Optymalizacja ciągła

Test zaliczeniowy, grupa: $\boxed{\text{A}}$

Imię i Nazwisko: .... **Numer indeksu:** ....

Uwaga: w pytaniach otwartych wpisuj gotowe odpowiedzi, a rozwiązania szukaj w brudnopisie!

1. `[3pkt]` Wyznacz analitycznie minimum funkcji jednej zmiennej na przedziale $[0, 1]$ zadanej wzorem:

$$f(x) = x - \sqrt{x + x^2}.$$

Wyznacz pochodną funkcji:

$f'(x) =$

Wyznacz kandydatów na minimum (pokaż obliczenia):

Odp: minimum w $x =$

Określ czy funkcja jest wypukła na przedziale $[0, 1]$:

2. `[3pkt]` Rozważ minimalizację funkcji jednej zmiennej:

$$f(x) = x^3 + 2x^2 - 7x + 13.75.$$

Rozpoczynając od przedziału $[-2, 14]$ wykonaj 3 kroki algorytmu bisekcji.

Wyznacz pochodną:

$f'(x) =$

Krok 1:

Krok 2:

Krok 3:

3. `[3pkt]` Używając operacji zachowujących wypukłość, pokaż, że funkcja $f : \mathbb{R}^3 \to \mathbb{R}$:

$$f(x, y, z) = (2x - 3y + 2z)^2 + e^{x-y},$$

jest wypukła.

Odp:

4. `[4pkt]` Rozważ minimalizację funkcji dwóch zmiennych:

$$f(x, y) = (x^2 - y^2)^2 + x^2 + y^2.$$

Rozpoczynając od punktu $(x_0, y_0) = (1, 1)$ wykonaj jeden krok:

- metodą spadku wzdłuż gradientu z ustaloną długością kroku $\alpha = 1$,
- metodą Newtona-Raphsona.

Wyznacz pochodne cząstkowe $f(x, y)$:

Stąd $\nabla f(x_0, y_0) =$

Wyznacz drugie pochodne cząstkowe $f(x,y)$:

Stąd $\nabla^2 f(x_0,y_0) =$

$\left(\nabla^2 f(x_0,y_0)\right)^{-1} =$

Rozwiązanie po pierwszym kroku algorytmu spadku wzdłuż gradientu (napisz wzór ogólny i wartość liczbową):

$\boldsymbol{x}_1 = (x_1, y_1) =$

Pierwszy krok algorytmu Newtona-Raphsona (napisz wzór ogólny i wartość liczbową):

$\boldsymbol{x}_1 = (x_1, y_1) =$

5. `[2pkt]` Przy każdym zdaniu dotyczącym metody gradientów sprzężonych (MGS) zaznacz, czy jest prawdziwe (`P`), czy fałszywe (`F`):

   (a) MGS znajduje dokładne rozwiązanie optymalne funkcji kwadratowej w n krokach, gdzie n to wymiar problemu.

   (b) Kierunki sprzężone są zawsze kierunkami poprawy, o ile tylko nie zostało osiągnięte rozwiązanie optymalne.

   (c) Metody Fletchera-Reevesa i Polaka-Ribière'a działają inaczej przy optymalizacji funkcji kwadratowej.

   (d) MGS wymaga policzenia hesjanu przy optymalizacji funkcji innej, niż kwadratowa.

   (e) Kierunki sprzężone są do siebie ortogonalne.

6. `[2pkt]` Przy każdym zdaniu dotyczącym omówionych metod spadkowych zaznacz, czy jest prawdziwe (`P`), czy fałszywe (`F`):

   (a) Żadna z metod nie gwarantuje znalezienia optymalnego rozwiązania funkcji niewypukłej.

   (b) Przy małej wartości wskaźnika uwarunkowania hesjanu w optimum metoda spadku wzdłuż gradientu będzie działać bardzo wolno.

   (c) Metoda Newtona-Raphsona zawsze zbiega do optymalnego rozwiązania dla funkcji wypukłych.

   (d) Liniowy rząd zbieżności oznacza, że odległość od optimum maleje odwrotnie proporcjonalnie do liczby iteracji k.

   (e) Kwadratowy rząd zbieżności oznacza, że odległość od optimum maleje podwójnie wykładniczo względem liczby iteracji k.

   (f) Metoda najszybszego spadku wzdłuż gradientu zbiega do optimum funkcji wypukłej z liniowym rzędem zbieżności.

7. `[2pkt]` Przy każdym zdaniu dotyczącym metody stochastycznego spadku wzdłuż gradientu (SGD) zaznacz, czy jest prawdziwe (`P`), czy fałszywe (`F`):

   (a) Jeżeli w problemie regresji liniowej z jedną zmienną ($\min_{a,b} \sum_{i=0}^{n} (y_i - (ax_i + b))^2$) zainicjalizujemy współczynnik kierunkowy a oraz wyraz wolny b na tę samą wartość to po jednej iteracji algorytmu SGD a nadal będzie równe b.

   (b) Jedna iteracja algorytmu SGD ma mniejszy koszt obliczeniowy niż jedna iteracja algorytmu spadku wzdłuż gradientu.

   (c) Technika mini-batch dla rozmiaru paczki równego 1 sprowadza się do algorytmu spadku wzdłuż gradientu.

   (d) Algorytm SGD zainicjalizowany w miniminum funkcji może oddalić się od niego w pierwszej iteracji.

8. `[3pkt]` Dana jest wypukła wielowymiarowa funkcja $f(\boldsymbol{x})$. Używając definicji funkcji wypukłej udowodnij, że jednowymiarowa funkcja $g(\alpha) = f(\boldsymbol{x} - \alpha \nabla f(\boldsymbol{x}))$ optymalizowana w czasie iteracji algorytmu najszybszego spadku gradientu dla danego $\boldsymbol{x}$ jest również wypukła.