MOwNiT laboratorium 9. - Sprawozdanie

Michał Szczurek
Informatyka, WIEiT

Grupa poniedziałek 12:50

1 Zadanie 1.

Wyznacz punkty krytyczne każdej z poniższych funkcji. Scharakteryzuj każdy znaleziony punkt jako minimum, maksimum lub punkt siodłowy. Dla każdej funkcji zbadaj, czy posiada minimum globalne lub maksimum globalne na zbiorze $\mathbb{R}^2$.

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

  $\nabla f_1(x,y) = \left[\frac{\partial f_1(x,y)}{\partial x}, \frac{\partial f_1(x,y)}{\partial y}\right] = [2x - 4y, 2y - 4x]$

  Wobec tego punkty krytyczne są rozwiązaniem poniższego układu równań:

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

  Rozwiązanie wygląda następująco:

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

  $$\begin{cases} x = 2y \\ 2y - 8y = 0 \end{cases}$$

  $$\begin{cases} x = 0 \\ y = 0 \end{cases}$$

  $$\nabla^2 f_1(x,y) = \begin{bmatrix} \frac{\partial f_1(x,y)}{\partial x^2} & \frac{\partial f_1(x,y)}{\partial xy} \\ \frac{\partial f_1(x,y)}{\partial yx} & \frac{\partial f_1(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & -4 \\ -4 & 2 \end{bmatrix}$$

  Wobec tego $\det \nabla^2 f_1(0,0) = \begin{vmatrix} 2 & -4 \\ -4 & 2 \end{vmatrix} = -12$

  Ponadto $\frac{\partial f_1(0,0)}{\partial x^2} = 2 > 0$, czyli punkt $(0,0)$ jest punktem siodłowym, jako że $\nabla^2 f_1(0,0)$ jest nieokreślony. $f_1(0,0) = 0$.

Funkcja nie posiada ekstremów lokalnych, a więc nie ma również ekstremów globalnych.

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

$\nabla f_2(x,y) = \left[\frac{\partial f_2(x,y)}{\partial x}, \frac{\partial f_2(x,y)}{\partial y}\right] = \left[4x^3 - 4y, 4y^3 - 4x\right]$

Wobec tego punkty krytyczne są rozwiązaniem poniższego układu równań:

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

Rozwiązanie wygląda następująco:

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

$$\begin{cases} x^3 = y \\ x(x^8 - 1) = 0 \end{cases}$$

$$\begin{cases} x = 0 \\ y = 0 \end{cases} \vee \begin{cases} x = 1 \\ y = 1 \end{cases} \vee \begin{cases} x = -1 \\ y = -1 \end{cases}$$

$$\nabla^2 f_2(x,y) = \begin{bmatrix} \frac{\partial f_2(x,y)}{\partial x^2} & \frac{\partial f_2(x,y)}{\partial xy} \\ \frac{\partial f_2(x,y)}{\partial yx} & \frac{\partial f_2(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{bmatrix}$$

$$\det \nabla^2 f_2(0,0) = \begin{vmatrix} 0 & -4 \\ -4 & 0 \end{vmatrix} = -16$$

$$\det \nabla^2 f_2(1,1) = \begin{vmatrix} 12 & -4 \\ -4 & 12 \end{vmatrix} = 128$$

$$\det \nabla^2 f_2(-1,-1) = \begin{vmatrix} 12 & -4 \\ -4 & 12 \end{vmatrix} = 128$$

jako że $\nabla^2 f_2(0,0)$ jest nieokreślony ($\det < 0$) punkt $(0,0)$ jest punktem siodłowym. $f_3(0,0) = 0$.

$\frac{\partial f_2(1,1)}{\partial x^2} = 12 > 0$, czyli punkt $(1,1)$ jest minimum lokalnym, jako że $\nabla^2 f(1,1)$ jest dodatnio określony. $f_2(1,1) = -2$.

$\frac{\partial f_2(-1,-1)}{\partial x^2} = 12 > 0$,czyli punkt $(-1,-1)$ jest minimum lokalnym, jako że $\nabla^2 f_2(-1,-1)$ jest dodatnio określony. $f_2(-1,-1) = -2$.
Funkcja nie posiada maksimów lokalnych, wobec czego nie posiada również maksimów globalnych. Aby sprawdzić, czy posiada minima globalne

policzę następujące granice:
$\lim_{x\to\pm\infty} f_2(x,y) = \lim_{x\to\pm\infty} x^4(1-\frac{4y}{x^3}+\frac{y^4}{x^4}) = \infty$
$\lim_{y\to\pm\infty} f_2(x,y) = \lim_{y\to\pm\infty} y^4(\frac{x^4}{y^4}+\frac{4x}{y^4}-1) = \infty$

Czyli funkcja nie przyjmuje wartości mniejszych od minimów lokalnych, wobec czego są to również minima globalne.

