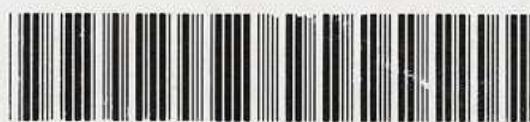


BIBLIOTEKA GŁÓWNA
POLITECHNIKI GDANSKIEJ
Magazyn Biblioteki Głównej

BG P 714873-00-00/05



STAJCZAK

METODY NUMERYCZNE

PRZYKŁADY I ZADANIA

WYDAWNICTWO POLITECHNIKI GDANSKIEJ

SPIS TREŚCI

PRZEDMOWA	5
1. WPROWADZENIE DO METOD NUMERYCZNYCH	7
1.1. Błędy modelowania matematycznego	7
1.2. Błędy danych	8
1.3. Błędy metody (błędy obcięcia)	9
1.4. Błędy zaokrągleń	11
1.5. Schemat Hornera	18
1.6. Sumowanie szeregów potęgowych	19
1.7. Zadania	21
2. ROZWIĄZYWANIE UKŁADÓW RÓWNAŃ LINIOWYCH	24
2.1. Wprowadzenie	24
2.2. Metoda wyznaczników Cramera	25
2.3. Metoda eliminacji Gaussa	27
2.4. Metoda eliminacji Gaussa z pełnym wyborem elementu podstawowego	31
2.5. Metoda eliminacji Jordana	33
2.6. Metoda LU	36
2.7. Metoda QR	38
2.8. Warianty metody iteracji prostej	43
2.8.1. Metoda iteracji prostej (metoda Jacobiego)	44
2.8.2. Metoda Gaussa-Seidela	46
2.9. Metoda Czebyszewa	48
2.10. Metoda najszybszego spadku	51
2.11. Metody rozwiązywania wielkich układów równań z macierzami rozrzedzonymi	54
2.12. Uwarunkowanie zadania rozwiązywania układu równań liniowych	59
2.13. Zadania	61
3. RÓWNANIA NIELINIOWE	78
3.1. Wstęp	78
3.2. Lokalizacja pierwiastków	78
3.3. Metoda bisekcji (metoda połowienia przedziału)	80
3.4. Metoda stycznych (metoda Newtona-Raphsona)	81
3.5. Metoda iteracji prostej	83
3.6. Efektywność metod przybliżonego obliczania pierwiastków funkcji	86
3.7. Zadania	87
4. ROZWIĄZYWANIE UKŁADÓW RÓWNAŃ NIELINIOWYCH	90
4.1. Wstęp	90
4.2. Lokalizacja rozwiązań	91
4.3. Metoda najszybszego spadku	92
4.4. Metoda Newtona-Raphsona	96
4.5. Metoda symulowanego wyżarzania	98
4.6. Zadania	99
5. INTERPOLACJA	114
5.1. Interpolacja w sensie Lagrange'a	114
5.2. Wzór Lagrange'a	116

5.3. Interpolacja Czebyszewa	118
5.4. Interpolacja trygonometryczna	120
5.5. Interpolacja w przypadku węzłów równoodległych	121
5.5.1. Wzór Stirlinga	121
5.5.2. Pierwszy wzór interpolacyjny Newtona	123
5.5.3. Drugi wzór interpolacyjny Newtona	125
5.6. Interpolacja funkcjami sklejonymi	126
5.7. Inne zagadnienia interpolacyjne	133
5.7.1. Zagadnienie interpolacyjne Taylora	133
5.7.2. Zagadnienie interpolacyjne Hermite'a	134
5.7.3. Zagadnienie interpolacyjne Abela-Gonczarowa	136
5.8. Zadania	138
6. APROKSYMACJA	145
6.1. Wprowadzenie	145
6.2. Aproksymacja średniokwadratowa dyskretna	146
6.3. Aproksymacja średniokwadratowa dyskretna za pomocą wielomianów	147
6.4. Aproksymacja średniokwadratowa dyskretna wielomianami ortogonalnymi	153
6.4.1. Aproksymacja za pomocą wielomianów p_k	154
6.4.2. Aproksymacja za pomocą wielomianów Grama	155
6.4.3. Aproksymacja za pomocą wielomianów trygonometrycznych	157
6.5. Aproksymacja średniokwadratowa całkowa	160
6.6. Aproksymacja średniokwadratowa całkowa za pomocą wielomianów	161
6.7. Aproksymacja średniokwadratowa całkowa wielomianami ortogonalnymi	162
6.7.1. Aproksymacja za pomocą wielomianów Legendre'a	163
6.7.2. Aproksymacja za pomocą wielomianów Czebyszewa pierwszego rodzaju	166
6.7.3. Aproksymacja za pomocą wielomianów Hermite'a	169
6.7.4. Aproksymacja za pomocą wielomianów trygonometrycznych	171
6.8. Aproksymacja za pomocą wzorów empirycznych	173
6.9. Zadania	174
7. RÓZNICZKOWANIE NUMERYCZNE	179
7.1. Sformułowanie zagadnienia	179
7.2. Różniczkowanie z wykorzystaniem interpolacyjnych funkcji sklejanych	179
7.3. Różniczkowanie z wykorzystaniem wzoru interpolacyjnego Stirlinga	182
7.4. Zadania	185
8. CAŁKOWANIE NUMERYCZNE	188
8.1. Sformułowanie zagadnienia	188
8.2. Kwadratury Newtona Cotesa	189
8.2.1. Złożony wzór trapezów	189
8.2.2. Złożony wzór parabol (wzór Simpsona)	190
8.3. Ekstrapolacja Richardsoна	192
8.3.1. Ekstrapolacja Richardsoña z wykorzystaniem wzoru trapezów	193
8.3.2. Ekstrapolacja Richardsoña z wykorzystaniem wzoru Simpsona	194
8.4. Kwadratury Gaussa	195
8.5. Zadania	196
DODATKI	201
A. Wzory sumacyjne	201
B. Normy wektorów i macierzy	202
C. Liczby i wielomiany Bernoulliego	204
D. Liczby Fibonacciego	206
E. Rozwinięcia funkcji elementarnych w szereg Maclaurina	207
F. Rozwinięcia funkcji elementarnych w ułamki łańcuchowe	209
LITERATURA	211

PRZEDMOWA

Metody numeryczne są działem matematyki stosowanej, zajmującym się opracowywaniem metod rozwiązywania problemów matematycznych, których albo nie można rozwiązać metodami dokładnymi, albo metody, których można użyć do rozwiązania, mają tak dużą złożoność obliczeniową, że są praktycznie nieużyteczne. Przymiotnik numeryczny znaczy „wyrażony liczbami” (*łac. numerus – liczba*). Zatem metody numeryczne zajmują się opracowywaniem algorytmów, których dane, wyniki pośrednie i ostateczne są liczbami. Obszar zastosowań metod numerycznych jest bardzo rozległy. Wszędzie tam, gdzie rozważanemu wycinkowi rzeczywistości przypisano model matematyczny w postaci funkcji, pochodnej, całki lub układu równań (liniowych, nieliniowych, różniczkowych, całkowych) rozwiązania zagadnienia poszukuje się z wykorzystaniem metod numerycznych.

Metody numeryczne są dziedziną ciągle rozwijającą się, zarówno jeśli chodzi o metody obliczeniowe, jak i obszar ich zastosowań. Takie pojęcia jak *uwarunkowanie zadania*, *stabilność algorytmu*, *podwójna precyzja* zostały wprowadzone w badaniach nad poszukiwaniem nowych metod numerycznych – metod realizujących obliczenia z większą dokładnością i poszukujących rozwiązania zadań o coraz większych rozmiarach. Motorem tego rozwoju jest szybki wzrost mocy obliczeniowej komputerów. W efekcie powstały już takie metody (np. metoda elementu skończonego), których realizację bez użycia komputerów trudno sobie wyobrazić.

