

## I. APROKSYMACJA I INTERPOLACJA FUNKCJI JEDNEJ ZMIENNEJ

Ogólnie zagadnienie aproksymacji można opisać następująco:

- Dane są punkty należące bądź to do wykresu funkcji bądź pochodzące z danych eksperymentalnych lub numerycznych (liczba punktów –  $n$ )

$$(x_i, f_i) \quad \text{dla } i = 1, 2, \dots, n$$

Odcięte  $x_i$  nazywamy węzłami aproksymacji, natomiast rzędne  $f_i$  wartościami węzłowymi.

- Przyjmuje się tzw. rząd aproksymacji  $m$  ( $m = 0, 1, \dots, n-1$ ). Jest to ilość niezależnych liniowo funkcji bazowych  $\varphi_i(x)$ , przyjmowanych na podstawie danego kryterium, a także ilość nieznanymi współczynników liczbowych  $a_i$ , które zostaną wyznaczone w dalszym ciągu zadania. Ogólny zapis funkcji aproksymującej:

$$p(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \dots + a_m\varphi_m(x) = \sum_{i=0}^m a_i\varphi_i(x) \quad (1)$$

lub w notacji macierzowej:

$$p(x) = \mathbf{a}^T \boldsymbol{\varphi}(x), \quad \text{gdzie: } \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \dots \\ a_m \end{bmatrix}, \quad \boldsymbol{\varphi}(x) = \begin{bmatrix} \varphi_0(x) \\ \varphi_1(x) \\ \dots \\ \varphi_m(x) \end{bmatrix}$$

- Przyjmuje się tzw. wagi  $w_i$  dla każdego węzła z osobna, które świadczą o odejściu krzywej aproksymacyjnej od oryginalnej wartości węzłowej wg zależności: im większa waga, tym bliżej tego właśnie punktu przejdzie krzywa. Wagi można dobierać np. według kryterium odległościowego od ustalonego z góry punktu. Wagi zbiera się do macierzy diagonalnej zwanej macierzą wagową.

$$\mathbf{W} = \text{diag}(w_i) = \begin{bmatrix} w_1 & 0 & 0 & 0 \\ 0 & w_2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & w_n \end{bmatrix}.$$

Oczywiście wprowadzanie wag nie jest konieczne. W takim przypadku:  
 $w_1 = w_2 = \dots = w_n = 1$ .

- Wyznacza się współczynniki liczbowe  $a_i$  z następującego układu równań:

$$\Phi_{(n \times m)} = \left| \varphi_j(x_i) \right| = \begin{bmatrix} \varphi_0(x_1) & \varphi_1(x_1) & \dots & \varphi_m(x_1) \\ \varphi_0(x_2) & \varphi_1(x_2) & \dots & \varphi_m(x_2) \\ \dots & \dots & \dots & \dots \\ \varphi_0(x_n) & \varphi_1(x_n) & \dots & \varphi_m(x_n) \end{bmatrix}, \quad \mathbf{F}_{(1 \times n)} = |f_i| = \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{bmatrix}$$

$$\Phi^T \mathbf{W} \Phi \mathbf{a} = \Phi^T \mathbf{W} \mathbf{F}, \Rightarrow \mathbf{a} = (\Phi^T \mathbf{W} \Phi)^{-1} \Phi^T \mathbf{W} \mathbf{F}$$

Na ich podstawie można budować aproksymację funkcji za pomocą wzoru (1).

- Oblicza się błąd aproksymacji na podstawie następujących wzorów:

- Dla aproksymacji ciągłej:  $\varepsilon = \int_{x_1}^{x_n} (p(x) - f(x)) dx$ ,

- Dla aproksymacji dyskretnej:

$$\varepsilon = \|p(x_i) - f_i\|_{i=1,2,\dots,n} = \begin{cases} \sqrt{\sum_{i=1}^n (p(x_i) - f_i)^2}, & \text{dla normy Euklidesa} \\ \max_i |p(x_i) - f_i|, & \text{dla normy maksymalnej} \end{cases}$$

Powyższy algorytm aproksymacji jest ogólny i prawdziwy dla dowolnej liczby węzłów, ilości i postaci funkcji bazowych. Wszystkie poniższe rodzaje aproksymacji można łatwo wyprowadzić korzystając z tego algorytmu. Jest on jednak dość uciążliwy zwłaszcza w obliczeniach ręcznych, stąd dla konkretnego rodzaju aproksymacji korzysta się z innych zależności, prostszych w zapisie i zastosowaniu.

## INTERPOLACJA FUNKCJI

Interpolacja funkcji to taka aproksymacja, w której funkcja  $p(x)$  przechodzi przez wszystkie punkty  $(x_i, f_i)$ ,  $i = 1, 2, \dots, n$  bez żadnego wyjątku. To znaczy, iż błąd liczony jak dla aproksymacji dyskretnej musi być w węzłach bezwarunkowo równy zero. Stąd warunek interpolacji formułuje się następująco:

$$p(x_i) = f_i, \quad \text{dla } i = 1, 2, \dots, n.$$

Implikuje to od razu postać funkcji interpolacyjnej:

$$p(x) = \sum_{i=1}^n a_i \varphi_i(x) \tag{2}$$

, tzn., że funkcji bazowych (oraz współczynników interpolacji) musi być dokładnie tyle, ile węzłów. Tak, więc zadanie interpolacji jest zadaniem jednoznacznym (jest tylko jedna krzywa interpolacyjna, która dla danego zestawu funkcji bazowych przechodzi ściśle przez wszystkie dane punkty). W zapisie macierzowym interpolacja wygląda następująco:

$$p(x) = \mathbf{a}^T \boldsymbol{\varphi}(x), \quad \text{gdzie: } \mathbf{a}_{(1 \times n)} = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix}, \quad \boldsymbol{\varphi}_{(1 \times n)}(x) = \begin{bmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \dots \\ \varphi_n(x) \end{bmatrix}.$$

Współczynniki  $a_i$  wyznacza się z następującego układu równań:

$$\Phi_{(n \times n)} = \left| \varphi_j(x_i) \right| = \begin{bmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \dots & \varphi_n(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \dots & \varphi_n(x_2) \\ \dots & \dots & \dots & \dots \\ \varphi_1(x_n) & \varphi_2(x_n) & \dots & \varphi_n(x_n) \end{bmatrix}, \quad \mathbf{F}_{(1 \times n)} = \left| f_i \right| = \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{bmatrix}$$

$$\Phi \mathbf{a} = \mathbf{F}, \Rightarrow \mathbf{a} = \Phi^{-1} \mathbf{F}$$

Powyższy układ równań ma jedno rozwiązanie, gdy macierz  $\Phi$  jest nieosobliwa, a to zachodzi wtedy, gdy węzły interpolacji nie pokrywają się (wyjściowe przyporządkowanie dyskretne jest funkcją).

W przypadku interpolacji zwężanie po liczbie funkcji bazowych nie jest konieczne, gdyż jest ona równa liczbie węzłów, a więc macierz współczynników  $\Phi$  jest od samego początku macierzą kwadratową. Nie ma sensu również wprowadzać wag, gdyż z założenia wynika, iż w węzłach krzywa ma mieć ustalone z góry wartości, a więc sterowanie jej przebiegiem w węzłach jest niemożliwe (wprowadzanie wag nie będzie miało żadnego wpływu na wynik końcowy). Po wyznaczeniu współczynników można budować krzywą wg wzoru (2). Błąd interpolacji zależy od wyboru funkcji bazowych.

Należy również nadmienić, iż interpolacja podana w tej postaci nie jest najlepszą z możliwych interpolacji, mimo iż przechodzi przez wszystkie dane punkty. Kosztem tego jest jej niestabilne i niczym nie kontrolowane zachowanie między węzłami. Interpolacja słabo więc odtwarza oryginalną funkcję. Im więcej węzłów, tym większych niestabilności można się spodziewać, zwłaszcza dla interpolacji wielomianowej. Poza tym, przejście funkcji przez wszystkie punkty ściśle wcale nie musi być najlepszym rozwiązaniem, zwłaszcza przy obróbce danych eksperymentalnych, gdy każdy wynik obarczony jest błędem zupełnie zaniedbywanym w wyniku zastosowania interpolacji.

## 1. Interpolacja jednomianowa

Jest to najprostsza, ale i najbardziej prymitywna z interpolacji (wymaga rozwiązywania dużych układów równań). Znana jest w klasycznej postaci: dane jest kilka punktów, przez które ma przejść krzywa. Zapisuje się więc jej wzór wielomianowy zależny od tylu współczynników, ile jest punktów, przez które ma ona przejść. Współczynniki znajduje się z układu równań, powstałego z zapisania jej przejścia ściślego przez wszystkie punkty. Np. dla dwóch punktów  $(x_1, f_1), (x_2, f_2)$  zapisuje się wzór funkcji liniowej  $p(x) = ax + b$ , a współczynniki  $a$  i  $b$  znajduje się z warunków  $p(x_1) = f_1$  oraz  $p(x_2) = f_2$ . Dokładnie to samo postępowanie wyniknie z ogólnego schematu interpolacji, tylko ze szczególną postacią funkcji bazowych w postaci kolejnych jednomianów:

$$\varphi_1(x) = 1, \quad \varphi_2(x) = x, \quad \varphi_3(x) = x^2, \quad \varphi_4(x) = x^3, \quad \dots, \quad \varphi_n(x) = x^{n-1}.$$

$$\text{Ogólnie: } \varphi_i(x) = x^{i-1}, \quad \text{dla } i = 1, 2, \dots, n.$$

Krzywą (2) znajduje się wtedy z układu równań:

$$\Phi a = F, \Rightarrow a = \Phi^{-1} F, \text{ gdzie: } \Phi = \begin{matrix} \begin{matrix} 1 & x_1 & \dots & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{n-1} \\ 1 & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^{n-1} \end{matrix} \\ (n \times n) \end{matrix}$$

Macierz  $\Phi$  przy interpolacji jednomianowej w literaturze nosi nazwę macierzy Van Der Monda. Podobnie jak przy ogólnym sformułowaniu interpolacji, macierz  $\Phi$  jest nieosobliwa ( $\det \Phi \neq 0$ ), gdy  $\forall_{i,j} x_i \neq x_j$ .

**Przykład 1**

Dany jest zbiór punktów:

|       |   |   |   |
|-------|---|---|---|
| $i$   | 1 | 2 | 3 |
| $x_i$ | 0 | 1 | 2 |
| $f_i$ | 0 | 1 | 4 |

Dokonać interpolacji jednomianowej.

Dobieramy trzy funkcje bazowe:  $\varphi_1(x)=1$ ,  $\varphi_2(x)=x$ ,  $\varphi_3(x)=x^2$ . Przyjmujemy postać interpolacji  $p(x) = \sum_{i=1}^3 a_i \varphi_i(x) = a_1 \varphi_1(x) + a_2 \varphi_2(x) + a_3 \varphi_3(x) = a_1 + a_2 x + a_3 x^2$ .

Budujemy macierz Van Der Monda:  $\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}$  oraz układ równań:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix} \Rightarrow \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Stąd:  $p(x) = a_1 + a_2 x + a_3 x^2 = 0 + 0 \cdot x + 1 \cdot x^2 = x^2$ .

Interpolacja idealnie odtworzyła pierwotną parabolę, z której zdjęte zostały punkty.

**2. Interpolacja Lagrange’a**

W przypadku, gdy funkcjami bazowymi są wielomiany coraz wyższych stopni, wynik końcowy (krzywa interpolacyjna) jest oczywiście taki sam. Natomiast można poszukiwać go na różne sposoby. Jeden z nich pozwala na ominięcie rozwiązywania układu równań zakładając specyficzną wielomianową postać funkcji bazowych. Otóż, jeżeli przyjmie się funkcje bazowe ( $\varphi_i(x) \equiv L_i(x)$ , tzw. *wielomiany Lagrange’a*) w zależności od rozłożenia

węzłów tak, że:  $L_i(x_j) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$ , to macierz współczynników  $\Phi$  przyjmie następującą postać:

$$\Phi = \begin{bmatrix} L_1(x_1) & L_2(x_1) & \dots & L_n(x_1) \\ L_1(x_2) & L_2(x_2) & \dots & L_n(x_2) \\ \dots & \dots & \dots & \dots \\ L_1(x_n) & L_2(x_n) & \dots & L_n(x_n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \mathbf{I}$$

Układ równań będzie miał rozwiązanie:  $\mathbf{Ia} = \mathbf{F} \Rightarrow \mathbf{a} = \mathbf{F}$ .

Tak więc w przypadku tej interpolacji (tzw. *interpolacji Lagrange'a*) przy odpowiednim doborze funkcji bazowych znane są od razu współczynniki krzywej interpolacyjnej – są nimi wartości węzłowe:

$$p(x) = \sum_{i=1}^n f_i L_i(x) = f_1 L_1(x) + f_2 L_2(x) + \dots + f_n L_n(x).$$

Jedyną trudność stanowi więc znalezienie wielomianów Lagrange'a. Jest ich tyle, ile węzłów. Dowolny,  $i$ -ty wielomian zeruje się we wszystkich węzłach oprócz węzła z numerem  $i$ -tym, w którym przyjmuje wartość 1. Oczywiście pomiędzy węzłami wielomian przyjmuje wartości niezerowe. Można go opisać wzorem (tzw. *wzór interpolacyjny Lagrange'a*):

$$L_i(x) = \frac{(x-x_1) \cdot (x-x_2) \cdot \dots \cdot (x-x_{i-1}) \cdot (x-x_{i+1}) \cdot \dots \cdot (x-x_n)}{(x_i-x_1) \cdot (x_i-x_2) \cdot \dots \cdot (x_i-x_{i-1}) \cdot (x_i-x_{i+1}) \cdot \dots \cdot (x_i-x_n)} = \frac{\prod_{\substack{j=1 \\ j \neq i}}^n (x-x_j)}{\prod_{\substack{j=1 \\ j \neq i}}^n (x_i-x_j)}.$$

Licznik jest iloczynem różnic  $(x-x_j)$  tworzonym z pominięciem węzła  $x_i$ . Pojawia się on za to w mianowniku, który jest licznikiem policzonym dla  $x = x_i$ .

Błąd interpolacji Lagrange'a dla dowolnego  $x$  można określić z następującego wzoru:

$$\varepsilon(x) = \frac{f^{(n)}(\xi) \cdot \prod_{i=1}^n (x-x_i)}{n!} \leq \left| \frac{f_{\max}^{(n)} \cdot \prod_{i=1}^n (x-x_i)}{n!} \right|, \quad x_1 \leq \xi \leq x_n.$$

$f^{(n)}$  oznacza pochodną  $n$ -tego rzędu, natomiast  $\xi$  jest punktem pośrednim z przedziału, w którym dokonuje się interpolacji.

Uogólnieniem interpolacji Lagrange'a jest *interpolacja l'Hermitte'a*, w której w węzłach obok wartości funkcji mogą być również dane wartości pochodnych.

## Przykład 2

Dany jest zbiór punktów:

|       |   |   |   |
|-------|---|---|---|
| $i$   | 1 | 2 | 3 |
| $x_i$ | 0 | 1 | 2 |
| $f_i$ | 0 | 1 | 4 |

Dokonać interpolacji Lagrange'a.

Budujemy kolejne wielomiany Lagrange'a:

$$L_1(x) = \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} = \frac{(x-1)(x-2)}{(0-1)(0-2)} = \frac{1}{2}(x-1)(x-2)$$

$$L_2(x) = \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} = \frac{(x-0)(x-2)}{(1-0)(1-2)} = -x(x-2)$$

$$L_3(x) = \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)} = \frac{(x-0)(x-1)}{(2-0)(2-1)} = \frac{1}{2}x(x-1)$$

Wzór interpolacyjny:

$$p(x) = f_1L_1(x) + f_2L_2(x) + f_3L_3(x)$$

$$p(x) = 0 \cdot \frac{1}{2}(x-1)(x-2) + 1 \cdot (-x)(x-2) + 4 \cdot \frac{1}{2}x(x-1) = -x^2 + 2x + 2x^2 - 2x = x^2$$

Błąd interpolacji jest równy 0 dla dowolnego  $x$  z uwagi, iż pochodna rzędu  $n = 3$  wyjściowej funkcji  $f(x) = x^2$  jest równa  $f'''(x) \equiv 0$ .

### Przykład 3

Dokonać interpolacji Lagrange'a funkcji ciągłej  $f(x) = \sin(x)$  w przedziale  $\langle 2, 4 \rangle$  stosując różne liczby węzłów. Wyznaczyć błąd interpolacji. Obliczyć wartość wielomianu interpolacyjnego dla  $x_0 = \pi$  dla  $i$  i porównać z wynikiem ścisłym.

W podanym przedziale dokonujemy dyskretyzacji funkcji za pomocą  $n = 3$  węzłów równomiernie rozłożonych. Otrzymujemy następujące punkty:

|                   |          |          |           |
|-------------------|----------|----------|-----------|
| $i$               | 1        | 2        | 3         |
| $x_i$             | 2        | 3        | 4         |
| $f_i = \sin(x_i)$ | 0.909297 | 0.141120 | -0.756802 |

Budujemy wielomiany Lagrange'a:

$$L_1(x) = \frac{(x-3)(x-4)}{(2-3)(2-4)} = \frac{1}{2}(x-3)(x-4), \quad L_2(x) = \frac{(x-2)(x-4)}{(3-2)(3-4)} = -(x-2)(x-4),$$

$$L_3(x) = \frac{(x-2)(x-3)}{(4-2)(4-3)} = \frac{1}{2}(x-2)(x-3)$$

Budujemy interpolację:

$$p(x) = 0.909297 \cdot \frac{1}{2}(x-3)(x-4) + 0.141120 \cdot (-1) \cdot (x-2)(x-4) - 0.756802 \cdot \frac{1}{2} \cdot (x-2)(x-3) =$$

$$= -0.06487x^2 - 0.443815x + 2.056416$$

Wartość interpolacji dla  $x_0 = \pi$ :  $p_0 = p(x_0 = \pi) = 0.021828$ .

Wartość ścisła dla  $x_0 = \pi$ :  $f_0 = 0.0$ .

Błąd bezwzględny wyniku:  $\varepsilon_0 = |p_0 - f_0| = |0 - 0.021828| = 0.021828$ .

Oszacowanie błędu interpolacji:

$$f'''(x) = -\sin(x), \quad f'''_{\max} = f'''(x=2) = -0.909297$$

$$\varepsilon(x) \leq \left| -0.909297 \frac{(x-2)(x-3)(x-4)}{6} \right| = |0.151550 \cdot (x-2)(x-3)(x-4)|$$

Błąd interpolacji dla  $x_0 = \pi$  wynosi:

$$\varepsilon_0 = \varepsilon(x_0 = \pi) \leq |0.151550 \cdot (\pi-2)(\pi-3)(\pi-4)| = 0.02103.$$

Wynik ulega istotnej poprawie dla większej liczby węzłów: dla  $n=4$   $p_0 = 0.000404$ , a dla  $n=5$   $p_0 = 0.000256$ .

