

# Algorytmy geometryczne

## Sprawozdanie z laboratorium 1.

Dawid Zawiślak gr. Czw. 13:00 A

### **Dane techniczne urządzenia oraz narzędzia za pomocą których wykonano ćwiczenia:**

- Laptop z systemem operacyjnym Pop! OS 22.04 LTS x64
- Procesor Intel Core i7-8750H
- RAM 16 GB

Do realizacji ćwiczenia użyto środowiska Jupyter Notebook z Pythonem w wersji 3.11.5, wykorzystując bibliotekę numpy(do obliczeń numerycznych), random(do generowania liczb pseudolosowych) oraz visualizer przygotowany przez KN BIT(do tworzenia wykresów).

# Realizacja ćwiczenia:

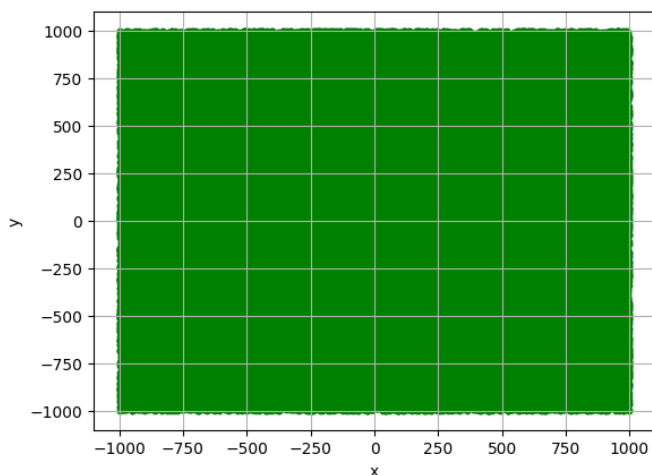
Celem ćwiczenia jest porównanie wyników klasyfikacji położenia punktu względem odcinka w zależności od metody obliczania wyznacznika, tolerancji przy porównywaniu z 0 oraz precyzji zmiennych użytych przy obliczeniach.

## 1. Generowanie zbiorów punktów do testów

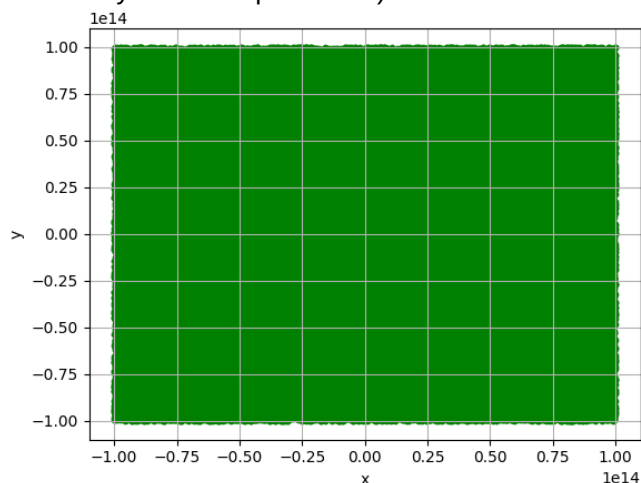
W celu realizacji ćwiczenia konieczne było przygotowanie odpowiednich danych testowych. Dane te składały się z 4 zbiorów punktów 2D których współrzędne były typu float, zostały one wylosowane za pomocą funkcji uniform z modułu random:

- $10^5$  losowych punktów o współrzędnych z przedziału  $[-1000, 1000]$  (Wyk. 1),
- $10^5$  losowych punktów o współrzędnych z przedziału  $[-10^{14}, 10^{14}]$  (Wyk. 2.),
- 1000 losowych punktów leżących na okręgu o środku  $(0,0)$  i promieniu  $R=100$  (Wyk. 3.)
- 1000 losowych punktów o współrzędnych z przedziału  $[-1000, 1000]$  leżących na prostej wyznaczonej przez wektor  $(\mathbf{a}, \mathbf{b})$  – przyjmujemy, że  $\mathbf{a} = [-1.0, 0.0]$ ,  $\mathbf{b} = [1.0, 0.1]$  (Wyk. 4.)

