

Algorytmy Geometryczne Lab1

Spis treści

1. Dane Podstawowe	1
2. Dane Techniczne	1
3. Realizacja ćwiczenia	2
3.1 Generowane zbiory punktów	2
3.2 Użyte wyznaczniki	2
3.3 Przyjęte progi tolerancji	2
3.4 Wizualizacja sklasyfikowanych punktów	2
4. Wyniki i analiza	3
4.1 Zbiór A	3
4.2 Zbiór B	4
4.3 Zbiór C	5
4.4 Zbiór D	6
4.5 float 32	8
5. Wnioski	9

1. Dane Podstawowe

Witold Nieć

<removed for privacy>

2. Dane Techniczne

Język: Python 3.12.0

Procesor: Intel core i7-8750H

System operacyjny: Windows 10

Precyzja: float64 oraz float32

Środowisko: Jupyter Notebook

3. Realizacja ćwiczenia

Celem ćwiczenia jest wizualizacja i klasyfikacja generowanych punktów w stosunku do prostej przechodzącej przez punkty $A(-1.0, 0.0)$ i $B(1.0, 0.1)$.

3.1 Generowane zbiory punktów

Punkty generowane były w 4 zbiorach:

- Zbiór A – 10^5 losowych punktów o współrzędnych z przedziału $[-1000, 1000]$,
- Zbiór B – 10^5 losowych punktów o współrzędnych z przedziału $[-1014, 1014]$,
- Zbiór C – 1000 losowych punktów leżących na okręgu o środku $(0,0)$ i promieniu $R=100$,
- Zbiór D – 1000 losowych punktów o współrzędnych z przedziału $[-1000, 1000]$ leżących na prostej wyznaczonej przez wektor (a, b) , przyjmij $a = [-1.0, 0.0]$, $b = [1.0, 0.1]$.

3.2 Użyte wyznaczniki

Wygenerowane punkty klasyfikowane były za pomocą 4 wyznaczników:

- 2×2 – wyznacznik 2×2 implementacji własnej,
- $2 \times 2_lib$ – wyznacznik 2×2 z biblioteki numpy,
- 3×3 – wyznacznik 3×3 implementacji własnej,
- $3 \times 3_lib$ – wyznacznik 3×3 z biblioteki numpy.

3.3 Przyjęte progi tolerancji

Przyjmowane jest 5 progów tolerancji (ε).

- $\varepsilon = 0$
- $\varepsilon = 10^{-14}$
- $\varepsilon = 10^{-12}$
- $\varepsilon = 10^{-10}$
- $\varepsilon = 10^{-8}$

3.4 Wizualizacja sklasyfikowanych punktów

Wyznacznik liczony był kolejno z dwóch punktów leżących na prostej oraz z punktu który chcemy sklasyfikować.

Poniżej przedstawiono sposób w jaki dany punkt jest klasyfikowany za pomocą wyznacznika oraz przyjęty kolor jakim dana grupa punktów będzie oznaczana.

Jeśli wyznacznik wychodzi ujemny \rightarrow punkt leży na lewo od prostej - **kolor niebieski**

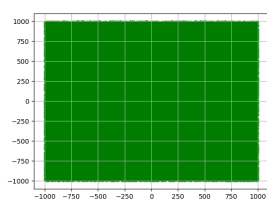
Jeśli wyznacznik dodatni \rightarrow punkt leży na prawo od prostej - **kolor zielony**

Jeśli wyznacznik = 0 \rightarrow punkt leży na prostej - **kolor czerwony**

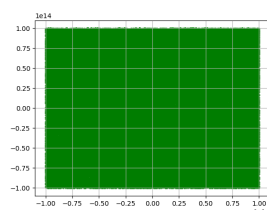
Niekiedy na wykresach oprócz punktów będzie się pojawiać również prosta AB.

Będzie ona oznaczana **kolorem fioletowym**.

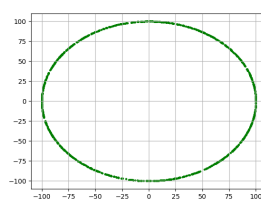
Przyjęty kierunek prostej to: $A \rightarrow B$



