Obraz zawierający tekst, Czcionka, biały, czarne i białe

Opis wygenerowany automatycznie

Sprawozdanie

Wprowadzenie do sztucznej inteligencji

Ćwiczenie nr. 1

Mikołaj Bańkowski

Numer albumu 310408

prowadzący

Grzegorz Rypeść

Warszawa 2023

Spis treści

[1. Treść ćwiczenia 3](#_Toc160480410)

[2. Pochodne cząstkowe funkcji 4](#_Toc160480411)

[3. Minima i maksima funkcji 5](#_Toc160480412)

[4. Czym jest gradient 14](#_Toc160480413)

[5. Jak punkt startowy wpływa wynik 15](#_Toc160480414)

[6. Jak wartość kroku uczącego wpływa na proces optymalizacji 19](#_Toc160480415)

[7. Jak można zwiększyć precyzję znalezionych ekstremów 27](#_Toc160480416)

# 1. Treść ćwiczenia

Proszę znaleźć minima oraz maksima funkcji wykorzystując algorytm spadku wzdłuż gradientu SGD (omówiony na wykładzie). Wzór na pochodne cząstkowe należy wyprowadzić ręcznie na kartce i jej skan/zdjęcie załączyć do raportu.

Odpowiedzieć na pytania:

1. Czym jest gradient
2. Jak punkt startowy algorytmu wpływa na wynik?
3. Jak wartość kroku uczącego wpływa na proces optymalizacji?
4. Jak można zwiększyć precyzję znalezionych ekstremów?

Odpowiedzi na pytania należy podeprzeć eksperymentami. Wyniki eksperymentów, odpowiedzi oraz wnioski należy zamieścić w raporcie

Uwagi:

Należy tę funkcję sobie zwizualizować korzystając przykładowo z:

https://c3d.libretexts.org/CalcPlot3D/index.html lub https://www.wolframalpha.com/.

# 2. Pochodne cząstkowe funkcji

Obraz zawierający tekst, pismo odręczne, Czcionka, numer

Opis wygenerowany automatycznie

# 3. Minima i maksima funkcji

# Do wstępnej analizy badanej funkcji i jej wizualizacji, wykorzystano narzędzie [WolframAplha](https://www.wolframalpha.com/input?i2d=true&i=Divide%5B9xy%2CPower%5Be%2CPower%5Bx%2C2%5D%2B0.5x%2BPower%5By%2C2%5D%5D%5D)

# Obraz zawierający diagram, origami, design, sztuka Opis wygenerowany automatycznie przy średnim poziomie pewności

# Obraz zawierający zrzut ekranu, krąg, Wielobarwność Opis wygenerowany automatycznie

# Na podstawie analizy poniższych wykresów funkcji, można zauważyć, że funkcja posiada dwa minima i dwa maksima.

# Minima funkcji znajdują się w II i IV ćwiartce układu współrzędnych. Są to kolejno

# min{f(x,y)} ≈ -2.43687 dla (x, y) ≈ (-0.84307, 0.707107)

# min{f(x,y)} ≈ -1.19715 dla (x, y) ≈ (0.59307, -0.707107)

# Maxima funkcji znajdują się w I i III ćwiartce układu współrzędnych. Są to kolejno

# max{f(x,y)} ≈ 1.19715 dla (x, y) ≈ (0.59307, 0.707107)

# max{f(x,y)} ≈ 2.43687 dla (x, y) ≈ (-0.84307, -0.707107)

# Minima funkcji znalezione za pomocą algorytm spadku wzdłuż gradientu SGD

# Dla danych wejściowych:

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione minimum |
| 0.1 | x = 0  y = -1 | x = 0.593070323659 y = -0.707108068995 | -1.197146871898 |

# Obraz zawierający tekst, diagram, zrzut ekranu, krąg Opis wygenerowany automatycznie

# Obraz zawierający zrzut ekranu, Wielobarwność, tekst, diagram Opis wygenerowany automatycznie

# Obraz zawierający diagram, zrzut ekranu, kwadrat, linia Opis wygenerowany automatycznie

# Dla danych wejściowych

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione minimum |
| 0.1 | x = 0  y = 1 | x = -0.843070208719 y = 0.707106781187 | -2.436868161523 |