Skrypt jest przeznaczony dla studentów Wydziału Elektroniki, Telekomunikacji i Informatyki jako pomoc dydaktyczna do przedmiotu „Metody numeryczne”. Przedmiot ten jest prowadzony na semestrze trzecim, po zapoznaniu studentów z podstawowymi działami matematyki wyższej (analiza matematyczna, algebra liniowa, geometria analityczna) oraz opanowaniu przez nich umiejętności programowania w języku wysokiego poziomu.

Z bogatej wiedzy dotyczącej metod numerycznych, do niniejszego skryptu zostały wybrane i przedstawione w nim następujące tematy:

- wiadomości o błędach występujących podczas obliczeń numerycznych (rozdział 1),
- najczęściej stosowane metody rozwiązywania układów równań liniowych (rozdział 2),
- metody obliczania przybliżonych wartości pierwiastków równań algebraicznych i kolejnych (rozdział 3),
- najczęściej stosowane metody rozwiązywania układów równań nieliniowych (rozdział 4),
- interpolacja funkcji (rozdział 5),
- aproksymacja średniokwadratowa (rozdział 6),
- różniczkowanie numeryczne (rozdział 7),
- całkowanie numeryczne (rozdział 8).

Treść ta została uzupełniona dodatkami zawierającymi wzory, definicje i własności pojęć matematycznych, które mogą okazać się przydatne przy rozwiązywaniu zadań numerycznych.

Pisząc niniejszy skrypt, starałem się zamierzony materiał przedstawić w sposób możliwie prosty. Doświadczenie dydaktyczne sugeruje, aby przekazywaniu wiedzy z metod

numerycznych na wykładzie przedmiotu towarzyszyło rozwiązywanie zadań na ćwiczeniach. Rozwiązywanie zadań jest dobrze znaną, dobrą i skutecną metodą uczenia się zagadnień matematycznych, prowadzącą do lepszego zrozumienia treści. Dlatego nie zamierzałem przedstawić Czytelnikowi obszernej (tym bardziej wyczerpującej) wiedzy z metod numerycznych. Moim celem było:

- a) opracowanie kompendium wiedzy z metod numerycznych na wykład przedmiotu, aby odciążyć słuchacza od notowania podstawowych wiadomości w trakcie wykładu,
- b) dostarczenie studentom szerokiego zbioru zadań z przykładami ich rozwiązań, aby dać możliwość uczenia się przedmiotu poprzez rozwiązywanie zadań.

Algorytmy numeryczne prezentuję, używając wersji publikacyjnej języka Pascal, gdyż umożliwia ona zapisanie algorytmów w bardziej zwięzły i precyzyjniejszy sposób.

Recenzentowi skryptu – Pani Doktor Jadwidze Kozłowskiej – serdecznie dziękuję za wnikliwe przeczytanie maszynopisu i cenne uwagi, które pozwoliły na usunięcie usterek i udoskonalenie tekstu.

Gdańsk, kwiecień 2006

Tadeusz Ratajczak

Wprowadzenie do metod numerycznych

Charakterystyczną cechą metod numerycznych jest wykonywanie obliczeń na liczbach przybliżonych; również rozwiązania zagadnień są wyrażone liczbami przybliżonymi. Należy jednak podkreślić, że wielkość błędu w procesie obliczeń numerycznych jest zawsze kontrolowana. Do określenia wielkości błędu służą dwa pojęcia:

$$\text{Błąd bezwzględny} = |\text{wartość przybliżona} - \text{wartość dokładna}| \quad (1.1)$$

$$\text{Błąd względny} = \frac{|\text{wartość przybliżona} - \text{wartość dokładna}|}{|\text{wartość dokładna}|} \quad (1.2)$$

Przy wykonywaniu obliczeń numerycznych obowiązuje zasada: *wartość przybliżona nie może różnić się od wartości dokładnej o więcej, niż z góry zadana dokładność obliczenia wielkości.*

W trakcie wykonywania obliczeń numerycznych wszechobecnymi są różnego rodzaju błędy, które musimy śledzić, aby otrzymać poprawne wyniki. Ważniejszymi z nich są:

- błędy modelowania,
- błędy danych,
- błędy metody,
- błędy zaokrągleń.

1.1. Błędy modelowania matematycznego

Błąd modelowania pojawia się wówczas, gdy przyjęty model matematyczny jest jedynie przybliżeniem zjawiska rzeczywistego. Istnieje zatem różnica między wartością w rzeczywistości a wartością obliczoną z przyjętego modelu matematycznego. Definicję zilustrujemy przykładem.

Przykład 1.1. Wzrost populacji często obliczamy ze wzoru

$$N(t) = N_0 e^{kt}, \quad (1.3)$$

w którym N_0 , k są stałymi dodatnimi (N_0 oznacza liczebność populacji w chwili $t = 0$). Dla małych wartości t , $N(t)$ obliczone ze wzoru (1.3) są zazwyczaj bliskie wartościom rzeczywistym. Dla dużych wartości t , wzór ten zwykle zawyża wyniki. Wynika to m.in. z faktu, że wzór ten opisuje środowisko, w którym wszystkie pokolenia mają dostęp do nieograniczonych zasobów pozwalających im na przeżycie. Jeśli więc ktoś poprosi o prognozę liczebności populacji w chwili t_n , a naszą odpowiedzią jest wynik obliczeń ze wzoru (1.3), to

prawdopodobnie będzie ona różniła się od rzeczywistej, dlatego że równanie (1.3) nieprecyjnie opisuje rzeczywistość. Ta różnica jest błędem modelowania.

1.2. Błędy danych

Błędy danych, to błędy jakimi są obarczone dane wejściowe zadania numerycznego. Błędy te pojawiają się głównie w dwóch przypadkach: jeśli dane wejściowe zadania są wynikami pomiarów oraz gdy w trakcie obliczeń używamy stałych będących liczbami niewymiernymi i zmuszeni jesteśmy używać ich przybliżeń.

Dane wejściowe są na ogół obarczone małymi błędami. Jednakże znane są przypadki, że małe błędy względne danych wejściowych zadania numerycznego mogą spowodować duże zmiany względne wyników. Własność ta została zilustrowana w poniższym przykładzie.

Przykład 1.2. Zadanie A: Rozwiązać następujący układ równań liniowych

$$\begin{aligned}x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 &= 32, \\ \frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 &= 27, \\ \frac{1}{3}x_1 + \frac{1}{4}x_2 + \frac{1}{5}x_3 &= 17.\end{aligned}\tag{1.4}$$

Zadanie B: Rozwiązać następujący układ równań liniowych

$$\begin{aligned}1.00x_1 + 0.50x_2 + 0.33x_3 &= 32, \\ 0.50x_1 + 0.33x_2 + 0.25x_3 &= 27, \\ 0.33x_1 + 0.25x_2 + 0.20x_3 &= 17.\end{aligned}\tag{1.5}$$

Rozwiązaniem zadania A jest $x_1 = 6$, $x_2 = 12$, $x_3 = 60$. Natomiast zadanie B ma następujące rozwiązanie: $x_1 = -1377.777778$, $x_2 = 7217.460317$, $x_3 = -6663.492063$. Możemy przyjąć, że zadania A i B opisują ten sam wycinek rzeczywistości; w zadaniu A równania zapisano za pomocą ułamków zwykłych, a w zadaniu B za pomocą ułamków dziesiętnych z dokładnością $5 \cdot 10^{-3}$. W konsekwencji przybliżono wartości trzech współczynników macierzy głównej układu: a_{13} , a_{22} , a_{31} . Zatem zadanie B możemy potraktować jako zadanie A, w którym zaburzono trzy (spośród dwunastu) dane. Błąd względny zaburzonych danych wynosi $\varepsilon(a_{13}) = \varepsilon(a_{22}) = \varepsilon(a_{31}) = 1.5\%$. Błędy względne wyników wynoszą: $\varepsilon(x_1) = 230.63 = 23063\%$, $\varepsilon(x_2) = -600.45 = -60045\%$, $\varepsilon(x_3) = 112.06 = 11206\%$.

