

Analiza Numeryczna (M) - Pracownia 2 - Zadanie

P2.10

Aproksymacja średniokwadratowa dyskretna

Prowadzący: dr hab. Paweł Woźny

Karolina Jeziorska

Wrocław, 16 kwietnia 2020

Spis treści

1. Wstęp

Aproksymacja dyskretna, a dokładniej aproksymacja jaka pojawiła się w zadaniu (na zaburzonych danych) jest ściśle związana z analizą danych eksperymentalnych. Badając zjawiska fizyczne, przyrodnicze, ekonomiczne itp... uzyskujemy konkretne wartości. Jednak w celu ogólniejszej interpretacji należy opisać te dane jakąś zależnością, funkcją. Wydaje się, że odpowiedzią na ten problem jest interpolacja, co jednak jest dość mylące. Interpolacja przechodzi przez podane węzły, ale między nimi może zachowywać się niekontrolowanie, szczególnie jeśli mamy dużo węzłów. Dodatkowo, nasze wartości zwykle nie są dokładne, pomiary zazwyczaj obarczone są błędem. Dlatego nie zależy nam, by funkcja opisująca dane przechodziła dokładnie przez nie. Widać, że jednak lepszym rozwiązaniem problemu jest aproksymacja, która zminimalizuje sumaryczny błąd, czyli przybliży nas do pierwotnej funkcji jaką opisywały dane. W tym sprawozdaniu opiszę aproksymację średniokwadratową dyskretną.

2. Teoretyczne omówienie problemu

2.1. Zadanie aproksymacji

Klasyczne zadanie aproksymacji możemy zdefiniować następująco: dla danej funkcji f spośród funkcji ustalonej klasy poszukujemy tej funkcji g , która w określonym sensie najlepiej przybliża f . Sens tej najlepszej aproksymacji zależy od wyboru normy.

2.2. Aproksymacja średniokwadratowa dyskretna

Powyższa definicja przy rozwiązywaniu zadania aproksymacji sprowadza się do rozwiązania zadania: szukamy takiego $w_n \in \Pi_n$ (gdzie Π_n to wszystkie wielomiany stopnia nie większego niż n), że dla każdego $w \in \Pi_n$ zachodzi:

$$\|f - w_n\| < \|f - w\|$$

Co jest równoważne stwierdzeniu:

$$\|f - w_n\| = \inf_{w \in \Pi_n} \|f - w\|$$

W aproksymacji średniokwadratowej dyskretnej nie mamy podanej funkcji f tylko jej wartości w skończonej liczbie punktów (przyjmijmy, że mamy m punktów). Dlatego też $\|f - w_n\|$ wyraża się wzorem:

$$\sum_{i=0}^m p(x_i) [f(x_i) - w_n(x_i)]^2 \quad (1)$$

Gdzie iloczyn skalarny to:

$$\langle f, g \rangle = \sum_{i=0}^m p(x_i) f(x_i) g(x_i) \quad (2)$$

Należy zwrócić też uwagę na $p(x)$ czyli funkcję wagową. Określa dokładność aproksymacji w danym punkcie. Jako, że w rozważanym zadaniu wszystkie punkty są jednakowo ważne, to $p(x) = 1$.

2.3. Rozwiązanie problemu aproksymacji

Wielomian w_n możemy przedstawić jako kombinację liniową wielomianów bazowych P_i :

$$w_n = \sum_{j=0}^n a_j P_j$$

Po podstawieniu do wzoru ?? dostajemy:

$$\|f - w_n\|^2 = \sum_{i=0}^m p(x_i) \left[f(x_i) - \sum_{j=0}^n a_j P_j(x_i) \right]^2 = D(a_0, a_1, \dots, a_n) \quad (3)$$

