Predykaty Geometryczne

1. Dane podstawowe

Franciszek Kuś, 07.10.2024, Predykaty geometryczne ćwiczenie 1.

2. Dane techniczne

Język Programowania	Python 3.11
Procesor	Intel Core i5-6300U 2.40GHz 4 rdzenie
RAM	16GB, 2133MHz
System operacyjny	Windows 11 22H2
Precyzja obliczeń	64-bit floating point

3. Realizacja ćwiczenia

Wygenerowane zostały 4 zbiory o współrzędnych z przedziału: $[-1000, 1000]$ (A),$[-10^{14}, 10^{14}]$ (B), na kole o środku w punkcie $(0, 0)$ i promieniu $R = 100$ (C), $[-1000, 1000]$ leżących na prostej przechodzącej przez punkty $a, b = (-1, 0) \ \text{ i } \ (1, 0.1)$ (D).

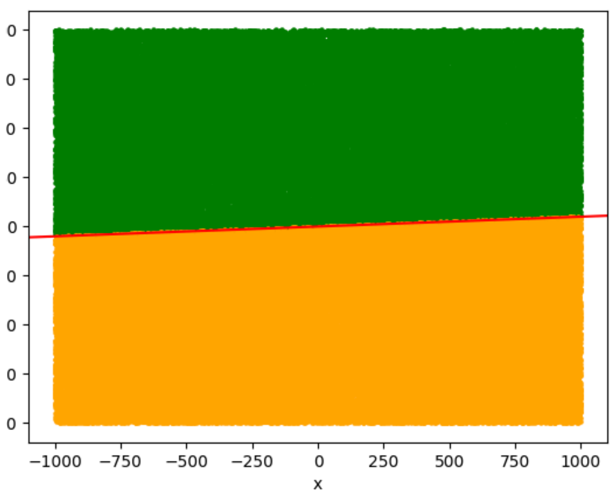
Prosta została wyznaczona przez punkty $(-1, 0) \text{ i } \ (1, 0.1)$. Pozycja punktu była względem linii przez wyliczenie wyznacznika macierzy: $\begin{pmatrix} a_x & a_y & 1 \\ b_x & b_y & 1 \\ c_x & c_y & 1 \end{pmatrix}, \begin{pmatrix} a_x - c_x & a_y - c_y \\ b_x - c_x & b_y - c_y \end{pmatrix}$ gdzie a, b to punkty wyznaczające prostą i c to punkt który sprawdzamy względem linii. Wyznacznik macierzy 2 na 2 był liczona standardowym wzorem, a 3 na 3 metodą Sarrusa. Wyznaczniki były również liczone za pomocą funkcji bibliotecznej `linalg.det` z biblioteki `numpy`. Tolerancje liczenia wyznacznika to: $\{0, 10^{-14}, 10^{-12}, 10^{-10}, 10^{-8}\}$.

4. Wyniki

W tej części gdy pokazane jest *2x2 numpy* lub *3x3 numpy* wyznacznik macierzy jest liczony przez funkcje biblioteczną. Inaczej jest on liczony wzorami bez skorzystania z biblioteki.