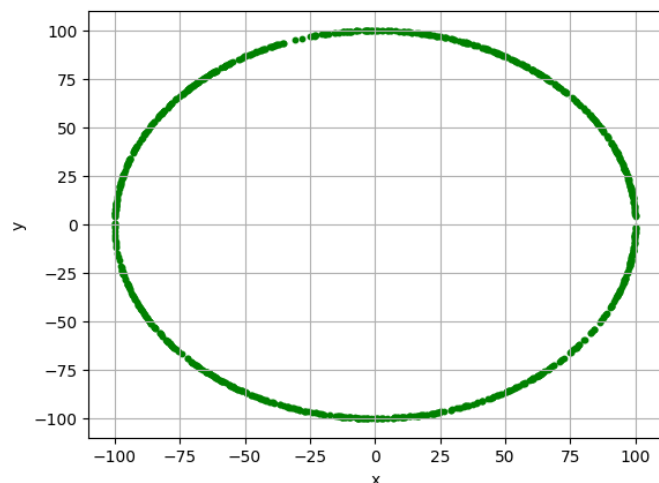
Wyk. 1. Zbiór punktów a)



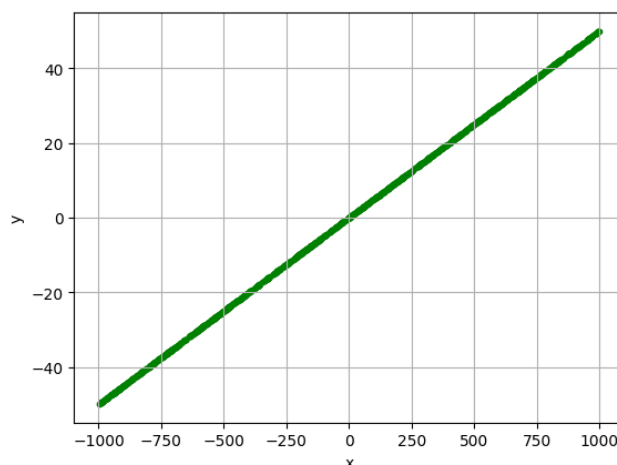
Wyk. 2. Zbiór punktów b)



Wyk. 3. Zbiór punktów c)



Wyk. 4. Zbiór punktów d)



## 2. Funkcje zastosowane do obliczania wyznacznika

Do przeprowadzenia doświadczenia zdefiniowano 4 funkcje do obliczania wyznacznika, po 2 dla wyznaczników 2x2 i 3x3 – własne implementacje oraz biblioteczne(np.linalg.det):

- a) mat\_det\_3x3
- b) mat\_det\_3x3\_lib
- c) mat\_det\_2x2
- d) mat\_det\_2x2\_lib

Rys. 1. Wyznacznik 3x3

$$\det(a, b, c) = \begin{vmatrix} a_x & a_y & 1 \\ b_x & b_y & 1 \\ c_x & c_y & 1 \end{vmatrix}$$

Rys. 2. Wyznacznik 2x2

$$\det(a, b, c) = \begin{vmatrix} a_x - c_x & a_y - c_y \\ b_x - c_x & b_y - c_y \end{vmatrix}$$

Gdzie **a**, **b** to punkty wyznaczające wektor prostej, a **c** to punkt którego położenie określamy.

## 3. Zastosowane wartości tolerancji dla 0 oraz precyzje float używane do obliczeń:

Przeprowadzono analizę wyników kategoryzacji dla następujących tolerancji dla 0:

- a)  $\varepsilon = 0$
- b)  $\varepsilon = 1e-14$
- c)  $\varepsilon = 1e-12$
- d)  $\varepsilon = 1e-10$
- e)  $\varepsilon = 1e-8$

Do obliczeń używano liczb zmiennoprzecinkowych o precyzji 64 oraz 32 bity.

## 4. Kategoryzacja punktów oraz wizualizacja wyników

Do kategoryzacji punktów użyto funkcji, która dla podanej metody obliczania wyznacznika zwraca 3 listy – odcinki zklasyfikowane jako znajdujące się nad prostą, na prostej oraz poniżej. Następnie wyniki zwizualizowano za pomocą modułu visualizer na wykrysach za pomocą 3 kolorów: **zielony** – na lewo od prostej, **czerwony** – na prostej, **żółty** – po prawo od prostej. Podano również wyniki ilościowe dla każdej klasyfikacji.

