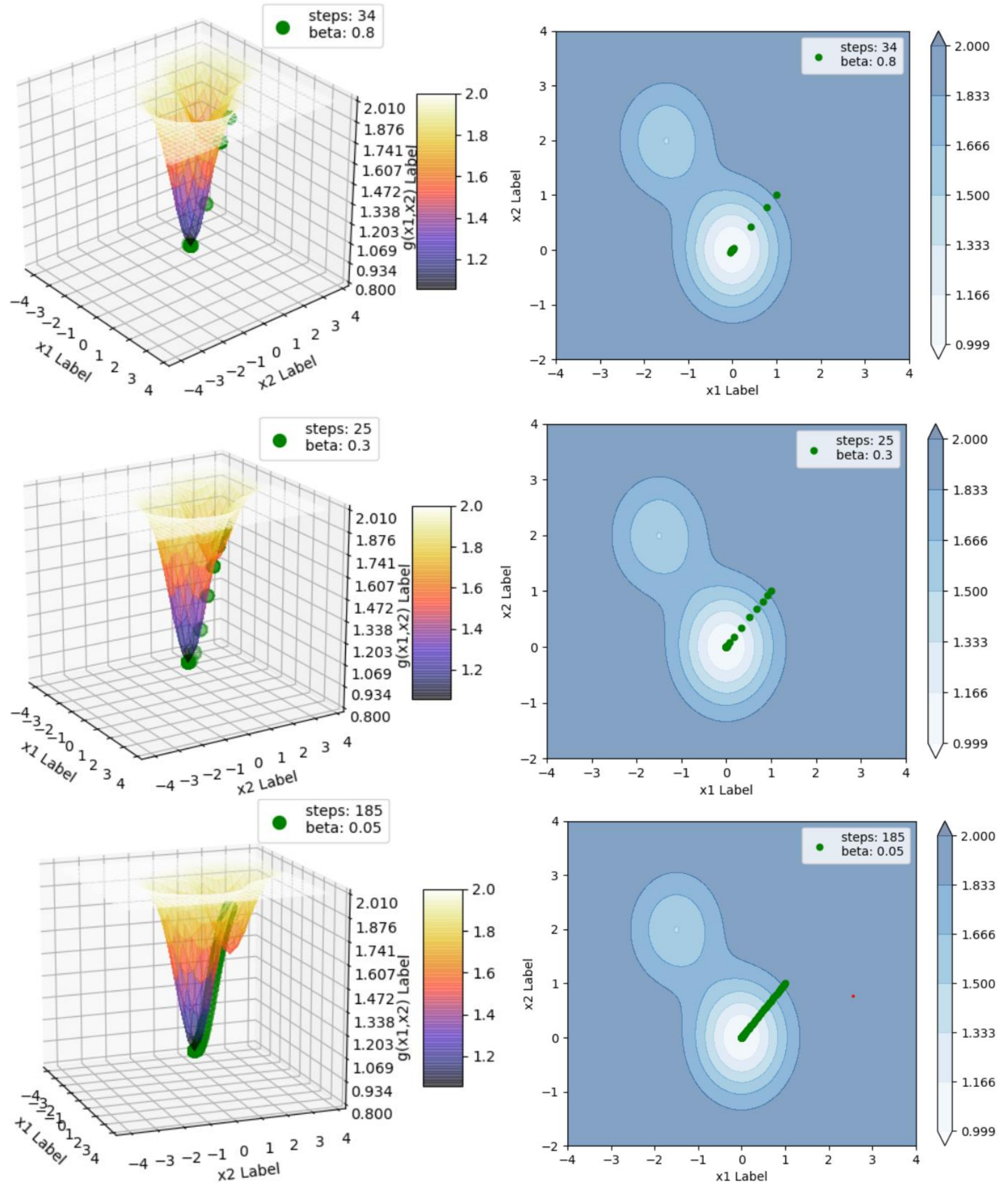
$$g(x) = 2 - \exp\left\{-x_1^2 - x_2^2\right\} - 0.5\exp\left\{-(x_1+1.5)^2 - (x_2-2)^2\right\}$$

$$\nabla g(x) = \begin{bmatrix} 2x_1\exp\left\{-x_1^2 - x_2^2\right\} + (x_1+1.5)\exp\left\{-(x_1+1.5)^2 - (x_2-2)^2\right\} \\ 2x_2\exp\left\{-x_1^2 - x_2^2\right\} + (x_2-2)\exp\left\{-(x_1+1.5)^2 - (x_2-2)^2\right\} \end{bmatrix}$$



