

DANES-PICTA.COM

A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19

AKADEMIA SZTABU GENERALNEGO
im. Generała Broni Karola Świerczewskiego

KATEDRA CYBERNETYKI

42

mgr Józef MARONSKI

ELEMENTY METOD NUMERYCZNYCH



4492

WARSZAWA

STYCZEŃ

1970



Colour Chart #13

DANES-PICTA.COM

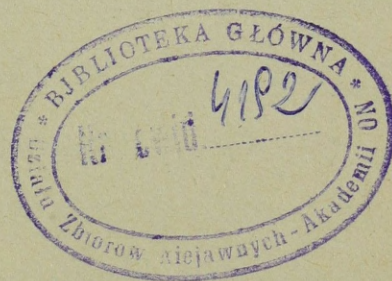
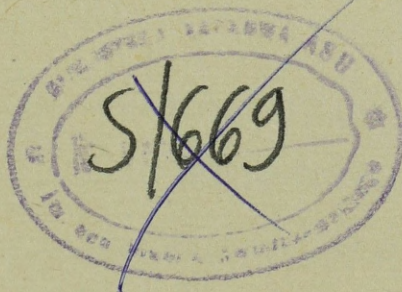
AKADEMIA SZTABU GENERALNEGO
im.gen.broni K.Świerczewskiego

KATEDRA CYBERNETYKI ASG

42

mgr Józef MAROŃSKI

ELEMENTY METOD NUMERYCZNYCH



WARSZAWA

STYCZEŃ

1970 r.

W S T Ę P

W różnorodnych dziedzinach współczesnej nauki i techniki coraz częściej spotykamy się z takimi zagadnieniami matematycznymi, dla których niemożliwe jest znajdowanie rozwiązań metodami klasycznej analizy matematycznej. Bardzo często spotykamy się z koniecznością rozwiązywania układów równań algebraicznych liniowych z dziesiątkami, a nawet setkami niewiadomych, z rozwiązywaniem równań różniczkowych nie całkowalnych w dziedzinie funkcji elementarnych, z obliczaniem pierwiastków równań algebraicznych wysokich stopni itp. Bardzo szybki wzrost ilości takich zadań z jednej strony, rozwój nowoczesnych środków obliczeniowych jakimi są maszyny cyfrowe, z drugiej, spowodowały gwałtowny rozwój pewnej dziedziny matematyki zajmującej się sprawnym operowaniem wielkościami liczbowymi, do których jak wiadomo daje sprowadzić się większość wspomnianych zagadnień. Dziedzina ta, to analiza numeryczna, zwana również często metodami numerycznymi, zajmująca się numerycznym rozwiązywaniem podstawowych zadań analizy matematycznej, algebry, geometrii itd.

Skrypt niniejszy stanowi wykład wybranych pojęć i metod analizy numerycznej. Jest on poświęcony takim zagadnieniom, jak działania na liczbach przybliżonych, interpolacja funkcji, aproksymacja i dobór wzorów empirycznych, numeryczne rozwiązywanie równań algebraicznych i przestępnych, metody numeryczne algebry liniowej, całkowanie numeryczne oraz numeryczne rozwiązywanie równań różniczkowych zwyczajnych.

Treść skryptu i dobór materiału pozwalają na korzystanie z niego czytelnikom posiadającym pewne minimum wiadomości z algebry liniowej oraz analizy matematycznej.

Większość twierdzeń występujących w skrypcie podana została bez dowodów. Czytelnikom zainteresowanym tymi dowodami, jak również czytelnikom bardziej zaawansowanym w matematyce wyższej polecam literaturę wymienioną na końcu skryptu.

W celu większej przejrzystości niektóre metody numeryczne sformułowane zostały w postaci reguł lub ujęte w schematy obliczeniowe. Dla łatwiejszego prześledzenia, większość przykładów została podana w bardzo prostej formie i ma charakter wyłącznie ilustracyjny.

Rozdział I

DZIAŁANIA NA LICZBACH PRZYBLIŻONYCH

§ 1. Klasyfikacja błędów

Numeryczne rozwiązanie jakiegoś zadania ręcznie, za pomocą arytmometru lub maszyny cyfrowej, z zasady nie może być dokonane z pełną dokładnością. Zarówno przy formułowaniu zadania, jak i podczas przeprowadzania rachunków, powstają pewne błędy, które powodują, że wyniki końcowe tylko z pewnym przybliżeniem reprezentują rzeczywiste wartości odpowiednich wielkości. Inaczej mówiąc, zdanie, w którym występują duże ilości działań na liczbach, może być rozwiązane prawie wyłącznie w sposób przybliżony z pewnym stopniem dokładności. W związku z powyższym, podczas formułowania zadania, zawsze należy podawać dokładność jego rozwiązania, tzn. należy podać maksymalny dopuszczalny błąd w całym procesie obliczeń.

Błędy dopuszczalne przy wyrażaniu zależności wśród zjawisk fizycznych za pomocą liczb przybliżonych spowodowane są całym szeregiem przyczyn. Do najważniejszych błędów można zaliczyć:

1. Błędy związane z niedokładnym odzwierciedleniem zjawisk za pomocą aparatu matematycznego, w związku z czym zamiast samego zjawiska rozpatruje się jego mniej lub bardziej wyidealizowany model matematyczny. Powstają wówczas błędy zwane błędami zagadnienia, jak również błędy metody.
2. Błędy związane z przybliżonym przedstawieniem wielkości fizycznych wchodzących do procesu obliczeń, a spowodowanych

częstokroć niedokładnością różnego rodzaju pomiarów. Są to tzw. błędy początkowe.

3. Błędy związane z zastąpieniem procesów nieskończonych w analizie matematycznej skończonymi ciągami działań. Są to błędy zwane najczęściej błędami obcięcia.
4. Błędy związane z pozycyjnym dziesiętnym lub innym przedstawieniem liczb. W obliczeniach może być wykorzystana jedynie skończona ilość cyfr, odrzucamy więc dalsze cyfry dokonując równocześnie okrąglenia danych wielkości. Błędy te nazywamy błędami okrąglenia.
5. Błędy związane z działaniami arytmetycznymi na liczbach przybliżonych, zwane krótko błędami działań.

Błędy dwóch pierwszych typów nie zależą od wykonującego obliczenia i nie można ich zmniejszyć w procesie obliczeń. Można je w pewnym stopniu zmniejszyć przed przystąpieniem do obliczeń przez uściślenie postawionego zadania i bardziej dokładne pomiary występujących w zadaniu wielkości. Błędy ostatnich typów mogą być zmniejszone podczas samych obliczeń poprzez wybór bardziej dokładnych metod obliczeniowych, jak również poprzez zwiększenie ilości cyfr w liczbach biorących udział w procesie obliczeń.

Błąd powstający w procesie obliczeń w wyniku skumulowania wszystkich błędów nazywamy błędem numerycznym.

Poniżej ograniczymy się w zasadzie do omówienia tylko niektórych z wymienionych rodzajów błędów. Zanim jednak przejdziemy do omówienia tych błędów, wymienimy pewne zasady, jakie powinny być przestrzegane w procesie obliczeń.

1. Określanie matematycznych charakterystyk dokładności przybliżonych wielkości.
2. Ocena dokładności wyniku przy znanej dokładności danych początkowych.
3. Znajdowanie dokładności danych początkowych zabezpieczającej żadaną dokładność wyników.
4. Uzgadnianie dokładności danych początkowych w celu zaniechania zbędnej pracy dla uzyskania dużej dokładności jednych wielkości jeśli inne dane początkowe obarczone są zbyt dużym błędem.
5. Śledzenie dokładności wyników pośrednich dla zabezpieczenia żadanej dokładności końcowych wyników z jednej strony oraz uproszczenia obliczeń z drugiej.

Radziecki matematyk Kryłow stwierdził, że wskutek nieznamości reguł działań na liczbach przybliżonych w pracach projektowych z jakimi spotykał się, około 90% rachunków wykonano całkowicie niepotrzebnie, dla uzyskania cyfr, które w rzeczywistości były bezwartościowe.

Bardzo ważnym elementem w procesie obliczeń jest tzw. zaokrąglanie liczb, które odbywa się według pewnych reguł. Przytoczymy tu jedną z nich. Liczby biorące udział w obliczeniach przedstawione są w określonym systemie pozycyjnym z zadana ilością cyfr i mają postać

$$x = \pm a^k \quad x_1 a^{-1} + x_2 a^{-2} + \dots + x_m a^{-m} \quad , \quad \text{I.}$$

gdzie: a - podstawa systemu rozwinięcia;

m - liczba całkowita dodatnia, określająca ilość cyfr;

x_i - cyfry ustalone w danym systemie rozwinięcia;

$$i=1,2,\dots,m$$

k - dowolna liczba całkowita charakteryzująca położenie przecinka.

Jeśli liczby początkowe lub wyniki pośrednie wymagają dla ich zapisu większej ilości cyfr, niż zostało ustalone dla danego procesu obliczeń, wówczas każdą z takich liczb zaokrągla się do ustalonej ilości cyfr. W ogólnym przypadku zaokrąglenie do m -tej cyfry przeprowadza się według następującej zasady:

1. odrzuca się wszystkie cyfry stojące na prawo od m -tej cyfry;
2. pozostałe cyfry pozostają bez zmiany, jeśli pierwsza z odrzuconych cyfr jest mniejsza od połowy podstawy danego rozwinięcia;
3. do ostatniej z pozostałych cyfr dodaje się jedynkę, jeśli pierwsza z odrzuconych cyfr jest większa lub równa połowie podstawy danego rozwinięcia.

Przykład: Zaokrąglić liczbę $=3.1415926535$ do czterech, trzech i dwóch cyfr po przecinku. Stosując powyższą regułę otrzymamy kolejno 3.1416, 3.142, 3.14.

§ 2. Zapis dziesiętny liczb przybliżonych.

Ilość cyfr dokładnych. Cyfry znaczące

Jak wiadomo, każdą liczbę dodatnią x można przedstawić w postaci skończonego lub nieskończonego ułamka dziesiętnego

$$x = x_m 10^m + x_{m-1} 10^{m-1} + x_{m-2} 10^{m-2} + \dots + x_{m-n+1} 10^{m-n+1} + \dots, \quad \text{I.2}$$

gdzie: x_i są cyframi liczby x $x_i = 0, 1, 2, 3, 4, \dots, 9$, m - pewna liczba całkowita, nazywana najwyższą potęgą dziesiętną liczby x i $x_m \neq 0$.

$$\text{Przykład: } 314159.265\dots = 3 \cdot 10^5 + 1 \cdot 10^4 + 4 \cdot 10^3 + 1 \cdot 10^2 + 5 \cdot 10^1 + 9 \cdot 10^0 + \\ + 2 \cdot 10^{-1} + 6 \cdot 10^{-2} + 5 \cdot 10^{-3} + \dots$$

Każda jedynek zapisana w określonym miejscu rozwinięcia dziesiętnego I.2 liczby x ma określoną wartość. Jedynek zapisana na pierwszym miejscu równa się 10^m , na drugim miejscu 10^{m-1} , na n -tym - 10^{m-n+1} itd.

Jeśli liczba x zadana ogólnie w postaci I.1 przedstawiająca wielkość, której rzeczywistą wartością jest x , jest taka, że wartość absolutna różnicy między x i x jest mniejsza lub równa połowie jednostki najmniej znaczącej cyfry

$$x - x \leq \frac{a^{k-m}}{2}, \quad \text{I.3}$$

to cyfry x_1, x_2, \dots, x_m nazywamy dokładnymi. W przypadkach, gdy nie wiadomo, czy nierówność I.3 jest spełniona, lub wiadomo, że nie jest spełniona ale różnica między x i x jest nie większa od dwóch jednostki najmniej znaczącej pozycji; wówczas ostatnią cyfrę x_m nazywamy niepewną.

Jeśli liczba x zapisana jest w systemie dziesiętnym I.2, to wszystkie cyfry w takim zapisie od pierwszej różnej od zera do ostatniej - nazywamy cyframi znaczącymi, jeśli przy tym zera stojące na końcu liczby nie zastępują cyfr nieznanymi.

Przykłady:

$$x = 7 \cdot 10^{-3} + 0 \cdot 10^{-4} + 1 \cdot 10^{-5} + 0 \cdot 10^{-6} = \underline{0.007010}$$

$$x = 2 \cdot 10^9 + 0 \cdot 10^8 + 0 \cdot 10^7 + 3 \cdot 10^6 + 0 \cdot 10^5 = 20030\underline{000000}.$$

Zera podkreślone nie są cyframi znaczącymi. Z powyższego punktu widzenia wynika, że liczby 0.002080 i 0.00208 nie mają tej samej wartości, ponieważ pierwsza ma cztery, a druga trzy cyfry znaczące.

W procesie obliczeń na liczbach przybliżonych powinniśmy korzystać z następującej bardzo praktycznej reguły: w wynikach pośrednich ilość cyfr znaczących może być większa od ilości cyfr dokładnych co najwyżej o jedną lub dwie cyfry. Wynik ostateczny może zawierać co najwyżej o jedną cyfrę znaczącą więcej niż jest w nim cyfr dokładnych.

§ 3. Ocena błędów

Ostateczny błąd obliczeń jest wynikiem złożonego wpływu błędów zarówno początkowych, jak i powstających w czasie obliczeń. Ze zwiększeniem ilości obliczeń, błędy mogą kumulować się i należy stwierdzić do jakiej granicy ta kumulacja będzie zachodzić.

We wszystkich przypadkach wartość bezwzględna błędu końcowego nie przewyższa sumy bezwzględnych wartości błędów składowych. Jednak w większości przypadków taka ocena jest zbyt gruba, jak też jest mało prawdopodobne, aby błąd całkowity osiągał taką wartość graniczną. Dokładne określenie błędu całkowitego jest zwykle niemożliwe, szczególnie przy dużej ilości obliczeń. Z tego względu szczególne znaczenie odgrywają metody oceny błędów pozwalające na określenie granic zmienności tych błędów.

Oceny zwykle dokonuje się za pomocą:

1. Błędu bezwzględnego
2. Błędu względnego
3. Szacowania ostatniego wyrazu
4. Ocen statystycznych

Błędem bezwzględnym liczby przybliżonej x nazywamy wartość bezwzględną różnicy pomiędzy liczbą dokładną i liczbą przybliżoną x

$$= x - x \quad \text{I.4}$$

Błąd bezwzględny nie charakteryzuje w pełni dokładności pomiarów czy obliczeń. Jeśli na przykład przy pomiarze długości dwóch prętów otrzymano wyniki $l_1 = 100.8 \text{ cm} \pm 0.1 \text{ cm}$ i $l_2 = 2.5 \text{ cm} \pm 0.1 \text{ cm}$, to niezależnie od tego, że błąd bezwzględny obu wielkości jest jednakowy, pomiar pierwszy jest dokładniejszy od drugiego.

Błędem względnym liczby przybliżonej x nazywamy stosunek błędu bezwzględnego danej liczby do wartości bezwzględnej liczby dokładnej x

$$= \frac{\quad}{x} \quad \text{I.5}$$

Błąd ten w istotny sposób charakteryzuje dokładność pomiarów i obliczeń.

Bardzo często liczba dokładna x jest nieznana, wówczas dla określenia granic błędu stosuje się zamiast równości I.4 i I.5 podwójne nierówności

$$\begin{aligned} x - x & \quad x \quad x + x ; \\ x \quad 1 - x & \quad x \quad x \quad 1 + x \end{aligned} \quad \text{I.6}$$

gdzie x i x oznaczają odpowiednio kres górny błędu bezwzględnego i kres górny błędu względnego liczby przybliżonej x .

Kresem górnym błędu bezwzględnego liczby przybliżonej x nazywamy możliwie najmniejszą liczbę ϵ spełniającą nierówność

$$x - \epsilon \leq x \leq x + \epsilon \quad (1.7)$$

Kresem górnym błędu względnego liczby przybliżonej x nazywamy możliwie najmniejszą liczbę δ spełniającą nierówność

$$\frac{\epsilon}{x} \leq \delta \quad (1.8)$$

Przykłady:

Określić kres górny błędu bezwzględnego liczby $x = 3.14$, przybliżającej liczbę π . Ponieważ zachodzi nierówność $3.14 < \pi < 3.15$, to $\epsilon = 0.01$, a więc możemy przyjąć $\epsilon = 0.01$. Jeśli uwzględnimy nierówność $3.14 < \pi < 3.142$, to otrzymamy oszacowanie znacznie lepsze $\epsilon = 0.002$.

Określić kres górny błędu względnego liczby $p = 999.847 = 0.001$. Oczywiście $\delta = 0.001$ i $p = 999.846$. Zatem

$$\delta = \frac{0.001}{999.846} \approx 10^{-4}\%$$

Przy określaniu stałej gazowej powietrza otrzymano $R = 29.25$. Wiedząc, że błąd względny tej wartości wynosi 1%, znaleźć przedział, w którym zawarte jest R . Ponieważ mamy $\delta = 0.001$, a więc $R - \delta \leq R \leq R + \delta$, czyli

$$29.22 \leq R \leq 29.28$$

Dla oceny w wielu metodach numerycznych błędów, pojawiających się w wyniku zastąpienia nieskończonych procesów obliczeniowych obliczaniem według formuł przybliżonych, zawierających skończoną liczbę działań, posługujemy się ostatnim wyrażeniem, który określa się jako różnicę między dokładnym i przyb-

liżonym obliczaniem danej wielkości. W wielu przypadkach dla oszacowania ostatniego wyrazu, wyprowadza się konkretne formuły, które pozwalają na ocenę tego wyrazu. Formuły takie będziemy podawać dla metod, dla których one istnieją.

Łącznie z kresem górnym błędu bezwzględnego i błędu względnego, wskazującymi na możliwie maksymalne rozłożenie liczb x i x , często stosuje się statystyczną ocenę błędu, która charakteryzuje stopień przybliżenia liczby x do rzeczywistej wielkości x zadaną miarą wiarygodności, tzn. z uwzględnieniem rozkładu różnych wartości różnicy $x - x$. W tym przypadku znajduje się prawdopodobieństwo tego, że błąd kumulujący się w procesie obliczeń, nie przewyższy pewnej zadanej wartości.

