Maciej Borowiec
09.04.2025

# Laboratorium 5
# Aproksymacja

# 1. Wprowadzenie

Celem ćwiczenia było zrobienie dwóch zadań dotyczących aproksymacji średniokwadratowej.

W pierwszym zadaniu należało dokonać aproksymacji średniokwadratowych punktowych dla wielomianów o różnych stopniach i porównać jakość tych aproksymacji.

W drugim zadaniu należało dokonać aproksymacji średniokwadratowej ciągłej danej funkcji.

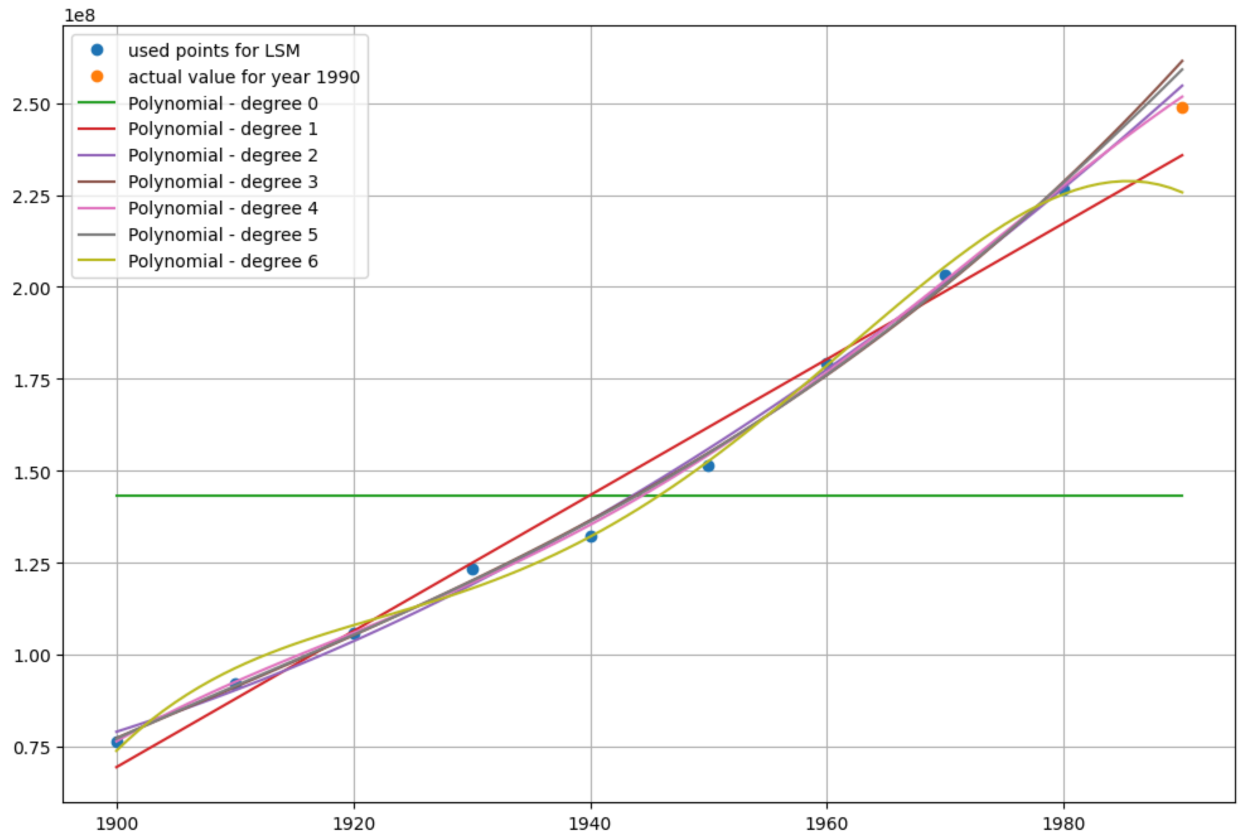
# 2. Zadania

## 2.1. Zadanie 1.

W tym zadaniu należało dokonać aproksymacji średniokwadratowej punktowej na danych populacji Stanów Zjednoczonych wielomianami stopniów $m,\ 0 \leq m \leq 6$. Za dane przyjęto te same dane co w laboratorium trzecim. Dla każdego wielomianu należało dokonać ekstrapolacji do roku 1990. Poniżej przedstawiona jest tabela z błędami względnymi otrzymanej ekstrapolacji oraz wykres otrzymanych wielomianów.

