

Metody Numeryczne – Sprawozdanie IV

Metody aproksymacji wielomianowej (wstęp teoretyczny)

Wstęp

Tematem tego sprawozdania są popularne metody aproksymacji funkcji. Praca ta, podobnie jak poprzednie, składa się z dwóch części – wstępu teoretycznego oraz praktycznej implementacji omawianych zagadnień w programie Mathcad. Stara się ona zilustrować zagadnienie aproksymacji w sposób zarówno matematyczny jak i intuicyjny (stąd przykłady i twierdzenia z pogranicza matematyki).

Szkic części teoretycznej

- Przedstawienie problemu
- Błąd średniokwadratowy
- Aproksymacja wielomianowa ogólna
- Aproksymacja wielomianami ortogonalnymi
- Aproksymacja wielomianami trygonometrycznymi
- Wnioski praktyczne
- Bibliografia

Przedstawienie problemu

Aproksymacja jest dziedziną matematyki o szerokich zastosowaniach w naukach technicznych i ekonomicznych. Ujmując problem językiem naukowym: mamy dane wartości nieznannej funkcji $f(x)$ w punktach x_0, x_1, \dots, x_n (węzłach aproksymacji) i poszukujemy funkcji $F(x)$, która w węzłach aproksymacji przyjmuje wartości bliskie funkcji $f(x)$, a w pozostałych punktach możliwie dobrze ją oddaje. Powyższe stwierdzenie może przywieść na pamięć zagadnienie interpolacji. Interpolacja jest szczególnym rodzajem aproksymacji - funkcja interpolująca musi przyjmować w węzłach identyczne wartości co funkcja interpolowana, podczas gdy w przypadku aproksymacji nie stawiamy takiego wymagania.

Samą ideę aproksymacji bardzo dobrze ilustruje również twierdzenie o trzech punktach i prostej:

Tw. 1: *Przez dowolne trzy punkty można przeprowadzić prostą, jeśli ta prosta jest odpowiednio gruba.*

Zadaniem aproksymacji jest taki dobór nachylenia prostej, by miała możliwie jak najmniejszą grubość i spełniała powyższe twierdzenie :). Zauważmy, że gdybyśmy chcieli zastosować w tym przypadku interpolację, to w ogólnym przypadku funkcja interpolacyjna nie opisywała by prostej, lecz krzywą (np. wielomian drugiego stopnia). Uwidacznia się tu zasadnicza różnica między

interpolacją a aproksymacją – **funkcje aproksymujące mogą być dużo prostsze od interpolujących** (przypomnijmy choćby interpolację wielomianową – stopień wielomianu dla n węzłów wynosi w ogólnym przypadku $n-1$, podczas gdy wielomian aproksymacyjny może mieć stopień dowolnie niższy).

Przypadek ciągły i dyskretny

W temacie aproksymacji możemy mieć do czynienia z dwoma jej rodzajami. Przypadek dyskretny oznacza, że dany mamy jedynie zbiór węzłów aproksymacji, ale nie wiemy nic na temat samej funkcji aproksymowanej. Obliczając błąd aproksymacji korzystamy jedynie z wartości w węzłach. Przypadek ciągły (aproksymacja integralna) oznacza, że przy wyliczaniu błędu aproksymacji możemy korzystać z informacji o funkcji aproksymowanej w całej dziedzinie. W tym sprawozdaniu rozpatrujemy jedynie przypadki dyskretne.

Ujęcie praktyczne

Nasuwa się pytanie – na co nam funkcja, która nie przechodzi przez znane węzły? Otóż same węzły nie zawsze są wiarogodne. Rozważmy prosty przypadek badania parametrów nieznanego opornika przez manipulację napięciem i pomiar przepływającego prądu. Jeśli przyjmiemy, że możemy ustalić napięcie z dokładnością do $\pm 0.05V$ (całkiem nieźle jak na warunki laboratoryjne) a pomiar prądu jest obarczony błędem rzędu $\pm 0.05A$, to po naniesieniu tak zmierzonych punktów na układ współrzędnych (prąd w zależności od napięcia) mogą one nie ułożyć się wzdłuż jednej prostej, choć z prawa Ohma wynika, że powinny. Wniosek z tego przykładu jest następujący: **nie ma sensu zmuszać funkcji, by dokładnie przechodziła przez niedokładne węzły**. Często natomiast interesuje nas bardziej charakter funkcji niż jej dokładność (w powyższym przykładzie chcemy wyznaczyć prostą najlepiej oddającą opór, patrz twierdzenie 1).

