# Kacper Kula

Zadanie numeryczne nr 8

## *Omówienie*

Problem polega na stworzeniu algorytmu aproksymującego działającego na podstawie pewnych danych. Zadany jest zbiór punktów $(x, y)$ oraz funkcja $F(x)$, która posłuży do modelowania tych punktów:
$F(x) = a * x^2 + b * \sin(x) + c * \cos(5x) + d * e^{-x}$

**(a)** W pierwszej części zadania należy znaleźć wartości współczynników a-d, które będą najlepiej opisywać dane w sensie metody najmniejszych kwadratów. Rezultat należy przedstawić graficznie, a przy rozwiązywaniu zadania nie wolno korzystać z funkcji bibliotecznych służących do aproksymacji.

**(b)** W drugiej części zadania należy zaproponować przykładowe funkcje modelujące i wygenerować zbiory postaci zbioru punktów o zaburzonych wartościach, dla różnych wartości zaburzenia $\delta$ i sprawdzić, czy napisany algorytm generuje poprawne, wcześniej ustalone współczynniki.

Narzędziem do napisania programu rozwiązującego zadanie jest język Python oraz biblioteka pythona służąca do tworzenia wykresów – matplotlib, a także biblioteka z zakresu algebry liniowej  - numpy.

# *Aproksymacja*

## *Wstęp*

W naukach przyrodniczych często wykonywane są eksperymenty polegające na pomiarach par wielkości, które, jak przypuszczamy, są ze sobą powiązane jakąś zależnością funkcyjną. Sensownym posunięciem jest znalezienie takiej krzywej, która możliwie najlepiej przybliża te punkty doświadczalne.

Do przybliżania funkcji, jak wiemy z poprzedniego zadania, służy nam również interpolacja. Jednak, kiedy do przybliżenia mamy funkcję między 20 danymi punktami i spodziewamy się zależności liniowej, to tworzenie wielomianu 19-ego stopnia do przybliżenia tej zależności funkcyjnej jest kompletnie bezsensowne.

Aby temu zapobiec, musimy zrezygnować z jednego z warunków interpolacji, mianowicie, że funkcja musi przechodzić przez wszystkie dane punkty. Wystarczy nam gwarancja, że funkcja będzie przebiegać jak najbliżej tych punktów i jednocześnie będzie najlepiej do nich dopasowana – w tym pomoże nam zagadnienie aproksymacji.

## *Zagadnienie aproksymacji metodą najmniejszych kwadratów*

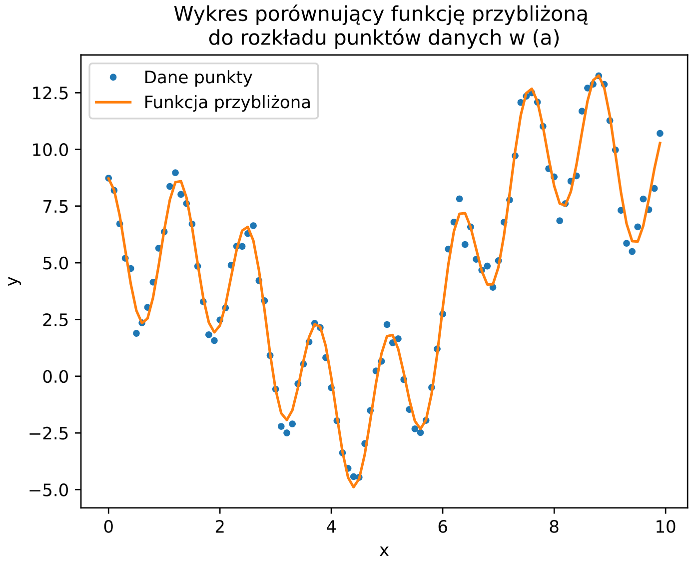
Ogólnie rzecz ujmując, aproksymacja polega na przybliżaniu pewnej funkcji $f(x)$, w obszarze jej określoności, za pomocą innej, prostszej funkcji $F(x)$ przybliżającej określonej w tym samym obszarze, której wartości zależą od pewnej liczby parametrów. Najczęściej jako funkcje $F(x)$ stosuje się wielomiany uogólnione w postaci $F(x)=a_1*\varphi_1(x)+a_2*\varphi_2(x)+...+a_n*\varphi_n(x)$, gdzie $\varphi_i(x)$ to funkcje bazowe.

Poszukiwanie parametrów takiej funkcji, przechodzącej jak najbliżej wszystkich danych punktów, polega na minimalizacji sumy $S(a)=\sum_i [y_i-F(x_i)]^2$. Jej minimum będzie określało krzywą najlepiej dopasowaną.