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

  $\nabla f_3(x,y) = \left[\frac{\partial f_3(x,y)}{\partial x}, \frac{\partial f_3(x,y)}{\partial y}\right] = \left[6x^2 - 6x - 12xy + 6y^2 + 6y, -6x^2 + 12xy + 6x\right]$

  Wobec tego punkty krytyczne są rozwiązaniem poniższego układu równań:

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

  Rozwiązanie wygląda następująco:

  $$\begin{cases} x^2 - x - 2xy + y^2 + y = 0 \\ -x^2 + 2xy + x = 0 \end{cases}$$
  $$\begin{cases} x^2 - x - 2xy + y^2 + y = 0 \\ x(-x + 2y + 1) = 0 \end{cases}$$

  Z drugiego równania widać, że $x = 0 \vee x = 2y + 1$
  Dla x = 0 z pierwszego równania wynika:
  $y(y+1) = 0$
  Dla $x = 2y + 1$ z pierwszego równania wynika:
  $y(y+1) = 0$
  Czyli

  $$\begin{cases} x = 0 \\ y = 0 \end{cases} \vee \begin{cases} x = 0 \\ y = -1 \end{cases} \quad \begin{cases} x = 1 \\ y = 0 \end{cases} \vee \begin{cases} x = -1 \\ y = -1 \end{cases}$$

  $$\nabla^2 f_3(x,y) = \begin{bmatrix} \frac{\partial f_3(x,y)}{\partial x^2} & \frac{\partial f_3(x,y)}{\partial xy} \\ \frac{\partial f_3(x,y)}{\partial yx} & \frac{\partial f_3(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 12x - 6 - 12y & -12x + 12y + 6 \\ -12x + 12y + 6 & 12x \end{bmatrix}$$

  $$\det \nabla^2 f_3(0,0) = \begin{vmatrix} -6 & 6 \\ 6 & 0 \end{vmatrix} = -36$$

  $$\det \nabla^2 f_3(0,-1) = \begin{vmatrix} 6 & -6 \\ -6 & 0 \end{vmatrix} = -36$$

  $$\det \nabla^2 f_3(1,0) = \begin{vmatrix} 6 & -6 \\ -6 & 12 \end{vmatrix} = 36$$

$$\det \nabla^2 f_3(-1,-1) = \begin{vmatrix} -6 & 6 \\ 6 & -12 \end{vmatrix} = 36$$

jako że $\nabla^2 f_3(0,0)$ jest nieokreślony ($\det < 0$) punkt $(0,0)$ jest punktem siodłowym. $f_3(0,0) = 0$.

jako że $\nabla^2 f_3(0,-1)$ jest nieokreślony ($\det < 0$) punkt $(0,-1)$ jest punktem siodłowym. $f_3(0,-1) = 0$.

$\frac{\partial f_3(1,0)}{\partial x^2} = 6 > 0$,czyli punkt $(1,0)$ jest minimum lokalnym, jako że $\nabla^2 f_3(1,0)$ jest dodatnio określony. $f_3(1,0) = -1$.

$\frac{\partial f_3(-1,-1)}{\partial x^2} = -6 < 0$,czyli punkt $(-1,-1)$ jest maksimum, jako że $\nabla^2 f_3(-1,-1)$ jest ujemnie określony. $f_3(-1,-1) = 1$.

Funkcja nie posiada jednak ekstremów globalnych, jako że:
$\lim_{x\to\infty} f_3(x,y) = \lim_{x\to\infty} x^3(2 - \frac{3}{x} - 6y(\frac{1}{x} - \frac{y}{x^2} - \frac{1}{x^2})) = \infty$
$\lim_{x\to-\infty} f_3(x,y) = \lim_{x\to-\infty} x^3(2 - \frac{3}{x} - 6y(\frac{1}{x} - \frac{y}{x^2} - \frac{1}{x^2})) = -\infty$
Czyli funkcja przyjmuje wartości mniejsze od znalezionego minimum lokalnego i większe od maksimum lokalnego. Wobec tego nie posiada ekstremów globalnych.

