WSI Zadanie 1. Krystian Piszczela

Nasza funkcja, przedstawiona za pomocą strony internetowej <https://c3d.libretexts.org/CalcPlot3D/index.html> wygląda następująco:

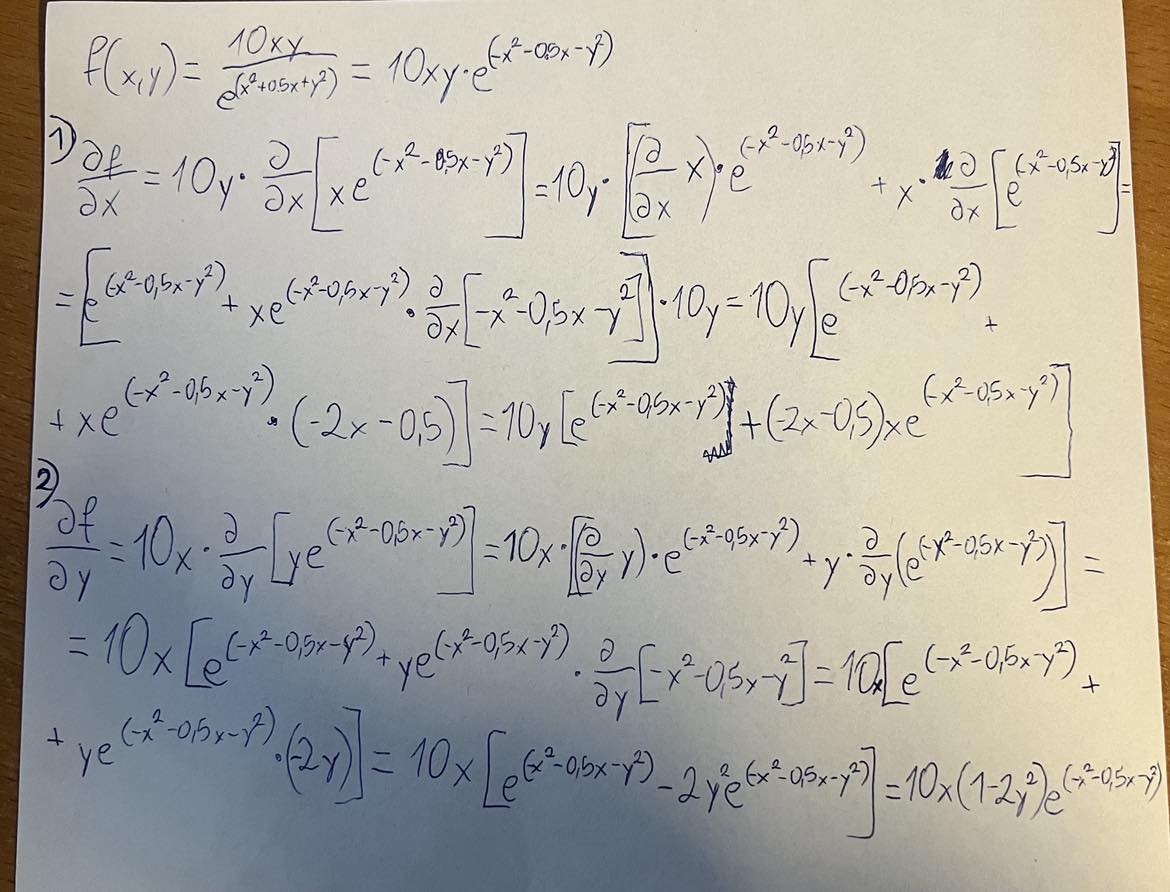
Obraz zawierający rysowanie, szkic, sztuka, design

