Przemysław Rola, Juliusz Wasieleski
Informatyka, III rok, grupa 6
październik 2023

Algorytmy macierzowe – rekurencyjne algorytmy macierzowe - sprawozdanie

1. Opis ćwiczenia

Naszym zadaniem było , po wybraniu naszego ulubionego języka, wygenerowanie losowych macierzy których elementy są z przedziału $(10^{-8}, 1)$ i zaimplementowanie algorytmów:

- Rekurencyjnego odwracania macierzy
- Rekurencyjnej LU faktoryzacji macierzy
- Rekurencyjnego obliczania wyznacznika.

Następnie, mieliśmy sprawdzić działanie naszych implementacji na losowo wygenerowanych macierzach rozmiarów $2^k \times 2^k$ gdzie $k \in (2, 3, 4, \dots 16)$.

2. Środowisko, biblioteki, założenia oraz użyte narzędzia

Ćwiczenie wykonaliśmy w języku Python przy użyciu Jupyer Notebooka. Do obliczeń, przechowywania danych użyliśmy bibliotek *numpy, pandas, scipy.*

Do rysowania wykresów użyliśmy biblioteki *matplotlib.*

Wszystkie obliczenia prowadziliśmy na komputerze Lenovo Y50-70 z systemem Windows 10 Pro w wersji 10.0.19045, procesor Intel Core i7-4720HQ 2.60GHz, 2601 MHz, rdzenie: 4, procesory logiczne: 8.

3. Implementacja algorytmów

3.1 Rekurencyjnego odwracania macierzy

3.1.1 Pseudokod

reverse_matrix*(A) :*

Jeżeli *A* ma rozmiar 1**:**

Zwróć *odwrotność elementu A*

W przeciwnym wypadku:

Podziel *A* na 4 równych rozmiarów mniejsze macierze $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$

Oblicz ${A_{11}}^{-1}$

Oblicz pomocniczą macierz $S = A_{22} - A_{21} * {A_{11}}^{-1} * A_{12}$

Zapisz macierz C jako:

$C_{11} = {A_{11}}^{-1} + {A_{11}}^{-1} * A_{12} * S^{-1} * A_{21} * {A_{11}}^{-1}$

$C_{12} = -{A_{11}}^{-1} * A_{12} * S^{-1}$

$C_{21} = -S^{-1} * A_{21} * {A_{11}}^{-1}$

$C_{22} = S^{-1}$

Zwróć C

3.1.2 Istotne fragmenty implementacji

Brak znaczących elementów do opisu, kod jest przepisaniem z pseudokodu na kod Pythona.

3.2 Rekurencyjnej LU faktoryzacji macierzy

3.2.1 Pseudokod

LU_factorise*(A,B):*

Jeżeli *A* ma rozmiar 1**:**

Zwróć *macierz jednostkową o rozmiarze 1 oraz A*

W przeciwnym wypadku:

Podziel *A na* równych rozmiarów mniejsze macierze $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$

Oblicz $L_{11}\ U_{11}$ poprzez rekurencyjne wywołanie LU_factorise(A_{11})

Oblicz ${U_{11}}^{-1}$

$L_{21} = A_{21} * {U_{11}}^{-1}$

Oblicz ${L_{11}}^{-1}$

$U_{12} = {L_{11}}^{-1} * A_{12}$

Oblicz pomocniczą macierz $S = A_{21} * {U_{11}}^{-1} * {L_{11}}^{-1} * A_{12}$

Oblicz $L_{22}\ U_{22}$ poprzez rekurencyjne wywołanie LU_factorise(A_{22})

Zwróć $L \quad oraz\ U$

3.2.2 Istotne fragmenty implementacji

Zakładając że to L ma jedynki na przekątnej, zwracamy najpierw macierz jednostkową dla warunku końcowego.

3.3 Rekurencyjnego obliczania wyznacznika

3.3.1 Pseudokod

```
recursive_det(A):
    L, U = LU_factiorise(A)
    Zwróć iloczyn elementów na przekątnej U
```

3.3.2 Istotne fragmenty implementacji

Brak znaczących elementów do opisu, kod jest przepisaniem z pseudokodu na kod Pythona.

4. Analiza wykonanych pomiarów

4.1 Pomiary rekurencyjnego odwracania macierzy

4.1.1 Pomiary

rozmiar	operacje addytywne	operacje multiplikatywne	wszystkie operacje zmiennoprzecinkowe	czas wykonania
4	126	102	228	0.013107
8	1588	726	2314	0.022753
16	14400	5010	19410	0.157311
32	114856	34542	149398	1.149132
64	862056	239202	1101258	8.047204
128	6270328	1663086	7933414	39.993987
256	44843400	11594370	56437770	288.084035
512	317722936	80967822	398690758	2047.599856

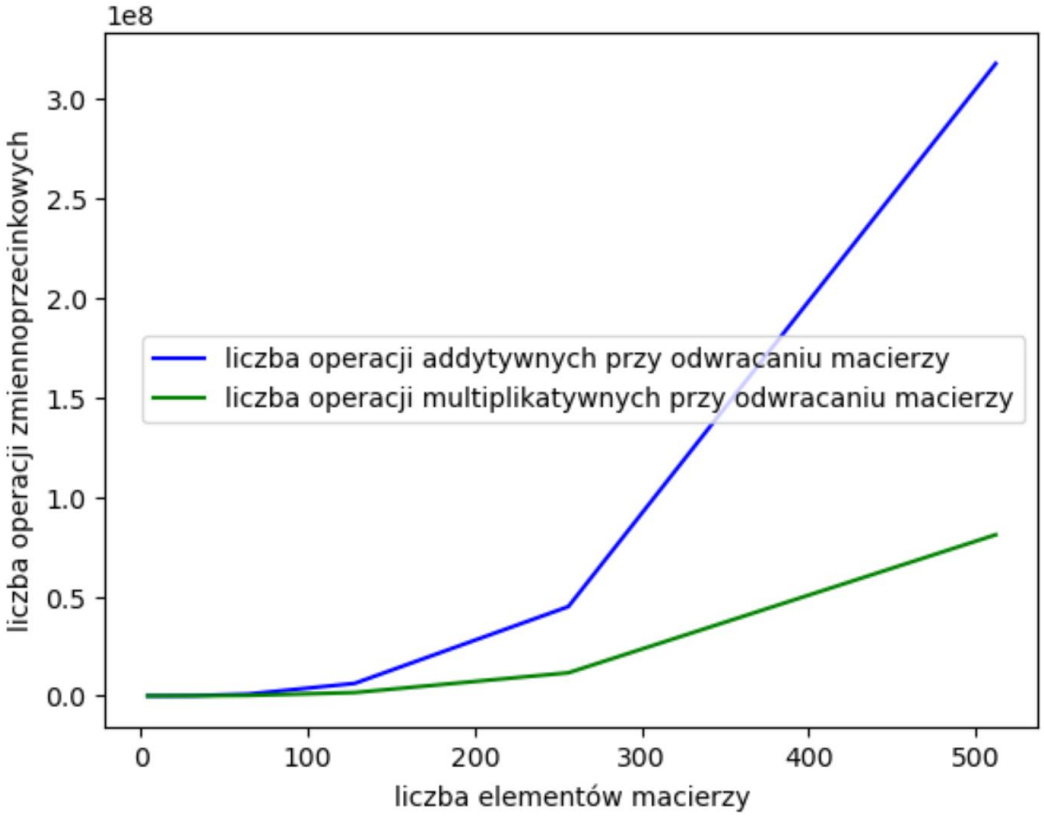
Tab. 1 Pomiary rekurencyjnego odwracania macierzy

Gdzie rozmiar macierzy to ilość elementów w pojedynczym wierszu.