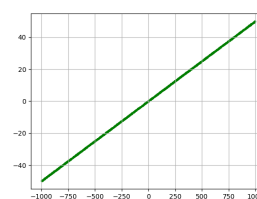
Rys. 0.1: zbiór A



Rys. 0.2: zbiór B



Rys. 0.3: zbiór C



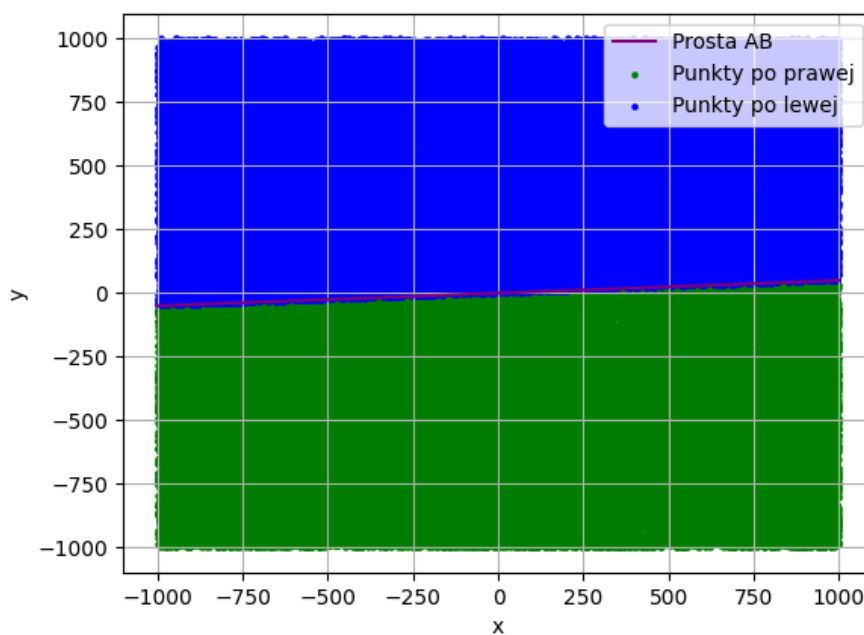
Rys. 0.4: zbiór D

4. Wyniki i analiza

4.1. Zbiór A

Tabela 1: Wyniki klasyfikacji punktów dla $\varepsilon = 10^{-8}, 10^{-10}, 10^{-12}, 10^{-14}, 0$

Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	48786	0	50214
2x2_lib	48786	0	50214
3x3	48786	0	50214
3x3_lib	48786	0	50214



Rys. 1: Klasyfikacja punktów dla zbioru A

Klasyfikacja wygenerowanych punktów dla zbioru A nie przyniosła zaskoczeń. Przy użyciu każdego wyznacznika oraz tolerancji ε , uzyskane wyniki były identyczne (Tabela 1). Prawdopodobieństwo wygenerowania pseudolosowo punktu na prostej jest bardzo małe, dlatego nie powinien dziwić fakt że nie udało się uzyskać takich punktów. Punkty zgodnie z oczekiwaniami podzieliły się w niemal równym rozkładzie po obu stronach prostej. Jak widać na Rysunku 1, algorytm klasyfikacji działa poprawnie i zgodnie z przyjętymi założeniami odnośnie wizualizacji podziału punktów.

4.2 Zbiór B

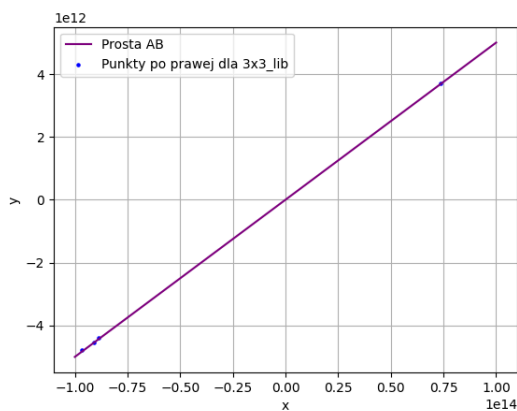
Tabela 2: Wyniki klasyfikacji punktów dla $\varepsilon = 0$

Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	49920	0	50080
2x2_lib	49920	0	50080
3x3	49916	0	50084
3x3_lib	49916	0	50084

Tabela 3: Wyniki klasyfikacji punktów dla $\varepsilon = 10^{-8}, 10^{-10}, 10^{-12}, 10^{-14}$

Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	49914	6	50080
2x2_lib	49914	6	50080
3x3	49916	0	50084
3x3_lib	49916	0	50084