### 3. Odwrotna interpolacja Lagrange'a

Zamiast budować interpolację na zmiennych niezależnych  $x$ , można odwrócić miejscami zmienne  $x$  z  $y$  i znaleźć w rezultacie wielomian interpolacyjny  $p(y)$ :

Dla danych punktów węzłowych:  $(x_i, f_i)$ ,  $i = 1, 2, \dots, n$  budujemy  $n$  wielomianów Lagrange'a, ale traktując  $y$  jako zmienną niezależną:

$$L_i(y) = \frac{(y-f_1) \cdot (y-f_2) \cdot \dots \cdot (y-f_{i-1}) \cdot (y-f_{i+1}) \cdot \dots \cdot (y-f_n)}{(f_i-f_1) \cdot (f_i-f_2) \cdot \dots \cdot (f_i-f_{i-1}) \cdot (f_i-f_{i+1}) \cdot \dots \cdot (f_i-f_n)} = \frac{\prod_{\substack{j=1 \\ j \neq i}}^n (y-f_j)}{\prod_{\substack{j=1 \\ j \neq i}}^n (f_i-f_j)}$$

oraz stosujemy zmodyfikowany wzór interpolacyjny Lagrange'a:

$$p(y) = \sum_{i=1}^n x_i L_i(y) = x_1 L_1(y) + x_2 L_2(y) + \dots + x_n L_n(y)$$

Teraz można odtworzyć, jaki oryginalnie  $x_0$  był przypisany danemu  $y_0$  poprzez obliczenie  $x_0 = p(y_0)$ . Metoda odwrotna może być też dobrym przybliżeniem metod iteracyjnych do znajdowania pierwiastka równania algebraicznego  $f(x) = 0$ . Wtedy budując interpolację odwrotną na zbiorze punktów funkcji  $f(x)$  w przedziale  $a \leq x \leq b$  można oszacować z dobrym przybliżeniem miejsce zerowe oryginalnej funkcji  $f(x)$  poprzez obliczenie  $x^* = p(y=0)$ .

Uwaga! Warunkiem rozwiązywalności zadania jest różnowartościowość funkcji  $f(x)$ .

#### Przykład 4

Znaleźć przybliżenie miejsca zerowego równania  $x - \sin(x) = 0$  w przedziale  $x \in (\frac{\pi}{2}, \frac{3}{2}\pi)$ .

W podanym przedziale wprowadzamy  $n=3$  węzły, dobierając wartości węzłowe na podstawie równania  $f(x) = x - \sin(x)$ .

|                |                     |       |                      |
|----------------|---------------------|-------|----------------------|
| $i$            | 1                   | 2     | 3                    |
| $x_i$          | $\frac{\pi}{2}$     | $\pi$ | $\frac{3}{2}\pi$     |
| $f_i = f(x_i)$ | $\frac{\pi}{2} - 1$ | $\pi$ | $\frac{3}{2}\pi + 1$ |

Budujemy na wartościach węzłowych wielomiany Lagrange'a:

$$L_1(y) = \frac{(y-f_2)(y-f_3)}{(f_1-f_2)(f_1-f_3)} = \frac{(y-\pi)(y-\frac{3}{2}\pi-1)}{(\frac{\pi}{2}-1-\pi)(\frac{\pi}{2}-1-\frac{3}{2}\pi-1)} = \frac{2}{(\pi+2)^2}(y-\pi)(y-\frac{3}{2}\pi-1)$$

$$L_2(y) = \frac{(y-f_1)(y-f_3)}{(f_2-f_1)(f_2-f_3)} = \frac{(y-\frac{\pi}{2}+1)(y-\frac{3}{2}\pi-1)}{(\pi-\frac{\pi}{2}+1)(\pi-\frac{3}{2}\pi-1)} = -\frac{4}{(2+\pi)^2}(y-\frac{\pi}{2}+1)(y-\frac{3}{2}\pi-1)$$

$$L_3(y) = \frac{(y-f_1)(y-f_2)}{(f_3-f_1)(f_3-f_2)} = \frac{(y-\pi)(y-\frac{\pi}{2}+1)}{(\frac{3}{2}\pi+1-\pi)(\frac{3}{2}\pi+1-\frac{\pi}{2}+1)} = \frac{2}{(\pi+2)^2}(y-\pi)(y-\frac{\pi}{2}+1)$$

oraz wzór interpolacyjny:

$$p(y) = x_1L_1(y) + x_2L_2(y) + x_3L_3(y)$$

$$p(y) = \frac{\pi}{2} \cdot \frac{2}{(\pi+2)^2}(y-\pi)(y-\frac{3}{2}\pi-1) + \pi \cdot (-1) \frac{4}{(2+\pi)^2}(y-\frac{\pi}{2}+1)(y-\frac{3}{2}\pi-1) + \frac{3}{2}\pi \cdot \frac{2}{(\pi+2)^2}(y-\pi)(y-\frac{\pi}{2}+1) = \pi \frac{y+2}{\pi+2}$$

Przybliżenie miejsca zerowego równania:  $x^* \approx p(0) = \frac{2\pi}{2+\pi} = 1.222031$ .

#### 4. Wielomiany Czebyszewa

Interpolacja wielomianowa funkcji dyskretnej daje wyniki ścisłe, gdy interpolowany jest wielomian, co najwyżej stopnia  $n-1$ . Dla stopni wyższych oraz dla wyjściowych funkcji niebędących wielomianami wyniki są w jakiś sposób przybliżone. Dla wysokich stopni interpolacji krzywe wielomianowe są niestabilne, tzn. mimo przejścia ścisłego przez wszystkie punkty między nimi zaczynają coraz bardziej się rozbiegać do nieskończoności. Aby zapewnić maksymalną stabilność takich wyników stosuje się jako funkcje bazowe wielomiany ortogonalne (lub ortogonalne z wagą) np. funkcje specjalne Lagrange'a (nie mylić z wcześniej omawianymi wielomianami Lagrange'a), l'Hermitte'a, Legendre'a czy Czebyszewa. Te ostatnie mają jeszcze jedną bardzo ważną dla aproksymacji własność: jeżeli mianowicie tak dobierze się węzły aproksymacji, aby były one równe miejscom zerowym odpowiedniego wielomianu Czebyszewa, to wtedy maksymalny błąd tak zbudowanej interpolacji wielomianowej zostanie zminimalizowany:

$$\text{Błąd maksymalny interpolacji: } \varepsilon(x) \leq \left| f_{\max}^{(n)} \right| \cdot \left| \prod_{i=1}^n (x-x_i) \right|$$

Znaleźć minimum maksymalnej wartości w przedziale  $\langle -1,1 \rangle$  z iloczynu  $\prod_{i=1}^n (x-x_i)$ ,

czyli:  $\min_x \max_{-1 \leq x \leq 1} \left| \prod_{i=1}^n (x-x_i) \right|$  - oryginalne zagadnienie Czebyszewa.



Wielomiany Czebyszewa można określić na dwa sposoby:

- Sposób iteracyjny:  $T_n(x) = \cos(n \cdot \arccos x)$ ,
- Sposób rekurencyjny: 
$$\begin{cases} T_0(x) = 1 \\ T_1(x) = x \\ T_n(x) = 2 \cdot x \cdot T_{n-1}(x) - T_{n-2}(x) \end{cases}$$

Powyższe wzory obowiązują w przedziale  $-1 \leq x \leq 1$ . To przedział, w którym wielomiany Czebyszewa są określone i w którym są ortogonalne.

W konkretnych zastosowaniach bardziej korzystny jest wzór rekurencyjny, gdzie dany wielomian oblicza się na podstawie dwóch poprzednich. Dla przykładu pokazano kilka następujących wielomianów Czebyszewa:

$$T_2(x) = 2 \cdot x \cdot T_1(x) - T_0(x) = 2 \cdot x \cdot x - 1 = 2x^2 - 1$$

$$T_3(x) = 2 \cdot x \cdot T_2(x) - T_1(x) = 2 \cdot x \cdot (2x^2 - 1) - x = 4x^3 - 3x$$

$$T_4(x) = 8x^4 - 8x^2 + 1$$

$$T_5(x) = 16x^5 - 20x^3 + 5$$

Aby znaleźć miejsca zerowe  $n$ -tego wielomianu Czebyszewa, nie trzeba rozwiązywać w tym celu równania  $T_n(x) = 0$ ; można posłużyć się gotowym wzorem:

$$x_i = \cos \frac{2 \cdot i + 1}{n} \frac{\pi}{2}, \quad i = 0, 1, \dots, n-1.$$

Własność ortogonalności wielomianów Czebyszewa z wagą  $\mu(x) = \frac{1}{\sqrt{1-x^2}}$  polega na tym, iż całka:

$$I_{ij} = \int_{-1}^1 \frac{T_i(x) \cdot T_j(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0, & i \neq j \\ \frac{\pi}{2}, & i = j \neq 0 \\ \pi, & i = j = 0 \end{cases}$$

Ponieważ w konkretnych zadaniach mamy do czynienia z dowolnym przedziałem  $x$ , dlatego też zachodzi często potrzeba transformacji wyjściowego przedziału do przedziału, w którym znane są wielomiany Czebyszewa i odwrotnie:

Niech  $z \in \langle a, b \rangle$ ,  $x \in \langle -1, 1 \rangle$ :

- Przejście  $z \rightarrow x$  :  $x = \frac{2z - (b+a)}{b-a}$ ,
- Przejście  $x \rightarrow z$  :  $z = \frac{1}{2}[(b-a) \cdot x + (b+a)]$ .

Uwaga! W zadaniach interpolacji można bazować na zadanej siatce węzłów a tylko jako funkcji bazowych użyć wielomianów Czebyszewa (tzw. *interpolacja Czebyszewa*), albo przyjąć węzły jako miejsca zerowe odpowiedniego wielomianu Czebyszewa a interpolować używając do tego jednej z poznanych metod (w tym także interpolacji Czebyszewa). To samo dotyczy także aproksymacji funkcji.

### Przykład 5

Dana jest funkcja dyskretna  $(z_i, f_i)$ ,  $i = 1, 2, 3$ , taka jak w przykładach 1 i 2:

|       |   |   |   |
|-------|---|---|---|
| $i$   | 1 | 2 | 3 |
| $z_i$ | 0 | 1 | 2 |
| $f_i$ | 0 | 1 | 4 |

Dokonać interpolacji Czebyszewa.

Węzłów nie wyznaczamy – są z góry podane. Do interpolacji na trzech węzłach potrzebne będą trzy wielomiany Czebyszewa (w przedziale  $x \in \langle -1, 1 \rangle$ ):

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1.$$

Wzory na transformację między przedziałami  $z \in \langle 0, 2 \rangle$ ,  $x \in \langle -1, 1 \rangle$ :  $x = z - 1$ ,  $z = x + 1$ .

Wielomiany Czebyszewa w przedziale  $z \in \langle 0, 2 \rangle$ :

$$T_0(z) = 1, \quad T_1(z) = z - 1, \quad T_2(z) = 2(z - 1)^2 - 1 = 2z^2 - 4z + 1.$$

Tworzymy układ równań:

$$\Phi = \begin{bmatrix} T_0(z_1) & T_1(z_1) & T_2(z_1) \\ T_0(z_2) & T_1(z_2) & T_2(z_2) \\ T_0(z_3) & T_1(z_3) & T_2(z_3) \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}$$

i rozwiązujemy:

$$\Phi \cdot \mathbf{a} = \mathbf{F} \Rightarrow \mathbf{a} = \begin{bmatrix} 1.5 \\ 2 \\ 0.5 \end{bmatrix}$$

Wzór interpolacyjny:

$$p(x) = a_0 T_0(z) + a_1 T_1(z) + a_2 T_2(z) = \frac{3}{2} \cdot 1 + 2 \cdot (z - 1) + \frac{1}{2} (2z^2 - 4z + 1) = z^2$$

Otrzymany wzór odtwarza pierwotną parabolę, tak samo jak w przypadku interpolacji jednomianowej i Lagrange'a.

### Przykład 6

Dokonać interpolacji funkcji  $f(z) = \sqrt{1+z^2}$  w przedziale  $z \in \langle 0, 5 \rangle$ . Jako funkcje bazowe przyjąć wielomiany Czebyszewa, a jako węzły interpolacji miejsca zerowe wielomianu  $T_3(x)$ .

Zacznijmy od węzłów interpolacji w przedziale  $x \in \langle -1, 1 \rangle$ . Wielomian  $T_3(x)$  ma trzy miejsca zerowe, co od razu implikuje trzy węzły a więc interpolację parabolą. Korzystamy ze wzoru na miejsca zerowe:

$$x_i = \cos \frac{2 \cdot i + 1}{3} \frac{\pi}{2}, \quad i = 0, 1, 2$$

$$\left\{ \begin{array}{l} x_0 = \cos \frac{2 \cdot 0 + 1}{3} \frac{\pi}{2} = \cos \frac{\pi}{6} = \frac{\sqrt{3}}{2} = 0.866025 \\ x_1 = \cos \frac{2 \cdot 1 + 1}{3} \frac{\pi}{2} = \cos \frac{\pi}{2} = 0 \\ x_2 = \cos \frac{2 \cdot 2 + 1}{3} \frac{\pi}{2} = \cos \frac{5\pi}{6} = -\frac{\sqrt{3}}{2} = -0.866025 \end{array} \right.$$

Natomiast wielomiany potrzebne do wzoru interpolacyjnego:

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1$$

Wzory na transformację między przedziałami  $z \in \langle 0, 5 \rangle$ ,  $x \in \langle -1, 1 \rangle$ :

$$x(z) = \frac{2}{5}z - 1, \quad z(x) = \frac{5}{2}(x + 1).$$

Miejsca zerowe i wielomiany w przedziale  $z \in \langle 0, 5 \rangle$ :

$$z_0 = \frac{5}{2}(x_0 + 1) = \frac{5}{4}(\sqrt{3} + 2) = 4.665064$$

$$z_1 = \frac{5}{2}(x_1 + 1) = \frac{5}{2} = 2.50$$

$$z_2 = \frac{5}{2}(x_2 + 1) = \frac{5}{4}(2 - \sqrt{3}) = 0.334936$$

$$T_0(z) = 1, \quad T_1(z) = \frac{2}{5}z - 1, \quad T_2(z) = 2 \cdot \left(\frac{2}{5}z - 1\right)^2 - 1 = \frac{8}{25}z^2 - \frac{8}{5}z + 1$$

Dyskretyzacja funkcji  $f(z) = \sqrt{1 + z^2}$  (węzły ułożono w kolejności rosnącej):

|                |          |          |          |
|----------------|----------|----------|----------|
| $i$            | 1        | 2        | 3        |
| $z_i$          | 0.334936 | 2.50     | 4.665064 |
| $f_i = f(z_i)$ | 1.054600 | 2.692582 | 4.771040 |

Budowa i rozwiązanie układu równań:

$$\Phi = \begin{bmatrix} T_0(z_0) & T_1(z_0) & T_2(z_0) \\ T_0(z_1) & T_1(z_1) & T_2(z_1) \\ T_0(z_2) & T_1(z_2) & T_2(z_2) \end{bmatrix} = \begin{bmatrix} 1 & -0.866025 & 0.5 \\ 1 & 0 & -1 \\ 1 & 0.866025 & 0.5 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} 1.054600 \\ 2.692582 \\ 4.771040 \end{bmatrix}$$

$$\Phi \cdot \mathbf{a} = \mathbf{F} \Rightarrow \mathbf{a} = \begin{bmatrix} 2.839407 \\ 2.145689 \\ 0.146825 \end{bmatrix}.$$

Wzór interpolacyjny:

$$p(z) = a_0 T_0(z) + a_1 T_1(z) + a_2 T_2(z) = 2.839407 \cdot 1 + 2.145689 \cdot \left(\frac{2}{5}z - 1\right) + 0.146825 \cdot \left(\frac{8}{25}z^2 - \frac{8}{5}z + 1\right) \\ = 0.046984 \cdot z^2 + 0.623355 \cdot z + 0.840544$$

Sprawdzenie własności interpolacyjnych wielomianu  $p(z)$ :

$$p_1 = p(z_1 = 0.334936) = 1.054600 = f_1,$$

$$p_2 = p(z_2 = 2.50) = 2.692582 = f_2,$$

$$p_3 = p(z_3 = 4.665064) = 4.771040 = f_3$$

Obliczenie średniego błędu interpolacji:

$$\varepsilon_{avr} = \int_0^5 [p(z) - f(z)] dz = \int_0^5 (0.046984 \cdot z^2 + 0.623355 \cdot z + 0.840544 - \sqrt{1+z^2}) dz = 0.048560$$

Oszacowanie maksymalnego błędu interpolacji:

$$f'''(z) = -3 \frac{z}{(1+z^2)^{\frac{5}{2}}}, \quad f'''_{\max} = f'''(z = \frac{1}{2}) = -0.85865$$

$$\varepsilon(z) \leq \left| -0.85865 \frac{(z-0.334936)(z-2.50)(z-4.665064)}{6} \right| = \\ = \left| 0.143108 \cdot (z-0.334936)(z-2.50)(z-4.665064) \right|$$

Np. dla  $z = 2$  oryginalna wartość funkcji wynosi  $f = \sqrt{1+2^2} = 2.236068$ , a ta pochodząca z interpolacji  $p = p(2) = 2.27519$ . Oszacowanie błędu  $\varepsilon(2) \leq 0.317521$ .

## 5. Interpolacja funkcjami sklejanymi (funkcje typu *spline*)

Przy wzroście liczby węzłów interpolacja daje niepożądane efekty międzywęzłowe w postaci coraz większych gradientów funkcji interpolującej. Aby temu zapobiec i jednocześnie zachować własności interpolacyjne, wprowadzono interpolację funkcjami sklejanymi. Polega ona na znalezieniu krzywej niskiego stopnia, składającej się z różnych kawałków, (czyli o różnych wzorach analitycznych) na przedziałach wyznaczonych przez kolejne pary węzłów. Dodatkowo wymaga się odpowiednich warunków ciągłości: funkcja sklejana (spline) rzędu  $k$  ma we wszystkich przedziałach wszystkie pochodne ciągłe aż do rzędu  $k-1$  włącznie.