§ 4. Błędy działań arytmetycznych

Przy ocenie błędów działań arytmetycznych należy brać pod uwagę następujące oszacowania:

1. Błąd bezwzględny sumy

Błąd bezwzględny skończonej sumy algebraicznej liczb przybliżonych

$$w = \pm x_1 \pm x_2 \pm \dots \pm x_n$$

nie przekracza sumy błędów bezwzględnych tych liczb

$$w = \sum_{i=1}^n x_i ;$$

Wniosek. Za kres górny błędu bezwzględnego sumy algebraicznej można przyjąć sumę kresów górnych błędów bezwzględnych poszczególnych składników

$$w = x_1 + x_2 + x_3 + \dots + x_n$$

Z powyższego wyniku następująca praktyczna reguła sumowania liczb przybliżonych o bliskich sobie wartościach bezwzględnych i różnych znakach; dla zachowania wymaganej dokładności obliczeń należy brać składowe z dużą ilością cyfr dokładnych, a jeśli to nie jest możliwe, przekształcić algorytm tak, aby odejmowanie liczb o bliskich wartościach bezwzględnych wyeliminować.

2. Błąd względny iloczynu

Błąd względny iloczynu liczb przybliżonych

$$w = x_1 \cdot x_2 \cdot x_3 \dots x_n$$

nie przekracza sumy błędów względnych poszczególnych czynników

$$w = \sum_{i=1}^n x_i ;$$

Uwaga. Jeśli w iloczynie n dziesiętnych liczb przybliżonych $n \geq 10$ każdy czynnik ma m cyfr dokładnych, to iloczyn ma z reguły $m-1$ cyfr dokładnych /nie mniej niż $m-2$ cyfry dokładne, a czasem i m cyfr/.

3. Błąd względny ilorazu

Błąd względny ilorazu liczb przybliżonych

$$w = \frac{x_1}{x_2}$$

nie przekracza sumy błędów względnych dzielnika i dzielnej

$$w = x_1 + x_2$$

Przykłady:

1. Znaleźć sumę liczb przybliżonych: 0.348; 0.1834; 345.4; 235.2; 11.75; 9.27; 0.0849; 0.0214; 0.000354, z których każda ma wszystkie cyfry dokładne.

Wybieramy liczby najmniej dokładne 345.4 i 235.2. Błąd bezwzględny tych liczb może być co najwyżej równy 0.1.

Zaokrąglając pozostałe liczby z dokładnością do 0.01 mamy

$$\begin{array}{r} 345.4 \\ 235.2 \\ 11.75 \\ 9.27 \\ 0.35 \\ 0.18 \\ 0.08 \\ 0.02 \\ \hline 0.00 \\ \hline 602.25 \end{array}$$

Z kolei, zgodnie z regułą zaokrąglania otrzymujemy wartość przybliżoną 602.2

2. Obliczyć z dokładnością do trzech cyfr różnicę

$$u = 2.01 - 2.$$

Ponieważ mamy

$$2.01 = 1.41774469\dots, \quad 2 = 1.41421356\dots,$$

więc szukany wynik wynosi

$$u = 0.00353 = 3.53 \cdot 10^{-3}.$$

3. Obliczyć iloczyn liczb przybliżonych $x_1 = 12.2$ i $x_2 = 73.56$, w których wszystkie napisane cyfry są dokładne.

$$\begin{aligned} \text{Po zaokrągleniu mamy } x_1 = 12.2 \text{ i } x_2 = 73.6. \text{ Stąd } x_1 \cdot x_2 = \\ = 12.2 \cdot 73.6 = 897.4 \quad 89.7 \cdot 10 \end{aligned}$$

Rozdział II

INTERPOLACJA FUNKCJI

§ 1. Sformułowanie zagadnienia interpolacji

Pod nazwą interpolacja funkcji rozumiano dawniej zagadnienie wyznaczania wartości funkcji odpowiadających pośrednim wartościom argumentów, nie występującym w danej tabelicy. Z zagadnieniem takim spotykamy się bardzo często w większości zagadnień techniki obliczeniowej.

Obecnie pod pojęciem interpolacji rozumiemy również zagadnienie budowania analitycznej postaci funkcji na podstawie tabelicy wartości tej funkcji.

Niech będzie dana funkcja $y = f(x)$, której znamy tylko tablice, tzn. wiadomo, że dla $x = x_0, x_1, x_2, \dots, x_n$ funkcja ma wartości $y = y_0, y_1, y_2, \dots, y_n$

Zagadnienie geometryczne wyznaczania funkcji $f(x)$ polega na wykreśleniu krzywej przechodzącej przez punkty $x_0, y_0, x_1, y_1, \dots, x_n, y_n$. Jak wiadomo przez punkty te można poprowadzić nieskończenie wiele krzywych. Tak więc zagadnienie nie jest określone jednoznacznie.

Oznaczmy przez $F x$ dowolną funkcję przechodzącą przez te punkty.

Jeśli założymy, że $F x$ musi spełniać dodatkowe warunki, zagadnienie będzie określone ściślej. Najczęściej żąda się, aby $F x$ była wielomianem stopnia mniejszego niż ilość znanych wartości funkcji.

Zagadnienie interpolacji wielomianem można sformułować następująco:

Mając wartości $x = x_0, x_1, \dots, x_n$ i $y = y_0, y_1, \dots, y_n$ wyznaczyć wielomian $y = F x$ stopnia nie większego niż n , spełniający warunki $F x_0 = y_0, F x_1 = y_1, \dots, F x_n = y_n$. II.1

Inaczej mówiąc, zagadnienie to polega na otrzymaniu analitycznego wyrażenia dla wielomianu, który w danych punktach przyjmuje dane wartości. Jest ono najczęściej spotykane ze względu na łatwość obliczania wartości wielomianów.

Punkty x_0, x_1, \dots, x_n nazywamy węzłami interpolacji.

Wielomian $F x$ spełniający warunki II.1 nazywamy wielomianem interpolacyjnym, a wzory służące do jego wyznaczenia wzorami interpolacyjnymi.

Ustalmy jeszcze główny cel stosowania wzorów interpolacyjnych. Zastępowanie funkcji wielomianem interpolacyjnym stosujemy przede wszystkim do znajdowania pośrednich wartości funkcji. Zastępowanie takie może być stosowane również wówczas, gdy znamy analityczną postać funkcji $f(x)$, ale jest ona bardzo złożona. Interpolację stosujemy również przy opracowywaniu wielkości doświadczalnych, kiedy postać analityczna funkcji nie jest znana.

§ 2. Interpolacja wielomianami

Wzór interpolacyjny Lagrange'a

Rozpatrzmy tablicę funkcji o danych wartościach

$y_i = f(x_i) \quad i = 0, 1, 2, \dots, n$, przy czym wszystkie x_i i y_i są znane. Należy wyznaczyć wielomian $y = F(x)$ stopnia nie większego niż n , dla którego

$$F(x_i) = f(x_i) \quad i = 0, 1, 2, \dots, n \quad \text{II.2}$$

Szukany wielomian możemy zapisać

$$F(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

Uwzględniając warunki II.2 otrzymujemy układ $n+1$ równań z $n+1$ niewiadomymi

$$a_0 + a_1 x_0 + a_2 x_0^2 + \dots + a_n x_0^n = y_0$$

$$a_0 + a_1 x_1 + a_2 x_1^2 + \dots + a_n x_1^n = y_1$$

.....

$$a_0 + a_1 x_n + a_2 x_n^2 + \dots + a_n x_n^n = y_n$$

Jeśli x_0, x_1, \dots, x_n są różne, wówczas powyższy układ posiada jedno rozwiązanie i możemy jednoznacznie wyznaczyć współczynniki a_i .

Przykład.

Dane są wartości $x_0=0$ $y_0=1$

$$x_1=1 \quad y_1=1$$

$$x_2=2 \quad y_2=3$$

Znaleźć współczynniki wielomianu interpolacyjnego $F(x)$ st.2, dla którego

$$F(0) = 1, \quad F(1) = 1 \quad \text{oraz} \quad F(2) = 3$$

Przyjmując

$$F(x) = a_0 + a_1 x + a_2 x^2 \quad \text{otrzymamy układ równań}$$

$$a_0 + a_1 \cdot 0 + a_2 \cdot 0^2 = 1$$

$$a_0 + a_1 \cdot 1 + a_2 \cdot 1^2 = 1$$

$$a_0 + a_1 \cdot 2 + a_2 \cdot 2^2 = 3$$

Po rozwiązaniu otrzymujemy $a_0 = 1, a_1 = -1, a_2 = 1$, czyli

$$F(x) = 1 - x + x^2$$

Spróbujmy zagadnienie rozwiązać bardziej ogólnie tzn. wyznaczyć współczynniki wielomianu interpolacyjnego bez rozwiązywania układów równań.

Wyznamy wielomian, który w punkcie $x=x_0$ przyjmuje wartość $y_0=1$, a w punktach x_1, x_2, \dots, x_n wartości $y_1=y_2=\dots=y_n=0$. Będzie on miał postać

$$F_0 x = \frac{x-x_1}{x_0-x_1} \frac{x-x_2}{x_0-x_2} \dots \frac{x-x_n}{x_0-x_n}$$

Wielomian przyjmujący w punkcie $x=x_0$ wartość y_0 , a dla $x=x_i$ $i=1, 2, \dots, n$ zerujący się ma postać

$$F_0 x = \frac{x-x_1}{x_0-x_1} \frac{x-x_2}{x_0-x_2} \dots \frac{x-x_n}{x_0-x_n} \cdot y_0$$

Uogólniając, można wyznaczyć $F_j x$ dla dowolnego j .

$$F_j x = \frac{x-x_0}{x_j-x_0} \frac{x-x_1}{x_j-x_1} \dots \frac{x-x_{j-1}}{x_j-x_{j-1}} \frac{x-x_{j+1}}{x_j-x_{j+1}} \dots \frac{x-x_n}{x_j-x_n} \cdot y_j$$

a szukany wielomian będzie równy sumie

$$F x = \sum_{j=0}^n F_j x$$

Po rozwinięciu otrzymamy

$$F x = \frac{x-x_1}{x_0-x_1} \frac{x-x_2}{x_0-x_2} \dots \frac{x-x_n}{x_0-x_n} \cdot y_0 + \frac{x-x_0}{x_1-x_0} \frac{x-x_2}{x_1-x_2} \dots \frac{x-x_n}{x_1-x_n} \cdot y_1 + \dots + \frac{x-x_0}{x_n-x_0} \frac{x-x_1}{x_n-x_1} \dots \frac{x-x_{n-1}}{x_n-x_{n-1}} \cdot y_n$$

II.3

Wzór ten nazywamy wzorem interpolacyjnym Lagrange'a.

Przykłady:

Przy $i=1$ mamy wzór Lagrange'a oparty na dwóch punktach

$$F x = \frac{x-x_1}{x_0-x_1} \cdot y_0 + \frac{x-x_0}{x_1-x_0} \cdot y_1$$

skąd po przekształceniu możemy otrzymać

$$F x = \frac{y_0-y_1}{x_0-x_1} \cdot x + \frac{x_0 y_1 - x_1 y_0}{x_0-x_1} \quad \text{II.4}$$

Jest to wzór na równanie prostej przechodzącej przez dwa punkty, służący do interpolacji liniowej. W analogiczny sposób dla $n=2$ można uzyskać wzór będący równaniem paraboli, służący do interpolacji kwadratowej.

Następująca formuła

$$R_n x = f x - F x \frac{M_{n+1}}{(n+1)!} w_{n+1} x \quad \text{II.5}$$

gdzie $M_{n+1} = \max_{a \leq x \leq b} f^{(n+1)} x$; $w_{n+1} x = (x-x_0)(x-x_1)\dots(x-x_n)$

służy do szacowania błędu bezwzględnego interpolacyjnego wzoru Lagrange'a.

§ 3. Wzory interpolacyjne Newtona

Niech będą dane wartości funkcji $f x$:

$$y_i = f x_i, \quad i=0, 1, 2, \dots, n$$

w punktach rozmieszczonych w jednakowych odstępach

$$x_0, x_1=x_0+h, x_2=x_0+2h, \dots, x_n=x_0+nh.$$

Jak wiadomo, istnieje jeden wielomian $F x$ stopnia n ,

taki, że

$$F x_k = y_k \quad k=0, 1, 2, \dots, n \quad \text{II.6}$$

Zapiszmy wielomian, który będzie tożsamościowo równy wielomianowi Lagrange'a w postaci

$$F x = a_0 + a_1 x - x_0 + a_2 x - x_0 x - x_1 + a_3 x - x_0 x - x_1 x - x_2 + \dots + a_n x - x_0 x - x_1 \dots x - x_{n-1} \quad \text{II.7}$$

Aby wyznaczyć współczynniki $a_0, a_1, a_2, \dots, a_n$ przyjmijmy najpierw

$$x = x_0 \quad \text{wówczas} \quad y_0 = F x_0 = a_0$$

$$x = x_1 \quad \text{wówczas} \quad y_1 = F x_1 = a_0 + a_1 x_1 - x_0, \text{ ponieważ}$$

$$x_1 - x_0 = h, \text{ więc } y_1 = y_0 + a_1 h \text{ skąd}$$

$$a_1 = \frac{y_1 - y_0}{h} = \frac{y_0}{h} .$$

Analogicznie możemy obliczyć pozostałe współczynniki

$$a_2 = \frac{y_0}{2!h^2}$$

$$a_3 = \frac{y_0}{3!h^3}$$

.....

$$a_k = \frac{y_0}{k!h^k}$$

Po podstawieniu tych wartości do wzoru II.7 otrzymamy

$$F x = y_0 + \frac{y_0}{h} x - x_0 + \frac{y_0}{2!h^2} x - x_0 x - x_1 + \dots + \frac{y_0}{n!h^n} x - x_0 x - x_1 \dots x - x_{n-1} \quad \text{II.8}$$

Jest to pierwszy wzór interpolacyjny Newtona.

Różnica między wzorami Lagrange'a i Newtona polega na tym, że we wzorze Lagrange'a każdy ze składników jest wielomianem stopnia n i wszystkie składniki mają jednakowe znaczenie, obliczenia muszą więc być prowadzone do końca. We wzorze Newtona składniki są wielomianami coraz to wyższego stopnia, a współczynniki ich są kolejnymi różnicami dzielonymi przez h w odpowiedniej potęgze, dzięki czemu kolejne wyrazy szybko maleją, obliczenia można przerwać w chwili uzyskania odpowiedniego rezultatu pośredniego.

Dla praktycznego posługiwania się tym wzorem dokonajmy następującego przekształcenia:

$$\frac{x-x_0}{h} = t \quad \text{czyli } x=x_0+th, \quad \text{a po podstawieniu}$$

do II.8 otrzymamy

$$F_{x_0+th} = y_0 + \frac{t}{1!} y_0' + \frac{t(t-1)}{2!} y_0'' + \frac{t(t-1)(t-2)}{3!} y_0''' + \dots$$

$$+ \frac{t(t-1)(t-2)\dots(t-n+1)}{n!} y_0^{(n)} \quad \text{II.9}$$

Wzór ten nazywamy często wzorem interpolacyjnym Newtona na interpolację w przód.

Przykład:

Dana jest tablica logarytmów dziesiętnych liczb od 1000 do 1050 z krokiem 10. Obliczyć logarytmy dziesiętne liczb od 1000 do 1010 co 1.

x	y	y	2_y	3_y
1000	3.0000000	43214	-426	8
1010	3.0043214	42788	-418	9
1020	3.0086002	42370	-409	8
1030	3.0128372	41961	-401	
1040	3.0170333	41560		
1050	3.0211893			

Trzecie różnice można przyjąć praktycznie za stałe.

Przyjmując $x_0 = 1000$ mamy $y_0 = 3.0000000$

$y_0 = 0.0043214$

${}^2y_0 = -0.0000426$

${}^3y_0 = 0.0000008$

Określmy występujący we wzorze II.9 parametr t . Ponieważ

$th = 10$, to dla $x_1 = 1001$ mamy $t_1 = 0.1$

dla $x_2 = 1002$ mamy $t_2 = 0.2$ itd.

t	x	$t y_0$	$\frac{t(t-1)}{2} {}^2y_0$	$\frac{t(t-1)(t-2)}{6} {}^3y_0$	y
0.0	1000	0	0	0	3.0000000
0.1	1001	43214	192	2	3.0004341
0.2	1002	86428	341	4	3.0008677
0.3	1003	129642	447	5	3.0013009
0.4	1004	172856	512	5	3.0017337
0.5	1005	216070	533	5	3.0021661
0.6	1006	259284	512	4	3.0025980
0.7	1007	302498	447	4	3.0030295
0.8	1008	345712	341	2	3.0034606
0.9	1009	388926	192	1	3.0038912
1.0	1010	432140	0	0	3.0043214

Porównując uzyskane wyniki z 7 - cyfrową tablicą logarytmów dziesiętnych okazuje się, że wszystkie cyfry uzyskane za pomocą interpolacji są dokładne. Aby zbudować tablicę w przedziale 1010×1020 przyjmujemy jako $x_0 = 1010$ oraz uwzględniamy drugi wiersz z tablicy różnic. Zastosowany w przykładzie wzór jest przydatny tylko na początku tablicy funkcji. Na końcu tablicy stosujemy inny wzór, który można wyprowadzić w analogiczny sposób.

$$F x = F x_{n+th} = y_{n+t} y_{n-1} + \frac{t t+1}{2!} y_{n-2} + \dots + \frac{t t+1 t+2 \dots t+n-1}{n!} y_0 \quad \text{II.10}$$

Jest to drugi wzór interpolacyjny Newtona, zwany często wzorem interpolacyjnym Newtona na interpolację wstecz.

Następujące formuły

$$R_n x = \frac{t t-1 t-2 \dots t-n}{n+1!} y_0 \quad \text{gdzie } t = \frac{x-x_0}{h}$$

$$R_n x = \frac{t t+1 t+2 \dots t+n}{n+1!} y_n \quad \text{gdzie } t = \frac{x-x_n}{h}$$

służą do szacowania błędu bezwzględnego odpowiednio pierwszego i drugiego wzoru interpolacyjnego Newtona.

Przykład:

W pięciocyfrowych tablicach logarytmicznych dane są logarytmy liczb całkowitych od 1000 do 10 000, przy czym kres górny błędu bezwzględnego tych logarytmów wynosi $\frac{1}{2} \cdot 10^{-5}$; zbadać, czy możliwa jest interpolacja liniowa z tą samą dokładnością.