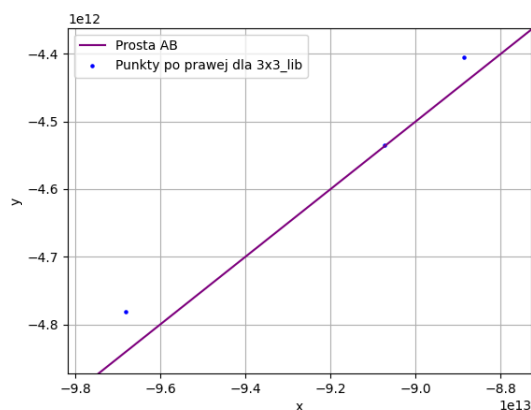
Ze względu na pewne podobieństwo zbioru B do zbioru A można by oczekiwać analogii wyników. Sytuacja jednak okazuje się bardziej złożona. Uzyskano różne wyniki dla wyznaczników macierzy 3x3 i 2x2. Dodatkowo wyniki dla wyznaczników 2x2 różnią się w zależności od wartości ε . Dla $\varepsilon = 10^{-8}, 10^{-10}, 10^{-12}, 10^{-14}$ uzyskano punkty na prostej. (Tabela 2, Tabela 3)



Rys. 2: Punkty różnie zakwalifikowane przez wyznaczniki: $\det_{2 \times 2_lib}$, $\det_{3 \times 3_lib}$ dla $\varepsilon = 0$

Rys. 3 dobrze obrazuje wyższość wyznacznika 3x3_lib nad wyznacznikiem 2x2_lib. W sytuacjach, gdy punkty są różnie klasyfikowane przez te wyznaczniki, poprawną klasyfikację zapewnia wyznacznik 3x3_lib. Dodatkowo przy wartości $\varepsilon > 0$, wyznaczniki 2x2 przyjmują punkty na prostej. Generowanie takich punktów dla zbioru B jest niemal niemożliwe, co świadczy o niskiej dokładności numerycznej tych wyznaczników.

Przy każdej wartości ε , wyniki dla wyznaczników bibliotecznych oraz ich samodzielnie zaimplementowanych odpowiedników były identyczne. Różnice wystąpiły jedynie pomiędzy wyznacznikami macierzy różnych wymiarów. Na Rys. 2 przedstawiono punkty różnie zakwalifikowane dla wyznaczników 2x2_lib oraz 3x3_lib. Liczność zbioru tych punktów to 4.



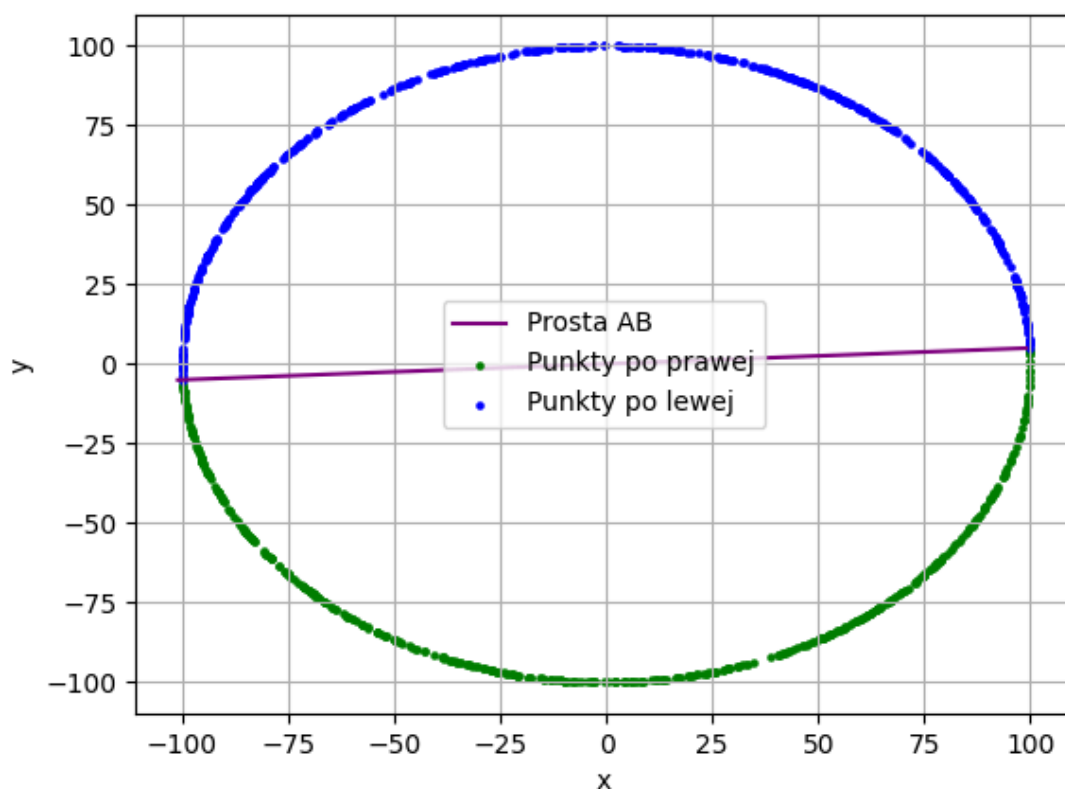
Rys. 3: Powiększony fragment Rys. 2

4.3 Zbiór C

Tabela 4: Wyniki klasyfikacji punktów dla $\varepsilon = 10^{-8}, 10^{-10}, 10^{-12}, 10^{-14}, 0$

Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	497	0	503
2x2_lib	497	0	503
3x3	497	0	503
3x3_lib	497	0	503

Dla losowo generowanego punktu równomiernie rozłożonego na obwodzie okręgu, prawdopodobieństwo, że znajdzie się on dokładnie na prostej, jest praktycznie zerowe. Wyniki w Tabeli 4 potwierdzają to zjawisko – żaden z wygenerowanych punktów nie leży na prostej, a podział na punkty po lewej i prawej stronie prostej jest niemal równomierny. Wyniki są identyczne dla każdego wyznacznika oraz każdej rozważanej wartości ε .



Rys. 4: Wizualizacja wyników klasyfikacji punktów dla zbioru C

Na Rys. 4 doskonale widać że przy odpowiednio małej wartości ε część wspólna okręgu i prostej redukuje się w zasadzie do dwóch punktów. Zatem szansa na trafienie na jeden z nich jest niezwykle niska.

4.4 Zbiór D

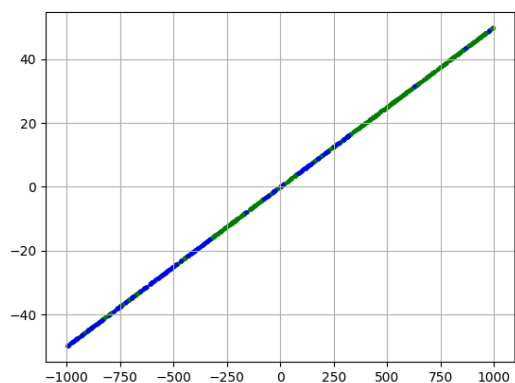
Prosta generowana z równania parametrycznego.

Tabela 5: Wyniki klasyfikacji punktów dla zbioru D

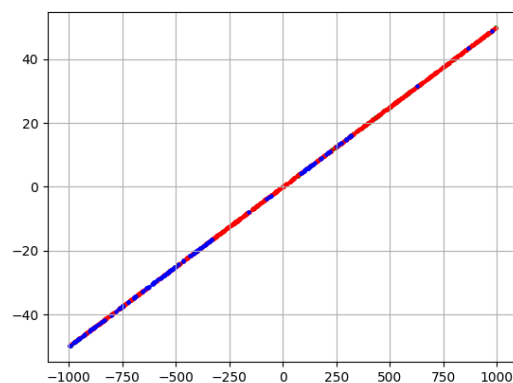
$\varepsilon = 0$			
Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	862	0	138
2x2_lib	848	0	152
3x3	681	0	319
3x3_lib	612	0	388
$\varepsilon = 10^{-14}$			
Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	114	755	131
2x2_lib	124	729	147
3x3	56	944	0
3x3_lib	130	829	41
$\varepsilon = 10^{-12}$			
Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	71	857	72
2x2_lib	86	804	110
3x3	0	1000	0
3x3_lib	0	1000	0
$\varepsilon = 10^{-10}, 10^{-8}$ - (taki sam wynik dla obu wartości ε)			
Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	0	1000	0
2x2_lib	0	1000	0
3x3	0	1000	0
3x3_lib	0	1000	0

Można by przypuszczać, że generując punkty na prostej, każdy wyznacznik z dowolną dokładnością powinien jednoznacznie je sklasyfikować. Okazuje się natomiast że wyniki prezentują zupełnie inne realia. Przy wartości $\varepsilon = 0$ żaden punkt nie zostaje wykryty jako „na prostej”. Dopiero przy dokładności $\varepsilon = 10^{-10}$ udaje się uzyskać zadowalający rezultat - wykrycia wszystkich punktów przez każdy wyznacznik. Warto zauważyć również, że po raz kolejny wyniki dowodzą przewagę wyznaczników trójwymiarowych które osiągnęły oczekiwany rezultat już przy dokładności $\varepsilon = 10^{-12}$.