Przeprowadzono doświadczenia na zbiorach testowych dla wymienionych wyżej wartości tolerancji zera a następnie powtórzono je dla floata o mniejszej precyzji 32-bit.

Tabela 1. Wyniki klasyfikacji dla pierwszego zbioru danych testowych

epsilon	wyznacznik	float64			float32		
		na lewo	na linii	po prawo	na lewo	na linii	po prawo
0	mat_det_3x3	50015	0	49985	50015	0	49985
	mat_det_3x3_lib	50015	0	49985	50015	0	49985
	mat_det_2x2	50015	0	49985	50015	0	49985
	mat_det_2x2_lib	50015	0	49985	50015	0	49985
1.00E-14	mat_det_3x3	50015	0	49985	50015	0	49985
	mat_det_3x3_lib	50015	0	49985	50015	0	49985
	mat_det_2x2	50015	0	49985	50015	0	49985
	mat_det_2x2_lib	50015	0	49985	50015	0	49985
1.00E-12	mat_det_3x3	50015	0	49985	50015	0	49985
	mat_det_3x3_lib	50015	0	49985	50015	0	49985
	mat_det_2x2	50015	0	49985	50015	0	49985
	mat_det_2x2_lib	50015	0	49985	50015	0	49985
1.00E-10	mat_det_3x3	50015	0	49985	50015	0	49985
	mat_det_3x3_lib	50015	0	49985	50015	0	49985
	mat_det_2x2	50015	0	49985	50015	0	49985
	mat_det_2x2_lib	50015	0	49985	50015	0	49985
1.00E-08	mat_det_3x3	50015	0	49985	50015	0	49985
	mat_det_3x3_lib	50015	0	49985	50015	0	49985
	mat_det_2x2	50015	0	49985	50015	0	49985
	mat_det_2x2_lib	50015	0	49985	50015	0	49985

Tabela 2. Wyniki klasyfikacji dla drugiego zbioru danych testowych

epsilon	wyznacznik	float64			float32		
		na lewo	na linii	po prawo	na lewo	na linii	po prawo
0	mat_det_3x3	49944	0	50056	49944	0	50056
	mat_det_3x3_lib	49944	0	50056	49944	0	50056
	mat_det_2x2	49940	6	50054	49939	9	50052
	mat_det_2x2_lib	49939	8	50053	49940	6	50054
1.00E-14	mat_det_3x3	49944	0	50056	49944	0	50056
	mat_det_3x3_lib	49944	0	50056	49944	0	50056
	mat_det_2x2	49940	6	50054	49939	9	50052
	mat_det_2x2_lib	49939	8	50053	49940	6	50054
1.00E-12	mat_det_3x3	49944	0	50056	49944	0	50056
	mat_det_3x3_lib	49944	0	50056	49944	0	50056
	mat_det_2x2	49940	6	50054	49939	9	50052
	mat_det_2x2_lib	49939	8	50053	49940	6	50054
1.00E-10	mat_det_3x3	49944	0	50056	49944	0	50056
	mat_det_3x3_lib	49944	0	50056	49944	0	50056
	mat_det_2x2	49940	6	50054	49939	9	50052
	mat_det_2x2_lib	49939	8	50053	49940	6	50054
1.00E-08	mat_det_3x3	49944	0	50056	49944	0	50056
	mat_det_3x3_lib	49944	0	50056	49944	0	50056
	mat_det_2x2	49940	6	50054	49939	9	50052
	mat_det_2x2_lib	49939	8	50053	49940	6	50054