Rozważmy zbiór punktów  $(x_i, f_i)$ ,  $i = 1, 2, \dots, n$ . Każdy spline rzędu  $k$  ma pierwszym odcinku

$x \in \langle x_1, x_2 \rangle$  wzór:  $p(x) = a_k \cdot x^{k-1} + a_{k-1} \cdot x^{k-2} + \dots + x \cdot a_2 + a_1 = \sum_{i=1}^{k+1} a_i x^{k+1-i}$ . Następnie wraz z przekraczaniem kolejnych węzłów dochodzą następujące składniki wielomianowe:

$$p(x) + b_2(x-x_2)^k \quad \text{dla } x \in \langle x_2, x_3 \rangle$$

$$p(x) + b_2(x-x_2)^k + b_3(x-x_3)^k \quad \text{dla } x \in \langle x_3, x_4 \rangle \text{ itd.}$$

Ogólnie spline rzędu  $k$  można zapisać jednym ogólnym wzorem:

$$s(x) = p(x) + \sum_{i=2}^{n-1} b_i(x-x_i)_+^k = \sum_{i=1}^{k+1} a_i x^{k+1-i} + \sum_{i=2}^{n-1} b_i(x-x_i)_+^k, \quad (x-x_i)_+^k = \begin{cases} (x-x_i)^k, & \text{dla } x > x_i \\ 0, & \text{dla } x \leq x_i \end{cases}$$

W każdym spinie są niewiadome współczynniki  $a_i$ ,  $i=1,2,\dots,k+1$  i  $b_i$ ,  $i=2,3,\dots,n-1$ . Razem niewiadomych jest  $n-1+k$ . Począwszy od  $k=2$  (kiedy niewiadomych jest  $n-1+2=n+1$ ) same równania pochodzące od punktów przez które krzywa ma przejść są niewystarczające. Wprowadza się więc dodatkowe warunki na pochodne spline'u w węzłach. I tak spline rzędu  $k=1$  (spline liniowy) nie wymaga znajomości żadnych dodatkowych warunków), spline rzędu  $k=2$  (spline kwadratowy, paraboliczny) wymaga znajomości wartości pochodnej w którymś z węzłów, tj.  $s'(x_j) = \alpha$ , natomiast spline rzędu  $k=3$  wymaga znajomości wartości pierwszej i drugiej pochodnej w wybranych dwóch węzłach (może być w tym samym), tj.  $s'(x_j) = \alpha$ ,  $s''(x_l) = \beta$  ( $j, l \in \{1, 2, \dots, n\}$ ). Jeżeli informacje o pochodnych są podane w węzłach pierwszego przedziału  $x \in \langle x_1, x_2 \rangle$  (tam gdzie obowiązuje przepis  $s(x) = p(x)$ ), to współczynniki  $a_i$  można wyznaczyć niezależnie (z układu równań) od współczynników  $b_i$  (ze wzoru rekurencyjnego). Jeżeli natomiast warunki brzegowe nie pozwalają na jednoznaczne wyznaczenie odcinka krzywej w przedziale  $x \in \langle x_1, x_2 \rangle$ , to wtedy nie można wyznaczyć rekurencyjnie współczynników  $b_i$ , lecz trzeba zbudować w ten sposób układ równań na niewiadome współczynniki  $a_i$  i  $b_i$ . Dalej rozważany będzie przypadek pierwszy: wszystkie wartości pochodnych dane są w pierwszym węźle ( $x = x_1$ ).

Ogólne wzory na spline (dla  $k = 1, 2, 3$ ):

- Spline liniowy:  $s(x) = a_1x + a_2 + \sum_{i=2}^{n-1} b_i(x-x_i)_+$ ,
- Spline kwadratowy:  $s(x) = a_1x^2 + a_2x + a_3 + \sum_{i=2}^{n-1} b_i(x-x_i)_+^2$ ,
- Spline sześcienny:  $s(x) = a_1x^3 + a_2x^2 + a_3x + a_4 + \sum_{i=2}^{n-1} b_i(x-x_i)_+^3$ .

Wyznaczenie współczynników  $a_i$ ,  $i = 1, 2, \dots, k+1$ :

- Poprzez zapisanie warunków interpolacji spline'u na pierwszym przedziale  $x \in \langle x_1, x_2 \rangle$  oraz poprzez wykorzystanie ewentualnych dodatkowych informacji o pochodnych w tych węzłach:

- Dla spline'u liniowego: 
$$\begin{cases} s(x_1) = f_1 \\ s(x_2) = f_2 \end{cases} \Rightarrow \begin{cases} a_1x_1 + a_2 = f_1 \\ a_1x_2 + a_2 = f_2 \end{cases} \Rightarrow \begin{cases} a_1 = \dots \\ a_2 = \dots \end{cases}$$

- Dla spline'u kwadratowego:

$$\begin{cases} s(x_1) = f_1 \\ s(x_2) = f_2 \\ s'(x_1) = \alpha \end{cases} \Rightarrow \begin{cases} a_1x_1^2 + a_2x_1 + a_3 = f_1 \\ a_1x_2^2 + a_2x_2 + a_3 = f_2 \\ 2a_1x_1 + a_2 = \alpha \end{cases} \Rightarrow \begin{cases} a_1 = \dots \\ a_2 = \dots \\ a_3 = \dots \end{cases}$$

- Dla spline'u sześciennego:

$$\begin{cases} s(x_1) = f_1 \\ s(x_2) = f_2 \\ s'(x_1) = \alpha \\ s''(x_1) = \beta \end{cases} \Rightarrow \begin{cases} a_1x_1^3 + a_2x_1^2 + a_3x_1 + a_4 = f_1 \\ a_1x_2^3 + a_2x_2^2 + a_3x_2 + a_4 = f_2 \\ 3a_1x_1^2 + 2a_2x_1 + a_3 = \alpha \\ 6a_1x_1 + 2a_2 = \beta \end{cases} \Rightarrow \begin{cases} a_1 = \dots \\ a_2 = \dots \\ a_3 = \dots \\ a_4 = \dots \end{cases}$$

Wyznaczenie współczynników  $b_i$ ,  $i = 2, 3, \dots, n-1$ :

- Ze wzoru rekurencyjnego niezależnie od rzędu spline'u; wzór wyprowadza się wykorzystując pozostałe warunki na spline począwszy od  $x = x_3$ :

$$\text{dla } x = x_3: s(x_3) = p(x_3) + b_2(x_3 - x_2)^k = f_3 \rightarrow b_2 = \frac{f_3 - p(x_3)}{(x_3 - x_2)^k}$$

$$\text{dla } x = x_4: s(x_4) = p(x_4) + b_2(x_4 - x_2)^k + b_3(x_4 - x_3)^k = f_4 \rightarrow b_3 = \frac{f_4 - p(x_4) - b_2(x_4 - x_2)^k}{(x_4 - x_3)^k}$$

itd. Ogólnie dla  $x = x_{j+1}$ ,  $j = 2, 3, \dots, n-1$ :

$$s(x_{j+1}) = p(x_{j+1}) + \sum_{i=2}^j b_i(x_{j+1} - x_i)^k = f_{j+1} \rightarrow p(x_{j+1}) + \sum_{i=2}^{j-1} b_i(x_{j+1} - x_i)^k + b_j(x_{j+1} - x_j)^k = f_{j+1}$$

$$b_j = \frac{f_{j+1} - p(x_{j+1}) - \sum_{i=2}^{j-1} b_i(x_{j+1} - x_i)^k}{(x_{j+1} - x_j)^k}.$$

### Przykład 7

Dla danych z poprzednich przykładów znaleźć spline liniowy.

|       |   |   |   |
|-------|---|---|---|
| $i$   | 1 | 2 | 3 |
| $x_i$ | 0 | 1 | 2 |
| $f_i$ | 0 | 1 | 4 |

Wzór ogólny spline'u:  $s(x) = p(x) + \sum_{i=2}^{3-1} b_i(x - x_i)_+ = a_1x + a_2 + b_2(x-1)_+$ .

Wyznaczenie współczynników  $a_1, a_2$ :

$$\begin{cases} s(0) = 0 \\ s(1) = 1 \end{cases} \Rightarrow \begin{cases} a_2 = 0 \\ a_1 + a_2 = 1 \end{cases} \Rightarrow \begin{cases} a_1 = 1 \\ a_2 = 0 \end{cases} \Rightarrow p(x) = x$$

Wyznaczenie współczynnika  $b_2$ :

$$s(2) = 4 \Rightarrow p(2) + b_2(2-1) = 4 \Rightarrow b_2 = 4 - 2 = 2$$

Wyznaczenie wzoru na spline:

$$s(x) = x + 2 \cdot (x-1)_+ = \begin{cases} x, & \text{dla } 0 \leq x \leq 1 \\ 3x - 2, & \text{dla } 1 < x \leq 2 \end{cases}$$

### Przykład 8

Dla danych z poprzedniego przykładu znaleźć spline kwadratowy.

|       |   |   |   |
|-------|---|---|---|
| $i$   | 1 | 2 | 3 |
| $x_i$ | 0 | 1 | 2 |
| $f_i$ | 0 | 1 | 4 |

Dołączamy informację o pochodnej spline'u dla  $x = 0 \rightarrow s'(0) = \alpha = 0$ .

$$\text{Wzór ogólny spline'u: } \begin{cases} s(x) = p(x) + \sum_{i=2}^{3-1} b_i (x-x_i)_+^2 = a_1 x^2 + a_2 x + a_3 + b_2 (x-1)_+^2 \\ s'(x) = p'(x) + 2 \sum_{i=2}^{3-1} b_i (x-x_i)_+ = 2a_1 x + a_2 + 2b_2 (x-1)_+ \end{cases}$$

Wyznaczenie współczynników  $a_1, a_2, a_3$ :

$$\begin{cases} s(0) = 0 \\ s(1) = 1 \\ s'(0) = 0 \end{cases} \Rightarrow \begin{cases} a_3 = 0 \\ a_1 + a_2 + a_3 = 1 \\ a_2 = 0 \end{cases} \Rightarrow \begin{cases} a_1 = 1 \\ a_2 = 0 \\ a_3 = 0 \end{cases} \Rightarrow p(x) = x^2$$

Wyznaczenie współczynnika  $b_2$ :

$$s(2) = 4 \Rightarrow p(2) + b_2 (2-1)^2 = 4 \Rightarrow b_2 = 4 - 2^2 = 0$$

Wyznaczenie wzoru na spline:

$$s(x) = x^2 + 0 \cdot (x-1)_+^2 = x^2 \quad \text{dla } 0 \leq x \leq 2.$$

W ostatnim przykładzie tylko pozornie interpolacja jest sklejana. Ponieważ dane pochodzą od funkcji kwadratowej, to spline kwadratowy przeistoczył się w oryginalną funkcję o jednym przepisie dla wszystkich  $x$ .

## 6. Najlepsza aproksymacja

Aproksymacja to takie dopasowanie krzywej  $p(x)$  stopnia  $m$ -tego ( $m \leq n-1$ ) do zestawu danych punktów  $(x_i, f_i)$ ,  $i = 1, 2, \dots, n$ , że krzywa aproksymacyjna w ogólności przez żaden punkt ściśle nie przejdzie, dopuszczając odchyłkę między oryginalną wartością  $f_i$ , a wartością na krzywej  $p(x_i) \neq f_i$ . Ogólnym założeniem podejścia *najlepszej aproksymacji* jest minimalizacja sumarycznego błędu (sumy odchyłek) w sensie jakiejś normy. Jeżeli zastosowaną normą jest norma Euklidesa (średnio kwadratowa) to metoda nazywa się *metodą najmniejszych kwadratów*.

$$\text{Aproksymacja: } p(x) = \sum_{i=0}^m a_i \varphi_i(x).$$

$$\text{Błąd aproksymacji: } \varepsilon(x) = f(x) - p(x), \quad \text{dla } x_1 \leq x \leq x_n.$$

$$\text{Najlepsza aproksymacja: } \min_{a_i} \|\varepsilon(x)\| = \min_{a_i} \left\| f(x) - \sum_{i=0}^m a_i \varphi_i(x) \right\|$$

- Metoda min-max:  $\|\varepsilon(x)\|_\infty = \max |\varepsilon(x)| \rightarrow \min_{a_i} \max_x |f(x) - p(x)|$ ,
- Metoda najmniejszych kwadratów:  $\min_{a_i} \|\varepsilon(x)\|_2$ :
  - Dla zbioru ciągłego:  $\|\varepsilon(x)\|_2 = \left( \int_{x_1}^{x_n} \varepsilon^2(x) dx \right)^{\frac{1}{2}}$ ,
  - Dla zbioru dyskretnego:  $\|\varepsilon(x)\|_2 = \left( \sum_{i=1}^n \varepsilon^2(x_i) \right)^{\frac{1}{2}}$ .

Najpopularniejszą bazę funkcji bazowej dla aproksymacji stanowią wielomiany, w tym najchętniej używa się funkcji ortogonalnych (lub przynajmniej ortogonalnych wag), takich

jak wielomiany Czebyszewa, Bessela, Legendre'a czy Hankela. Korzysta się też z bazy jednomianowej, zwłaszcza dla aproksymacji dyskretnej. O jednomianach jako funkcjach bazowych będzie dalej mowa. Funkcja aproksymująca będzie miała wtedy postać:

$$p(x) = \sum_{i=0}^m a_i x^{m-i} = a_0 x^m + a_1 x^{m-1} + \dots + a_{m-1} x + a_m$$

Współczynniki liczbowe  $a_i$ ,  $i=0, 2, \dots, m$  należy wyznaczyć na podstawie minimalizacji sumarycznego błędu w każdym z węzłów w sensie normy średnio kwadratowej.

Układamy funkcjonal zbierający informacje o wszystkich węzłach do jednego wzoru:

$$B(a_0, a_1, \dots, a_m) = (p(x_1) - f_1)^2 + (p(x_2) - f_2)^2 + \dots + (p(x_n) - f_n)^2 = \sum_{i=1}^n (p(x_i) - f_i)^2$$

$$B(a_0, a_1, \dots, a_m) = \sum_{i=1}^n (p(x_i) - f_i)^2 = \sum_{i=1}^n \left( \sum_{j=0}^m a_j x_i^{m-j} - f_i \right)^2$$

W celu wyznaczenia niewiadomych współczynników układamy równania będące pochodnymi powyższego funkcjonału względem każdego z nich:

$$\frac{\partial}{\partial a_k} B(a_0, a_1, \dots, a_m) = 2 \sum_{i=1}^n \left( \sum_{j=0}^m a_j x_i^{m-j} - f_i \right) x_i^{m-k} = 0, \quad \text{dla } k=0, 1, 2, \dots, m$$

Z układu równań  $(m+1) \times (m+1)$  wyznaczamy współczynniki, a następnie wyznaczamy  $p(x)$ :

$$\sum_{i=1}^n \left( \sum_{j=0}^m a_j x_i^{m-j} \right) \cdot x_k^{m-k} = \sum_{i=1}^n f_i \cdot x_k^{m-k} \Rightarrow \begin{cases} a_0 = \dots \\ a_1 = \dots \\ \dots \\ a_m = \dots \end{cases} \Rightarrow p(x) = \sum_{i=0}^m a_i x^{m-i}.$$

Zmodyfikowana metoda ważona polega na przypisaniu każdemu z węzłów liczby (wagi)  $w_i$ ,  $i=1, 2, \dots, n$  świadczącej o stopniu odejścia krzywej od wartości węzłowej: im waga większa waga, tym w rezultacie bliżej krzywa przejdzie obok punktu z tą wagą. Funkcjonał wzbogacony o wagi wygląda następująco:

$$B(a_0, a_1, \dots, a_m) = w_1 \cdot (p(x_1) - f_1)^2 + w_2 \cdot (p(x_2) - f_2)^2 + \dots + w_n \cdot (p(x_n) - f_n)^2 = \sum_{i=1}^n w_i \cdot (p(x_i) - f_i)^2$$

Dalsze operacje są identyczne, co prowadzi do układu równań ( $k=0, 1, 2, \dots, m$ ):

$$\sum_{i=1}^n \left( w_i \cdot \sum_{j=0}^m a_j x_i^{m-j} \right) \cdot x_k^{m-k} = \sum_{i=1}^n w_i \cdot f_i \cdot x_k^{m-k} \Rightarrow \begin{cases} a_0 = \dots \\ a_1 = \dots \\ \dots \\ a_m = \dots \end{cases} \Rightarrow p(x) = \sum_{i=0}^m a_i x^{m-i}.$$

O błędzie aproksymacji decyduje wartość funkcjonału dla policzonych współczynników. Informuje o maksymalnej odchyłce dla danego zestawu węzłów.



### Przykład 9

Dla danych z poprzedniego przykładu znaleźć aproksymację liniową. Rozpatrzeć dwa przypadki: metodę zwykłą i ważoną przypisując każdemu z węzłów jego numer jako wagę.

|       |   |   |   |
|-------|---|---|---|
| $i$   | 1 | 2 | 3 |
| $x_i$ | 0 | 1 | 2 |
| $f_i$ | 0 | 1 | 4 |

Przyjmujemy funkcję liniową:  $p(x) = a \cdot x + b$ .

*I. Metoda zwykła*

Układamy funkcjonal:

$$B(a,b) = \sum_{i=1}^3 (a \cdot x_i + b - f_i)^2 = (a \cdot 0 + b - 0)^2 + (a \cdot 1 + b - 1)^2 + (a \cdot 2 + b - 4)^2.$$

Różniczkujemy po zmiennych  $a$  i  $b$ :

$$\begin{cases} \frac{\partial}{\partial a} B(a,b) = 2 \cdot (a+b-1) + 2 \cdot 2 \cdot (2a+b-4) = 0 \\ \frac{\partial}{\partial b} B(a,b) = 2 \cdot b + 2 \cdot (a+b-1) + 2 \cdot (2a+b-4) = 0 \end{cases}$$
$$\begin{cases} 5a + 3b = 9 \\ 3a + 3b = 5 \end{cases} \Rightarrow \begin{cases} a = 2 \\ b = -\frac{1}{3} \end{cases} \Rightarrow p(x) = 2x - \frac{1}{3}.$$

Wyniki zestawiono w tabelce.

| $i$ | $x_i$ | $f_i$ | $p_i = p(x_i)$ | $\varepsilon_i = p_i - f_i$ | $\varepsilon_i^2$ |
|-----|-------|-------|----------------|-----------------------------|-------------------|
| 1   | 0     | 0     | -0.333333      | -0.333333                   | 0.111111          |
| 2   | 1     | 1     | 1.666667       | 0.666667                    | 0.444444          |
| 3   | 2     | 4     | 3.666667       | -0.333333                   | 0.111111          |

$$\text{Błąd maksymalny } B_{\max} = \sum_{i=1}^3 \varepsilon_i^2 = 0.666667.$$

*II. Metoda ważona*

Wagi:  $w_1 = 1$ ,  $w_2 = 2$ ,  $w_3 = 3$ .

$$B(a,b) = \sum_{i=1}^3 w_i (a \cdot x_i + b - f_i)^2 = 1 \cdot (a \cdot 0 + b - 0)^2 + 2 \cdot (a \cdot 1 + b - 1)^2 + 3 \cdot (a \cdot 2 + b - 4)^2.$$

Różniczkujemy po zmiennych  $a$  i  $b$ :

$$\begin{cases} \frac{\partial}{\partial a} B(a,b) = 2 \cdot 2 \cdot (a+b-1) + 2 \cdot 2 \cdot 3 \cdot (2a+b-4) = 0 \\ \frac{\partial}{\partial b} B(a,b) = 1 \cdot 2 \cdot b + 2 \cdot 2 \cdot (a+b-1) + 2 \cdot 3 \cdot (2a+b-4) = 0 \end{cases}$$
$$\begin{cases} 14a + 8b = 26 \\ 8a + 6b = 14 \end{cases} \Rightarrow \begin{cases} a = 2.2 \\ b = -0.6 \end{cases} \Rightarrow p(x) = 2.2x - 0.6.$$

| $i$ | $x_i$ | $f_i$ | $p_i = p(x_i)$ | $\varepsilon_i = p_i - f_i$ | $\varepsilon_i^2$ |
|-----|-------|-------|----------------|-----------------------------|-------------------|
| 1   | 0     | 0     | -0.60          | -0.60                       | 0.360             |
| 2   | 1     | 1     | 1.60           | 0.60                        | 0.360             |
| 3   | 2     | 4     | 3.80           | 0.20                        | 0.04              |

$$B_{\max} = \sum_{i=1}^3 w_i \varepsilon_i^2 = 1 \cdot 0.36 + 2 \cdot 0.36 + 3 \cdot 0.04 = 1.20.$$

Widać poprawę tam gdzie waga była największa: dla węzła  $x_3 = 2$ .