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

  $\nabla f_4(x,y) = \left[\frac{\partial f_4(x,y)}{\partial x}, \frac{\partial f_4(x,y)}{\partial y}\right] =$
  $[4x^3 - 12x^2y + 12xy^2 - 4y^3 + 2x - 2, -4x^3 + 12x^2y - 12xy^2 + 4y^3 - 2y + 2]$

  Wobec tego punkty krytyczne są rozwiązaniem poniższego układu równań:

  $$\begin{cases} 4x^3 - 12x^2y + 12xy^2 - 4y^3 + 2x - 2 = 0 \\ -4x^3 + 12x^2y - 12xy^2 + 4y^3 - 2y + 2 = 0 \end{cases}$$

  Rozwiązanie wygląda następująco:

  $$\begin{cases} 4x^3 - 12x^2y + 12xy^2 - 4y^3 = 2x - 2 \\ -4x^3 + 12x^2y - 12xy^2 + 4y^3 - 2y + 2 = 0 \end{cases}$$
  $$\begin{cases} -4x^3 + 12x^2y - 12xy^2 + 4y^3 = -2x + 2 \\ -2x + 2 - 2y + 2 = 0 \end{cases}$$

  Z drugiego równania widać, że $x = 2 - y$
  Podstawiając x do 1. równania:
  $-32y^3 + 96y^2 - 98y + 34 = 0$

  Czyli

$$\begin{cases} x = 1 \\ y = 1 \end{cases}$$

$$\nabla^2 f_4(x,y) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x^2} & \frac{\partial f(x,y)}{\partial xy} \\ \frac{\partial f(x,y)}{\partial yx} & \frac{\partial f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 12x^2 - 24xy + 12y^2 + 2 & -12x^2 + 24xy - 12y^2 \\ -12x^2 + 24xy - 12y^2 & 12x^2 - 24xy + 12y^2 - 2 \end{bmatrix}$$

$$\det \nabla^2 f_4(1,1) = \begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix} = -4$$

Jako że $\nabla^2 f_4(x,y)$ jest nieokreślony ($\det < 0$) punkt $(1,1)$ jest punktem siodłowym. $f_4(1,1) = 1$.

Funkcja nie posiada ekstremów lokalnych, wobec czego nie posiada również ekstremów globalnych.

2 Zadanie 2.

Napisz program znajdujący minimum funkcji Rosenbrocka

$$f(x) = 100(x_2 + x_1^2)^2 + (1 - x_1)^2$$

implementując następujące metody optymalizacji:

- metodę największego spadku (ang. steepest descent)
- metodę Newtona

Przetestuj obie metody z następującymi punktami startowymi:

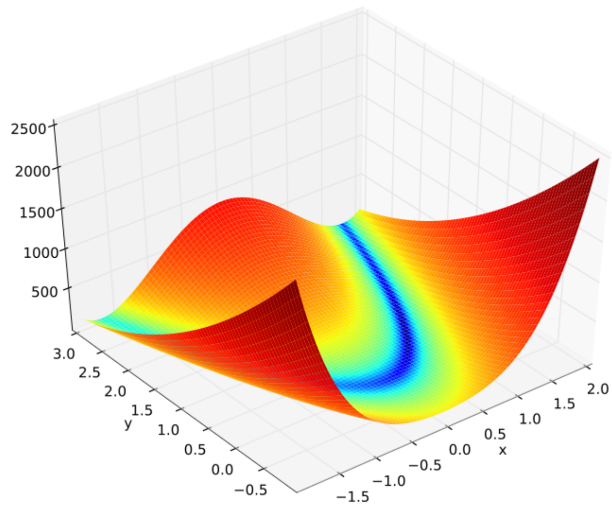
$$x_0 = [-1, 1]^T$$

$$x_0 = [0, 1]^T$$

$$x_0 = [2, 1]^T$$

Każdą metodę wykonaj przez 10 iteracji i porównaj wyniki z wynikami otrzymanymi dla pozostałych punktów startowych. Czy metody zachowują się zgodnie z oczekiwaniami?

Funkcja Rosenbrocka to funkcja niewypukła używana w optymalizacji jako test dla algorytmów optymalizacji. Minimum globalne funkcji znajduje się wewnątrz długiego, parabolicznego wgłębienia funkcji – w punkcie $(x_1, x_2) = (1, 1)$ dla którego funkcja przyjmuje wartość $f(x_1, x_2) = 0$.



Rysunek 1: Funkcja Rosenbrocka. Źródło: Wikipedia.org

Do znalezienia minimum wykorzystano dwie metody:

- Metodę Newtona polegającą na wykorzystaniu w każdej iteracji wzoru:

  $$x_{i+1} = x_i - V \cdot g$$

  Gdzie V to macierz odwrotna do Hesjanu, a g to gradient funkcji. Metoda wymaga, by V była dodatnio określona.

