

AKADEMIA GÓRNICZO-HUTNICZA

WYDZIAŁ INFORMATYKI  
KIERUNEK INFORMATYKA



METODY OBLICZENIOWE W NAUCE I TECHNICE

---

## Laboratorium 11

Optymalizacja

---

Wojciech Michaluk, Kyrylo Iakymenko

Kraków, 7 czerwca 2024

# 1 Wprowadzenie

W ramach zadań podczas dzisiejszego laboratorium będziemy się zajmować optymalizacją. Optymalizacja to szeroka klasa zagadnień, ale w tym przypadku głównie skupiamy się na znajdowaniu minimum funkcji wielu zmiennych. W przypadku problemu optymalizacji często utożsamiamy go właśnie z **minimalizacją**, gdyż najczęściej dążymy do zminimalizowania wartości określonej funkcji, zwanej *funkcją kosztu*. Istnieje wiele technik optymalizacyjnych, natomiast my skupimy się na zastosowaniu wyznaczania ekstremów funkcji danej wzorem analitycznym oraz wykorzystaniu metody największego spadku.

W zadaniu pierwszym zajmiemy się wyznaczaniem tzw. *punktów krytycznych* dla podanych funkcji. Są to punkty, wśród których można spodziewać się minimum funkcji, natomiast nie ma na to gwarancji - sprawdzając dodatkowe warunki, można zweryfikować rodzaj punktu krytycznego.

W drugim zadaniu zastosujemy technikę optymalizacji - metodę największego spadku - w ciekawym zagadnieniu praktycznym, modelując ruch robota. Zadanie polega na znalezieniu najkrótszej ścieżki ruchu robota między dwoma wybranymi punktami, ale należy unikać przeszkód usytuowanych na trasie robota.

## 2 Zadanie 1

### 2.1 Opis zadania

Zadanie polega na wyznaczeniu punktów krytycznych oraz ich charakteryzacji (określeniu ich typu) dla każdej z badanych funkcji:

- $f_1(x, y) = x^2 - 4xy + y^2$
- $f_2(x, y) = x^4 - 4xy + y^4$
- $f_3(x, y) = 2x^3 - 3x^2 - 6xy(x - y - 1)$
- $f_4(x, y) = (x - y)^4 + x^2 - y^2 - 2x + 2y + 1$

Również zbadamy, czy powyższe funkcje posiadają minimum globalne lub maksimum globalne na zbiorze  $\mathbb{R}^2$ .

### 2.2 Opracowanie zadania

Przedstawmy dokładnie, jak rozpoznać, że mamy do czynienia z punktem krytycznym. Punkt krytyczny to punkt, który albo jest tzw. *punktem stacjonarnym*, czyli takim, w którym funkcja jest różniczkowalna oraz wszystkie pochodne przyjmują wartość 0, co można zapisać dla rozważanej funkcji  $F$  jako:

$$\nabla F(x) = 0,$$

albo jest punktem, w którym pochodna nie istnieje<sup>1</sup>. Czasami jednak utożsamiamy pojęcie punktu krytycznego z punktem stacjonarnym.

Klasyfikujemy je w sposób następujący:

1. maksimum lokalne,
2. minimum lokalne,
3. tzw. *punkt siodłowy*.

Rodzaj punktu krytycznego determinujemy na podstawie kolejnych wiodących minorów głównych **macierzy Hessego** dla danej funkcji  $F$ . Macierz Hessego to macierz drugich pochodnych cząstkowych funkcji, oznaczana także jako  $\nabla^2 F$ .

Jeśli wszystkie wartości kolejnych wyznaczników są dodatnie, mamy do czynienia z **minimum**. Z kolei jeśli znaki wyznaczników występują na przemian, począwszy od minusa,

---

<sup>1</sup>Za wikipedią: [https://pl.wikipedia.org/wiki/Punkt\\_krytyczny\\_\(matematyka\)](https://pl.wikipedia.org/wiki/Punkt_krytyczny_(matematyka))

to w punkcie jest **maksimum**. Pozostaje przypadek, gdzie znaki są zachowane, ale występują wartości zerowe - wtedy trzeba dokonać dalszej, dogłębnej analizy - oraz każdy inny przypadek, kiedy w punkcie występuje **punkt siodłowy**.

Aby sprawdzić, czy funkcja posiada ekstrema globalne, zbadamy jak zachowuje się ona na krańcach dziedziny oraz porównamy wartości w znalezionych ekstremach lokalnych.

Poniżej, w kolejnych sekcjach badamy każdą z wymienionych funkcji.

**Funkcja**  $f_1(x, y) = x^2 - 4xy + y^2$

**Pierwsze pochodne cząstkowe:**

$$\begin{aligned}f_{1x} &= \frac{\partial}{\partial x}(x^2 - 4xy + y^2) = 2x - 4y \\f_{1y} &= \frac{\partial}{\partial y}(x^2 - 4xy + y^2) = -4x + 2y\end{aligned}$$

**Wyznaczenie punktów krytycznych:**

Rozwiązujemy układ równań:

$$\begin{cases} 2x - 4y = 0 \\ -4x + 2y = 0 \end{cases}$$

Dzieląc pierwsze równanie przez 2, otrzymujemy  $x = 2y$ . Podstawiając  $x = 2y$  do drugiego równania:

$$-4(2y) + 2y = 0 \implies -8y + 2y = 0 \implies -6y = 0 \implies y = 0$$

Zatem  $x = 0$ . Stąd punkt krytyczny to  $(0, 0)$ .