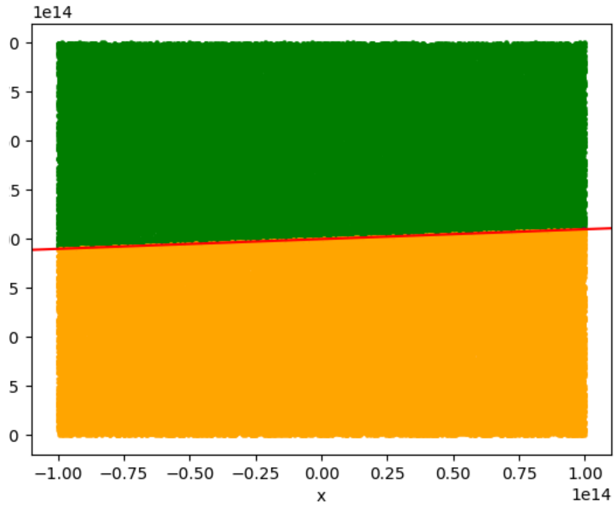
4.1 Zbiór punktów A

Dla zbioru punktów ze współrzędnymi miedzy $[-1000, 1000]$ nie zostało zaobserwowane żadne wyjątkowe zachowanie. Żaden punkt nie został zaklasyfikowany jako będący na linii. Nie było też żadnych różnic w ilości punktów określonych jako na lewo i prawo od linii miedzy różnymi metodami.
49997 punktów było na lewo od linii, 50003 na prawo od linii.



4.2 Zbiór punktów B

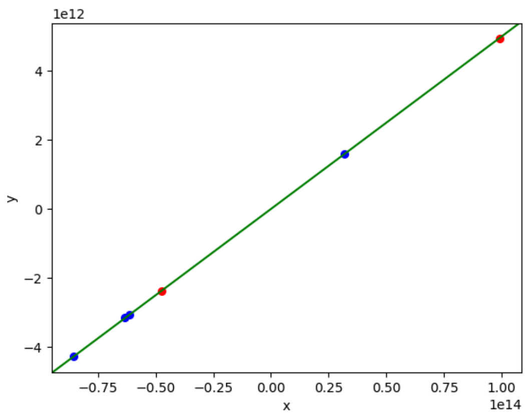
Zbiór punktów ze współrzędnymi od $[-10^{14}, 10^{14}]$ posiadał kilka punktów które zostały zakwalifikowywane jako będące na linii jednak były różnice miedzy różnymi metodami liczenia wyznacznika. Nie było różnic miedzy wartościami ε. Tabela 1 wskazuje ile punktów było zakwalifikowanej do której kategorii.



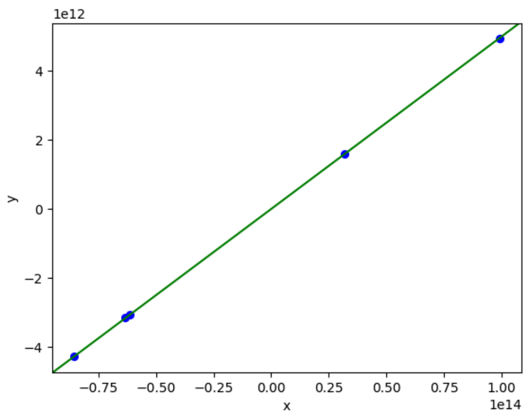
	Lewo	Linia	Prawo
2x2 wzór	49940	5	50055
2x2 `numpy`	49944	2	50054
3x3 wzór 3x3 `numpy`	49944	0	50056

Tabela 1: Wyniki dla zbioru B

Obie metody które wskazały że istnieją punkty które leżą na linii wyznaczały wyznacznik macierzy 2x2. Gdy wyznacznik był wyliczany metodą posługującą się macierzą 3x3 żaden punkt nie był wskazany jako będący na linii.
Poniższe wykresy pokazują różnice między metodami wyznaczania wyznacznika.



Wykres 1: Punkty zakwalifikowane przez *2x2 wzór* (niebieskie), a nie *2x2 numpy* (czerwone).