Tabela 3. Wyniki klasyfikacji dla trzeciego zbioru danych testowych

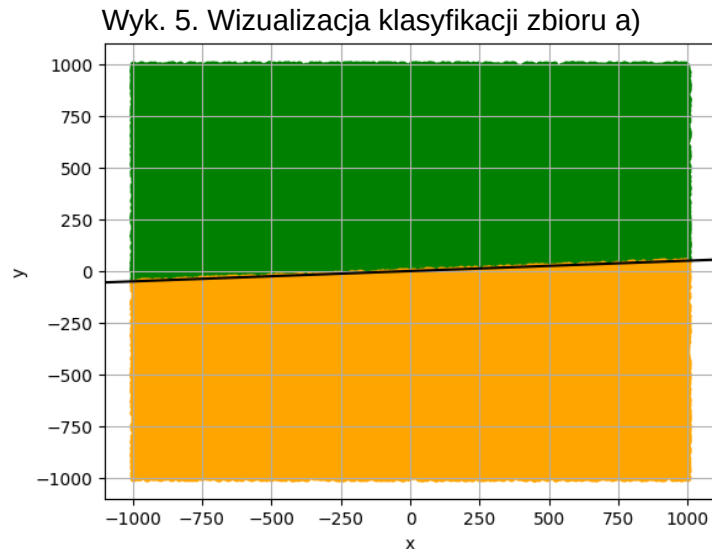
epsilon	wyznacznik	float64			float32		
		na lewo	na linii	po prawo	na lewo	na linii	po prawo
0	mat_det_3x3	520	0	480	520	0	480
	mat_det_3x3_lib	520	0	480	520	0	480
	mat_det_2x2	520	0	480	520	0	480
	mat_det_2x2_lib	520	0	480	520	0	480
1.00E-14	mat_det_3x3	520	0	480	520	0	480
	mat_det_3x3_lib	520	0	480	520	0	480
	mat_det_2x2	520	0	480	520	0	480
	mat_det_2x2_lib	520	0	480	520	0	480
1.00E-12	mat_det_3x3	520	0	480	520	0	480
	mat_det_3x3_lib	520	0	480	520	0	480
	mat_det_2x2	520	0	480	520	0	480
	mat_det_2x2_lib	520	0	480	520	0	480
1.00E-10	mat_det_3x3	520	0	480	520	0	480
	mat_det_3x3_lib	520	0	480	520	0	480
	mat_det_2x2	520	0	480	520	0	480
	mat_det_2x2_lib	520	0	480	520	0	480
1.00E-08	mat_det_3x3	520	0	480	520	0	480
	mat_det_3x3_lib	520	0	480	520	0	480
	mat_det_2x2	520	0	480	520	0	480
	mat_det_2x2_lib	520	0	480	520	0	480

Tabela 4. Wyniki klasyfikacji dla czwartego zbioru danych testowych

epsilon	wyznacznik	float64			float32		
		na lewo	na linii	po prawo	na lewo	na linii	po prawo
0	mat_det_3x3	69	739	192	479	99	422
	mat_det_3x3_lib	354	298	348	492	71	437
	mat_det_2x2	148	719	133	435	171	394
	mat_det_2x2_lib	153	683	164	462	126	412
1.00E-14	mat_det_3x3	0	828	172	463	153	384
	mat_det_3x3_lib	16	885	99	432	177	391
	mat_det_2x2	141	731	128	435	173	392
	mat_det_2x2_lib	143	702	155	460	130	410
1.00E-12	mat_det_3x3	0	1000	0	426	190	384
	mat_det_3x3_lib	0	1000	0	426	190	384
	mat_det_2x2	84	840	76	426	190	384
	mat_det_2x2_lib	112	779	109	426	190	384
1.00E-10	mat_det_3x3	0	1000	0	426	190	384
	mat_det_3x3_lib	0	1000	0	426	190	384
	mat_det_2x2	0	1000	0	426	190	384
	mat_det_2x2_lib	0	1000	0	426	190	384
1.00E-08	mat_det_3x3	0	1000	0	425	194	381
	mat_det_3x3_lib	0	1000	0	425	194	381
	mat_det_2x2	0	1000	0	425	194	381
	mat_det_2x2_lib	0	1000	0	425	194	381