**Drugie pochodne cząstkowe i macierz Hessego:**

$$\begin{aligned}f_{1xx} &= \frac{\partial^2}{\partial x^2}(x^2 - 4xy + y^2) = 2 \\f_{1yy} &= \frac{\partial^2}{\partial y^2}(x^2 - 4xy + y^2) = 2 \\f_{1xy} &= f_{1yx} = \frac{\partial^2}{\partial x \partial y}(x^2 - 4xy + y^2) = -4\end{aligned}$$

Macierz Hessego to:

$$H = \begin{bmatrix} 2 & -4 \\ -4 & 2 \end{bmatrix}.$$

**Wartości własne macierzy Hessego:**

Liczymy kolejne wiodące minory główne macierzy Hessego:

$$\begin{aligned}\det([2]) &= 2 \\ \det(H) &= 2 \cdot 2 - (-4) \cdot (-4) = 4 - 16 = -12\end{aligned}$$

Ponieważ pierwszy wyznacznik jest dodatni, a drugi ujemny, macierz Hessego jest nie-określona. Punkt  $(0, 0)$  jest punktem siodłowym.

Funkcja nie ma żadnego ekstremum lokalnego, a na krańcach dziedziny ucieka zarówno do  $+\infty$ , jak i  $-\infty$ , więc nie posiada ekstremów globalnych.

**Funkcja**  $f_2(x, y) = x^4 - 4xy + y^4$

**Pierwsze pochodne cząstkowe:**

$$\begin{aligned}f_{2x} &= \frac{\partial}{\partial x}(x^4 - 4xy + y^4) = 4x^3 - 4y \\f_{2y} &= \frac{\partial}{\partial y}(x^4 - 4xy + y^4) = -4x + 4y^3\end{aligned}$$

**Wyznaczenie punktów krytycznych:**

Rozwiązujemy układ równań:

$$\begin{cases} 4x^3 - 4y = 0 \\ -4x + 4y^3 = 0 \end{cases}$$

Dzieląc pierwsze równanie przez 4, otrzymujemy  $x^3 = y$ . Podstawiając  $x^3 = y$  do drugiego równania:

$$-4x + 4(x^3)^3 = 0 \implies -4x + 4x^9 = 0 \implies 4x(x^8 - 1) = 0$$

Zatem  $x = 0$  lub  $x^8 = 1$ . Stąd  $x = 0$  lub  $x = \pm 1$ . Rozpisujemy wszystkie przypadki:

1. Dla  $x = 0$ :  $y = 0$ , więc punkt krytyczny to  $(0, 0)$ .
2. Dla  $x = 1$ :  $y = 1$ , więc punkt krytyczny to  $(1, 1)$ .
3. Dla  $x = -1$ :  $y = -1$ , więc punkt krytyczny to  $(-1, -1)$ .

**Drugie pochodne cząstkowe i macierz Hessego:**

$$\begin{aligned}f_{2xx} &= \frac{\partial^2}{\partial x^2}(x^4 - 4xy + y^4) = 12x^2 \\f_{2yy} &= \frac{\partial^2}{\partial y^2}(x^4 - 4xy + y^4) = 12y^2 \\f_{2xy} &= f_{2yx} = \frac{\partial^2}{\partial x \partial y}(x^4 - 4xy + y^4) = -4\end{aligned}$$

Macierz Hessego to:

$$H = \begin{bmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{bmatrix}.$$

**Wartości własne macierzy Hessego:**

Liczymy kolejne wiodące minory główne macierzy Hessego w punkcie  $(0, 0)$ :

$$\begin{aligned}\det([0]) &= 0 \\ \det(H(0, 0)) &= 12 \cdot 0 \cdot 12 \cdot 0 - (-4) \cdot (-4) = 0 - 16 = -16\end{aligned}$$

Ponieważ pierwszy wyznacznik jest zerowy, a drugi jest ujemny, macierz Hessego jest nieokreślona. Punkt  $(0, 0)$  jest punktem siodłowym.

Analogiczne obliczenia w punktach  $(1, 1)$  i  $(-1, -1)$ :

$$\begin{aligned}\det([12]) &= 12 \\ \det(H(1, 1)) &= \det(H(-1, -1)) = 12 \cdot 1 \cdot 12 \cdot 1 - (-4) \cdot (-4) = 144 - 16 = 128\end{aligned}$$

Ponieważ obydwa wyznaczniki są dodatnie, punkty  $(1, 1)$  i  $(-1, -1)$  są minimami lokalnymi.

Funkcja  $f_2$  na krańcach dziedziny ucieka do nieskończoności, stąd stwierdzamy, że nie posiada maksimum globalnych. Posiada natomiast **minima globalne**: są to wyznaczone przez nas przed chwilą minima lokalne. Funkcja przyjmuje najmniejszą wartość -2 w punktach  $(1, 1)$  oraz  $(-1, -1)$ .