Zadanie numeryczne, w którym niewielkie zmiany względne danych zadania powodują duże zmiany jego rozwiązania, nazywamy zadaniem *źle uwarunkowanym*. Zadanie A z przykładu 1.2 jest zadaniem źle uwarunkowanym. Oczywiście nie jest tak, że wszystkie zadania numeryczne dzielą się na dobrze uwarunkowane i źle uwarunkowane. Do rozważań wprowadzamy wskaźnik uwarunkowania zadania, oznaczany $cond(\text{zadanie})$, który charakteryzuje liczbowo wpływ zaburzeń danych zadania na zaburzenia rozwiązania. Wskaźnik uwarunkowania zadania $cond(\text{zadanie})$ jest to najmniejsza liczba rzeczywista, dodatnia, spełniająca nierówność

$$\text{błąd względny(wyniki zadania)} \leq cond(\text{zadanie}) \cdot \text{błąd względny(dane zadania)}. \tag{1.6}$$

Warto podkreślić, że uwarunkowanie zadania jest właściwością zadania, a nie sposobu jego rozwiązania. Zadanie źle uwarunkowane zachowa tę własność niezależnie od metody i narzędzi, których użyjemy do rozwiązania zadania.

1.3. Błędy metody (błędy obcięcia)

Błędy metody (błędy obcięcia) wynikają z zastąpienia działań nieskończonych działaniami skończonymi lub działań na wielkościach nieskończonymi małymi działaniami na wielkościach skończonymi małymi. Ten rodzaj błędów pojawia się wtedy, gdy należy obliczyć wartość pojęcia zdefiniowanego za pomocą granicy.

Przykład 1.3. Wiadomo, że spośród działań arytmetycznych arytmometr komputera wykonuje tylko cztery podstawowe działania. Jak zatem obliczyć wartość funkcji

$$y = \sin x$$

dla $0 \leq x \leq 2\pi$?

Rozwiązanie. Jedną z metod rozwiązywania powyższego problemu jest rozwinięcie funkcji $y = \sin x$ w szereg potęgowy

$$f_\infty(x) = \sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (1.7)$$

Szereg (1.7) nie definiuje algorytmu, ponieważ nakazuje wykonanie nieskończonej liczby operacji. Należy więc sumę (1.7) zastąpić sumą skońzoną

$$f_n(x) = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad (1.8)$$

z tym że do obliczeń należy wziąć tyle składników, aby dla zadanej dokładności ε zachodziła nierówność

$$|f_\infty(x) - f_n(x)| = \left| \sum_{k=n+1}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \right| \leq \varepsilon. \quad (1.9)$$

Zbudowanie algorytmu realizującego nierówność (1.9) nie jest trudne w przypadku, gdy szereg potęgowy jest naprzemienny (a w taki szereg rozwija się funkcja $y = \sin x$). Dla szeregiów naprzemiennych bowiem prawdziwe jest następujące twierdzenie.

Twierdzenie 1.1. Jeżeli szereg $\sum_{k=0}^{\infty} a_k$ jest naprzemienny i jego wyrazy co do bezwzględnej wartości zmierzają monotonicznie do zera, to dla każdego naturalnego n prawdziwa jest nierówność

$$\left| \sum_{k=0}^{\infty} a_k - \sum_{k=0}^n a_k \right| = \left| \sum_{k=n+1}^{\infty} a_k \right| < |a_{n+1}|. \quad (1.10)$$

Opierając się na powyższym twierdzeniu możemy zapisać warunek

$$|f_\infty(x) - f_n(x)| = \left| \sum_{k=n+1}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \right| < \frac{x^{2n+3}}{(2n+3)!}, \quad (1.11)$$

który będzie wykorzystany jako kryterium zakończenia algorytmu.

Algorytm 1.1. Obliczanie wartości $y = \sin x$ dla $0 \leq x \leq 2\pi$ zadaną dokładnością ε .
 $Sin(x, \varepsilon)$

```

1       $k := 0$ 
2       $y := 0$ 
3       $S_0 := x$ 
4      repeat
5       $y := y + S_k$ 
6       $k := k + 1$ 
7       $S_k := S_{k-1} \cdot \left( -\frac{x^2}{2k(2k+1)} \right)$ 
8      until  $|S_k| < \varepsilon$ 

```

Przykład 1.4. Najprostszym przykładem wykonywania działań na wielkościach nieskończonymi małymi jest obliczanie pochodnej funkcji. Rozpatrzmy zatem funkcję $y = f(x)$, określoną i różniczkowalną w przedziale $[a; b]$. W obliczeniach numerycznych możemy wykonywać działania na wielkościach skończenie małych. Zatem zamiast obliczać $f'(x)$, obliczamy dla małych h iloraz różnicowy

$$D_h f(x) = \frac{f(x+h) - f(x-h)}{2h}. \quad (1.12)$$

Postępując tak, jesteśmy świadomi, że obliczona wartość jest obarczona błędem równym

$$|f'(x) - D_h f(x)| = \left| f'(x) - \frac{f(x+h) - f(x-h)}{2h} \right|. \quad (1.13)$$

Obliczanie pochodnej według wzoru (1.12) można uznać za metodę numeryczną tylko wtedy, gdy dla zadanej dokładności ε można wyznaczyć takie $h > 0$, że dla wszystkich $x \in [a+h; b-h]$ będzie zachodziła nierówność

$$|f'(x) - D_h f(x)| = \left| f'(x) - \frac{f(x+h) - f(x-h)}{2h} \right| \leq \varepsilon. \quad (1.14)$$

Dla wyznaczenia h spełniającego warunek (1.14) skorzystamy z równości

$$f'(x) - \frac{f(x+h) - f(x-h)}{2h} = -\frac{h^2}{6} f'''(c),$$

zachodzącej dla pewnego $x-h \leq c \leq x+h$ (równość została udowodniona w [2]). Stąd mamy

$$\left| f'(x) - D_h f(x) \right| = \left| f'(x) - \frac{f(x+h) - f(x-h)}{2h} \right| = \left| -\frac{h^2}{6} f'''(c) \right| \leq \frac{h^2}{6} M_3,$$

gdzie

$$a \leq c \leq b, \quad M_3 = \max_{a \leq x \leq b} |f'''(x)|.$$

Jeśli więc $y = f'''(x)$ nie jest tożsamościowo równa zeru na $[a; b]$, to dla zadanej dokładności ε możemy h określić z nierówności

$$\frac{h^2}{6} M_3 \leq \varepsilon.$$

Przyjmując

$$h = \sqrt{\frac{6\varepsilon}{M_3}},$$

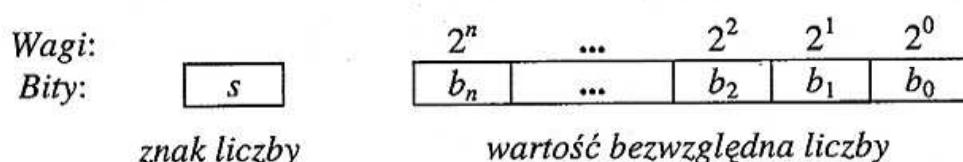
mamy gwarancję, że na całym odcinku $[a; b]$ zachodzi nierówność

$$|f'(x) - D_h f(x)| \leq \varepsilon.$$

1.4. Błędy zaokrągleń

Błędy zaokrągleń wynikają z faktu, że obliczenia wykonujemy na liczbach o skończonym rozwinięciu pozycyjnym. Najpierw omówimy sposób przedstawiania liczb w pamięci komputera. Komputery zapamiętują liczby albo jako *stałoprzecinkowe* (*liczby stałopozycyjne*, ang. *fixed-point numbers*) albo jako *zmiennoprzecinkowe* (*liczby zmiennopozycyjne*, ang. *floating-point numbers*).