## Wizualizacja wyników

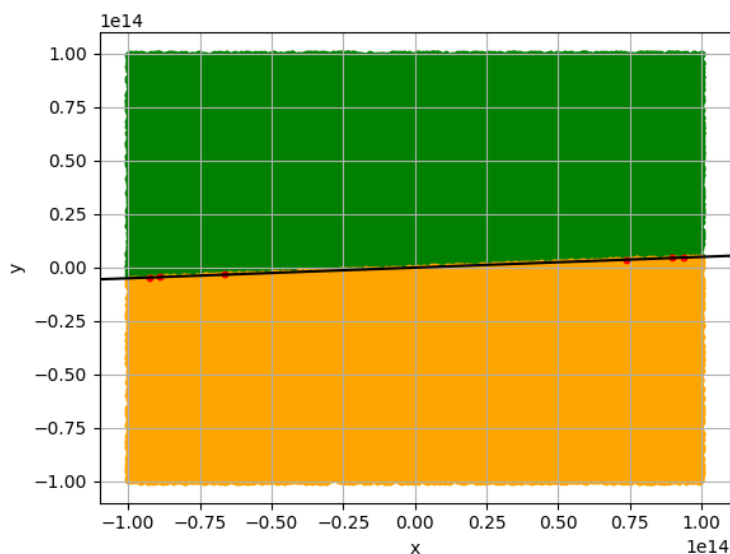
Zbiór punktów a) został sklasyfikowany tak samo dla każdej metody wyznacznika, przyjętej tolerancji wartości 0, oraz użytej precyzji obliczeń (Tabela 1.). Poniżej wizualizacja:



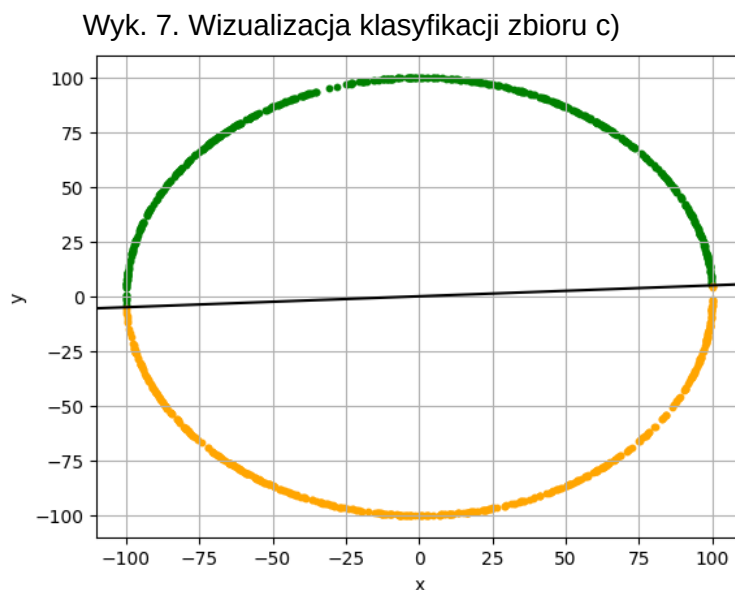
Na zbiorze punktów b) widać już nieznaczne różnice w klasyfikacji w zależności od metody wyznaczania wyznacznika i precyzji obliczeń (Tabela 2.).

Za to wyniki nie zależą od zmiennej tolerancji 0, zapewne dlatego, że punkty losowane są z bardzo dużego przedziału  $[-10^{14}, 10^{14}]$ . Możemy zauważyć, że punkty są klasyfikowane inaczej głównie ze względu na użytą metodę wyznacznika 2x2 oraz 3x3. Punkty klasyfikowane inaczej mają dużą wartość bezwzględną na współrzędnej x i małą wartość bezwzględną współrzędnej y. Poniżej wizualizacja:

Wyk. 6. Wizualizacja klasyfikacji zbioru b) dla funkcji wyznacznika `mat_det_2x2`, precyzji `float64` oraz  $\varepsilon = 1e-14$