Niech $y = \log x$, przyjmując we wzorach na szacowanie błędu bezwzględnego zależność

$${}_{n+1}y_0 \quad f^{n+1} \quad ,$$

wówczas $y' = \frac{1}{x}$

$$y'' = -\frac{1}{x^2}$$

$$f'' = \frac{0.5}{10^6} = \frac{1}{2} \cdot 10^{-6}$$

więc

$$R_1 \approx \frac{t(t-1)}{2!} \cdot \frac{1}{2} \cdot 10^{-6}$$

ponieważ $0 < t < 1$, $t(t-1) = \frac{1}{4} - \frac{1}{2}t^2 + \frac{1}{4}$

$$R_1 \approx \frac{1}{4} \cdot \frac{1}{2} \cdot 10^{-6} = 10^{-5}$$

czyli interpolacja liniowa jest możliwa.

Oprócz podanych tu wzorów interpolacyjnych Lagrange'a i Newtona istnieją dość często spotykane wzory Gaussa, Stirlinga, Bessela, Aitkena i inne.

§ 4. Ekstrapolacja

Rozważaliśmy dotychczas zastosowanie wzorów interpolacyjnych do wyznaczania wartości funkcji odpowiadających wartościom argumentów leżącym pomiędzy zadanymi punktami. Wzory interpolacyjne można stosować również do wyznaczania wartości funkcji dla argumentów znajdujących się poza granicami zadanej tablicy. Takie zastosowanie wzorów interpolacyjnych nazywamy ekstrapolacją.

Przy ekstrapolacji należy pamiętać, że występują tu na ogół większe błędy niż przy interpolacji oraz o tym, że zakres ekstrapolacji jest bardzo ograniczony.

Do ekstrapolacji najczęściej stosuje się wzory interpolacyjne Newtona oraz Lagrange'a, nie nadają się natomiast zupełnie wzory Stirlinga i Bessela. Pierwszy wzór interpolacyjny

Newtona stosujemy do ekstrapolacji wstecz, a drugi do ekstrapolacji w przód.

§ 5. Wybór wzoru interpolacyjnego

Wyniki obliczeń interpolacyjnych, jak wiadomo, nie zależą od wyboru wzoru interpolacyjnego, wykazaliśmy bowiem, że wielomian interpolacyjny jest jeden. Podobnie przedstawia się sprawa z obliczaniem nie wielomianu, a bezpośrednio jego wartości w określonym punkcie. Stosowanie różnych wzorów interpolacyjnych podyktowane jest wyłącznie prostotą i dążeniem do skrócenia czasu obliczeń, jakie należy wykonać dla wyznaczenia wielomianu, lub jego wartości.

Wśród zagadnień interpolacyjnych wyróżniamy cztery zasadnicze typy:

- I. Poszukiwanie wielomianu stopnia co najwyżej n , przy czym argumenty funkcji są równoległe, tzn. $x_i = x_0 + ih$ $i=1, 2, 3, \dots, n$.
- II. Poszukiwanie wartości wielomianu w danym punkcie, określonego w zagadnieniu typu I.
- III. Poszukiwanie wielomianu stopnia co najwyżej n , przy czym argumenty funkcji nie są równoodległe.
- IV. Poszukiwanie wartości wielomianu interpolacyjnego w danym punkcie, określonego w zagadnieniu typu III.

W przypadku zagadnienia typu I, jeśli dysponujemy tablicami różnic aż do n -tej, najwygodniej jest korzystać z wzorów Newtona, jeśli nie mamy różnic, korzystamy z wzoru interpolacyjnego Lagrange'a.

W zagadnieniach typu III wybór wzoru interpolacyjnego zależy od położenia punktu, dla którego poszukujemy wartości wielomianu. Wybieramy zazwyczaj jeden z dwóch wzorów Newtona.

Do zagadnień typu III najczęściej wykorzystuje się wzór Lagrange'a lub Newtona oparty na ilorazach różnicowych.

W zagadnieniach typu IV wykorzystujemy wzór Lagrange'a lub Aitkena, a najczęściej wzór Newtona oparty na ilorazach różnicowych.

Wzory interpolacyjne Gaussa, Stirlinga, Bessela lub inne możemy stosować do zagadnień wszystkich czterech typów, ale tylko wówczas, kiedy dysponujemy tablicami, lub możemy pozwolić sobie na zbudowanie tablic różnic centralnych.

Przykład

Określić z jaką dokładnością można obliczyć 115 za pomocą wzoru interpolacyjnego dla funkcji $y = x$, jeśli wiadomo,

$$\text{że } x_0 = 100$$

$$x_1 = 121$$

$$x_2 = 144$$

Kolejne pochodne funkcji $y = x$ będą:

$$y' = \frac{1}{2}x^{-\frac{1}{2}}$$

$$y'' = -\frac{1}{4}x^{-\frac{3}{2}}$$

$$y''' = \frac{3}{8}x^{-\frac{5}{2}}$$

Wówczas otrzymujemy

$$M_3 = \max y''' = \frac{3}{8} \frac{1}{100^{\frac{5}{2}}} = \frac{3}{8} \cdot 10^{-5}$$

dla 100 x 144, wobec czego

$$R_2 = \frac{3}{8} \cdot 10^{-5} \cdot \frac{1}{3!} \cdot 115-100 \cdot 115-121 \cdot 115-144 = \frac{1}{16} \cdot 10^{-5} \cdot 15 \cdot 6 \cdot 29 = 1.6 \cdot 10^{-3}$$

Możemy więc dysponując trzema wartościami funkcji obliczyć 115 z dokładnością do $1.6 \cdot 10^{-3}$

Rozdział III

Aproksymacja

§ 1. Sformułowanie zagadnienia.

Przybliżenia jednostajne

W rachunkach numerycznych bardzo często spotykamy się z potrzebą zastępowania danej funkcji $y = F(x)$, inną znacznie wygodniejszą do obliczeń funkcją $y = f(x)$. Funkcję $f(x)$ nazywamy wówczas aproksymacją lub przybliżeniem funkcji $F(x)$.

Postępowanie takie powoduje powstawanie w procesie obliczeń pewnych błędów, które nazywamy błędami aproksymacji. Aproksymacja ma sens tylko wówczas, gdy podaje się sposób oszacowania błędu tej aproksymacji, który zależy w znacznym stopniu od doboru metody aproksymacyjnej.

Przykład

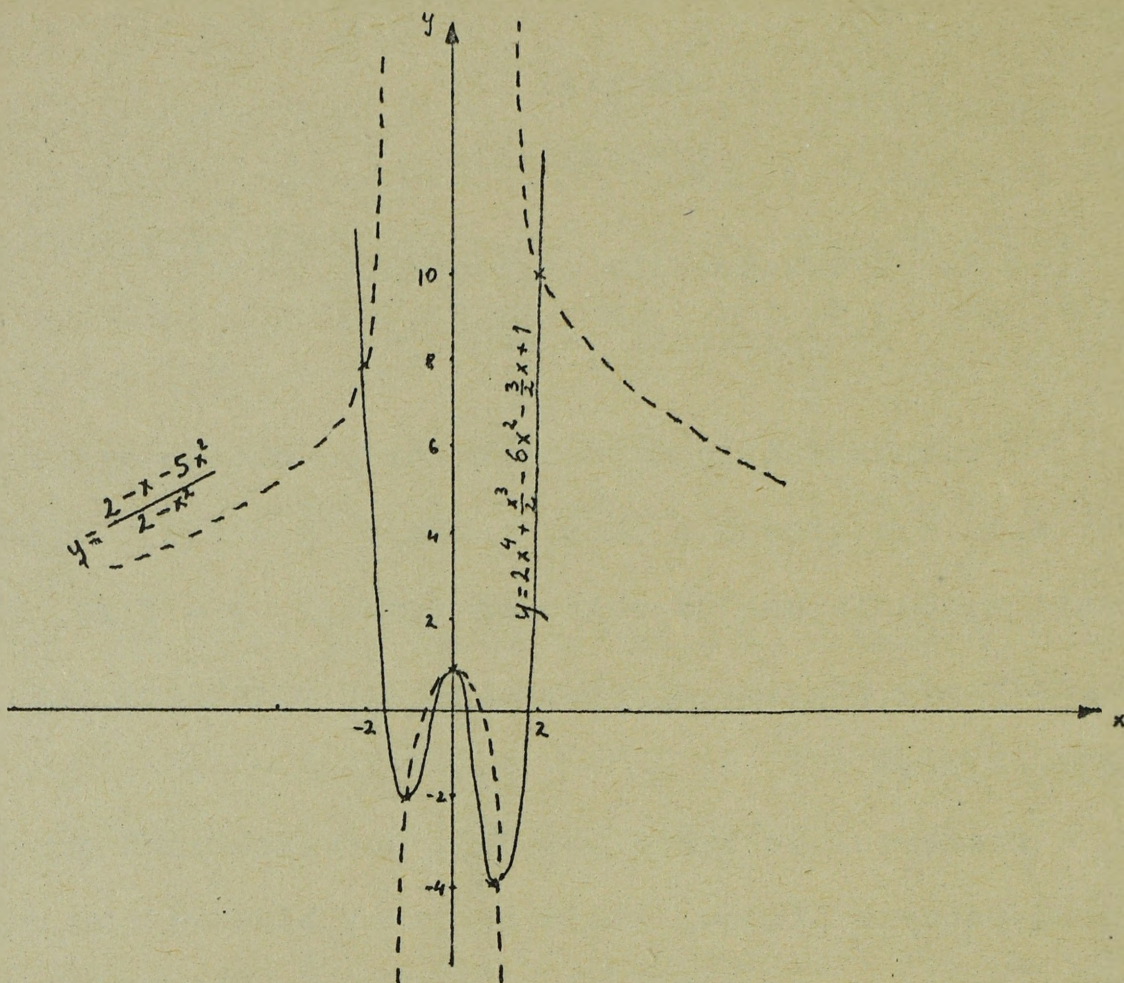
Dana jest funkcja $y = F(x)$, która w punktach $-2, -1, 0, 1, 2$, przybiera wartości $8, -2, 1, -4, 10$.

Żądając przybliżenia tej funkcji wielomianem możliwie niskiego stopnia, który w odpowiednich punktach posiadałby odpowiednie wartości, otrzymamy

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

Żądając przybliżenia funkcją wymierną otrzymamy

$$f^x(x) = \frac{2 - x - 5x^2}{2 - x^2}$$



Niech będą dane dwie funkcje $F(x)$ i $f(x)$ określone w przedziale domkniętym $[a, b]$. Funkcję $f(x)$ nazywać będziemy przybliżeniem jednostajnym funkcji $F(x)$ w przedziale $[a, b]$ z dokładnością do ξ , jeżeli

$$|F(x) - f(x)| < \xi \quad (\text{III.1})$$

dla wszystkich $x \in [a, b]$.

Najważniejszym przypadkiem przybliżenia jednostajnego jest aproksymacja funkcji ciągłej wielomianem algebraicznym. Następujące twierdzenie Weierstrassa jest podstawą tej aproksymacji:

Niech $F(x)$ będzie funkcją ciągłą w przedziale domkniętym $[a, b]$. Dla każdej liczby dodatniej ξ istnieje taki wielomian $\omega(x)$, że

$$|F(x) - \omega(x)| < \xi \quad (\text{III.2})$$

dla wszystkich $x \in [a, b]$.

Liczbę

$$\Delta = \max_{a \leq x \leq b} |\omega(x) - F(x)|$$

nazywamy odległością wielomianu $\omega(x)$ od funkcji $F(x)$ w przedziale $[a, b]$. Każdą liczbę nie mniejszą od Δ nazywamy maksymalnym błędem aproksymacji funkcji $F(x)$ w przedziale $[a, b]$ przez wielomian $\omega(x)$.

§ 2. Metoda szeregów potęgowych

Niech będzie dana funkcja $F(x)$, którą można rozwinąć w otoczeniu punktu $x = c$ w szereg Taylora o promieniu zbieżności r .

Niech również będzie

$$c-r < a < b < c+r$$

wówczas wielomian stopnia n

$$\omega(x) = F(c) + \frac{F'(c)}{1!}(x-c) + \frac{F''(c)}{2!}(x-c)^2 + \dots + \frac{F^{(n)}(c)}{n!}(x-c)^n \quad (\text{III.4})$$

przybliża jednostajnie $F(x)$ w przedziale $[a, b]$.

Jak wynika ze wzoru Taylora

$$F(x) - \omega(x) = \frac{F^{(n+1)}(\xi)}{(n+1)!} (x-c)^{n+1} \quad (\text{III.5})$$

gdzie $c < \xi < x$ gdy $c < x$

lub $x < \xi < c$ gdy $x < c$

$$\text{czyli } |F(x) - \omega(x)| \leq \frac{M_{n+1}}{(n+1)!} h^{n+1}$$

gdzie $|F^{(n+1)}(x)| \leq M_{n+1}$ dla $a < x < b$, $|c-a| \leq h$, $|c-b| \leq h$

Jest to oszacowanie błędu aproksymacji funkcji $F(x)$ przez wielomian $\omega(x)$ w przedziale $[a, b]$.

Rozwijalność funkcji w szereg Taylora pozwala nam przybliżać ją wielomianami dowolnego stopnia.

Przykład

Znaleźć za pomocą szeregów potęgowych wielomian pierwszego stopnia przybliżający funkcję $F(x) = \sin x$ w przedziale $[0, \frac{\pi}{6}]$.

Przyjmijmy $c = \frac{\pi}{12}$, wówczas

$$\omega(x) = \sin \frac{\pi}{12} + \cos \frac{\pi}{12} (x - \frac{\pi}{12}) = \cos \frac{\pi}{12} x + (\sin \frac{\pi}{12} - \frac{\pi}{12} \cos \frac{\pi}{12})$$

$$\omega(x) = 0.96593x + 0.00594.$$

Obliczmy jeszcze błąd aproksymacji. W przedziale $0, \frac{\pi}{6}$ mamy

$$\left| F^{(n+1)}(x) \right| = \left| F''(x) \right| = \left| -\sin x \right| = \sin x \leq 0.5, \text{ więc}$$

$$M_2 = 0.5. \text{ Ponieważ } a = 0, \quad b = \frac{\pi}{6}, \quad c = \frac{\pi}{12} \text{ to } h = \frac{\pi}{12}$$

a więc

$$\left| F(x) - \omega(x) \right| < \frac{0.5}{2} \frac{\pi}{12} < 0.0172$$

§ 3. Metoda najmniejszych kwadratów

Niech będzie dana funkcja $y = F(x)$, która w różnych punktach $x_0, x_1, x_2, \dots, x_m$ przyjmuje wartości $y_0, y_1, y_2, \dots, y_m$. Należy znaleźć takie przybliżenie funkcji $y = F(x)$ wielomianem co najwyżej stopnia ($n \leq m$)

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n \quad (\text{III.6})$$

aby błąd średni M aproksymacji osiągnął minimum tzn. aby funkcja

$$\begin{aligned}
 \varphi(a_0, a_1, a_2, \dots, a_n) &= \\
 &= (y_0 - a_0 - a_1 x_0 - a_2 x_0^2 - \dots - a_n x_0^n + \\
 &+ (y_1 - a_0 - a_1 x_1 - a_2 x_1^2 - \dots - a_n x_1^n)^2 + \quad \text{(III.7)} \\
 &\quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
 &+ (y_m - a_0 - a_1 x_m - a_2 x_m^2 - \dots - a_n x_m^n)^2 = \\
 &= (m+1)M^2
 \end{aligned}$$

osiągała minimum.

Przy takim postawieniu zagadnienia minimalizujemy tylko błąd średni M z jakim wartości wielomianu (III.6) w punktach x_0, x_1, \dots, x_m przybliżają wartości y_0, y_1, \dots, y_m funkcji $F(x)$. O błędzie z jakim wielomian (III.6) przybliża $F(x)$ poza punktami $x_0, x_1, x_2, \dots, x_m$ nie potrafimy na ogół nic powiedzieć.

Taką metodę aproksymacji nazywamy metodą najmniejszych kwadratów. W przypadku przybliżeń interpolacyjnych stosuje się wielomian stopnia n równego ilości punktów m , natomiast w przypadku metody najmniejszych kwadratów stosujemy wielomian stopnia n mniejszego od ilości punktów m .

Metodę najmniejszych kwadratów można również stosować w przypadku $n = m$, ale wówczas proces obliczeń trwa znacznie dłużej niż za pomocą innych metod.

W metodzie najmniejszych kwadratów szukamy minimum funkcji $\varphi(a_0, a_1, \dots, a_n)$. Muszą więc być spełnione warunki

$$\frac{\partial \varphi}{\partial a_0} = \frac{\partial \varphi}{\partial a_1} = \dots = \frac{\partial \varphi}{\partial a_n} = 0 \quad \text{(III.8)}$$

Otrzymujemy stąd $n+1$ równań liniowych z $n+1$ niewiadomymi.

Przykład:

Dana jest funkcja $y = F(x)$, która w punktach $0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}$ przyjmuje wartości $0, \frac{1}{2}, \frac{\sqrt{2}}{2}, \frac{\sqrt{3}}{2}, 1$.

Znaleźć metodą najmniejszych kwadratów przybliżenie funkcji $F(x)$ wielomianem $\omega(x)$ co najwyżej drugiego stopnia.

$$\text{Mamy } \omega(x) = a_0 + a_1 x + a_2 x^2$$

oraz $n = 2, m = 4$.