4.1.2 Analiza wyników

Zależność operacji addytywnych od multiplikatywnych

Powyższą zależność przedstawia wykres numer 1. Można zauważyć, że operacji addytywnych jest znacznie więcej niż multiplikatywnych, z czego można wysnuć wniosek, że chcąc przyspieszyć działanie odwracania macierzy lepiej niskopoziomowo przyspieszać dodawanie.



Wykres 1

Szacunek złożoności obliczeniowej

Złożoność obliczeniową szacowaliśmy empirycznie przy użyciu funkcji curve_fit z modułu scipy.optimize, która aproksymuje funkcję przy użyciu metody najmniejszych kwadratów. My próbowaliśmy aproksymować dane do funkcji postaci:

$$y = a \cdot x^k \tag{1}$$

Gdzie próbowaliśmy oszacować a oraz k.

Na podstawie czasu rekurencyjnego odwracania macierzy otrzymaliśmy:

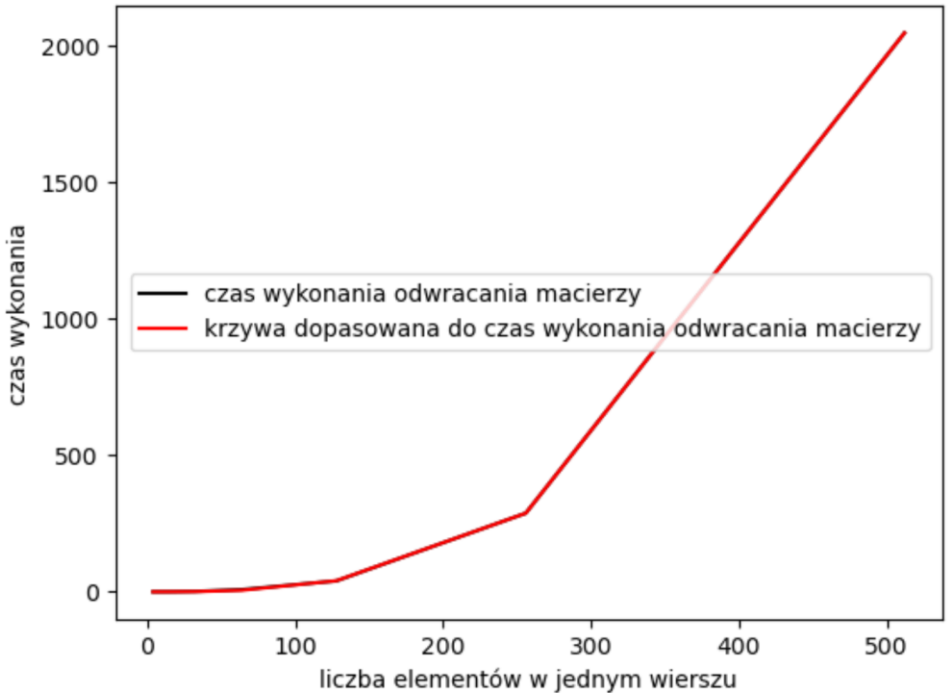
$$a = 4{,}422 \cdot 10^{-5}$$
$$k = 2{,}894 \tag{2}$$

A na podstawie liczby operacji zmiennoprzecinkowych rekurencyjnego odwracania macierzy otrzymaliśmy:

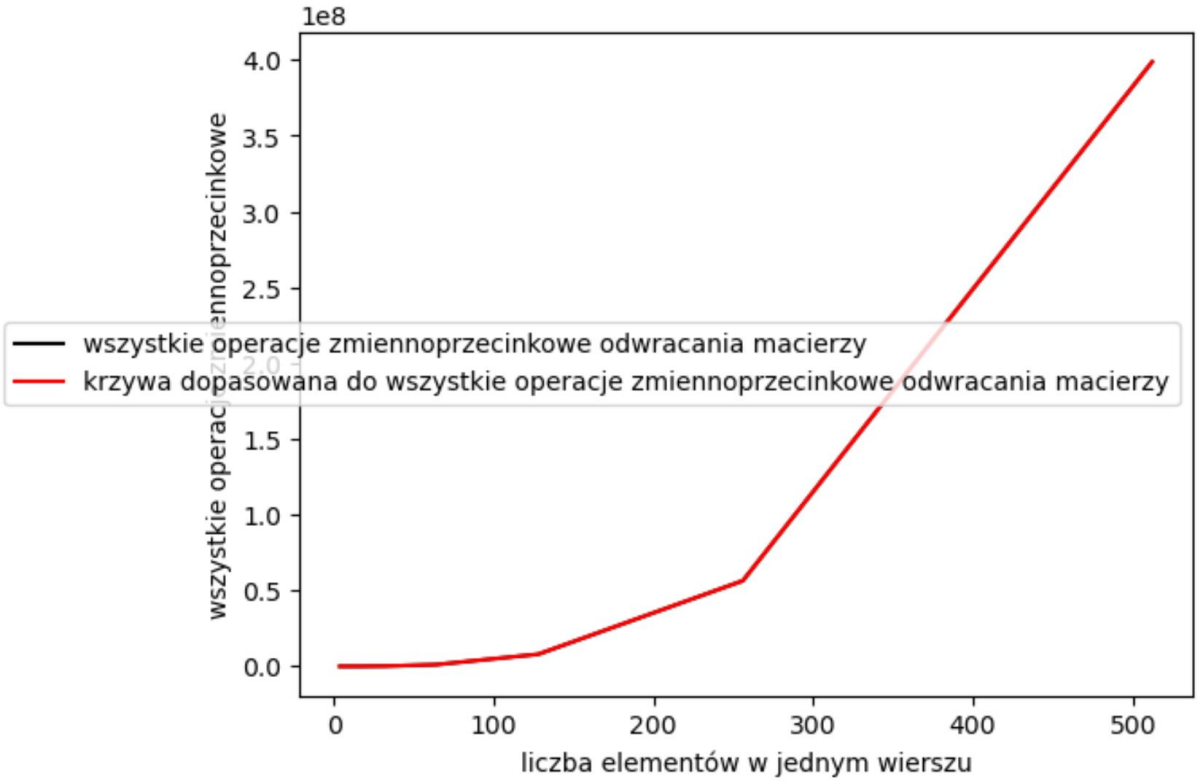
$$a = 9{,}076$$
$$k = 2{,}821 \tag{3}$$

Nasze dopasowane funkcje przedstawiliśmy graficznie razem z oryginalnymi danymi na wykresach numer 2 i 3. Na podstawie tych wykresów możemy stwierdzić, że udało nam się dość dobrze oszacować prawdziwą złożoność, ponieważ oba wykresy pokrywają się ze sobą. Dodatkowo szacunki te pokrywają się z teoretyczną złożonością, która jest ograniczona mnożeniem macierzy. Ponieważ do tego celu korzystaliśmy z algorytmu Strassena, który ma złożoność $O(n^{2{,}807})$ to widzimy, że wyszła nam trochę gorsza złożoność Dlaczego? Domniemamy, że trochę większa złożoność na podstawie liczby operacji zmiennoprzecinkowych wynika z błędów pomiarowych. W przeciwieństwie do niej,

większa złożoność na podstawie czasu działania, wynika ze strat podczas kopiowania macierzy w pamięci i przechowywaniem ich .



Wykres nr. 2



Wykres nr. 3

Zależność liczby operacji zmiennoprzecinkowych od czasu wykonania

Patrząc na fakt, który już wspomnieliśmy, można dostrzec, że szacunek złożoności na podstawie czasu wykonania jest gorszy niż na podstawie liczby operacji zmiennoprzecinkowych ponieważ wyniki są przekłamane ze wsklędu na przechowywanie i przesyłanie macierzy w pamięci.

4.2 Pomiary LU faktoryzacji macierzy

4.2.1 Pomiary

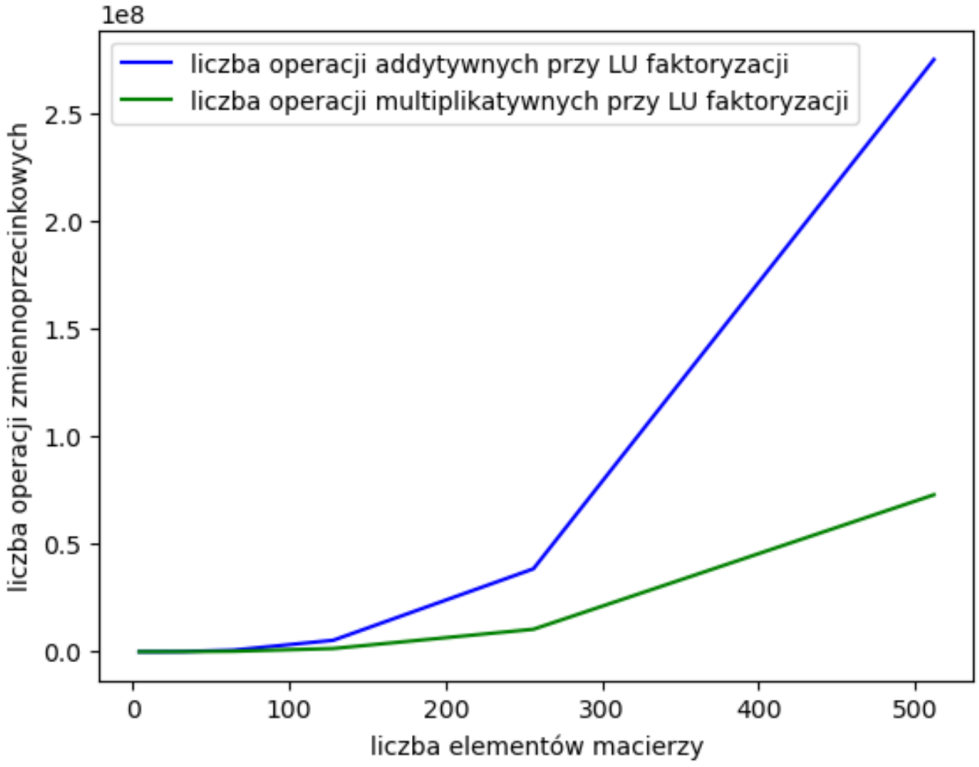
rozmiar	operacje addytywne	operacje multiplikatywne	wszystkie operacje zmiennoprzecinkowe	czas wykonania
4	62	69	131	0.001145
8	1018	587	1605	0.010178
16	10576	4341	14917	0.086105
32	90962	30707	121669	0.565409
64	712536	214533	927069	4.449985
128	5310786	1495715	6806501	32.035126
256	38514632	10435317	48949949	225.179765
512	275067634	72883379	347951013	1554.568073

Tab. 2 Pomiary LU faktoryzacji

4.2.2 Analiza wyników

Zależność operacji addytywnych od multiplikatywnych

Powyższą zależność przedstawia wykres numer 4. Można zauważyć, że operacji addytywnych jest znacznie więcej niż multiplikatywnych, z czego można wysnuć wniosek, że chcąc przyspieszyć działanie LU faktoryzacji lepiej niskopoziomowo przyspieszać dodawanie.



Wykres 4

Szacunek złożoności obliczeniowej

Złożoność obliczeniową szacowaliśmy empirycznie zgodnie ze wzorem nr 1, tak jak w punkcie 4.1.2.

Na podstawie czasu LU faktoryzacji otrzymaliśmy:

$$a = 4{,}334 \cdot 10^{-5}$$
$$k = 2{,}788$$

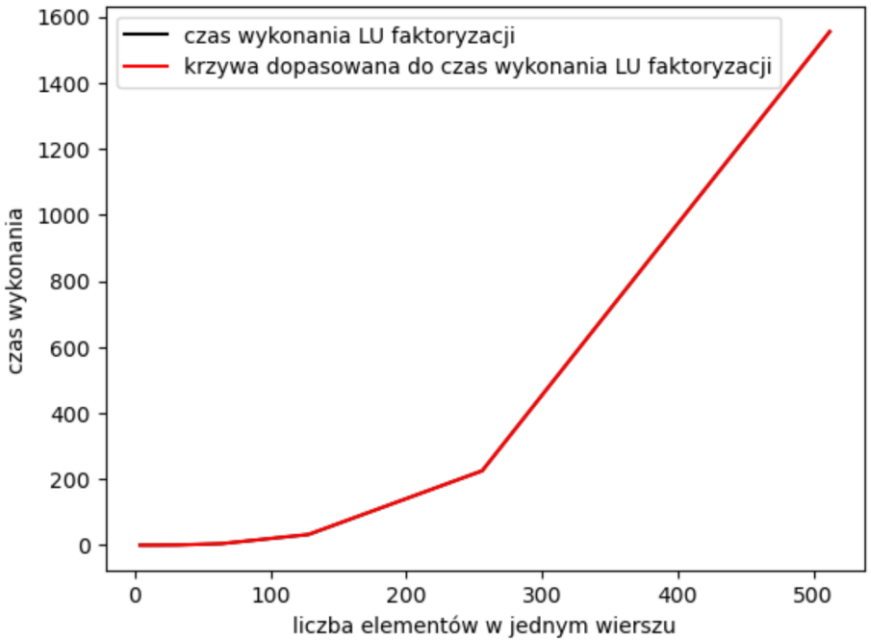
(4)