- Metodę największego spadku polegającą na odejmowaniu w każdej iteracji od funkcji jej gradientu przemnożonego przez gdzie α jest taka dobrane, aby:

  $$f(x_k - \alpha_k g_k) = min_{\alpha>0} f(x_k - \alpha_k g_k)$$

  W tym znalezienia α spełniającego powyższy warunek użyłem metody złotego przedziału z przedziałem początkowym $(0, 4)$ z tolerancją 10^{-9}.

Poniżej znajdują się przeprowadzone testy i wyniki:

- Dla punktu początkowego $x_0 = [-1, 1]^T$
  Metoda największego spadku (Minimum po i-tej iteracji):

```
1 : [1. 1.]
2 : [1. 1.]
3 : [1. 1.]
4 : [1. 1.]
5 : [1. 1.]
6 : [1. 1.]
```

```
7 : [1. 1.]
8 : [1. 1.]
9 : [1. 1.]
10 : [1. 1.]
```

Metoda Newtona

```
1 : [ 1. -3.]
2 : [1. 1.]
3 : [1. 1.]
4 : [1. 1.]
5 : [1. 1.]
6 : [1. 1.]
7 : [1. 1.]
8 : [1. 1.]
9 : [1. 1.]
10 : [1. 1.]
```

Obie metody po stosunkowo małej liczbie iteracji dały poprawny wynik.

- Dla punktu początkowego $x_0 = [0, 1]^T$
  Metoda największego spadku (Minimum po i-tej iteracji):

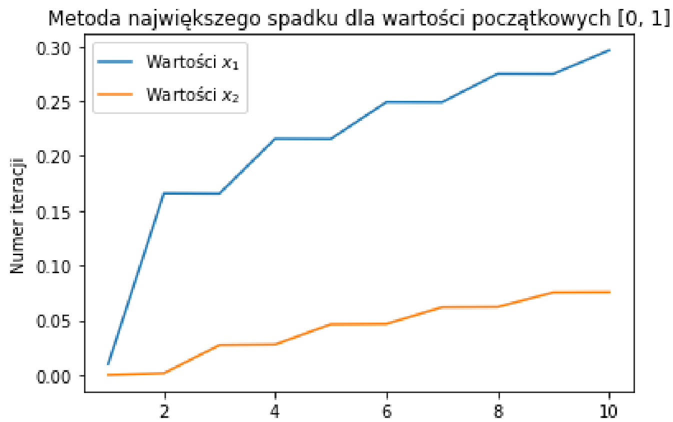
```
1 : [9.99999378e-03 6.22167055e-07]
2 : [0.16583412 0.00156522]
3 : [0.16557542 0.02733171]
4 : [0.21555392 0.02783354]
5 : [0.21536845 0.04630513]
6 : [0.24905667 0.04664336]
7 : [0.24890372 0.06187803]
8 : [0.27490455 0.06213907]
9 : [0.27477114 0.07542677]
10 : [0.29618873 0.0756418 ]
```

Metoda Newtona

```
1 : [-0.00502513  0.        ]
2 : [ 0.99494975 -0.01002475]
3 : [0.99497488 0.989975  ]
4 : [1.         0.99997475]
5 : [1. 1.]
6 : [1. 1.]
```

```
7 : [1. 1.]
8 : [1. 1.]
9 : [1. 1.]
10 : [1. 1.]
```

Metoda Newtona wymagała nieco więcej iteracji niż poprzednio by dać poprawny Wynik. Metoda największego spadku dała z kolei wynik dość odległy od oczekiwanego. Dla tej metody sporządzono wykres: Wykres



wskazuje na to, że wartości zbiegają do oczekiwanych, jednak zdecydowanie wolniej niż miało to miejsce w przypadku drugiej metody.

- Dla punktu początkowego $x_0 = [2, 1]^T$
  Metoda największego spadku (Minimum po i-tej iteracji):

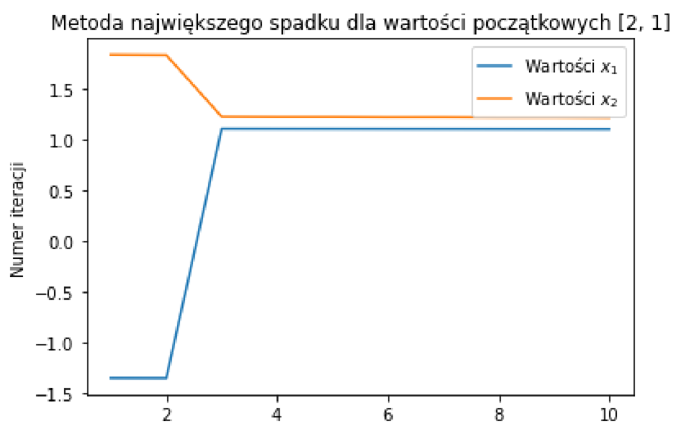
```
1 : [-1.35192908  1.83728453]
2 : [-1.35283053  1.83363257]
3 : [1.10722931 1.2263924 ]
4 : [1.1066757  1.22415029]
5 : [1.10626394 1.22425196]
6 : [1.10571601 1.22203289]
7 : [1.10530773 1.2221337 ]
8 : [1.10476549 1.21993757]
9 : [1.10436069 1.22003752]
10 : [1.10382402 1.21786395]
```