Metoda ta nosi nazwę metody najmniejszych kwadratów.

# *Podpunkt (a)*

## *Funkcja modelująca F(x)*




Dla funkcji modelującej $F(x) = a*x^2 + b*\sin(x) + c*\cos(5x) + d*e^{-x}$ znalazłem wartości współczynników a-d i sporządziłem wykres funkcji $F(x)$ z podstawionymi, przybliżonymi wartościami współczynników.

$a = 0.10093369$, $b = 4.02305946$, $c = 3.08874327$, $d = 5.63283974$

Jak można zauważyć na wykresie, natura funkcji $F(x)$ została dobrze oddana przez przybliżoną funkcję, co potwierdza skuteczność opisanej metody aproksymacyjnej.

## *Podpunkt (b)*

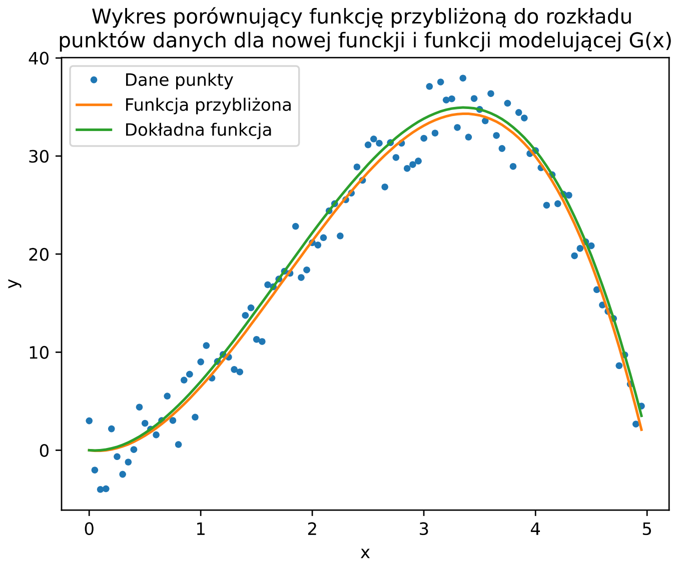
### *Funkcja modelująca G(x)*

Aby potwierdzić skuteczność opisanej metody aproksymacyjnej sprawdziłem ją również dla funkcji modelującej $G(x)=a*x^4+b*x^3+c*x^2+d*x$ o współczynnikach: $a=0.01$, $b=-2$, $c=10$, $d=-1$.

W wszystkich przyszłych eksperymentach generowane punkty były od siebie równoodległe.

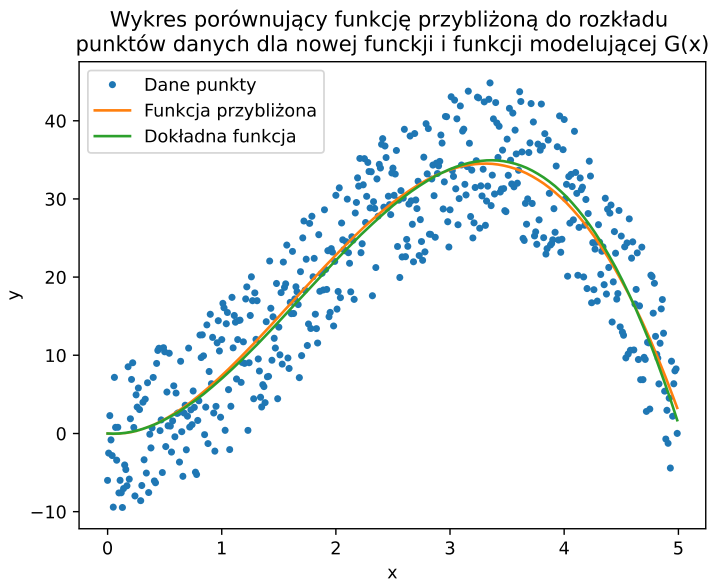
Wygenerowałem 100 zaburzonych punków z przedziału $x\in[0,5]$, $y=G(x)+\delta$ dla $\delta\in[-4,4]$.

Rezultat metody aproksymacyjnej przedstawiłem na wykresie.



