Laboratorium przedmiotu Metody Numeryczne		
Sprawozdanie nr 2: Aproksymacja		
Data: 19/03/2019	Ćwiczenie wykonał: Bartosz Michałowski, Jan Starczewski	Ćwiczenie prowadził: Dr Inż. Bartosz Chaber
Grupa dziekańska: 1	Ocena:	

Wprowadzenie

Zadanie aproksymacji polega na znalezieniu pewnej funkcji danego rodzaju (np. wielomianu, albo funkcji wykładniczej), która przebiega „jak najbliżej" aproksymowanego zbioru punktów. Aby określić, co oznacza „jak najbliżej" można użyć różnych kryteriów, jednak najczęściej stosowanym jest kryterium minimalizacji błędu średniokwadratowego. Oznacza to, że poszukujemy takiej funkcji aproksymującej $\boldsymbol{f(x)}$, dla której suma odległości wszystkich $\boldsymbol{N}$ aproksymowanych punktów od przebiegu $\boldsymbol{f(x)}$ jest możliwie najmniejsza. Matematycznie zapisujemy to jako funkcję błędu średniokwadratowego (w tym przykładzie wykorzystujemy funkcję kwadratową):

$$\boldsymbol{E(a_0, a_1, a_2) = \sum_{i=1}^{N} \left(a_0 + a_1 x_i + a_2 x_i^2 - y_i\right)^2},$$

gdzie $\boldsymbol{f(x) = a_0 + a_1 x + a_2 x^2}$. Warto zwrócić uwagę, że dziedziną funkcji błędu średniokwadratowego są współczynniki funkcji aproksymującej.

Znalezienie takiego zbioru współczynników $\boldsymbol{a_0, a_1, a_2}$ dla których $\boldsymbol{E(a_0, a_1, a_2)}$ jest możliwie najmniejsze, polega na wykorzystaniu pochodnych funkcji błędu względem jej argumentów. Najmniejszą wartość błędu otrzymamy, gdy $\boldsymbol{\frac{\partial E}{\partial a_0} = \frac{\partial E}{\partial a_1} = \frac{\partial E}{\partial a_2} = 0}$. Jest to analogiczna metoda do poszukiwania minimum funkcji jednej zmiennej poprzez znalezienie miejsca, gdzie pochodna tej funkcji równa się zeru. Należy zwrócić uwagę, że ogólnie szukając współczynników, dla których pochodna błędu jest zero możemy znaleźć ekstremum tej funkcji (minimum, albo maksimum), natomiast intuicyjnie wiemy, że nie ma konkretnej wartości maksymalnego błędu (może być on nieskończenie wielki). Z tego powodu zakładamy, że znaleziony zestaw współczynników zapewnia nam minimalny błąd średniokwadratowy – czyli funkcję najlepiej opisującą aproksymowany zbiór punktów.

Na podstawie $\boldsymbol{\frac{\partial E}{\partial a_0} = \frac{\partial E}{\partial a_1} = \frac{\partial E}{\partial a_2} = 0}$ (w przypadku aproksymacji funkcją kwadratową) powstaną nam trzy równania. Możemy zapisać je w postaci macierzowej $\boldsymbol{A \cdot x = b}$, gdzie $\boldsymbol{A}$ jest macierzą 3×3, której wartości są wyliczone na podstawie zbioru punktów podlegających aproksymacji, $\boldsymbol{x = [a_0, a_1, a_2]^T}$ jest wektorem kolumnowym nieznanych współczynników, natomiast $\boldsymbol{b}$ jest również wektorem kolumnowym o trzech elementach, wyliczonych na podstawie zbioru punktów podlegających aproksymacji. Wartości elementów $\boldsymbol{A}$ i $\boldsymbol{b}$ zależą od rodzaju funkcji aproksymującej oraz od współrzędnych ze zbioru punktów aproksymowanych. Wyznaczenie wzorów i programu budującego takie macierze pozostawia się do wykonania studentom. Wymagana wiedza to: umiejętność obliczenia pochodnej złożonej $\boldsymbol{\frac{\partial E}{\partial a_j}}$, umiejętność rozbicia sumy $\boldsymbol{\sum_{i=1}^{N} b\, x_i + c\, x_i^2}$ jako $\boldsymbol{\sum_{i=1}^{N} b\, x_i + \sum_{i=1}^{N} c\, x_i^2}$ oraz