# 

# Obraz zawierający tekst, zrzut ekranu, diagram, krąg Opis wygenerowany automatycznie

# Obraz zawierający zrzut ekranu, tekst, Wielobarwność Opis wygenerowany automatycznie

# Obraz zawierający zrzut ekranu, diagram, kwadrat, Prostokąt Opis wygenerowany automatycznie

# Maxima funkcji znalezione za pomocą algorytm spadku wzdłuż gradientu SGD

# Dla danych wejściowych

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione maximun |
| 0.1 | x = 0  y = 1 | x = 0.593070323659 y = 0.707108068995 | 1.197146871898 |

# Obraz zawierający tekst, diagram, zrzut ekranu, krąg Opis wygenerowany automatycznie

# Obraz zawierający zrzut ekranu, tekst, Wielobarwność Opis wygenerowany automatycznie

# Obraz zawierający diagram, zrzut ekranu, kwadrat, design Opis wygenerowany automatycznie

# Dla danych wejściowych

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione maximum |
| 0.1 | x = 0  y = -1 | x = -0.843070208719 y = -0.707106781187 | 2.436868161523 |

# Obraz zawierający tekst, diagram, zrzut ekranu, krąg Opis wygenerowany automatycznie

# Obraz zawierający zrzut ekranu, Wielobarwność, tekst Opis wygenerowany automatycznie

# Obraz zawierający diagram, zrzut ekranu, kwadrat, linia Opis wygenerowany automatycznie

# 4. Czym jest gradient

# Gradient to wektor, który w danym punkcie przestrzeni wskazuje kierunek i wartość najszybszego wzrostu funkcji skalarnej w tym punkcie. W jego skład wchodzą pochodne cząstkowe funkcji względem każdej ze zmiennych niezależnych.

# W kontekście optymalizacji funkcji celu, gradient jest kluczowym narzędziem, które pomaga algorytmom optymalizacyjnym znajdować minimum lub maksimum funkcji. Dzięki niemu możemy określić kierunek najszybszego wzrostu (gdzie wartość funkcji rośnie najszybciej) oraz przeciwnie - kierunek najszybszego spadku (gdzie wartość funkcji maleje najszybciej).

# Algorytm spadku wzdłuż gradientu SGD używa gradientu do aktualizacji parametrów modelu w celu minimalizacji funkcji kosztu. Algorytmy te iteracyjnie poruszają się w kierunku przeciwnym do gradientu, aż osiągną zbieżność do lokalnego minimum lub maksimum

# 5. Jak punkt startowy wpływa wynik

# Szukanie minimum funkcji za pomocą SGD w zależności od punktu startowego

# Algorytm spadku wzdłuż gradientu został uruchomiony 20 razy w celu znalezienia minimum funkcji, punkt startowy x i punkt startowy y, za każdym razem były wybierane losowo z przedziału [-2,2]S

# Tabela przedstawia rezultat eksperymentu

# Obraz zawierający tekst, zrzut ekranu, menu Opis wygenerowany automatycznie

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione minimum |
| 0.1 | x = 1.2  y = -1.2 | x = -2.331571045911 y = -2.835618989202 | 0.000267833261 |

# Obraz zawierający tekst, zrzut ekranu, diagram, krąg Opis wygenerowany automatycznie

# Obraz zawierający zrzut ekranu, Wielobarwność, tekst Opis wygenerowany automatycznie

# Obraz zawierający diagram, zrzut ekranu, kwadrat, Prostokąt Opis wygenerowany automatycznie

# Szukanie maximum funkcji za pomocą SGD w zależności od punktu startowego

# Algorytm spadku wzdłuż gradientu został uruchomiony 20 razy w celu znalezienia maximum funkcji, punkt startowy x i punkt startowy y, za każdym razem były wybierane losowo z przedziału [-2,2]

# Tabela przedstawia rezultat eksperymentu

# Obraz zawierający tekst, zrzut ekranu, menu Opis wygenerowany automatycznie

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione maksimum |
| 0.1 | x = 1.2  y = -1.2 | x = 2.470023196368 y = -2.170725183170 | -0.000282576685 |

# Obraz zawierający diagram, tekst, zrzut ekranu, krąg Opis wygenerowany automatycznie

# Obraz zawierający zrzut ekranu, Wielobarwność, tekst Opis wygenerowany automatycznie

# Obraz zawierający diagram, zrzut ekranu, kwadrat, Prostokąt Opis wygenerowany automatycznie

# Wybór różnych punktów startowych może prowadzić do różnych lokalnych minimów/maksimów

# Punkt startowy blisko minimum może przyspieszyć zbieżność, ale nie zawsze jest łatwy do znalezienia.

# Wybór punktu startowego może wpływać na liczbę iteracji wymaganych do zbieżności algorytmu.

# Dobranie nieodpowiedniego punktu startowego może doprowadzić, że algorytm źle zadziała, zamiast zbliżać się do oczekiwanego rozwiązania będzie się od niego oddalać

