

1 Interpolacja ogólna i wielomianami

1.1 Na czym polega interpolacja?

Interpolacja polega na budowaniu tzw. **funkcji interpolujących** $\varphi(x)$ na podstawie zadanych wartości $f(x)$ zapamiętanych w postaci dyskretnej. Postać dyskretną tworzą dwa zbiory liczbowe:

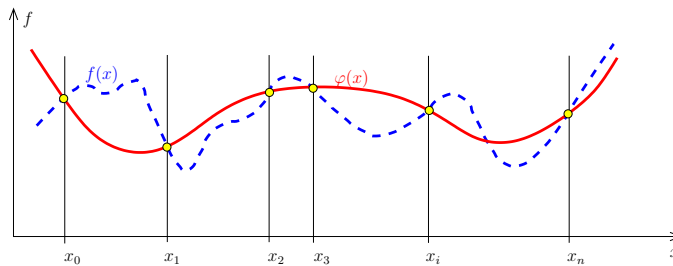
$$\mathbf{X} = \{x_0, x_1, \dots, x_n\}$$

które nazywamy **węzłami interpolacji**, oraz wartości pewnej funkcji interpolowanej:

$$\mathbf{F} = \{f_0, f_1, \dots, f_n\}$$

gdzie dla wygody zapisu $f_i \equiv f(x_i)$.

Istotę interpolacji wyjaśnia rysunek na którym pokazano wykres funkcji interpolowanej $f(x)$ i interpolującej $\varphi(x)$.



Obie funkcje: **interpolowana** $f(x)$ i **interpolująca** $\varphi(x)$ na zbiorze węzłów \mathbf{X} mają **dokładnie takie same wartości liczbowe**. Są zatem spełnione warunki

$$\varphi(x_i) = f(x_i), \quad i = 0, 1, \dots, n$$

na podstawie których konstruuje się funkcję interpolującą $\varphi(x)$.

1.2 Postać ogólna wielomianu interpolacyjnego

Funkcja $\varphi(x)$ w przypadku ogólnym jest przedstawiana najczęściej w postaci **wielomianu uogólnionego**, utworzonego z odpowiednio dobranych tzw. **funkcji bazowych** $\phi_i(x)$, $i = 0, 1, \dots, n$.

$$\varphi(x) = a_0 \phi_0(x) + a_1 \phi_1(x) + \dots + a_n \phi_n(x) = \sum_{i=0}^n a_i \phi_i(x) = \mathbf{\Phi}(x) \mathbf{a},$$

gdzie wprowadzono oznaczenia:

$\mathbf{\Phi}(x) = [\phi_0(x), \phi_1(x), \dots, \phi_n(x)]$ – **baza interpolacyjna**

$\mathbf{a} = [a_0, a_1, \dots, a_n]^T$ – **wektor współczynników (mnożników)** funkcji $\varphi(x)$ względem bazy $\mathbf{\Phi}(x)$.

1.3 Sformułowanie zagadnienia interpolacji funkcji jednej zmiennej

Każda baza $\Phi(x)$ ma spełniać podstawowy warunek: funkcje $\phi_i(x)$, $i = 0, 1, \dots, n$ muszą być liniowo niezależne, czyli gdy żadnej z nich nie da się przedstawić w postaci

$$\phi_i(x) = \sum_{k=1}^{i-1} a_k \phi_k(x) + \sum_{k=i+1}^n a_k \phi_k(x), \quad i = 1, 2, \dots, n.$$

W celu utworzenia konkretnej funkcji interpolującej $\varphi(x)$, należy:

- przyjąć odpowiednią bazę $\Phi(x)$
- wyznaczyć wartości liczbowe wszystkich jej współczynników a_i , $i = 0, 1, \dots, n$
- sformułować wielomian $\varphi(x_i)$

Wielomian ma postać:

$$\varphi(x_i) = \Phi(x_i) \mathbf{a} = f_i, \quad i = 0, 1, \dots, n \quad \text{lub} \quad \mathbf{U} \mathbf{a} = \mathbf{f}. \quad (1)$$

Po wprowadzeniu oznaczenia $u_{ij} = \phi_i(x_j)$, $i, j = 0, 1, 2, \dots, n$ otrzymujemy

$$\mathbf{U} = \begin{bmatrix} u_{00} & u_{10} & \cdots & u_{n0} \\ u_{01} & u_{11} & \cdots & u_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{0n} & u_{1n} & \cdots & u_{nn} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}, \quad (2)$$

