

X. RÓWNANIA RÓŻNICZKOWE ZWYCZAJNE

10.1. Wprowadzenie

Rozważmy układ równań różniczkowych

$$y'_i = f_i(t, y), \quad t_0 = a \leq t \leq b, \quad (10.1)$$

z warunkami początkowymi

$$y_i(t_0) = y_{i0}, \quad i = 1, 2, \dots, n. \quad (10.2)$$

Zagadnienie (10.1) – (10.2) nazywa się *zagadnieniem początkowym*. Naszym zadaniem jest znalezienie funkcji $y_i(t)$ będących rozwiązaniem ww. układu. W dalszym ciągu będziemy pomijać indeks i .

Idea numerycznego rozwiązywania zagadnienia początkowego jest następująca: wychodząc od znanego warunku początkowego y_0 , „przebiegamy” przedział $[a, b]$ obliczając przybliżone wartości rozwiązania dla niektórych wartości zmiennej niezależnej t , aż zostanie osiągnięty koniec przedziału.

Niech t_k ($k = 1, 2, \dots$) oznaczają pewne punkty przedziału $[a, b]$, przy czym $t_k < t_{k+1}$ oraz $t_0 = a$. Niech każdemu punktowi t_k odpowiada liczba y_k będąca przybliżeniem wartości dokładnego rozwiązania $y(t_k)$.

Rozwiązanie dokładne jest nieznane, ale spodziewamy się, że przybliżenia y_1, y_2, \dots są na tyle dobre, że odtwarzają one w sposób wierny kształt funkcji $y(t)$. Nasuwają się zatem co najmniej trzy pytania:

- w jaki sposób należy określać punkty t_1, t_2, \dots ,
- w jaki sposób należy obliczać przybliżenia y_1, y_2, \dots ,
- jak można stwierdzić, że wartość y_k jest dobrym przybliżeniem dokładnej wartości $y(t_k)$, skoro jest ona nieznana?

Numeryczne rozwiązywanie zagadnienia początkowego określamy inaczej jako całkowanie zagadnienia początkowego. *Długość kroku całkowania* jest określona następująco:

$$h_k = t_k - t_{k-1}, \quad k = 1, 2, \dots$$

Jeżeli dla wszystkich punktów t_1, t_2, \dots mamy

$$t_k - t_{k-1} = h = \text{const}, \quad k = 1, 2, \dots,$$

to mówimy, że obliczenia są wykonywane ze *stałą* długością kroku całkowania. W przeciwnym przypadku mówimy, że obliczenia są wykonywane ze *zmienną* długością tego kroku.

Cechą charakterystyczną całkowania numerycznego zagadnienia początkowego jest to, że przybliżenia y_1, y_2, \dots są obliczane kolejno dla kolejno następujących po sobie punktów t_1, t_2, \dots . Niech y_1, y_2, \dots, y_{k-1} oznaczają obliczone już przybliżenia dla punktów kolejno t_1, t_2, \dots, t_{k-1} . Obliczenie przybliżenia y_k dla następnego punktu t_k nazywamy *jednym krokiem obliczeń*. Metoda numeryczna jest to wyrażony za pomocą wzoru sposób obliczenia wartości przybliżenia rozwiązania w jednym kroku obliczeń.

Metody numeryczne rozwiązywania zagadnienia początkowego dzielimy na:

- metody jednokrokowe (w celu wykonania jednego kroku obliczeń wykorzystujemy tylko przybliżenie obliczone w bezpośrednio poprzedzającym kroku), wśród których wyróżniamy metody wynikające z bezpośredniego zastosowania rozwinięcia Taylora, metody Rungego-Kutty i metody specjalne,
- metody wielokrokowe (w celu wykonania jednego kroku obliczeń wykorzystujemy przybliżenia obliczone w kilku kolejnych, bezpośrednio poprzedzających, krokach), wśród których wyróżniamy liniowe metody wielokrokowe, nieliniowe metody wielokrokowe (metody specjalne) i metody predyktor-korektor,
- metody ekstrapolacyjne.

Definicja 10.1. Różnicę

$$r_k(h) = y(t_{k-1} + h) - y_k$$

nazywamy *błędem aproksymacji* metody w kroku $t_k = t_{k-1} + h$.

Mówimy, że metoda numeryczna jest *rzędu p* , jeżeli dla każdego zagadnienia początkowego mamy

$$r_k(0) = 0, r'_k(0) = 0, \dots, r_k^{(p)}(0) = 0 \text{ oraz } r_k^{(p+1)}(0) \neq 0.$$

Jeżeli metoda jest rzędu p , to błąd aproksymacji posiada rozwinięcie

$$r_k(h) = h^{p+1} \frac{r_k^{(p+1)}(0)}{(p+1)!} + \dots$$

Pierwszy wyraz tego rozwinięcia nazywa się *częścią główną błędu aproksymacji*.

Definicja 10.2. Błąd całkowity odpowiadający punktowi $t_k = t_k(h) = a + kh$ jest określony wzorem

$$e_k = e_k(h) = y(t_k) - y_k(h).$$

Definicja 10.3. Metoda numeryczna jest *zbieżna*, jeśli dla dowolnego $\varepsilon > 0$ istnieje stała $H = H(\varepsilon) > 0$, że jeśli wartość $h < H$, to dla wszystkich punktów $t_k = t_k(h) = a + kh$ jednocześnie zachodzi

$$|e_k(h)| = |y(t_k) - y_k(h)| < \varepsilon, \quad k = 1, 2, \dots, N = N(h).$$

Sens tej definicji jest następujący: jeśli dla $t \in [a, b]$ wykres rozwiązania dokładnego $y(t)$ otoczmy paskiem $[y(t) - \varepsilon, y(t) + \varepsilon]$, gdzie ε oznacza dowolnie małą liczbę dodatnią, to przybliże-

nia uzyskane dla dostatecznie gęstych podziałów odcinka $[a, b]$ zawierają się w tym epsilonowym otoczeniu.