### Przykład 10

Dla danych z poprzedniego zadania zastosować aproksymację kwadratową.

|       |   |   |   |
|-------|---|---|---|
| $i$   | 1 | 2 | 3 |
| $x_i$ | 0 | 1 | 2 |
| $f_i$ | 0 | 1 | 4 |

Funkcja aproksymująca:  $p(x) = a \cdot x^2 + b \cdot x + c$ .

$$B(a, b, c) = \sum_{i=1}^3 (a \cdot x_i^2 + b \cdot x_i + c - f_i)^2 = (c-0)^2 + (a+b+c-1)^2 + (4a+2b+c-4)^2.$$

$$\begin{cases} \frac{\partial B}{\partial a} = (a+b+c-1) + 4 \cdot (4a+2b+c-4) = 0 \\ \frac{\partial B}{\partial b} = (a+b+c-1) + 2 \cdot (4a+2b+c-4) = 0 \\ \frac{\partial B}{\partial c} = c + (a+b+c-1) + (4a+2b+c-4) = 0 \end{cases}$$

$$\begin{cases} \frac{\partial B}{\partial a} = (a+b+c-1) + 4 \cdot (4a+2b+c-4) = 0 \\ \frac{\partial B}{\partial b} = (a+b+c-1) + 2 \cdot (4a+2b+c-4) = 0 \\ \frac{\partial B}{\partial c} = c + (a+b+c-1) + (4a+2b+c-4) = 0 \end{cases}$$

$$\begin{cases} \frac{\partial B}{\partial a} = (a+b+c-1) + 4 \cdot (4a+2b+c-4) = 0 \\ \frac{\partial B}{\partial b} = (a+b+c-1) + 2 \cdot (4a+2b+c-4) = 0 \\ \frac{\partial B}{\partial c} = c + (a+b+c-1) + (4a+2b+c-4) = 0 \end{cases}$$

$$\begin{cases} 17a+9b+5c=17 \\ 9a+5b+3c=9 \\ 5a+3b+3c=5 \end{cases} \Rightarrow \begin{cases} a=1 \\ b=0 \\ c=0 \end{cases} \Rightarrow p(x) = x^2.$$

Jest to przypadek szczególny: budowanie aproksymacji kwadratowej na trzech węzłach daje w rezultacie interpolację: otrzymaliśmy wyjściową parabolę. Nie ma sensu stosować metody ważonej.

## II. NUMERYCZNE RÓŻNICZKOWANIE FUNKCJI

Wynikiem numerycznego różniczkowania nie jest analityczny wzór na pochodną, ale jej wartość w wybranym węźle zwanym węzłem centralnym. Zadanie sprowadza się do wyznaczenia tzw. wzoru różnicowego, czyli wzoru liczącego określoną pochodną w węźle centralnym na podstawie wartości dyskretnych funkcji w innych węzłach, np.:

Dane są wartości funkcji  $w_0, w_1, w_2$  w równych odstępach  $h$ . Należy zbudować wzory różnicowe na pierwszą i drugą pochodną w węźle centralnym  $w_1$ .

Najbardziej oczywistym sposobem, ale najbardziej prymitywnym jest dokonanie interpolacji (ogólnie: aproksymacji) w podanych punktach, a następnie na podstawie otrzymanego wzoru interpolacyjnego (np. wielomianowego) określić wzór na pochodną i w końcu policzyć

wartość pochodnej w żądanym węźle. Jest to dość złożony proces, gdyż wymaga przejścia z wartości dyskretnych funkcji do wzoru ciągłego a następnie ponowne przejście na wartości dyskretne. Można tego uniknąć, skoro i tak wychodząc od wartości w punktach, szukamy również wartości dyskretnej. Najlepszą metodą do tego celu jest *metoda współczynników nieoznaczonych* bazująca na rozwijaniu wszystkich wartości węzłowych w *szereg Taylora*.

Przyjmujemy lokalny układ współrzędnych w węźle centralnym  $w_1$ . Teraz odległości od pozostałych węzłów wynoszą odpowiednio  $-h$  oraz  $h$ . Rozwijamy każdą z wartości w *szereg Taylora* wokół węzła centralnego zachowując tyle wyrazów ile niewiadomych będzie w końcowym układzie równań. Liczba niewiadomych jest równa ilości informacji, na jakich budujemy wzór różnicowy (w tym przypadku zachowamy trzy wyrazy). Wzoru różnicowego szukamy jako kombinacji liniowej wartości węzłowych i nieznanymi (nieoznaczonymi – stąd nazwa metody) współczynników liczbowych.

- Dla pierwszej pochodnej:  $w'(x) \approx \sum_{i=0}^2 a_i w_i = a_0 w_0 + a_1 w_1 + a_2 w_2$ ,
- Dla drugiej pochodnej:  $w''(x) \approx \sum_{i=0}^2 b_i w_i = b_0 w_0 + b_1 w_1 + b_2 w_2$ .

Dla obydwu pochodnych wypisujemy rozwinięcia w poszczególnych węzłach:

$$\begin{cases} w_0 = w_1 - h \cdot w_1 + \frac{1}{2} h^2 w_1 + \dots \\ w_1 \equiv w_1 \\ w_2 = w_1 + h \cdot w_1 + \frac{1}{2} h^2 w_1 + \dots \end{cases}$$

Rozwinięcia mnożymy przez współczynniki stojące we wzorach różnicowych. Następnie sumujemy je ze sobą, porządkując wyrazy stojące przy odpowiednich pochodnych. Układ równań powstaje przez porównanie współczynników stojących przy odpowiednich pochodnych: ściślej pochodnej i wzoru różnicowego.

- Dla pierwszej pochodnej:

$$\overbrace{a_0 w_0 + a_1 w_1 + a_2 w_2}^{\approx w_1'} = w_1(a_0 + a_1 + a_2) + w_1'(-h \cdot a_0 + h \cdot a_2) + w_1''\left(\frac{1}{2} h^2 a_0 + \frac{1}{2} h^2 a_2\right)$$

$$w_1' \approx w_1(a_0 + a_1 + a_2) + w_1'(-h \cdot a_0 + h \cdot a_2) + w_1''\left(\frac{1}{2} h^2 a_0 + \frac{1}{2} h^2 a_2\right) \quad ,$$

- Dla drugiej pochodnej:

$$\overbrace{b_0 w_0 + b_1 w_1 + b_2 w_2}^{\approx w_1''} = w_1(b_0 + b_1 + b_2) + w_1'(-h \cdot b_0 + h \cdot b_2) + w_1''\left(\frac{1}{2} h^2 b_0 + \frac{1}{2} h^2 b_2\right)$$

$$w_1'' \approx w_1(b_0 + b_1 + b_2) + w_1'(-h \cdot b_0 + h \cdot b_2) + w_1''\left(\frac{1}{2} h^2 b_0 + \frac{1}{2} h^2 b_2\right) .$$

Dla obydwu przypadków powstaje układ równań z tą samą macierzą współczynników, ale z innymi prawymi stronami:

$$\begin{bmatrix} 1 & 1 & 1 \\ -h & 0 & h \\ \frac{1}{2}h^2 & 0 & \frac{1}{2}h^2 \end{bmatrix} \cdot \begin{bmatrix} a_0 & b_0 \\ a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} a_0 & b_0 \\ a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2h} & \frac{1}{h^2} \\ 0 & -\frac{2}{h^2} \\ \frac{1}{2h} & \frac{1}{h^2} \end{bmatrix}$$

Stąd:

$$w_1' \approx \frac{1}{2h}(w_2 - w_0)$$

$$w_1'' \approx \frac{1}{h^2}(w_0 - 2 \cdot w_1 + w_2)$$

Obliczenie dokładności takich wzorów polega na przywróceniu pierwszego z odrzuconych niezerowych wyrazów w każdym z rozwinięć, przemnożeniu przez odpowiedni współczynnik a następnie zsumowaniu.

- Dla pierwszej pochodnej (wyrazy trzeciego rzędu):

$$\varepsilon(h) = a_0 \cdot \left(-\frac{1}{6}h^3 w_1'''\right) + a_2 \cdot \frac{1}{6}h^3 w_1''' = \frac{1}{6}h^3 w_1''' (a_2 - a_0) = \frac{1}{6}h^3 w_1''' \left(\frac{1}{2h} + \frac{1}{2h}\right) = \frac{1}{6}h^2 w_1'''$$

- Dla drugiej pochodnej (wyrazy czwartego rzędu):

$$\varepsilon(h) = b_0 \cdot \frac{1}{24}h^4 w_1^{IV} + b_2 \cdot \frac{1}{24}h^4 w_1^{IV} = \frac{1}{24}h^4 w_1^{IV} (b_2 - b_0) = \frac{1}{24}h^4 w_1^{IV} \left(\frac{1}{h^2} + \frac{1}{h^2}\right) = \frac{1}{12}h^2 w_1^{IV}$$

Sprawdzenie powyższych wzorów może odbyć się dla wielomianów, dla których wzory dają jeszcze wynik ścisły. W tym przypadku będą to wielomiany rzędu drugiego.

Przyjmijmy funkcję  $f(x) = x^2$  oraz następujące węzły:

|                |   |   |   |
|----------------|---|---|---|
| $i$            | 1 | 2 | 3 |
| $x_i$          | 0 | 1 | 2 |
| $f_i = f(x_i)$ | 0 | 1 | 4 |

Węzły są równooddalone, ich odległość wynosi  $h = 1$ .

- Ścisłe wartości analityczne pochodnych:

$$f(x) = x^2 \rightarrow f'(x) = 2x \rightarrow f''(x) = 2, \text{ stąd: } f_1' = 2, \quad f_1'' = 2.$$

- Wartości numeryczne pochodnych (ze wzorów różnicowych):

$$w_1' \approx \frac{4-0}{2 \cdot 1} = 2, \quad w_1'' \approx \frac{0-2 \cdot 1+4}{1^2} = 2.$$

Wniosek:  $w_1' = f_1' = 2, \quad w_1'' = f_1'' = 2.$

Wyprowadzone wyżej wzory należą do tzw. centralnych wzorów różnicowych. Oprócz nich istnieją też tzw. poboczne wzory różnicowe, o wiele mniej dokładne, np. dla pierwszej pochodnej:

- tzw. iloraz „wprzód”:  $w_1' \approx \frac{w_1 - w_0}{h},$
- tzw. iloraz „wstecz”:  $w_1' \approx \frac{w_2 - w_1}{h}.$

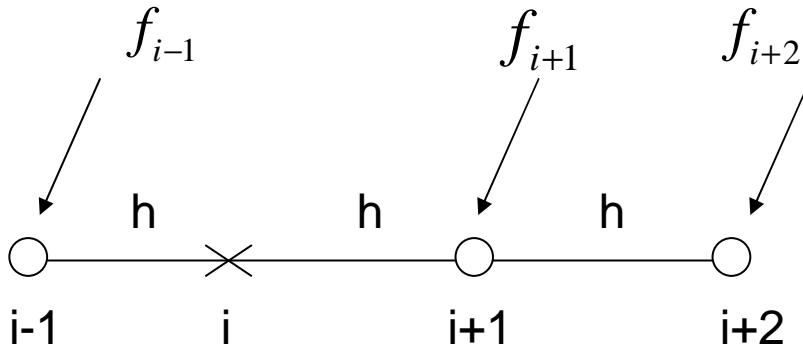
Dają one wyniki ścisłe w ramach pierwszego rzędu wielomianowej aproksymacji. Dla wyżej testowanej funkcji nie dałyby wyników ścisłych, tylko przybliżone.

### Przykład 11

Znaleźć przedstawienie operatora drugiej pochodnej w postaci:

$$f_i'' = \alpha_i f_{i-1} + \beta_i f_{i+1} + \gamma_i f_{i+2}.$$

Zakładając, iż odstęp między węzłami są stałe i wynoszą  $h$ , dana konfiguracja węzłów wygląda następująco:



W punkcie  $(i)$  nie jest dana żadna informacja (a więc nie jest on węzłem), a mimo wszystko poszukuje się w nim wartości numerycznej na drugą pochodną funkcji.

Rozwijamy każdą wartość funkcyjną względem punktu  $(i)$  w *szereg Taylora* zachowując tyle wyrazów rozwinięcia, ile nieznanych współczynników należy wyznaczyć (3). W takim przypadku uzyskamy interpolację, czyli przeprowadzimy lokalną krzywą paraboliczną przez wszystkie wartości węzłowe.

$$f_{i-1} = f_i - h f_i' + \frac{1}{2} h^2 f_i''$$

$$f_{i+1} = f_i + h f_i' + \frac{1}{2} h^2 f_i''$$

$$f_{i+2} = f_i + 2h f_i' + \frac{1}{2} (2h)^2 f_i''$$

Dalej mnożymy każde z rozwinięć przez niewiadomy współczynnik stojący przy rozwiniętej wartości funkcyjnej we wzorze różnicowym.

$$f_{i-1} = f_i - h f_i' + \frac{1}{2} h^2 f_i'' \quad / \times \alpha_i$$

$$f_{i+1} = f_i + h f_i' + \frac{1}{2} h^2 f_i'' \quad / \times \beta_i$$

$$f_{i+2} = f_i + 2h f_i' + \frac{1}{2} (2h)^2 f_i'' \quad / \times \gamma_i$$

Teraz dodajemy stronami powyższe rozwinięcia, pamiętając o mnożeniu ich przez współczynniki  $\alpha_i$ ,  $\beta_i$ ,  $\gamma_i$ .

$$\underbrace{\alpha_i f_{i-1} + \beta_i f_{i+1} + \gamma_i f_{i+2}}_{\approx f_i''} = f_i(\alpha_i + \beta_i + \gamma_i) + f_i'(-h \cdot \alpha_i + h \cdot \beta_i + 2h \cdot \gamma_i) + f_i''\left(\frac{1}{2}h^2 \alpha_i + \frac{1}{2}h^2 \beta_i + 2h^2 \gamma_i\right)$$

Ponieważ wyrażenie lewej strony to wyjściowy wzór różnicowy na drugą pochodną, można zastąpić je wartością drugiej pochodnej.

$$f_i'' \approx f_i(\alpha_i + \beta_i + \gamma_i) + f_i'(-h \cdot \alpha_i + h \cdot \beta_i + 2h \cdot \gamma_i) + f_i''\left(\frac{1}{2}h^2 \alpha_i + \frac{1}{2}h^2 \beta_i + 2h^2 \gamma_i\right).$$

Aby zachodziła równość między lewą i prawą stroną, współczynniki przy funkcji i jej kolejnych pochodnych muszą być sobie równe.

$$\begin{cases} \alpha_i + \beta_i + \gamma_i = 0 \\ -h \cdot \alpha_i + h \cdot \beta_i + 2h \cdot \gamma_i = 0 \\ \frac{1}{2}h^2 \alpha_i + \frac{1}{2}h^2 \beta_i + 2h^2 \gamma_i = 1 \end{cases}$$

W ten sposób powstaje układ równań na współczynniki  $\alpha_i, \beta_i, \gamma_i$ . Po jego rozwiązaniu otrzymujemy

$$\begin{cases} \alpha_i = \frac{1}{3h^2} \\ \beta_i = -\frac{1}{h^2} \\ \gamma_i = \frac{2}{3h^2} \end{cases}$$

Końcowy wzór różnicowy

$$f_i'' = \frac{1}{3h^2}(f_{i-1} - 3f_{i+1} + 2f_{i+2})$$

Dokładność wzoru można oszacować zbierając pierwsze odrzucone wyrazy rozwinięcia.

$$\varepsilon(h) = -\alpha_i \frac{1}{6}h^3 f_i''' + \beta_i \frac{1}{6}h^3 f_i''' + \gamma_i \frac{1}{6}(2h)^3 f_i''' = \frac{1}{18}h \cdot f_i'''(-1-3+8 \cdot 2) = \frac{2}{3}h \cdot f_i'''.$$

Wzór jest ścisły dla wielomianów rzędu, co najwyżej drugiego. Sprawdzenie będzie polegać na policzeniu pochodnej numerycznej dla funkcji testowej oraz porównanie ze ścisłą wartością. Przyjęto rozstaw węzłów:  $x_{i-1} = 0$ ,  $x_{i+1} = 2$ ,  $x_{i+2} = 3$ . Rozstaw  $h = 1$ .

- Funkcja testowa:  $f(x) = x^2$ 
  - Wartości węzłowe:  $f_{i-1} = 0^2 = 0$ ,  $f_{i+1} = 2^2 = 4$ ,  $f_{i+2} = 3^2 = 9$ .

- o Wartość numeryczna drugiej pochodnej:  $f_i'' \approx \frac{1}{3 \cdot 1^2} (0 - 3 \cdot 4 + 2 \cdot 9) = 2$ .
- o Wartość ścisła drugiej pochodnej:  $f(x) = x^2 \Rightarrow f''(x) = 2 \Rightarrow f_i'' = 2$ .

Numeryczna wartość jest wartością ścisłą. Nie jest to przypadek, gdyż faktycznie dla funkcji parabolicznej  $f'''(x) = 0$ , a więc błąd wyniku  $\varepsilon \equiv 0$ .

- Funkcja testowa:  $f(x) = x^3$ 
  - o Wartości węzłowe:  $f_{i-1} = 0^2 = 0$ ,  $f_{i+1} = 2^3 = 8$ ,  $f_{i+2} = 3^3 = 27$ .
  - o Wartość numeryczna drugiej pochodnej:  $f_i'' \approx \frac{1}{3 \cdot 1^2} (0 - 3 \cdot 8 + 2 \cdot 27) = 10$ .
  - o Wartość ścisła drugiej pochodnej:  $f(x) = x^3 \Rightarrow f''(x) = 6x \Rightarrow f_i'' = 6$ .

Numeryczna wartość nie jest wartością ścisłą. Błąd wyniku  $\varepsilon = \frac{2}{3} \cdot 1 \cdot (f_i''' = 6) = 4$  jest w tym przypadku różnicą bezwzględną między wartością numeryczną i ścisłą.