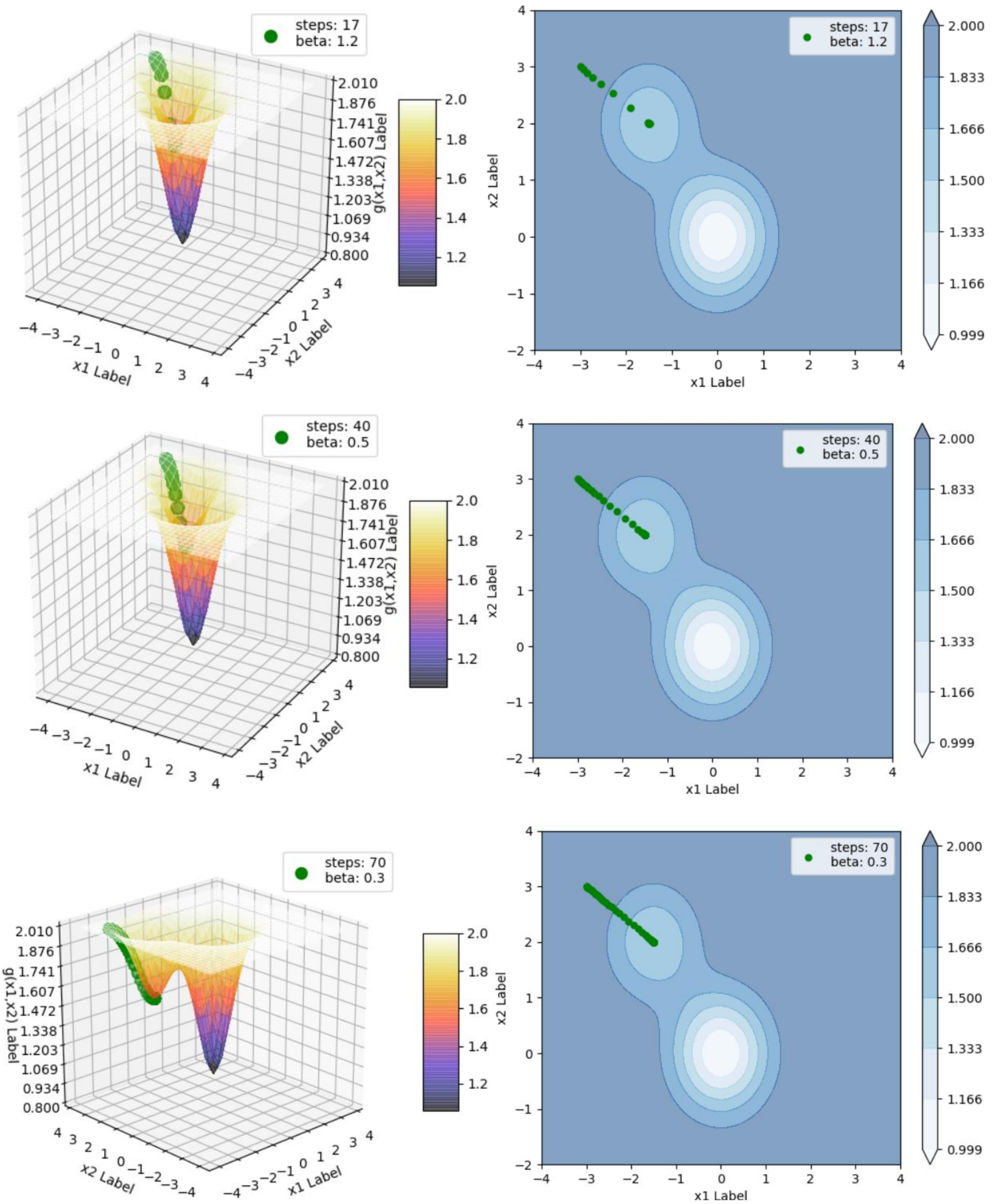


Funkcja g posiada minimum globalne oraz minimum lokalne. Pierwszym badanym punktem będzie punkt dla argumentu (1, 1). Punkt jest położony blisko minimum globalnego i celem badania będzie zmniejszanie współczynnika β i obserwacja ilości iteracji.



Dobranie za małego współczynnika β skutkuje zwiększoną ilością iteracji co spowalnia poszukiwania minimum. Z drugiej strony duży współczynnik β ma problem w końcowej fazie algorytmu podczas znajdywania dokładnego punktu. W wyżej pokazanym przykładzie najbardziej optymalny jest średni współczynnik β (β=0.3).

Drugim badanym punktem będzie punkt dla argumentów (-3, 3). Punkt położony bliżej minimum lokalnego celem badania będzie zmniejszania współczynnika β i obserwacja ilości iteracji.



W drugim przykładzie gradient odnajduję minimum lokalne. Widocznie widać, że wraz z zmniejszaniem współczynnika β rośnie liczba iteracji algorytmu.

4. Wnioski

Aby zoptymalizować wyszukiwanie minimum funkcji kluczową rolę odgrywa współczynnik β. W momencie, gdy parametr jest za mały dojście do minimum może trwać długo, lecz gdy ten parametr jest za duży algorytm nie może odnaleźć dokładnego punktu minimum przez za duże „skoki" pomiędzy argumentami.
Drugim ważnym parametrem będzie ε tj. długość wektora pochodnej w dokładnym minimum powinien być równy zero, lecz akceptując pewne przybliżenie minimum może on wynosić małą wartość np. (1e-7) co również przyspieszy odnajdywanie minimum.

Łukasz Wójcicki, 318746