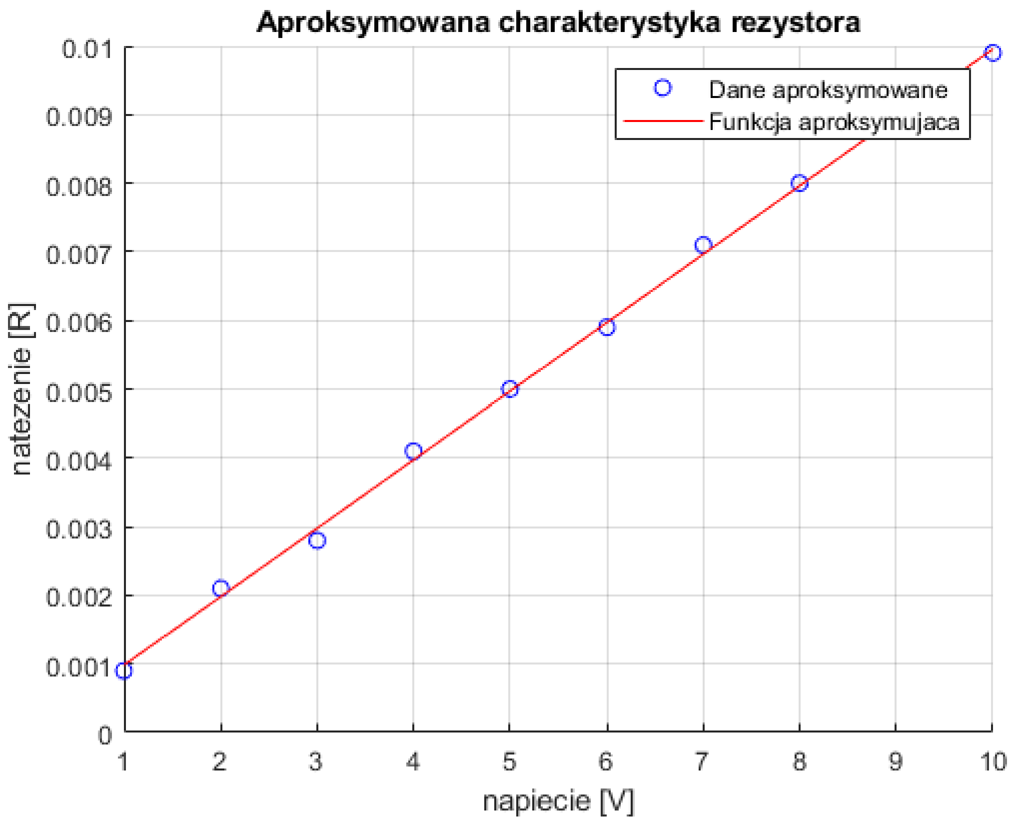
umiejętność zapisania sumy $\boldsymbol{a_0 p_0 + a_1 p_1 + a_2 p_2 - r = 0}$ w postaci macierzowej $[\boldsymbol{p_0}, \boldsymbol{p_1}, \boldsymbol{p_2}] \cdot [\boldsymbol{a_0}, \boldsymbol{a_1}, \boldsymbol{a_2}]^{\boldsymbol{T}} = \boldsymbol{r}$.
W powyższym przykładzie, poszukiwaliśmy funkcji wielomianowej. Funkcje zmiennej $\boldsymbol{x}$ przy współczynnikach nazywamy *funkcjami bazowymi*. Wielomianowe funkcje bazowe nie są jedynymi, które możemy wykorzystać. Innymi funkcjami bazowymi mogą być wielomiany Legendre, wielomiany Czebyszewa czy funkcje trygonometryczne.