**Funkcja**  $f_3(x, y) = 2x^3 - 3x^2 - 6xy(x - y - 1)$

**Pierwsze pochodne cząstkowe:**

$$\begin{aligned} f_{3x} &= \frac{\partial}{\partial x}(2x^3 - 3x^2 - 6xy(x - y - 1)) = 6x^2 - 6x - 6y(x - y - 1) - 6xy \\ &= 6x^2 - 6x - 6yx + 6y^2 + 6y - 6xy = 6x^2 - 6x - 12xy + 6y^2 + 6y \\ f_{3y} &= \frac{\partial}{\partial y}(2x^3 - 3x^2 - 6xy(x - y - 1)) = -6x(x - y - 1) + 6xy \\ &= -6x^2 + 6xy + 6x + 6xy = -6x^2 + 6x + 12xy \end{aligned}$$

**Wyznaczenie punktów krytycznych:**

Rozwiązujemy układ równań:

$$\begin{cases} 6x^2 - 6x - 12xy + 6y^2 + 6y = 0 \\ -6x^2 + 6x + 12xy = 0 \end{cases}$$

Dodajemy równania stronami, otrzymując:

$$6y^2 + 6y = 0.$$

Rozwiązujemy to równanie.

$$6y^2 + 6y = 0 \implies 6y(y + 1) = 0 \implies y = 0 \text{ lub } y = -1$$

Wyliczamy możliwe wartości  $x$ , korzystając z drugiego równania.

Dla  $y = 0$ :

$$\begin{aligned} -6x^2 + 6x + 12x(0) = 0 &\implies -6x^2 + 6x = 0 \implies -6x(x - 1) = 0 \\ &\implies x = 0 \text{ lub } x = 1 \end{aligned}$$

Punkty krytyczne:  $(0, 0)$  oraz  $(1, 0)$ .

Dla  $y = -1$ :

$$\begin{aligned} -6x^2 + 6x + 12x(-1) = 0 &\implies -6x^2 + 6x - 12x = 0 \\ &\implies -6x^2 - 6x = 0 \implies -6x(x + 1) = 0 \\ &\implies x = 0 \text{ lub } x = -1 \end{aligned}$$

Punkty krytyczne:  $(0, -1)$  oraz  $(-1, -1)$ .

**Drugie pochodne cząstkowe i macierz Hessego:**

$$\begin{aligned} f_{3xx} &= \frac{\partial^2}{\partial x^2}(2x^3 - 3x^2 - 6xy(x - y - 1)) = 12x - 6 - 12y \\ f_{3yy} &= \frac{\partial^2}{\partial y^2}(2x^3 - 3x^2 - 6xy(x - y - 1)) = 12x \\ f_{3xy} &= f_{3yx} = \frac{\partial^2}{\partial x \partial y}(2x^3 - 3x^2 - 6xy(x - y - 1)) = -12x + 12y + 6 \end{aligned}$$

Macierz Hessego to:

$$H = \begin{bmatrix} 12x - 6 - 12y & -12x + 12y + 6 \\ -12x + 12y + 6 & 12x \end{bmatrix}.$$

**Analiza punktów krytycznych:**

W punkcie  $(0, 0)$ :

$$H(0, 0) = \begin{bmatrix} -6 & 6 \\ 6 & 0 \end{bmatrix}$$

Wyznacznik macierzy Hessego:

$$\det(H(0, 0)) = (-6)(0) - (6)(6) = -36$$

Macierz Hessego jest nieokreślona. Punkt  $(0, 0)$  jest punktem siodłowym.

W punkcie  $(1, 0)$ :

$$H(1, 0) = \begin{bmatrix} 6 & -6 \\ -6 & 12 \end{bmatrix}$$

Kolejne minory wiodące główne macierzy Hessego:

$$\det([6]) = 6$$

$$\det(H(1, 0)) = (6)(12) - (-6)(-6) = 36$$

Oba wyznaczniki są dodatnie, zatem punkt  $(1, 0)$  jest minimum lokalnym.

W punkcie  $(0, -1)$ :

$$H(0, -1) = \begin{bmatrix} 6 & -6 \\ -6 & 0 \end{bmatrix}$$

Wyznacznik macierzy Hessego:

$$\det(H(0, -1)) = (6)(0) - (-6)(-6) = -36$$

Macierz Hessego jest nieokreślona. Punkt  $(0, -1)$  jest punktem siodłowym.

W punkcie  $(-1, -1)$ :

$$H(-1, -1) = \begin{bmatrix} -6 & 6 \\ 6 & -12 \end{bmatrix}$$

Kolejne minory wiodące główne macierzy Hessego:

$$\det([-6]) = -6$$

$$\det(H(-1, -1)) = (-6)(-12) - (6)(6) = 36$$

Pierwszy wyznacznik jest ujemny, a drugi dodatni, zatem punkt  $(-1, -1)$  jest maksimum lokalnym.

Funkcja na krańcach dziedziny ucieka zarówno do  $+\infty$ , jak i  $-\infty$ , więc nie posiada ekstremów globalnych.

**Funkcja**  $f_4(x, y) = (x - y)^4 + x^2 - y^2 - 2x + 2y + 1$

**Pierwsze pochodne cząstkowe:**

$$f_{4x} = \frac{\partial}{\partial x}((x - y)^4 + x^2 - y^2 - 2x + 2y + 1) = 4(x - y)^3 + 2x - 2$$

$$f_{4y} = \frac{\partial}{\partial y}((x - y)^4 + x^2 - y^2 - 2x + 2y + 1) = -4(x - y)^3 - 2y + 2$$

### Wyznaczenie punktów krytycznych:

Rozwiązujemy układ równań:

$$\begin{cases} 4(x-y)^3 + 2x - 2 = 0 \\ -4(x-y)^3 - 2y + 2 = 0 \end{cases}$$

Dodajemy równania stronami, otrzymując:

$$4(x-y)^3 + 2x - 2 - 4(x-y)^3 - 2y + 2 = 0 \implies 2x - 2y = 0 \implies x = y$$

Podstawiając  $y = x$  do pierwszego równania:

$$4(x-x)^3 + 2x - 2 = 0 \implies 2x - 2 = 0 \implies x = 1$$

Zatem  $y = 1$ . Punkt krytyczny to  $(1, 1)$ .