*Metoda współczynników nieoznaczonych*, oparta na rozwinięciu w szereg Taylora ma wiele zalet. Jedną z nich jest możliwość łatwego oszacowania błędu wzoru różnicowego. Metoda pozwala również na budowanie operatorów różniczkowych dowolnej postaci, np.

$$\mathcal{L} = a \cdot \frac{d^2}{dx^2} + b \cdot \frac{d}{dx} + c, \quad a, b, c \in \mathfrak{R}$$

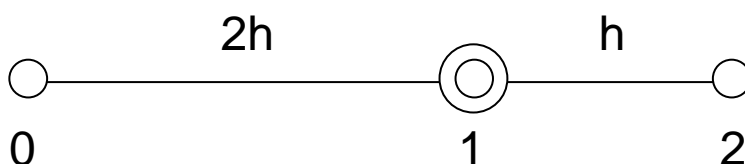
poprzez przybliżanie ich wartości w węźle ( $i$ ) wzorem interpolacyjnym opartym na trzech węzłach:

$$\mathcal{L}u_i \approx Lu_i = \alpha_{i-1}u_{i-1} + \beta_i u_i + \gamma_{i+1}u_{i+1}.$$

Inną cechą tak budowanych wzorów różnicowych jest to, iż mogą one bazować nie tylko na wartościach samej funkcji w węzłach, ale także ich kolejnych pochodnych (byle nie wyższych niż najwyższy rząd pochodnej występującej w operatorze różniczkowym). Wartości pochodnych funkcji w węzłach (lub nawet wartości całych operatorów różniczkowych) nazywane są uogólnionymi stopniami swobody.

### Przykład 12

Znaleźć numeryczną wartość operatora różniczkowego  $\mathcal{L}u_1 = -2u_1 + 4u_1' - 3u_1''$  za pomocą następującego wzoru różnicowego  $Lu_1 = \alpha u_0 + \beta u_1 + \gamma u_2'$  dla zadanej konfiguracji węzłów jak na rys. Wzór sprawdzić dla funkcji testowych  $x^2$ ,  $x^3$ . Określić dokładność takiego wzoru.



Rozwinięcie wartości węzłowych w szereg Taylora i przemnożenie rozwinięć przez odpowiedni współczynnik:

$$u_0 = u_1 - 2h \cdot u_1' + \frac{1}{2}(-2h)^2 u_1'' + \dots \quad / \times \alpha$$

$$u_1 = u_1 \quad / \times \beta$$

$$u_2' = u_1' + h \cdot u_1'' + \dots \quad / \times \gamma$$

Ostatnie równanie to rozwinięcie wartości pierwszej pochodnej. Znajduje się je poprzez rozwinięcie samej wartości funkcji:  $u_2 = u_1 + h \cdot u_1' + \frac{1}{2}h^2 \cdot u_1'' + \dots$  a następnie różniczkuje się je stronami (tak, aby otrzymać po stronie lewej pierwszą pochodną) opuszczając wyrazy rzędu wyższego niż drugi.

Dodanie rozwinięć stronami:

$$\alpha \cdot u_0 + \beta \cdot u_1 + \gamma \cdot u_2' = u_1(\alpha + \beta) + u_1'(-2h \cdot \alpha + \gamma) + u_1''(2h^2 \cdot \alpha + h \cdot \gamma)$$

oraz zastąpienie (w przybliżeniu) wzoru różnicowego (lewa strona) wartością operatora różniczkowego:

$$\underbrace{\alpha \cdot u_0 + \beta \cdot u_1 + \gamma \cdot u_2'}_{\mathcal{L}u_1 = -2u_1 + 4u_1' - 3u_1''} = u_1(\alpha + \beta) + u_1'(-2h \cdot \alpha + \gamma) + u_1''(2h^2 \cdot \alpha + h \cdot \gamma) \Rightarrow$$

$$\Rightarrow -2u_1 + 4u_1' - 3u_1'' \cong u_1(\alpha + \beta) + u_1'(-2h \cdot \alpha + \gamma) + u_1''(2h^2 \cdot \alpha + h \cdot \gamma) \quad \sim$$

prowadzi, po przyrównaniu współczynników przy funkcji i odpowiednich pochodnych do końcowego układu równań algebraicznych:

$$\begin{cases} \alpha + \beta = -2 \\ -2h \cdot \alpha + \gamma = 4 \\ 2h^2 \cdot \alpha + h \cdot \gamma = -3 \end{cases} \Rightarrow \begin{cases} \alpha = \frac{-3-4h}{4h^2} \\ \beta = \frac{-8h^2+3+4h}{4h^2} \\ \gamma = \frac{-3+4h}{2h} \end{cases}$$

Końcowa postać wzoru różnicowego:

$$\mathcal{L}u_1 = -2u_1 + 4u_1' - 3u_1'' \approx Lu_1 = \frac{-3-4h}{4h^2}u_0 + \frac{-8h^2+3+4h}{4h^2}u_1 + \frac{-3+4h}{2h}u_2'$$

Dokładność wzoru:

$$\varepsilon(h) = \frac{1}{6}(-2h)^3 \cdot u_1''' \cdot \alpha + \frac{1}{2}h^2 u_1''' \cdot \gamma = \frac{h}{12} u_1''' (3 + 28h).$$



Sprawdzenie dla jednomianów:

(przyjęto:  $x_0 = 1, x_1 = 3, x_2 = 4 \rightarrow h = 1$ )

- dla  $u(x) = x^2$  ( $u'(x) = 2x, u''(x) = 2, u'''(x) = 0$ )

Wartość ścisła:  $u_1 = 9, u_1' = 6, u_1'' = 2 \rightarrow Lu_1 = -2 \cdot 9 + 4 \cdot 6 - 3 \cdot 2 = 0$

Wartość numeryczna:

$$u_0 = 1, u_1 = 9, u_2' = 8 \rightarrow Lu_1 = \frac{-3-4}{4} \cdot 1 + \frac{3+4-8}{4} \cdot 9 + \frac{4-3}{2} \cdot 8 = 0.$$

$$\text{Błąd wyniku: } \varepsilon = \frac{1}{12} \cdot 0 \cdot (3+28) = 0.$$

- dla  $u(x) = x^3$  ( $u'(x) = 3x^2, u''(x) = 6x, u'''(x) = 6$ )

Wartość ścisła:  $u_1 = 27, u_1' = 27, u_1'' = 18 \rightarrow Lu_1 = -2 \cdot 27 + 4 \cdot 27 - 3 \cdot 18 = 0$

Wartość numeryczna:

$$u_0 = 1, u_1 = 27, u_2' = 48 \rightarrow Lu_1 = \frac{-3-4}{4} \cdot 1 + \frac{3+4-8}{4} \cdot 27 + \frac{4-3}{2} \cdot 48 = 15.5.$$

$$\text{Błąd wyniku: } \varepsilon = \frac{1}{12} \cdot 6 \cdot (3+28) = 15.5$$

Operatory różnicowe można też budować metodami aproksymacji funkcji dyskretnej, np. *najlepszej aproksymacji*. Wyniki mogą się różnić od wyników pochodzących z interpolacji (zwłaszcza, jeżeli w tzw. *gwiazdzie*, – czyli konfiguracji węzłów – jest nadmiar węzłów w stosunku do niezbędnej liczby informacji potrzebnej do zbudowania odpowiedniego operatora). Technika powszechnie używaną w metodach dyskretnych do rozwiązywania równań różniczkowych brzegowych (zwłaszcza w beziatkowej metodzie różnic skończonych *BMRS*) służącą do generacji kompletów wzorów różnicowych jest technika aproksymacji *MWLS* (ang. *Moving Weighted Least Squares*) – *technika najmniejszych ważonych kroczących kwadratów*.

### III. NUMERYCZNE CAŁKOWANIE FUNKCJI

Tak jak wynikiem numerycznego różniczkowania była wartość dowolnej pochodnej w konkretnym węźle (lub w dowolnym punkcie), tak wynikiem całkowania numerycznego nie jest funkcja analityczna, a jedynie wartość liczbowa całki. Stąd oczywisty wniosek, iż numeryka pozwala na obliczanie przede wszystkim całek oznaczonych (liczb) w dowolnym przedziale  $(a, b)$ . Wzory całkowania numerycznego, zwane kwadraturami, pozwalają na obliczenie (w przybliżeniu) wartości całki:

$$I = \int_a^b f(x) dx$$

Na początek zakładamy, iż granice całkowania są skończone, a funkcja podcałkowa nie ma w przedziale  $(a, b)$  osobliwości (jest ciągła) – tzw. całka właściwa. Wzory te dzielimy na dwie główne grupy:

- *kwadratury Newtona – Cotesa*, polegające na zastąpieniu funkcji podcałkowej wielomianami coraz to wyższych rzędów w przedziale podzielonym na odcinki równej długości,
- *kwadratury Gaussa*, polegające na zastąpieniu funkcji podcałkowej wielomianami ortogonalnymi w taki sposób, aby wzór był ścisły dla wielomianu możliwie najwyższego rzędu.

Po zastąpieniu funkcji podcałkowej wielomianem, łatwym do scałkowania, otrzymujemy wzór całkowania, bazujący na wartościach funkcji w przedziale  $(a, b)$ .

### Kwadratury Newtona – Cotesa

Funkcja podcałkowa jest aproksymowana przez wielomiany coraz to wyższych rzędów w przedziale  $(a, b)$  podzielonym na odcinki o równej długości (podział równomierny).

Założenie:  $x_{i+1} - x_i = x_i - x_{i-1} = h = \text{const}$ .

Przedział  $(a, b)$  dzielimy na podprzedziały o równej długości punktami  $x_i$ ,  $i = 0, 1, 2, \dots, n$ ,  
 $a = x_0 + ph$ ,  $b = x_0 + qh$ ,  $p \geq 0$ ,  $q \leq n$ .

Budujemy wielomiany *Lagrange'a*:

$$I = \int_a^b f(x) dx \approx \int_a^b \sum_{j=0}^n L_j^{(n)}(x) f_j dx = I_{(n+1)} = \sum_{j=0}^n \int_a^b L_j^{(n)}(x) f_j dx.$$

Wprowadzamy indeks  $s$  tak, że  $x = x_0 + sh$ .

$$I_{(n+1)} = \sum_{j=0}^n \int_p^q \prod_{\substack{k=0 \\ k \neq j}}^n \frac{(x_0 + sh) - (x_0 + kh)}{(x_0 + jh) - (x_0 + kh)} f_j h ds = h \sum_{j=0}^n \int_p^q \prod_{\substack{k=0 \\ k \neq j}}^n \frac{s-k}{j-k} f_j ds = h \sum_{j=0}^n f_j \prod_{\substack{k=0 \\ k \neq j}}^n \frac{1}{j-k} \int_p^q \prod_{\substack{k=0 \\ k \neq j}}^n \frac{s-k}{s-j} ds$$

Wprowadzamy współczynniki liczbowe  $\alpha_j$

$$\alpha_j = \frac{h}{\prod_{\substack{k=0 \\ k \neq j}}^n \frac{1}{j-k} \int_p^q \prod_{\substack{k=0 \\ k \neq j}}^n \frac{s-k}{s-j} ds} = h \frac{(-1)^{n-j}}{n!} \binom{n}{j} \int_p^q \prod_{k=0, k \neq j}^n \frac{s-k}{s-j} ds, \quad j = 0, 1, \dots, n.$$

Ostateczna postać kwadratury

$$I = \int_a^b f(x) dx \approx \sum_{j=0}^n \alpha_j f_j + E, \quad E = I - I_{(n+1)},$$

gdzie błąd  $E$  wyniku wyraża się wzorem:

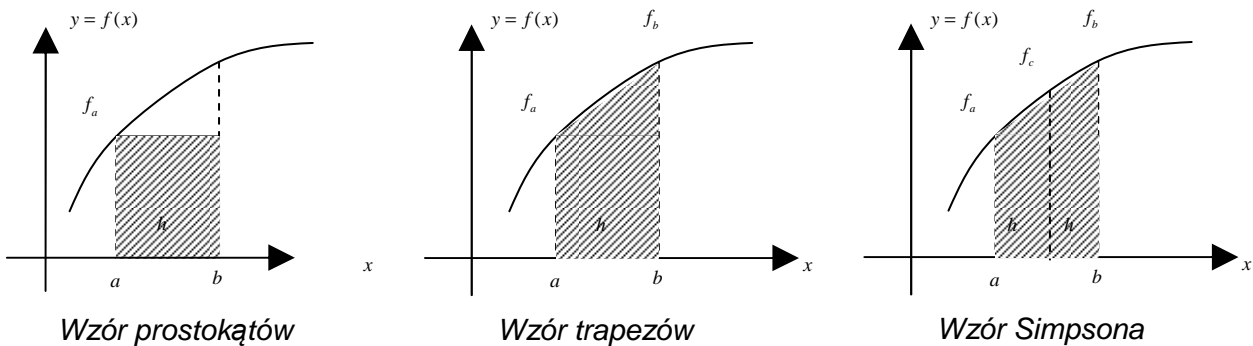
$$E = \begin{cases} \frac{2h^{n+2}}{(n+1)!} f^{(n+1)}(\xi) \int_0^2 \prod_{k=\frac{-2r+1}{2}}^{\frac{2r-1}{2}} (s-k) ds, & \text{dla } n \text{ nieparzystych } (\xi \in \langle a, b \rangle) \\ \frac{2h^{n+3}}{(n+2)!} f^{(n+2)}(\eta) \int_0^m \prod_{k=0}^{k=r} (s^2 - k^2) ds, & \text{dla } n \text{ parzystych } (\eta \in \langle a, b \rangle) \end{cases}$$

Tabela współczynników  $\frac{\alpha_i}{h}$  wzorów Newtona – Cotesa

| $n$ | $j=0$         | $j=1$         | $j=2$         | $j=3$         | błąd                            | nazwa wzoru   |
|-----|---------------|---------------|---------------|---------------|---------------------------------|---------------|
| 1   | $\frac{1}{2}$ | $\frac{1}{2}$ | —             | —             | $-\frac{1}{12} h^3 f''(\xi)$    | wzór trapezów |
| 2   | $\frac{1}{3}$ | $\frac{4}{3}$ | $\frac{1}{3}$ | —             | $-\frac{1}{90} h^5 f^{IV}(\xi)$ | wzór Simpsona |
| 3   | $\frac{3}{8}$ | $\frac{9}{8}$ | $\frac{9}{8}$ | $\frac{3}{8}$ | $-\frac{3}{80} h^5 f^{IV}(\xi)$ | —             |

Szczególnie korzystny w zastosowaniach jest wzór *Simpsona* ze względu na podwyższoną dokładność.

Trzy pierwsze kwadratury Newtona – Cotesa to najpowszechniej używane wzory całkowania numerycznego.



- Wzór prostokątów

$$I = \int_a^b f(x) dx \approx \int_a^b f(a) dx = f_a x \Big|_a^b = f_a(b-a) = f_a h.$$

- Wzór trapezów

$$I = \int_a^b f(x) dx \approx \int_a^b f(a) \frac{h-x}{x} + f(b) \frac{x}{h} dx = \frac{h}{2} (f_a + f_b).$$

- Wzór Simpsona

$$I = \int_a^b f(x) dx \approx \int_a^b f(a)L_0^{(2)}(x) + f\left(\frac{a+b}{2}\right)L_1^{(2)}(x) + f(b)L_2^{(2)}(x) dx = \frac{h}{3}(f_a + 4 \cdot f_c + f_b)$$

$$c = \frac{a+b}{2}, \quad h = \frac{b-a}{2}$$

W praktyce nie używa się już wzorów wyższego rzędu, natomiast stosuje się powyższe trzy wzory niskich rzędów (zwłaszcza *wzór Simpsona*) w podprzedziałach wynikających z podziału wyjściowego przedziału  $(a, b)$ . Powstają w ten sposób tzw. *wzory złożone całkowania*. Ilość podziałów nie jest z góry założona, należy ją dobrać iteracyjnie ze względu na żadaną dokładność wyników  $\varepsilon$ .

### Przykład 12

Podaną całkę  $I = \int_0^1 \sqrt{1+x} dx$  obliczyć numerycznie stosując *wzory Newtona – Cotesa* proste i złożone (dwa podziały). Za każdym razem porównać otrzymany wynik numeryczny z rozwiązaniem analitycznym.

Wynik analityczny

$$I = \int_0^1 \sqrt{1+x} dx = \frac{2}{3} (1+x)^{\frac{3}{2}} \Big|_0^1 = \frac{2}{3} \cdot 2^{\frac{3}{2}} - \frac{2}{3} = 1.218951.$$

Proste wzory całkowania (1 przedział)

$$I_{p_1} = f(a) \cdot h = \sqrt{1+0} \cdot 1 = 1, \quad \varepsilon = \left| \frac{I_{p_1} - I}{I} \right| \cdot 100\% = \left| \frac{1 - 1.218951}{1.218951} \right| \cdot 100\% = 18\%$$

$$I_{t_1} = \frac{f(a) + f(b)}{2} \cdot h = \frac{1}{2} \cdot 1 \cdot (\sqrt{1+0} + \sqrt{1+1}) = \frac{1 + \sqrt{2}}{2} = 1.207107,$$

$$\varepsilon = \left| \frac{I_{t_1} - I}{I} \right| \cdot 100\% = \left| \frac{1.207107 - 1.218951}{1.218951} \right| \cdot 100\% = 1\%$$

$$I_{s_1} = \frac{b-a}{6} (f(a) + 4f(c) + f(b)) = \frac{1}{6} \cdot (\sqrt{1+0} + 4\sqrt{1+\frac{1}{2}} + \sqrt{1+1}) = \frac{1 + 4\sqrt{\frac{3}{2}} + \sqrt{2}}{6} = 1.218866,$$

$$\varepsilon = \left| \frac{I_{s_1} - I}{I} \right| \cdot 100\% = \left| \frac{1.218866 - 1.218951}{1.218951} \right| \cdot 100\% = 0.007\%$$

Złożone wzory całkowania (2 równe przedziały,  $h = \frac{1}{2}$ )

$$I_{p_2} = f(0) \cdot \frac{1}{2} + f(1) \cdot \frac{1}{2} = \sqrt{1+0} \cdot \frac{1}{2} + \sqrt{1+1} \cdot \frac{1}{2} = \frac{1}{2} + \frac{1}{2} \sqrt{2} = 1.112372,$$

$$\varepsilon = \left| \frac{I_{p_2} - I}{I} \right| \cdot 100\% = \left| \frac{1.112372 - 1.218951}{1.218951} \right| \cdot 100\% = 8.7\%$$

$$\begin{aligned}
I_{t_2} &= \left[ f(0) + f\left(\frac{1}{2}\right) \right] \cdot \frac{1}{2} \cdot \frac{1}{2} + \left[ f\left(\frac{1}{2}\right) + f(1) \right] \cdot \frac{1}{2} \cdot \frac{1}{2} = \left( \sqrt{1+0} + \sqrt{1+\frac{1}{2}} \right) \cdot \frac{1}{4} + \left( \sqrt{1+\frac{1}{2}} + \sqrt{1+1} \right) \cdot \frac{1}{4} = \\
&= \frac{1}{4} (1 + 2\sqrt{\frac{3}{2}} + \sqrt{2}) = 1.215926, \quad \varepsilon = \left| \frac{I_{t_2} - I}{I} \right| \cdot 100\% = \left| \frac{1.215926 - 1.218951}{1.218951} \right| \cdot 100\% = 0.2\% \\
I_{s_2} &= \left[ f(0) + 4f\left(\frac{1}{4}\right) + f\left(\frac{1}{2}\right) \right] \cdot \frac{1}{4 \cdot 3} + \left[ f\left(\frac{1}{2}\right) + 4f\left(\frac{3}{4}\right) + f(1) \right] \cdot \frac{1}{4 \cdot 3} = \left( 1 + 4\sqrt{\frac{5}{4}} + 2\sqrt{\frac{3}{2}} + 4\sqrt{\frac{7}{4}} + \sqrt{2} \right) \cdot \frac{1}{12} = \\
&= 1.218945, \quad \varepsilon = \left| \frac{I_{s_2} - I}{I} \right| \cdot 100\% = \left| \frac{1.218945 - 1.218951}{1.218951} \right| \cdot 100\% = 0.0005\%
\end{aligned}$$

## Kwadratury Gaussa

We wzorach Gaussa zastępujemy całkę analityczną w przedziale  $\langle -1, 1 \rangle$  kombinacją liniową wartości funkcji podcałkowej  $f(x_i)$  w tzw. *punktach Gaussa*  $x_i$  (węzły całkowania) oraz wag liczbowych  $\omega_i$ .

$$\int_{-1}^1 f(x) dx \approx \sum_{i=1}^N \omega_i f(x_i)$$

$N$  oznacza ilość *punktów Gaussa* (jak i również wag).