Aproksymację stosujemy wtedy, gdy posiadamy zbiór punktów, które chcemy przybliżyć ciągłą funkcją o znanej postaci (natomiast nieznanych współczynnikach). Jest to częsty przypadek w inżynierskich zagadnieniach, kiedy posiadamy dane pomiarowe (obarczone niepewnościami pomiarowymi). Przykładem z elektrotechniki może być wyznaczenie rezystancji elementu obwodu elektrycznego na podstawie serii pomiarów spadku napięcia na tym elemencie przy znanym prądzie płynącym przez niego. Analogiczne zadanie informatyczne, korzystające z aproksymacji to analiza czasu działania algorytmu w celu przybliżenia jego złożoności czasowej.

Przykład zadania nr 1
Celem tego ćwiczenia wykorzystanie **samodzielnie** zaimplementowanej aproksymacji średniokwadratowej do znalezienia współczynników funkcji liniowej $\boldsymbol{f(x) = a_0 + a_1 x}$ aproksymującej zbiór danych `dane_apx0.mat` znajdującego się w iSod. W tym zbiorze danych mamy zdefiniowaną macierz `s` o wymiarach 30×2, gdzie każdy wiersz opisuje dwuwymiarowy punkt (pierwsza kolumna zawiera wartość współrzędnej $\boldsymbol{x}$, druga – $\boldsymbol{y}$).

Po wyznaczeniu współczynników funkcji aproksymującej należy: wyznaczyć wartość błędu średniokwadratowego oraz wyliczyć wybrany przez prowadzącego parametr (np. wartość funkcji aproksymującej dla wybranego $\boldsymbol{x}$, punkt przecięcia wyznaczonej prostej z osią $\boldsymbol{X}$ czy kąt nachylenia funkcji liniowej).

Należy zamieścić wykres (z opisami osi i legendą) zawierający: punkty ze zbioru aproksymowanego oznaczone wyłącznie punktami (format `'o'`, `'*'`, `'.'` lub `'s'`) oraz przebieg funkcji aproksymującej, której wartości wyznaczone zostaną dla 100 współrzędnych $\boldsymbol{x}$ (równooddalonych) z zakresu współrzędnych $\boldsymbol{x}$ wczytanego zbioru danych. We wnioskach należy opisać swoje obserwacje dotyczące wierności aproksymacji.



RYSUNEK 1. Wynik aproksymacji funkcją liniową zbioru danych `dane_resistor.mat`

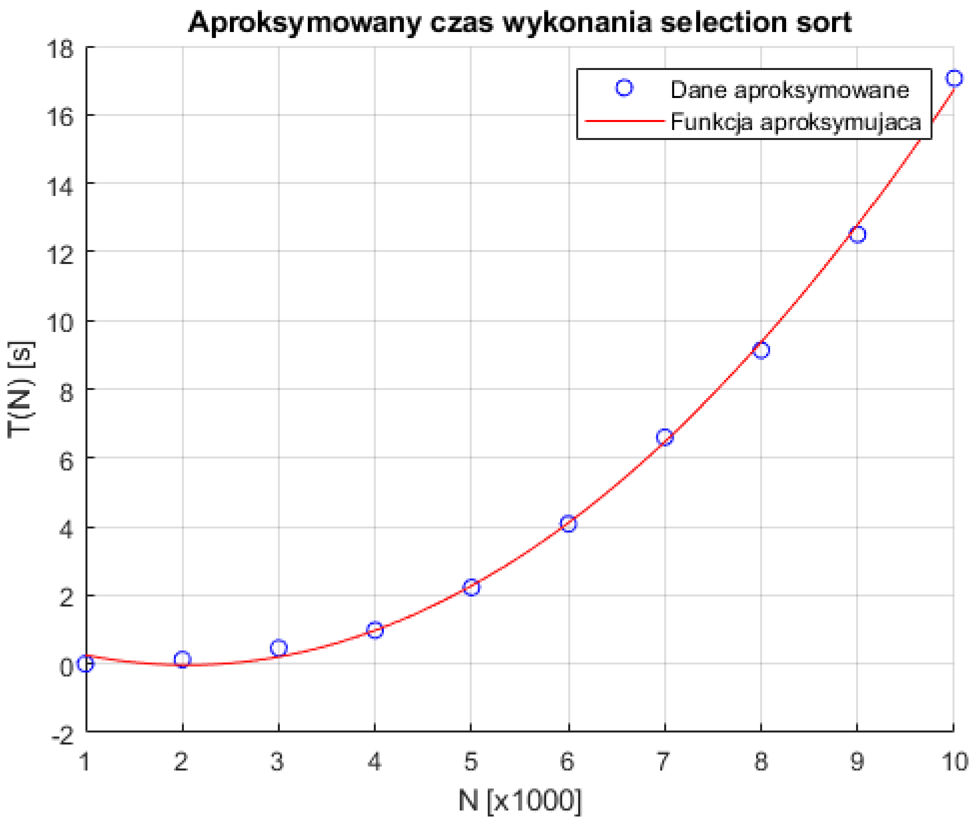
OBLICZONA WARTOŚĆ REZYSTANCJI JEST RÓWNA 1003.0395 Ω