### Drugie pochodne cząstkowe i macierz Hessego:

$$f_{4xx} = \frac{\partial^2}{\partial x^2}((x-y)^4 + x^2 - y^2 - 2x + 2y + 1) = 12(x-y)^2 + 2$$

$$f_{4yy} = \frac{\partial^2}{\partial y^2}((x-y)^4 + x^2 - y^2 - 2x + 2y + 1) = 12(x-y)^2 - 2$$

$$f_{4xy} = f_{4yx} = \frac{\partial^2}{\partial x \partial y}((x-y)^4 + x^2 - y^2 - 2x + 2y + 1) = -12(x-y)^2$$

Macierz Hessego to:

$$H = \begin{bmatrix} 12(x-y)^2 + 2 & -12(x-y)^2 \\ -12(x-y)^2 & 12(x-y)^2 - 2 \end{bmatrix}.$$

### Analiza punktu krytycznego:

W punkcie  $(1, 1)$ :

$$H(1, 1) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

Wyznacznik macierzy Hessego:

$$\det(H(1, 1)) = (2)(-2) - (0)(0) = -4$$

Macierz Hessego jest nieokreślona. Punkt  $(1, 1)$  jest punktem siodłowym.

Podobnie jak i funkcja  $f_1$ , tak i funkcja  $f_4$  nie ma żadnych ekstremów lokalnych, więc nie posiada także ekstremów globalnych.

## 3 Zadanie 2

### 3.1 Opis zadania

Celem zadania jest wyznaczenie najkrótszej ścieżki robota między dwoma ustalonymi punktami  $x^{(0)}$  i  $x^{(n)}$ . Na trasie robota znajdują się przeszkody, których należy unikać. Ścieżka złożona z  $n+1$  punktów jest opisana poprzez macierz  $X \in \mathbb{R}^{(n+1) \times 2}$ . Każdy z tych punktów:  $x^{(0)}, x^{(1)}, \dots, x^{(n)}$  posiada 2 współrzędne, czyli  $x^{(i)} \in \mathbb{R}^2, i = 0, 1, \dots, n$ . Z kolei punkty z przeszkodami, które również mają dwie współrzędne, są przedstawione w macierzy przeszkód  $R \in \mathbb{R}^{k \times 2}$ .

Aby zoptymalizować ścieżkę robota, należy użyć metody największego spadku oraz przyjąć funkcję celu  $F(x^{(0)}, x^{(1)}, \dots, x^{(n)})$  zdefiniowaną następująco:

$$F(x^{(0)}, x^{(1)}, \dots, x^{(n)}) = \lambda_1 \sum_{i=0}^n \sum_{j=1}^k \frac{1}{\epsilon + \|x^{(i)} - r^{(j)}\|_2^2} + \lambda_2 \sum_{i=0}^{n-1} \|x^{(i+1)} - x^{(i)}\|_2^2, \quad (1)$$

gdzie poszczególne symbole określają:

- stała  $\lambda_1$  - wpływ (wagę) składnika zapobiegającego zbyt niemu zbliżaniu się do przeszkody
- stała  $\lambda_2$  - wpływ (wagę) składnika zapobiegającego tworzeniu bardzo długich ścieżek
- $n$  - liczbę odcinków
- $n + 1$  - liczbę punktów na trasie robota
- $k$  - liczbę przeszkód, których robot musi unikać
- $\epsilon$  - pewną (małą) dodatnią liczbę, której dodanie do mianownika zapobiega dzieleniu przez zero

Chcąc zminimalizować funkcję kosztu, sprowadzamy problem nieliniowej optymalizacji z ograniczeniami do problemu nieograniczonej optymalizacji.

W tym zadaniu należy po kolei:

1. Wyprowadzić wyrażenie na gradient  $\nabla F$  funkcji celu  $F$  względem  $x^{(i)}$ :

$$\nabla F = \left[ \frac{\partial F}{\partial x^{(0)}}, \dots, \frac{\partial F}{\partial x^{(n)}} \right],$$

przy czym wzór należy wyrazić poprzez wektory  $x^{(i)}$  i ich składowe, wektory  $r^{(j)}$  i ich składowe oraz niektóre z wielkości wymienionych powyżej.

2. Opisać matematycznie i zaimplementować kroki algorytmu wykorzystującego metodę największego spadku z przeszukiwaniem liniowym (ang. *line search*), który służy do minimalizacji funkcji celu  $F$ . Należy użyć metody złotego podziału (ang. *golden section search*).

Aby formalnie móc skorzystać z tej metody, trzeba założyć, że funkcja  $F$  jest unimodalna, mimo że w rzeczywistości tak nie jest oraz że można ustalić początkowy przedział, w którym znajduje się minimum.

3. Używając tego algorytmu, należy znaleźć najkrótszą ścieżkę robota dla podanych przykładowych wartości parametrów. Aby nie zmieniać położenia początku i końca ścieżki, należy wyzerować gradient funkcji  $F$  względem tych punktów. Ponadto, trzeba przedstawić obliczenia dla 5 losowych inicjalizacji punktów wewnątrz ścieżki  $x^{(1)}, \dots, x^{(n-1)}$  oraz narysować przykładowy wykres wartości funkcji  $F$  w zależności od iteracji.