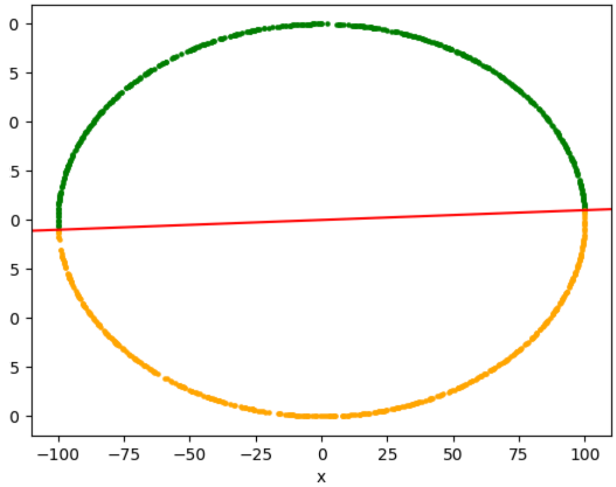


Wykres 2: Punkty zakwalifikowane przez *2x2 wzór*, a nie przez metody używające macierzy 3x3.

Jak widać na wykresach 1 i 2 te punkty znajdowały się dość daleko od $(0, 0)$ więc ilość bitów składających się na mantysę była mała. mogło to wpłynąć na precyzję obliczeń i przez to metoda która wymaga najmniej mnożenia zakwalifikowała je.

4.3 Zbiór punktów C

Zbiór punktów będący kołem o środku w punkcie $(0, 0)$ i średnicy 100 zachował się podobnie do zbioru A. Nie było żadnych punktów na linii dla każdego ε i sposobu wyznaczania wyznacznika. 523 punkty były na lewo od linii, a 477 na prawo. Można jednak zauważyć na wykresie 3 że ten zbiór nie wygenerował się równomiernie tylko widoczne są miejsca bez punktów. Moim zdaniem jest to efekt używania funkcji `pi, sin, cos` z biblioteki `math`. Ograniczona precyzja tych funkcji i tych liczb mogła spowodować te miejsca.



Wykres 3

4.4 Zbiór punktów D

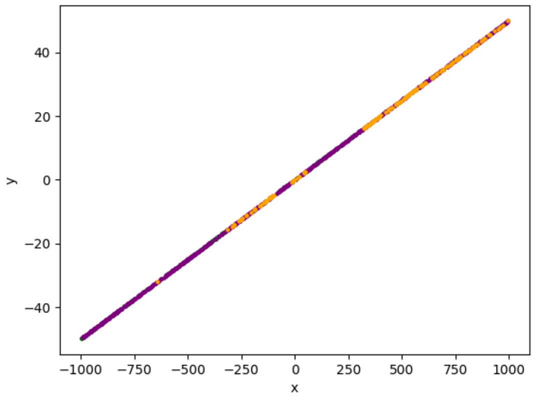
Zbiór punktów na prostej zawierającej $(-1, 0)$ i $(1, 0.1)$ okazał się wyjątkową różnorodnością między wartościami ε i sposobami liczenia wyznacznika.

ε	2x2 wzór			2x2 numpy			3x3 wzór			3x3 numpy		
	Lewo	Linia	Prawo	Lewo	Linia	Prawo	Lewo	Linia	Prawo	Lewo	Linia	Prawo
0	145	701	154	158	668	174	173	399	428	346	300	354
10^{-14}	139	713	148	150	685	165	0	1000	0	16	863	121
10^{-12}	82	831	87	114	770	116	0	1000	0	0	1000	0
10^{-10}	0	1000	0	0	1000	0	0	1000	0	0	1000	0

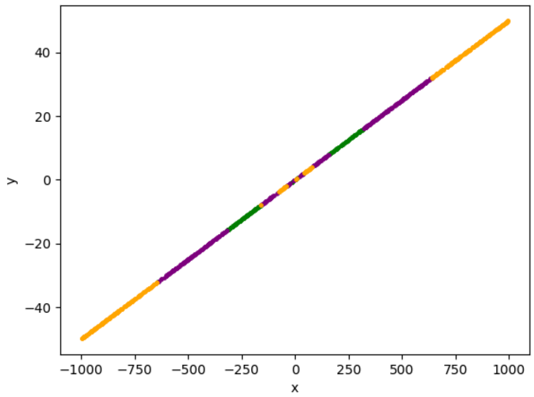
Tabela 2: Wartości dla $\varepsilon = 10^{-8}$ zostały pominięte gdyż są identyczne do wartości dla $\varepsilon = 10^{-10}$.