Jeżeli na reprezentację liczby stałoprzecinkowej przeznacza się $n + 2$ bity (1 bit na znak liczby oraz $n + 1$ bitów na wartość bezwzględną liczby – rys. 1.1),



Rys. 1.1. Struktura liczby stałoprzecinkowej

to tę strukturę interpretuje się następująco:

$$l = s \cdot \sum_{k=0}^n b_k 2^k, \quad (1.16)$$

gdzie:

- $s = 1$ albo $s = -1$ (znak liczby);
- $b_k \in \{0; 1\}$ (bity).

Zatem na $n+2$ bitach można zapisywać liczby całkowite z przedziału $[-2^{n+1} + 1; 2^{n+1} - 1]$. Liczby stałoprzecinkowe są podzbiorem liczb całkowitych. Podzbiór ten jest tym większy,

im większe jest n . Języki programowania wysokiego poziomu oferują kilka typów liczb stałoprzecinkowych: na reprezentację liczb typu *Integer* przeznacza się 16 bitów (2 kolejne bajty), na reprezentację liczb typu *LongInt* – 4 bajty, na typ *ShortInt* – 1 bajt. Liczby całkowite $p > 2^{n+1} - 1$ nie mogą być reprezentowane w komputerze. Jeżeli w trakcie obliczeń otrzymamy taką liczbę, to wystąpi sytuacja wyjątkowa nazywana *nadmiarem stałoprzecinkowym*.

Przykład 1.5. Dysponujemy komputerem, w którym liczba stałoprzecinkowa zapisywana jest na 6 bitach:

- znak liczby s na 1 bicie,
- wartość bezwzględna liczby na 5 bitach.

Pokazać, jak w tym komputerze jest zapisana liczba -22 .

Rozwiążanie. Mamy $(22)_{10} = (10110)_2$. Zatem

Wagi:	$2^4 \quad 2^3 \quad 2^2 \quad 2^1 \quad 2^0$	
Bity:	<table border="1" style="display: inline-table; vertical-align: middle; border-collapse: collapse;"> <tr> <td style="padding: 2px 10px;">-1</td> </tr> </table> $1 \quad 0 \quad 1 \quad 1 \quad 0$	-1
-1		
	<i>znak</i> <i>wartość bezwzględna</i>	

Liczby zmiennoprzecinkowe są reprezentowane za pomocą trzech obszarów bitów: *znaku liczby s* , *mantysy m_t* – zapisanej na t bitach, oraz *cechy c_n* – zapisanej w formacie stałoprzecinkowym na $n + 2$ bitach – rys. 1.2 (s_c oznacza znak cechy).

Wagi:	$2^{-1} \quad 2^{-2} \quad \dots \quad 2^{-t}$	$2^n \quad \dots \quad 2^2 \quad 2^1 \quad 2^0$	
Bity:	<table border="1" style="display: inline-table; vertical-align: middle; border-collapse: collapse;"> <tr> <td style="padding: 2px 10px;">s</td> </tr> </table> $b_{-1} \quad b_{-2} \quad \dots \quad b_{-t}$	s	$s_c \quad b_n \quad \dots \quad b_2 \quad b_1 \quad b_0$
s			
	<i>Znak liczby</i> <i>mantysa m_t</i>	<i>cecha c_n</i>	

Rys. 1.2. Struktura liczby zmiennoprzecinkowej

Reprezentacja zmiennoprzecinkowa liczby x – oznaczana $rd(x)$ – jest liczbą

$$rd(x) = s \cdot m_t \cdot 2^{c_n}, \quad (1.17)$$

gdzie:

$$\begin{aligned} &— s = 1 \text{ albo } s = -1 && (\text{znak liczby } x), \\ &— m_t = 1 + \sum_{k=1}^t b_{-k} 2^{-k} \quad (b_{-1}=1) && (\text{mantysa znormalizowana}), \end{aligned} \quad (1.18)$$

$$m_t = 1 + \sum_{k=1}^t b_{-k} 2^{-k} + b_{-(t+1)} \cdot 2^{-t} \quad (b_{-1} = 1) \text{ (mantysa znormalizowana zaokrąglona do } t \text{ cyfr binarnych),} \quad (1.19)$$

$$— c_n = s_c \cdot \sum_{k=0}^n b_k 2^k \quad (\text{cecha}). \quad (1.20)$$

Przykład 1.6. Dysponujemy komputerem, w którym liczba zmiennoprzecinkowa zapisywana jest następująco:

- znak liczby s na 1 bicie,
 - mantysa znormalizowana m_6 na 6 bitach,
 - cecha c_6 na 6 bitach (bit znaku cechy s_c oraz 5 bitów bezwzględnej wartości cechy).
- Pokazać, jak w tym komputerze zapisana jest liczba 13.75.

Rozwiązańe. Mamy $(13.75)_{10} = (1101.11)_2$. Ponieważ w postaci znormalizowanej mantysy kropka znajduje się po pierwszej jedynce liczby w systemie dwójkowym, więc

$$(1101.11)_2 = (1.10111)_2 \cdot 2^{+(11)_2}.$$

Stąd $m_6 = (1.10111)_2$, $c_6 = +(00011)_2$. Ostatecznie mamy

Wagi:	s	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}	s_c	2^4	2^3	2^2	2^1	2^0
Bity:	<input type="checkbox"/>	1	0	1	1	1	0	<input type="checkbox"/>	1	0	0	1	1
Znak liczby	<i>mantysa</i>						<i>cecha</i>						

Przykład 1.7. Dysponujemy komputerem, jak w przykładzie 1.6. Pokazać, jak w tym komputerze zapisana będzie liczba 5.3.

Rozwiązańe. Mamy

$$(5.3)_{10} = (101.010011001\ldots)_2 = (1.01010011001\ldots)_2 \cdot 2^{(10)_2}.$$

Liczba 5.3 nie ma skończonego rozwinięcia w systemie dwójkowym. Liczba ta ma rozwinięcie okresowe z okresem (1001). Zatem zapamiętamy jej przybliżenie. Komputer może zapamiętywać wyniki działań albo *bez zaokrąglenia* albo *z zaokrągleniem*.

Jeśli komputer zapamiętuje liczby bez zaokrąglenia, to mantysa będzie pamiętaana w postaci (1.18), a cała liczba 5.3 będzie reprezentowana jako:

Wagi:	s	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}	s_c	2^4	2^3	2^2	2^1	2^0
Bity:	<input type="checkbox"/>	1	0	1	0	1	0	<input type="checkbox"/>	1	0	0	0	1
Znak liczby	<i>mantysa</i>						<i>cecha</i>						

Zatem zapamiętana zostanie nie liczba 5.3, ale 5.25.