Po obliczeniu otrzymujemy

$$\begin{aligned} S_0 &= 5 & T_0 &= \frac{3 + \sqrt{2} + \sqrt{3}}{2} \\ S_1 &= \frac{5}{4} & T_1 &= \frac{\pi (14 + 3\sqrt{2} + 4\sqrt{3})}{24} \\ S_2 &= \frac{65\pi^2}{144} & T_2 &= \frac{\pi^2 (76 + 9\sqrt{2} + 16\sqrt{3})}{288} \\ S_3 &= \frac{35\pi^3}{192} \\ S_4 &= \frac{1649\pi^4}{20736} \end{aligned}$$

Po zamianie $a_0 = A_0$

$$\pi a_1 = A_1$$

$$\pi^2 a_2 = A_2$$

i podstawieniu do (III.10) otrzymujemy

$$\text{układ } 720 A_0 + 180 A_1 + 65 A_2 = 72 (3 + \sqrt{2} + \sqrt{3})$$

$$720 A_0 + 260 A_1 + 105 A_2 = 24 (14 + 3\sqrt{2} + 4\sqrt{3})$$

$$9360 A_0 + 3780 A_1 + 1649 A_2 = 72 (76 + 9\sqrt{2} + 16\sqrt{3})$$

Po rozwiązaniu otrzymujemy

$$A_0 = \frac{56 - 8\sqrt{2} - 27\sqrt{3}}{420}, \quad A_1 = \frac{120\sqrt{2} + 111\sqrt{3} - 105}{70},$$

$$A_2 = \frac{42 - 24\sqrt{2} - 18\sqrt{3}}{7}$$

skąd

$$\omega(x) \approx -0.004950 + 3.670904 \frac{x}{\pi} - 3.302577 \left(\frac{x}{\pi}\right)^2$$

Błąd średni

$$M^2 = \frac{1}{5} (0.00495^2 + 0.01513^2 + 0.00075^2 + 0.01430^2 + 0.00486^2) \approx \\ \approx 0.00009642$$

stąd $M \approx 0.00982$.

Błąd ten można zmniejszyć przez przybliżenie funkcji $F(x)$ wielomianem stopnia większego niż 2, albo aproksymować $F(x)$ na mniejszych przedziałach.

§ 4. Dobór wzorów empirycznych

Zagadnienie znajdowania wzoru empirycznego polega na poprowadzeniu krzywej postaci

$$\tilde{y} = \tilde{f}(x)$$

należącej do pewnej klasy K i leżącej "możliwie blisko" układu punktów $M_i(x_i, y_i)$ ($i = 0, 1, 2, \dots, n$).

Różnica pomiędzy interpolacją a doбором wzoru empirycznego polega w zasadzie na ilości danych punktów jakimi dysponujemy.

Konstrukcja wzoru empirycznego przebiega w dwóch etapach:

- 1) określenie ogólnej postaci wzoru
- 2) znalezienie najlepszych parametrów.

Punkt 1) w dużym stopniu zależy od doświadczenia opracowującego dane doświadczalnie. W większości przypadków ograniczamy się do wielomianów, ale dość często stosujemy również inne funkcje elementarne (wymierne, potęgowe, wykładnicze itp.).

Najczęściej spotykanymi w praktyce wzorami empirycznymi są:

$$y = ax + b$$

$$y = ax^b$$

$$y = ab^x$$

$$y = a + \frac{b}{x}$$

$$y = \frac{1}{ax + b}$$

$$y = \frac{x}{ax + b}$$

$$y = a \ln x + b$$

gdzie a, b są parametrami podlegającymi optymalizacji.

§ 5. Wybór metody aproksymacji

Zagadnienia związane z aproksymacją można podzielić na dwie klasy w zależności od teorii błędów, na której opieramy aproksymację.

Do pierwszej klasy należą zagadnienia oparte na teorii błędów maksymalnych. Zakładamy tu, że błąd aproksymacji nie może w żadnym punkcie rozpatrywanego przedziału przekroczyć określonej wartości, czyli poszukujemy przybliżeń jednostajnych. Do klasy tej należą zagadnienia związane z aproksymacją funkcji danych w postaci analitycznej przez inne funkcje. Stosujemy tu najczęściej metodę szeregów potęgowych albo metodę interpolacyjną.

Do drugiej klasy należą zagadnienia oparte na statystycznej teorii błędów. Błąd aproksymacji określamy przez podanie błędu średniego. W związku z tym błędy aproksymacji mogą być lokalnie bardzo duże, ale prawdopodobieństwo ich wystąpienia może być bardzo małe, co uzyskujemy przez założenie, że błąd średni ma być mniejszy od pewnej ustalonej

wartości. Do klasy tej należą zagadnienia doboru wzorów empirycznych. Ponieważ wyniki eksperymentu mają jedynie znaczenie statystyczne i prowadzenie dokładnych krzywych nie ma większego sensu, stosujemy tu najczęściej metodę najmniejszych kwadratów.

Rozdział IV

Przybliżone rozwiązywanie równań algebraicznych
nieliniowych i przestępnych

§ 1. Sformułowanie zagadnienia

Klasyfikacja metod

W rozdziale tym omówimy metody przybliżonego rozwiązywania równań postaci

$$f(x) = 0,$$

gdzie $f(x)$ - zadana funkcja jednej zmiennej, ciągła w przedziale $a < x < b$. Dodatkowo założymy, że funkcja $f(x)$ posiada $f'(x)$ oraz $f''(x)$.

Każdą wartość ξ taką, że $f(\xi) = 0$ nazywamy pierwiastkiem równania $f(x) = 0$ lub zerem funkcji $f(x)$.

Funkcje $f(x)$ mogą być wielomianami stopnia n , wówczas równania nazywamy algebraicznymi nieliniowymi lub funkcjami przestępnymi, wówczas równania nazywamy przestępnymi.

W teorii równań algebraicznych dowodzi się, że równanie stopnia n

$$f(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n = 0 \quad (\text{IV.1})$$

posiada n pierwiastków, wśród których mogą być zarówno pierwiastki rzeczywiste jak i zespolone.

Równanie przestępne

$$f(x) = 0 \quad (\text{IV.2})$$

może w ogóle nie mieć pierwiastków, jak również może posiadać skończoną lub nieskończoną liczbę pierwiastków (rzeczywistych i zespolonych).

Rozwiązanie równań algebraicznych nieliniowych lub przestępnych sprowadza się do znalezienia przybliżonych wartości pierwiastków, dla których zadana funkcja $f(x)$ zeruje się z zadaną dokładnością.

Znajdowanie pierwiastków równania (IV.1) lub (IV.2) można rozbić na dwa etapy:

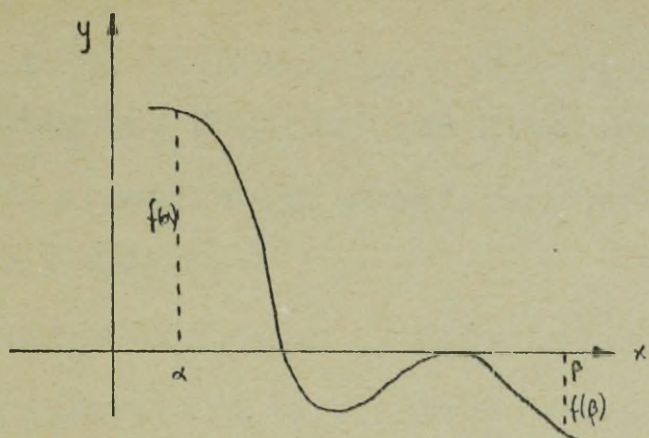
1. lokalizacja i rozdzielanie pierwiastków, tzn. znajdowanie dostatecznie małych przedziałów, w których znajduje się dokładnie po jednym pierwiastku;
2. zwiększenie dokładności określonych w pierwszym etapie pierwiastków przybliżonych.

W następnych paragrafach omówimy pewne metody lokalizacji i oddzielania pierwiastków oraz metody uściślenia pierwiastków równań algebraicznych nieliniowych oraz równań przestępnych, mające szerokie zastosowanie w technice obliczeniowej.

§ 2. Lokalizacja i rozdzielanie pierwiastków

W celu rozdzielania pierwiastków równania (IV.1) lub (IV.2) należy przede wszystkim określić granice przedziałów, w których rozłożone są pierwiastki odpowiedniego równania. Bardzo przydatne przy tym okazuje się twierdzenie Bolzano-Cauchy'ego.

Jeżeli funkcja ciągła $f(x)$ ma na końcach przedziału domkniętego $[\alpha, \beta]$ różne znaki, tzn. $f(\alpha) \cdot f(\beta) < 0$, to wewnątrz tego przedziału istnieje co najmniej jeden pierwiastek równania $f(x) = 0$, tzn. istnieje do najmniej jedna liczba $\zeta \in (\alpha, \beta)$ dla której $f(\zeta) = 0$.



Jeżeli w przedziale otwartym (α, β) istnieje i nie zmienia znaku pochodna $f'(x)$, tzn. jeżeli $f'(x) > 0$ lub $f'(x) < 0$ dla $\alpha < x < \beta$, to ξ jest jedynym pierwiastkiem. Lokalizację pierwiastków rozpoczynamy zwykle od ustalenia znaków na końcach przedziału, a następnie ten przedział dzielimy na mniejsze podprzedziały $\alpha_1, \alpha_2, \dots, \alpha_n$, których wybór zależy od osobliwości funkcji $f(x)$. Jeśli okaże się, że dla pewnego podprzedziału zachodzi $f(\alpha_k) \cdot f(\alpha_{k+1}) < 0$, to na podstawie twierdzenia Bolzano-Cauchy'ego pierwiastki leżą pomiędzy α_k i α_{k+1} .

Istnieje wiele różnych metod określania początkowych przybliżeń pierwiastków, z których kilka przytoczymy.

1. Metoda graficzna

Do rozwiązywania równań algebraicznych nieliniowych i przestępnych możemy stosować jedną z metod graficznych.

Równanie (IV.1) lub (IV.2) można zapisać w postaci

$$f(x) = 0$$

lub
$$\varphi(x) = \psi(x)$$

Rysujemy kolejno krzywe odpowiadające poszczególnym funkcjom $f(x)$ lub $\varphi(x)$ i $\psi(x)$.

Współrzędne punktów przecięcia krzywej $f(x)$ z osią x lub

punktów wzajemnego przecięcia się krzywych odpowiadających $\varphi(x)$ i $\psi(x)$ stanowią początkowe przybliżenia pierwiastków równania.

Przykłady.

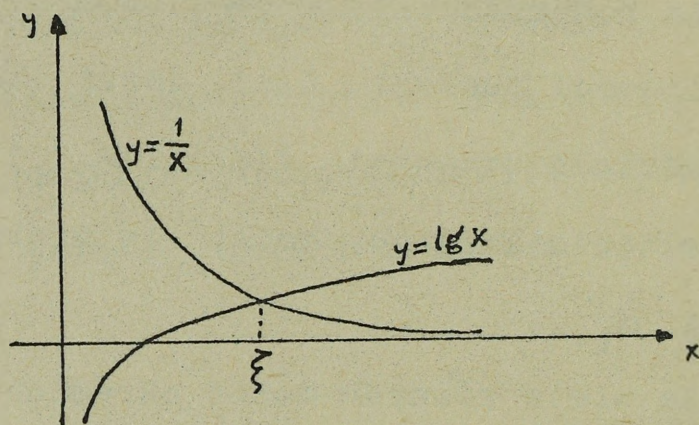
1) Określić graficznie pierwiastki równania

$$x - x \log x - 1 = 0$$

Po przekształceniu równanie to możemy zapisać w formie następującej: $\log x = \frac{1}{x}$, wówczas

$$\varphi(x) = \log x$$

$$\psi(x) = \frac{1}{x}$$



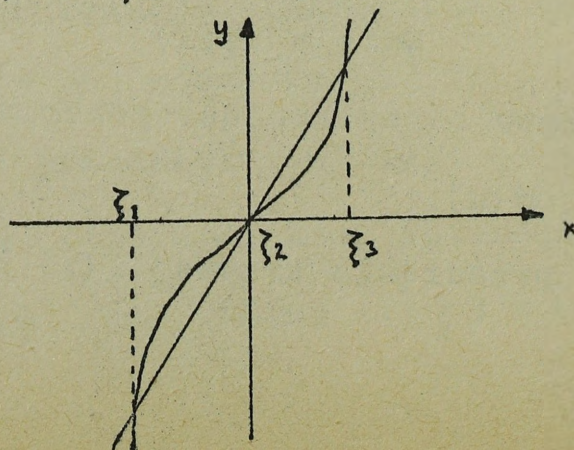
Tak wyznaczona liczba stanowi przybliżenie początkowe równania.

2) Określić metodą graficzną pierwiastki równania

$$\varphi = x^3 - 1,75x + 0,75 = 0$$

Możemy tu dokonać następującego podziału $\varphi(x) = x^3$

oraz $\psi(x) = -1,75x + 0,75$



ξ_1, ξ_2 i ξ_3 stanowią przybliżenie początkowe równania
 $y = x^3 - 1,75x + 0,75$

2. Metoda Maclaurina

Granice rozłożenia pierwiastków równania (IV.1) o współczynnikach rzeczywistych przy $a_0 > 0$ można określić

$$\frac{1}{1 + \sqrt[s]{\frac{|b_r|}{b_0}}} < x < 1 + \sqrt[m]{\frac{|a_k|}{a_0}} \quad (\text{IV.3})$$

gdzie

m - indeks pierwszego ujemnego współczynnika a_m w ciągu a_1, a_2, \dots, a_n ;

a_k - największy co do bezwzględnej wartości ujemny współczynnik tego ciągu;

$b_i = (-1)^n a_{n-i}$ ($i = 0, 1, 2, \dots, n$), tzn. współczynniki

b_i są współczynnikami równania

$$a_0 + a_1 y + a_2 y^2 + \dots + a_n y^n = 0$$

pomnożonego przez $(-1)^n$

s - indeks pierwszego ujemnego współczynnika b_s ;

b_r - największy co do bezwzględnej wartości ujemny współczynnik w ciągu b_1, b_2, \dots, b_n .

Wzór (IV.3) służy do określania dolnej i górnej granicy pierwiastków rzeczywistych.

W celu znalezienia granic pierwiastków rzeczywistych ujemnych, należy w równaniu (IV.1) dokonać zamiany $x = -z$ i pomnożyć równanie przez $(-1)^n$.

Metodę tę stosuje się najczęściej w tych przypadkach, kiedy wystarczy bardzo gruba ocena granic pierwiastków rzeczywistych.

Oprócz podanych tu metod istnieje wiele innych metod rozwiniętych szczególnie dla równań algebraicznych. W celu znalezienia pierwszych przybliżeń pierwiastków dla równań przestępnych możemy posłużyć się metodami dla równań algebraicznych przez rozwinięcie funkcji występującej w równaniu $f(x)=0$ w szereg potęgowy.

§ 3. Metoda Newtona

Założmy, że pierwiastek równania

$$f(x) = 0$$

jest zlokalizowany w przedziale domkniętym $[a, b]$ oraz $f'(x)$ $f''(x)$ istnieją w tym przedziale i nie zmieniają znaku. Założmy również, że znalezione zostało pewne przybliżenie pierwiastka $x_n \approx \zeta$. W celu zwiększenia dokładności tego przybliżenia możemy posłużyć się następującą metodą zwaną metodą Newtona albo metodą stycznych.

Niech

$$\zeta = x_n + h_n$$

gdzie h_n jest pewnym małym przyrostem.

Rozwijając funkcję $f(x)$ w szereg Taylora w otoczeniu punktu x_n , mamy

$$0 = f(x_n + h_n) \approx f(x_n) + h_n f'(x_n)$$

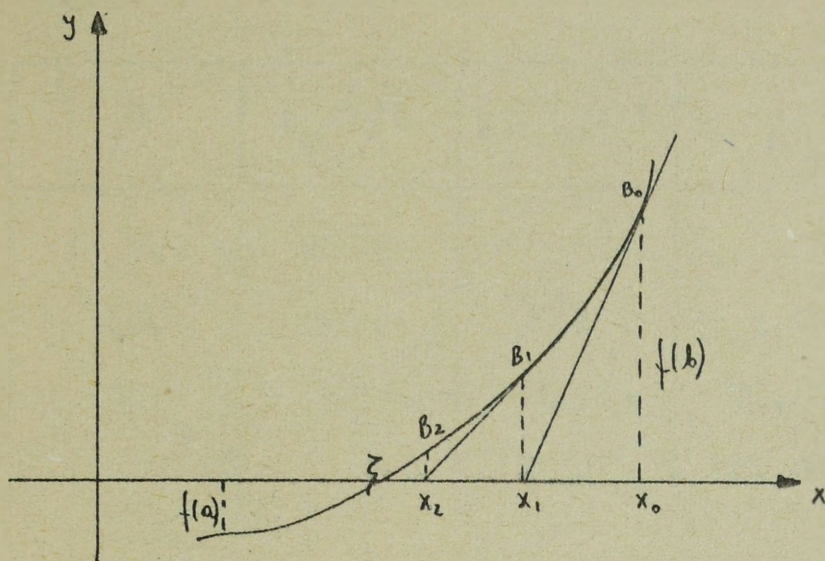
skąd

$$h_n = -\frac{f(x_n)}{f'(x_n)}$$

Otrzymujemy więc

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n = 0, 1, 2, \dots) \quad (\text{IV.4})$$

Interpretacja geometryczna metody Newtona przedstawia się następująco:



Zastępujemy kolejno łuki krzywej $y = f(x)$, B_0B_1 przez styczną w punkcie x_0 , B_1B_2 przez styczną w punkcie x_1 itd.

Do szacowania metody Newtona służy następująca formuła:

$$|\xi - x_n| \leq \frac{|f(x_n)|}{m_1} \quad (\text{IV.5})$$

gdzie $m_1 \geq f'(x)$ w przedziale $[a, b]$.

Przykład:

Znaleźć z dokładnością do pięciu cyfr ujemny pierwiastek równania

$$x^4 - 3x^2 + 75x - 10000 = 0$$

Po zastosowaniu jednej z metod § 2 możemy stwierdzić, że $\xi \in (-11, -10)$. Spełnione są też nierówności $f'(x) < 0$ i $f''(x) > 0$. Możemy więc przyjąć $x_0 = 11$ i wykonać kilka iteracji metodą Newtona

n	x_n	$f(x_n)$	$f'(x_n)$	h_n
0	-11	3453	-5183	0,7
1	-10,3	134,3	-4234	0,03
2	-10,27	37,8	-4196	0,009
3	-10,261	0,2		

Pierwiastkiem równania jest więc $\xi \approx -10,261$.

§ 4. Reguła fałsi

Oprócz metody Newtona bardzo popularną jest metoda zwana regułą fałsi albo metodą siecznych.