Wagi i węzły całkowania ustala się według zasady, by wzór całkowania przybliżony był wzorem ścisłym dla wielomianu możliwie wysokiego stopnia.

$$\sum_{i=1}^N \omega_i f(x_i) \approx \sum_{i=1}^N \omega_i x_i^k = \int_{-1}^1 x^k dx = \frac{1}{k+1} [1 - (-1)^{k+1}], \quad k = 0, 1, 2, \dots, 2N-1.$$

Np. dla  $N = 2$  (wzór *dwupunktowy Gaussa*)

$$\begin{aligned}
k=0 &\rightarrow \omega_1 \cdot 1 + \omega_2 \cdot 1 = 2 \\
k=1 &\rightarrow \omega_1 \cdot x_1 + \omega_2 \cdot x_2 = 0 \\
k=2 &\rightarrow \omega_1 \cdot x_1^2 + \omega_2 \cdot x_2^2 = \frac{2}{3} \\
k=3 &\rightarrow \omega_1 \cdot x_1^3 + \omega_2 \cdot x_2^3 = 0
\end{aligned}
\Rightarrow \begin{cases} \omega_1 = \omega_2 = 1 \\ x_1 = -\frac{1}{\sqrt{3}}, \quad x_2 = \frac{1}{\sqrt{3}} \end{cases}$$

W praktyce wagi i *punktów Gaussa* nie znajduje się w powyższego warunku. Pochodzą one mianowicie od rodziny pewnych wielomianów ortogonalnych (z wagą) w przedziale  $\langle -1, 1 \rangle$ . Wtedy *punkty Gaussa* są ich miejscami zerowymi. Powyższe liczby pochodzą od tzw. *wielomianów Legendre'a* – wtedy wzory *Gaussa* nazywane są *wzorami Gaussa – Legendre'a*. Wartości wag i miejsc zerowych tych wielomianów są tablicowane, tak jak innych kwadratur wykorzystujących wielomiany ortogonalne.

Tablica rodziny wzorów Gaussa – Legendre'a

| Stopień wielomianu Legendre'a | Miejsca zerowe wielomianów Legendre'a<br>$x_i$ | Wagi $\omega_i$                         |
|-------------------------------|--|---|
| 1                             | 0  | 2                                       |
| 2                             | $\pm \frac{1}{\sqrt{3}}$                       | 1, 1                                    |
| 3                             | $0, \pm \sqrt{\frac{3}{5}}$                    | $\frac{8}{9}, \frac{5}{9}, \frac{5}{9}$ |
| 4                             | $\pm 0.3399810436$<br>$\pm 0.8611363116$       | 0.6521451549<br>0.3478548451            |

W ogólnym przypadku liczymy całkę z dowolnego przedziału  $(a, b)$ . Konieczna jest więc transformacja liniowa między danym przedziałem a przedziałem  $\langle -1, 1 \rangle$ , tak aby można było zastosować powyższe dane z tabeli, obowiązujące tylko w tym przedziale.

Niech  $I = \int_a^b f(z) dz, \quad z \in (a, b)$ .

Wzory na transformację  $(a, b) \Rightarrow \langle -1, 1 \rangle$

$$x = \frac{2z - (a+b)}{b-a} \quad \rightarrow \quad z = \frac{b-a}{2}x + \frac{a+b}{2}$$

$$dx = \frac{2}{b-a} dz \quad dz = \frac{b-a}{2} dx$$

$$I = \int_a^b f(z) dz = \int_{-1}^1 f(z(x)) \frac{dz}{dx} dx = \frac{b-a}{2} \int_{-1}^1 f\left(\frac{b-a}{2}x + \frac{a+b}{2}\right) dx = \frac{b-a}{2} \int_{-1}^1 f(x) dx \approx \sum_{i=1}^N \omega_i f(x_i)$$

**Przykład 13**

Obliczyć całkę z poprzedniego przykładu  $I = \int_0^1 \sqrt{1+z} dz$  stosując wzory Gaussa: dwupunktowy i trzypunktowy.

$$a=0, \quad b=1 \quad \rightarrow \quad \begin{cases} z = \frac{1-0}{2}x + \frac{1+0}{2} = \frac{1}{2}x + \frac{1}{2} \\ dx = \frac{1}{2} dx \end{cases}$$

$$I = \int_0^1 \sqrt{1+z} dz = \frac{1}{2} \int_{-1}^1 \sqrt{\frac{1}{2}x + \frac{3}{2}} dx.$$

Wzór dwupunktowy:

$$I_{G_2} = \frac{1}{2} \left[ 1 \cdot \sqrt{\frac{1}{2} \cdot \frac{1}{\sqrt{3}} + \frac{3}{2}} + 1 \cdot \sqrt{\frac{1}{2} \cdot \left( -\frac{1}{\sqrt{3}} \right) + \frac{3}{2}} \right] = 1.219008,$$

$$\varepsilon = \left| \frac{I_{G_2} - I}{I} \right| \cdot 100\% = \left| \frac{1.219008 - 1.218951}{1.218951} \right| \cdot 100\% = 0.005\%$$

Wzór trzypunktowy:

$$I_{G_3} = \frac{1}{2} \left[ \frac{8}{9} \cdot \sqrt{0 + \frac{3}{2}} + \frac{5}{9} \cdot \sqrt{\frac{1}{2} \cdot \frac{\sqrt{3}}{\sqrt{5}} + \frac{3}{2}} + \frac{5}{9} \cdot \sqrt{\frac{1}{2} \cdot \left( -\frac{\sqrt{3}}{\sqrt{5}} \right) + \frac{3}{2}} \right] = 1.218952,$$

$$\varepsilon = \left| \frac{I_{G_3} - I}{I} \right| \cdot 100\% = \left| \frac{1.218952 - 1.218951}{1.218951} \right| \cdot 100\% = 0.00007\%$$

Wszystkie omawiane powyżej wzory całkowania numerycznego dotyczyły przypadków nieosobliwych, tzn. tzw. całej właściwych. Istnieją też całki niewłaściwe, gdy jedna z granic całkowania to nieskończoność (niewłaściwość I rodzaju) lub istnieje osobliwość funkcji podcałkowej w jednej z granic (niewłaściwość II rodzaju). W tych przypadkach na ogół nie da się stosować bezpośrednio wzorów całkowania numerycznego, należy dodatkowo przekształcić całkę analitycznie. W przypadku nieskończoności w jednej z granic wzory Newtona i Gaussa są bezużyteczne, bo nie da się wprowadzić węzłów całkowania do przedziału nieskończonego. W drugim przypadku osobliwości funkcji podcałkowej w jednej z granic nie można stosować wzorów Newtona – gdyż wymagają one znajomości wartości funkcji podcałkowej w jednej z granic – a jest ona równa nieskończoności. Wzory Gaussa można stosować, gdyż węzły całkowania pochodzą wtedy z wnętrza przedziału i nie natrafiają na punkt osobliwy.

*Całki niewłaściwe I rodzaju*

Można je przedstawić w postaci ogólnej  $\int_a^{\infty} f(x) dx$ .

Analityczne rozwiązanie wymaga liczenia granicy  $\int_a^{\infty} f(x) dx = F(x) \Big|_a^{\infty} = \lim_{x \rightarrow \infty} F(x) - F(a)$ .

Numeryczne rozwiązanie wymaga podstawienia typu  $t = \frac{1}{x}$ . Wtedy korzystając z twierdzenia o zmianie granic otrzymuje się nowe, skończone granice całkowania  $t_1 = t(\infty) = \frac{1}{\infty} = 0$ ,  $t_2 = t(a) = \frac{1}{a}$ . Postać całki nadaje się już do całkowania numerycznego.

$$I = \int_a^{\infty} f(x) dx = \int_{\frac{1}{a}}^0 f(x(t)) \frac{dx}{dt} dt = \dots$$

Wyjątkowo „złośliwa” jest następująca całka osobliwa  $I = \int_0^{\infty} f(x)dx$ . Proponowane podstawienie nie odniesie żądanego skutku, dlatego iż na powrót dostaniemy granicę całkowania równą nieskończoności  $t_1 = t(\infty) = \frac{1}{\infty} = 0$ ,  $t_2 = t(a=0) = \frac{1}{0} = \infty(!)$ . Dlatego też należy np. rozłożyć całkę na dwie całki składowe, tak, aby całka niewłaściwa miała drugą granicę różną od zera.

$$I = \int_0^{\infty} f(x)dx = \int_0^1 f(x)dx + \int_1^{\infty} f(x)dx.$$

Pierwszą całkę obliczamy numerycznie bezpośrednio, drugą składową po opisanym wyżej podstawieniu.

### Całki niewłaściwe II rodzaju

Ogólna postać całki:  $I = \int_a^b \frac{f(x)}{(x-a)^k} dx$ ,  $k \in \mathfrak{R}$ ,  $k \neq 0$ .

Aby pozbyć się osobliwości, należy usunąć ją z mianownika funkcji podcałkowej. Można to zrobić również przez podstawienie, ale łatwiejsze będzie w tym przypadku zastosowanie twierdzenia o całkowaniu przez części.

Całkowanie przez części:  $I = \int_a^b f(x)g'(x)dx = \begin{vmatrix} f(x) & f'(x) \\ g'(x) & g(x) \end{vmatrix} = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)dx$ .

W omawianym przypadku dla  $k = \frac{1}{2}$

$$I = \int_a^b \frac{f(x)}{\sqrt{x-a}} dx = \begin{vmatrix} f(x) & f'(x) \\ g(x) = \frac{1}{\sqrt{x-a}} & 2\sqrt{x-a} \end{vmatrix} = [2f(x)\sqrt{x-a}]_a^b - 2 \int_a^b f'(x)\sqrt{x-a} dx.$$

Ostatnia całkę można policzyć numerycznie bez żadnych trudności. Dla innych wartości  $k$  należy powtórzyć całkowanie przez części tak, aby otrzymać w końcu całkę właściwą. Przypadek szczególny  $k = 1$  doprowadzi do funkcji logarytmicznej, która będzie miała znowu osobliwość dla  $x = a$ . Taką całkę należy obliczać *kwadraturami Gaussa*.

### Przykład 14

Obliczyć numerycznie następujące całki niewłaściwe  $I = \int_a^{\infty} \frac{dz}{z^2}$  oraz  $I = \int_0^1 \frac{x}{\sqrt{1-x}}$

Wyniki analityczne

$$I = \int_1^{\infty} \frac{dz}{z^2} = \left. \frac{z^{-1}}{-1} \right|_1^{\infty} = -\left. \frac{1}{z} \right|_1^{\infty} = (-0+1) = 1.$$

$$I = \int_0^1 \frac{x}{\sqrt{1-x}} dx = -\int_0^1 \frac{\sqrt{(1-x)^2 - 1}}{\sqrt{1+x}} dx = -\int_0^1 \sqrt{1-x} dx + \int_0^1 \frac{1}{\sqrt{1-x}} dx = \left[ -\frac{2}{3}(1-x)^{\frac{3}{2}} + 2(1-x)^{\frac{1}{2}} \right]_0^1 = 1.333333$$



Przekształcenia analityczne (dla obliczeń numerycznych)

$$I = \int_1^{\infty} \frac{dz}{z^2} = \left| \begin{array}{l} t = \frac{1}{z^2} \rightarrow z = \frac{1}{\sqrt{t}} \quad t(\infty) = 0 \\ dz = -\frac{1}{2} t^{-\frac{3}{2}} dt \quad t(1) = 1 \end{array} \right| = \int_1^0 t \cdot \left(-\frac{1}{2}\right) \cdot t^{-\frac{3}{2}} dt = \frac{1}{2} \int_0^1 \frac{dt}{\sqrt{t}}.$$

$$I = \int_0^1 \frac{x}{\sqrt{1-x}} = \left| \begin{array}{l} f(x) = x \quad f'(x) = 1 \\ g'(x) = \frac{1}{\sqrt{1-x}} \quad 2\sqrt{1-x} \end{array} \right| = \underbrace{(2x\sqrt{1-x})}_0^1 - 2 \int_0^1 \sqrt{1-x} dx = -2 \int_0^1 \sqrt{1-x} dx.$$

Obliczenia numeryczne (wzór dwupunktowy Gaussa)

$$I = \frac{1}{2} \int_0^1 \frac{dt}{\sqrt{t}} = \left| \begin{array}{l} t(x) = \frac{1}{2}x + \frac{1}{2} \\ dt = \frac{1}{2} dx \end{array} \right| = \frac{1}{4} \int_{-1}^1 \frac{dx}{\sqrt{\frac{1}{2}x + \frac{1}{2}}} \approx \frac{1}{4} \left( 1 \cdot \frac{1}{\sqrt{\frac{1}{2} \cdot \frac{1}{\sqrt{3}} + \frac{1}{2}}} + 1 \cdot \frac{1}{\sqrt{\frac{1}{2} \cdot \left(-\frac{1}{\sqrt{3}}\right) + \frac{1}{2}}} \right) = 0.825340$$

$$\varepsilon = \left| \frac{I_{G_2} - I}{I} \right| \cdot 100\% = \left| \frac{0.825340 - 1.0}{1.0} \right| \cdot 100\% = 17.4\%$$

$$I = -2 \int_0^1 \sqrt{1-t} dt = 2 \int_1^0 \sqrt{1-t} dt = \left| \begin{array}{l} t(x) = -\frac{1}{2}x + \frac{1}{2} \\ dt = -\frac{1}{2} dx \end{array} \right| = \int_{-1}^1 \sqrt{\frac{1}{2} + \frac{1}{2}x} dx \approx 1 \cdot \sqrt{\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{\sqrt{3}}} + 1 \cdot \sqrt{\frac{1}{2} + \frac{1}{2} \cdot \left(-\frac{1}{\sqrt{3}}\right)} = 1.347775$$

$$\varepsilon = \left| \frac{I_{G_2} - I}{I} \right| \cdot 100\% = \left| \frac{1.347775 - 1.333333}{1.333333} \right| \cdot 100\% = 1.1\%$$

## IV. NUMERYCZNE ROZWIĄZYWANIE PROBLEMÓW POCZĄTKOWYCH

Ogólne sformułowanie problemu początkowego

$$\frac{d^{(n)}y}{dx^{(n)}} = f(x, y, y', \dots, y^{(n-1)}), \quad x \in (a, b)$$

$$y(x_0) = y_0, \quad y'(x_0) = y'_0, \quad \dots, \quad y^{(n-1)}(x_0) = y_0^{(n-1)}; \quad x_0 \in (a, b)$$

Szczególnym przypadkiem problemu początkowego jest równanie różniczkowe rzędu pierwszego z warunkiem na niewiadomą funkcję. Równania wyższych rzędów sprowadza się do równania rzędu pierwszego i rozwiązuje niezależnie.

$$\frac{dy}{dx} = f(x, y), \quad x \in (a, b), \quad y(x_0) = y_0; \quad x_0 \in (a, b)$$

Metody numeryczne pozwalają na wyznaczenie zbioru wartości dyskretnej funkcji niewiadomej  $y = y(x)$  począwszy od punktu początkowego  $x_0$ . Zbiór par  $(x_i, y_i)$  wyznacza się z następujących zależności (dla węzłów równoodległych  $x_{i+1} - x_i = x_i - x_{i-1} = h = \text{const.}$ )

$$\begin{cases} x_{i+1} = x_i + h = x_0 + i \cdot h \\ y_{i+1} = y_i + \int_{x_i}^{x_{i+1}} f(t, y) dt = y_i + \Delta y_i, \quad y_0 = y(x_0) \end{cases}$$

Całkę oznaczoną przez  $\Delta y_i$  oblicza się numerycznie na różne sposoby. W zależności od sposobu jej obliczania metody numeryczne do rozwiązywania zadań początkowych można podzielić na jednokrokowe  $\Delta y_i = \Delta y_i(f_i)$ ,  $f_i \equiv f(x_i, y_i)$  (wartość delty zależy tylko od jednego punktu wstecz) i wielokrokowe  $\Delta y_i = \Delta y_i(f_i, f_{i-1}, f_{i-2}, \dots)$  (wartość delty zależy od kilku punktów wstecz). Inna klasyfikacja dotyczy tzw. jawności metod. Przedstawione wyżej wzory dotyczyły metod jawnych (otwartych, ekstrapolacyjnych) – wartość  $y_{i+1}$  liczona jest na podstawie znanych wartości funkcji danych lub obliczonych wcześniej w poprzednich punktach -  $\Delta y_i = \Delta y_i(f_i, f_{i-1}, f_{i-2}, \dots)$ . Natomiast inną grupę metod stanowią bardzo dokładne metody niejawne (zamknięte, interpolacyjne), gdzie wartość  $y_{i+1}$  zależna jest od siebie samej poprzez deltę  $\Delta y_i = \Delta y_i(f_{i+1}, f_i, f_{i-1}, \dots)$ . Oblicza się ją stosując metody iteracyjne, startujące ze wstępnego określenia wartości  $y_{i+1}^{(0)}$  znanego z metody jedno- lub wielokrokowej otwartej.