Jeśli komputer zapamiętuje liczby z zaokrągleniem, to mantysa będzie pamiętaana w postaci (1.19), a cała liczba 5.3 będzie reprezentowana jako

Wagi:	s	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}	s_c	2^4	2^3	2^2	2^1	2^0
Bity:	<input type="checkbox"/>	1	0	1	0	1	0	<input type="checkbox"/>	1	0	0	0	1
Znak liczby	<i>mantysa</i>						<i>cecha</i>						

W tym przypadku pamiętaana będzie liczba 5.3125.

Odrodumy najważniejsze własności liczb zmiennoprzecinkowych. Z definicji (1.18) wynika, że

$$1 \leq m_t \leq 2 - 2^{-t} \quad (1.21)$$

(często mówimy, że wartości mantysy należą do przedziału domknięto-otwartego [1 ; 2)). Cecha jest liczbą całkowitą z przedziału

$$-2^{n+1} + 1 \leq c_n \leq 2^{n+1} - 1. \quad (1.22)$$

Z równań (1.17) \div (1.20) wynika, że

$$2 \cdot 2^{-2^{n+1}} \leq |rd(x)| < 2^{2^{n+1}}. \quad (1.23)$$

Jeżeli reprezentacja zmennoprzecinkowa liczby x jest opisana za pomocą wzorów (1.17) \div (1.20), to liczbę x określa wzór

$$x = s \cdot m_{\infty} \cdot 2^{c_n}, \quad (1.24)$$

gdzie

$$m_{\infty} = 1 + \sum_{k=1}^{\infty} b_{-k} \cdot 2^{-k}. \quad (1.25)$$

Ze wzorów (1.18) oraz (1.25) mamy

$$|m_{\infty} - m_t| \leq 2^{-t}. \quad (1.26)$$

Błąd względny reprezentacji zmennoprzecinkowej daje się oszacować następująco:

$$\left| \frac{x - rd(x)}{x} \right| \leq 2^{-t}. \quad (1.27)$$

Liczba 2^{-t} jest parametrem charakteryzującym dokładność komputera. Parametr ten nie zależy od x ; zależy tylko od t - liczby bitów przeznaczonych na zapis mantysy. Nierówność (1.27) jest równoważna równości

$$rd(x) = x(1 + \varepsilon),$$

gdzie ε jest błędem względnym reprezentacji $rd(x)$, spełniającym nierówność $|\varepsilon| \leq 2^{-t}$.

Z pomocą liczb zmennoprzecinkowych reprezentowane są w komputerze liczby rzeczywiste. Z powyższych rozważań widzimy, że liczby rzeczywiste nie są na ogół reprezentowane dokładnie, ale w sposób przybliżony z błędem względnym ε nie większym od 2^{-t} . Biorąc pod uwagę równość (1.17) oraz nierówności (1.21) możemy stwierdzić, że liczby zmennoprzecinkowe są ograniczonym podzbiorem liczb wymiernych. Liczby $|x| < 2 \cdot 2^{-2^{n+1}}$ (w tym liczba 0) nie mają reprezentacji zmennoprzecinkowej w komputerze. Jeśli w trakcie obliczeń uzyskamy liczbę x , taką że $|x| < 2 \cdot 2^{-2^{n+1}}$, to sytuację taką nazywamy *niedomiarem zmennoprzecinkowym* (ang. *underflow*). Podobnie nie mogą być reprezentowane w komputerze liczby $|x| \geq 2^{2^{n+1}}$. Sytuację, w której wystąpi taka liczba w trakcie obliczeń nazywamy *nadmiarem zmennoprzecinkowym* (ang. *overflow*).

W kolejnych przykładach pokażemy wpływ sposobu reprezentowania liczb na obliczenia numeryczne.

Przykład 1.8. Liczba zespolona $z = a + bi$ ($i = \sqrt{-1}$) zapamiętywana jest jako uporządkowana para $[a ; b]$ liczb zmennoprzecinkowych. Przedyskutujmy algorytm obliczania modułu $|z|$ tej liczby.

Rozwiązanie. W klasycznym algorytmie

Algorytm 1.2. Obliczanie modułu liczby zespolonej $z = a + bi$.

Modull(a, b)

$$1 \quad |z| := \sqrt{a^*a + b^*b}$$

wystąpi nadmiar zmiennoprzecinkowy, gdy a lub b będzie większe od 2^{2^n} (porównaj nierówności (1.23)). Przypadek ten nie wystąpi, gdy obliczenia będziemy wykonywali według poniższego algorytmu.

Algorytm 1.3. Obliczanie modułu liczby zespolonej $z = a + bi$.

Modul2(a, b)

$$1 \quad \text{if } (|a| > 0) \text{ or } (|b| > 0)$$

$$2 \quad \text{then if } |a| > |b|$$

$$3 \quad \text{then } |z| := |a| * \sqrt{1 + \left(\frac{b}{a}\right)^2}$$

$$4 \quad \text{else } |z| := |b| * \sqrt{1 + \left(\frac{a}{b}\right)^2}$$

$$5 \quad \text{else } |z| := 0$$

Innym praktycznym zagadnieniem jest pytanie o *największą liczbę całkowitą w zapisie dziesiętnym, która może być dokładnie reprezentowana w komputerze*. Ta największa liczba – oznaczmy ją przez N – jest w pamięci komputera zapisana zgodnie z formatem (1.17) \div (1.20), a więc

$$N = \left(1 + 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2^2} + \dots + 1 \cdot \frac{1}{2^t}\right) \cdot 2^t \cong 2^{t+1}. \quad (1.28)$$

Najczęściej odpowiedzi udzielamy, podając maksymalną liczbę cyfr dziesiętnych, z których składa się ta liczba. Oto przykłady formatów zmiennoprzecinkowych.

- W standardzie arytmetyki IEEE mamy $N = 2^{24} = 16777216$. Zatem każda liczba całkowita 7-cyfrowa jest w komputerze reprezentowana dokładnie.
- Na komputerach CDC, $N = 2^{48} = 2.81 \cdot 10^{14}$. Zatem każda liczba całkowita, zapisana za pomocą 14 cyfr lub mniej, jest w komputerze reprezentowana dokładnie.

Języki programowania wysokiego poziomu oferują co najmniej dwa typy zmiennoprzecinkowe, np.

- FORTRAN – *real* oraz *double precision*,
- Pascal – *real* oraz *double*,
- C ++ – *float*, *double* oraz *long double*.

Typ *double* powstaje na bazie typu *real (float)* w ten sposób, że na reprezentację mantysy przeznacza się podwójną liczbę bitów, czyli $2t$ (podwójna precyzja). W tym typie liczby rzeczywiste są reprezentowane z błędem względny 2^{-2t} , a największa liczba całkowita, która może być dokładnie reprezentowana w komputerze, wynosi 2^{2t+2} .

Przykład 1.9. W tabeli 1.1 przedstawiono iteracje algorytmu 1.1, obliczania wartości funkcji $y = \sin x$ dla $x = 1.2$ na komputerach z precyzją:

- 6 cyfr dziesiętnych,
- 12 cyfr dziesiętnych.

Dodajmy, że $\sin(1.2) = 0.9320390859672$. Aby móc obliczyć wartość $\sin(1.2)$ z dokładnością większą od $5 \cdot 10^{-6}$, należy obliczenia wykonywać na komputerze z większą precyzją.