Definicja 10.3 może być też sformułowana równoważnie inaczej.

Definicja 10.4. *Metoda numeryczna jest zbieżna, jeżeli dla każdego ciągu h_m ($m = 1, 2, \dots$) spełniającego warunki $h_1 > h_2 > \dots$ oraz $h_m \rightarrow 0$ przy $m \rightarrow \infty$, mamy*

$$\max_{1 \leq k \leq N(m)} |e_k^{(m)}| \rightarrow 0, \text{ gdy } m \rightarrow \infty,$$

gdzie $e_k^{(m)} = y(t_k^{(m)}) - y_k(h_m)$.

10.2. Bezpośrednie zastosowanie rozwinięcia Taylora

Rozważmy skalarne zagadnienie początkowe postaci

$$y' = f(y), \quad y(a) = y_0. \quad (10.3)$$

Niech $t \in [a, b]$. Zakładając istnienie pochodnych rozwiązania $y = y(t)$ do rzędu $k + 1$, ze wzoru Taylora otrzymujemy

$$y(t+h) = y(t) + hy'(t) + \frac{h^2}{2!} y''(t) + \dots + \frac{h^k}{k!} y^{(k)}(t) + R(t, h), \quad (10.4)$$

gdzie $R(t, h) = O(h^{k+1})$ oznacza resztę. Korzystając z równania (10.3) oraz zakładając istnienie pochodnych funkcji f dostatecznie wysokiego rzędu, można wyrazić pochodne $y^{(j)}(t)$ za pomocą funkcji $y(t)$ oraz pochodnych funkcji f :

$$y'(t) = f(y(t)),$$

$$y''(t) = \frac{df(y(t))}{dy} f(y(t)),$$

$$y'''(t) = \frac{d^2 f(y(t))}{dy^2} f^2(y(t)) + \left[\frac{df(y(t))}{dy} \right]^2 f(y(t))$$

itd..

Uwaga! Jeśli równania (10.3) przedstawiają układ n równań, to wyrażenia $\frac{df}{dy}$, $\frac{d^2 f}{dy^2} f$ itd.. oznaczają macierze kwadratowe stopnia n .

Jeśli w równaniu (10.4) uwzględnimy tylko pochodną rzędu pierwszego, to otrzymamy

$$y(t+h) = y(t) + hf(y(t)) + O(h^2),$$

skąd po zaniedbaniu reszty mamy metodę postaci

$$y_{k+1} = y_k + hf(y_k), \quad k = 0, 1, \dots \quad (10.5)$$

Jest to *metoda Eulera*.

Korzystając z wyrażeń dla pochodnych pierwszego, drugiego i trzeciego rzędu, z równania (10.4) mamy

$$y(t+h) = y(t) + hf(y(t)) + \frac{h^2}{2!} \frac{df(y(t))}{dy} f(y(t)) + \frac{h^3}{3!} \left\{ \frac{d^2 f(y(t))}{dy^2} f(y(t)) + \left[\frac{df(y(t))}{dy} \right]^2 \right\} f(y(t)) + O(h^4),$$

skąd otrzymujemy metodę postaci

$$y_{k+1} = y_k + hf(y_k) + \frac{h^2}{2!} \frac{df(y_k)}{dy} f(y_k) + \frac{h^3}{3!} \left\{ \frac{d^2 f(y_k)}{dy^2} f(y_k) + \left[\frac{df(y_k)}{dy} \right]^2 \right\} f(y_k).$$

Obie ww. metody są jednokrokowe. Oprócz metody Eulera (10.5), wszystkie metody otrzymywane w ten sposób wymagają niewygodnego w praktyce i pracochłonnego obliczania pochodnych funkcji f , rzędu tym wyższego, im więcej wyrazów rozwinięcia Taylora uwzględnimy przy konstrukcji metody. Wyjątkiem są tu metody, w których pochodne te można obliczyć w sposób rekurencyjny.

10.3. Metody wielokrokowe

Rozważmy zagadnienie początkowe postaci

$$y'(t) = f(t, y), \quad y(a) = y_0. \quad (10.6)$$

Niech $\{t_k\}$ oznacza ciąg punktów, takich że $t_k < t_{k+1}$, $k = 1, 2, \dots$. Załóżmy, że $t_k = a + kh$, gdzie $h > 0$. Oznacza to, że punkty t_k są równoodległe, czyli że obliczenia są wykonywane ze stałym krokiem całkowania.

Dla ustalonego k ($k = 1, 2, \dots$) Zastosowanie metody k -krokowej wymaga znajomości wielkości y_0, y_1, \dots, y_{k-1} , będących przybliżeniami wartości dokładnych $y(t_0), y(t_1), \dots, y(t_{k-1})$, gdzie $t_0 = a, y_0 = y(t_0) = y(a)$.