## Opracowanie zadania

### 3.2 Wyprowadzenie wyrażenia na gradient

Wyprowadzimy wzór na pochodną cząstkową funkcji celu  $F$  względem składowej  $x^{(i)}$ , korzystając ze **wskazówki**:  $\frac{\partial \|x\|^2}{\partial x} = 2x$ . Następnie wyprowadzimy wzór na gradient funkcji celu względem  $x$ .

Zauważmy, że we wzorze na funkcję celu (1) wyrazy występujące w sumowaniu, kiedy liczymy pochodną cząstkową, zerują się poza tym wyrazem, względem którego liczymy pochodną.

Najpierw obliczamy pomocnicze pochodne:

$$\frac{\partial}{\partial x^{(i)}} \left( \frac{1}{\epsilon + \|x^{(i)} - r^{(j)}\|_2^2} \right) = \frac{-2(x^{(i)} - r^{(j)})}{(\epsilon + \|x^{(i)} - r^{(j)}\|_2^2)^2} \quad (2)$$



oraz

$$\frac{\partial}{\partial x^{(i)}} \left( \sum_{i=0}^{n-1} \|x^{(i+1)} - x^{(i)}\|_2^2 \right) = 2(x^{(i)} - x^{(i-1)}) - 2(x^{(i+1)} - x^{(i)}) \quad (3)$$

$$= 2(2x^{(i)} - x^{(i+1)} - x^{(i-1)}) \quad (4)$$

Teraz możemy przejść do wyprowadzenia wzoru na pochodną cząstkową względem zmiennej  $x^{(i)}$ :

$$\frac{\partial F}{\partial x^{(i)}} = -2\lambda_1 \sum_{j=1}^k \frac{x^{(i)} - r^{(j)}}{(\epsilon + \|x^{(i)} - r^{(j)}\|_2^2)^2} \quad (5)$$

$$+ 2\lambda_2(x^{(i)} - x^{(i-1)}) \quad (6)$$

$$- 2\lambda_2(x^{(i+1)} - x^{(i)}) \quad (7)$$

Dla  $i = 0$  nie występuje składnik (6), natomiast dla  $i = n$  nie występuje składnik (7). W pozostałych, pośrednich przypadkach wszystkie składniki we wzorze na pochodną występują. Te szczególne przypadki nie wymagają od nas dodatkowego obsłużenia, gdyż i tak zerujemy gradient względem punktów skrajnych.

Zatem, mając pochodne cząstkowe, gradient jest przedstawiony wyrażeniem:

$$\nabla F = \left[ \frac{\partial F}{\partial x^{(0)}}, \dots, \frac{\partial F}{\partial x^{(n)}} \right]. \quad (8)$$

### 3.3 Opis i implementacja algorytmu

Algorytm minimalizujący funkcję kosztu  $F$  składa się z następujących kroków:

#### 1. Inicjalizacja:

- Ustawienie parametrów: liczba punktów  $n$ , liczba przeszkód  $k$ , pozycje punktów ścieżki: początkowego  $x^{(0)}$  i końcowego  $x^{(n)}$ , współczynniki wagowe  $\lambda_1$  i  $\lambda_2$ , parametr  $\epsilon$  oraz liczba iteracji.
- Wylosowanie początkowych pozycji przeszkód oraz punktów pośrednich ścieżki, tzn. między  $x^{(0)}$  i  $x^{(n)}$ .

#### 2. Funkcja kosztu:

- Zdefiniowanie funkcji kosztu  $F$  zgodnie ze wzorem (1), która składa się z dwóch składników: pierwszego, związanego z odległością od przeszkód oraz drugiego, związanego z długością ścieżki.

#### 3. Gradient funkcji kosztu:

- Wyznaczenie gradientu funkcji kosztu  $\nabla F$ , który jest wektorem pochodnych cząstkowych funkcji  $F$  względem wszystkich punktów ścieżki  $x^{(i)}$ .

#### 4. Algorytm największego spadku:

- Inicjalizacja punktów ścieżki oraz przeszkód.
- Powtarzanie następujących kroków przez określoną liczbę iteracji:
  - (a) Obliczenie gradientu  $\nabla F$  w aktualnym punkcie,
  - (b) Wyznaczenie optymalnej wartości kroku  $\alpha$  za pomocą metody złotego podziału (*golden section search*).
  - (c) Aktualizacja punktów ścieżki:  $X = X + \alpha \cdot \nabla F$  (metoda największego spadku).
  - (d) Sprawdzenie warunku stopu: jeżeli norma gradientu jest mniejsza od  $\epsilon$ , zakończ działanie algorytmu (przed wykonaniem wszystkich zamierzonych iteracji).

Metoda złotego podziału jest jedną z użytecznych metod optymalizacyjnych. Opiera się na wykorzystaniu szczególnych własności funkcji unimodalnej, tzn. *takiej, która na danym przedziale  $[a; b]$  ma tylko jedno minimum*. W tej metodzie przyjmujemy współczynnik  $t \in (0; 1)$ , czyli *współczynnik redukcji* po każdym etapie. Mając początkowy przedział  $[a; b]$  postępujemy w każdym kroku następująco:

1. Obliczamy nową długość przedziałów  $d = t(b - a)$ ,
2. Przyjmujemy punkty  $x_L = b - d, x_R = a + d$ ,
3. Porównujemy wartości funkcji: jeśli  $f(x_L) > f(x_R)$ , to wnioskujemy, że szukane minimum jest w przedziale  $[x_L; b]$ , zatem przyjmujemy dla "nowego" przedziału  $a = x_L$ ; w przeciwnym przypadku, rozumując analogicznie,  $b = x_R$ ,
4. Procedurę powtarzamy aż do osiągnięcia żądanej zbieżności.