Tabela 1.1. Obliczanie wartości $\sin(1.2)$ z dokładnością komputerową

Obliczenia na komputerze z precyzją				
	6 cyfr dziesiętnych		12 cyfr dziesiętnych	
k	y	S_k	y	S_k
0	0.00000	1.20000	0.000000000000	1.200000000000
1	1.20000	-0.28800	1.200000000000	-0.288000000000
2	0.91200	0.02074	0.912000000000	0.02073600000
3	0.93274	-0.00071	0.932736000000	-0.00071094857
4	0.93203	0.00001	0.93202505143	0.00001421897
5	0.93204	0.00000	0.93203927040	-0.00000018614
6	0.93204	0.00000	0.93203908426	0.000000000172
7	0.93204	0.00000	0.93203908598	-0.000000000001
8	0.93204	0.00000	0.93203908597	0.000000000000

Kolejnym problemem numerycznym jest możliwość utraty cyfr znaczących liczby podczas obliczeń. Problem wyjaśnimy na przykładzie.

Przykład 1.10. Zagadnienie polega na obliczeniu wartości wyrażenia

$$f(x) = x(\sqrt{x+1} - \sqrt{x}), \quad (1.29)$$

dla nieujemnego, rzeczywistego x .

Rozwiązanie. W tabeli 1.2 zostały podane wartości funkcji (1.29), obliczone za pomocą kalkulatora wykonującego działania na 6 cyfrach znaczących oraz za pomocą komputera wykonującego działania arytmetyczne na 14 cyfrach znaczących.

Tabela 1.2. Obliczenia wartości funkcji (1.29) dla wybranych argumentów

x	Obliczenia wykonane za pomocą				
	kalkulatora z 6 cyframi znaczącymi				komputera
	$\sqrt{x+1}$	\sqrt{x}	$\sqrt{x+1} - \sqrt{x}$	$f(x)$	$f(x)$
1	1.41421E+00	1.00000E+00	4.14210E-01	0.41421	0.41421356
10	3.31662E+00	3.16228E+00	1.54340E-01	1.54340	1.54347130
100	1.00499E+01	1.00000E+01	4.99000E-02	4.99000	4.98756211
1000	3.16386E+01	3.16228E+01	1.58000E-02	15.80000	15.80743743
10000	1.00005E+02	1.00000E+02	5.00000E-03	50.00000	49.99875006
100000	3.16229E+02	3.16228E+02	1.00000E-03	100.00000	158.11348773

Porównując wyniki obliczeń wykonanych za pomocą dwóch wspomnianych narzędzi, stwierdzamy, że dla małych wartości argumentu x wartości funkcji $y = f(x)$ różnią się mało, a dla dużych wartości argumentu x różnią się istotnie. Ażeby zrozumieć to zjawisko, zauważmy najpierw, że dla dużych wartości argumentu x , wartości $\sqrt{x+1}$ oraz \sqrt{x} w przybliżeniu są równe, co oznacza, że na pozycjach najbardziej znaczących cyfry są równe, a na pozycjach mniej znaczących cyfry mogą różnić się. Utrata cyfr znaczących następuje podczas obliczenia różnicy $\sqrt{x+1} - \sqrt{x}$. Ostatecznie tak zredukowaną liczbę mnożymy przez dużą wartość x , otrzymując wynik nieprawidłowy.

W wielu przypadkach tego rodzaju mankament obliczeniowy możemy wyeliminować, przekształcając wyrażenie arytmetyczne w inne, równoważne, w którym zjawisko utraty cyfr znaczących nie występuje. Przykładowo, wyrażenie (1.29) możemy przekształcić następująco:

$$f(x) = x(\sqrt{x+1} - \sqrt{x}) = x \cdot \frac{\sqrt{x+1} - \sqrt{x}}{1} \cdot \frac{\sqrt{x+1} + \sqrt{x}}{\sqrt{x+1} + \sqrt{x}} = \frac{x}{\sqrt{x+1} + \sqrt{x}}. \quad (1.30)$$

Błędy zaokrągleń są przyczyną niestabilności algorytmów. *Algorytm* nazywamy *stabilnym*, jeśli błąd względny wyniku niewiele różni się od błędu względnego danych. *Algorytm* nazywamy *niestabilnym*, jeśli

$$\text{Błąd względny(wyniki)} >> \text{Błąd względny(dane)}.$$

Przykładem algorytmu niestabilnego jest następujący

Algorytm 1.4. Obliczanie $y = a^n$, dla zadanego rzeczywistego $a > 0$ oraz naturalnego n .

Power(n, a)

```

1      y := 1.0
2      for k := 1 to n
3          do y := y * a

```

Jeśli podstawa potęgi a jest obarczona błędem ε , to mamy

$$\text{Błąd względny(dane)} = \frac{|a - (a + \varepsilon)|}{|a|} = \frac{\varepsilon}{a}. \quad (1.31)$$

$$\text{Błąd względny(wyniki)} = \frac{|a^n - (a + \varepsilon)^n|}{|a^n|} = \frac{\left| a^n - (a^n + \binom{n}{1} a^{n-1} \varepsilon + \binom{n}{2} a^{n-2} \varepsilon^2 + \dots + \varepsilon^n) \right|}{a^n}.$$

Zaniedbując w liczniku ostatniego wyrażenia składniki zawierające czynnik ε^k dla $k \geq 2$ (jako bardzo małe i niemające znaczącego wpływu na wartość sumy), otrzymujemy

$$\text{Błąd względny(wyniki)} \cong \frac{\left| a^n - (a^n + \binom{n}{1} a^{n-1} \varepsilon) \right|}{a^n} = n \cdot \frac{\varepsilon}{a}.$$

Porównując ostatni wzór ze wzorem (1.31), odnotowujemy następujące spostrzeżenia:

- a) błąd względny funkcji $y = a^n$ jest n razy większy od błędu względnego podstawy potęgowania a ,
 b) błąd względny potęgi a^n rośnie wraz ze wzrostem wykładnika potęgi n .

Przykład 1.11. W tabeli 1.3 zestawiono wartości kolejnych potęg liczby $a = 7/3$ oraz potęg liczby 2.33 (podstawy potęgi a obarczonej błędem względnym 1/700 i błędem bezwzględnym równym 1/300).

Tabela 1.3. Propagacja błędów zaokrągleń w procesie obliczania potęg liczby $a = 7/3$

n	$(7/3)^n$	$(2.33)^n$
0	1.000000000	1.00
1	2.333333333	2.33
2	5.444444444	5.44
3	12.703703704	12.69
4	29.641975309	29.61
5	69.164609053	69.09
6	161.384087791	161.21
7	376.562871513	376.16
8	878.646700198	877.71
9	2050.175633796	2047.99
10	4783.743145523	4778.64

1.5. Schemat Hornera

Wiele zadań numerycznych sprowadza się do obliczania wartości wielomianu

$$w_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0, \quad (1.32)$$

dla zadanego argumentu x . Wyznaczenie wartości wielomianu ze wzoru (1.32) wymaga wykonania $n(n+1)/2$ mnożeń i n dodawań. Algorytm ten może spowodować wystąpienie nadmiaru zmiennoprzecinkowego przy próbie obliczenia x^n dla dużych wartości x .

Zauważmy, że wyrażenie (1.32) możemy zapisać w postaci

$$w_n(x) = (\dots ((a_n x + a_{n-1}) x + a_{n-2}) x + \dots + a_1) x + a_0. \quad (1.33)$$

Algorytm obliczania wartości wielomianu według wzoru (1.33) nosi nazwę *schematu Hornera*. Wymaga on wykonania n mnożeń i n dodawań; zatem ma mniejszą złożoność obliczeniową. Dodajmy, że podczas obliczeń według schematu Hornera zmniejsza się liczba przypadków, w których wystąpi nadmiar zmiennoprzecinkowy.

Algorytm 1.5. Obliczanie wartości wielomianu według schematu Hornera.