Baza $\Phi(x)$ utworzona jest z funkcji liniowo niezależnych $\det(\mathbf{U}) \neq 0$. Układ równań (1) ma jednoznaczne rozwiązanie: wyraża się wzorem:

$$\mathbf{a} = \mathbf{U}^{-1} \mathbf{f}. \quad (3)$$

1.4 Bazy interpolacji

Najczęściej stosowanymi, a zarazem najprostszymi bazami interpolacji są:

1. **baza jednomianowa** dla ciągłych funkcji na skończonym odcinku $[a, b]$

$$\Phi(x) = [1, x, x^2, x^3, \dots, x^n],$$

2. **baza trygonometryczna** dla okresowych funkcji $f(x)$ na odcinku $[-\pi, \pi]$

$$\Phi(x) = [1/\sqrt{2}, \sin x, \cos x, \dots, \sin nx, \cos nx],$$

Nie każda baza interpolacyjna nadaje się równie dobrze do rozwiązywania konkretnych zadań. Dlatego staje się niezbędne dostosowanie bazy do specyfiki rozwiązywanego problemu.

1.5 Interpolacja wielomianowa (algebraiczna)

Ten rodzaj aproksymacji jest najprostszy. Ma on jednak ograniczone zastosowanie, zwłaszcza gdy zachodzi konieczność posługiwania się wielomianami $n > 10$. W tym przypadku algorytm metody staje się niestabilny i wrażliwy na błędy zaokrągleń. Dla wielomianów stopni $n > 10$ w obliczeniach pojawia tzw. efekt Rungego, polegający na tym, że wielomian interpolacyjny przybliża funkcję interpolowaną w sposób niejednostajny. Wielomian ten wykazuje oscylacje względem funkcji $f(x)$ w pobliżu jego końców osiągając tym większe wartości im wyższy jest stopień n wielomianu.

Przyjmujemy bazę jednomianową

$$\Phi(x) = [1, x, x^2, x^3, \dots, x^n],$$

i budujemy wielomianową funkcję interpolującą

$$\varphi(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots + a_n x^n \quad (4)$$

Układ równań $\mathbf{U}\mathbf{a} = \mathbf{f}$ przyjmuje teraz postać:

$$\begin{bmatrix} 1 & x_0 & x_0^2 & x_0^3 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix},$$

z którego w sposób jednoznaczny można obliczyć współczynniki \mathbf{a} wielomianu (4).

1.6 Przejście do bazy Lagrange'a

Stosowanie wzorów:

$$\mathbf{a} = \mathbf{U}^{-1} \mathbf{f} \quad \longrightarrow \quad \varphi(x) = \Phi(x) \mathbf{a} = \Phi(x) \mathbf{U}^{-1} \mathbf{f} = \mathbf{N}(x) \mathbf{f},$$

w celu zbudowania nowej bazy:

$$\mathbf{N}(x) \equiv \Phi(x) \mathbf{U}^{-1} = [L_0(x), L_1(x), \dots, L_n(x)]$$

w przypadku ogólnym jest niewygodne ze względu na dużą złożoność obliczeń. Dlatego też elementy tej bazy nazywanej **bazą Lagrange'a** zostaną wygenerowane w inny sposób.

Przyjmujemy funkcję interpolującą stopnia n w postaci **wielomianu interpolacyjnego Lagrange'a**

$$N(x) \equiv L^n(x) = f_0 L_0^n(x) + f_1 L_1^n(x) + \dots + f_n L_n^n(x). \quad (5)$$

Jeżeli założymy, że

$$\mathbf{f} = [0, 0, \dots, 1_i, \dots, 0, 0,]^T, \quad (6)$$

to na podstawie (5) otrzymamy

$$L^n(x) \equiv L_i^n(x). \quad (7)$$

czyli w szczególnym przypadku funkcja interpolująca $L^n(x)$ jest identyczna z bazową $L_i^n(x)$ i musi zgodnie z (6)

- zerować się we wszystkich węzłach x_j dla których $j \neq i$,
- przyjmować wartość jednostkową w węźle x_i , czyli

$$L_i^n(x_j) = \delta_{ij}, \quad \begin{cases} \delta_{ij} = 1, & \text{gdy } i = j, \\ \delta_{ij} = 0, & \text{gdy } i \neq j. \end{cases}$$

Z powyższego wynika, że funkcja bazowa $L_i^n(x)$ musi mieć postać:

$$L_i^n(x_j) = C_i (x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n). \quad (8)$$