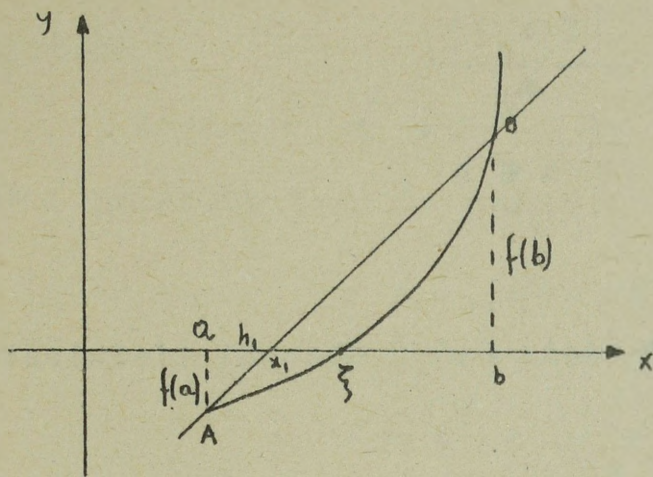
Jeśli mamy zlokalizowany przedział domknięty $[a, b]$, w którym znajduje się pierwiastek ξ równania $f(x) = 0$, tzn. $f(a) \cdot f(b) < 0$, wówczas dzielimy przedział $[a, b]$ na dwie części w stosunku $f(a):f(b)$. Otrzymamy wartość przybliżoną pierwiastka

$$x_1 = a + h_1$$

$$\text{gdzie } h_1 = \frac{-f(a)}{-f(a) + f(b)} (b-a) = -\frac{f(a)}{f(b)-f(a)} (b-a)$$

Następnie stosujemy to samo postępowanie do tego z przedziałów $[a, x_1]$ i $[x_1, b]$, na końcach którego funkcja $f(x)$ ma różne znaki itd.

Geometrycznie oznacza to, że zastępujemy łuk krzywej $y = f(x)$ prostą (cięciwą) łączącą punkty A i B.



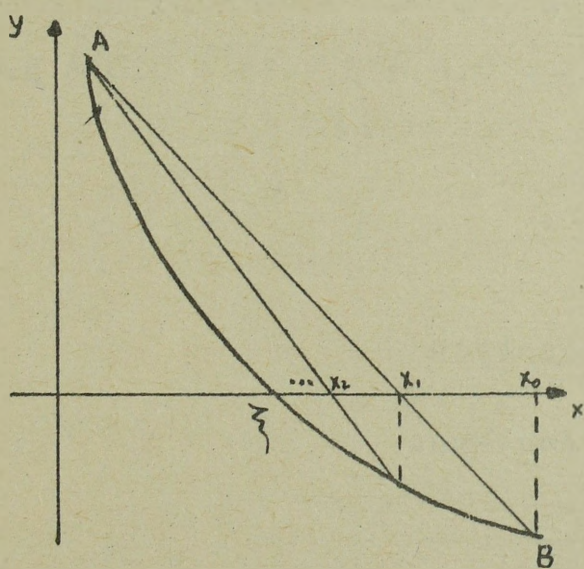
Na podstawie równania prostej przechodzącej przez dwa punkty mamy

$$x_1 = a - \frac{f(a)}{f(b) - f(a)} (b - a) \quad (\text{IV.6})$$

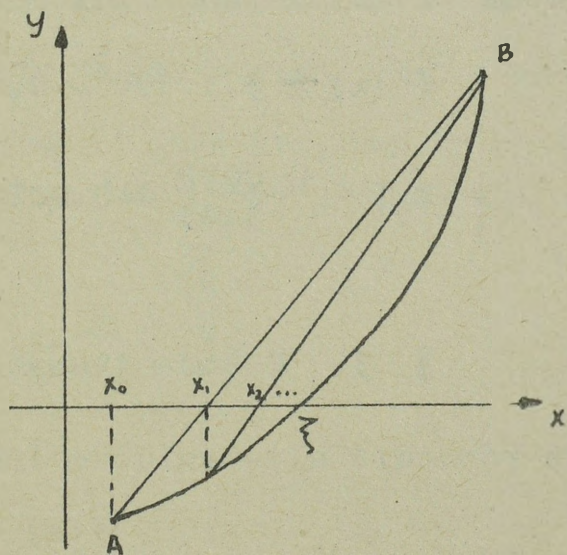
Przy założeniu, że $f''(x) > 0$, krzywa $y = f(x)$ nie posiada punktów przegięcia w przedziale $a \leq x \leq b$ możemy rozpatrzeć dwa przypadki:

1) $f(a) > 0$

2) $f(a) < 0$



$$x_0 = b$$



$$x_0 = a$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f(x_n) - f(a)} (x_n - a) \quad x_{n+1} = x_n - \frac{f(x_n)}{f(b) - f(x_n)} (b - x_n) \quad (\text{IV.7})$$

Metodę siecznych szacujemy za pomocą wzoru:

$$|\xi - x_n| \leq \frac{|f(x_n)|}{m_1} \quad (\text{IV.8})$$

gdzie $|f'(x)| \geq m_1$ dla $a \leq x \leq b$.

Przykład:

Obliczyć z dokładnością 0.002 dodatni pierwiastek równania

$$x^3 - 0.2x^2 - 0.2x - 1.2 = 0$$

Po dokonaniu analizy możemy stwierdzić, że pierwiastek ξ leży w przedziale $[1, 1.5]$. Obliczamy $f(1) = -0.6$ i $f(1.5) = 1.425$.

Stosujemy więc drugi wzór z (IV.7)

$$x_1 = 1 + \frac{0.6}{1.425 + 0.6} (1.5 - 1) = 1 + 0.15 = 1.15$$

$$x_2 = 1.190$$

$$x_3 = 1.198$$

Ponieważ $f'(x) = 3x^2 - 0.4x - 0.2$ dla $x_3 < x < 1.5$ jest

$$f'(x) \geq 3 \cdot 1.198^2 - 0.4 \cdot 1.5 - 0.2 = 3.1.43 - 0.8 = 3.49$$

więc $|\xi - x_3| < \frac{0.0072}{3.43} \approx 0.002$

§ 5. Metoda iteracji prostych

Każde równanie algebraiczne lub przestępne

$$f(x) = 0$$

możemy przekształcić do postaci

$$x = \psi(x) \quad (\text{IV.9})$$

Przekształcenia tego można dokonać różnymi sposobami, uzyskując różne postaci (IV.9).

Przykład

$$\text{Równanie } x^5 + 2x^4 - 3x^3 + 8x^2 - 7x + 11 = 0$$

można przekształcić do postaci

$$x = \frac{1}{7} (x^5 + 2x^4 - 3x^3 + 8x^2 + 11)$$

lub

$$x = x^5 + 2x^4 - 3x^3 + 8x^2 - 6x + 11$$

albo

$$x = \frac{-2x^4 - 8x^2 - 11}{x^4 - 3x^2 - 7}$$

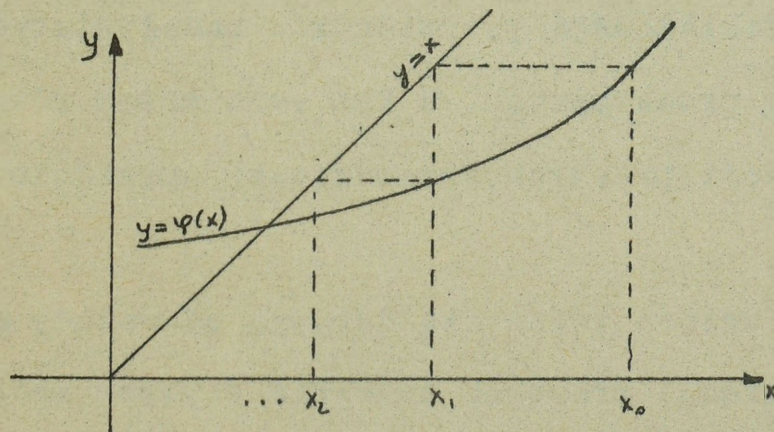
itd.

Po dokonaniu przekształcenia (IV.9) przyjmujemy przybliżenie x_0 możliwie bliskie pierwiastkowi równania $f(x) = 0$ i budujemy ciąg

$$x_{k+1} = \varphi(x_k) \quad (k = 0, 1, 2, \dots) \quad (\text{IV.10})$$

Przy pewnych warunkach ciąg tak otrzymanych liczb x_k jest zbieżny do jednego z pierwiastków równania.

Geometrycznie metoda ta interpretuje pierwiastki równania jako punkty przecięcia prostej $y = x$ oraz krzywej $y = \varphi(x)$



Proces zbieżności tej metody zależy od indywidualnych własności funkcji $f(x)$ i często jest niezbyt szybki.

Przykład

Obliczyć największy pierwiastek dodatni równania

$$x^3 + x = 1000$$

z dokładnością do 10^{-4} .

Jako punkt startowy przyjmijmy $x_0 = 10$.

Dokonajmy przekształcenia

$$x = \sqrt[3]{1000 - x}$$

Wówczas otrzymujemy ciąg

$$x_0 = 10$$

$$x_1 = 9.96655$$

$$x_2 = 9.96666$$

$$x_3 = 9.96667$$

który jest zbieżny i z dokładnością 10^{-4} można przyjąć

$$\xi = 9.96667$$

Metodę iteracji prostych stosuje się zazwyczaj do obliczania pierwiastków z niewielką dokładnością lub w połączeniu z innymi szybko zbieżnymi metodami.

§ 6. Porównanie metod

Metodę siecznych i metodę Newtona stosuje się zwykle do zwiększenia dokładności pierwiastków rzeczywistych równań algebraicznych i przestępnych. W tym celu można również stosować metodę iteracji prostych, sprawdzwszy uprzednio warunki zbieżności.

Metoda Newtona różni się zarówno od metody siecznych jak i metody iteracji prostych znacznie szybszą zbieżnością kwadratową, przy której błąd maleje proporcjonalnie do przybliżania się obliczanego pierwiastka do pierwiastka rzeczywistego. Jednakże na pierwszych etapach, kiedy dokładność przybliżenia jest jeszcze niewielka, metoda Newtona może dawać gorsze

wyniki niż metody siecznych i iteracji prostych.

Niekiedy bardzo pożyteczne jest kombinowanie metody Newtona z metodą siecznych lub iteracji prostych: na pierwszych etapach wykorzystuje się jedną z tych dwóch metod, a przy końcu jeden lub dwa etapy wykonuje się metodą Newtona. Pozwala to, znając $f'(x)$, obliczyć względny i bezwzględny błąd pierwiastka.

Każdy etap metody siecznych, oprócz pierwszego wymaga zwykle w przybliżeniu takiej samej ilości obliczeń co metoda iteracji prostych. Wykonanie jednego etapu metodą Newtona wymaga średnio dwukrotnie więcej obliczeń.

W tych przypadkach, kiedy utrzymanie kilku zbędnych cyfr niezbyt komplikuje obliczenia, celowe jest posługiwanie się metodami przyspieszenia zbieżności. Pozwala to, szczególnie dla równań wysokiego stopnia, lub złożonych równań przestępnych, skrócić ilość obliczeń wielokrotnie.

Rozwiązywanie układów równań algebraicznych liniowych

§ 1. Klasyfikacja metod

Do rozwiązywania układów równań algebraicznych liniowych zapisanych w formie rozwiniętej

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \dots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned} \tag{V.1}$$

lub w formie macierzowej

$$Ax = b \tag{V.2}$$

stosuje się różne metody, które można podzielić na dwie zasadnicze grupy:

1. dokładne,
2. iteracyjne.

Metody dokładne pozwalają w wyniku skończonej, z góry znanej ilości operacji arytmetycznych, otrzymać szukane rozwiązanie. Nazywamy je dokładnymi ze względu na istnienie możliwości otrzymania za ich pomocą rozwiązania dokładnego. Jednakże w praktyce, metody te nie dają wyniku dokładnego wskutek uwzględnienia błędów początkowych oraz błędów zaokrągleń.

Dla metod iteracyjnych charakterystyczne jest to, że wymagają one określenia przybliżeń początkowych nieznanych wielkości, rozwiązania poszukuje się według schematu iteracyjnego, w postaci ciągu stopniowo polepszanych przybliżeń, liczba których może nie być znana wcześniej, dla

uzyskania szukanego rozwiązania wymagane jest aby proces iteracyjny był zbieżny. Metody iteracyjne pozwalają uzyskać rozwiązanie przybliżone z praktycznie dowolnym stopniem dokładności, chociaż dokładnego rozwiązania w skończonej ilości kroków w tym przypadku uzyskać nie można. Dokładny wynik jest granicą nieskończonego ciągu przybliżeń. W praktyce proces iteracyjny trwa dopóty, dopóki dwa kolejne przybliżenia $x^{(k)}$ i $x^{(k+1)}$ nie pokryją się z żadaną ilością cyfr.

Liczne metody rozwiązywania układów (V.2) związane są z odwracaniem macierzy, ponieważ wartości niewiadomych x_i ($i = 1, 2, \dots, n$) spełniają równanie macierzowe

$$x = A^{-1}b,$$

gdzie A^{-1} oznacza macierz odwrotną do macierzy A . Zakłada się przy tym, że macierz współczynników jest nieosobliwa.

§ 2. Metoda Gaussa

Podstawową ideą metody Gaussa, zwanej też metodą eliminacji, jest kolejna eliminacja niewiadomych z równań i doprowadzenie układu (V.1) do formy trójkątnej, wygodnej do bezpośredniego znajdowania wartości niewiadomych.

Na pierwszym etapie tej metody eliminujemy niewiadomą x_1 ze wszystkich równań układu (V.1) poczynając od drugiego. W tym celu wszystkie współczynniki pierwszego równania, łącznie z wyrazem wolnym, dzielimy przez a_{11} po czym otrzymane równanie

$$x_1 + a_{12}^{(1)}x_2 + \dots + a_{1n}^{(1)}x_n = b_1^{(1)}$$

mnożymy kolejno przez $a_{21}, a_{31}, \dots, a_{n1}$ i odejmujemy od drugiego, trzeciego i wszystkich pozostałych równań. W wyniku tego otrzymujemy układ równań $n-1$ rzędu

$$a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \dots + a_{2n}^{(1)}x_n = b_2^{(1)}$$

$$a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 + \dots + a_{3n}^{(1)}x_n = b_3^{(1)}$$

.....

$$a_{n2}^{(1)}x_2 + a_{n3}^{(1)}x_3 + \dots + a_{nn}^{(1)}x_n = b_n^{(1)} \quad (V.3)$$

gdzie $a_{ij}^{(1)} = a_{ij} - a_{i1} \frac{a_{1j}}{a_{11}}, \quad (i, j \geq 2)$

$$b_i^{(1)} = b_i - a_{i1} \frac{b_1}{a_{11}}$$

Na drugim etapie w ten sam sposób eliminujemy niewiadomą x_2 ze wszystkich równań układu (V.3) oprócz pierwszego.

Na k -tym etapie otrzymamy równania

$$x_k + a_{kk+1}^{(k)}x_{k+1} + \dots + a_{kn}^{(k)}x_n = b_k^{(k)}$$

$$a_{k+1, k+1}^{(k)}x_{k+1} + \dots + a_{k+1, n}^{(k)}x_n = b_{k+1}^{(k)}$$

(V.4)

.....

$$a_{n, k+1}^{(k)}x_{k+1} + \dots + a_{nn}^{(k)}x_n = b_n^{(k)}$$

w których poczynając od drugiego nie występują niewiadome $x_j \quad (j = 1, 2, \dots, k)$.

Współczynniki przy niewiadomych i wyrazy wolne układu (V.4) obliczymy ze wzorów:

$$a_{kj}^{(k)} = \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}} \quad (j \geq k+1)$$

$$b_k^{(k)} = \frac{b_k^{(k-1)}}{a_{kk}^{(k-1)}} \quad (V.5)$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - a_{ik}^{(k-1)} a_{kj}^{(k)} \quad (i, j \geq k+1)$$

$$b_i^{(k)} = b_i^{(k-1)} - a_{ik}^{(k-1)} b_k^{(k)} \quad (i \geq k+1)$$

oczywiście przy założeniu, że $a_{kk}^{(k)} \neq 0$.

Po zebraniu wszystkich pierwszych równań otrzymanych na każdym z n etapów uzyskujemy trójkątny układ równań, który jest równoważny układowi (V.1).

$$\begin{aligned} x_1 + a_{12}^{(1)} x_2 + a_{13}^{(1)} x_3 + \dots + a_{1n}^{(1)} x_n &= b_1^{(1)} \\ x_2 + a_{23}^{(2)} x_3 + \dots + a_{2n}^{(2)} x_n &= b_2^{(2)} \\ x_3 + \dots + a_{3n}^{(3)} x_n &= b_3^{(3)} \\ &\dots \dots \dots \\ x_n &= b_n^{(n)} \end{aligned} \quad (V.6)$$

Wartość niewiadomych otrzymujemy z tego układu przez proces odwrotny, który można zapisać

$$x_i = b_i^{(i)} - \sum_{k=i+1}^n a_{ik}^{(i)} x_k \quad (i = n, n-1, \dots, 1) \quad (V.7)$$

Dla lepszego zilustrowania metody Gaussa zapiszmy dla układu czterech równań odpowiednie elementy na poszczególnych etapach w następującej tabelicy

x_1	x_2	x_3	x_4	Wyrazy wolne
a_{11}	a_{12}	a_{13}	a_{14}	a_{15}
a_{21}	a_{22}	a_{23}	a_{24}	a_{25}
a_{31}	a_{32}	a_{33}	a_{34}	a_{35}
a_{41}	a_{42}	a_{43}	a_{44}	a_{45}
1	b_{12}	b_{13}	b_{14}	b_{15}
	a_{22}	a_{23}	a_{24}	a_{25}
	a_{32}	a_{33}	a_{34}	a_{35}
	a_{42}	a_{43}	a_{44}	a_{45}
	1	b_{23}	b_{24}	b_{25}
		a_{33}	a_{34}	a_{35}
		a_{43}	a_{44}	a_{45}
		1	b_{34}	b_{35}
			a_{44}	a_{45}
			1	b_{45}
			1	x_4
		1		x_3
	1			x_2
1				x_1

Błędy działań arytmetycznych oraz błędy okragleń powstałe podczas obliczeń można korygować w następujący sposób:

Założmy, że mamy rozwiązanie x_0 układu

$$Ax = b$$

Przyjmując $x = x_0 + \delta$ gdzie $\delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_n \end{bmatrix}$ jest

poprawką x_0 , otrzymamy układ równań

$$A(x_0 + \delta) = b \text{ lub } A\delta = \xi$$

gdzie $\xi = b - Ax_0$.