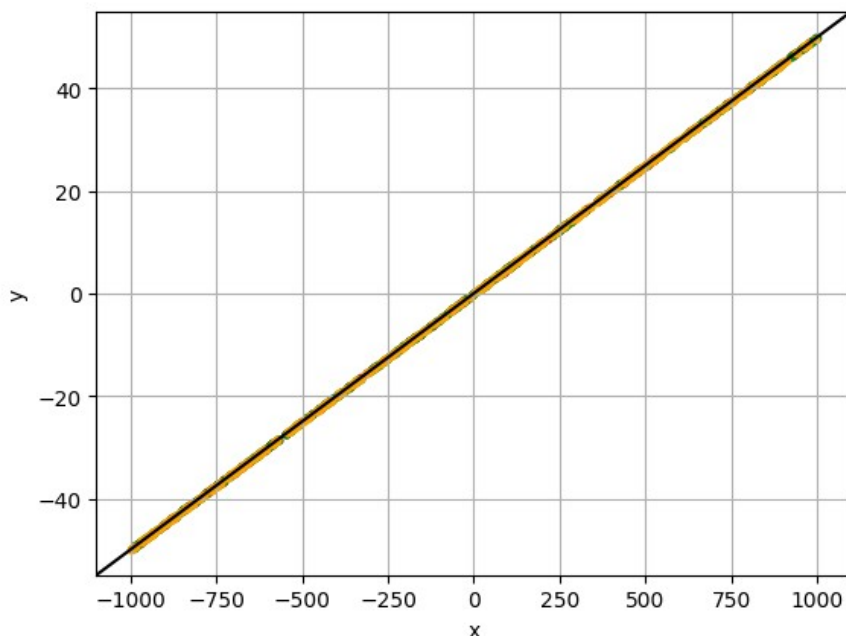


Dla zbioru c) wyniki klasyfikacji są takie same niezależnie od zmiennych (Tabela 3). Jest to zrozumiałe, ponieważ tylko 2 punkty na całym okręgu mogą leżeć na prostej, dlatego jest bardzo małe prawdopodobieństwo że wylosujemy któryś z nich. Wizualizacja na Wyk. 7.

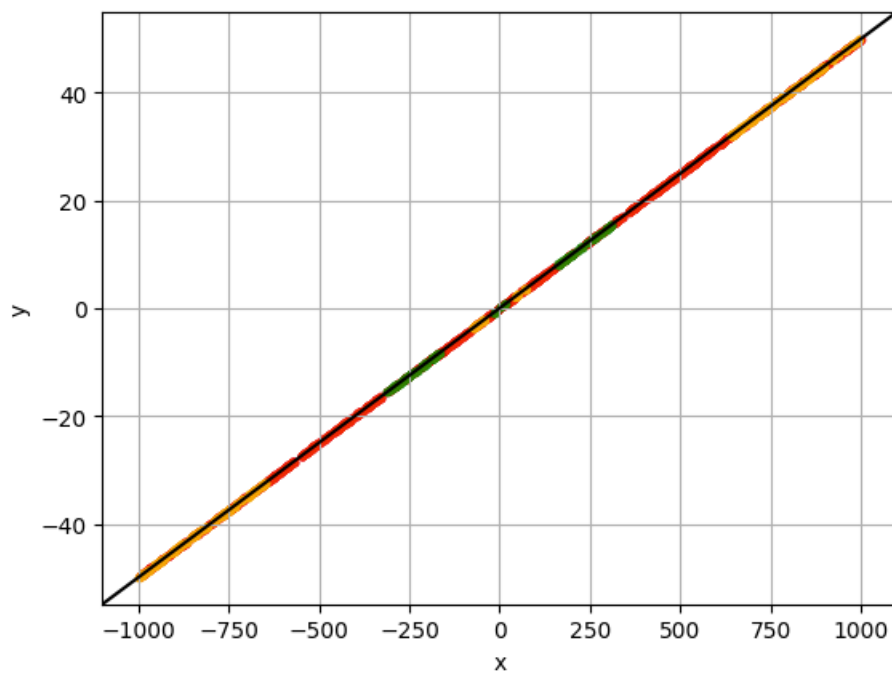


Po testach na zbiorze c) jesteśmy w stanie stwierdzić czy zwracane wyniki klasyfikacji są poprawne, bo punkty w tym zbiorze testowym z zależności znajdują się na linii. Wyniki klasyfikacji różnią się za względu na wszystkie 3 zmienne (Tabela 4.). Dla precyzji float32 są podobne, ale bardzo niedokładne (7-20% skuteczności) – Wyk. 8. W klasyfikacji na ogół lepiej radzą sobie metody wyznaczania wyznacznika własnej implementacji (porównanie na Wyk. 9. i 10.). Można również zauważyć, że wyniki są coraz dokładniejsze dla rosnącej tolerancji dla 0. Wszystkie punkty są klasyfikowane jako leżące na prostej dla  $\varepsilon = 1e-12$  i wyznacznika 3x3 własnej implementacji (Wyk. 11). Również widać zależność jak na zbiorze b) związaną z wyznacznikami 2x2 i 3x3 (wyk. 11. i 12.)

