# Lab3 - Filip Jedrzejewski

## Opis problemu

Populacja Stanów Zjednoczonych na przestrzeni lat przedstawiała sie nastepujaco:

| Rok | Populacja |
|---|---|
| 1900 | 76 212 168 |
| 1910 | 92 228 496 |
| 1920 | 106 021 537 |
| 1930 | 123 202 624 |
| 1940 | 132 164 569 |
| 1950 | 151 325 798 |
| 1960 | 179 323 175 |
| 1970 | 203 302 031 |
| 1980 | 226 542 199 |

W celu wyznaczenia wielomianu, który interpoluje powyższe dziewieć punktów rozważono nastepujace zbiory funkcji bazowych $\phi(t)$, $j = 1, 2, ..., 9$:

$$\phi_j(t) = t^{j-1} \tag{1}$$

$$\phi_j(t) = (t - 1900)^{j-1} \tag{2}$$

$$\phi_j(t) = (t - 1940)^{j-1} \tag{3}$$

$$\phi_j(t) = \left(\frac{t - 1940}{40}\right)^{j-1} \tag{4}$$

## Wyznaczanie wielomianu za pomoca macierzy

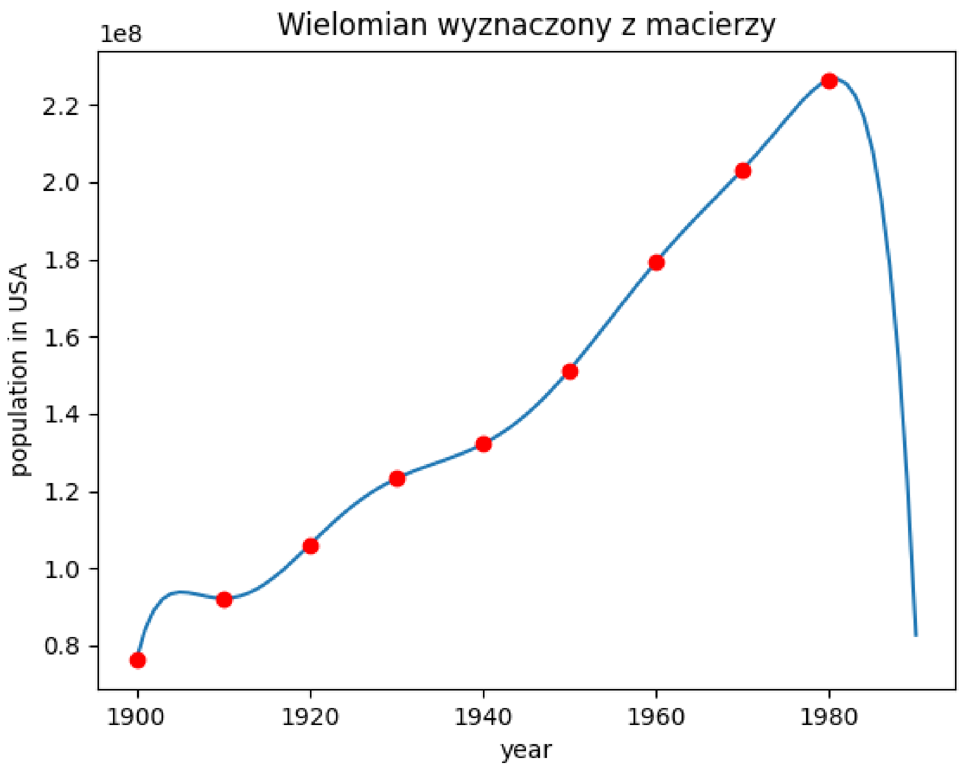
Dla każdego z czterech zbiorów funkcji bazowych utworzono macierz Vandermonde'a, a nastepnie korzystajac z funkcji `numpy.linalg.cond` obliczono współczynniki uwarunkowania każdej z nich. Wyniki przedstawiono w tabeli:

| Numer funkcji bazowej | Współczynnik uwarunkowania macierzy |
|---|---|
| 1 | $5.031 \cdot 10^{26}$ |
| 2 | $6.307 \cdot 10^{15}$ |
| 3 | $9.316 \cdot 10^{12}$ |
| 4 | $1.605 \cdot 10^{3}$ |