Mamy funkcję $n + 1$ zmiennych i szukamy jej minimum, czyli dla każdego a_k ($k = 0, 1, \dots, n$) liczymy pochodną i przyrównujemy ją do 0:

$$\begin{aligned} \frac{\partial D}{\partial a_k} &= 0 \\ -2 \sum_{i=0}^m p(x_i) \left[f(x_i) - \sum_{j=0}^n a_j P_j(x_i) \right] P_k(x_i) &= 0 \\ \sum_{i=0}^m p(x_i) f(x_i) P_k(x_i) &= \sum_{i=0}^m p(x_i) P_k(x_i) \sum_{j=0}^n a_j P_j(x_i) \\ \langle f, P_k \rangle &= \sum_{j=0}^n a_j \sum_{i=0}^m p(x_i) P_k(x_i) P_j(x_i) \\ \langle f, P_k \rangle &= \sum_{j=0}^n a_j \langle P_k, P_j \rangle \end{aligned}$$

Biorąc wszystkie równania możemy układ przedstawić w postaci macierzowej:

$$\begin{bmatrix} \langle P_0, P_0 \rangle & \langle P_0, P_1 \rangle & \dots & \langle P_0, P_n \rangle \\ \langle P_1, P_0 \rangle & \langle P_1, P_1 \rangle & \dots & \langle P_1, P_n \rangle \\ \vdots & & & \vdots \\ \langle P_n, P_0 \rangle & \langle P_n, P_1 \rangle & \dots & \langle P_n, P_n \rangle \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \langle f, P_0 \rangle \\ \langle f, P_1 \rangle \\ \vdots \\ \langle f, P_n \rangle \end{bmatrix} \quad (4)$$

Mamy $n+1$ równań i $n+1$ niewiadomych. Dla dużych n liczenie rozwiązania może być skomplikowane i czasochłonne - o ile układ ma jednoznaczne rozwiązanie.

Jak spojrzymy na podany układ to łatwo zauważyć, że rozwiązanie będzie oczywiste i jednoznaczne, gdy macierz iloczynów skalarnych funkcji bazowych poza przekątną będzie miała wartości zerowe. Uzyskamy taką sytuację, gdy wielomiany bazowe będą ortogonalne, czyli $\langle P_i, P_j \rangle = 0$, dla $i \neq j$. Wtedy nasz układ będzie miał postać:

$$\begin{bmatrix} \langle P_0, P_0 \rangle & 0 & \dots & 0 \\ 0 & \langle P_1, P_1 \rangle & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & \langle P_n, P_n \rangle \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \langle f, P_0 \rangle \\ \langle f, P_1 \rangle \\ \vdots \\ \langle f, P_n \rangle \end{bmatrix} \quad (5)$$

Od razu widać rozwiązanie:

$$a_i = \frac{\langle f, P_i \rangle}{\langle P_i, P_i \rangle}$$

Wracając do początkowego przedstawienia w_n i podstawiając wyliczone a_k wielomian optymalny możemy wyrazić wzorem:

$$w_n = \sum_{i=0}^n \frac{\langle f, P_i \rangle}{\langle P_i, P_i \rangle} P_i$$

2.4. Wielomiany bazowe

Cała trudność zadania sprowadza się do wyznaczenia wielomianów ortogonalnych względem ustalonej normy i funkcji wagowej. Problem ten można rozwiązać na dwa sposoby. Jeżeli mamy liniowo niezależny układ wielomianów (bazę), to możemy go zortogonalizować za pomocą ortogonalizacji Grama-Schmidta. Jest jednak ona czasochłonna oraz do wyznaczenia kolejnego wielomianu ortogonalnego trzeba pamiętać wszystkie poprzednie. Jednak łatwiejszym i szybszym sposobem jest wyznaczanie od razu wielomianów ortogonalnych za pomocą zależności rekurencyjnej. Będziemy wyznaczać standardowe wielomiany ortogonalne, czyli takie o współczynniku wiodącym jeden. Spełniają one zależność:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= (x - c_1)P_0 \\ P_k(x) &= (x - c_k)P_{k-1}(x) - d_k P_{k-2}(x) \end{aligned}$$

gdzie

$$c_k = \frac{\langle xP_{k-1}, P_{k-1} \rangle}{\langle P_{k-1}, P_{k-1} \rangle}$$

$$d_k = \frac{\langle P_{k-1}, P_{k-1} \rangle}{\langle P_{k-2}, P_{k-2} \rangle}$$

Dzięki tej zależności wystarczy pamiętać tylko dwa poprzednie wielomiany bazowe, by wyznaczyć kolejny.