Można zauważyć w tabeli 2 że metody korzystające z metody wyznaczania wyznacznika macierzy 2x2 przy $\varepsilon = 0$ poprawnie kategoryzowało większość punktów, a metody używające macierzy 3x3 wyjątkowo mało. Jednak metody korzystające z wyznacznika 3x3 przy $\varepsilon > 0$ wykazały się bardzo szybkim wzrostem dokładności a te korzystające z 2x2 pozostały w tyle.

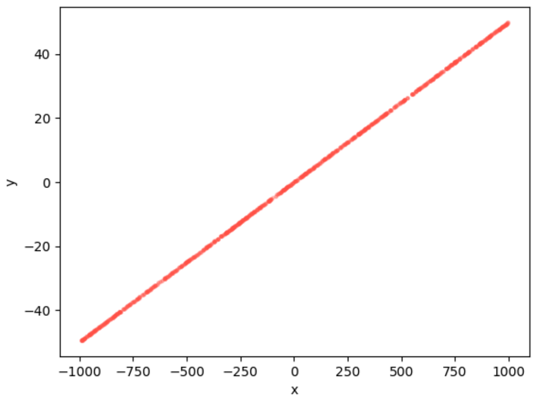
Funkcje biblioteczne wykazały się też gorszą dokładnością niż funkcje własne. Moim zdaniem jest to spowodowane tym że funkcje z biblioteki są przygotowane do liczenia wyznaczników dla bardzo dużych macierzy oraz dla bardzo dużych liczb. Z tego powodu wypadają gorzej w tym konkretnym zastosowaniu.



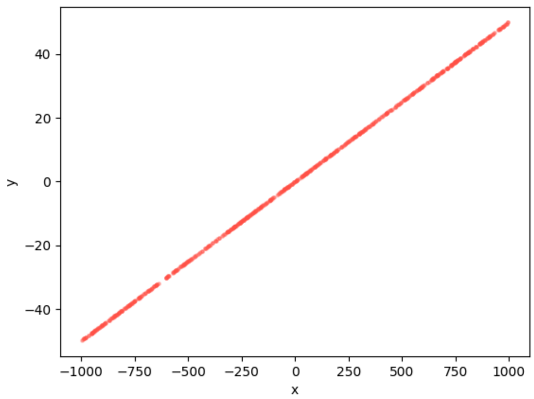
Wykres 4: 2x2 wzór dla $\varepsilon = 0$



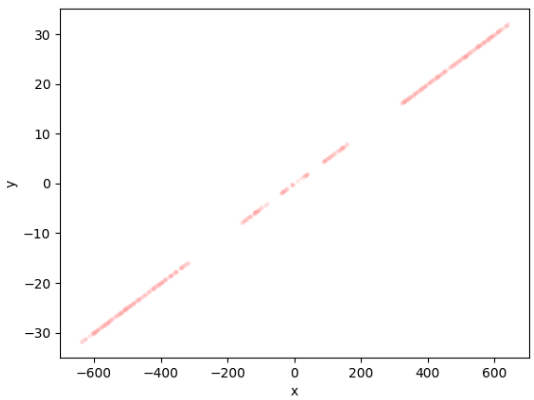
Wykres 5: 3x3 wzór dla $\varepsilon = 0$