Rozwiązując układ równań

$$A\delta = \xi$$

otrzymamy wartości poprawek, które należy dodać do rozwiązania x_0 .

Przykład

Rozwiązać metodą Gaussa układ

$$6x_1 - x_2 - x_3 = 11.33$$

$$-x_1 + 6x_2 - x_3 = 32$$

$$-x_1 - x_2 + 6x_3 = 42$$

z dokładnością do trzech cyfr dziesiętnych, a następnie skorygować pierwiastki tak aby uzyskać dokładność 10^{-4} .

x_1	x_2	x_3	Wyrazy wolne	ϵ
6	-1	-1	11.33	-0.02
-1	6	-1	32	0
-1	-1	6	42	-0.01
=====				
1	-0.167	-0.167	1.89	-0.0033
=====				
	5.83	-1.17	33.9	-0.0033
	-1.17	5.83	43.9	-0.0133
=====				
	1	-0.200	5.80	-0.0006
=====				
		5.60	50.7	-0.0140
=====				
		1	9.05	-0.0025
=====				
	1		7.62	-0.0011
=====				
1			4.67	-0.0039

Poprawione pierwiastki będą wyglądały następująco:

$$x_1 = 9.05 - 0.0025 = 9.0475$$

$$x_2 = 7.62 - 0.0011 = 7.6189$$

$$x_3 = 4.67 - 0.0039 = 4.6661$$

§ 3. Metoda iteracji

Niech będzie dany układ równań

$$A \cdot x = b \tag{V.8}$$

Zakładając $a_{ii} \neq 0$ rozwiążmy kolejne równania względem kolejnych niewiadomych

$$\begin{aligned}
 x_1 &= \beta_1 + \mathcal{L}_{12}x_2 + \mathcal{L}_{13}x_3 + \dots + \mathcal{L}_{1n}x_n \\
 x_2 &= \beta_2 + \mathcal{L}_{21}x_1 + \mathcal{L}_{23}x_3 + \dots + \mathcal{L}_{2n}x_n \\
 &\dots \\
 x_n &= \beta_n + \mathcal{L}_{n1}x_1 + \mathcal{L}_{n2}x_2 + \dots + \mathcal{L}_{nn-1}x_{n-1}
 \end{aligned} \tag{V.9}$$

gdzie $\beta_i = \frac{b_i}{a_{ii}}$ $\alpha_{ij} = -\frac{a_{ij}}{a_{ii}}$ dla $i \neq j$
 $\alpha_{ii} = 0$ dla $i = j$

Układ równań (V.8) można zapisać w postaci

$$x = \beta + \alpha x$$

i rozwiązać według (V.9) przyjmując jako przybliżenie początkowe $x^{(0)} = \beta$. Otrzymamy wówczas

$$\begin{aligned} x^{(1)} &= \beta + \alpha x^{(0)} \\ x^{(2)} &= \beta + \alpha x^{(1)} \\ &\dots \\ x^{(k+1)} &= \beta + \alpha x^{(k)} \end{aligned} \tag{V.10}$$

Jeśli ciąg liczb $x_0, x_1, x_2, \dots, x_k, \dots$ posiada granicę $x = \lim_{k \rightarrow \infty} x^{(k)}$, to jest ona rozwiązaniem układu (V.8).

Główną zaletą tej metody jest prostota schematu obliczeniowego, mała ilość równocześnie zapamiętywanych wyników pośrednich oraz własność autokorekcji. Wadą metody jest natomiast jej powolna zbieżność. Dlatego zaleca się stosowanie tej metody łącznie z metodami przyspieszającymi zbieżność.

§ 4. Metoda Seidela

Metoda ta stanowi pewnego rodzaju modyfikację metody iteracyjnej. Modyfikacja polega na tym, że przy obliczaniu $(k+1)$ -go przybliżenia niewiadomej x_i bierze się pod uwagę obliczone uprzednio $(k+1)$ -sze przybliżenia niewiadomych x_1, x_2, \dots, x_{i-1} .

Zapiszemy układ równań w postaci

$$x_i = \beta_i + \sum_{j=1}^n \alpha_{ij} x_j \quad (i = 1, 2, \dots, n)$$

$$x_1^{(1)} = 1.2 - 0.1 \cdot 0 - 0.1 \cdot 0 = 1.2$$

$$x_2^{(1)} = 1.3 - 0.2 \cdot 1.2 - 0.1 \cdot 0 = 1.06$$

$$x_3^{(1)} = 1.4 - 0.2 \cdot 1.2 - 0.2 \cdot 1.06 = 0.948$$

Zapiszmy wyniki kolejnych kilku iteracji w tablicy

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	1.000	0.0000	0.0000
1	1.200	1.0600	0.9480
2	0.9992	1.0054	0.9991
3	0.9996	1.0001	1.0001
4	1.0000	1.0000	1.0000

Wartościami dokładnymi pierwiastków są $x_1 = 1$, $x_2 = 1$, $x_3 = 1$.

§ 5. Wybór metody

Wybór metody rozwiązywania układu równań algebraicznych liniowych zależy od wielu czynników, które często trudno ze sobą pogodzić. Samo sformułowanie zagadnienia może stawiać metodzie różne wymagania. Ta sama metoda może być bardzo dobra przy rozwiązaniu jednego układu równań, a dająca zupełnie niezadowolające wyniki przy innym układzie. Przy tym istotne jest i to, jaki problem fizyczny lub matematyczny opisuje dany układ równań.

Przy wyborze metody ważną rolę odgrywa stopień uwarunkowania układu. Układ uważamy za dobrze uwarunkowany jeśli niewielkie zmiany współczynników układu i wyrazów wolnych nie powodują dużych zmian w rozwiązaniach tego układu. Warunek ten jest równoważny stwierdzeniu, że przy małych zmianach elementów

macierzy A nieznacznie zmieniają się elementy macierzy odwrotnej A^{-1} . Jeśli małe zmiany macierzy A prowadzą do istotnych zmian macierzy odwrotnej A^{-1} , to układ nazywamy źle uwarunkowanym.

Często uwarunkowanie układu równań ocenia się wielkością wyznacznika $|A|$. Taka ocena nie może jednak w pełni charakteryzować stopnia uwarunkowania układu. Znacznie doskonalszą metodą oceny uwarunkowania jest metoda oceny wielkości elementów macierzy odwrotnej A^{-1} . Jeśli te elementy są duże, to układ jest źle uwarunkowany. Innymi kryteriami oceny uwarunkowania układu są liczby

$$M = n \max_{ij} |a_{ij}| \max_{ij} |a_{ij}^{-1}|;$$

lub
$$P = \frac{\max |\lambda_i|}{\min |\lambda_i|}$$

gdzie λ_i są wartościami własnymi macierzy A .

Liczby te nazywamy liczbami uwarunkowania układu równań.

Na wybór metody ma również bardzo duży wpływ stopień układu równań, błąd dopuszczalny, złożoność schematu obliczeniowego itp. Nie może więc istnieć jedna uniwersalna metoda dla wszystkich układów równań. Nawet duża ilość różnych metod nie chroni przed różnymi niespodziankami spotykanymi w praktyce.

Rozdział VI

Wyznaczanie wartości własnych i wektorów własnych.

§ 1. Sformułowanie zagadnienia.

Klasyfikacja metod

Ogromna ilość zagadnień z różnych dziedzin nauki zarówno teoretycznych, jak i praktycznych, wymaga wyznaczania wartości własnych i wektorów własnych macierzy określonych w sposób następujący:

Wiadomo, że jednorodny układ równań algebraicznych liniowych zapisany w formie macierzowej

$$AX = \lambda X \quad (\text{VI.1})$$

posiada rozwiązanie niezerowe wtedy i tylko wtedy, kiedy wielkość skalarna λ stanowi pierwiastek równania algebraicznego

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = \\ &= (-1)^n [\lambda^n + b_1 \lambda^{n-1} + \dots + b_{n-1} \lambda + b_n] = 0 \quad (\text{VI.2}) \end{aligned}$$

gdzie A - macierz kwadratowa o współczynnikach a_{ij} ,

I - macierz jednostkowa,

b_i - współczynniki wielomianu otrzymanego z rozwinięcia wyznacznika (VI.2).

Równanie (VI.2) nazywamy równaniem charakterystycznym, a jego pierwiastki λ_i ($i = 1, 2, \dots, n$) - wartościami własnymi macierzy A . Wektor X , będący nietrywialnym rozwiązaniem układu (VI.1), nazywamy wektorem własnym macierzy A .

Każdej wartości własnej λ_i odpowiada wektor własny $X^{(i)}$, spełniający równanie

$$(A - \lambda_i I)X^{(i)} = 0.$$

Do wyznaczania wartości własnych i wektorów własnych macierzy stosuje się różne metody, które można podzielić na dwie zasadnicze grupy:

- 1/ metody pozwalające rozwiązywać pełne zagadnienie własne, tzn. pozwalające wyznaczyć wszystkie wartości własne i wektory własne;
- 2/ metody pozwalające na częściowe rozwiązywanie zagadnienia własnego, tzn. pozwalające wyznaczyć jedną lub kilka wartości własnych i odpowiadających im wektorów własnych.

Oprócz takiego podziału metod wyznaczania wartości własnych i wektorów własnych istnieje też inny podział, w którym wyróżnia się:

- 1/ metody pośrednie, związane z poszukiwaniem współczynników wielomianu charakterystycznego (VI.2), a następnie wyznaczaniem jego pierwiastków;
- 2/ metody bezpośrednie, pozwalające wyznaczać wartości własne i odpowiadające im wektory własne bez rozwijania wyznacznika (VI.2).

Pierwsze z tych metod często nazywamy metodami dokładnymi, a drugie metodami iteracyjnymi. Przez metody dokładne rozumiemy takie metody, które przy dokładnych wartościach elementów macierzy i dokładnych obliczeniach pozwalają znaleźć dokładne wartości współczynników wielomianu charakterystycznego. W metodach iteracyjnych wartości własne i składowe odpowiadających im wektorów własnych są granicami

pewnych ciągów liczbowych.

§ 2. Metody dokładne

W paragrafie tym omówimy dwie najczęściej stosowane metody pozwalające na obliczanie współczynników wielomianów charakterystycznych.

1. Metoda Kryłowa

Podstawowa idea metody Kryłowa polega na tym, że za pomocą szeregu przekształceń wyznacznik (VI.2) sprowadza się do postaci

$$\begin{vmatrix} c_{11} - \lambda & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} - \lambda & \dots & c_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{vmatrix}$$

gdzie λ występuje tylko w pierwszej kolumnie. Wyznacznik taki jest wygodniejszy z tego względu, że pozwala obliczać współczynniki wielomianu charakterystycznego b_1 bez obliczania kolejnych minorów macierzy. Do tego celu wykorzystuje się twierdzenie Cayley'a - Hamiltona, dzięki któremu otrzymuje się układ n równań algebraicznych liniowych z n niewiadomymi współczynnikami wielomianu charakterystycznego. Układ ten można zapisać w formie równania wektorowego

$$C_n + C_{n-1}b_1 + C_{n-2}b_2 + \dots + C_1b_{n-1} + C_0b_n = 0 \quad (\text{VI.3})$$

gdzie

$$C_1 = A^1 C_0 \quad (\text{VI.4})$$

gdzie C_0 - dowolny wektor początkowy.

W taki sposób wyznaczenie wartości własnych macierzy A sprowadza się do rozwiązania układu równań algebraicznych

liniowych (VI.3) oraz wyznaczenia pierwiastków wielomianu charakterystycznego (VI.2).

Dokładność otrzymanych tą metodą wartości własnych w dużym stopniu zależy od dokładności uzyskanych współczynników wielomianu. Dlatego też przy rozwiązywaniu układu (VI.3) należy posługiwać się metodami dającymi bardzo dokładne wyniki.

Metoda Kryłowa pozwala po znalezieniu C_i , b_i i λ_i obliczyć wektory własne $X^{(i)}$ z następującego wzoru

$$X^{(i)} = \sum_{j=0}^{n-1} \alpha_{ij} C_{n-1-j} \quad (\text{VI.5})$$

gdzie α_{ij} są określone za pomocą formuły rekurencyjnej

$$\alpha_{i0} = 1 \quad \alpha_{ij} = \lambda_i \alpha_{i,j-1} - b_j \quad (j = 1, 2, \dots, n-1).$$

Metodę Kryłowa stosuje się najczęściej dla macierzy niezbyt wysokiego rzędu, kiedy z niedużą dokładnością chcemy znaleźć wszystkie wartości i wektory własne.

2. Metoda Danilewskiego

Metoda Danilewskiego polega na sprowadzeniu macierzy wyjściowej $A = A^{(1)}$ za pomocą $(n-1)$ przekształceń podobieństwa do tzw. macierzy Frobeniusa.

$$A^{(n)} = \begin{pmatrix} 0 & 0 & 0 & \dots & b_n \\ 1 & 0 & 0 & \dots & b_{n-1} \\ 0 & 1 & 0 & \dots & b_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & b_1 \end{pmatrix} \quad (\text{VI.6})$$

Macierz ta zawiera w ostatniej kolumnie w jawnej postaci szukane współczynniki wielomianu charakterystycznego (VI.2).

Elementy macierzy $A^{(k+1)}$ otrzymanej na k -tym etapie przekształceń oblicza się z elementów macierzy $A^{(k)}$ na podstawie wzorów

$$c_{k+1,j}^{(k)} = \frac{1}{a_{k+1,k}^{(k)}} \cdot a_{k+1,j}^{(k)}$$

$$c_{ij}^{(k)} = a_{ij}^{(k)} - a_{ik}^{(k)} c_{k+1,j}^{(k)} \quad (i \neq k+1)$$

$$a_{ij}^{(k+1)} = c_{ij}^{(k)} \quad (j \neq k+1) \quad (\text{VI.7})$$

$$a_{i,k+1}^{(k+1)} = \sum_{j=1}^n c_{ij}^{(k)} a_{jk}^{(k)}$$

gdzie $c_{ij}^{(k)}$ i $c_{k+1,j}^{(k)}$ oznaczają elementy macierzy pomocniczej C .

W celu eliminacji przypadków zwyrodnienia procesu (co może mieć miejsce przy $a_{n+1,k}^{(k)}$ bliskim zeru) i w celu zwiększenia dokładności obliczeń, jako $a_{k+1,k}^{(k)}$ wybiera się największy co do modułu element $a_{jk}^{(k)}$ przy $j \geq k+1$. Dokonuje się równocześnie w macierzy $A^{(k)}$ przestawienia $(k+1)$ -go i j -go wiersza oraz j -tej i $(k+1)$ -szej kolumny.

Metoda Danilewskiego pozwala również obliczyć na podstawie znajomości wartości własnych λ_i odpowiadające im wektory własne $X^{(i)}$.

$$y_i = \lambda^{n-i} - \sum_{j=1}^{n-1} b_j^{n-i-j} \quad (i = n, n-1, \dots, 2, 1)$$

$$y_i^{(k)} = y_i^{(k+1)} + y_{k+1}^{(k+1)} a_{ik}^{(k)} \quad (i \neq k+1) \quad (\text{VI.8})$$

$$y_{k+1}^{(k)} = y_{k+1}^{(k+1)} a_{k+1,k}^{(k)}$$

gdzie $y_i = y_i^{(n)}$, $y_i^{(1)} = X^{(1)}$, $k = n-1, n-2, \dots, 1$

Metodę Danilewskiego stosuje się w tych przypadkach, kiedy należy znaleźć z dużą dokładnością wszystkie wartości i wektory własne macierzy A , której rząd jest nie większy niż 20.

§ 3. Metody iteracyjne

Omówimy tu jedną z najczęściej stosowanych metod tzw. metodą obrotów. Metoda ta dotyczy wprowadzie macierzy symetrycznych, ale z macierzami takimi w praktyce spotykamy się najczęściej.

Idea metody obrotów zawiera się w tym, że drogą ciągu przekształceń podobieństwa macierz symetryczną A sprowadza się do macierzy diagonalnej, gdzie wszystkie jej wartości własne występują w jawnej postaci. Wspomniany ciąg przekształceń daje nam w wyniku ciąg macierzy

$A = A^{(0)}, A^{(1)}, A^{(2)}, \dots$, z których każdą otrzymuje się z poprzedniej w wyniku zamiany dwóch wierszy i dwóch kolumn.

W ten sposób ze wzrostem indeksu k macierz $A^{(k)}$ coraz bardziej przybliża się do macierzy diagonalnej. Miarą tego przybliżenia może być suma kwadratów elementów poza-diagonalnych.

Proces przekształceń dokonuje się za pomocą elementarnych macierzy T_{ij} zwanych macierzami obrotów.

$$c = \sqrt{\frac{1}{2} \left(1 + \frac{|a_{ii}^{(k)} - a_{jj}^{(k)}|}{d} \right)};$$

$$d = \sqrt{(a_{ii}^{(k)} - a_{jj}^{(k)})^2 + 4(a_{ij}^{(k)})^2};$$

$$s = \text{sign } a_{ij}^{(k)} (a_{ii}^{(k)} - a_{jj}^{(k)}) \sqrt{\frac{1}{2} \left(1 - \frac{|a_{ii}^{(k)} - a_{jj}^{(k)}|}{d} \right)}$$

Następnie znajdujemy macierz

$$A^{(k+1)} = T_{ij}^{(k)} A^{(k)} T_{ij}^{(k)} \tag{VI.11}$$

$$a_{ti}^{(k+1)} = ca_{ti}^{(k)} + sa_{tj}^{(k)} \quad (t = 1, 2, \dots, n)$$

$$a_{tj}^{(k+1)} = -sa_{ti}^{(k)} + ca_{tj}^{(k)}.$$

Obliczone według formuł (VI.11) elementy i -tej i j -tej kolumny wstawia się na miejsce odpowiednich kolumn macierzy $A^{(k)}$. Analogicznie postępuje się z wierszami, które obliczamy ze wzorów:

$$a_{it}^{(k+1)} = ca_{it}^{(k)} + sa_{jt}^{(k)}$$

$$a_{jt}^{(k+1)} = -sa_{it}^{(k)} + ca_{jt}^{(k)} \quad (t = 1, 2, \dots, n)$$

$$a_{ij}^{(k+1)} = a_{ji}^{(k+1)} = 0$$

Następne kroki tworzy się podobnie jak dotychczas dopóty, dopóki nie przekroczą wszystkich granic. Diagonalne elementy macierzy otrzymanej w ostatnim kroku takiego postępowania będą wartościami własnymi macierzy A .