A na podstawie liczby operacji zmiennoprzecinkowych LU faktoryzacji otrzymaliśmy:

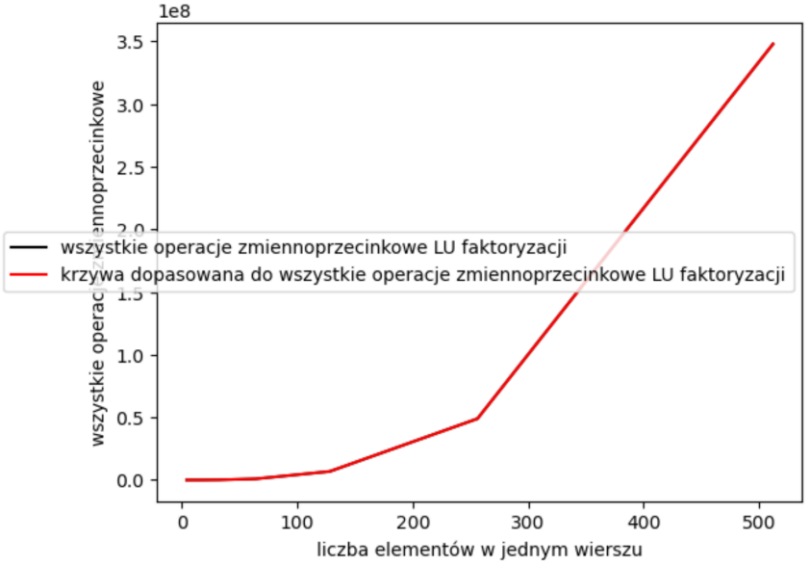
$$a = 7{,}478$$
$$k = 2{,}830$$

(5)

Nasze dopasowane funkcje przedstawiliśmy graficznie razem z oryginalnymi danymi na wykresach numer 5 i 6. Na podstawie tych wykresów możemy stwierdzić, że udało nam się dość dobrze oszacować prawdziwą złożoność, ponieważ oba wykresy pokrywają się ze sobą. Dodatkowo szacunki te pokrywają się z teoretyczną złożonością, która jest ograniczona mnożeniem macierzy. Ponieważ oparliśmy naszą LU faktoryzację na algorytmie mnożenia macierzy metodą Strassena, który ma złożoność $O(n^{2{,}807})$ to możemy zobaczyć że empiryczne szacunki złożoności pokrywają się z teoretycznymi.



Wykres nr. 5



Wykres nr. 6

Zależność liczby operacji zmiennoprzecinkowych od czasu wykonania

Patrząc na fakt, że złożoność wynikająca z czasu wykonania wyszła mniejsza niż ta z liczby operacji zmiennoprzecinkowych możemy wywnioskować, że tak jak dla algorytmu z analizowanego w podpunkcie 4.1.2 przechowywanie macierzy pogarszało złożoność tak tutaj jego czas skalował się lepiej niż czas wykonania operacji zmiennoprzecinkowych i dzięki temu otrzymaliśmy lepszy wynik.

4.3 Rekurencyjnego obliczania wyznacznika

4.3.1 Pomiary

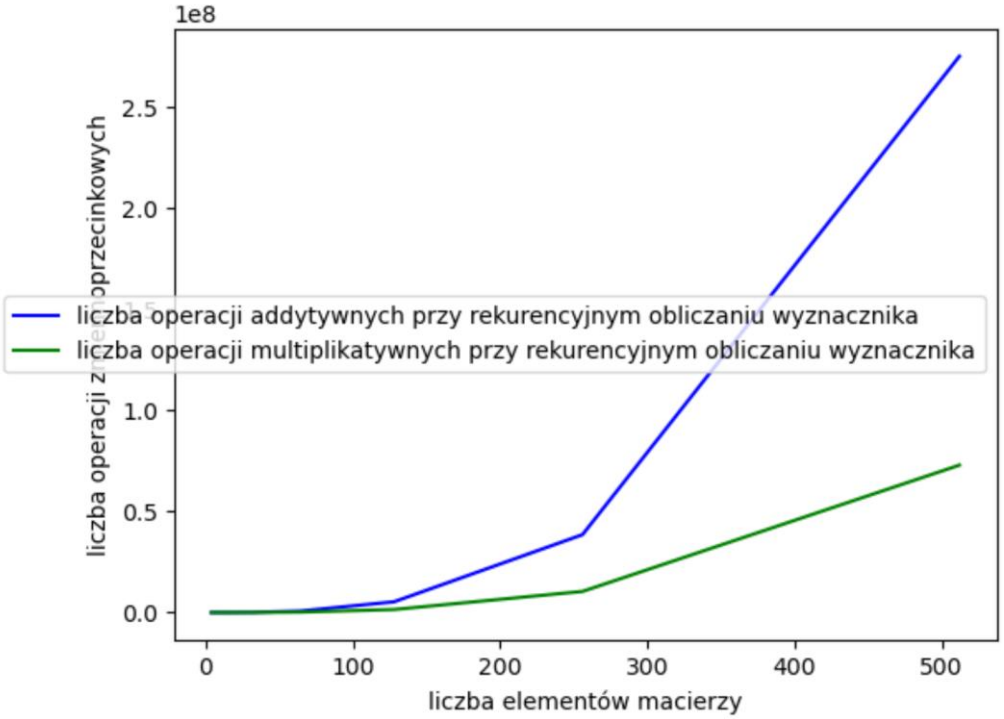
rozmiar	operacje addytywne	operacje multiplikatywne	wszystkie operacje zmiennoprzecinkowe	czas wykonania
4	62	72	134	0.002701
8	1018	594	1612	0.015897
16	10576	4356	14932	0.101679
32	90962	30738	121700	1.445360
64	712536	214596	927132	5.064898
128	5310786	1495842	6806628	35.039661
256	38514632	10435572	48950204	255.147330
512	275067634	72883890	347951524	1771.125482

Tab. 3 Wyniki pomiarów dla rekurencyjnego obliczania wyznacznika macierzy

4.3.2 Analiza wyników

Zależność operacji addytywnych od multiplikatywnych

Powyższą zależność przedstawia wykres numer 1. Można zauważyć, że operacji addytywnych jest znacznie więcej niż multiplikatywnych, z czego można wysnuć wniosek, że chcąc przyspieszyć działanie rekurencyjnego obliczania wyznacznika lepiej niskopoziomowo przyspieszać dodawanie.