Przyjmijmy nieco ogólniej, że dla pewnego $n \geq k$ są znane wielkości $y_{n-k}, y_{n-k+1}, \dots, y_{n-1}$ będące przybliżeniami wartości dokładnych $y(t_{n-k}), y(t_{n-k+1}), \dots, y(t_{n-1})$. Algorytm metody k -krokowej jest określony następująco:

$$y_n = \sum_{j=1}^k \alpha_j y_{n-j} + h \sum_{j=0}^k \beta_j f_{n-j}, \quad n = k, k+1, \dots, \quad (10.7)$$

gdzie $f_{n-j} = f(t_{n-j}, y_{n-j})$, a współczynniki α_j i β_j są odpowiednio dobranymi liczbami.

Definicja 10.5. Metoda k -krokowa jest jawna, gdy $\beta_0 = 0$. Metoda k -krokowa jest niejawna, gdy $\beta_0 \neq 0$.

Algorytm (10.7) dla metody jawnej ma postać

$$y_n = \sum_{j=1}^k \alpha_j y_{n-j} + h \sum_{j=1}^k \beta_j f_{n-j}, \quad (10.8)$$

a dla metody niejawnej –

$$y_n = \sum_{j=1}^k \alpha_j y_{n-j} + h \sum_{j=1}^k \beta_j f_{n-j} + h \beta_0 f(t_n, y_n). \quad (10.9)$$

W przypadku metody (10.9) w każdym kroku całkowania musimy rozwiązać na ogół nieliniowe równanie (lub układ równań nieliniowych). Okazuje się, że pomimo większej pracochłonności, zastosowanie metody niejawnej jest celowe.

Wzór (10.7) określa metodę k -krokową w sposób czysto formalny. Zajmiemy się teraz wyprowadzeniami metod tej postaci.

Przepiszmy zagadnienie początkowe (10.6) w równoważnej postaci całkowej

$$y(t) = y(t_{n-1}) + \int_{t_{n-1}}^t f(x, y(x)) dx, \quad t > t_{n-1}.$$

Stąd

$$y(t_n) = y(t_{n-1}) + \int_{t_{n-1}}^{t_n} f(x, y(x)) dx. \quad (10.10)$$

Przybliżając funkcję podcałkową za pomocą odpowiedniego wielomianu interpolacyjnego, a następnie całkując otrzymany wzór otrzymamy metodę postaci (10.7).

Niech $W(x)$ oznacza wielomian interpolacyjny stopnia $k-1$, taki że

$$W(t_{n-j}) = f(t_{n-j}, y(t_{n-j})), \quad j = 1, 2, \dots, k.$$

Oznaczając

$$f(t_{n-j}) = f(t_{n-j}, y(t_{n-j})), \quad j = 1, 2, \dots, k,$$

oraz dokonując zamiany zmiennych $x = t_{n-1} + th$, wielomian $W(x)$ zapisujemy w postaci Newtona za pomocą różnic wstecznych:

$$W(t_{n-1} + th) = f(t_{n-1}) + t \nabla f(t_{n-1}) + \dots + \frac{t(t+1) \dots (t+k-2)}{(k-1)!} \nabla^{k-1} f(t_{n-1}).$$

Funkcję podcałkową we wzorze (10.10) przybliżamy wielomianem $W(x)$, tj.

$$f(x, y(x)) = W(x) + r(x),$$

gdzie $r(x)$ oznacza błąd interpolacji dany wzorem

$$r(x) = r(t_{n-1} + th) = f^{(k)}(\xi(t)) h^k \frac{t(t+1) \dots (t+k-1)}{k!}.$$

Całkowanie wielomianu interpolacyjnego nie nastęrcza trudności, a wynik całkowania błędu interpolacji uzyskujemy na podstawie twierdzenia o wartości średniej dla całek, wykorzystując fakt, że funkcja $t(t+1) \dots (t+k-1)$ ma stały znak w przedziale $(0, 1)$.

Jeżeli oznaczymy

$$\gamma_0 = 1, \quad \gamma_j = \frac{1}{j!} \int_0^1 t(t+1) \dots (t+j-1) dt, \quad j \geq 1,$$

to otrzymamy

$$y(t_n) = y(t_{n-1}) + h \sum_{j=0}^{k-1} \gamma_j \nabla^j f(t_{n-1}) + h^{k+1} \gamma_k y^{(k+1)}(\xi_n). \quad (10.11)$$

Zależność (10.11) jest dokładna i jest ona równoważna z zależnością (10.10), która z kolei jest rozwiązaniem zagadnienia początkowego na przedziale $[t_{n-1}, t_n]$.

Wykorzystując równość

$$\nabla^j f(t_{n-1}) = \sum_{m=0}^j (-1)^m \binom{j}{m} f(t_{n-1-m}),$$

otrzymujemy wzór (10.11) w innej postaci:

$$y(t_n) = y(t_{n-1}) + h \sum_{j=1}^k \beta_{kj} f(t_{n-j}) + h^{k+1} \gamma_k y^{(k+1)}(\xi_n), \quad (10.12)$$

gdzie

$$\beta_{kj} = (-1)^{j-1} \sum_{m=j-1}^{k-1} \binom{m}{j-1} \gamma_m, \quad j = 1, 2, \dots, k.$$

Zaniedbując błąd interpolacji oraz przyjmując zamiast dokładnych wartości $y(t_{n-j})$ i $f(t_{n-j}) = f(t_{n-j}, y(t_{n-j}))$ wartości przybliżone odpowiednio y_{n-j} i $f_{n-j} = f(t_{n-j}, y_{n-j})$, otrzymujemy algorytm metody Adamsa-Bashfortha postaci

$$y_n = y_{n-1} + h \sum_{j=0}^{k-1} \gamma_j \nabla^j f_{n-1}$$

lub

$$y_n = y_{n-1} + h \sum_{j=1}^k \beta_{kj} f_{n-j}. \quad (10.13)$$

Zauważmy, że wzór (10.13) ma postać wzoru (10.8) i określa metodę jawną.