Ponieważ w węźle x_i wielomian ten przyjmuje wartość równą 1 otrzymujemy

$$C_i (x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n) = 1.$$

Stąd otrzymujemy

$$C_i = \frac{1}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}. \quad (9)$$

Po podstawieniu do (8) i -ta funkcja bazowa Lagrange'a ma postać

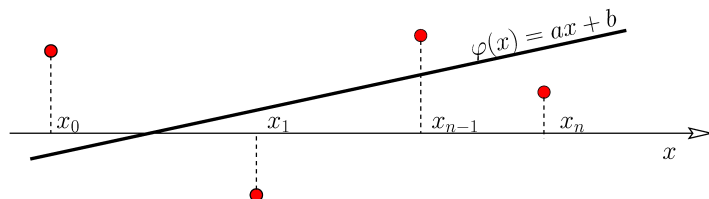
$$L_i^n = \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}$$

$$L_i^n = \prod_{j=0, j \neq i}^{j=n} \frac{x - x_j}{x_i - x_j}. \quad (10)$$

2 Aproksymacja wielomianami i ogólna

2.1 Zasadnicza różnica pomiędzy interpolacją a aproksymacją

Aproksymacja różni się od interpolacji funkcji tym, że dla wyznaczania współczynników wielomianu aproksymacyjnego (nie interpolacyjnego!) nie korzysta się z warunków $\varphi_i(x) = f_i$. Oznacza to, że wielomian aproksymacyjny na zbiorze X **nie musi** przyjmować wartości funkcji aproksymowanej.



2.2 Aproksymacja optymalna

Stopień wielomianu aproksymacyjnego nie ma więc związku z liczbą elementów zbioru \mathbf{X} , a wyznaczenie niewiadomych a_i musi być realizowane inną metodą.

Wybór tego sposobu decyduje o własnościach stosowanej aproksymacji, która w pewnym określonym sensie powinna być najlepsza (optymalna). Rozwiązanie problemu **aproksymacji optymalnej** wymaga:

- przyjęcia odpowiedniej bazy funkcyjnej
- ustalenia kryterium oceny jakości aproksymacji, które służy do jednoznacznego określenia wartości a_i .

Można stwierdzić, że aproksymacja funkcji $f(x)$ za pomocą funkcji $\varphi(x)$ jest tym lepsza, im te dwie funkcje mniej różnią się od siebie.

2.3 Kryterium jakości aproksymacji

Jeśli funkcja $f(x)$ określona jest tylko za pomocą skończonego zbioru jej rzędnych to wartości różnicy $|f(x) - \varphi(x)|$ mogą być obliczone tylko na zbiorze \mathbf{X} i wynoszą $|f(x_i) - \varphi(x_i)|$, $i = 1, 2, \dots, n$.

Najprostsza postać kryterium oceniającego jakość aproksymacji jest następująca

$$\min R = \min \sum_{i=0}^n [f(x_i) - \varphi(x_i)]^2 \quad (11)$$

Wartość funkcji R jest pewną miarą odchylenia funkcji aproksymującej $\varphi(x)$ od aproksymowanej $f(x)$.

Ten sposób postępowania nosi nazwę metody najmniejszych kwadratów.

2.4 Aproksymacja wielomianami algebraicznymi

Wielomian aproksymacyjny stopnia $m = 1, 2, \dots, n$ ma postać:

$$\varphi(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_m x^m = \sum_{k=0}^m a_k x^k \quad (12)$$

W przypadku gdy $m = n$ mamy do czynienia z interpolacją, dla której $R = 0$.

Gdy $m < n$, wówczas $R > 0$ i odchylenie $\varphi(x)$ od $f(x)$ jest najmniejsze gdy R osiąga wartość minimalną.

Obliczenie wartości a_i polega na wykorzystaniu warunków stacjonarności funkcji R :

$$\frac{\partial R}{\partial a_k} = 0, \quad k = 0, 1, 2, \dots, m. \quad (13)$$

2.5 Warunki stacjonarności dla aproksymacji wielomianami. Metoda najmniejszych kwadratów

Warunków tych jest $m + 1$, a więc tyle ile niewiadomych współczynników a_i ma wielomian aproksymujący.