Metodę obrotów stosujemy najczęściej w tych przypadkach, kiedy:

- 1/ poszukujemy wszystkich wartości własnych macierzy symetrycznej A z bardzo dużą dokładnością;
- 2/ macierz A jest macierzą wysokiego rzędu.

Rozdział VII

Całkowanie numeryczne

§ 1. Sformułowanie zagadnienia
całkowania numerycznego

Zagadnienie przybliżonego obliczenia pojedynczej całki oznaczonej można sformułować w następujący sposób:

Należy obliczyć wartość całki

$$\int_a^b f(x) dx \quad (\text{VII.1})$$

jeśli znane są wartości funkcji podcałkowej $f(x)$ oraz wartości jej pochodnych $f'(x)$, $f''(x)$, ..., $f^{(i)}(x)$, w ustalonych punktach x_1, x_2, \dots, x_n przedziału całkowania $[a, b]$. Punkty x_1, x_2, \dots, x_n nazywamy często węzłami całkowania. Formuły pozwalające obliczyć przybliżoną wartość całki jako funkcję $f'(x_k)$, $f''(x_k)$, ..., $f^{(i)}(x_k)$, gdzie $k = 1, 2, 3, \dots, n$ nazywamy formułami kwadraturowymi lub krótko kwadraturami.

Najprostszymi są formuły kwadraturowe, w których całka wyraża się za pomocą liniowej kombinacji $f(x)$ w punktach węzłowych x_k , tzn. formuły postaci

$$\int_a^b f(x) dx = \sum_{k=1}^n A_k f(x_k) + R_n \quad (\text{VII.2})$$

gdzie A_1, A_2, \dots, A_n są pewnymi stałymi współczynnikami niezależnymi od $f(x)$, a R_n oznacza dokładność formuły kwadraturowej, równą różnicy między dokładną wartością całki a sumą stojącą po prawej stronie (VII.2).

W praktyce człon R_n z reguły odrzuca się. Formułę kwadraturową nazywamy dokładną jeśli $R_n = 0$, tzn. jeśli suma stojąca po prawej stronie (VII.2) daje dokładną wartość całki.

Formuła (VII.2) jest niezbyt wygodna ze względu na to,

że współczynniki A_k oraz węzły x_k zależą od przedziału całkowania $[a, b]$. W celu uwolnienia się od tej zależności przez podstawienie $x = a + (b-a)t$ zamieniamy przedział $[a, b]$ na przedział $[0, 1]$.

Wówczas

$$\begin{aligned} \int_a^b f(x) dx &= (b-a) \int_0^1 f[a + (b-a)t] dt = \\ &= \sum_{k=1}^n A_k f[a + (b-a)t_k] + R_n \end{aligned} \quad (\text{VII.3})$$

albo

$$\int_0^1 f[a + (b-a)t] dt = \sum_{k=1}^n C_k f[a + (b-a)t_k] + \frac{R}{b-a}$$

gdzie $C_k = \frac{A_k}{b-a}$, $0 \leq t_k \leq 1$ ($k = 1, 2, \dots, n$)

Współczynniki C_k oraz węzły t_k nie zależą od przedziału całkowania. Ogólnie możemy teraz zapisać

$$\int_a^b f(x) dx = (b-a) \sum_{k=1}^n C_k f[a + (b-a)t_k] + R_n \quad (\text{VII.4})$$

W praktyce posługujemy się zwykle formułami kwadratowymi postaci (VII.4). Różne formuły kwadraturowe tego rodzaju różnią się tylko liczbą n oraz sposobem wyboru współczynników C_k i węzłów t_k .

§ 2. Metoda Newtona-Cotesa i jej szczególne przypadki

Metoda Newtona-Cotesa należy do grupy metod zwanych interpolacyjnymi. W formułach tego typu funkcję podcałkową aproksymuje się wielomianem

$$P_{n-1}(x) = a_0 x^{n-1} + a_1 x^{n-2} + \dots + a_{n-2} x + a_{n-1}$$

który w punktach węzłowych przyjmuje te same wartości co funkcja podcałkowa $f(x)$, tzn. spełnia warunek

$$P_{n-1}(x_k) = f(x_k) \quad (k = 1, 2, \dots, n).$$

Wielomian interpolacyjny tworzy się za pomocą wzoru interpolacyjnego Lagrange'a. Formułę kwadraturową uzyskujemy przez całkowanie wielomianu $P_{n-1}(x)$.

W formułach Newtona-Cotesa

$$\int_a^b f(x) dx = (b-a) \sum_{k=1}^n C_k f[a + (b-a)t_k] + R_n, \quad (\text{VII.5})$$

n może przyjmować różne wartości w zależności od postaci funkcji podcałkowej. Przy wybranym n

$$t_k = \frac{k-1}{n-1} \quad (k = 1, 2, \dots, n), \quad (\text{VII.6})$$

co odpowiada podziałowi przedziału całkowania na $n-1$ równych części oraz aproksymacji funkcji podcałkowej wielomianem stopnia $n-1$. Współczynniki C_k we wzorze (VII.5) wyrażają się następująco:

$$C_k = \frac{\int_0^1 \prod_{\substack{i=1 \\ i \neq k}}^n (z - \frac{i-1}{n-1}) dz}{\prod_{\substack{i=1 \\ i \neq k}}^n \frac{k-i}{n-1}} \quad (\text{VII.7})$$

Tak obliczone współczynniki C_k charakteryzują się następującymi własnościami:

1/ są liczbami wymiernymi

$$2/ \sum_{k=1}^n C_k = 1$$

3/ $C_{n+1-k} = C_k$ tzn. że współczynniki odpowiadające węzłom położonym symetrycznie względem środka przedziału całkowania są sobie równe.

Błędy formuł Newtona-Cotesa wyrażają się wzorami

$$R_n = \frac{f^{(n+1)}(\xi)}{(n+1)} \int_a^b w(x) dx \quad \text{przy } n \text{ nieparzystym} \quad (\text{VII.8})$$

$$R_n = \frac{f^{(n)}(\xi)}{n} \int_a^b w(x) dx \quad \text{przy } n \text{ parzystym}$$

gdzie $a < \xi < b$ a $w(x) = (x-x_1)(x-x_2)\dots(x-x_n)$.

Wynika stąd, że formuły Newtona-Cotesa będą dokładne jeżeli funkcja podcałkowa będzie wielomianem stopnia $n-1$ przy n parzystym oraz stopnia n przy n nieparzystym.

Przyjmując dla formuły (VII.5) konkretne wartości otrzymamy:

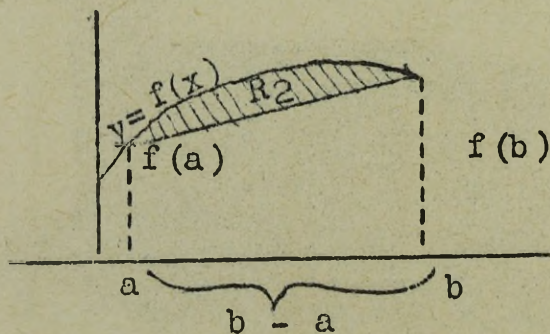
1/ Dla $n = 2$ metodą trapezów dla której

$$t_1 = 0, \quad t_2 = 1, \quad c_1 = c_2 = \frac{1}{2}$$

W tym przypadku funkcję podcałkową aproksymuje się wielomianem stopnia pierwszego

$$P_1(x) = f(a) + \frac{f(b) - f(a)}{b - a} (x - a),$$

co geometrycznie oznacza, że pole powierzchni figury krzywoliniowej zastępujemy polem trapezu prostokątnego o podstawach $f(a)$ i $f(b)$ oraz wysokości $b-a$.



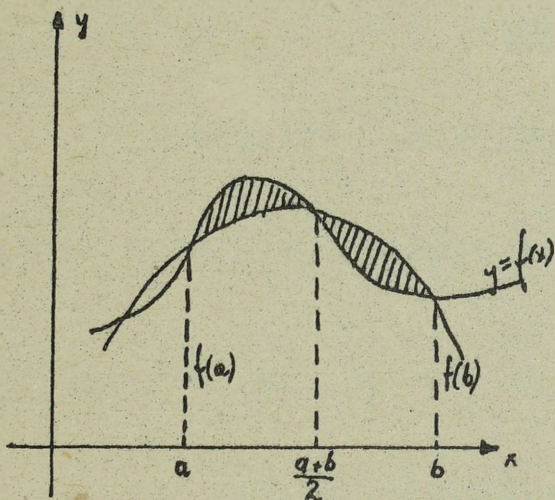
Formułę trapezów można zapisać w następujący sposób:

$$\int_a^b f(x) dx = \frac{b-a}{2} [f(a) + f(b)] + R_2 \quad (\text{VII.9})$$

gdzie $R_2 = -\frac{(b-a)^3}{12} f''(\xi)$, $a < \xi < b$

2/ Gdy $n = 3$ otrzymujemy metodę Simpsona, dla której
 $t_1 = 0$, $t_2 = \frac{1}{2}$, $t_3 = 1$, $c_1 = c_3 = \frac{1}{6}$, $c_2 = \frac{2}{3}$.

W tym przypadku funkcję podcałkową aproksymuje się wielomianem stopnia drugiego, czyli zastępuje się krzywą $y = f(x)$ odcinkiem paraboli.



Formuła Simpsona wyraża się więc wzorem

$$\int_a^b f(x) dx = \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] + R_3 \quad (\text{VII.10})$$

gdzie $R_3 = -\frac{(b-a)^5}{180 \cdot 2^4} f^{(IV)}(\xi)$, $a < \xi < b$

Przykład

Za pomocą formuły Simpsona obliczyć całkę

$$I = \int_0^1 \frac{dx}{1+x}$$

$$\int_0^1 \frac{dx}{1+x} \approx \frac{1-0}{6} \left[\frac{1}{1} + 4 \cdot 0,66667 + \frac{1}{2} \right] = \frac{1}{6} [1 + 2,66668 + 0,5] = \\ = \frac{1}{6} \cdot 4,16668 = 0,69444$$

W celu zwiększenia dokładności możemy przedział całkowania podzielić na pewną ilość odcinków i do nich stosować metodę Simpsona. Zakładając, że przedział całkowania podzielimy punktami $x_0, x_1, x_2, \dots, x_{10}$ na 10 równych odcinków, wówczas formułę Simpsona możemy zapisać

$$\int_a^b f(x) dx = \frac{b-a}{3 \cdot 10} \left[f(x_0) + f(x_{10}) + 4(f(x_1) + f(x_3) + f(x_5) + f(x_7) + f(x_9)) \right. \\ \left. + 2(f(x_2) + f(x_4) + f(x_6) + f(x_8)) \right]$$

Stosując tę formułę do przykładu otrzymamy

$$\int_0^1 \frac{dx}{1+x} = \frac{1}{30} [1 + 0,5 + 4 \cdot 3,45955 + 2 \cdot 2,72818] = 0,69315$$

Dokładna wartość tej całki wynosi $\ln 2 \approx 0,693114118 \dots$

§ 3. Metoda Gaussa

W celu skonstruowania formuły całkowania Gaussa

$$\int_{-1}^1 f(t) dt = \sum_{k=1}^n c_k f(t_k) \quad (\text{VII.11})$$

współczynniki c_k oraz odcięte t_k wybieramy tak, aby była ona dokładna dla dowolnego wielomianu algebraicznego stopnia nie większego niż $2n - 1$. Wybór taki jest optymalny i może być dokonany dla dowolnego n . Odcięte Gaussa t_k związane są z pierwiastkami z_k wielomianu Legendre'a

$$P_n(z) = \frac{1}{2^n n!} \frac{d^n}{dz^n} [(z^2 - 1)^n] \quad (\text{VII.12})$$

zależnością

$$t_k = \frac{1}{2} (1 + z_k), \quad (\text{VII.13})$$

a współczynniki Gaussa c_k określa się wzorem

$$c_k = \frac{1}{(1-z_k^2) [P_n'(z_k)]^2} \quad (\text{VII.14})$$

Odcięte t_k i współczynniki c_k występujące w formule Gaussa charakteryzują się następującymi własnościami:

1/ $t_{n+1-k} = 1-t_k$, tzn. odcięte t_k są rozłożone symetrycznie względem punktu $t = \frac{1}{2}$

2/ $c_{n+1-k} = c_k$, tzn. współczynniki odpowiadające symetrycznie rozłożonym węzłom są sobie równe.

3/ $\sum_{k=1}^n c_k = 1$

Błąd formuły Gaussa wyraża się wzorem

$$R_n = M_n (b-a)^{2n+1} f^{(2n)}(\xi), \quad \text{gdzie } M_n = \frac{(n!)^4}{(2n+1) [(2n)!]^3};$$

$$a < \xi < b \quad (\text{VII.15})$$

Zaletą formuły Gaussa jest to, że daje ona dość dużą dokładność, przy stosunkowo małej liczbie węzłów, nawet jeśli na końcach przedziału całkowania funkcja podcałkowa jest nieokreślona.

Przykład

Obliczyć wartość całki

$$I = \int_0^1 \frac{dx}{1+x}$$

metodą Gaussa przyjmując $n = 2$ oraz $n = 3$.

Aby zastosować formułę Gaussa należy w pierwszej kolejności dokonać przekształcenia całki I za pomocą podstawienia

$$x = \frac{1}{2} (b+a) + \frac{1}{2} (b-a)t$$

gdzie a i b oznaczają granice całkowania $(0,1)$.

Dla $n = 2$ otrzymujemy wielomian Legendre'a

$$P_2(z) = \frac{1}{2}(3z^2 - 1),$$

którego pierwiastkami są

$$z_1 = 0,577350$$

$$z_2 = 0,577350$$

Na podstawie (VII.13) i (VII.14) otrzymamy

$$t_1 = 0,788675$$

$$t_2 = 0,211325$$

oraz

$$c_1 = c_2 = 0,5$$

Możemy więc na podstawie (VII.11) obliczyć

$$\int_0^1 \frac{dt}{1+t} = \sum_{k=1}^2 c_k f(t_k) =$$

$$= 0,5 \frac{1}{1+0,788675} + 0,5 \frac{1}{1+0,211325} = 0,6923076$$

Dla $n = 3$ otrzymamy odpowiednio

$$t_1 = 0,11270$$

$$t_2 = 0,50000$$

$$t_3 = 0,88730$$

$$c_1 = 0,27778$$

$$c_2 = 0,44444$$

$$c_3 = 0,27778$$

$$\int_0^1 \frac{dt}{1+t} = 0.27778 \cdot \frac{1}{1+0.11270} + 0.44444 \frac{1}{1+0.5} + \\ + 0.27778 \frac{1}{1+0.88730} = 0.693122$$

§ 4. Wybór metody

Najdokładniejszą z opisanych w niniejszym rozdziale metod całkowania numerycznego jest metoda Gaussa; jest ona bowiem oparta na wielomianie interpolacyjnym stopnia $2n-1$, podczas gdy w metodzie Newtona-Cotesa występuje wielomian stopnia co najwyżej n . Pewną wadą formuły Gaussa jest trudność szacowania błędu przez szacowanie wartości pochodnych wysokich rzędów.

W przypadku gdy dobór wartości zmiennej niezależnej w równych odstępach wyraźnie ułatwia obliczanie wartości funkcji, korzystamy zazwyczaj ze wzorów Newtona-Cotesa, a w szczególności z formuły trapezów lub Simpsona.

Metodę Newtona-Cotesa można stosować dwojako. Albo dobierać wzory możliwie dokładne dla całego przedziału całkowania, albo dzielić cały przedział na części i stosować do każdej z tych części wzory niższego rzędu. Jeśli weźmiemy pod uwagę, że trudność szacowania błędu w metodzie Newtona-Cotesa wzrasta bardzo szybko ze wzrostem wartości n , stwierdzimy, że najwygodniej jest stosować metodę dla małych wartości n , ale za to nie do całego przedziału całkowania, lecz do jego części.

Metodę Newtona-Cotesa dla $n > 2$ stosuje się zazwyczaj w przypadku wartości funkcji podcałkowych danych z doświadczenia, które z różnych względów są dane w równoległych punktach. Zastosowanie formuły Gaussa jest wówczas niemożliwe.

Oprócz podanych tu metod w praktyce spotyka się również formułę Czebyszewa, metodę szeregów potęgowych i inne.

Rozdział VIII

Rozwiązywanie równań różniczkowych zwyczajnych

§ 1. Uwagi ogólne

Równaniem różniczkowym zwyczajnym rzędu n nazywamy wyrażenie

$$F(x, y, y', y'', \dots, y^{(n)}) = 0 \quad (\text{VIII.1})$$

gdzie x jest zmienną niezależną, a y funkcją niewiadomą.

Wiadomo, że rozwiązanie ogólne równania (VIII.1) można zapisać w sposób następujący

$$G(x, y, C_1, C_2, \dots, C_n) = 0 \quad (\text{VIII.2})$$

gdzie C_1, C_2, \dots, C_n oznaczają dowolne stałe.

Rozwiązania, otrzymane z rozwiązania ogólnego (VIII.2) przy ustalonych wartościach C_1, C_2, \dots, C_n nazywamy rozwiązaniem szczególnym równania (VIII.1).

Aby otrzymać z rozwiązania ogólnego rozwiązanie szczególne, należy zadać n warunków, które pozwolą jednocześnie określić C_1, C_2, \dots, C_n . Warunki te można postawić określając wartości poszukiwanej funkcji i jej pochodnych w pewnym punkcie x_0 , tzn. podając

$$y_0 = y(x_0), y'_0 = y'(x_0), \dots, y_0^{(n)} = y^{(n)}(x_0). \quad (\text{VIII.3})$$

Tak postawione zagadnienie całkowania równań różniczkowych nazywamy zagadnieniem Cauchy'ego, a warunki (VIII.3), warunkami początkowymi.

W teorii równań różniczkowych zwyczajnych rozwiązanie przedstawia się za pomocą pewnego wyrażenia analitycznego $y = y(x)$, które podstawione do równania (VIII.1) spełnia je tożsamościowo.