Wartości współczynników γ_j i β_{kj} są następujące:

j	0	1	2	3	4	5
γ_j	1	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	$\frac{251}{720}$	$\frac{95}{288}$

k	β_{k1}	β_{k2}	β_{k3}	β_{k4}	β_{k5}
1	1				
2	$\frac{3}{2}$	$-\frac{1}{2}$			
3	$\frac{23}{12}$	$-\frac{16}{12}$	$\frac{5}{12}$		
4	$\frac{55}{24}$	$-\frac{59}{24}$	$\frac{37}{24}$	$-\frac{9}{24}$	
5	$\frac{1901}{720}$	$-\frac{2774}{720}$	$\frac{2616}{720}$	$-\frac{1274}{720}$	$\frac{251}{720}$

W szczególności z zależności (10.13) otrzymujemy następujące wzory Adamsa-Bashfortha:

- $k = 1$ (metoda Eulera – por. (10.5))

$$y_n = y_{n-1} + hf_{n-1},$$

- $k = 2$

$$y_n = y_{n-1} + \frac{h}{2}(3f_{n-1} - f_{n-2}),$$

- $k = 3$

$$y_n = y_{n-1} + \frac{h}{12}(23f_{n-1} - 16f_{n-2} + 5f_{n-3}),$$

- $k = 4$

$$y_n = y_{n-1} + \frac{h}{24}(55f_{n-1} - 59f_{n-2} + 37f_{n-3} - 9f_{n-4}).$$

Niech teraz $\bar{W}(x)$ oznacza wielomian interpolacyjny stopnia k , taki że

$$\bar{W}(t_{n-j}) = f(t_{n-j}, y(t_{n-j})), \quad j = 0, 1, \dots, k.$$

Postępując podobnie jak poprzednio otrzymamy

$$y(t_n) = y(t_{n-1}) + h \sum_{j=0}^k \bar{\gamma}_j \nabla^j f(t_n) + h^{k+2} \bar{\gamma}_{k+1} y^{(k+2)}(\xi_n), \quad (10.14)$$

gdzie

$$\bar{\gamma}_0 = 1, \quad \bar{\gamma}_j = \frac{1}{j!} \int_{-1}^0 t(t+1) \dots (t+j-1) dt, \quad j \geq 1.$$

Zależność (10.14) jest dokładna i równoważna zależności (10.10), a więc także zagadnieniu początkowemu (10.6) na przedziale $[t_{n-1}, t_n]$. Uwzględniając, że

$$\nabla^j f(t_n) = \sum_{m=0}^j (-1)^m \binom{j}{m} f(t_{n-m}),$$

z wzoru (10.14) otrzymamy równość dokładną

$$y(t_n) = y(t_{n-1}) + h \sum_{j=0}^k \bar{\beta}_{kj} f(t_{n-j}) + h^{k+2} \bar{\gamma}_{k+1} y^{(k+2)}(\xi_n), \quad (10.15)$$

gdzie

$$\bar{\beta}_{kj} = (-1)^j \sum_{m=j}^k \binom{m}{j} \bar{\gamma}_m, \quad j = 0, 1, \dots, k.$$

Zastępując we wzorach (10.14) i (10.15) dokładne wartości przybliżonymi i pomijając resztę otrzymujemy metody Adamsa-Moultona postaci

$$y_n = y_{n-1} + h \sum_{j=0}^k \bar{\gamma}_j \nabla^j f_n$$

lub

$$y_n = y_{n-1} + h \sum_{j=0}^k \bar{\beta}_{kj} f_{n-j}, \quad (10.16)$$

gdzie $f_{n-j} = f(t_{n-j}, y_{n-j})$. Zauważmy, że wzór (10.16) ma postać wzoru (10.9) i określa metodę niejawną.

Wartości współczynników $\bar{\gamma}_j$ i $\bar{\beta}_{kj}$ podano w poniższych tabelach.

j	0	1	2	3	4
$\bar{\gamma}_j$	1	$-\frac{1}{2}$	$-\frac{1}{12}$	$-\frac{1}{24}$	$-\frac{19}{720}$

k	$\bar{\beta}_{k0}$	$\bar{\beta}_{k1}$	$\bar{\beta}_{k2}$	$\bar{\beta}_{k3}$	$\bar{\beta}_{k4}$
1	$\frac{1}{2}$	$\frac{1}{2}$			
2	$\frac{5}{12}$	$\frac{8}{12}$	$-\frac{1}{12}$		
3	$\frac{9}{24}$	$\frac{19}{24}$	$-\frac{5}{24}$	$\frac{1}{24}$	
4	$\frac{251}{720}$	$\frac{646}{720}$	$-\frac{264}{720}$	$\frac{106}{720}$	$-\frac{19}{720}$

W szczególności z zależności (10.16) otrzymujemy następujące wzory Adamsa-Moultona:

- $k = 1$ (wzór trapezów)

$$y_n = y_{n-1} + \frac{h}{2}(f_n + f_{n-1}),$$

- $k = 2$

$$y_n = y_{n-1} + \frac{h}{12}(5f_n + 8f_{n-1} - f_{n-2}),$$

- $k = 3$

$$y_n = y_{n-1} + \frac{h}{24}(9f_n + 19f_{n-1} - 5f_{n-2} + f_{n-3}).$$

Gdyby zastosować wielomian interpolacyjny stopnia zerowego, tj. $W(x) = f(t_n, y(t_n))$, to otrzymamy jeszcze jedną metodę Adamsa-Moultona postaci

$$y_n = y_{n-1} + hf_n.$$

Jest to tzw. *wsteczna metoda Eulera*.