2.5. Algorytm Clenshawa

Uogólniony algorytm Clenshawa pozwala obliczyć wielomian optymalny, czyli sumę

$$s_n = a_0P_0 + a_1P_1 + \dots + a_nP_n$$

gdy wielomiany P_i spełniają zależność:

$$P_0(x) = \alpha_0$$

$$P_1(x) = (\alpha_1x - \beta_1)P_0(x)$$

$$P_k(x) = (\alpha_kx - \beta_k)P_{k-1}(x) - \gamma_kP_{k-2}(x)$$

Widać, że jest to ta sama zależność, którą spełniają wielomiany ortogonalne, gdzie:

$$\alpha_i = 1 \quad \beta_i = c_i \quad \gamma_i = d_i$$

Wtedy liczymy pomocnicze wartości V_k ($k = n, n-1, \dots, 0$) za pomocą wzoru:

$$V_k = a_k + (\alpha_{k+1}x - \beta_{k+1})V_{k+1} - \gamma_{k+2}V_{k+2}$$

gdzie $V_{n+2} = V_{n+1} = 0$

Wtedy wartość $s_n = \alpha_0V_0$

3. Część praktyczna

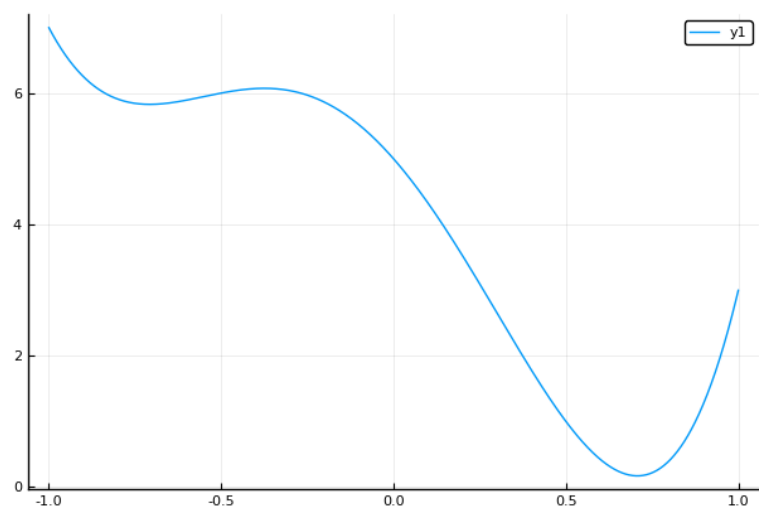
3.1. Wybór wielomianu do aproksymacji

Weźmy przykładowy wielomian 4 stopnia:

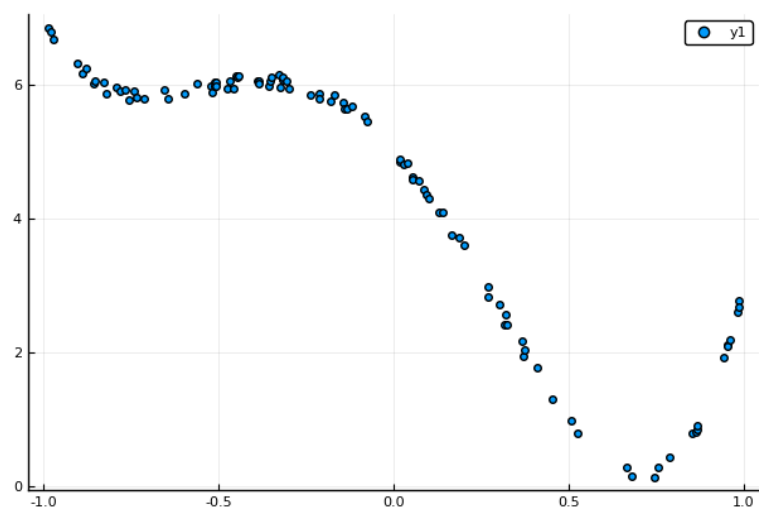
$$f(x) = x^4 + x^3 + x^2 + x + 1 \tag{6}$$

Następnie obliczymy wartości wielomianu w 100 losowych punktach z przedziału $[-1, 1]$ i zaburzmy te wyniki dodając do nich losowe wartości z przedziału $[-\frac{1}{10}, \frac{1}{10}]$

Jest to dokładnie problem opisany we wstępie - mamy zestaw danych, częściowo zaburzonych. Chcemy jak najlepiej odwzorować pierwotną funkcję.