# 6. Jak wartość kroku uczącego wpływa na proces optymalizacji

# Szukanie minimum funkcji za pomocą SGD w zależności od kroku uczącego

# Algorytm spadku wzdłuż gradientu został uruchomiony 7 razy, z różnym krokiem uczącym w celu znalezienia minimum funkcji, dla punktu startowego x = 0, y = -1 oraz x = 0, y = 1

# Obraz zawierający tekst, zrzut ekranu, Czcionka Opis wygenerowany automatycznie

# Obraz zawierający tekst, zrzut ekranu, Czcionka Opis wygenerowany automatycznie

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione minimum |
| 0.001 | x = 0  y = -1 | x = 0.591573675756 y = -0.711839163416 | -1.197086871197 |

# Obraz zawierający tekst, zrzut ekranu, diagram, Wykres Opis wygenerowany automatycznie

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione minimum |
| 0.1 | x = 0  y = -1 | x = 0.593070323659 y = -0.707108068995 | -1.197146871898 |

# Obraz zawierający tekst, zrzut ekranu, diagram, krąg Opis wygenerowany automatycznie

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione minimum |
| 1 | x = 0  y = -1 | x = 3.772484966940 y = -0.428943649562 | -0.000001211866 |

# Obraz zawierający tekst, zrzut ekranu, diagram, Wykres Opis wygenerowany automatycznie

# Szukanie maximum funkcji za pomocą SGD w zależności od kroku uczącego

# Algorytm spadku wzdłuż gradientu został uruchomiony 7 razy, z różnym krokiem uczącym w celu znalezienia maximum funkcji, dla punktu startowego x = 0, y = -1 oraz x = 0, y = 1

# Tabela przedstawia rezultat eksperymentu

# Obraz zawierający tekst, zrzut ekranu, Czcionka, numer Opis wygenerowany automatycznie

# Obraz zawierający tekst, zrzut ekranu, Czcionka Opis wygenerowany automatycznie

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione maksimum |
| 0.001 | x = 1.2  y = -1.2 | x = -0.842806526532 y = -0.707163165748 | 2.436867857116 |

# Obraz zawierający tekst, zrzut ekranu, diagram, Wykres Opis wygenerowany automatycznie

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione maksimum |
| 0.1 | x = 0 y = -1 | x = -0.843070208719 y = -0.707106781187 | 2.436868161523 |

# Obraz zawierający tekst, zrzut ekranu, diagram, krąg Opis wygenerowany automatycznie

|  |  |  |  |
| --- | --- | --- | --- |
| Krok uczący | Punkt startowy | Punkt końcowy | Znalezione maximum |
| 1 | x = 0 y = -1 | x = -3.850166167893 y = .369230749634 | -0.000027924101 |

# Obraz zawierający tekst, zrzut ekranu, diagram, Wykres Opis wygenerowany automatycznie

# Wartość kroku uczącego wpływa na szybkość zbieżności algorytmu. Zbyt mała wartość kroku może spowodować wolną zbieżność, podczas gdy zbyt duża może prowadzić do oscylacji lub braku zbieżności.

# Zbyt duży krok uczący może sprawić, że algorytm "przeskakuje" minimum, podczas gdy zbyt mały krok może spowolnić zbliżanie się do minimum lub jego nieosiągnięcie

# 7. Jak można zwiększyć precyzję znalezionych ekstremów Poprzez dobranie odpowiedniego punktu startowego, im bliżej punkt startowy znajduje się szukanego minimum/maksimum tym większa szansa, że znajdziemy szukane rozwiązanie.

# Poprzez dobranie odpowiedniego kroku uczącego, bardzo mały krok uczący wymaga większej liczby iteracji, ale może prowadzić do precyzyjniejszych wyników. Dobór wartości kroku to kompromis pomiędzy szybkością zbieżności a precyzją

# Możemy jednak połączyć dobranie odpowiedniego punktu startowego z bardzo małym krokiem. Punkt startowy położony bliżej szukanego rozwiązania zmniejszy nam ilość iteracji, natomiast mniejszy krok uczenia pozwoli nam na otrzymanie dokładniejszego wyniku.