Błąd średniokwadratowy

W tym sprawozdaniu rozpatrywać będziemy następujące przybliżenie aproksymacyjne:

$$F(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + \dots + c_m\varphi_m(x) = \sum_{i=0}^m c_i\varphi_i(x) \quad (1)$$

Tak zapisana funkcję nazywamy kombinacją liniową znanych nam funkcji bazowych $\varphi_i(x)$. Współczynniki c_i są nam nieznane; zadanie aproksymacyjne polega na ich wyliczeniu. Można zauważyć, że rozwiązań takiego zadania jest nieskończenie wiele – aproksymację można zrealizować w dowolny sposób. Pożądane są jednak rozwiązania o konkretnych właściwościach. Jedną z nich jest **minimalizacja normy średniokwadratowej**. Zdefiniujmy pojęcie *odchylenia średniokwadratowego funkcji $F(x)$ od funkcji $f(x)$* w przypadku dyskretnym:

$$E = \sum_{j=0}^n (f(x_j) - F(x_j))^2 \quad (2)$$

Można powiedzieć, iż jest to suma kwadratów różnicy $F(x_i)$ i $f(x_i)$ (odchylenia $F(x_i)$ od $f(x_i)$ - stąd nazwa). Rozpatrując różne metody aproksymacji będziemy dążyli do znalezienia takich

współczynników c_i , które zminimalizują to odchylenie.

Układ równań normalnych

Łącząc ze sobą równania (1) i (2) otrzymujemy równanie (3):

$$E = \sum_{j=0}^n \left[f(x_j) - \sum_{i=0}^m c_i \varphi_i(x_j) \right]^2 \quad (3)$$

Chcemy odnaleźć współczynniki c_i minimalizujące tę wartość, więc interesuje nas, kiedy wszystkie pochodne cząstkowe $\frac{\partial E}{\partial c_k} = 0$, $k = 1, 2, \dots, m$. Kolejne pochodne cząstkowe opisuje wzór:

$$\frac{\partial E}{\partial c_k} = -2 \sum_{j=0}^n \left[f(x_j) - \sum_{i=0}^m c_i \varphi_i(x_j) \right] \varphi_k(x_j) \quad (4)$$

Przyrównując równania na pochodne cząstkowe do zera i wstawiając węzły x_j uzyskamy tzw. *układ równań normalnych*. Jest to układ nadokreślony – mamy $n+1$ równań na $m+1$ niewiadomych, przy $m \leq n$. Dokładniejsze omówienie tego zagadnienia można znaleźć w [1] i [2]. Współczynniki będące rozwiązaniami powyższego układu równań minimalizują odchylenie średniokwadratowe.

Aproksymacja wielomianowa ogólna

Przyjmując jako funkcje bazowe $\varphi_i(x) = x^i$ możemy układ równań normalnych przekształcić (dokładne omówienie tego przekształcenia w [1]) do postaci:

$$\begin{aligned} \sum_{i=0}^m c_i g_{i,k} &= \rho_k, \text{ gdzie:} \\ g_{i,k} &= \sum_{j=0}^n x_j^{i+k} \\ \rho_k &= \sum_{j=0}^n f(x_j) x_j^k \end{aligned} \quad (5)$$

