Sprawozdanie NUM8

Szymon Tomaszewski

28 grudnia 2024

1. Wstęp

1.1. Treść zadania

Zaproponuj wielomian uogólniony w postaci , gdzie ilość parametrów m. Następnie:

Zaproponuj wielomian uogólniony w postaci $F(x) = \sum_{j=1}^{m} a_j \phi_j(x)$, gdzie ilość parametrów $m \geq 3$, a $\phi_j(x)$ są pewnymi funkcjami. Zdefiniuj siatkę punktów x_i oraz (dla pewnego ustalonego zestawu parametrów a_j) wygeneruj dane w postaci $\{(x_i, y_i)\}$, gdzie i = 1,...,n, a $y_i = F(x_i) + \delta y_i$. Zaburzenia δyi należy losować z rozkładu normalnego z odchyleniem standardowym σ.

(a) Znajdź wartości współczynników a_j, dla których funkcja F(x) najlepiej opisuje zaburzone dane w sensie metody najmniejszych kwadratów. Rezultat przedstaw graficznie dla kilku wyborów wielkości siatki n oraz odchylenia standardowego σ.

(b) Przeanalizuj różnicę pomiędzy wcześniej ustalonymi współczynnikami, a ich wartościami uzyskanymi w procedurze aproksymacji przeprowadzonej dla zaburzonych danych

UWAGA:Rozwiązując to zadanie nie można korzystać z procedur bibliotecznych służących do aproksymacji. Poza tym, użycie procedur z zakresu algebry liniowej jest dozwolone.

1.2. Wprowadzenie

Termin aproksymacja w tym przypadku będzie związany z aproksymacją punktów, mając n punktów staramy się znaleźć funkcję należącą do znanej kategorii, która będzie przebiegać możliwie "najbliżej" tych punktów. Należy podkreślić, że funkcja jest znana co do swojego kształtu, to znaczy, że np. jest wielomianem określonego stopnia złożonym z kombinacji funkcji trygonometrycznych a nieznane są jej parametry.

Posiadamy n par punktów $\{(x_i, y_i)\}$ gdzie, x_i jest dokładną wartością argumentu a y_i jest właśnie obarczoną jakimś błędem wartością funkcji w tym argumencie.

Każdej obarczonej błędem wartości y_i odpowiada wartość teoretyczna $\overline{y_i}$.

Przyjmujemy, że wartość teoretyczna jest kombinacją liniową pewnych znanych funkcji

$$\overline{y_i} = a_1 * f_1(x_i) + a_2 * f_2(x_i) + \dots + a_s * f_s(x_i)$$

Zatem możemy to przedstawić w postaci $\overline{y_i} = Ap$

gdzie

$$\mathbf{A} = \begin{bmatrix} f_1(x_1) & f_2(x_1) & f_3(x_1) & \cdots & f_s(x_1) \\ f_1(x_2) & f_2(x_2) & f_3(x_2) & \cdots & f_s(x_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & f_3(x_n) & \cdots & f_s(x_n) \end{bmatrix} \in \mathbb{R}^{n \times s} \qquad \mathbf{p} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_s \end{bmatrix} \in \mathbb{R}^s$$

Problem więce sprowadza się do znalezienia najlepszego wektora parametrów p.

Metoda najmniejszych kwadratów polega na dopasowaniu danych w taki sposób aby zminimalizować sumę kwadratów błędów. Wówczas problem minimalizacji przedstawiamy jako $Q = \sum_{i=1}^{n} (y_i - \overline{y_i})^2$ gdzie p spełnia układ równań $A^T A p = A^T y$ gdzie y to wektor wartości zaburzonych.

Generowanie danych z szumem odbywa się poprzez $y_i = F(x_i) + \delta y_i$ gdzie δy_i to zaburzenie losowanie z rozkładu normalnego z odchyleniem standardowym σ.