Przepiszmy teraz zagadnienie początkowe (10.6) w następującej równoważnej postaci:

$$y(t) = y(t_{n-2}) + \int_{t_{n-2}}^t f(x, y(x)) dx, \quad t > t_{n-2},$$

skąd otrzymamy

$$y(t_n) = y(t_{n-2}) + \int_{t_{n-2}}^{t_n} f(x, y(x)) dx. \quad (10.17)$$

Niech $W(x)$ oznacza wielomian interpolacyjny stopnia $k-1$, jak w metodzie Adamsa-Bashfortha. Funkcję podcałkową we wzorze (10.17) przybliżamy tym wielomianem i wykonujemy całkowanie. Otrzymamy

$$y(t_n) = y(t_{n-2}) + h \sum_{j=0}^{k-1} v_j \nabla^j f(t_{n-1}) + \dots,$$

gdzie

$$v_0 = 2, \quad v_j = \frac{1}{j!} \int_{-1}^1 t(t+1) \dots (t+j-1) dt, \quad j \geq 1,$$

Pomijając błąd przybliżenia i zastępując wartości dokładne ich przybliżeniami otrzymujemy jawne k -krokowe metody Nyströma postaci

$$y_n = y_{n-2} + h \sum_{j=0}^{k-1} v_j \nabla^j f_{n-1}. \quad (10.18)$$

W szczególności dla $k = 1$ z wzoru (10.18) mamy

$$y_n = y_{n-2} + 2hf_{n-1}.$$

Jest to tzw. *metoda punktu środkowego*.

Jeżeli weźmiemy wielomian $\bar{W}(x)$ stopnia k taki, jak w metodzie Adamsa-Moultona i funkcję podcałkową w zależności (10.17) przybliżymy tym wielomianem, to po wykonaniu całkowania dostaniemy

$$y(t_n) = y(t_{n-2}) + h \sum_{j=0}^k \bar{v}_j \nabla^j f(t_n) + \dots,$$

gdzie

$$\bar{v}_0 = 2, \quad \bar{v}_j = \frac{1}{j!} \int_{-2}^0 t(t+1) \dots (t+j-1) dt, \quad j \geq 1.$$

Stąd, postępując podobnie jak poprzednio, otrzymujemy niejawne metody Milne'a-Simpsona postaci

$$y_n = y_{n-2} + h \sum_{j=0}^k \bar{v}_j \nabla^j f_n.$$

W szczególności dla $k = 2$ mamy dwukrokową metodę Milne'a:

$$y_n = y_{n-2} + \frac{h}{3} (f_n + 4f_{n-1} + f_{n-2}).$$

Do metod wielokrokowych zalicza się także metody wstecznego różniczkowania, które polegają na przybliżeniu rozwiązania zagadnienia początkowego za pomocą odpowiedniego wielomianu interpolacyjnego, a następnie na przybliżeniu pochodnej rozwiązania za pomocą pochodnej tego wielomianu. Wśród metod wstecznego różniczkowania mamy też metody jawne i niejawne.

Okazuje się, że przy tej samej liczbie kroków niejawne metody wielokrokowe mają wyższy rząd zbieżności niż metody jawne. Połączenie metod jawnych i niejawnych prowadzi do metod predyktor-korektor, w których predyktorem jest metoda jawna, a korektorem metoda niejawna. Za pomocą predyktora uzyskuje się początkowe przybliżenie rozwiązania w określonym punkcie, które następnie jest wykorzystywane w procesie iteracyjnym dla metody niejawnej.

10.4. Metody Rungego-Kutty

Jak pamiętamy, zależność całkowa

$$y(t) = y(t_n) + \int_{t_n}^t f(\tau, y(\tau))d\tau, \quad t > t_n, \quad (10.19)$$

jest równoważna z równaniem różniczkowym dla $t > t_n$. Przyjmując $t_{n+1} = t_n + h$ zamiast t , z równania (10.19) otrzymujemy

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t))dt. \quad (10.20)$$

W metodach Rungego-Kutty dokonujemy przybliżenia całki występującej w powyższym wzorze za pomocą sumy o takiej postaci, że każdy składnik tej sumy wyraża się przez $y(t_n)$. W ten sposób wartość $y(t_{n+1})$, stojąca po lewej stronie, będzie wyrażała się jedynie poprzez wartość $y(t_n)$.