### Metody jednokrokowe

- *Metoda Eulera* (metoda ta zakłada stałość funkcji  $y(x)$  na odcinku  $(x_i, x_{i+1})$ ).

$$y_{i+1} = y_i + h \cdot f(x_i, y_i)$$

- *Metoda ulepszona Eulera*

$$\begin{cases} y_{i+1} = y_i + h \cdot f(x_i, y_i) \\ \tilde{f}_i = \frac{f_i + f_{i+1}}{2} \\ y_{i+1} = y_i + h \cdot \tilde{f}_i \end{cases}$$

- *Metoda Rungego – Kutty II rzędu*

$$\begin{aligned} K_1 &= h \cdot f(x_i, y_i) \\ K_2 &= h \cdot f(x_i + h, y_i + K_1) \\ y_{i+1} &= y_i + \frac{1}{2}(K_1 + K_2) \end{aligned}$$

- *Metoda Rungego – Kutty IV rzędu*

$$\begin{aligned} K_1 &= h \cdot f(x_i, y_i) \\ K_2 &= h \cdot f(x_i + \frac{1}{2}h, y_i + \frac{1}{2}K_1) \end{aligned}$$

$$K_3 = h \cdot f\left(x_i + \frac{1}{2}h, y_i + \frac{1}{2}K_2\right)$$

$$K_4 = h \cdot f(x_i + h, y_i + K_3)$$

$$y_{i+1} = y_i + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)$$

### Metody wielokrokowe

- *Metoda Adamsa – Bashfortha* (metoda otwarta)

$$y_{i+1} = y_i + h \cdot \sum_{j=n}^{j=0} f_{i-j} \cdot L_{i-j}^{(n)} = y_i + h \cdot (f_{i-n} \cdot L_{i-n}^{(n)} + \dots + f_{i-1} \cdot L_{i-1}^{(n)} + f_i \cdot L_i^{(n)})$$

**Tabela współczynników wzorów Adamsa – Bashfortha /·h**

| $n/k$ | 0               | 1                | 2               | 3               |
|-------|-----------------|------------------|-----------------|-----------------|
| 0     | 1               |                  |                 |                 |
| 1     | $\frac{3}{2}$   | $-\frac{1}{2}$   |                 |                 |
| 2     | $\frac{23}{12}$ | $-\frac{10}{12}$ | $\frac{5}{12}$  |                 |
| 3     | $\frac{55}{24}$ | $-\frac{59}{24}$ | $\frac{37}{24}$ | $-\frac{9}{24}$ |

Np. dla  $n = 2$ :  $y_{i+1} = y_i + \frac{h}{12} \cdot (23f_i - 16f_{i-1} + 5f_{i-2})$ .

- *Metoda Adamsa – Moultona* (metoda zamknięta)

$$y_{i+1} = y_i + h \cdot \sum_{j=n}^{j=0} f_{i-j+1} \cdot L_{i-j+1}^{(n)} = y_i + h \cdot (f_{i-n+1} \cdot L_{i-n+1}^{(n)} + \dots + f_i \cdot L_i^{(n)} + f_{i+1} \cdot L_{i+1}^{(n)})$$

**Tabela współczynników wzorów Adamsa – Moultona /·h**

| $n/k$ | 0              | 1               | 2               | 3              |
|-------|----------------|-----------------|-----------------|----------------|
| 0     | 1              |                 |                 |                |
| 1     | $\frac{1}{2}$  | $\frac{1}{2}$   |                 |                |
| 2     | $\frac{5}{12}$ | $\frac{8}{12}$  | $-\frac{1}{12}$ |                |
| 3     | $\frac{9}{24}$ | $\frac{19}{24}$ | $-\frac{5}{24}$ | $\frac{1}{24}$ |

Np. dla  $n = 2$ :  $y_{i+1} = y_i + \frac{h}{12} \cdot (5f_{i+1} + 8f_i - f_{i-1})$ .

### Przykład 15

Znaleźć wartość funkcji  $f(1)$ , jeżeli

$$f' = f^2 + 2 + x, \quad f(0) = 1, \quad h = 1$$

metodą Rungego - Kutty 4 rzędu.

$$x_0 = 0, \quad f_0 = f(x_0) = f(0) = 1, \quad F(x, f) = f^2 + 2 + x, \quad h = 1$$

$$K_1 = h \cdot F(x_0, f_0) = 1 \cdot F(0, 1) = 1 + 2 + 0 = 3$$

$$K_2 = h \cdot F(x_0 + \frac{1}{2}h, f_0 + \frac{1}{2}K_1) = 1 \cdot F(0 + \frac{1}{2} \cdot 1, 1 + \frac{1}{2} \cdot 3) = \left(\frac{5}{2}\right)^2 + 2 + \frac{1}{2} = \frac{35}{4}$$

$$K_3 = h \cdot F(x_0 + \frac{1}{2}h, f_0 + \frac{1}{2}K_2) = 1 \cdot F(0 + \frac{1}{2} \cdot 1, 1 + \frac{1}{2} \cdot \frac{35}{4}) = \left(\frac{43}{8}\right)^2 + 2 + \frac{1}{2} = \frac{2009}{64}$$

$$K_4 = h \cdot F(x_0 + h, f_0 + K_3) = 1 \cdot F(0 + 1, 1 + \frac{2009}{64}) = \left(\frac{2073}{64}\right)^2 + 2 + 1 = \frac{4309617}{64 \cdot 64}$$

$$f_{i+1} = f_i + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4) = 1 + \frac{1}{6}\left(3 + \frac{35}{4} + \frac{2009}{64} + \frac{4309617}{64 \cdot 64}\right) = 183.548869$$

Praktyczne stosowanie metod zamkniętych (zazwyczaj wielokrokowych) wiąże się z następującym algorytmem iteracyjnym zwanym zwyczajowo *metodą predyktor – korektor*. Polega ona na znalezieniu kilku pierwszych wartości funkcji metodą jednokrokową wysokiego rzędu (np. *metodą Rungego – Kutty IV rzędu*), a następnie wstępnego określenia (predykcji – stąd nazwa „predyktor”) szukanej, następnej z kolei wartości funkcyjnej za pomocą wzoru otwartego wielokrokowego. Wartość ta służy jako punkt startowy dla metody wielokrokowej zamkniętej, która iteracyjnie poprawia (stąd nazwa „korektor”) szukaną wartość aż do osiągnięcia wymaganej dokładności.

Dla przykładu rozważmy równanie początkowe rzędu pierwszego

$$\frac{dy}{dx} = f(x, y), \quad x \in (a, b), \quad y(x_0) = y_0; \quad x_0 \in (a, b).$$

Dwie pierwsze wartości funkcyjne znaleziono stosując *metodę Rungego – Kutty rzędu IV*.

$$y_0 = y(x_0) \rightarrow \text{z warunku początkowego}$$

$$\left. \begin{array}{l} y_1 = y_0 + \Delta y_0 \\ y_2 = y_1 + \Delta y_1 \end{array} \right\} \rightarrow \text{z metody Rungego - Kutty IV rzędu}$$

Wartość  $y_3$ , a dokładniej jej zerowe przybliżenie znaleziono korzystając z *metody Adamsa – Bashfortha rzędu II*.

$$y_3^{(0)} = y_2 + \frac{h}{12} \cdot (23f_2 - 16f_1 + 5f_0), \quad f_i \equiv f(x_i, y_i), \quad i = 0, 1, 2.$$

Następnie posłużono się odpowiednim schematem zamkniętym (*metoda Adamsa – Moultona rzędu II*) układając w ten sposób procedurę iteracyjną, kontrolowaną przed warunek zbieżności na podstawie znanej dokładności wyniku  $\varepsilon$ .

$$y_3^{(k+1)} = y_2 + \frac{h}{12} \cdot (5f_3^{(k)} + 8f_2 - f_1), \quad \begin{cases} f_i \equiv f(x_i, y_i), & i = 1, 2; \\ f_i^{(k)} = f(x_i, y_i^{(k)}). \end{cases}, \quad \text{gdzie dla } k = 0 \text{ wynik } y_3^{(0)}$$

pochodzi z poprzedniej metody (z predyktora). Wynik poprawiamy sprawdzając na każdym kroku warunek zbieżności

$$\left\| \frac{y_3^{(k+1)} - y_3^{(k)}}{y_3^{(k+1)}} \right\| \leq \varepsilon.$$

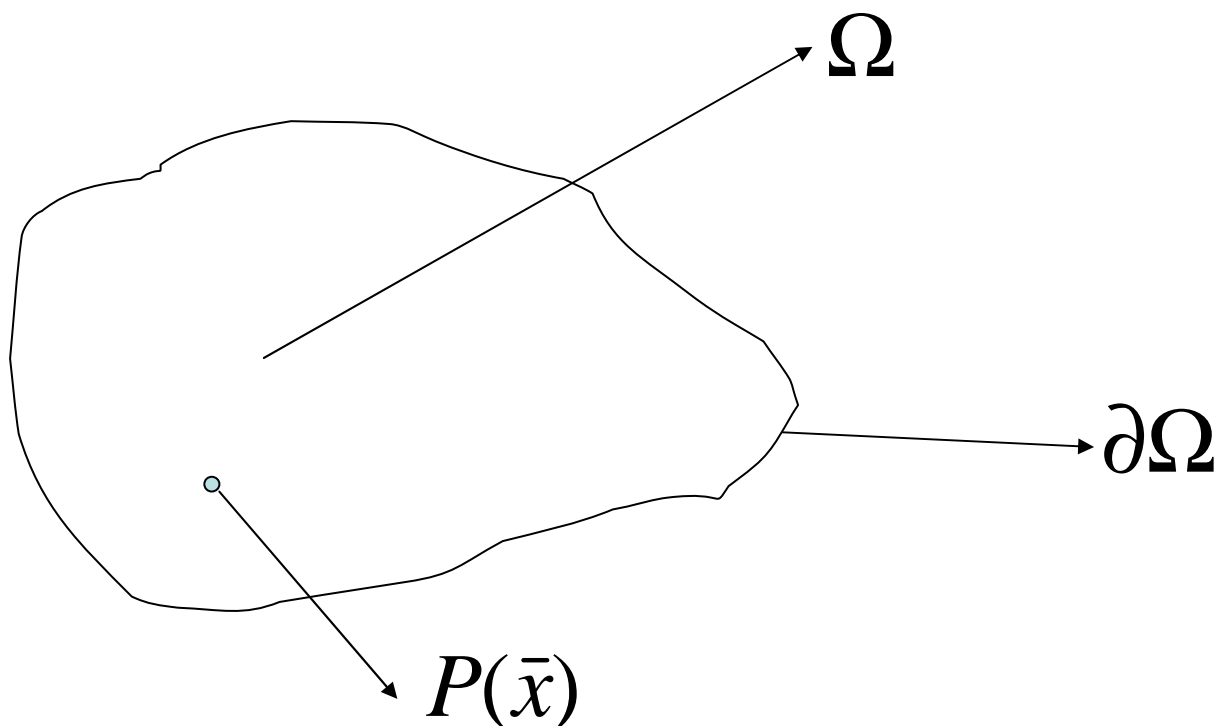
Gdy wynik się ustabilizuje, można przejść do obliczania następnej wartości funkcji  $y_4$  w ten sam sposób, co powyżej.

## V. NUMERYCZNE ROZWIĄZYWANIE PROBLEMÓW BRZEGOWYCH

Podstawową różnicą między problemem początkowym i brzegowym jest sposób określenia warunków. W problemie początkowym warunki (początkowe) nałożone były na funkcję niewiadomą i jej kolejne pochodne aż do odpowiedniego rzędu w jednym, wybranym punkcie obszaru. W problemach brzegowych na ogół mamy do czynienia ze zbiorem punktów, w których dane są wartości funkcji lub jej pochodnych. Metody numeryczne do rozwiązywania obydwu problemów diametralnie różnią się od siebie. Problemy początkowe numerycznie prowadziły do znalezienia tablicy wartości funkcji punkt po punkcie zaczynając od punktu z warunkiem początkowym. W metodach dyskretnych do analizy zadań brzegowych otrzymujemy dla zadanego zbioru punktów (węzłów) układ równań, z którego jednocześnie otrzymujemy wartości we wszystkich niewiadomych węzłach.

Niezwykle ważną rzeczą jest sposób sformułowania problemu brzegowego. Ogólnie każdy zapis problemu, w którym występuje nieznaną funkcją jest dopuszczalny, ale w zagadnieniach fizyki i mechaniki funkcjonują od lat dwa zasadnicze typy sformułowań brzegowych – lokalne i globalne. Również od sformułowania zależy sposób otrzymania i jakość wyniku różnicowego.

*Zagadnienie (problem) brzegowe:* dany jest obszar  $\Omega$ , w którym poszukiwane jest rozwiązanie, układ równań różniczkowych cząstkowych oraz warunki początkowo – brzegowe nałożone na zbiór punktów należących do brzegu  $\partial\Omega$  obszaru.



W rozważanym obszarze poszukiwana jest funkcja  $u(\bar{x})$  w każdym punkcie  $P(\bar{x})$ . Można stosować następujące sformułowania zagadnień brzegowych:

- *Sformułowanie lokalne (mocne, silne)*: szukane jest rozwiązanie układu równań różniczkowych w każdym z punktów obszaru osobno:

$$\begin{aligned} \mathcal{L}u &= f \quad \text{dla } P \in \Omega \\ \mathcal{B}u &= g \quad \text{dla } P \in \partial\Omega \end{aligned}$$

gdzie  $\mathcal{L}$  i  $\mathcal{B}$  są operatorami różniczkowymi odpowiednio w obszarze i na jego brzegu. Równanie  $\mathcal{B}u = g$  dla  $P \in \partial\Omega$  nosi nazwę warunków brzegowych. Jeżeli są one nałożone na funkcję (tzn.  $\mathcal{B} \equiv 1$ ), noszą nazwę podstawowych *warunków brzegowych Dirichleta*, natomiast dowolna kombinacja warunków brzegowych złożona z pochodnych nosi nazwę *warunków brzegowych Neumanna*.

- *Sformułowanie globalne*: może być formułowane jako problem optymalizacji funkcjonału lub jako zasada wariacyjna.
  - *Minimalizacja funkcjonału*:

$$I(u) = \frac{1}{2} \mathcal{B}(u, u) - \mathcal{L}(u)$$

W funkcjonałach energetycznych pierwszy składnik prezentuje energię wewnętrzną układu, podczas gdy drugi jest równy pracy wykonanej przez siły zewnętrzne. Nieznana funkcja  $u(P)$  może przedstawiać sobą przemieszczenia  $u$ , odkształcenia  $\varepsilon$ , naprężenia  $\sigma$  lub wszystkie z nich. Funkcja  $u$  realizująca ekstremum (minimum, punkt stacjonarny) funkcjonału  $\min_{(u)} I(u)$  jest szukana.

Można rozważać problem optymalizacji funkcjonału bez ograniczeń (w całej

przestrzeni rozwiązań dopuszczalnych) lub z ograniczeniami (ekstremum jest szukane w podprzestrzeni narzuconych ograniczeń).

- o *Zasada wariacyjna*

$$\mathcal{B}(u, \partial u) = \mathcal{L}(\partial u) \quad \text{dla} \quad \partial u \in V$$

W mechanice powyższe równanie może mieć sens np. zasady prac wirtualnych. Sformułowanie wariacyjne (tzw. słabe) ma podstawowe znaczenie przy konstruowaniu rozwiązań przybliżonych. Można go uzyskać ze sformułowania mocnego w czterech krokach:

- Przemnożenie równania różniczkowego przez dowolną funkcję (tzw. funkcja testująca),
- Przecałkowanie wyniku po rozważanym obszarze  $\Omega$ ,
- Całkowanie przez części z wykorzystaniem *twierdzenia Greena* w celu zredukowania pochodnych do minimalnego rzędu,
- Wprowadzenie do funkcjonału *warunków brzegowych Neumanna*.

Sformułowania globalne wymagają dodatkowego całkowania po obszarze. Sformułowanie wariacyjne jest ogólniejsze, gdyż możliwe jest w przypadku wszystkich zagadnień brzegowych, podczas gdy ułożenie funkcjonału możliwe jest tylko dla niektórych zadań mechaniki, np. dla zadań liniowej sprężystości (*funkcjonał Lagrange'a, Hamiltona, Reissnera, Castigliano*, itp.).

- Możliwe są również podejścia mieszane, polegające np. na podziale obszaru  $\Omega$  na podobszary, gdzie stosuje się różne sformułowania wraz z odpowiednimi warunkami ograniczającymi.

Budowa rozwiązania przybliżonego problemu brzegowego zależy przede wszystkim od wybranej metody dyskretnej. Można wyróżnić dwie główne koncepcje:

- Rozwiązanie dyskretne w postaci kombinacji liniowej współczynników liczbowych oraz funkcji bazowych:

$$p(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \dots + a_n\varphi_n(x) = \sum_{i=1}^n a_i\varphi_i(x).$$

Funkcje bazowe (najczęściej: wielomiany, funkcje trygonometryczne, funkcje specjalne) muszą być liniowo niezależne, odpowiednio ciągłe oraz muszą spełniać jednorodne warunki brzegowe rozważanego problemu (jednorodne warunki to takie, w których po prawej stronie stoi 0, (np.  $u(x_0) = 0$ ,  $u'(x_0) = 0$ ). Przy takim zapisie postaci rozwiązania przybliżonego można szukać budując odpowiednie residua, (czyli wyrażenia świadczące o spełnieniu przez rozwiązanie przybliżone wyjściowych równań różniczkowych) odpowiednio w obszarze i na brzegu:

$$\varepsilon_d(x) = \mathcal{L}p(x) - f, \quad \varepsilon_b(x) = \mathcal{B}p(x) - g$$

Funkcjonał wazący powyższe wyrażenia ma postać:

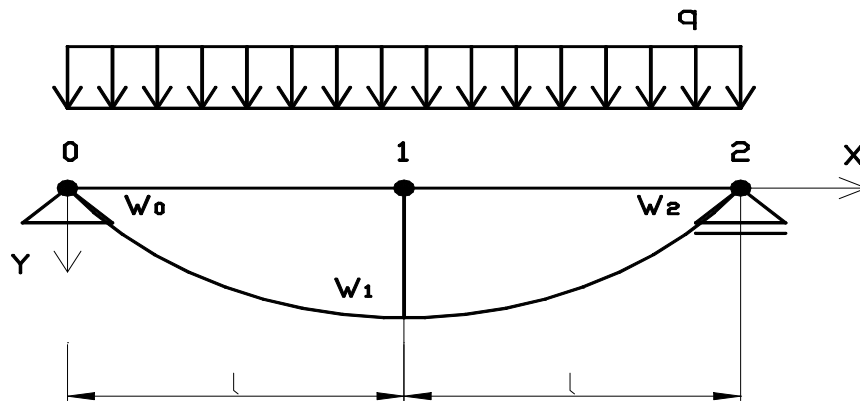
$$I(p) = \int_{\Omega} \varepsilon_d w_d d\Omega + \int_{\partial\Omega} \varepsilon_b w_b d\partial\Omega.$$

Wagi  $w_a$  i  $w_b$  świadczą o odejściu  $p(x)$  od wyniku ścisłego odpowiednio w obszarze i na jego brzegi. Dla *metod residuów ważonych (metoda Bubnowa - Galerkina, metoda najmniejszych kwadratów, metoda kolokacji)* i *metod energetycznych (metoda Rayleigha – Ritz)* zakłada się błąd na brzegu  $\varepsilon_b = 0$  (ściśle spełnienie warunków brzegowych) i rozważa jedynie  $I(p) = \int_{\Omega} \varepsilon_d w_d d\Omega$ . Odmianą koncepcję prezentują tzw. *metody Trefza*, w których zakłada się ściśle spełnienie równania wewnątrz obszaru a rozwiązań przybliżonych poszukuje na jego brzegu.