Najmniejszy współczynnik uwarunkowania miała baza określona wzorem (4), zatem została ona wybrana do wynaczenia wielomianu interpolacyjnego. W tym celu rozwiazano nastepujace równanie macierzowe:

$$M \cdot C = Y \tag{5}$$

przy czym: $M$ - macierz Vandermonde'a utworzona na najlepiej uwarunkowanej bazie, $C$ - szukana macierz współczynników, $Y$ - macierz wartości wielomianu dla danych punktów (pierwsza tabela).
Do rozwiazania tego równania użyto funkcji `numpy.linalg.solve`, która zwraca rozwiazanie równania macierzowego typu $AX = B$, gdzie $A$ i $B$ sa dane, a $X$ szukana.
W kolejnym kroku utworzono wykres otrzymanego wielomianu na przedziale $year \in [1900, 1990]$, na który naniesiono wezły interpolacji.




Wykres wielomianu interpolacyjnego przechodzi przez wszystkie dziewieć wezłow interpolacji, co potwierdza poprawność wykonanej interpolacji.
Kolejna czynnościa było dokonanie ekstrapolacji wielomianu do 1990 roku. Otrzymano wartość 82749141, która w porównaniu prawdziwej wartości populacji dla Stanów Zjednoczonych w 1990 roku, równej 248709873, miała bład wzgledny ekstrapolacji równy:

$$relativeExtrapolationError = 0.6672864651416437 \approx 66.73\% \tag{6}$$

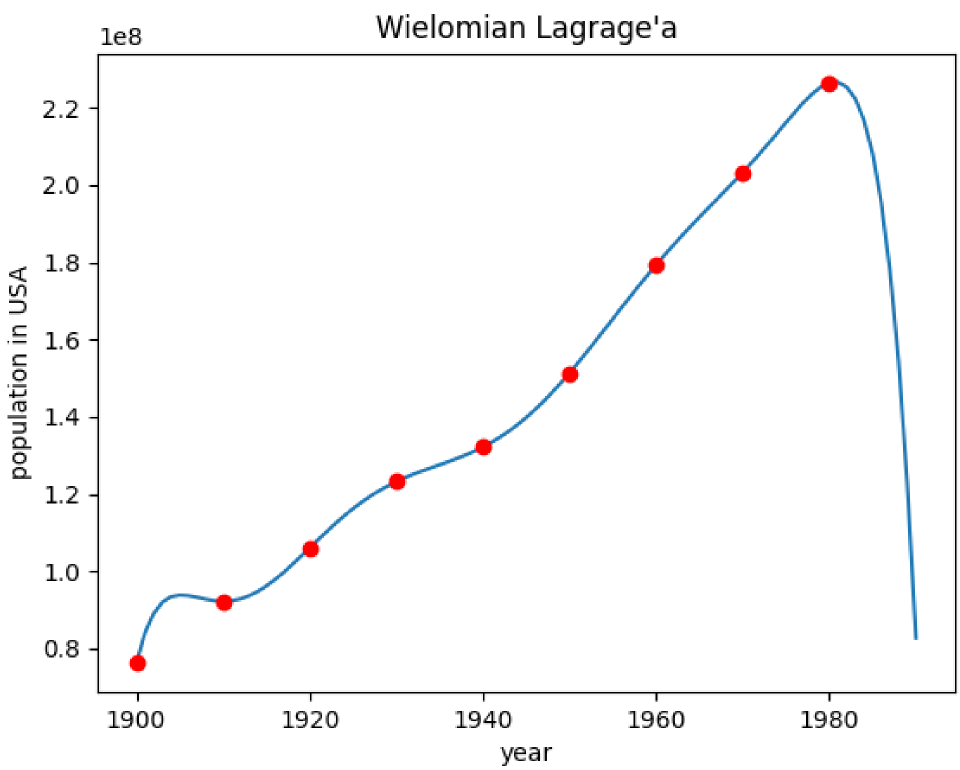
## Wyznaczanie wielomianu Lagrange'a

W celu wyznaczenia wielomianu interpolacyjnego Lagrange'a użyto nastepujacych wzorów:

$$l_j(t) = \prod_{k=1, k\neq j}^{n} \frac{t - t_k}{t_j - t_k} \tag{7}$$

$$p(t) = \sum_{i=1}^{n} y_i l_i(t) \tag{8}$$

Na ich podstawie, korzystajac z funkcji `lambda`, wyznaczono wielomian interpolacyjny Lagrange'a. Nastepnie obliczono wartości tego wielomianu dla takiego samego przedziału jak w poprzednim podpunkcie z krokiem co jeden rok. Na podstawie uzyskanych wartości stworzono wykres:




Powyższy wykres jest identyczny jak ten stworzony na podstawie macierzy, co potwierdza poprawność wyznaczenia wielomianu metoda Lagrange'a.