Nazwa **złoty podział** wzięła się stąd, iż chcąc zminimalizować liczbę ewaluacji funkcji  $f$ , łatwo pokazać, że  $t = \frac{1}{\phi}$ , gdzie  $\phi$  to złota liczba.

Wspomnijmy także pokrótce o metodzie największego spadku z przeszukiwaniem liniowym. W ogólnym przypadku, mając dane funkcję  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  oraz  $x_k \in \mathbb{R}^n$  i  $\alpha_k \in \mathbb{R}$ , ustalamy wartość w następnym kroku  $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$ .

W przeszukiwaniu liniowym przyjmujemy minimalną wartość prawej strony powyższego równania po wartościach  $\alpha_k$ .

Na potrzeby tego laboratorium użyliśmy funkcji bibliotecznej do metody złotego podziału: `scipy.optimize.golden`. Napisanie własnej wersji tej procedury, choć nie było skomplikowane w teorii, w praktyce sprawiało wiele problemów. Przedstawiamy implementację algorytmu opisanego powyżej.

```
def line_search(X):
    # robimy kopię punktów początkowych
    copyX = copy.deepcopy(X)

    # metoda pomocnicza, którą będziemy przekazywać do metody złotego podziału
    def helper(xs, grad):
        return lambda alpha: cost_function(xs + alpha * grad)

    for i in range(iterations):
        # liczymy gradient
        grad = fun_grad(copyX)

        # znajdujemy współczynnik alpha minimalizujący funkcję kosztu
        alpha = golden(helper(copyX, grad))

        # przesuwamy punkty
        copyX += alpha * grad

        # jeśli kroki stają się zbyt małe - zwracamy aktualną wartość
        # jeszcze przed przebiegiem wszystkich iteracji
        if np.linalg.norm(grad * alpha) < eps:
            break

    return copyX
```

Listing 1: Implementacja algorytmu

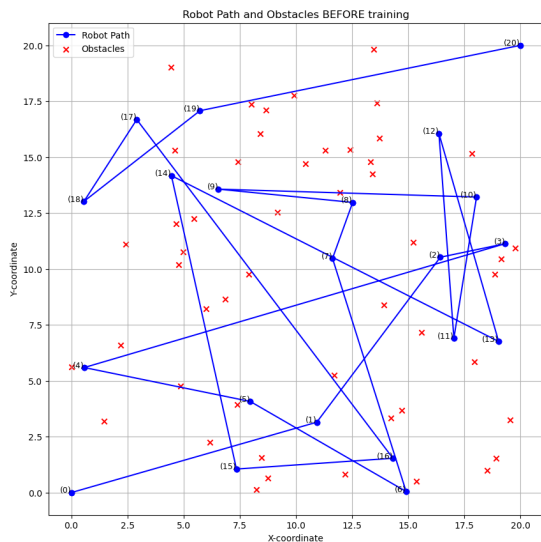
Pojawiające się w kodzie oznaczenia to odpowiednio: *cost\_function*, które odpowiada funkcji kosztu oraz *fun\_grad*, czyli funkcja obliczająca gradient dla podanych wartości  $X$ . Obie te funkcje są zaimplementowane zgodnie ze wzorami omówionymi wcześniej.

### 3.4 Przykładowy wynik działania algorytmu

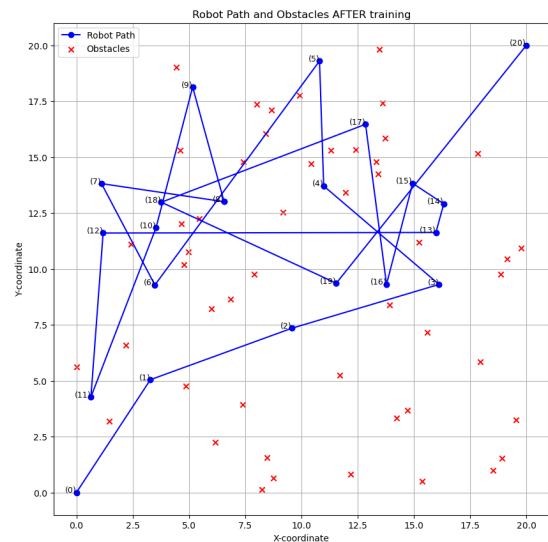
Aby przetestować działanie algorytmu, zgodnie z poleceniem przyjmujemy następujące wartości parametrów:

- $n = 20$  oraz  $k = 50$ ,
- $x^{(0)} = [0, 0]$  oraz  $x^{(n)} = [20, 20]$ ,
- $r^{(i)} \sim \mathcal{U}(0, 20) \times \mathcal{U}(0, 20)$ ,
- $\lambda_1 = \lambda_2 = 1$ ,
- $\epsilon = 10^{-13}$ ,
- liczba iteracji wynosi 400.

Uruchomiliśmy optymalizację na pięciu zbiorach  $x$ -ów, losując za każdym razem nowe punkty startowe (oprócz pierwszego i ostatniego, które są niezmiennie). Poniżej przedstawiamy przykładowe wykresy ścieżek robota przed optymalizacją i po niej - dla jednego z wylosowanych zbiorów punktów startowych.