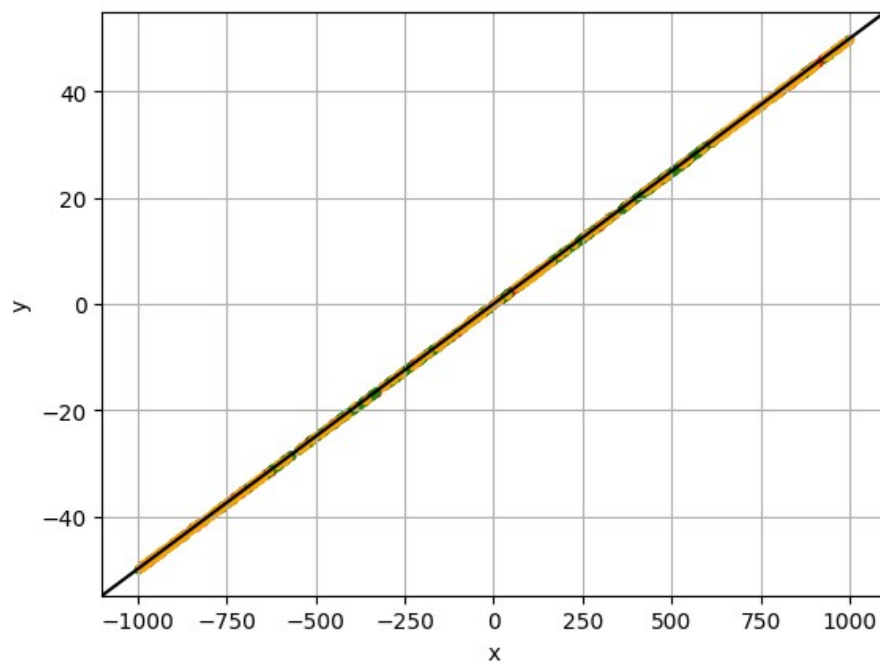
Wyk. 8. Wizualizacja klasyfikacji zbioru d) dla precyzji float32 oraz  $\varepsilon = 1e-10$



Wyk. 9. Wizualizacja klasyfikacji zbioru d) dla precyzji float64,  $\varepsilon = 0$  i funkcji mat\_det\_3x3

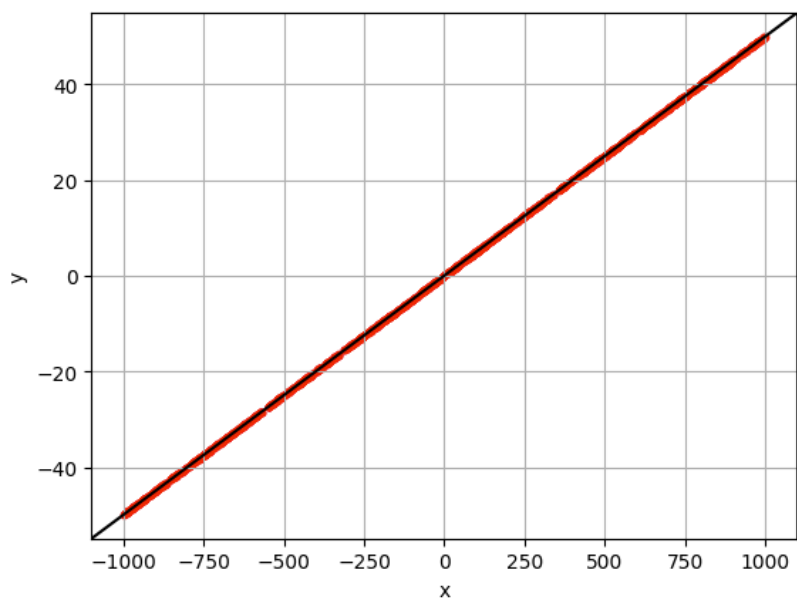


Wyk. 10. Wizualizacja klasyfikacji zbioru d) dla precyzji float64,  $\varepsilon = 0$  i funkcji mat\_det\_3x3\_lib