Rozwiązanie równania różniczkowego (VIII.1) metodami

przybliżonymi oznacza znalezienie dla danego ciągu argumentów $x_0 < x_1 < x_2 < \dots < x_m$ wartości funkcji $y_k = y(x_k)$

($k = 0, 1, 2, \dots, m$), będących analitycznym rozwiązaniem (VIII.1). Rozwiązanie analityczne dla funkcji $y = y(x)$, z zasady nieznane, może nie być znalezione w skończonej formie.

Wielkość $h_k = x_{k+1} - x_k$ nazywamy krokiem całkowania.

Krok całkowania nazywamy przemiennym jeśli dla różnych k wielkości h_k są różne. Jeśli $h_k = \text{const.}$, mówimy, że całkowanie odbywa się ze stałym krokiem i $x_k = x_0 + kh$ ($k = 0, 1, 2, \dots, m$).

Metody numerycznego rozwiązywania zagadnienia Cauchy'ego najbardziej zostały rozwinięte dla równań różniczkowych, rozwikłanych względem najwyższej pochodnej, tzn. dla równań

$$y^{(n)} = f(x, y, y', y'', \dots, y^{(n-1)}) \quad (\text{VIII.4})$$

Jeśli mamy równanie dane w postaci (VIII.1), to posługując się metodami rozwiązywania równań z jedną niewiadomą, można je zawsze sprowadzić do postaci (VIII.4).

W zasadzie wszystkie formuły numerycznego rozwiązywania zagadnienia Cauchy'ego są formułami rekurencyjnymi, pozwalają bowiem na podstawie jednej lub kilku wartości funkcji i jej pochodnych obliczyć kolejną wartość funkcji. Formuły rekurencyjne równań różniczkowych w przeciwieństwie do innych charakteryzują się tym, że są mało dokładne, a na każdym kroku obliczeń dokładność zależy od metody, kroku całkowania oraz własności szukanej funkcji $y = f(x)$.

W dalszym ciągu rozpatrzemy kilka metod rozwiązywania równań różniczkowych 1-go rzędu.

§ 2. Metody rozwiązywania równań różniczkowych

1-go rzędu

Niech będzie dane równanie

$$y' = f(x, y) \quad (\text{VIII.5})$$

Dla zadanych wartości argumentu x_1, x_2, \dots, x_m założonych w porządku rosnącym, należy obliczyć wartości $y_k = y(x_k)$ ($k = 1, 2, \dots, m$) funkcji $y = y(x)$ będącej rozwiązaniem równania (VIII.5), przy założeniu, że

$$y_0 = y(x_0) \quad (\text{VIII.6})$$

1/ Metoda potęgowa

Założmy, że szukana funkcja $y = y(x)$ posiada ciągłe pochodne aż do rzędu $p+1$ włącznie. Możemy wówczas $y = y(x)$ rozłożyć w szereg Taylora w otoczeniu punktu $x = x_k$:

$$y(x_k + h) = y_k + \frac{h}{1!} y'_k + \frac{h^2}{2!} y''_k + \dots + \frac{h^p}{p!} y_k^{(p)} + \frac{h^{p+1}}{(p+1)!} y^{(p+1)}(x_k + \theta h); \quad (\text{VIII.7})$$

gdzie $0 < \theta < 1$.

Ponieważ pochodna $y^{(p+1)}$ jest ciągła, a więc ograniczona, to można dobrać tak małe h , że ostatni człon rozwinięcia (VIII.7) można zaniedbać.

Wówczas

$$y(x_k + h) = y_k + \frac{h}{1!} y'_k + \frac{h^2}{2!} y''_k + \dots + \frac{h^p}{p!} y_k^{(p)} \quad (\text{VIII.8})$$

Jeśli krok $h_k = x_{k+1} - x_k$, z którym odbywa się całkowanie równania (VIII.5) jest dostatecznie mały, to przyjmując $h = h_k$ otrzymamy

$$y_{k+1} = y_k + \frac{h_k}{1!} y'_k + \frac{h_k^2}{2!} y''_k + \dots + \frac{h_k^p}{p!} y_k^{(p)} \quad (\text{VIII.9})$$

Wzór ten wykorzystuje się do całkowania równania (VIII.5) przyjmując kolejno $k = 0, 1, 2, \dots, m$. Kolejne pochodne $y', y'', \dots, y^{(p)}$ można znaleźć przez różniczkowanie równania (VIII.5).

Praktycznie metodę potęgową dla $p > 1$ wykorzystuje się bardzo rzadko, ponieważ obliczanie pochodnych dość wysokich rzędów bywa uciążliwe. Zamiast formuły (VIII.9) wykorzystujemy metody nie zawierające pochodnych funkcji $y = y(x)$.

2/ Metoda Eulera

Jeśli we wzorze (VIII.9) pozostawić tylko dwa pierwsze człony, to otrzymamy formułę

$$y_{k+1} = y_k + h_k f(x_k, y_k), \quad (\text{VIII.10})$$

którą nazywamy formułą Eulera.

Błąd tej formuły wyraża się wzorem

$$R_k = \frac{h_k^2}{2} y''(\xi), \quad x_k \leq \xi \leq x_{k+1}. \quad (\text{VIII.11})$$

Jak widać błąd formuły Eulera jest bardzo duży na każdym kroku całkowania. Znacznie lepszą dokładność daje ulepszona metoda Eulera, wyrażona formułami:

$$y_{k+1} = y_k + h_k f(x_k, y_k);$$
$$y_{k+1}^* = y_k + \frac{1}{2} h_k [f(x_k, y_k) + f(x_{k+1}, y_{k+1})] \quad (\text{VIII.12})$$

Z pierwszej z tych formuł oblicza się pierwsze, bardzo "grube" przybliżenie wartości y_{k+1} , a następnie drugą formułą zwiększa się dokładność uzyskując y_{k+1}^* .

Błąd tej metody można zapisać

$$R_k = \frac{\theta}{4} h_k^3 y'''(\xi) \quad \text{gdzie } -1 < \theta < 1 \text{ oraz } x_k \leq \xi \leq x_{k+1} \quad (\text{VIII.13})$$

Podczas całkowania ze stałym krokiem h wygodnie jest obliczyć z formuł (VIII.12) tylko y_1 , a następnie stosować wzór zwany zmodyfikowaną formułą Eulera

$$y_{k+1} = y_{k-1} + 2hf(x_k, y_k) \quad (\text{VIII.14})$$

którego błąd wyraża się wzorem

$$R_k = \frac{h^3}{3} y'''(\xi), \quad x_{k-1} \leq \xi \leq x_{k+1} \quad (\text{VIII.15})$$

Metodę Eulera we wszystkich jej odmianach stosuje się w tych przypadkach, kiedy wymagana jest niewielka dokładność wyników. Zwiększenie dokładności bowiem można uzyskać tylko przez zmniejszenie kroku h , co z kolei powoduje bardzo znaczny wzrost ilości obliczeń.

3/ Metoda Runge - Kutta

Metoda Runge-Kutta polega na zastąpieniu odcinka szeregu Taylora dla funkcji $y = y(x)$ wyrażeniami nie zawierającymi pochodnych.

Metodę Runge-Kutta nazywamy metodę p -go rzędu jeśli odpowiada ona odcinkowi szeregu Taylora, zawierającemu pochodne funkcji $y = y(x)$ do p -go rzędu włącznie.

Metoda Runge-Kutta pierwszego rzędu pokrywa się z metodą Eulera (VIII.10), a metoda drugiego rzędu z ulepszoną metodą Eulera (VIII.12). Dla $p > 2$ istnieje kilka postaci formuł Runge-Kutta, z których dwie przedstawiamy.

Formuły Runge-Kutta trzeciego rzędu:

$$y_{k+1} = y_k + \frac{1}{6} (k_1 + 4k_2 + k_3);$$
$$k_1 = hf(x_k, y_k); \quad (\text{VIII.16})$$

$$k_2 = h_k f(x_k + \frac{1}{2} h_k, y_k + \frac{1}{2} k_1);$$

$$k_3 = h_k f(x_k + h_k, y_k - k_1 + 2k_2).$$

Formuły Runge-Kutta czwartego rzędu:

$$y_{k+1} = y_k + \frac{1}{6} (k_1 + 2k_3 + k_4);$$

$$k_1 = h_k f(x_k, y_k);$$

$$k_2 = h_k f(x_k + \frac{1}{2} h_k, y_k + \frac{1}{2} k_1);$$

(VIII.17)

$$k_3 = h_k f(x_k + \frac{1}{2} h_k, y_k + \frac{1}{2} k_2);$$

$$k_4 = h_k f(x_k + h_k, y_k + k_3).$$

Formuły te są najpowszechniej stosowanymi formułami Runge-Kutta jako że formuły wyższych rzędów są dość złożone i praktycznie nie wykorzystywane.

W każdej z formuł (VIII.16) i (VIII.17) najpierw oblicza się kolejne k_1, k_2, \dots , po czym znajduje się y_{k+1} .

Dla formuł Runge-Kutta rzędu $p > 2$ brak jest ścisłych wzorów określających ich błąd; wiadomo tylko, że formuła rzędu p daje błąd rzędu h^{p+1} .

Oprócz wymienionych tu metod można spotkać w literaturze dość często stosowane formuły Adamsa, Stirlinga, Gaussa i inne.

§ 3. Zagadnienie brzegowe

Dla równania różniczkowego zwyczajnego rzędu n można sformułować zagadnienie brzegowe w następujący sposób:

Dane równanie różniczkowe

$$F(x, y, y', \dots, y^{(n)}) = 0 \quad (\text{VIII.18})$$

gdzie $n \geq 2$.

Należy znaleźć rozwiązanie $y = y(x)$ tego równania, które w żadnych punktach $x_1 \leq x_2 \leq \dots \leq x_k$ spełnia n warunków brzegowych

$$V_\nu(y_1, y_1', \dots, y_1^{(n)}; y_2, y_2', \dots, y_2^{(n)}; \dots, y_k, y_k', \dots, y_k^{(n)}) = 0 \quad (\text{VIII.19})$$

gdzie $\nu = 1, 2, \dots, n$; $y_i^{(s)} = y^{(s)}(x_i)$ ($i = 1, 2, \dots, k$);

F i V są zadanymi na ogół nieliniowymi funkcjami. Zakładamy przy tym, że funkcje V_ν zależą od funkcji y lub jej pochodnych w co najmniej dwóch punktach, ponieważ w przeciwnym razie mielibyśmy do czynienia z zagadnieniem Cauchy'ego.

W zależności od postaci funkcji F i V_ν , zagadnienie brzegowe może posiadać jedno, kilka, nieskończenie wiele lub nie posiadać rozwiązań w ogóle.

Zagadnienie brzegowe nazywamy liniowym, jeśli zarówno równanie (VIII.18) jak i warunki (VIII.19) są liniowe.

W najprostszym liniowym zagadnieniu brzegowym do warunków brzegowych (VIII.19) wchodzi tylko dwie współrzędne $x_1 = a$ i $x_2 = b$, tzn. rozwiązuje się równanie różniczkowe

$$L[y] \equiv \sum_{\nu=0}^n p_\nu(x) y^{(\nu)} = f(x) \quad (\text{VIII.20})$$

z warunkami brzegowymi

$$\sum_{k=0}^{n-1} [\alpha_{ik} y^{(k)}(a) + \beta_{ik} y^{(k)}(b)] = \gamma_i \quad (i=1, 2, \dots, n) \quad (\text{VIII.21})$$

gdzie $p_\nu(x)$ i $f(x)$ są zadanymi funkcjami, a α_{ik} , β_{ik} , γ_i - zadanymi stałymi.

Równanie (VIII.20) nazywamy jednorodnym jeśli w przedziale $a \leq x \leq b$ funkcja $f(x) \equiv 0$; w przeciwnym przypadku

równanie to nazywamy niejednorodnym.

Analogicznie, warunki brzegowe (VIII.21) nazywamy jednorodnymi jeśli wszystkie $\gamma_i = 0$, a niejednorodnymi w przypadku przeciwnym.

Zagadnienie brzegowe nazywamy jednorodnym, jeśli są jednorodne zarówno równania różniczkowe jak i warunki brzegowe.

Omówimy teraz jedną z najbardziej ogólnych metod rozwiązywania zagadnień brzegowych polegającą na sprowadzaniu tego zagadnienia do zagadnienia Cauchy'ego. Jest to jedna z najbardziej wygodnych i najbardziej dokładnych metod. W ogólnym przypadku, kiedy dane jest równanie n -go rzędu (VIII.18) z warunkami brzegowymi (VIII.19), jego rozwiązanie można sprowadzić do rozwiązania nie więcej niż $n+1$ zagadnień Cauchy'ego w następujący sposób:

1. Poszukujemy n funkcji y_1, y_2, \dots, y_n , będących rozwiązaniem równań

$$L[y_k] = 0 \quad (\text{VIII.22})$$

z warunkami początkowymi

$$y_k^{(\nu)}(a) = \begin{cases} 0, & \text{jeśli } \nu \neq k-1 \\ 1, & \text{jeśli } \nu = k-1 \end{cases} \quad (\text{VIII.23})$$

gdzie $k = 1, 2, \dots, n$; $\nu = 0, 1, 2, \dots, n-1$.

2. Znajdujemy funkcję y_0 , jako rozwiązanie niejednorodnego równania różniczkowego

$$L[y_0] = f(x)$$

z warunkami początkowymi

$$y_0^{(\nu)}(a) = 0 \quad (\nu = 0, 1, 2, \dots, n-1)$$

Wówczas na mocy liniowości równania różniczkowego (VIII.20) szukane rozwiązanie może być zapisane w postaci

$$y(x) = y_0(x) + \sum_{k=1}^n c_k y_k(x) \quad (\text{VIII.24})$$

gdzie stałe c_k określone są warunkami brzegowymi (VIII.21).

Rozpatrzmy bardziej szczegółowo przypadek gdy $n = 2$, kiedy dane jest zagadnienie brzegowe

$$L[y] = p_0(x)y + p_1(x)y' + p_2(x)y'' = f(x) \quad (\text{VIII.25})$$

$$\begin{aligned} \alpha_{10}y(a) + \alpha_{11}y'(a) + \beta_{10}y(b) + \beta_{11}y'(b) &= \gamma_1 \\ \alpha_{20}y(a) + \alpha_{21}y'(a) + \beta_{20}y(b) + \beta_{21}y'(b) &= \gamma_2 \end{aligned} \quad (\text{VIII.26})$$

Należy wówczas rozwiązać trzy zagadnienia Cauchy'ego

$$L[y_0] = f(x), \quad y_0(a) = 0, \quad y_0'(a) = 0;$$

$$L[y_1] = 0, \quad y_1(a) = 1, \quad y_1'(a) = 0; \quad (\text{VIII.27})$$

$$L[y_2] = 0, \quad y_2(a) = 0, \quad y_2'(a) = 1$$

Rozwiązanie zagadnienia brzegowego zapisze się w postaci

$$y(x) = y_0(x) + c_1 y_1(x) + c_2 y_2(x) \quad (\text{VIII.28})$$

przy czym

$$y'(x) = y_0'(x) + c_1 y_1'(x) + c_2 y_2'(x).$$

Przyjmując w formułach (VIII.28) $x = a$, $x = b$ oraz podstawiając je do warunków brzegowych (VIII.26) otrzymamy układ dwóch algebraicznych równań liniowych dla określenia stałych c_1 i c_2 . Po ich obliczeniu i podstawieniu do (VIII.28) otrzymamy szukane rozwiązanie danego zagadnienia brzegowego.

Oprócz omówionej tu metody sprowadzania zagadnienia brzegowego do zagadnień początkowych dość często stosuje się do rozwiązywania zagadnienia brzegowego liniowego i nieliniowego takie metody jak różnic skończonych, metodę "progonki", metodę Ritca itd.

S p i s t r e ś c i

W s t ę p

Rozdział I

Działania na liczbach przybliżonych	str.
§ 1. Klasyfikacja błędów	3
§ 2. Zapis dziesiętny liczb przybliżonych. Ilość cyfr dokładnych. Cyfry znaczące	6
§ 3. Ocena błędów	8
§ 4. Błędy działań arytmetycznych	11

Rozdział II

Interpolacja funkcji

§ 1. Sformułowanie zagadnienia interpolacji	14
§ 2. Interpolacja wielomianami. Wzór interpolacyjny Lagrange'a	16
§ 3. Wzory interpolacyjne Newtona	19
§ 4. Ekstrapolacja	24
§ 5. Wybór wzoru interpolacyjnego	25

Rozdział III

Aproksymacja

§ 1. Sformułowanie zagadnienia. Przybliżenia jednostajne	27
§ 2. Metoda szeregów potęgowych	29
§ 3. Metoda najmniejszych kwadratów	30
§ 4. Dobór wzorów empirycznych	34
§ 5. Wybór metody aproksymacji	35

Rozdział IV

Przybliżone rozwiązywanie równań algebraicznych nieliniowych i przestępnych

	str.
§ 1. Sformułowanie zagadnienia. Klasyfikacja metod	37
§ 2. Lokalizacja i rozdzielanie pierwiastków	38
§ 3. Metoda Newtona	42
§ 4. Reguła fałsi	44
§ 5. Metoda iteracji prostych	46
§ 6. Porównanie metod	48

Rozdział V

Rozwiązywanie układów równań algebraicznych liniowych

§ 1. Klasyfikacja metod	50
§ 2. Metoda Gaussa	51
§ 3. Metoda iteracji	56
§ 4. Metoda Seidela	57
§ 5. Wybór metody	59

Rozdział VI

Wyznaczanie wartości własnych i wektorów własnych

§ 1. Sformułowanie zagadnienia. Klasyfikacja metod	61
§ 2. Metody dokładne	63
§ 3. Metody iteracyjne	66

Rozdział VII

Całkowanie numeryczne

§ 1. Sformułowanie zagadnienia całkowania numerycznego	69
§ 2. Metoda Newtona-Cotesa i jej szczególne przypadki	70

	str.
§ 3. Metoda Gaussa	74
§ 4. Wybór metody	77

Rozdział VIII

Rozwiązywanie równań różniczkowych zwyczajnych

§ 1. Uwagi ogólne	78
§ 2. Metody rozwiązywania równań różniczkowych 1-go rzędu	80
§ 3. Zagadnienia brzegowe	83

Wykonano w 50 egz.

Egz. 1-50 Bibl. Szkol.
/Dział jawny/

Wyk. mgr MAROŃSKI
Nr ks. 858/390/WW
Druk ASG-O-XV-4426