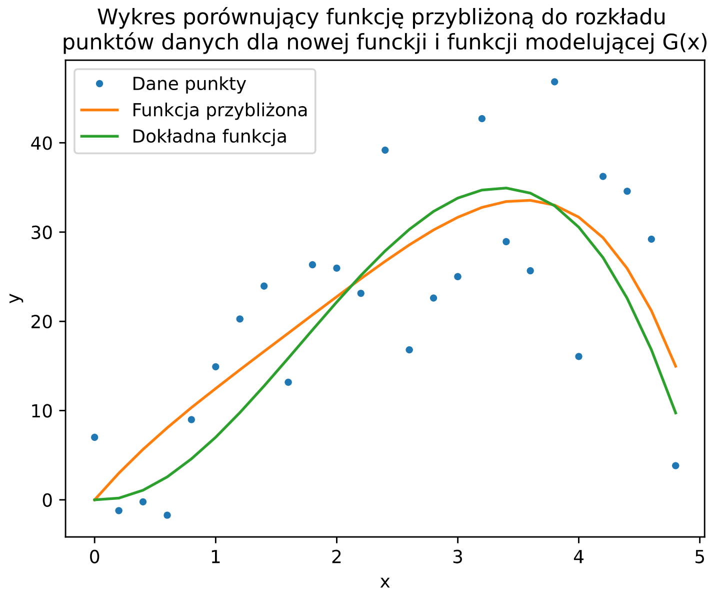

Przybliżona funkcja jest bardzo dokładna i zachowuje się niemal identycznie do funkcji dokładnej. Zastosowanie metody aproksymacyjnej przyniosło pozytywne rezultaty.

Poeksperymentowałem również z liczbą punktów oraz z wielkością zaburzenia $\delta$.
Poniższy wykres przedstawia porównanie dla 500 punktów z przedziału $x\in[0,5]$ oraz $\delta\in[-10,10]$.



Jak widać przybliżenie funkcji nie straciło na wartości mimo zwiększenia zaburzenia $\delta$.

Porównanie sporządziłem również dla 25 punktów z przedziału $x \in [0,5]$ oraz $\delta \in [-15,15]$.



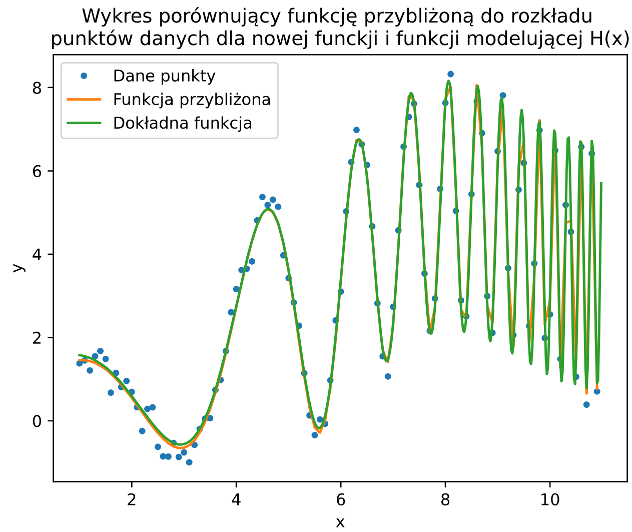

Dla tak znaczącego zaburzenia $\delta$ i niewielkiej liczby punktów nasze przybliżenie okazało się znacznie mniej dokładne, niż dla poprzednich przybliżeń.

### *Funkcja modelująca H(x)*

Wybrałem również funkcję modelującą $H(x)=a*\sin(x)+b*\cos\left(e^{\left(\frac{2}{5}*x\right)}\right)+c*x^{-2}+d*\ln(x)$ dla współczynników o wartościach: $a=1$, $b=3$, $c=\frac{1}{2}$, $d=2$ i sprawdziłem działanie mojej metody aproksymacyjnej.

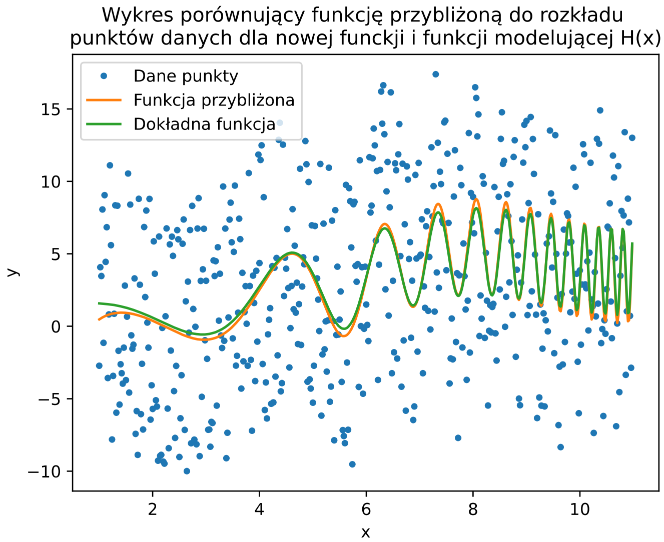
Wygenerowałem 100 zaburzonych punków postaci $x\in[1,11]$, $y=H(x)+\delta$ dla $\delta\in[-5,5]$.