Horner(x)

- 1 $w := a_n$
- 2 **for** $k := n - 1$ **downto** 0
- 3 **do** $w := w * x + a_k$

Powróćmy jeszcze na moment do algorytmu 1.1 obliczania wartości funkcji $y = \sin x$. Tam już wykorzystany został schemat Hornera. Gdyby składnik

$$S_k = (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

obliczano według algorytmu

$$L := x^{2k+1}$$

$$M := (2k+1)!$$

$$S_k := (-1)^k \cdot \frac{L}{M},$$

to istniałaby możliwość wystąpienia przy niewielkich k :

- nadmiaru zmienoprzecinkowego podczas obliczania wartości L ,
- nadmiaru stałoprzecinkowego podczas obliczania wartości M .

Dodatkowo trzeba zwrócić uwagę na dużą złożoność obliczeniową tego algorytmu; wymaga on bowiem wykonania $4k + 2$ mnożeń i dzielenia. Natomiast wzór zaproponowany w algorytmie 1.1 istotnie redukuje wszystkie wymienione wady.

1.6. Sumowanie szeregów potęgowych

Często obliczamy wartość funkcji $y = f(x)$ dla zadanego argumentu x , korzystając z rozwinięcia tej funkcji w szereg Maclaurina

$$f(x) = \sum_{k=0}^{\infty} a_k x^k, \quad (1.34)$$

gdzie

$$a_k = \frac{f^{(k)}(0)}{k!}.$$

Algorytm obliczania przybliżonej wartości szeregu (1.34), w przypadku gdy jest on naprzemienny, został omówiony w przykładzie 1.3. Gdy szereg (1.34) nie jest naprzemienny, to nie zachodzi nierówność (1.10), która w algorytmie 1.1 została wykorzystana do skonstruowania kryterium zakończenia sumowania. W ogólnym przypadku zapiszemy szereg Maclaurina z resztą Lagrange'a

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + \frac{f^{(n+1)}(\theta x)}{(n+1)!} x^{n+1}, \text{ dla pewnego } (0 < \theta < 1).$$

Stąd

$$\left| f(x) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k \right| = \frac{|f^{(n+1)}(\theta x)|}{(n+1)!} \cdot |x|^{n+1} \leq \frac{M_{n+1}}{(n+1)!} \cdot |x|^{n+1},$$

gdzie

$$M_{n+1} = \max_{0 \leq t \leq x} |f^{(n+1)}(t)|.$$

Zatem, jeśli chcemy obliczyć wartość funkcji $y = f(x)$ z dokładnością ε , to jako kryterium zakończenia sumowania można przyjąć najmniejsze n spełniające nierówność

$$\frac{M_{n+1}}{(n+1)!} \cdot |x|^{n+1} < \varepsilon. \quad (1.35)$$

Przykład 1.12. Opracować algorytm obliczania wartości funkcji

$$f(x) = e^x$$

dla zadanej wartości argumentu x , zakładając że potrafimy wykonywać tylko podstawowe działania arytmetyczne.

Rozwiązanie. Funkcja $f(x) = e^x$ (patrz dodatek E) ma następujące rozwinięcie w szereg MacLaurina:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$$

Dla określenia kryterium (1.35) wyznaczamy

$$M_2 = \max_{0 \leq t \leq x} (e^t)^{(n+1)} = e^x < (2.718281829)^x.$$

Uwzględniając uwagi z rozdziału 1.5, możemy zapisać algorytm.

Algorytm 1.6. Obliczanie wartości $y = e^x$ z zadaną dokładnością ε .

$Exp(x, \varepsilon)$

- 1 $M_2 := (2.718281829)^x$
- 2 $k := 0$
- 3 $S_0 := 1$
- 4 $y := 1$
- 5 **while** $M_2 S_k \frac{|x|}{k+1} \geq \varepsilon$
- 6 **do** $S_{k+1} := S_k \cdot \frac{x}{k+1}$
- 7 $y := y + S_{k+1}$
- 8 $k := k + 1$ **end**

1.7. Zadania

Zad. 1.1. Wzór

$$N(t) = \frac{N_c}{1 + e^{-bt}} \quad \text{dla } t \geq 0$$

przy pewnych stałych dodatnich N_c , b jest dokładniejszym od wzoru (1.2) modelem opisującym wzrost populacji. Jaka jest zasadnicza różnica między tym modelem a modelem wyrażonym wzorem (1.2)? Co oznacza stała N_c ?

Zad. 1.2. Dysponujemy komputerem, w którym liczba zmiennoprzecinkowa zapisywana jest następująco (patrz przykład 1.6):

- znak liczby s na 1 bicie,
- mantysa znormalizowana m_6 na 6 bitach,
- cecha c_6 na 6 bitach (bit znaku cechy s_c oraz 5 bitów bezwzględnej wartości cechy).

Na tym komputerze zostanie wykonany następujący program:

Program ReadWrite;

```
var
  a      : Real;
begin {ReadWrite}
  read(a);
  write(a)
end {ReadWrite}.
```

Wyznaczyć:

- 1) binarną wartość zmiennej a po wykonaniu instrukcji $read(a)$,
- 2) liczbę wydrukowaną przez instrukcję $writeln(a)$.

Zmiennej a nadać wartość jednej z poniższych liczb:

- a) 23,
- b) -14,
- c) -19,
- d) 17,
- e) -29,
- f) 6.5,
- g) 5.25,
- h) -12.875,
- i) -13.75,
- j) 31.25,
- k) 1562.5,
- l) 0.3,
- m) 6.7,
- n) 25.5,
- o) 23.2,
- p) -18.75,
- q) -15.85,
- r) 21.67,
- s) 7.6,



P II 714873/5

- t) 0.38,
 u) -0.03,
 v) 0.125,
 w) 0.09,
 x) 100,
 y) -120.

Zad. 1.3. W pewnych sytuacjach możemy uniknąć utraty cyfr znaczących podczas wykonywania obliczeń, przekształcając wyrażenie arytmetyczne w sposób podobny do (1.30) w przykładzie 1.10. Wykonaj przekształcenia następujących wyrażeń arytmetycznych dla wyeliminowania zjawiska utraty cyfr znaczących podczas obliczeń.

- a) $f(x) = \frac{1 - \cos x}{x^2}$,
 b) $f(x) = \log(x+1) - \log x$, dla dużych x ,
 c) $f(x) = \sin(x+\alpha) - \sin \alpha$,
 d) $f(x) = \sqrt[3]{1+x} - 1$.

Zad. 1.4. Opracuj algorytmy obliczania wartości funkcji dla zadanego argumentu x , korzystając z rozwinięć funkcji w szereg potęgowy.

- a) $\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$,
 b) $\arctgx = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}$ dla $x^2 < 1$,
 c) $\ln x = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(x-1)^k}{k}$ dla $0 < x \leq 2$,
 d) $\sinh x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}$,
 e) $\text{Arsinh } x = \sum_{k=0}^{\infty} (-1)^k \frac{(2k)!}{2^{2k}(k!)^2(2k+1)} \cdot x^{2k+1}$ dla $x^2 < 1$,
 f) $\sqrt{1-x} = \sum_{k=0}^{\infty} (-1)^k \left(\frac{1}{2} \right)_k x^k$ dla $|x| \leq 1$.

Zad. 1.5. Opracuj algorytmy obliczania poniższych całek z dokładnością ε , dla zadanego $x > 0$.