## Wyznaczanie wielomianu Newtona

W celu wyznaczenia wielomianu interpolacyjnego Newtona użyto nastepujacych wzorów:

$$\pi(t) = \prod_{k=1}^{j-1} (t - t_k) \tag{9}$$

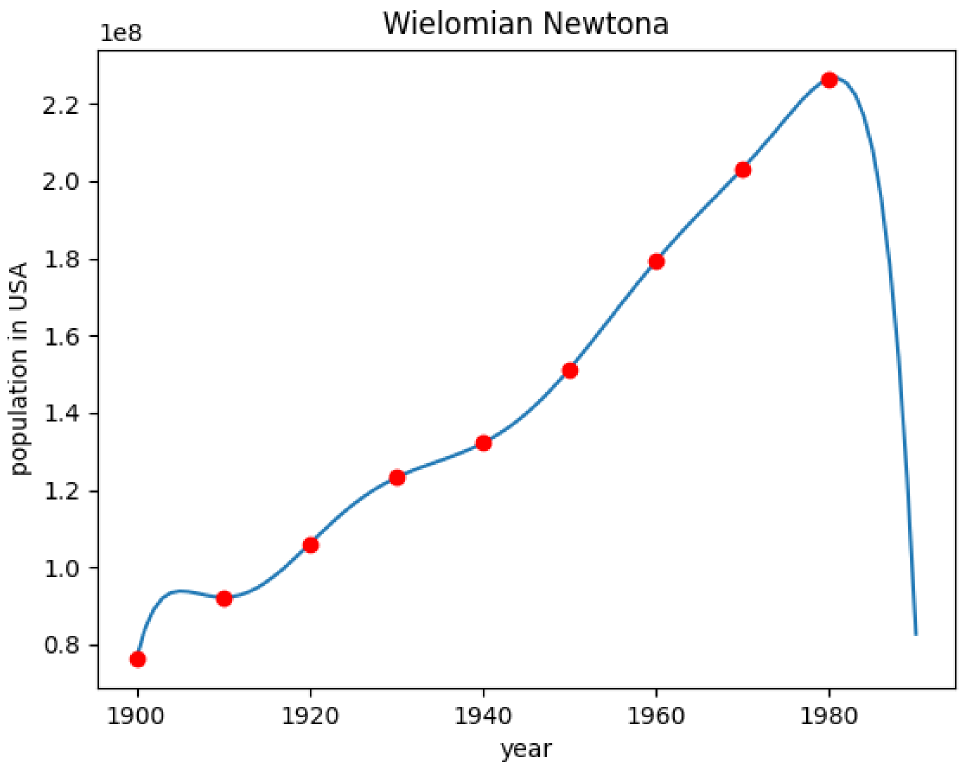
$$p(t) = \sum_{i=1}^{n} f[t_1, ..., t_i] \cdot \pi_i(t) \tag{10}$$

$$f[x_i] = f(x_i) \tag{11}$$

$$f[x_i, x_{i+1}] = \frac{f[x_{i+1}] - f[x_i]}{x_{i+1} - x_i} \tag{12}$$

$$f[x_0, ..., x_i] = \frac{f[x_1, ..., x_i] - f[x_0, ..., x_{i-1}]}{x_i - x_0} \tag{13}$$

Na ich podstawie, korzystajac z funkcji `lambda` oraz rekurencyjnie obliczajac wartości ilorazów różnicowych, wyznaczono wielomian interpolacyjny Newtona. Nastepnie obliczono wartości tego wielomianu dla takiego samego przedziału jak w pierwszym podpunkcie z krokiem co jeden rok. Na podstawie uzyskanych wartości stworzono wykres:




Powyższy wykres jest identyczny jak ten stworzony na podstawie macierzy oraz metoda Lagrange'a, co potwierdza poprawność wyznaczonego wielomianu metoda Newtona.