Wykres 7

Szacunek złożoności obliczeniowej

Złożoność obliczeniową szacowaliśmy empirycznie zgodnie ze wzorem nr 1, tak jak w punkcie 4.1.2.

Na podstawie czasu mnożenia macierzy metodą zaproponowaną przez sztuczną inteligencję otrzymaliśmy:

$$a = 4{,}656 \cdot 10^{-5}$$
$$k = 2{,}798$$

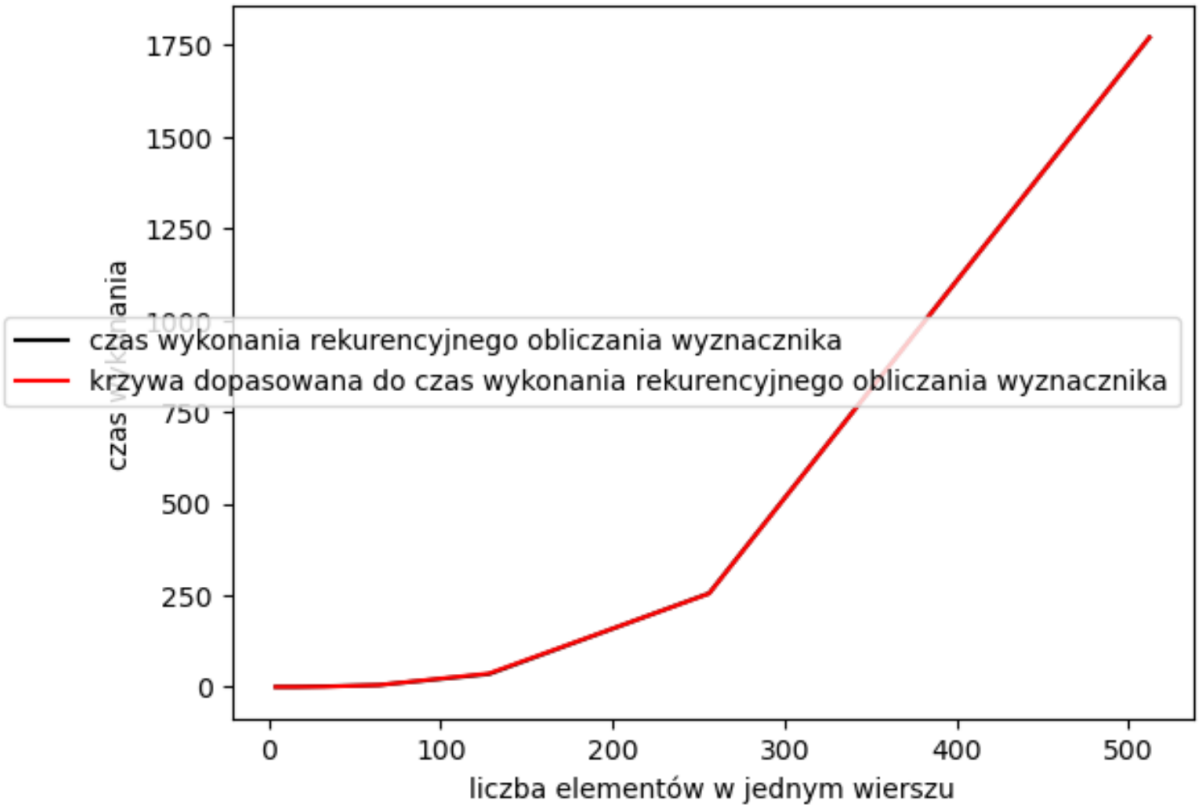
(6)

A na podstawie liczby operacji zmiennoprzecinkowych mnożenia macierzy metodą zaproponowaną przez sztuczną inteligencję otrzymaliśmy:

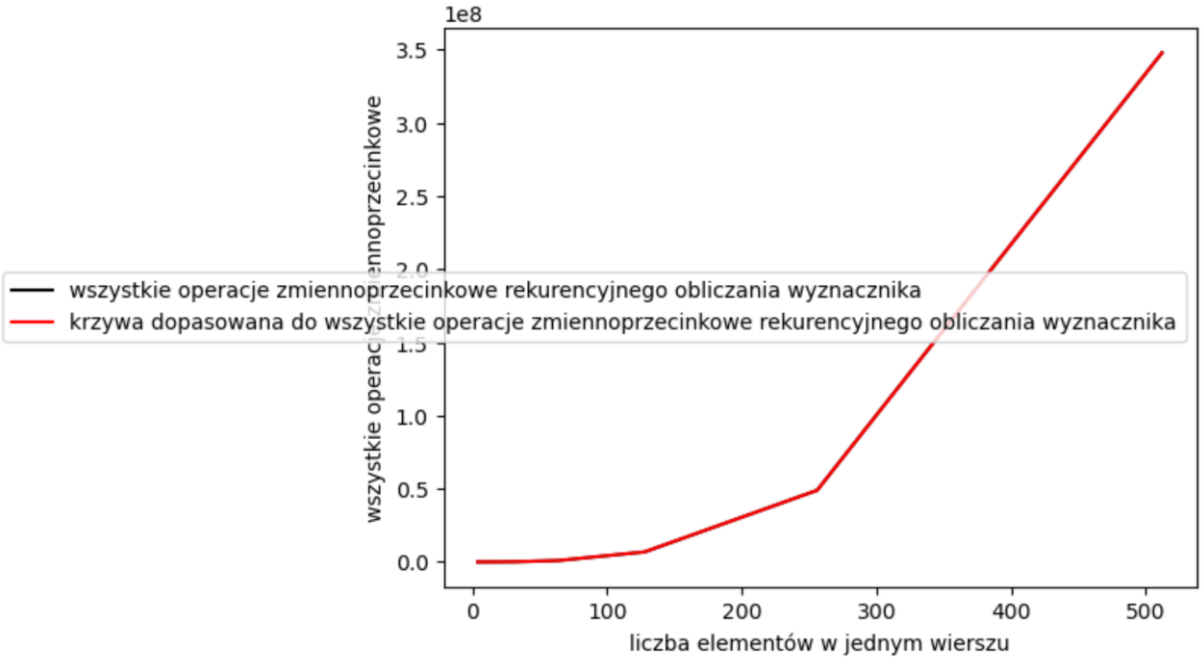
$$a = 7{,}478$$
$$k = 2{,}830$$

(7)

Nasze dopasowane funkcje przedstawiliśmy graficznie razem z oryginalnymi danymi na wykresach numer 5 i 6. Na podstawie tych wykresów możemy stwierdzić, że udało nam się dość dobrze oszacować prawdziwą złożoność, ponieważ oba wykresy pokrywają się ze sobą. Dodatkowo szacunki te pokrywają się z teoretyczną złożonością, która jest ograniczona mnożeniem macierzy. Ponieważ oparliśmy rekurencyjne obliczanie wyznacznika na naszej implementacji LU faktoryzacji (która to z kolei jest oparta na algorytmie mnożenia macierzy metodą Strassena, który ma złożoność $O(n^{2{,}807})$) to możemy zobaczyć że empiryczne szacunki złożoności pokrywają się z teoretycznymi.