Rysunek 1: Wielomian $f(x) = x^4 + x^3 + x^2 + x + 1$ na przedziale $[-1, 1]$



Rysunek 2: Zaburzone wartości wielomianu

3.2. Konstrukcja wielomianów bazowych

Jak było napisane wcześniej, najlepszym sposobem na rozwiązanie tego problemu jest wcześniejsze wyznaczenie bazy ortogonalnej. Jednak, gdy mamy zbiór skończonej liczby punktów, to nie musimy wyznaczać konkretnego wzoru na P_i a jedynie wyznaczyć wartości w punktach. Pozwoli nam to przy późniejszych obliczeniach wykorzystać już gotowe wartości. Należy także zapamiętać wielkości c_k i d_k - będą potrzebne przy liczeniu wartości wielomianu optymalnego algorytmem Clenshawa.

Żeby wykorzystać trójzależność wielomianów ortogonalnych najlepiej też od razu liczyć wartości współczynników a_k , wystarczy wtedy pamiętać trzy wielomiany bazowe zamiast wszystkich.

3.3. Konstrukcja wielomianów optymalnych

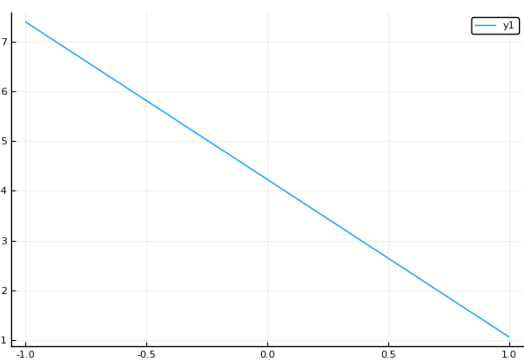
Jak było powiedziane w poprzednim podrozdziale współczynniki a_k , c_k i d_k zostały już policzone. Zostało teraz tylko wyznaczenie wielomianu optymalnego. Ponownie nie musimy wyznaczać ogólnej postaci wielomianu w_n - wystarczy znać jego wartość w punktach, co sprowadza się do wyliczenia za pomocą algorytmu Clenshawa. W celu prezentacji wykresów wartości w_n zostały obliczone w 1000 równoodległych punktach.

Na wykresie ?? zostało przedstawionych kilka pierwszych wielomianów optymalnych

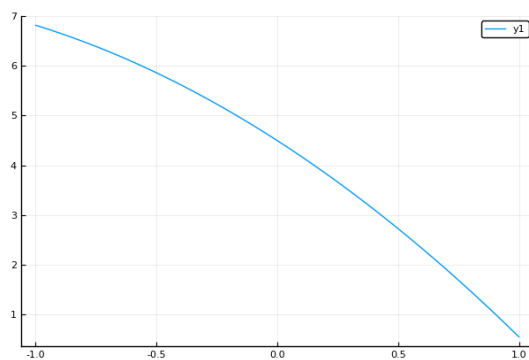
Widać, że już czwarty wielomian optymalny przypomina wybraną funkcję czwartego stopnia, a piąty i szósty wyglądają podobnie. Wydawałoby się, że czym wyższy stopień, tym wielomian aproksymujący będzie lepszy - nie jest to prawda. Jeżeli spojrzymy na trzydziesty wielomian optymalny na rysunku ??, to odbiega on znacznie od oryginalnego. Są tego dwie przyczyny. Wylosowane węzły nie są dokładnie wielomianem pierwotnym, więc próbując przybliżyć się do nich możemy oddalać się od oryginalnej funkcji. Co więcej, czym wyższy stopień wielomianu, tym więcej obliczeń, co prowadzi do kumulacji błędów zaokrągleń.

Warto zauważyć, że gdy będziemy rozważać 99-ty wielomian optymalny, to będzie to dokładnie interpolacja - mamy 100 punktów i szukamy takiego wielomianu, którego odległość od tych punktów jest jak najmniejsza. Możemy zbudować taki wielomian, który dokładnie przez te punkty przejdzie. Jednak jest on mocno niestabilny - jak widać na wykresie ?? mocno maleje w okolicy -1, przez co ciężko porównać wykres do pierwotnej funkcji.