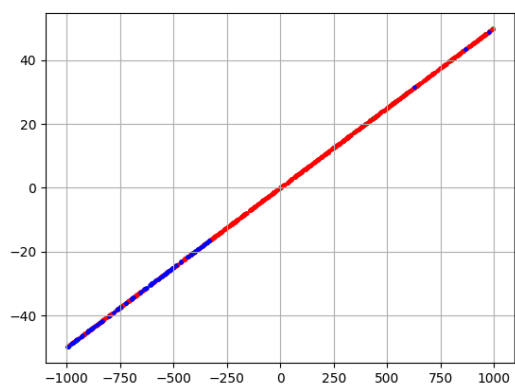
Poniżej przedstawiono wykresy dla klasyfikacji zbioru D za pomocą wyznacznika $2 \times 2_{lib}$.



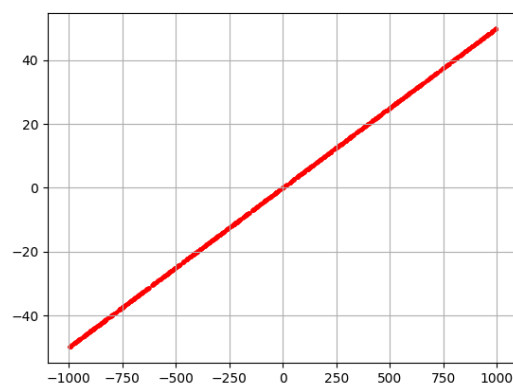
Rys. 5: Klasyfikacja zbioru D dla $\varepsilon = 0$



Rys. 6: Klasyfikacja zbioru D dla $\varepsilon = 10^{-14}$



Rys. 7: Klasyfikacja zbioru D dla $\varepsilon = 10^{-12}$



Rys. 8: Klasyfikacja zbioru D dla $\varepsilon = 10^{-10}, 10^{-8}$

Jak pokazano na Rys. 5-8, wraz ze wzrostem wartości ε , rośnie liczba punktów zaliczanych jako leżących na prostej. Początkowo nie sklasyfikowano w ten sposób żadnych punktów, lecz po uzyskaniu odpowiednio dużego ε , wyznacznik osiąga pełną skuteczność i prawidłowo klasyfikuje wszystkie punkty.

4.5 float 32

Zbiór D jest przekonwertowany z formatu float64 na format float32

Tabela 6: Wyniki klasyfikacji punktów dla zbioru D w formacie float32

$\varepsilon = 0$, float32			
Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	573	0	427
2x2_lib	547	0	453
3x3	522	0	478
3x3_lib	520	0	480
$\varepsilon = 10^{-14}$, float32			
Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	415	158	427
2x2_lib	437	110	453
3x3	410	171	419
3x3_lib	418	155	427
$\varepsilon = 10^{-12}$, float32			
Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	410	171	419
2x2_lib	429	128	443
3x3	410	171	419
3x3_lib	410	171	419
$\varepsilon = 10^{-10}, 10^{-8}$ - (taki sam wynik dla obu wartości ε), float32			
Wyznacznik	Liczba punktów po lewej	Liczba punktów na prostej	Liczba punktów po prawej
2x2	410	171	419
2x2_lib	410	171	419
3x3	410	171	419
3x3_lib	410	171	419

W przeciwieństwie do zbioru D dla formatu float64, w tym przypadku dla przyjętych wartości ε nie udaje się uzyskać bezbłędnej klasyfikacji punktów. Przez konwersję tracimy część informacji o liczbie rzeczywistej. Skutkiem tego jest dalsza utrata dokładności w reprezentacji punktów. Być może dla większych wartości ε , udało by się uzyskać oczekiwane wyniki.

5. Wnioski

Ćwiczenie to świetnie prezentuje iluzoryczną dokładność reprezentacji liczb rzeczywistych w komputerze. Na podstawie przeprowadzonych analiz mogę z przekonaniem stwierdzić, że sposób wyliczania wyznacznika z macierzy trójwymiarowej stanowił znacznie dokładniejszy sposób klasyfikacji punktów jednak nadal nie w pełni idealny.

Na podstawie wyników dla zbiorów A i C można by wnioskować, że precyzja obliczeń ani dobór wyznacznika nie mają znaczenia, ponieważ wyniki są identyczne. Jednak analiza pozostałych zbiorów ujawnia trudności, jakie komputer napotyka przy przetwarzaniu liczb zmiennoprzecinkowych. Szczególnie zaskakujące okazały się wyniki dla zbioru D mimo że punkty zostały wygenerowane jako leżące na prostej, wyznacznik nie klasyfikował ich jako należących do tej prostej. To zjawisko wyraźnie ilustruje ograniczenia wynikające z binarnej reprezentacji liczb, które wpływają na dokładność obliczeń numerycznych.