Dokonując w zależności (10.20) zamiany zmiennych

$$c = \frac{t - t_n}{t_{n+1} - t_n},$$

dostajemy

$$y(t_{n+1}) = y(t_n) + h \int_0^1 f(t_n + ch, y(t_n + ch))dc.$$

Istnieje wiele sposobów obliczenia całki występującej w powyższym wzorze. Zastosowanie ich polega na zastąpieniu całki odpowiednią sumą:

$$y(t_{n+1}) = y(t_n) + h \sum_{i=1}^m w_i f(t_n + c_i h, y(t_n + c_i h)) + E_m(h), \quad (10.21)$$

gdzie w_i oraz c_i oznaczają dla ustalonego $m \geq 1$ współczynniki zależne od przyjętego sposobu przybliżenia całki, a $E_m(h)$ oznacza błąd przybliżenia. Zależność (10.21) jest tylko formalna i nie przedstawia poszukiwanego wyniku, gdyż wartości $y(t_n + c_i h)$ są nieznane. Wartości te zastępuje się pierwszymi dwoma wyrazami odpowiedniego rozwinięcia w szereg Taylora.

Oznaczając

$$k_1(h) = f(t_n, y(t_n)),$$

$$k_i(h) = f\left(t_n + c_i h, y(t_n) + h \sum_{j=1}^{i-1} a_{ij} k_j(h)\right), \quad i > 1,$$

zależność (10.21) możemy napisać w postaci

$$y(t_{n+1}) = y(t_n) + h \sum_{i=1}^m w_i k_i(h) + R_m(h),$$

gdzie $R_m(h)$ oznacza błąd $E_m(h)$ wynikający z zastąpienia całki sumą i powiększony o błędy wynikające z zastąpienia wartości $y(t_n + c_i h)$ pierwszymi dwoma wyrazami rozwinięcia w szereg Taylora. Z rozwinięcia $y(t_n + c_i h)$ względem poprzedniego punktu $t_n + c_{i-1} h$ wynikają przy tym zależności

$$a_{i1} = c_2, \quad a_{i2} = c_3 - c_2, \quad \dots, \quad a_{i,i-1} = c_i - c_{i-1},$$

skąd

$$c_i = \sum_{j=1}^{i-1} a_{ij} \quad \text{dla } i > 1 \text{ i } c_1 = 0.$$

Zastępując wartość dokładną $y(t_n)$ przybliżoną wartością y_n i pomijając błąd $R_m(h)$ otrzymujemy

$$y_{n+1} = y_n + h \sum_{i=1}^m w_i k_i, \quad (10.22)$$

gdzie

$$\begin{aligned} k_1 &= f(t_n, y_n), \\ k_i &= f\left(t_n + c_i h, y_n + h \sum_{j=1}^{i-1} a_{ij} k_j\right), \quad i > 1, \end{aligned} \quad (10.23)$$

przy czym

$$c_i = \sum_{j=1}^{i-1} a_{ij}, \quad i > 1. \quad (10.24)$$

Powyższe wzory zostały wyprowadzone przy założeniu, że zastosowano określony sposób przybliżenia całki. Wzory (10.22) – (10.24) przyjmuje się często za podstawę definicji metod Rungego-Kutty bez odwoływania się do przybliżonego całkowania. O współczynnikach w_i , c_i i a_{ij} zakłada się, że są niezależnymi parametrami spełniającymi warunek (10.24). Współczynniki te wygodnie jest przedstawiać w postaci poniższej tablicy.

0					
c_2	a_{21}				
c_3	a_{31}	a_{32}			
\vdots	\vdots	\vdots			
c_m	a_{m1}	a_{m2}	\dots	$a_{m, m-1}$	
	w_1	w_2	\dots	w_{m-1}	w_m

Definicja 10.6. *Błędem aproksymacji* odpowiadającym punktowi t_{n+1} nazywamy wyrażenie

$$r_{n+1}(h) = y(t_n + h) - \left(y(t_n) + h \sum_{i=1}^m w_i k_i(h) \right). \quad (10.25)$$

Definicja 10.7. Mówimy, że metoda Rungego-Kutty jest rzędu p , jeśli dla każdego zagadnienia początkowego mamy

$$r_{n+1}(0) = 0, \quad r_{n+1}^{(i)}(0) = 0 \quad \text{dla } i = 1, 2, \dots, p \quad \text{oraz } r_{n+1}^{(p+1)}(0) \neq 0,$$

gdzie $r_{n+1}(h)$ jest określone wzorem (10.25).

Okazuje się, że w ogólności równań wiążących współczynniki w_i , c_i i a_{ij} , które gwarantują, że metoda jest rzędu p jest mniej niż liczba tych współczynników. Stąd dla danej liczby m we wzorze (10.22), którą nazywa się *liczbą etapów* metody Rungego-Kutty otrzymuje się różne rodziny metod, w których niektóre ze wspomnianych współczynników przyjmuje się za parametry. Na przykład, dla $m = 2$ i rzędu $p = 2$ (nie istnieją dwuetapowe metody wyższego rzędu) mamy warunki

$$w_1 + w_2 = 1, \quad w_2 c_2 = \frac{1}{2}.$$

Z drugiego równania wynika, że przypadki $w_2 = 0$ oraz $c_2 = 0$ są wykluczone. Jeśli wartość $c_2 \neq 0$ obierzemy jako parametr, to

$$w_1 = \frac{2c_2 - 1}{2c_2}, \quad w_2 = \frac{1}{2c_2}$$

i stwierdzamy, że istnieje nieskończenie wiele metod dwuetapowych rzędu drugiego. Przykładami takich metod mogą być:

- *ulepszona metoda Eulera*

$$\begin{aligned} k_1 &= f(t_n, y_n), \\ k_2 &= f\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_1\right), \\ y_{n+1} &= y_n + hk_2, \end{aligned}$$