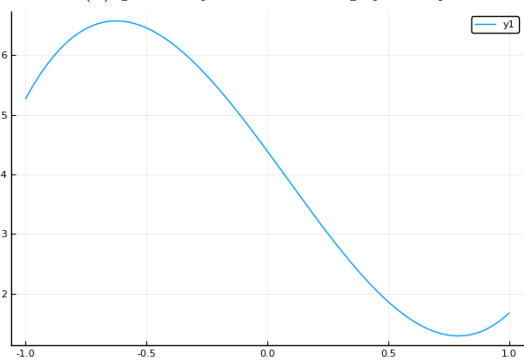
Widząc to ciekawe zachowanie wielomianów optymalnych można, przyjrzeć się maksymalnemu odchyleniu we wcześniej losowanych 100 węzłach. Tabela ?? pokazuje właśnie to maksymalne odchylenie w stosunku do pierwotnych wartości funkcji, jak i tych zaburzonych. Jak widać najmniejsze maksymalne odchylenie jest przy czwartym wielomianie optymalnym (czyli ten sam stopień co oryginalny wielomian). Późniejsze wyniki są już tylko mniej dokładne.



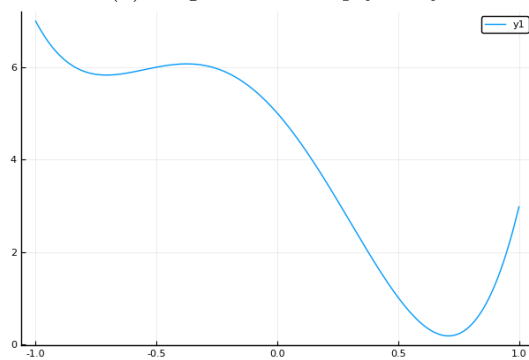
(a) pierwszy wielomian optymalny



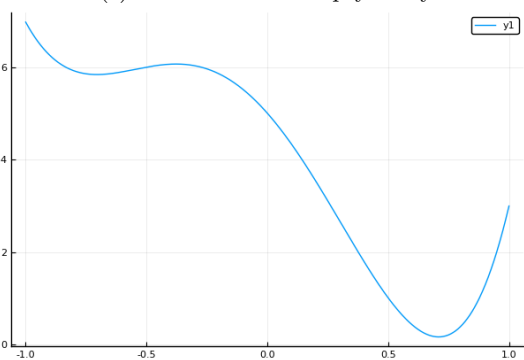
(b) drugi wielomian optymalny



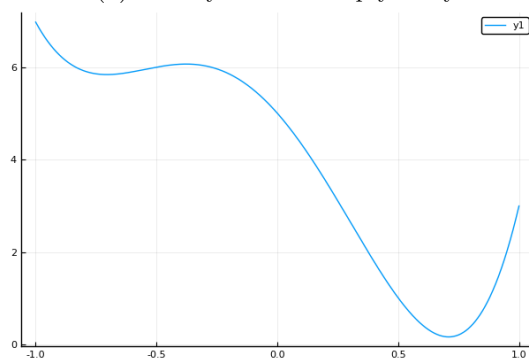
(c) trzeci wielomian optymalny



(d) czwarty wielomian optymalny

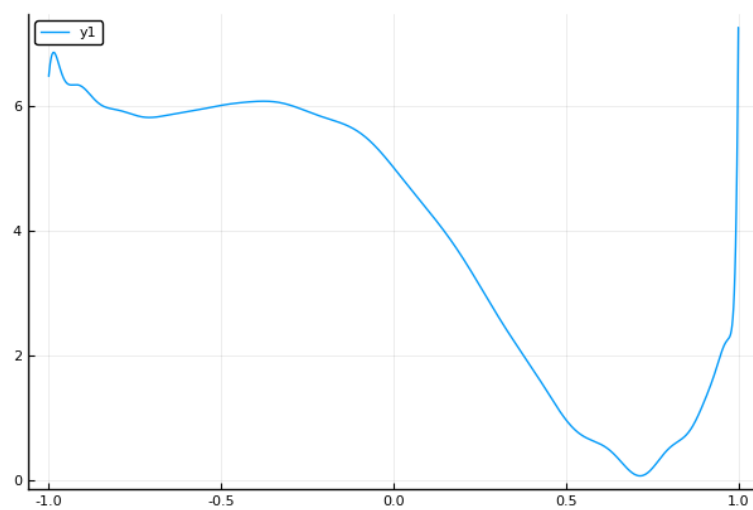


(e) piąty wielomian optymalny

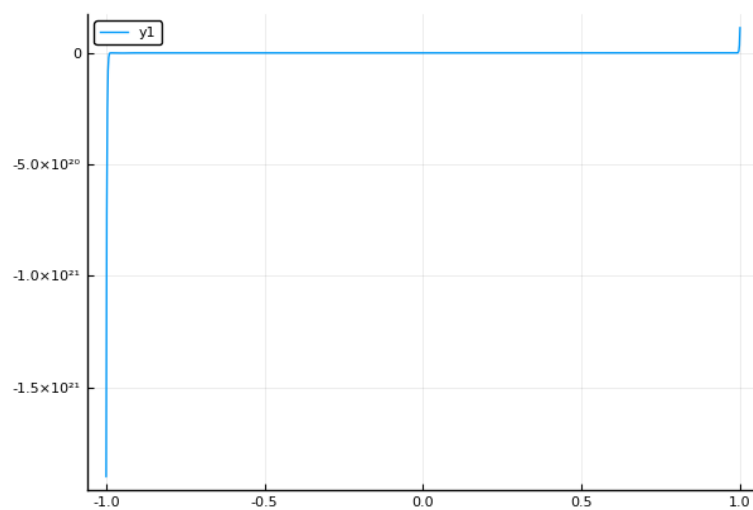


(f) szósty wielomian optymalny

Rysunek 3: Wielomiany optymalne



Rysunek 4: 30-ty wielomian optymalny



Rysunek 5: 99-ty wielomian optymalny

max stopień	wartości oryginalne	wartości zaburzone
1	1.91038405	1.92009947
2	2.07132017	2.14547981
3	1.48622452	1.48810183
4	0.01332262	0.10600691
5	0.02046311	0.11333113
6	0.02063537	0.11352204
⋮	⋮	⋮
30	0.10492267	0.10629689
⋮	⋮	⋮
96	15.13903492	15.13715761
97	47.13903492	47.13715761
98	25.13903492	25.13715761