Wykres nr. 8

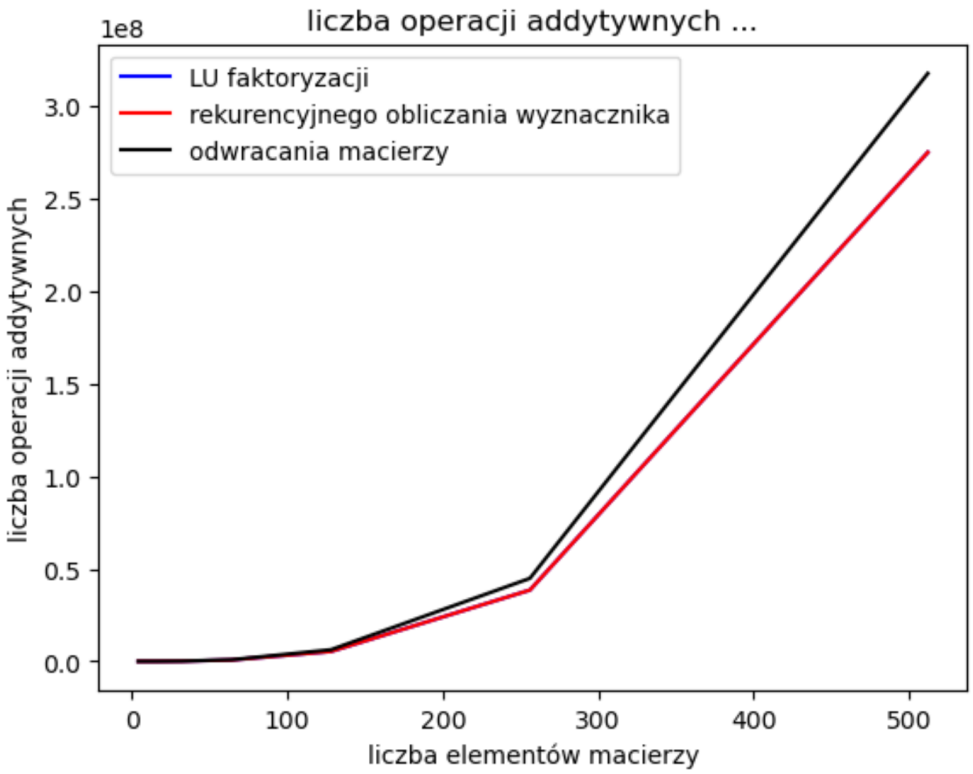


Wykres nr. 9

4.4 Porównanie wyników trzech powyższych algorytmów

4.4.1 Operacje addytywne

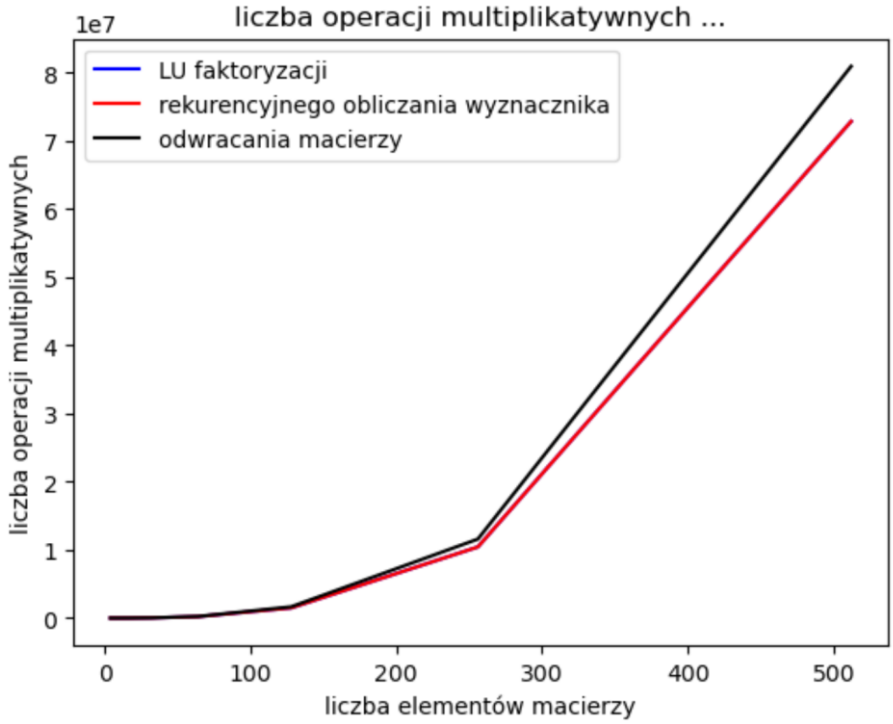
Wykres numer 10 przedstawia graficzne porównanie liczby operacji addytywnych dla wszystkich algorytmów zaimplementowanych przez nas w tym zadaniu. Możemy na nim dostrzec, że LU faktoryzacja wykonuje tyle samo operacji co rekurencyjne obliczanie wyznacznika (co nie jest zaskoczeniem ponieważ ten drugi algorytm mocno bazuje na LU faktoryzacji). Dodatkowo widzimy, że odwracanie macierzy potrzebuje więcej operacji addytywnych niż LU faktoryzacja.