2. Kod

Na potrzeby zadania została zdefiniowana funkcja

$F(x) = \sum_{j=1}^{5} a_j \phi_j(x)$ gdzie

$\phi_1(x) = e^{-5x} * sin(40x)$ $\quad a_1 = -0.55$

$\phi_2(x) = ln(x + 1.51)$ $\quad a_2 = 1.5$

$\phi_3(x) = sin(cos(3x))$ $\quad a_3 = 2$

$\phi_4(x) = tanh(2x)$ $\quad a_4 = -1.4$

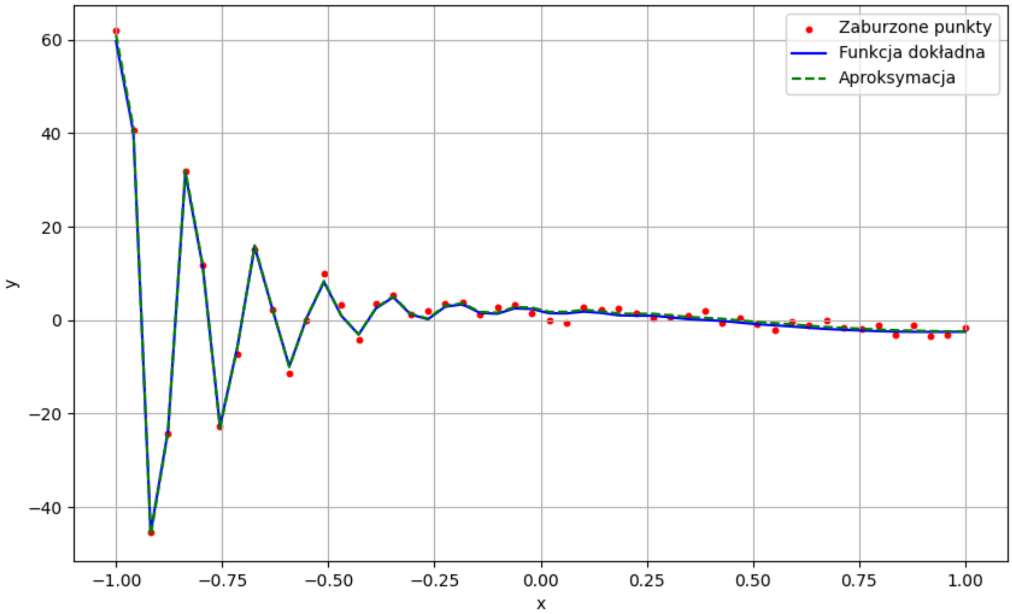
$\phi_5(x) = cos(e^x + 1)$ $\quad a_5 = 1.0$

3. Wyniki

a) Dla n=50 i σ=1

Dokładne: [-0.55, 1.5, 2.0, -1.4, 1.0]

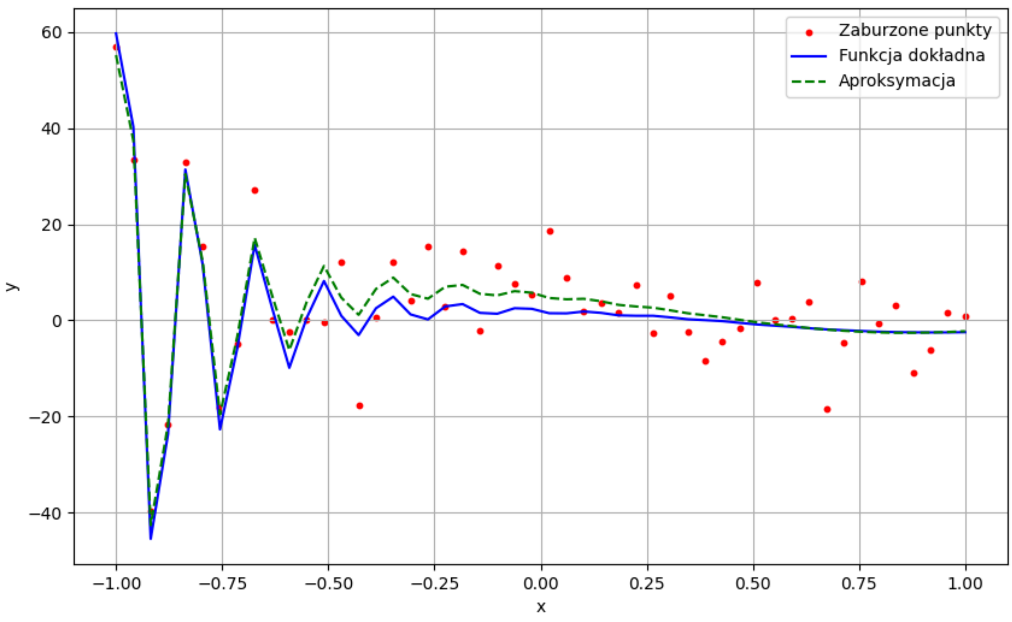
Aproksymacja: [-0.55862532 -0.94488018 2.32184235 -0.70689474 -0.7447105]



b) Dla n=50 i σ=7

Dokładne: [-0.55, 1.5, 2.0, -1.4, 1.0]