Rysunek 1: Ścieżka robota PRZED optymalizacją



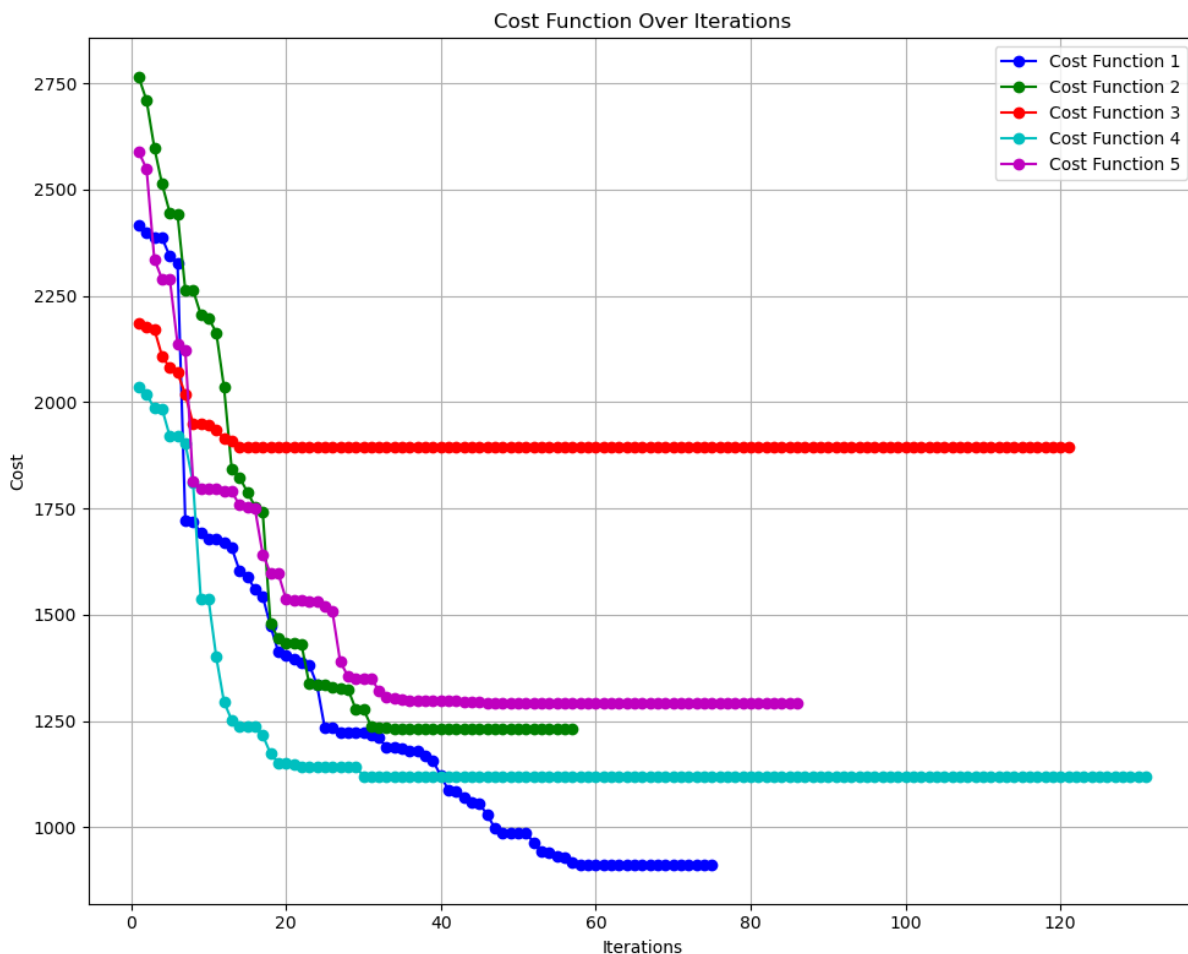
Rysunek 2: Ścieżka robota PO optymalizacji

Jak widzimy, nasz algorytm częściowo poradził sobie z postawionym zadaniem i w pewnym stopniu zoptymalizował ścieżkę. Otrzymana ścieżka jest mniej "zapętlona" w porównaniu do ścieżki początkowej, co było jednym z naszych celów przy pracy nad tym zadaniem. Natomiast oczekiwanym wynikiem byłaby ścieżka idąca mniej więcej wzdłuż prostej poprowadzonej od punktu  $x^{(0)}$  do punktu  $x^{(n)}$ .

Zakładamy, że implementacja biblioteczna metody złotego podziału jest prawidłowa, więc prawdopodobnie błąd może leżeć po naszej stronie, np. wystąpiła pomyłka przy implementacji funkcji kosztu czy wyznaczaniu gradientu. Nie możemy także wykluczyć, że nasza funkcja kosztu "utknęła" w pewnym minimum lokalnym.

Ponadto nie możemy powiedzieć, że optymalizacja ścieżki względem odległości od przeszkód była tak samo skuteczna. Jednym z możliwych powodów może być dosyć duża koncentracja przeszkód na wykresach, co utrudnia oszacowanie odległości "na oko", ale nie patrząc na to, możemy zauważyć, że otrzymana ścieżka w kilku miejscach przechodzi bardzo blisko którejś z przeszkód. Jednym z możliwych sposobów na poprawienie tego parametru ścieżki wynikowej byłoby zwiększenie współczynnika wagowego odpowiadającego za odległości od przeszkód w funkcji kosztu ( $\lambda_1$ ).