| **Prawdziwa wartość populacji w 1990r.** | 248 709 873 | |
|---|---|---|
| **Stopień wielomianu** | **Estymowana populacja** | **Błąd względny** |
| 0 | 143 369 177 | 0.4235 |
| 1 | 235 808 109 | 0.0519 |
| 2 | 254 712 945 | 0.0241 |
| 3 | 261 439 111 | 0.0512 |
| 4 | 251 719 359 | 0.0121 |
| 5 | 259 115 342 | 0.0418 |
| 6 | 225 702 990 | 0.0925 |

**Tab. 1.** Błędy względne ekstrapolacji wielomianów




**Rys. 1.** Wykres otrzymanych wielomianów z aproksymacji

Jak widać z tabeli (tab. 1) i wykresu (rys. 1) najmniejszy błąd dla ekstrapolacji zaszedł w przypadku wielomianu stopnia czwartego. Jest tak, gdyż wielomiany o małych stopniach nie są w stanie dobrze uwzględnić zmienności danych, a wielomiany o dużych stopniach uwzględnia szum i błędy danych.

## Kryterium informacyjne Akaikego

Kryterium informacyjne Akaikego jest estymatorem błędu dla aproksymacji. Dzięki niemu możemy z pewną dozą pewności przewidzieć najlepszy stopień wielomianu jaki należy użyć do aproksymacji - im mniejsza wartość, tym mniejszy błąd aproksymacji. Kryterium definiujemy następująco:

$$AIC = 2k + n\ln\left(\frac{\sum_{i=1}^{n}[y_i - \hat{y}(x_i)]^2}{n}\right)$$, gdzie:

- $k = m + 1$ ($m$ to stopień wielomianu)
- $n$ to liczba punktów użytych do aproksymacji
- $y_i - \hat{y}(x_i)$ to różnice wartości prawdziwych od wartości aproksymowanych dla punktów użytych do aproksymacji ($i = 1, 2, \ldots, n$)

Jednak w naszym przypadku rozmiar próbki jest niewielki ($n/k < 40$) zatem należy zmodyfikować kryterium o składnik korygujący:

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1}$$

Poniżej przedstawione zostały wartości $AIC_c$ dla otrzymanych wielomianów.

| **Stopień wielomianu** | **Wartość** $AIC_c$ |
|---|---|
| 0 | 321.011 |
| 1 | 289.056 |
| 2 | 279.453 |
| 3 | 284.880 |
| 4 | 292.672 |
| 5 | 319.828 |
| 6 | 392.074 |

**Tab. 2.** Wartości $AIC_c$ dla danych wielomianów

Jak widać z danych tabel (tab. 1, tab. 2) wartość kryterium informacyjnego Akaikego nie wyznaczyła w tym przypadku najdokładniejszego wielomianu. Może to być spowodowane błędem obliczeniowym, bądź po prostu nie stuprocentową dokładnością tego kryterium. Jednakże wyznaczony wielomian (stopnia drugiego) w rzeczywistości nadal jest drugą najlepszą aproksymacją, a różnica błędów jest na tyle mała, że użycie prostszego modelu będzie tu bardziej preferowane.

## 2.2. Zadanie 2.

W tym zadaniu należało dokonać aproksymacji średniokwadratowej ciągłej dla funkcji $f(x) = \sqrt{x}$ na przedziale $[0; 2]$ wielomianem drugiego stopnia. Należało wykorzystać wielomiany Czebyszewa:

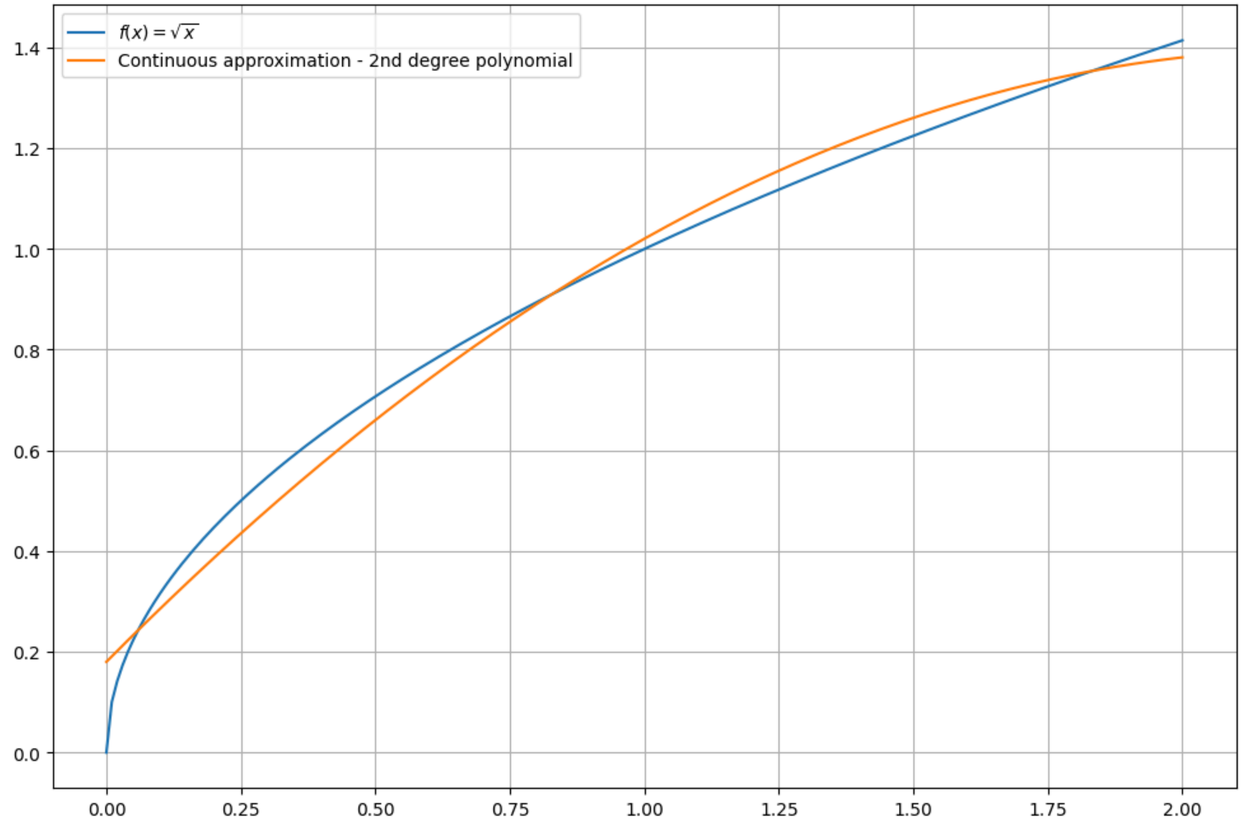
$$T_k(x) = 2x \cdot T_{k-1}(x) - T_{k-2}(x)$$

$$T_0 \equiv 1, T_1 \equiv x$$

Są one ortogonalne oraz:

$$< T_i, T_i > = \pi \cdot 2^{0^i - 1}$$

Poniżej przedstawiony jest wykres funkcji $f$ oraz otrzymanej aproksymacji $p$.




**Rys. 2.** Wykres funkcji $f$ oraz aproksymacji $p$ tej funkcji

Możemy również obliczyć błąd tej funkcji jako całka po wartościach bezwzględnych różnicy wartości funkcji $f$ oraz wartości aproksymacji $p$.

$$\int_0^2 |f(x) - p(x)| dx \approx 0.0628$$

Obliczmy też błąd względny:

$$\int_0^2 \frac{|f(x) - p(x)|}{f(x)} dx \approx 0.1196$$

Z podanego wykresu (rys. 2) oraz obliczonej wartości błędu możemy stwierdzić, że otrzymana aproksymacja jest w miarę zgodna z rzeczywistością.

# 5. Podsumowanie

Zadania wykonane w tym laboratorium pokazały użycie średniokwadratowej aproksymacji. Z pierwszego zadania łatwo wywnioskować, że stopień użytego wielomianu do aproksymacji ma duży wpływ na dokładność tej aproksymacji. Kryterium informacyjne Akaikego jest dobrym sposobem na wyznaczenie dobrego wielomianu do aproksymacji. Nie należy zakładać, że zawsze wybierze wielomian najbardziej dokładny, jednakże uwzględniając użycie prostszego modelu, ten wielomian nadal może być preferowany. Zadanie drugie pokazało, że aproksymacja wersji ciągłej może być dość skutecznie użyta do aproksymacji danych funkcji. Nie będzie on niebywale dokładny, ale przedstawienie skomplikowanej funkcji za pomocą dużo prostszego od niej wielomianu może być bardzo przydatne.

# 6. BIbliografia


- Materiały zamieszczone wraz zadaniem
- Materiał z wikipedii nt. kryterium informacyjnego Akaikego