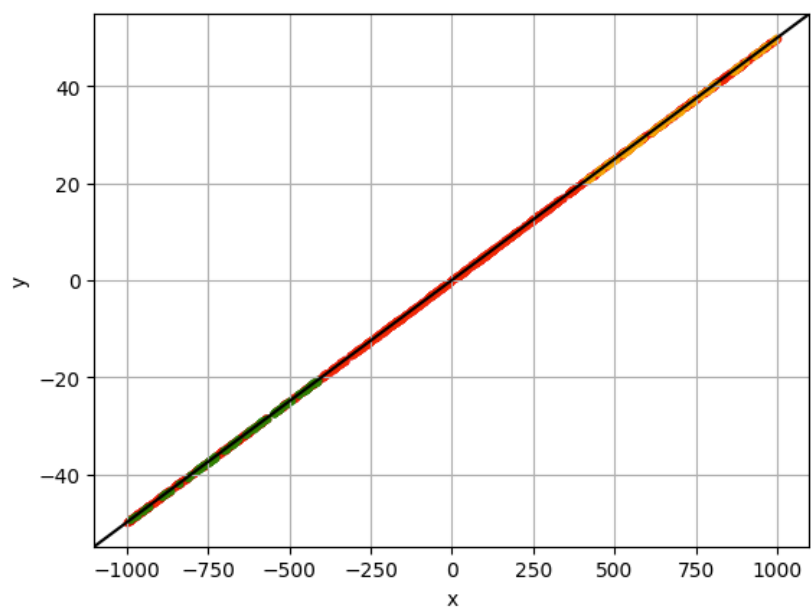




Wyk. 11. Wizualizacja klasyfikacji zbioru d) dla precyzji float64,  $\varepsilon = 1e-12$  i funkcji mat\_det\_3x3



Wyk. 12. Wizualizacja klasyfikacji zbioru d) dla precyzji float64,  $\varepsilon = 1e-12$  i funkcji mat\_det\_2x2



## 5. Wnioski

Dla pierwszych 3 zbiorów danych wyniki nieznacznie się różniły dla wszystkich zmiennych (precyzji obliczeń, metody wyznacznika, wartości tolerancji dla 0). (Tabele 1-3) Również nie jesteśmy w stanie w łatwy sposób określić poprawności klasyfikacji, ponieważ dane generowane są losowo w dużych zakresach.

Jadynie za pomocą testów na zestawie d) (Tabela 4.) jesteśmy w stanie ocenić trafność klasyfikacji przy różnych parametrach, ponieważ punkty z założenia leżą na prostej.

Widzimy, że najbardziej na dokładność klasyfikacji wpływa precyzja obliczeń (użycie float32 znacząco pogarsza trafność) (Wyk. 9. i 10.). Również metody wyznaczania wyznacznika wpływają na zwracane wyniki. Najdokładniejsze wyniki otrzymujemy dla wyznacznika 3x3 własnej implementacji (Wyk. 11), na ogół wyznaczniki biblioteczne dają gorsze wyniki od tych własnej implementacji.

Wyznacznik 2x2 częściej daje niedokładne wyniki od wyznacznika 3x3, wynika to ze specyfiki obliczania tego wyznacznika i niepoprawnej klasyfikacji dla niektórych punktów, których wartość bezwzględna współrzędnej x jest duża, a współrzędnej y mała (Wyk. 11. i 12. oraz Wyk. 6.). Różnice wynikają zapewne ze sposobu obliczania wyznacznika 2x2, gdzie przy takich współrzędnych każdy składnik we wzorze ma bardzo małą wartość, a float ma ograniczoną dokładność i może dawać złe wyniki dla operacji na bardzo małych liczbach.

Jeśli chodzi o wartość tolerancji dla zera, to w pełni poprawną klasyfikację otrzymujemy już dla  $\epsilon = 1e-12$  przy własnej implementacji wyznacznika 3x3 (Wyk. 11.) lub przy wartości  $1e-10$  dla wszystkich wyznaczników.

Po przeprowadzonym doświadczeniu możemy stwierdzić, że najbardziej trafne wyniki klasyfikacji przy zachowaniu małej wartości tolerancji dla zera otrzymamy dla kombinacji:

- precyzja obliczeń – float 64 bitowy
- $\epsilon = 1e-12$
- metoda obliczania wyznacznika – wyznacznik 3x3 własnej implementacji (mat\_det\_3x3)