## Zaokraglone wartości populacji, a wynik interpolacji

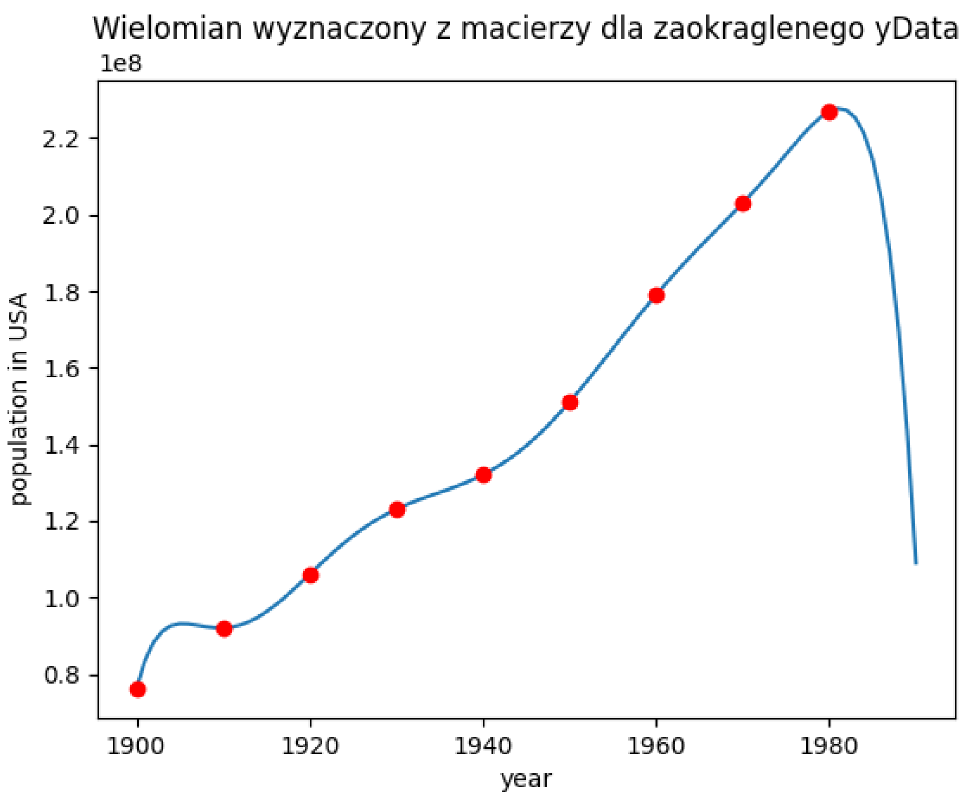
Dane z pierwszej tabeli zaokraglono do jednego miliona:

| Rok | Zaokraglona populacja |
|---|---|
| 1900 | 76 000 000 |
| 1910 | 92 000 000 |
| 1920 | 106 000 000 |
| 1930 | 123 000 000 |
| 1940 | 132 000 000 |
| 1950 | 151 000 000 |
| 1960 | 179 000 000 |
| 1970 | 203 000 000 |
| 1980 | 227 000 000 |

Nastepnie rozwiazano rownanie (5), przy czym macierza $Y$ w tym przypadku były dane z powyższej tabeli. Otrzymane wartości współczynnikow (Wsp-Z) porównano w tabeli z wartościami współczynników obliczonych dla niezaokraglonych danych (Wsp-NZ).

| Numer wspolczynnika | Wsp-NZ | Wsp-Z |
|---|---|---|
| 1 | $1.32164569 \cdot 10^8$ | $1.32000000 \cdot 10^8$ |
| 2 | $4.61307656 \cdot 10^7$ | $4.59571429 \cdot 10^7$ |
| 3 | $1.02716315 \cdot 10^8$ | $1.00141270 \cdot 10^8$ |
| 4 | $1.82527130 \cdot 10^8$ | $1.81111111 \cdot 10^8$ |
| 5 | $-3.74614715 \cdot 10^8$ | $-3.56755556 \cdot 10^8$ |
| 6 | $-3.42668456 \cdot 10^8$ | $-3.38488889 \cdot 10^8$ |
| 7 | $6.06291250 \cdot 10^8$ | $5.70311111 \cdot 10^8$ |
| 8 | $1.89175576 \cdot 10^8$ | $1.86920635 \cdot 10^8$ |
| 9 | $-3.15180235 \cdot 10^8$ | $-2.94196825 \cdot 10^8$ |

Na podstawie tych współczynników wyznaczono wielomian interpolujacy zaokraglone punkty i wykonano jego wykres:



Jak można zauważyć, wykres ten jest bardzo podobny do wykresów powstałych na podstawie niezaokraglonych danych. Ponadto tabela znajdujaca sie powyżej pokazuje, że wartości otrzymanych współczynników sa bardzo zbliżone do wartości odpowiadajacych im współczynników obliczonych z niezaokraglonych danych.