- *metoda Eulera-Cauchy'ego*

$$\begin{aligned} k_1 &= f(t_n, y_n), \\ k_2 &= f(t_n + h, y_n + hk_1), \\ y_{n+1} &= y_n + \frac{1}{2}h(k_1 + k_2). \end{aligned}$$

W przypadku $m = 4$ i $p = 4$ (nie istnieją czteroetapowe metody wyższego rzędu) mamy następujących osiem równań wiążących współczynniki:

$$\begin{aligned}
w_1 + w_2 + w_3 + w_4 &= 1, \\
w_2c_2 + w_3c_3 + w_4c_4 &= \frac{1}{2}, \\
w_2c_2^2 + w_3c_3^2 + w_4c_4^2 &= \frac{1}{3}, \\
w_3a_{32}c_2 + w_4a_{42}c_2 + w_4a_{43}c_3 &= \frac{1}{6}, \\
w_2c_2^3 + w_3c_3^3 + w_4c_4^3 &= \frac{1}{4}, \\
w_3c_3a_{32}c_2 + w_4c_4a_{42}c_2 + w_4c_4a_{43}c_3 &= \frac{1}{8}, \\
w_3a_{32}c_2^2 + w_4a_{42}c_2^2 + w_4a_{43}c_3^2 &= \frac{1}{12}, \\
w_4a_{43}a_{32}c_2 &= \frac{1}{24}.
\end{aligned}$$

Z równań tych wynika, że możemy otrzymać dwu- i jednoparametrowe rodziny czteroetapowych metod Rungego-Kutty. W szczególnym przypadku dla następującej tablicy parametrów:

0				
$\frac{1}{2}$	$\frac{1}{2}$			
$\frac{1}{2}$	0	$\frac{1}{2}$		
1	0	0	1	
	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

otrzymujemy jedną z najbardziej popularnych metod Rungego-Kutty rzędu czwartego, zwaną po prostu *metodą Rungego-Kutty* postaci

$$\begin{aligned}
k_1 &= f(t_n, y_n), \\
k_2 &= f\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_1\right), \\
k_3 &= f\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_2\right), \\
k_4 &= f(t_n + h, y_n + hk_3),
\end{aligned}$$

$$y_{n+1} = y_n + \frac{1}{6}h(k_1 + 2k_2 + 2k_3 + k_4).$$

Wykazano, że dla $m = 1, 2, 3, 4$ istnieje nieskończenie wiele metod m -etapowych oraz że dla metody m -etapowej maksymalny rząd jest równy $p = m$. Dla $m = 5, 6, 7$ maksymalny rząd jest równy $m - 1$, dla $m = 8, 9$ mamy $p = m - 2$, a dla $m \geq 10$ dla maksymalnego rzędu zachodzi nierówność $p \leq m - 2$.

XI. RÓWNANIA RÓŻNICZKOWE CZĄSTKOWE

W równaniach różniczkowych cząstkowych występują pochodne cząstkowe (różnych rzędów) funkcji, którą należy znaleźć. Oprócz tego dla szukanej funkcji są określone różne warunki brzegowe lub początkowo-brzegowe. Istnieje wiele metod numerycznego rozwiązywania takich równań. Tutaj ograniczymy się do równania rzędu drugiego i metody różnicowej jego rozwiązania.

Niech Ω oznacza pewien obszar na płaszczyźnie xy , a Γ – brzeg tego obszaru. Załóżmy, że w obszarze Ω jest zadane równanie różniczkowe cząstkowe drugiego rzędu postaci

$$a \frac{\partial^2 u}{\partial x^2} + 2b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2} + 2d \frac{\partial u}{\partial x} + 2e \frac{\partial u}{\partial y} + gu = f, \quad (11.1)$$

gdzie $a = a(x, y)$, $b = b(x, y)$, $c = c(x, y)$, $d = d(x, y)$, $e = e(x, y)$, $f = f(x, y)$ i $g = g(x, y)$ oznaczają dane funkcje określone w obszarze $\Omega + \Gamma$.

Jeżeli $b^2 - ac < 0$ dla każdego $x, y \in \Omega$, to równanie (11.1) nazywa się *eliptycznym* w obszarze Ω . Gdy $b^2 - ac = 0$, to równanie nazywamy *parabolicznym*, a gdy $b^2 - ac > 0$ – *hiperbolicznym*.

W ogólności zarówno obszar Ω , jak i jego brzeg Γ mogą być dowolne. W dalszym ciągu założymy, że obszar Ω jest prostokątem postaci

$$\Omega = \{(x, y): 0 \leq x \leq \alpha, 0 \leq y \leq \beta\}$$

o brzegu określonym wzorem

$$\Gamma = \{(x, y): x = 0, \alpha \text{ i } 0 \leq y \leq \beta \text{ lub } 0 \leq x \leq \alpha \text{ i } y = 0, \beta\}.$$

Niech dla szukanej funkcji $u = u(x, y)$, gdzie $0 \leq x \leq \alpha$ oraz $0 \leq y \leq \beta$, będą określone następujące warunki brzegowe:

$$u|_{\Gamma} = \varphi(x, y) = \begin{cases} \varphi_1(y), & \text{dla } x = 0, \\ \varphi_2(x), & \text{dla } y = 0, \\ \varphi_3(y), & \text{dla } x = \alpha, \\ \varphi_4(x), & \text{dla } y = \beta, \end{cases} \quad (11.2)$$

przy czym

$$\varphi_1(0) = \varphi_2(0), \quad \varphi_2(\alpha) = \varphi_3(0), \quad \varphi_3(\beta) = \varphi_4(\alpha), \quad \varphi_4(0) = \varphi_1(\beta).$$