- Rozwiązanie dyskretne w wybranych punktach obszaru (lub/i jego brzegu) zwanych węzłami. W tej koncepcji niezbędna jest dyskretyzacja obszaru (na węzły, elementy itp.), gdzie zastępuje się wielkości ciągłe wielkościami dyskretnymi. Numeryczne wyniki dyskretne można aproksymować funkcją ciągłą w ramach tzw. *postprocesingu*. Do tych metod należą: *metoda różnic skończonych (MRS, zamiana operatorów różniczkowych na różnicowe, poszukiwanie wartości węzłowych funkcji szukanej, aproksymacja metodami najmniejszych kwadratów)*, *metoda elementów skończonych (MES, podział na elementy i aproksymacja funkcjami kształtu)* oraz *metod elementów brzegowych (MEB, podział brzegu na odcinki, obliczanie całek brzegowych)*.

### Przykład 16

Belka swobodnie podparta obciążona obciążeniem ciągłym równomiernie rozłożonym.



Sformułowanie lokalne:

$$\mathcal{L}(x) = \frac{d^2}{dx^2} w(x) = f(x) \quad f(x) = -\frac{M(x)}{EJ} \quad 0 \leq x \leq 2l$$

$$M(x) = \frac{1}{2} qx(2l - x) \quad w(0) = 0 \quad w(2l) = 0$$

Sformułowanie globalne:

- W postaci funkcjonału:

$$\min_w I(w) \Rightarrow I(w) = \int_0^{2l} \left[ \frac{1}{2} \left( \frac{dw}{dx} \right)^2 - \frac{M(x)}{EJ} w \right] dx, \quad w(0) = w(2l) = 0$$



- W postaci zasady wariacyjnej:

$$\int_0^{2l} \left[ \frac{dw^2}{dx^2} + \frac{M(x)}{EJ} \right] v(x) dx = 0, \quad w(0) = w(2l) = 0$$

$v(x)$  – funkcja próbna, odpowiednio ciągła, spełnia warunki brzegowe:  $v(0) = v(2l) = 0$

lub po przecałkowaniu przez części (sformułowanie słabe):

$$\int_0^{2l} \left[ \frac{dw}{dx} \frac{dv}{dx} - \frac{M(x)}{EJ} v(x) \right] dx = 0, \quad w(0) = w(2l) = 0, \quad v(0) = v(2l) = 0$$

Rozwiązanie przybliżone dla *metod residualnych*:

- Funkcje bazowe:  $\varphi_1(x) = x(x-2l)$ ,  $\varphi_2(x) = x^2(x-2l)$ ,
- Rozwiązanie próbne:  $p(x) = a\varphi_1(x) + b\varphi_2(x) = a x(x-2l) + b x^2(x-2l)$ ,
- Residuum w obszarze:  $\varepsilon(x) = \frac{d^2 p(x)}{dx^2} + \frac{M(x)}{EJ} = 2a + b(6x-4l) - \frac{1}{2}qx(2l-x)$
- Dla *metody Bubnowa - Galerkina*:

$$\begin{cases} \int_0^{2l} \varepsilon(x) \cdot \varphi_1(x) dx = 0 \\ \int_0^{2l} \varepsilon(x) \cdot \varphi_2(x) dx = 0 \end{cases} \Rightarrow \begin{cases} \int_0^{2l} \left[ 2a + b(6x-4l) - \frac{1}{2}qx(2l-x) \right] x(x-2l) dx = 0 \\ \int_0^{2l} \left[ 2a + b(6x-4l) - \frac{1}{2}qx(2l-x) \right] x^2(x-2l) dx = 0 \end{cases} \Rightarrow a, b = \dots$$

- Dla *metody najmniejszych kwadratów*:

$$I(a, b) = \int_0^{2l} \varepsilon(x) \cdot \varepsilon(x) dx \Rightarrow \min_{(a, b)} I(a, b)$$

$$I(a, b) = \int_0^{2l} \left[ 2a + b(6x-4l) - \frac{1}{2}qx(2l-x) \right]^2 dx \Rightarrow \begin{cases} \frac{\partial}{\partial a} I(a, b) = 0 \\ \frac{\partial}{\partial b} I(a, b) = 0 \end{cases} \Rightarrow a, b = \dots$$

- Dla *metody kolokacji* (punkty kolokacji:  $x_1 = \frac{l}{3}$ ,  $x_2 = \frac{2l}{3}$ ):

$$\begin{cases} \int_0^{2l} \varepsilon(x) \cdot \delta(x-x_1) dx = 0 \\ \int_0^{2l} \varepsilon(x) \cdot \delta(x-x_2) dx = 0 \end{cases} \Rightarrow \begin{cases} \varepsilon(x_1) = 0 \\ \varepsilon(x_2) = 0 \end{cases} \Rightarrow \begin{cases} 2a + b(6x_1-4l) - \frac{1}{2}qx_1(2l-x_1) = 0 \\ 2a + b(6x_2-4l) - \frac{1}{2}qx_2(2l-x_2) = 0 \end{cases} \Rightarrow a, b = \dots$$

W *metodzie różnic skończonych MRS* wprowadzono w ramach dyskretyzacji obszaru 3 węzły (patrz: rysunek). Z trzech wartości węzłowych dwie z nich stanowią warunki brzegowe:

$w_0 = w_2 = 0$ , pozostaje do obliczenia wartość  $w_1$ . Przy sformułowaniu lokalnym zamianie na operator różnicowy ulega operator różniczkowy na drugą pochodną:  $\frac{d^2 w}{dx^2} \Big|_{x=l} = w_1'' \approx Lw_1 = \frac{w_0 - 2w_1 + w_2}{l^2}$ . Równania różnicowe generuje się metodą kolokacji (ściśle spełnienie równania w węzłach obszaru):

$$Lw_1 = f_1 \Rightarrow \frac{\overset{0}{w_0} - 2w_1 + \overset{0}{w_2}}{l^2} = \frac{1}{2} \frac{ql^2}{EJ} \Rightarrow w_1 = \dots$$

W sformułowaniu globalnym można ułożyć funkcjonal energii potencjalnej układu. Po jego dyskretyzacji (całkowanie *kwadraturą Newtona-Cotesa* między węzłami) otrzymuje się:

$$I(w_0, w_1, w_2) = \frac{1}{2} \left[ \left( \frac{w_1 - w_0}{l} \right)^2 l + \left( \frac{w_2 - w_1}{l} \right)^2 l - (M_0 w_0 + M_1 w_1) \frac{l}{2EJ} - (M_1 w_1 + M_2 w_2) \frac{l}{2EJ} \right]$$

Niewiadomą  $w_1$  (oczywiście  $w_0 = w_2 = 0$ ) otrzymuje się minimalizując powyższy funkcjonal względem  $w_1$ :

$$\frac{d}{dw_1} I(w_1) = 0 \Rightarrow w_1 = \dots$$

Przy sformułowaniu wariacyjnym słabym (funkcja testowa:  $v(0) = v(2l) = 0$ ) od razu otrzymuje się gotowe równanie różnicowe:

$$\frac{w_1 - w_0}{l} \frac{v_1 - v_0}{l} l + \frac{w_2 - w_1}{l} \frac{v_2 - v_1}{l} l - (M_0 w_0 + M_1 w_1) \frac{l}{2EJ} - (M_1 w_1 + M_2 w_2) \frac{l}{2EJ} = 0.$$

Podstawiając  $w_0 = w_2 = 0$  oraz  $v_0 = v_2 = 0$  i przyrównując wyrażenie stojące przy dowolnym  $v_1$  do zera otrzymuje się wartość  $w_1$ .

### Przykład 17

Rozwiązać równanie

$$w'' + w = 1 \quad w(0) = 0, \quad w(4) = 1$$

*metodą różnic skończonych* i analitycznie. Wynik sprawdzić analitycznie dla  $x = 2$  (obliczyć normę błędu)

Rozwiązanie analityczne

$$CO/RJ: \quad w'' + w = 0 \rightarrow r^2 + 1 = 0 \rightarrow w_0(x) = A \sin(x) + B \cos(x) \text{ - całka ogólna}$$

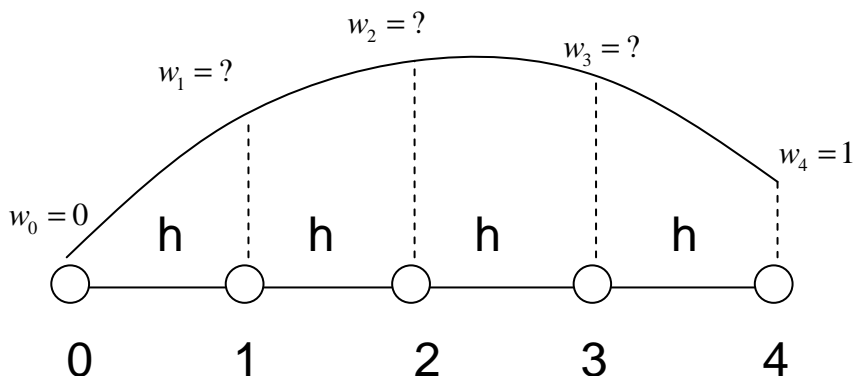
$$CS/RNJ \quad w_s(x) = C \rightarrow w_s''(x) + w_s(x) = 1 \rightarrow C = 1 \rightarrow w_s(x) = 1 \text{ - całka szczególna}$$

$$w(x) = w_0(x) + w_s(x) = A \sin(x) + B \cos(x) + 1$$

$$\begin{cases} w(0) = 0 \\ w(4) = 1 \end{cases} \Rightarrow \begin{cases} B + 1 = 0 \\ A \sin(4) + B \cos(4) + 1 = 1 \end{cases} \Rightarrow \begin{cases} A = 0.863691 \\ B = -1 \end{cases}$$

$$\Rightarrow w(x) = 0.863691 \cdot \sin(x) - 1 \cdot \cos(x) + 1$$

Rozwiązanie numeryczne (*metoda różnic skończonych MRS*)



Wprowadzono do obszaru zadania  $x \in \langle 0, 4 \rangle$  pięć równoodległych węzłów ( $h = 1$ ). Warunki brzegowe  $w_0 = w(x_0 = 0) = 0$ ,  $w_4 = w(x_4 = 4) = 1$ . Przyjęto klasyczny operator różnicowy na drugą pochodną (zbudowany na trzech węzłach).

$$w_i'' \approx \frac{w_{i-1} - 2w_i + w_{i+1}}{h^2}.$$

Generacja równań różnicowych (techniką *kolokacji*)

$$\begin{cases} \frac{w_0 - 2w_1 + w_2}{1^2} + w_1 = 1 \\ \frac{w_1 - 2w_2 + w_3}{1^2} + w_2 = 1 \\ \frac{w_2 - 2w_3 + w_4}{1^2} + w_3 = 1 \\ w_0 = 0, \quad w_4 = 1 \end{cases} \Rightarrow \begin{cases} -2w_1 + w_2 + w_1 = 1 \\ w_1 - 2w_2 + w_3 + w_2 = 1 \\ w_2 - 2w_3 + 1 + w_3 = 1 \end{cases}$$

$$\begin{cases} -w_1 + w_2 = 1 \\ w_1 - w_2 + w_3 = 1 \\ w_2 - w_3 = 0 \end{cases} \Rightarrow \begin{cases} w_1 = 1 \\ w_2 = 2 \\ w_3 = 2 \end{cases}$$

Ścisłe wartości węzłowe (z rozwiązania analitycznego)  $w(1) = 1.186469$ ,  $w(2) = 2.201499$ ,  $w(3) = 2.111877$ .

Norma błędów wyniku numerycznego dla  $x = 2$ :

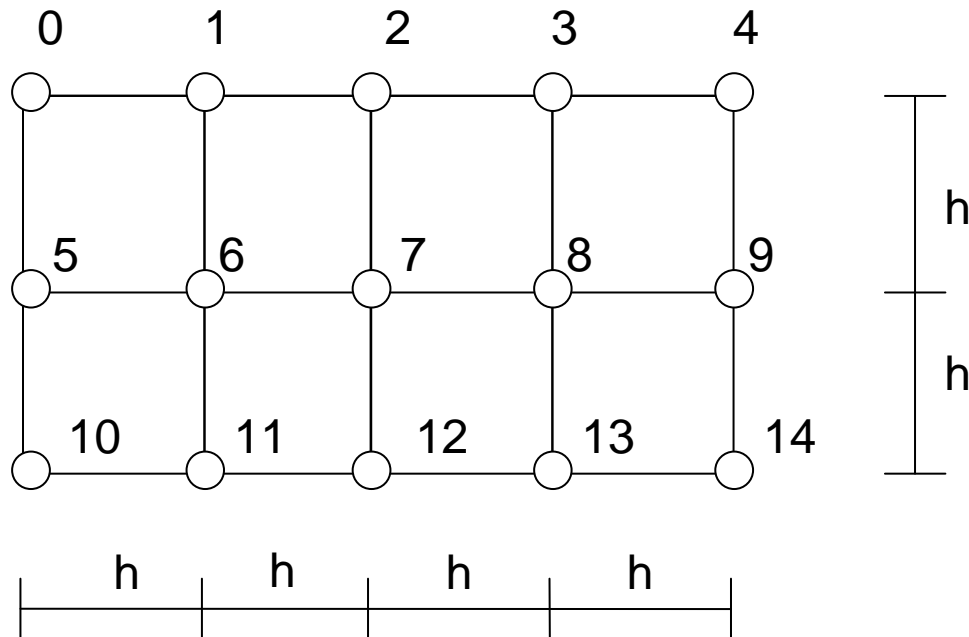
$$\varepsilon = \left| \frac{w(2) - w_2}{w(2)} \right| \cdot 100\% = \left| \frac{2.201499 - 2.0}{2.201499} \right| \cdot 100\% = 9.2\%.$$

### Przykład 18

Znaleźć wartości węzłowe dla równania

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)f = 1, \quad h = 1$$

przy zerowych warunkach brzegowych na funkcję.



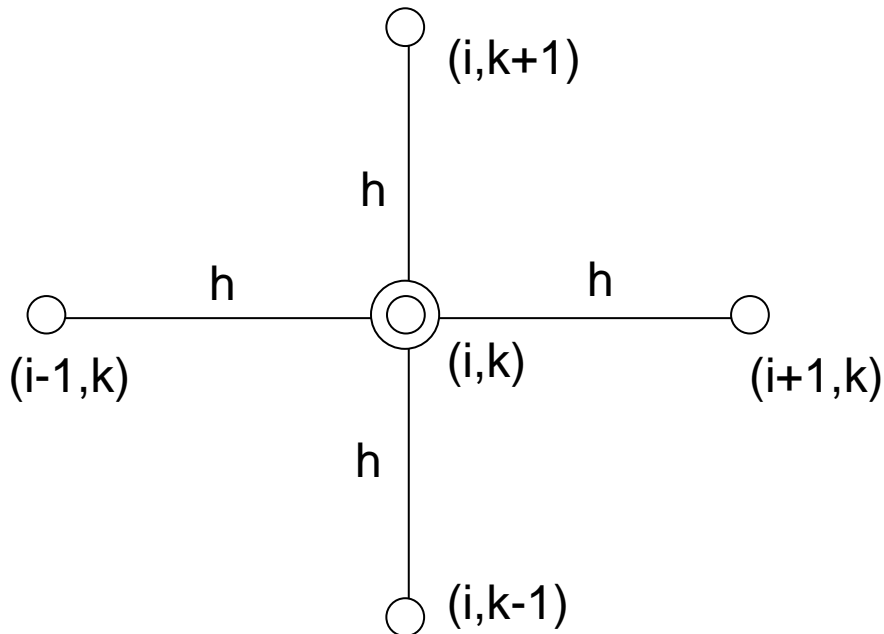
Zadanie brzegowe należy do dziedziny zadań dwuwymiarowych, typu *eliptycznego*. Występujący w sformułowaniu problemu operator różniczkowy zwie się *operatorem Laplace'a*. Mimo to metodologia postępowania jest identyczna jak w zadaniach jednowymiarowych. Obszar zadania podlega dyskretyzacji – wprowadzono 15 węzłów numerowanych od 0 do 14 równomiernie rozłożonych w obszarze (obszarze obydwu kierunkach  $h = 1$ ). 12 z nich to węzły brzegowe, w których z warunków zadania wiadomo, że  $f = 0$ . Pozostałe węzły zawierają niewiadome węzłowe wartości. W zadaniu można skorzystać z warunków symetrii (symetria wynika z geometrii obszaru, postaci warunków brzegowych i postaci funkcji prawej strony równania różniczkowego w obszarze). Uwzględniając symetrię można zapisać

$$f_0 = f_1 = f_2 = f_3 = f_4 = f_5 = f_9 = f_{10} = f_{11} = f_{12} = f_{13} = f_{14} = 0$$
$$f_8 = f_6$$

Liczba niewiadomych została więc zredukowana do dwóch ( $f_6, f_7 = ?$ ).

Kolejnym krokiem analizy numerycznej problemu brzegowego metodą **MRS** jest zastąpienie w węzłach obszaru operatora różniczkowego odpowiednim operatorem różnicowym. Operator różnicowy *Laplace'a* można wygenerować dowolną metodą do budowania schematów różnicowych (np. *metodą współczynników nieoznaczonych* omawianą w **rozdziale**

II). Można również, korzystając z jego prostoty, stworzyć go za pomocą kompozycji odpowiednich składowych operatorów go tworzących. Ostateczne rozwiązanie



to następujący wzór różnicowy

$$\nabla f_{(i,k)} = \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) f_{(i,k)} \approx \frac{1}{h^2} (f_{(i-1,k)} + f_{(i+1,k)} + f_{(i,k-1)} + f_{(i,k+1)} - 4f_{(i,k)}).$$

Po raz kolejny stosujemy technikę kolokacji do generacji układu równań różnicowych. Przykładamy operator *Laplace'a* w węzłach z niewiadomymi wartościami funkcji – (6) i (7).

$$\begin{cases} \frac{1}{1^2} \left( \overset{0}{f_1} + f_7 + \overset{0}{f_{11}} + \overset{0}{f_5} - 4f_6 \right) = 1 \\ \frac{1}{1^2} \left( \overset{0}{f_2} + \overset{f_6}{f_8} + \overset{0}{f_{12}} + f_6 - 4f_7 \right) = 1 \end{cases} \Rightarrow \begin{cases} f_7 - 4f_6 = 1 \\ 2f_6 - 4f_7 = 1 \end{cases} \Rightarrow \begin{cases} f_6 = -\frac{5}{14} \\ f_7 = -\frac{6}{14} \end{cases}.$$