Wykres nr. 10

4.4.2 Operacje multiplikatywne

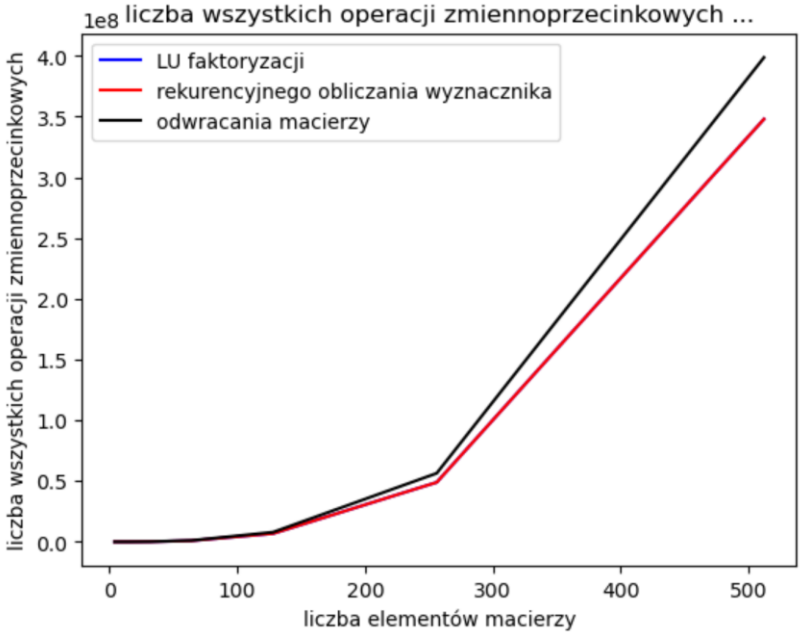
Wykres numer 11 przedstawia graficzne porównanie liczby operacji multiplikatywnych dla wszystkich algorytmów zaimplementowanych przez nas w tym zadaniu. Możemy na nim dostrzec, że podobnie jak w podpunkcie 4.4.1, LU faktoryzacja wykonuje tyle samo operacji co rekurencyjne obliczanie wyznacznika (co nie jest zaskoczeniem ponieważ ten drugi algorytm mocno bazuje na LU faktoryzacji). Dodatkowo widzimy, że odwracanie macierzy potrzebuje więcej operacji addytywnych niż LU faktoryzacja.



Wykres nr.11

4.4.3 Wszystkie operacje zmiennoprzecinkowe

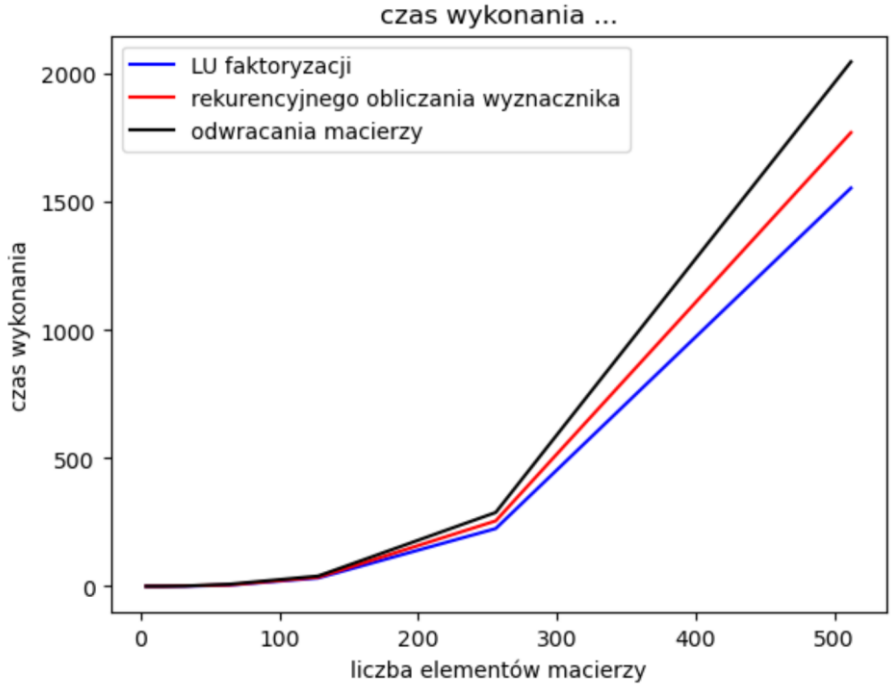
Wykres numer 11 przedstawia graficzne porównanie liczby operacji multiplikatywnych dla wszystkich algorytmów zaimplementowanych przez nas w tym zadaniu. Możemy na nim dostrzec, że podobnie jak w podpunkcie 4.4.1, LU faktoryzacja wykonuje tyle samo operacji co rekurencyjne obliczanie wyznacznika (co nie jest zaskoczeniem ponieważ ten drugi algorytm mocno bazuje na LU faktoryzacji). Dodatkowo widzimy, że odwracanie macierzy potrzebuje więcej operacji addytywnych niż LU faktoryzacja.



Wykres nr. 12

4.4.4 Czas działania

Wykres numer 13 przedstawia graficzne porównanie czasu działania dla wszystkich algorytmów z tego zadania. Na nim możemy dostrzec, że najlepiej wypadają po kolei algorytm: LU faktoryzacji, rekurencyjnego obliczania wyznacznika i odwracania macierzy. Takie wyniki nie są zaskoczeniem ponieważ wyznacznik obliczamy na podstawie LU faktoryzacji. W naszej implementacji, uzyskawszy macierz U obliczamy wyznacznik na podstawie jej przekątnej. Te operacje dają jak widać narzut na czas wykonania (na liczbie operacji zmiennoprzecinkowych nie widzieliśmy zbytnio tego efektu)