Przykład zadania nr 2

Celem tego ćwiczenia wykorzystanie **samodzielnie** zaimplementowanej aproksymacji średniokwadratowej do znalezienia współczynników funkcji liniowej $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{a_0} + \boldsymbol{a_1}\boldsymbol{x} + \boldsymbol{a_2}\boldsymbol{x}^2$ aproksymującej zbiór danych `dane_apx3.mat` znajdującego się w iSod. W tym zbiorze danych mamy zdefiniowaną macierz `s` o wymiarach 30×2, gdzie każdy wiersz opisuje dwuwymiarowy punkt (pierwsza kolumna zawiera wartość współrzędnej $\boldsymbol{x}$, druga – $\boldsymbol{y}$).

Po wyznaczeniu współczynników funkcji aproksymującej należy: wyznaczyć wartość błędu średniokwadratowego oraz wyliczyć wybrany przez prowadzącego parametr (np. wartość funkcji aproksymującej dla wybranego $\boldsymbol{x}$ czy współrzędne wierzchołka paraboli).

Należy zamieścić wykres (z opisami osi i legendą) zawierający: punkty ze zbioru aproksymowanego oznaczone wyłącznie punktami (format `'o'`, `'*'`, `'.'` lub `'s'`) oraz przebieg funkcji aproksymującej, której wartości wyznaczone zostaną dla 100 współrzędnych $\boldsymbol{x}$ (równooddalonych) z zakresu współrzędnych $\boldsymbol{x}$ wczytanego zbioru danych. We wnioskach należy opisać swoje obserwacje dotyczące wierności aproksymacji.



RYSUNEK 2. Wynik aproksymacji funkcją kwadratową zbioru danych `dane_selectionsort.mat`

APROKSYMOWANY CZAS WYKONANIA SORTOWANIA DLA N = 18 000 JEST RÓWNA 67.4765 S

Co podlega głównej ocenie

Najważniejszym elementem oceny jest umiejętność formułowania **własnych** wniosków z przeprowadzonych ćwiczeń oraz zdolność do samodzielnego zaimplementowania wskazanych metod aproksymacji. Kolejnymi elementami oceny są również: staranność przygotowanego kodu oraz zamieszczonych ilustracji.

Wnioski

Dane dostarczone na potrzeby zadań pozwoliły nam zbudować funkcje aproksymujące, które wydają się wiernie odwzorowywać charakterystykę danych. Biorąc pod uwagę nabytą wiedzę jesteśmy świadomi, że dla bardziej chaotycznych danych aproksymacja funkcjami bazowymi stanowiącymi kolejne potęgi x okazała by się mniej dokładna niż aproksymacja wielomianami Czebyszewa bądź początkowymi wyrazami szeregu Taylora. Nie mniej jednak metoda z której skorzystaliśmy jest odpowiednia dla tego typu przykładów, szczególnie ze względu na łatwość implementacji.