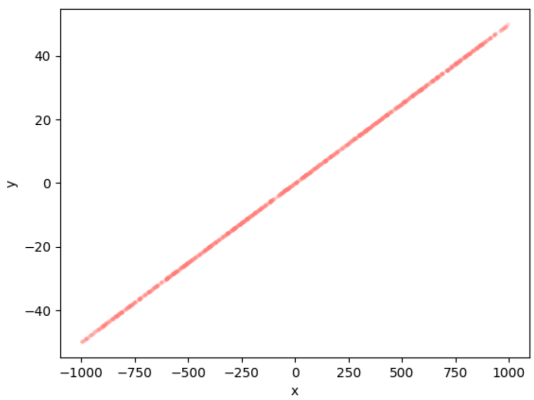
Wykres 6: 2x2 wzór



Wykres 7: 2x2 numpy



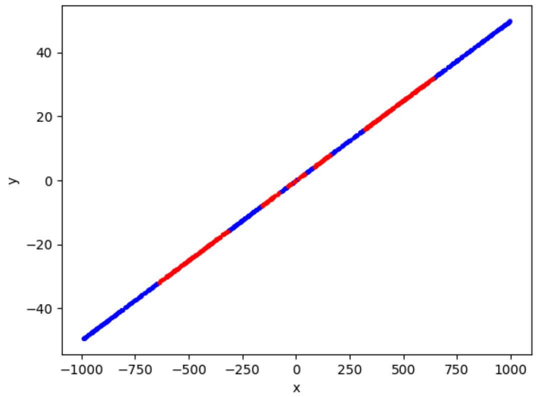
Wykres 8: 3x3 wzór



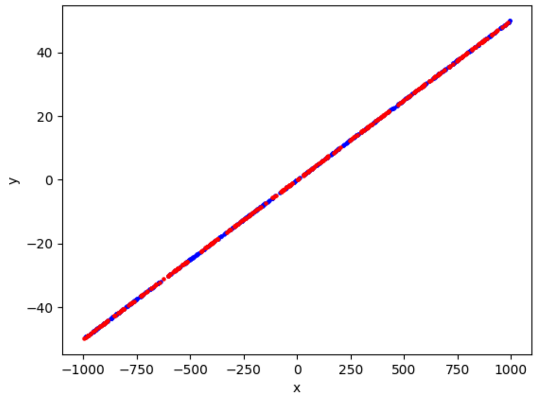
Wykres 9: 3x3 numpy

Na wykresach 6,7,8,9 są przedstawione zależności między ε i miejscem punktu, bardziej czerwone punkty zostały zaklasyfikowane dla większych wartości ε.

Na wykresie 5 okazuje się bardzo dziwna zależność. Metoda *3x3 wzór* nie zakwalifikowała punktów znajdujących się w okolicy -250 i 250 na osi x. Natomiast patrząc na wykresy 8 i 9 można wywnioskować że taki charakterystyczny układ punktów nie wystąpił dla metody *3x3 numpy*. Należy sprawdzić różnice między nimi i innej metody.



Wykres 10: Porównanie *2x2 wzór* i *3x3 wzór*. Czerwone to punkty z *3x3 wzór* a niebieskie z *2x2 wzór*



Wykres 11: Porównanie *2x2 `numpy`* i *3x3 `numpy`*. Czerwone to punkty z *3x3 `numpy`* a niebieskie z *2x2 `numpy`*

Patrząc na wykres 10 można bardzo łatwo zobaczyć że *3x3 wzór* zalicza punkty będące na linii w dość wyjątkowym wzorze. Ten wzór nie występuje w *3x3 `numpy`* i jest to bardzo widoczne na wykresach 9 i 11. Ten wzór może być powodowany przez bardzo dużą ilość mnożenia przy obliczaniu *3x3 wzór*.

5. Wnioski

Najbardziej dokładną metodą wydaje się być stosowanie wzoru Sarrusa dla macierzy 3x3. Poprawnie wytypowała wszystkie punkty już przy $\varepsilon = 10^{-14}$ podczas gdy inne warianty potrzebowały większej tolerancji dla zera. Metoda korzystająca z macierzy 2x2 i wzoru wypada dobrze gdy $\varepsilon = 0$. Funkcje biblioteczne wypadły trochę gorzej ponieważ są przystosowane do większych macierzy.