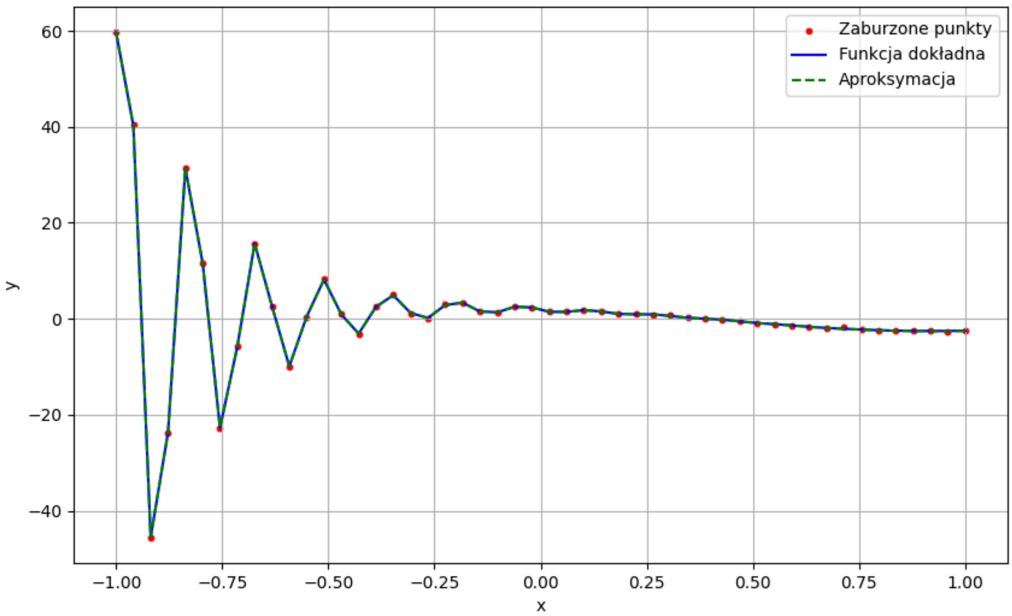
Aproksymacja: [-0.54577112 6.6884157   1.5245289 -1.07193105 7.52062754]



c) Dla n=50 i σ=0.1

Dokładne: [-0.55, 1.5, 2.0, -1.4, 1.0]

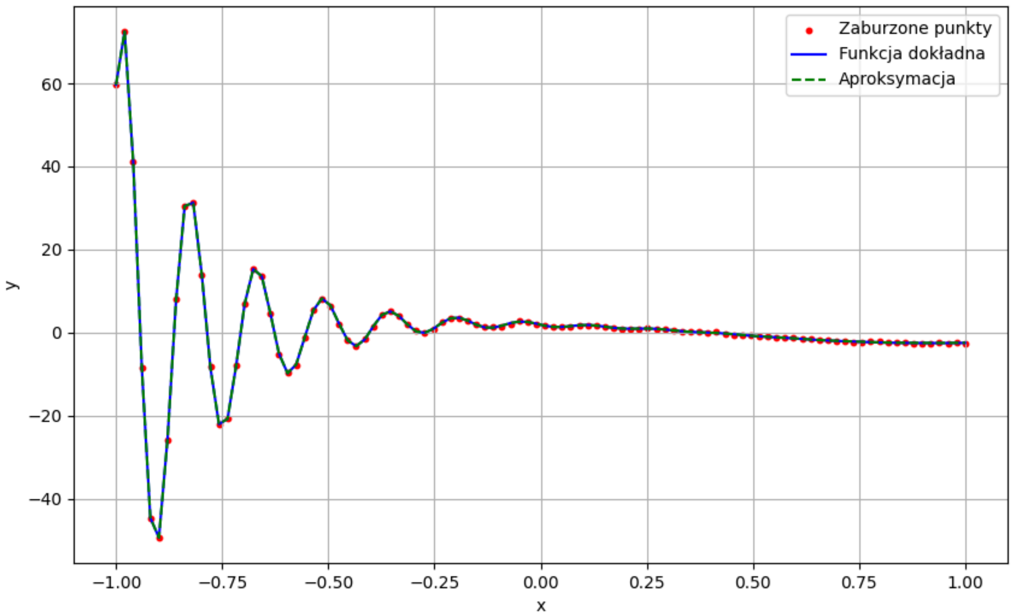
Aproksymacja: [-0.54950103 1.41614115 1.98606794 -1.37954673 0.92193047]



d) Dla n=100 i σ=0.5

Dokładne: [-0.55, 1.5, 2.0, -1.4, 1.0]