Opis wygenerowany automatycznie

Jak można zauważyć – posiada ona dwa maksima oraz dwa minima. Algorytm stochastycznego spadku wzdłuż gradientu SGD pozwala nam na znalezienie jednego z wyżej wymienionych ekstremów. Gradient jest to wektor, który wskazuje kierunek, w jakim wzrost wartości pola skalarnego odbywa się najszybciej.  
Punkt startowy algorytmu decyduje o tym, czy dotrzemy do minimum (lub maksimum w zależności od naszego celu poszukiwań) lokalnego czy globalnego. Niestety powyższa funkcja nie daje nam gwarancji, iż wynik jest naszym minimum/maksimum. Przykładem może być sytuacja, kiedy punkt startowy zostanie wylosowany na przykład w jednym z rogów „za górką” widoczną na powyższym rysunku (zakładając, że patrzymy z punktu [0,0]). Może się okazać, że nasz wynik utknie w okolicy punktu startowego, gdyż pozostała część funkcji jest w przybliżeniu płaską powierzchnią i każda z potencjalnych dróg do uzyskania prawdziwego minimum znajduje się za wzniesieniem, którego algorytm niestety nie przeskoczy. Z tego też powodu, ograniczyłem rozważaną przez nas przestrzeń za pomocą następujących równań: oraz .   
Wielkość kroku uczącego może mieć znaczący wpływ na nasz wynik. Zbyt duży krok może spowodować, iż „przeskoczymy” nasze ekstremum, przez co go nie odnajdziemy, natomiast zbyt mały może spowodować, iż nie uda nam się nic odnaleźć w założonym przez nas limicie ilości iteracji.  
W celu potwierdzenia tych słów, przeprowadzony został mały eksperyment dla różnych wartości kroku uczącego oraz stałego punktu startowego - (0, 0). Aby poprawić precyzję znalezionego ekstremum możemy zmniejszyć krok uczenia. W niektórych przypadkach również zwiększenie limitu iteracji (lub zmniejszenie kryterium tolerancji stopu przy nieco innej implementacji funkcji SGD niż moja) może skutkować otrzymaniem lepszego wyniku.  
W celu potwierdzenia tych słów, przeprowadzony został mały eksperyment dla różnych wartości kroku uczącego oraz stałego punktu startowego (-0.9, 0.7). Naszym celem jest odnalezienie minimum globalnego przy limicie 1000 iteracji.

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący: | Odnalezione minimum: | Pozycja x: | Pozycja y: |
| 0.00001 | -2.695143401976056 | -0.8951268972495346 | 0.7007293405486688 |
| 0.0001 | -2.705169654095363 | -0.8661153950666174 | 0.7047034630202397 |
| 0.001 | -2.7076312904393296 | -0.8430758736679288 | 0.7071066488654649 |
| 0.01 | -2.7076312905811304 | -0.8430703308172541 | 0.707106781186547 |
| 0.1 | -2.7076312905811304 | -0.8430703308172536 | 0.7071067811865475 |
| 1 | -7.916810227557549e-07 | -4.327688239373855 | 0.31607289115646975 |

Jako, że punkt startowy znajdował się dość blisko oczekiwanej pozycji możemy tu zobaczyć wszystkie z wyżej wymienionych przypadków. W pierwszej sytuacji krok uczący jest zbyt mały – skutkuje to tym, że nasza pozycja nie została znacząco zmieniona i nie udało się dotrzeć do minimum. Dziesięciokrotne zwiększenie kroku poprawiło wynik, lecz nadal nieco odbiega on od właściwego. Następne trzy przypadki są już blisko odpowiedniego wyniku. Natomiast krok równy 1 idealnie prezentuje sytuację, gdzie wynik został „przeskoczony”.

Wyprowadzenie wzorów na pochodne cząstkowe użyte w programie:

Znalezione minima:

1. Globalne:   
   Punkt startowy to: -0.5400551209847606, -0.09125022112952053

Wartość x, dla której f(x, y) jest minimalne: -0.843070330817253

Wartość y, dla której f(x, y) jest minimalne: 0.707106781186547

Minimum funkcji: -2.7076312905811313

1. Lokalne:   
   Punkt startowy to: 1.4344749746232863, -0.4817995041182699

Wartość x, dla której f(x, y) jest minimalne: 0.5930703308172545

Wartość y, dla której f(x, y) jest minimalne: -0.7071067811865466

Minimum funkcji: -1.3301631910020266

Znalezione maksima:

1. Globalne:

Punkt startowy to: -0.12828593958629497, -1.6726572562873443

Wartość x, dla której f(x, y) jest minimalne: -0.843070330817253

Wartość y, dla której f(x, y) jest minimalne: -0.707106781186548

Minimum funkcji: 2.707631290581131

1. Lokalne:

Punkt startowy to: -0.5764121267333113, 1.3499547051752199

Wartość x, dla której f(x, y) jest minimalne: 0.5930703308172528

Wartość y, dla której f(x, y) jest minimalne: 0.7071067811865486

Minimum funkcji: 1.3301631910020268