Przeanalizujemy również wykres wspólny wszystkich funkcji kosztu dla każdego z wylosowanych zbiorów punktów początkowych.



Rysunek 3: Wykres funkcji kosztu w zależności od kroku optymalizacji

Z powyższego wykresu możemy wywnioskować parę ciekawych własności o charakterze funkcji kosztu w naszym zadaniu:

1. Charakter funkcji kosztu jest mocno uzależniony od konfiguracji startowej. Zmiana pozycji pewnych punktów może spowodować, że nasz algorytm optymalizacyjny zatrzyma się na pewnym minimum lokalnym, co spowoduje, że nie znajdzie minimum globalnego funkcji kosztu (albo innego minimum lokalnego o mniejszej wartości). Oczywiście warto było się spodziewać podobnego zachowania, gdyż nasza funkcja kosztu przyjmuje  $(n+1) \times 2 = 21 \times 2$  współrzędnych punktów ścieżki i  $2 \times k = 2 \times 50$  współrzędnych punktów przeszkód.
2. Również zauważyliśmy, że (przynajmniej dla wylosowanych konfiguracji) początkowa wartość funkcji kosztu nie jest dobrym predyktorem końcowego sukcesu. Przykładowo dla drugiego wylosowanego zbioru wartość funkcji kosztu jest największa ze wszystkich testowanych, ale końcowy wynik optymalizacji jest stosunkowo dobry, w tym przypadku sprowadza koszt do trzeciej najniższej wartości ze wszystkich badanych (czyli w połowie stawki). Odwrotną sytuację widzimy w przypadku zbioru trzeciego, kiedy bardzo niska wartość początkowa funkcji kosztu (w porównaniu do pozostałych przypadków) nie zagwarantowała nam dobrego wyniku końcowego, a nawet na odwrót - uzyskaliśmy ścieżkę optymalizowaną najgorzej ze wszystkich badanych przypadków. Zatem kwestia predykcji wyniku badanego algorytmu dla danej konfiguracji punktów wymagałaby bardziej dogłębnej analizy.

Na koniec przedstawiamy dokładne wartości funkcji kosztu przed optymalizacją i po niej dla każdego z badanych przypadków.

Nr przypadku, $i$	Wartość funkcji kosztu przed optymalizacją	Wartość funkcji kosztu po optymalizacji
1	2415.720	912.582
2	2763.881	1230.866
3	2187.125	1895.791
4	2035.320	1120.319
5	2588.369	1293.408

Tabela 1: Porównanie wartości funkcji kosztu

Wartości na ogół zmniejszały się kilkukrotnie, ale wciąż są bardzo wysokie.

## 4 Podsumowanie i wnioski

W zadaniu pierwszym korzystaliśmy z zagadnień teorii matematyki, aby określić ewentualne minima oraz maksima lokalne funkcji, a następnie zbadać, czy funkcja ma także ekstrema globalne. Obliczaliśmy pochodne cząstkowe pierwszego i drugiego rzędu, aby znaleźć punkty krytyczne oraz zbudować macierz Hessego dla funkcji, na podstawie której mogliśmy już zbadać własności tych punktów.

Celem zadania drugiego było wyznaczenie najkrótszej ścieżki robota między dwoma ustalonymi punktami  $x^{(0)}$  i  $x^{(n)}$ , z uwzględnieniem przeszkód, których należy unikać. Opisaliśmy ścieżkę robota za pomocą macierzy  $X \in \mathbb{R}^{(n+1) \times 2}$ , zawierającej współrzędne punktów ścieżki, a współrzędne przeszkód analogicznie poprzez macierz  $R \in \mathbb{R}^{k \times 2}$ .

Do optymalizacji ścieżki zastosowaliśmy metodę największego spadku, z funkcją kosztu  $F(x^{(0)}, x^{(1)}, \dots, x^{(n)})$  zdefiniowaną jako suma dwóch składników:

1. Składnik związany z odległością od przeszkód, kontrolowany przez parametr  $\lambda_1$ .
2. Składnik związany z długością ścieżki, kontrolowany przez parametr  $\lambda_2$ .

Wyprowadziliśmy wzór na gradient funkcji kosztu, a następnie wykorzystaliśmy go w algorytmie największego spadku z przeszukiwaniem liniowym, używając metody złotego podziału.

Przeprowadziliśmy potem testy algorytmu na pięciu losowych zestawach punktów, a wyniki zaprezentowaliśmy na wykresach, przedstawiających wybraną ścieżkę przed optymalizacją i po niej. Otrzymana ścieżka była mniej zapętlona, ale wciąż daleka od ideału i często przechodziła blisko przeszkód.

Analizując funkcje kosztu dla wszystkich przypadków zauważyliśmy, że jej wartość początkowa nie jest dobrym predyktorem sukcesu końcowego optymalizacji, a charakter funkcji kosztu jest mocno uzależniony od konfiguracji startowej.

Podsumowując, badanemu algorytmowi częściowo udało się zoptymalizować ścieżkę robota, choć dalsze modyfikacje parametrów, chociażby takich jak zwiększenie współczynnika  $\lambda_1$ , mogą poprawić jego efektywność w unikaniu przeszkód.

## Literatura

- [1] Materiały pomocnicze do laboratorium zamieszczone na platformie Teams w katalogu *lab11*.
- [2] Treści przedstawione na wykładzie dotyczące zagadnień optymalizacji.