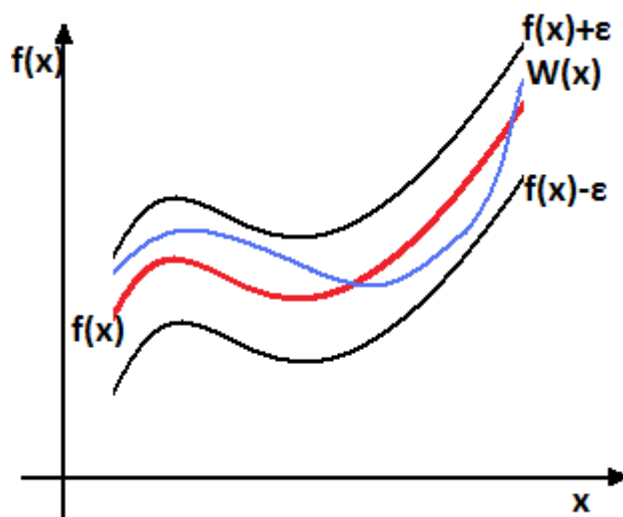
Można wykazać, że jeśli węzły aproksymacji $x_0 \dots x_n$ są różne oraz gdy $m \leq n$, to wyznacznik powyższego układu równań jest różny od zera, a więc układ ten posiada jednoznaczne rozwiązanie (w szczególności dla $m = n$ otrzymujemy wielomian interpolacyjny). W ten sposób funkcja aproksymująca uzyskuje postać: $F(x) = c_0 + c_1 x + c_2 x^2 + \dots + c_m x^m$ - stąd nazwa 'aproksymacja wielomianowa'.

W związku z powyższymi rozważaniami nasuwają się dwa pytania:

- Czy każdą funkcję można możliwie dokładnie aproksymować wielomianem?
- Jak dobierać stopień wielomianu aproksymacyjnego?

Na pierwsze z nich odpowiada nam twierdzenie Weierstrassa:

Tw. 2 (Weierstrassa): Jeśli $f(x)$ jest funkcją określoną i ciągłą w przedziale $[a, b]$ i dane jest $\varepsilon > 0$, to wówczas istnieje wielomian $W(x)$, określony w $[a, b]$ taki, że $|f(x) - W(x)| < \varepsilon$ dla każdego $x \in [a, b]$.



rys. 1. Twierdzenie Weierstrassa

Zgodnie z tym twierdzeniem dla każdej funkcji jesteśmy w stanie odnaleźć wielomian aproksymujący ją z dowolną dokładnością.

Drugie pytanie – jak dobierać stopień wielomianu aproksymacyjnego – wymaga nieco dłuższego omówienia. W ogólnym przypadku zależy nam, by wielomian miał jak najmniejszy stopień, gdyż wraz z jego wzrostem zwiększają się też błędy związane z metodą (układ normalny jest źle uwarunkowany) jak i obliczeniami zmiennopozycyjnymi. Z drugiej jednak strony zależy nam, by wielomian dobrze przybliżał aproksymowaną funkcję i wygładzał ewentualne błędy pomiarów węzłów aproksymacji. Poniżej zamieszczam kilka wytycznych:

- Jeśli znamy postać funkcji aproksymowanej, możemy dopasować stopień wielomianu aproksymacyjnego do tej funkcji. W szczególności jeśli sama funkcja jest wielomianem, to możemy przyjąć taki sam stopień, jak stopień funkcji.
- Stopień wielomianu aproksymacyjnego może wynikać z szczególnych wymagań dotyczących problemu; przykładowo możemy potrzebować wielomianu stopnia pierwszego lub drugiego (takie potrzeby często zachodzą np. w statystyce), i te zapotrzebowania narzucają nam stopień.
- Możemy zwiększać stopień wielomianu i badać błąd aproksymacji. Ponieważ rozwiązanie układu normalnego jest operacją kosztowną, sensowne jest ograniczenie żądanej dokładności – na przykład jeśli błąd między stopniem n i $n+1$ nie zmalał znacząco (kilkadziesiąt procent), należy zatrzymać się na stopniu n -tym.
- W przypadku, gdy stopień wielomianu jest o jeden mniejszy od liczby węzłów, uzyskamy wielomian interpolacyjny.