99	15.13903492	15.13715761

Tabela 1: Tabela błędów (odchyłeń) w 100 wcześniej losowanych punktach

4. Uwagi

Można jeszcze bardziej zaoszczędzić pracy przy liczeniu wielomianów ortogonalnych. Jeśli wybierzemy równoodległe węzły i odpowiednią funkcję wagową to wielomiany ortogonalne są już wyznaczone co pozwala pominąć wszystkie obliczenia z nimi związane. Jednak w przypadku tego zadania punkty są całkowicie losowe, nie można więc z tego ułatwienia skorzystać.

5. Podsumowanie

Praktyczna część tego zadania pokazuje, że aproksymacja jest dobrym rozwiązaniem problemu przybliżania zbioru danych funkcją (wielomianem). Mając nawet wiele punktów możemy przybliżyć je wielomianem o stosunkowo niskim stopniu. Jak widzieliśmy wcześniej zwiększanie stopnia pogarszało nawet wyniki. Pozwala nam to na łatwiejszą analizę danych i opisywanie zjawisk za pomocą funkcji, które są znacznie łatwiejsze w interpretacji niż zbiór danych.

Literatura

- [1] David Kincaid, Ward Cheney *Analiza Numeryczna*, Warszawa, WNT, 2006.
- [2] Janina i Michał Jankowscy *Przegląd metod i algorytmów numerycznych*, Warszawa, WNT, 1981.
- [3] Sławomir Milewski *Metody numeryczne część druga*, Samodzielny Zakład Metod Komputerowych w Mechanice L6, WIL, PK.
- [4] Zenon Fortuna, Bohdan Macukow, Janusz Wąsowski *Metody Numeryczne*, Warszawa, WNT, 2001.