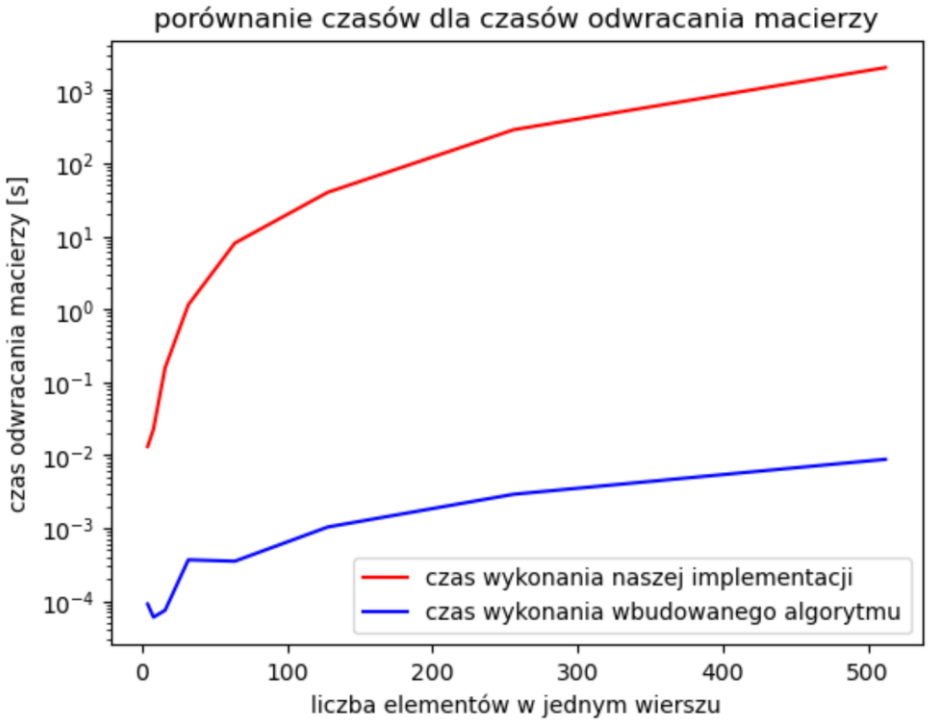
Wykres nr. 13

4.5 Porównanie z numpy

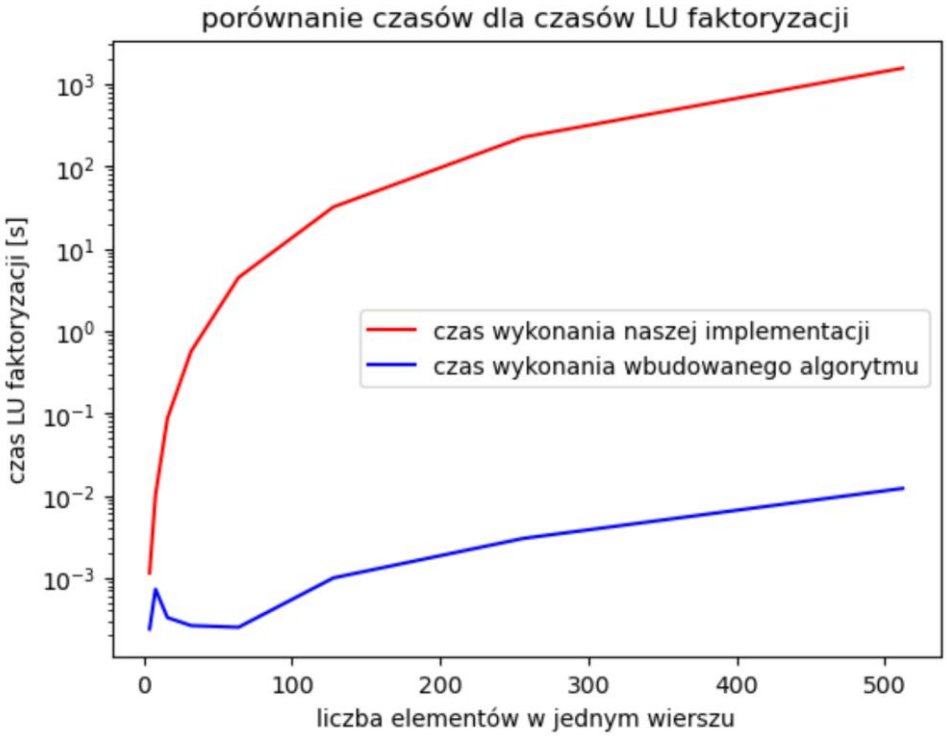
Na poniższych wykresach (14-16) przedstawiliśmy porównanie naszych implementacji z istniejącymi już w bibliotekach do pythona implementacjami, odpowiednio np.linalg.inv, scipy.linalg.lu, np.linalg.det. Implementacje te w rzeczywistości są napisane w językach C i C++ co czyni je porównywalne do języków przeznaczonych do mnożenia macierzy takich jak Matlab czy Octave.

Trzy poniższe wykresy mają na osi Y skalę logarytmiczną, ponieważ w innym wypadku nie bylibyśmy w stanie wyczytać z nich niczego poza tym, że się sporo różnią.

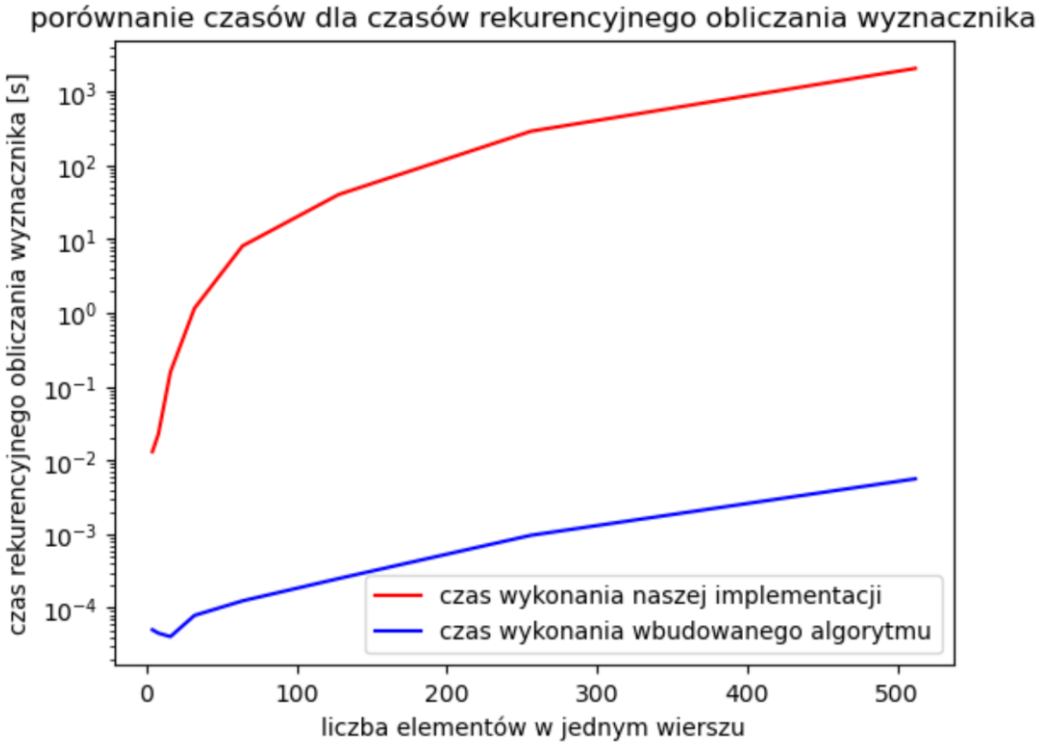
Zauważmy, dla każdego z algorytmów, jak bardzo się różnią od siebie te wykresy. Czasy wykonania algorytmów wbudowanych to ułamki sekund, w momencie kiedy nasze implementacje sięgają nawet czasów 30- kilku minut.



Wykres 14



Wykres 15



Wykres 16

5. Wnioski

- Każdy z implementowanych przez nas algorytmów był zależny od algorytmu mnożenia macierzy na którym go oparliśmy.
- Ze względu na fakt, że wybraliśmy do każdego metodę Strassena, każdy z algorytmów zaimplementowanych przez nas cechował się większą liczbą operacji addytywnych od multiplikatywnych, a złożoność obliczeniowa policzona przez nas empirycznie dla każdego oscylowała niedaleko ($O^{2,807}$) co jest złożonością metody Strassena.
- Czasy działania naszych implementacji są nieporównanie dłuższe od czasów wykonania tych samych algorytmów wbudowanych w pythona. Z tego względu samodzielne implementowanie tych algorytmów w pythonie nie ma innego sensu oprócz dydaktycznego.