Aproksymacja wielomianami ortogonalnymi

Wielomiany ortogonalne (słowo *ortogonalne* można w dużym uproszczeniu odczytywać jako *wzajemnie prostopadłe*) to szczególny rodzaj wielomianów. Dwa wielomiany $W(x)$ i $Q(x)$ są wzajemnie ortogonalne, jeśli ich iloczyn skalarny jest równy 0:

$$W(x) \bullet Q(x) = 0 \quad (6)$$

Iloczyn skalarny dla funkcji w postaci dyskretnej możemy zdefiniować analogicznie jak dla wektorów:

$$f(x) \bullet g(x) = \sum_{i=0}^n f(x_i) \cdot g(x_i) \quad (7)$$

przy założeniu, że nie wszystkie punkty x_i są miejscami zerowymi dla funkcji f i g (warto zauważyć, że dwa niezerowe wektory są prostopadłe wtedy i tylko wtedy, gdy ich iloczyn skalarny równy jest 0 – dlatego termin *ortogonalny* można rozumieć jako *prostopadły*).

W aproksymacji korzystamy ze specjalnej grupy wielomianów ortogonalnych, zwanych wielomianami ortogonalnymi Grama **dla równo odległych węzłów aproksymacyjnych**. Użycie wielomianów ortogonalnych zapewnia nam, że macierz współczynników w układzie normalnym jest macierzą diagonalną, dzięki czemu nie musimy prowadzić skomplikowanych obliczeń w celu uzyskania współczynników wielomianu (zapewnia to też dobre uwarunkowanie układu normalnego, dzięki czemu minimalizujemy wpływ błędów pomiaru węzłów aproksymacji). Po stosownych obliczeniach opisanych w [1] uzyskujemy następującą postać rodziny wielomianów Grama:

$$G_k^{(n)}(q) = \sum_{s=0}^k (-1)^s \binom{k}{s} \binom{k+s}{s} \frac{q^{[s]}}{n^{[s]}} \quad (8)$$

We wzorze tym $G_k^{(n)}$ oznacza wielomian Grama stopnia k na $n+1$ **węzłach równo odległych**, symbol $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ to dwumian Newtona, a symbol $x^{[n]} = x \cdot (x-1) \cdot \dots \cdot (x-n+1) = \frac{x!}{(x-n)!}$.

Odpowiednik wzoru (1) dla wielomianów Grama ma postać:

$$F_m(x) = \sum_{k=0}^m \frac{c_k}{s_k} \cdot G_k^{(n)}(q) = \sum_{k=0}^m \frac{c_k}{s_k} \cdot G_k^{(n)}\left(\frac{x-x_0}{h}\right), \text{ gdzie:}$$

$$q = \frac{x-x_0}{h}$$

$$c_k = f \bullet G_k^{(n)} = \sum_{i=0}^n f(x_i) \cdot G_k^{(n)}(x_i)$$

$$s_k = \sum_{q=0}^n [G_k^{(n)}(q)]^2 \quad (9)$$

Gdzie F_m to wielomian aproksymacyjny stopnia m , a $h = x_{i+1} - x_i$ to odległość między sąsiadującymi węzłami (stała, bo węzły z założenia są równo odległe). Zasady doboru stopnia

wielomianu aproksymacyjnego są identyczne jak przy aproksymacji wielomianami ogólnymi.

Aproksymacja wielomianami trygonometrycznymi

Gdy dokonujemy aproksymacji funkcji okresowej często lepsze (dokładniejsze, mniejszym kosztem) rezultaty uzyskamy z pomocą rodziny wielomianów trygonometrycznych. W przypadku wielomianów algebraicznych ogólnych za kolejne funkcje bazowe $\varphi_i(x)$ z wzoru (1) przyjmowaliśmy jednomiany x^i . Dla wielomianów trygonometrycznych taka baza (ciąg funkcji bazowych) to:

$$(\varphi_i(x)) = 1, \sin(x), \cos(x), \sin(2x), \cos(2x), \dots, \sin(mx), \cos(mx) \quad (10)$$

Można pokazać, że gdy przyjmujemy $n+1$ równo odległych węzłów aproksymacji opisanych wzorem $x_i = n \cdot i \cdot \frac{\pi}{2}$, to kolejne elementy tej bazy będą do siebie ortogonalne, tj.:

$$\varphi_i(x) \bullet \varphi_{i+1}(x) = 0, \quad i = 0, 1, \dots, n \quad (11)$$

Daje nam to te same korzyści co w przypadku ortogonalnych wielomianów algebraicznych – układ normalny jest dobrze uwarunkowany a jego policzenie jest zadaniem trywialnym, gdyż jedyne niezerowe elementy znajdują się na diagonalu macierzy współczynników.