Obliczając pochodne cząstkowe funkcji R otrzymujemy

$$\frac{\partial R}{\partial a_k} = \frac{\partial}{\partial a_k} \left\{ \sum_{i=0}^n [f(x_i) - \varphi(x_i)]^2 \right\} =$$

$$2 \sum_{i=0}^n [f(x_i) - \varphi(x_i)] \frac{\partial}{\partial a_k} \varphi(x_i) = 0, \quad \text{dla } k = 0, 1, \dots, m \quad (14)$$

Gdy podstawimy $\frac{\partial}{\partial a_k} \varphi(x_i) = x_i^k$, to:

$$\sum_{i=0}^n [\Phi(x_i) \mathbf{a} - f(x_i)] \cdot x_i^k = \sum_{i=0}^n [\Phi(x_i) \cdot x_i^k \mathbf{a} - f(x_i) x_i^k] =$$

$$\sum_{i=0}^n \left\{ [x_i^k \ x_i^{k+1} \ \dots \ x_i^{k+m}] \mathbf{a} - f_i x_i^k \right\} = 0, \quad k = 0, 1, \dots, m.$$

Układ równań dla aproksymacji wielomianami

Wprowadzając oznaczenia

$$s_k = \sum_{i=0}^n x_i^k, \quad t_k = \sum_{i=0}^n f_i x_i^k \quad (15)$$

otrzymujemy ostatecznie

$$[s_k, s_{k+1}, \dots, s_{k+m}] \mathbf{a} = t_k, \quad k = 0, 1, 2, \dots, m \quad (16)$$

lub

$$\mathbf{S} \mathbf{a} = \mathbf{t} \quad (17)$$

gdzie:

$$\mathbf{S} = \begin{bmatrix} s_0 & s_1 & s_2 & \cdots & s_m \\ s_1 & s_2 & s_3 & \cdots & s_{m+1} \\ s_2 & s_3 & s_4 & \cdots & s_{m+2} \\ \dots & \dots & \dots & \dots & \dots \\ s_m & s_{m+1} & s_{m+2} & \cdots & s_{2m} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \cdots \\ a_m \end{bmatrix}, \quad \mathbf{t} = \begin{bmatrix} t_0 \\ t_1 \\ t_2 \\ \cdots \\ t_m \end{bmatrix}.$$

2.6 Aproksymacja wielomianami uogólnionymi

Założmy, że $\phi_i(x)$, $i = 0, 1, \dots, m$ jest układem funkcji bazowych.

Poszukujemy wielomianu uogólnionego $\varphi(x)$, będącego najlepszym przybliżeniem średniokwadratowym funkcji $f(x)$ na zbiorze \mathbf{X} , tj. funkcji:

$$\varphi(x) = a_0 \phi_0(x) + a_1 \phi_1(x) + \dots + a_m \phi_m(x) = \sum_{k=0}^m a_k \phi_k(x). \quad (18)$$

Współczynniki a_k są tak określone aby wyrażenie

$$R = \sum_{i=0}^n |f(x_i) - \varphi(x_i)|^2. \quad (19)$$

było minimalne.

Oznaczmy:

$$R = \sum_{i=0}^n \left[f(x_i) - \sum_{k=0}^m a_k \phi_k(x_i) \right]^2. \quad (20)$$

Z warunków:

$$\frac{\partial R}{\partial a_j} = 0, \quad j = 0, 1, \dots, m, \quad (21)$$

otrzymamy układ $m + 1$ równań liniowych z $m + 1$ niewiadomymi a_j :

$$\frac{\partial R}{\partial a_j} = \sum_{i=0}^n \left[f(x_i) - \sum_{k=0}^m a_k \phi_k(x_i) \right] \phi_j(x_i) = 0, \quad j = 0, 1, \dots, m. \quad (22)$$

2.7 Układ równań dla aproksymacji

W zapisie macierzowym układ (22) przyjmuje postać:

$$\mathbf{D}^T \mathbf{D} \mathbf{a} = \mathbf{D}^T \mathbf{f} \quad (23)$$

gdzie

$$\mathbf{D} = \begin{bmatrix} \phi_0(x_0) & \phi_1(x_0) & \dots & \phi_m(x_0) \\ \phi_0(x_1) & \phi_1(x_1) & \dots & \phi_m(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \dots & \phi_m(x_2) \\ \dots & \dots & \dots & \dots \\ \phi_0(x_n) & \phi_1(x_0) & \dots & \phi_m(x_n) \end{bmatrix} \quad \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \dots \\ a_m \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \\ \dots \\ f(x_n) \end{bmatrix}. \quad (24)$$

Macierz współczynników układu (23) jest macierzą symetryczną i dodatnio określoną co zapewnia jednoznaczność rozwiązania.