Pochodne cząstkowe występujące w równaniu (11.1) mogą być przybliżone za pomocą ilorazów różnicowych. Mamy

$$\begin{aligned}\frac{\partial^2 u}{\partial x^2} &\approx \frac{u(x-h, y) - 2u(x, y) + u(x+h, y)}{h^2}, \\ \frac{\partial^2 u}{\partial x \partial y} &\approx \frac{u(x-h, y-k) - u(x-h, y+k) + u(x+h, y+k) - u(x+h, y-k)}{4hk}, \\ \frac{\partial^2 u}{\partial y^2} &\approx \frac{u(x, y+k) - 2u(x, y) + u(x, y-k)}{k^2}, \\ \frac{\partial u}{\partial x} &\approx \frac{u(x+h, y) - u(x-h, y)}{2h}, \\ \frac{\partial u}{\partial y} &\approx \frac{u(x, y+k) - u(x, y-k)}{2k},\end{aligned}$$

W celu rozwiązania zagadnienia (11.1) – (11.2) w obszarze Ω definiuje się siatkę:

$$\left\{ (x, y): x = ih, y = jk, \text{ gdzie } i = 0, 1, \dots, n, j = 0, 1, \dots, m, h = \frac{\alpha}{n}, k = \frac{\beta}{m} \right\},$$

na której konstruuje się dla równania (11.1), uwzględniając przybliżenia pochodnych cząstkowych, schemat różnicowy postaci

$$\begin{aligned}a_{ij} \frac{u_{i-1,j} - 2u_{ij} + u_{i+1,j}}{h^2} + b_{ij} \frac{u_{i-1,j-1} - u_{i-1,j+1} + u_{i+1,j+1} - u_{i+1,j-1}}{2hk} \\ + c_{ij} \frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{k^2} + d_{ij} \frac{u_{i+1,j} - u_{i-1,j}}{h} \\ + e_{ij} \frac{u_{i,j+1} - u_{i,j-1}}{k} + g_{ij} u_{ij} = f_{ij}, \\ i = 1, 2, \dots, n-1, \quad j = 1, 2, \dots, m-1,\end{aligned}\tag{11.3}$$

przy czym, zgodnie z warunkami (11.2), mamy

$$\begin{aligned}u_{0j} = \varphi_1(jk), \quad u_{nj} = \varphi_3(jk), \quad j = 0, 1, \dots, m, \\ u_{i0} = \varphi_2(ih), \quad u_{im} = \varphi_4(ih), \quad i = 1, 2, \dots, n-1.\end{aligned}\tag{11.4}$$

Uwzględniając warunki (11.4), z zależności (11.3) otrzymujemy układ $(n-1)(m-1)$ równań liniowych, które w zapisie blokowym można przedstawić następująco:

$$\begin{bmatrix} \mathbf{B}_1 & \mathbf{C}_1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ \mathbf{A}_2 & \mathbf{B}_2 & \mathbf{C}_2 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{A}_3 & \mathbf{B}_3 & \mathbf{C}_3 & \dots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \mathbf{A}_{n-3} & \mathbf{B}_{n-3} & \mathbf{C}_{n-3} & 0 & \mathbf{u}_{n-3} \\ 0 & 0 & 0 & 0 & \dots & 0 & \mathbf{A}_{n-2} & \mathbf{B}_{n-2} & \mathbf{C}_{n-2} & \mathbf{u}_{n-2} \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \mathbf{A}_{n-1} & \mathbf{B}_{n-1} & \mathbf{u}_{n-1} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \\ \vdots \\ \mathbf{u}_{n-3} \\ \mathbf{u}_{n-2} \\ \mathbf{u}_{n-1} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \\ \vdots \\ \mathbf{r}_{n-3} \\ \mathbf{r}_{n-2} \\ \mathbf{r}_{n-1} \end{bmatrix},$$

gdzie

$$\mathbf{u}_i = \begin{bmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{i,m-1} \end{bmatrix}, \quad \mathbf{r}_i = \begin{bmatrix} r_{i1} \\ r_{i2} \\ \vdots \\ r_{i,m-1} \end{bmatrix}, \quad i = 1, 2, \dots, n-1,$$

a \mathbf{A}_i , \mathbf{B}_i oraz \mathbf{C}_i oznaczają macierze trójdiagonalne. Wzory na współczynniki tych macierzy i składowe wektorów \mathbf{r}_i wynikają z rozważanego układu, który można rozwiązać na przykład metodą eliminacji Gaussa z pełnym wyborem elementu podstawowego.

LITERATURA UZUPEŁNIAJĄCA

- [1] J. i M.. Jankowscy, *Przegląd metod i algorytmów numerycznych, Cz. 1 i 2*, WNT.
- [2] J. Stoer, R. Bulirsch, *Wstęp do analizy numerycznej*, PWN.
- [3] Z. Fortuna, B. Macukow, J. Wąsowski, *Metody numeryczne*, WNT.
- [4] A. Ralston, *Wstęp do analizy numerycznej*, PWN.
- [5] A. Marciniak, D. Gregulec, J. Kaczmarek, *Podstawowe procedury numeryczne w języku Turbo Pascal*, NAKOM.
- [6] D. Kincaid, W. Cheney, *Analiza numeryczna*, WNT.