Ostateczne wzory przybliżające szukaną funkcję wielomianem trygonometrycznym (na podstawie [1]):

$$\begin{aligned} F_m(x) &= \frac{1}{2} \cdot a_0 + \sum_{j=1}^m (a_j \cdot \cos(j \cdot x) + b_j \cdot \sin(j \cdot x)) \\ a_j &= \frac{2}{n} \cdot \sum_{i=0}^{n-1} f(x_i) \cdot \cos(j \cdot x_i) \\ b_j &= \frac{2}{n} \cdot \sum_{i=0}^{n-1} f(x_i) \cdot \sin(j \cdot x_i) \end{aligned} \quad (12)$$

Wielomianami trygonometrycznymi można aproksymować dowolną funkcję okresową, co wynika pośrednio z twierdzenia Weierstrassa dla funkcji okresowych:

Tw. 3 (Weierstrassa): Jeśli $f(x)$ jest funkcją określoną i ciągłą w przedziale $[a, b]$ oraz okresową o okresie równym 2π i dane jest $\varepsilon > 0$, to wówczas istnieje wielomian trygonometryczny $W(x)$, określony w $[a, b]$ taki, że $|f(x) - W(x)| < \varepsilon$ dla każdego $x \in [a, b]$.

Twierdzenie to można zilustrować w taki sam sposób jak twierdzenie 2.

Zasady doboru 'stopnia' wielomianu aproksymacyjnego różnią się od tych podanych przy okazji wielomianów algebraicznych. W przypadku wielomianów trygonometrycznych możemy od razu przyjąć najwyższy dopuszczalny stopień, równy $m = \left\lfloor \frac{n-1}{2} \right\rfloor$ (symbol $\lfloor n \rfloor$ oznacza największą liczbę całkowitą mniejszą lub równą n). Próba przyjęcia wyższego stopnia sprawia, że problem staje się

źle uwarunkowany.

Wnioski praktyczne

- Odchylenie średniokwadratowe w przypadku aproksymacji trygonometrycznej jest znacznie lepsze od pozostałych rodzajów aproksymacji, dlatego warto ją stosować tam, gdzie to tylko możliwe.
- W przeciwieństwie do pozostałych omówionych metod, w aproksymacji wielomianami trygonometrycznymi wzrost liczby węzłów aproksymacji oraz stopnia wielomianu aproksymacyjnego nie powoduje zwiększenia się błędów obliczeniowych.
- Aproksymacja wielomianami ogólnymi może być źle uwarunkowana, co zawęży jej zastosowanie. Problem ten nie występuje w przypadku pozostałych omówionych metod.
- Z twierdzeń 2 i 3 (twierdzenia Weierstrassa) wynika, że każdą funkcję ciągłą na danym przedziale możemy aproksymować wielomianami z dowolną żadaną dokładnością na tym przedziale.
- Ze względów wydajności i dokładności obliczeń należy starać się minimalizować stopień wielomianu aproksymacyjnego.
- Przed wybraniem sposobu aproksymacji danej funkcji warto poświęcić czas na uzyskanie jak największej ilości informacji na temat całego problemu oraz samej funkcji aproksymowanej – pomoże to dobrać optymalną metodę i dokładność aproksymacji (pewne problemy mogą narzucać specyficzną postać lub zadaną dokładność funkcji aproksymującej).

Bibliografia

1. Z. Fortuna, B. Macukow, J. Wąsowski, „Metody Numeryczne”, wydanie III, wyd. WNT, Warszawa 1982, 1993
2. Åke Björck, Germund Dahlquist, „Metody Numeryczne”, wydanie II, wyd. PWN, Warszawa 1987