Metoda Newtona

```
1 : [1.99833611 3.99334443]
```

```
2 : [1.00055248 0.0055331 ]
3 : [1.00054972 1.00109974]
4 : [1.         0.9999997]
5 : [1. 1.]
6 : [1. 1.]
7 : [1. 1.]
8 : [1. 1.]
9 : [1. 1.]
10 : [1. 1.]
```

Podobnie jak w poprzednim przykładzie - metoda Newtona dała poprawny wynik, a metoda największego spadku od pewnej iteracji(3), daje wyniki zbiegające do wartości oczekiwanej bardzo wolno. Obrazuje to poniższy wykres. Wykres wskazuje na to, że wartości zbiegają do oczekiwanych,

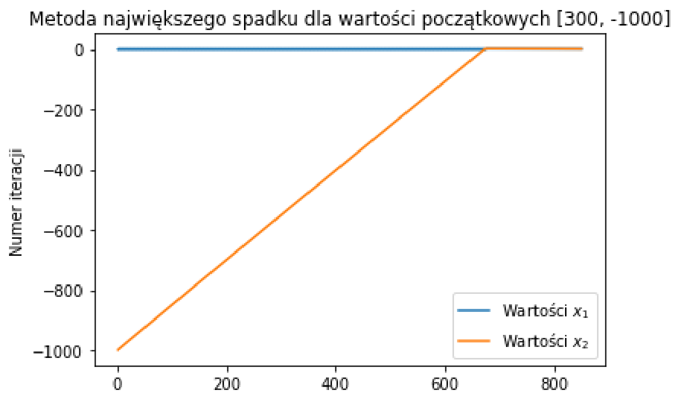


jednak zdecydowanie wolniej niż miało to miejsce w przypadku drugiej metody.

Postanowiłem przeprowadzić również testy dla innych, losowych wartości. Poniżej znajdują się wyniki uzyskane po 10 iteracjach

- Dla punktu początkowego $x_0 = [3, 4]^T$
  Wynik uzyskany metodą największego spadku: $[-2.19, 4.78]$
  Wynik uzyskany metodą Newtona spadku: $[1, 1]$
- Dla punktu początkowego $x_0 = [-4, 5]^T$
  Wynik uzyskany metodą największego spadku: $[-1.91, 3.63]$
  Wynik uzyskany metodą Newtona $[1, 1]$
- Dla punktu początkowego $x_0 = [3, 4]^T$
  Wynik uzyskany metodą największego spadku: $[2.93, 8.60]$
  Wynik uzyskany metodą Newtona spadku: $[1, 1]$

Jak widać metoda Newtona zawsze dawała poprawny wynik, czego nie można powiedzieć o metodzie największego spadku. Postanowiłem wykonać jeszcze jeden test, by zobaczyć jak metody radzą sobie z wyjątkowo złym punktem startowym. Za punkt początkowy przyjąłem $x_0 = [300, -1000]^T$.
Wynik uzyskany metodą największego spadku to $[0.976, -984]$.
Wynik uzyskany metodą Newtona spadku to $[1, 1]$.
Metoda Newtona poradziła sobie również z tak odległym punktem początkowym. Postanowiłem sprawdzić, eksperymentalnie, czy dla pewnego n metoda największego spadku da wynik dokładny. Odpowiedź na to pytanie jest twierdząca. Dla $n = 850$ metoda dała wynik $[1.00076781, 1.0016137]$, a dla $n = 890$ $[1, 1]$. Poniżej znajduje się wykres wartości dla iteracji z przedziału 1 do 850. Jak widać wartość x_1 została już dość dobrze wyznaczona po 1 iteracji. Następ-



nie przez około 650 iteracji x_2 zbliża się do wyniku w sposób liniowy. Potem następuje wolniejsze zbieganie obu wartości do $[1, 1]$.

W przeprowadzonych przeze mnie testach jednoznacznie lepsza okazała się być metoda Newtona.