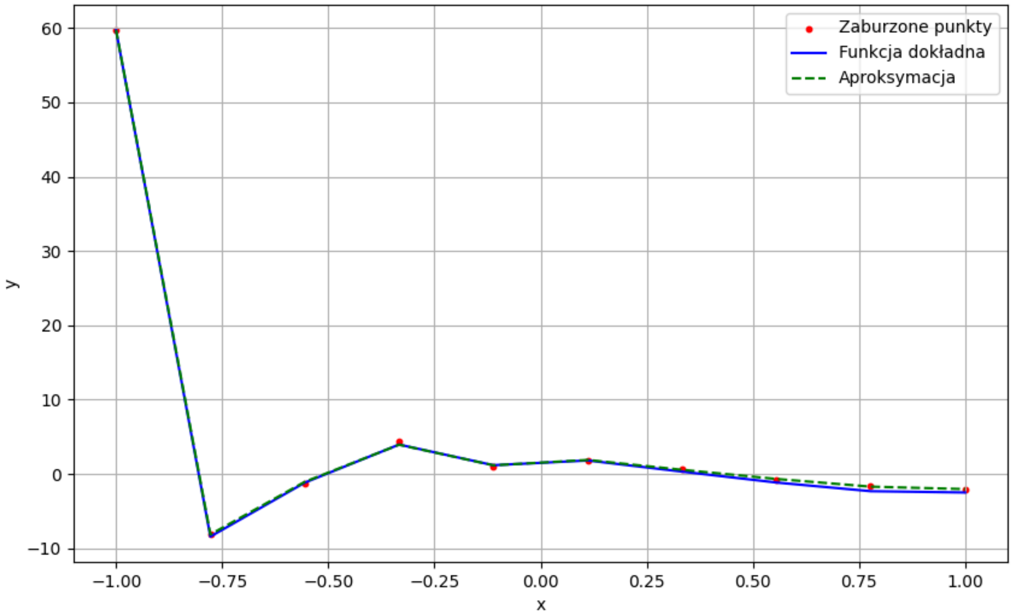
Aproksymacja: [-0.54999229 1.56595791 1.9826567 -1.41601874 1.02693197]



e) Dla n=10 i σ=0.4

Dokładne: [-0.55, 1.5, 2.0, -1.4, 1.0]

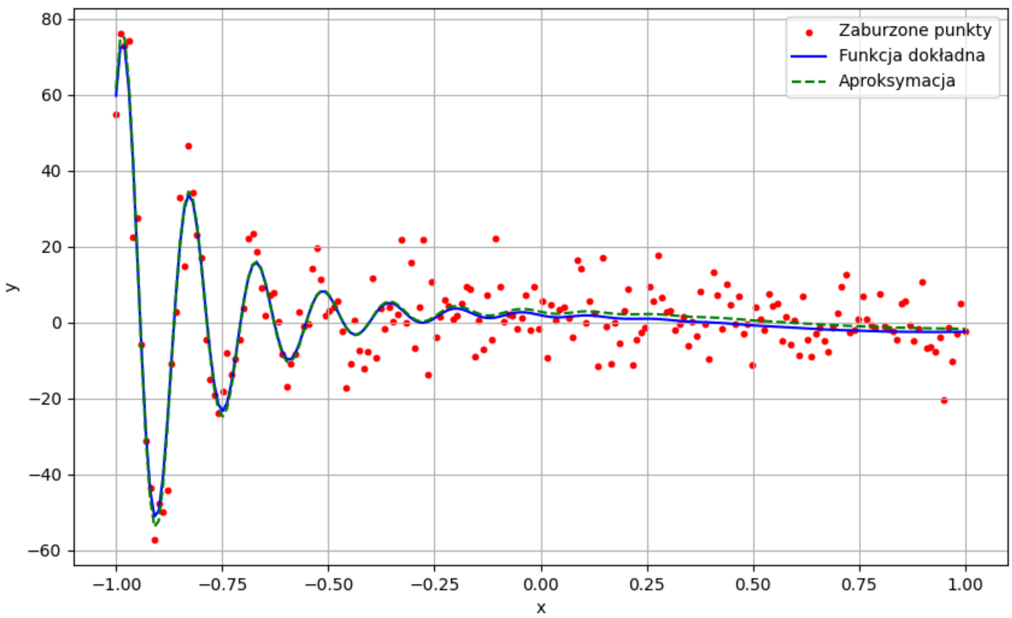
Aproksymacja: [-0.54681409 1.02633257 1.78872899 -1.46785912 0.06789482]



f) Dla n=200 i σ=7

Dokładne: [-0.55, 1.5, 2.0, -1.4, 1.0]

Aproksymacja: [-0.57276098 1.12314166 2.47371059 -1.21389341 -0.59726564]



4. Dyskusja i Wnioski

4.1. Wpływ σ

Odpowiada za szum w danych. Wraz z wzrostem dane coraz bardziej oddalają się od wartości teoretycznej, metoda najmniejszych kwadratów zaczyna działać gorzej.

Dla **σ<1** uzyskane wyniki są bliskie teoretycznym.

4.2. Wpływ n

Zwiększając liczbę punktów aproksymacja jest coraz bardziej zbliżona do teoretycznej funkcji. Dodatkowo zwiększenie liczby punktów może kompensować duże **σ.**

Dla małych wartości n współczynnika a_i diametralnie różni się od wartości dokładnej.

Podsumowując metoda najmniejszych kwadratów nie jest dobrym nardzędziem do aproksymacji jeśli dane na których operujemy mają duże błędy pomiarowe lub mamy mało punktów owych pomiarów.