Rezultat metody aproksymacyjnej przedstawiłem na wykresie.



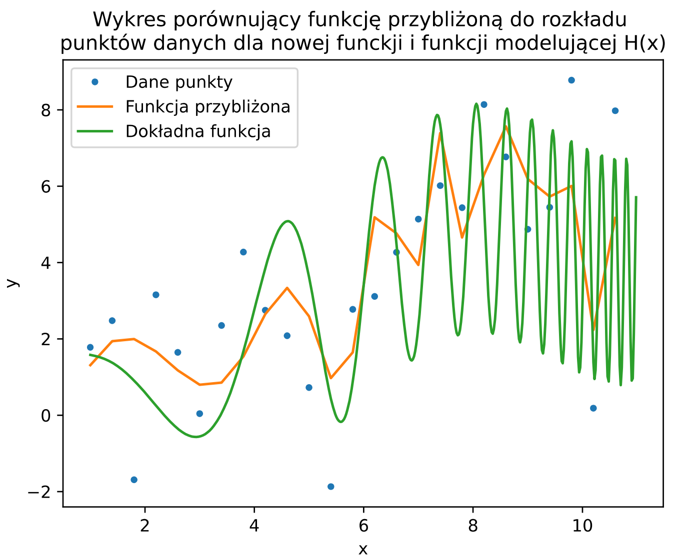

Przybliżona funkcja jest bardzo dokładna i zachowuje się prawie identycznie do funkcji dokładnej. Zastosowanie metody aproksymacyjnej przyniosło pozytywne rezultaty.

Przy tej funkcji również poeksperymentowałem z liczbą punktów oraz z wielkością zaburzenia $\delta$. Poniższy wykres przedstawia porównanie dla 500 punktów z przedziału $x\in[1,11]$ oraz $\delta\in[-10,10]$.



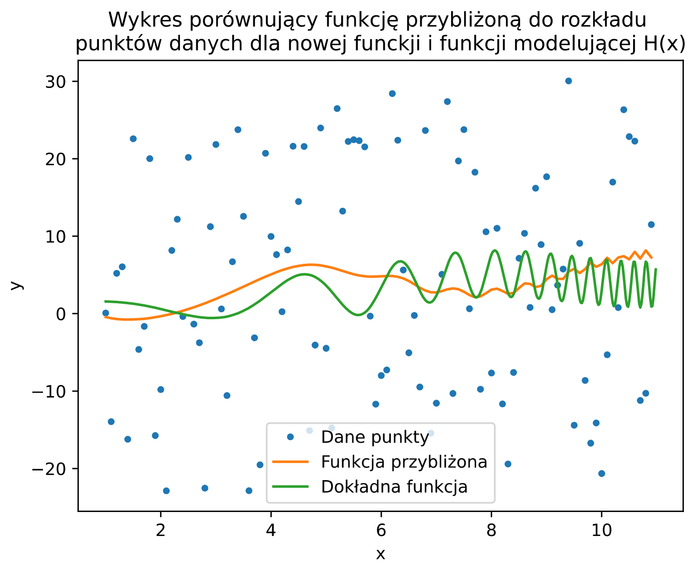
Jak widać przybliżenie funkcji nie straciło wiele na wartości mimo podwojenia zaburzenia $\delta$, chociaż trzeba zauważyć, że to przybliżenie jest gorsze od poprzedniego.

Porównanie sporządziłem również dla 25 punktów z przedziału $x \in [1, 11]$ oraz $\delta \in [-3, 3]$.




Dla tak małej liczby punktów, mimo niewielkiego zaburzenia $\delta$, nasze przybliżenie okazało się poważnie niedokładne.

Aby zauważyć, że nasze przybliżenie jest bardzo czułe na zwiększenie zaburzenia $\delta$. Powtórzyłem aproksymację dla 100 punktów, ale zwiększyłem zaburzenie do $\delta \in [-25, 25]$




Nasze przybliżenie znacznie odbiega od dokładnej funkcji i jest bezwartościowe, mimo stosunkowo dużej liczby punktów.