- a) $\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ (funkcja gęstości rozkładu normalnego),
 b) $S(x) = \frac{2}{\sqrt{2\pi}} \int_0^x \sin t^2 dt$ (całka Frensela $S(x)$),

-
- c) $C(x) = \frac{2}{\sqrt{2\pi}} \int_0^x \cos t^2 dt$ (całka Frensela $C(x)$),
- d) $\gamma = \int_0^1 \left(\frac{1}{\ln t} + \frac{1}{1-t} \right) dt$ (stała Eulera γ),
- e) $lix = \int_0^x \frac{dt}{\ln t}$ (logarytm całkowy lix),
- f) $six = -\frac{\pi}{2} + \int_0^x \frac{\sin t}{t} dt$ (sinus całkowy six),
- g) $cix = \gamma + \ln x + \int_0^x \frac{\cos t - 1}{t} dt$ (cosinus całkowy cix).

Rozwiązywanie układów równań liniowych

2.1. Wprowadzenie

W bieżącym rozdziale będziemy zajmowali się metodami rozwiązywania układu n równań liniowych z n niewiadomymi

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= a_{10}, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= a_{20}, \\ \dots &\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= a_{n0}. \end{aligned} \tag{2.1}$$

Wprowadzając oznaczenia macierzowe

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad A_0 = \begin{bmatrix} a_{10} \\ a_{20} \\ \dots \\ a_{n0} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix},$$

układ równań (2.1) możemy zapisać w postaci

$$AX = A_0. \tag{2.2}$$

Zakładamy, że elementy a_{ij} są liczbami rzeczywistymi. Macierz A nazywamy *macierzą główną* układu (2.1), wektor A_0 *kolumną wyrazów wolnych*, a wektor X *wektorem niewiadomych*. *Rozwiązaniem układu równań* (2.1) nazywamy każdy wektor

$$\bar{X} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \dots \\ \bar{x}_n \end{bmatrix}$$

o składowych \bar{x}_j rzeczywistych, spełniający równanie

$$A\bar{X} = A_0. \tag{2.3}$$

Twierdzenie 2.1 formuluje warunek konieczny i wystarczający na to, aby układ równań (2.1) miał dokładnie jedno rozwiązanie.

Twierdzenie 2.1 (Cramer)

Układ równań liniowych (2.1) ma dokładnie jedno rozwiązanie wtedy i tylko wtedy, gdy $\det(A) \neq 0$.

Metody rozwiązywania układów równań liniowych dzielimy na dwie klasy:

- *metody skończone* pozwalające uzyskać rozwiązanie po wykonaniu skończonej liczby działań arytmetycznych. Przykładami metod skończonych są: *metoda wyznaczników Cramera, metoda eliminacji Gaussa, metoda LU, metoda QR*. Rozwiązanie wyznaczone metodą skończoną jest obarczone tylko błędami zaokrągleń;
- *metody iteracyjne*, polegające na wyznaczaniu ciągu wektorów $X^{(0)}, X^{(1)}, \dots, X^{(k)}, \dots$ zbieżnego do rozwiązania układu. Najczęściej stosowanymi metodami iteracyjnymi są: *metoda Jacobiego, metoda Gaussa-Seidela, metoda najszybszego spadku, metoda Czebyszewa*. Rozwiązanie wyznaczone metodą iteracyjną jest obarczone błędem metody oraz błędem zaokrągleń – jest więc przybliżone. Ale metody iteracyjne pozwalają wyznaczyć rozwiązanie z dowolną, z góry zadaną dokładnością.

2.2. Metoda wyznaczników Cramera

Rozwiązanie układu (2.1) można otrzymać, stosując wzory Cramera

$$x_k = \frac{\begin{vmatrix} a_{11} & \dots & a_{1,k-1} & a_{10} & a_{1,k+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2,k-1} & a_{20} & a_{2,k+1} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,k-1} & a_{n0} & a_{n,k+1} & \dots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}} \quad \text{dla } k = 1, 2, \dots, n. \quad (2.4)$$

Przykład 2.1. Stosując wzory Cramera, wyprowadzimy wzory na rozwiązanie układu dwóch równań liniowych z dwiema niewiadomymi x_1, x_2

$$a_{11}x_1 + a_{12}x_2 = b_1,$$

$$a_{21}x_1 + a_{22}x_2 = b_2.$$

Rozwiązanie

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{b_1 a_{22} - b_2 a_{12}}{a_{11} a_{22} - a_{21} a_{12}},$$

$$x_2 = \frac{\begin{vmatrix} a_{12} & b_1 \\ a_{21} & b_2 \\ \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{b_2 a_{12} - b_1 a_{21}}{a_{11} a_{22} - a_{21} a_{12}}.$$

Przykład 2.2. Stosując wzory Cramera, wyprowadzimy wzory na rozwiązańe układu trzech równań z trzema niewiadomymi x_1, x_2, x_3 :

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3. \end{aligned}$$

Rozwiązanie

$$\begin{aligned} x_1 &= \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} = \frac{b_1 a_{22} a_{33} + b_2 a_{32} a_{13} + b_3 a_{12} a_{23} - b_3 a_{22} a_{13} - a_{32} a_{23} b_1 - a_{33} b_2 a_{12}}{a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13} - a_{32} a_{23} a_{11} - a_{33} a_{21} a_{12}}, \\ x_2 &= \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} = \frac{a_{11} b_2 a_{33} + a_{21} b_3 a_{13} + a_{31} b_1 a_{23} - a_{31} b_2 a_{13} - b_3 a_{23} a_{11} - a_{33} a_{21} b_1}{a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13} - a_{32} a_{23} a_{11} - a_{33} a_{21} a_{12}}, \\ x_3 &= \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} = \frac{a_{11} a_{22} b_3 + a_{21} a_{32} b_1 + a_{31} a_{12} b_2 - a_{31} a_{22} b_1 - a_{32} b_2 a_{11} - b_3 a_{21} a_{12}}{a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13} - a_{32} a_{23} a_{11} - a_{33} a_{21} a_{12}}. \end{aligned}$$

Wyznaczmy złożoność obliczeniową metody wyznaczników Cramera. Wzory Cramera nakazują obliczenie $n + 1$ wyznaczników stopnia n . Zatem problem sprowadza się do znalezienia efektywnego algorytmu obliczania wyznacznika macierzy.

Jeśli decydujemy się na obliczanie wyznacznika, korzystając bezpośrednio z definicji

$$\det A = \sum_{\text{Perm}\{k_1, k_2, \dots, k_n\}} (-1)^{I_{\text{Perm}}} a_{1k_1} a_{2k_2} \cdots a_{nk_n} \quad (2.5)$$

(gdzie sumowanie rozciąga się na wszystkie permutacje $\{k_1, k_2, \dots, k_n\}$ zbioru $\{1, 2, \dots, n\}$, a I_{perm} oznacza liczbę inwersji w permutacji $\{k_1, k_2, \dots, k_n\}$), to musimy wykonać:

- $n!$ dodawań oraz
- $n \cdot n!$ mnożeń.

Łącznie musimy wykonać $(n+1)!$ działań arytmetycznych zmennoprzecinkowych. Tabela 2.1 daje wyobrażenie o czasie realizacji tego algorytmu przy założeniu, że do obliczeń używamy komputera wykonującego 10^8 operacji zmennoprzecinkowych na sekundę.

Tabela 2.1. Czas obliczania wyznacznika według wzoru (2.5)

n	Liczba operacji arytmetycznych	Czas wykonania	Jedn. czasu
4	120	1.20E-06	s
10	39916800	3.99E-01	s
20	5.10909E+19	5.11E+11	lat
50	1.55112E+66	1.55E+58	lat

Aby rozwiązać układ równań (2.1), trzeba obliczyć $n+1$ wyznaczników stopnia n oraz wykonać n dzielen. Zatem metoda ta wymaga wykonania $(n+1) \cdot (n+1)! + n$ działań arytmetycznych. Wobec tak dużej złożoności obliczeniowej, wzory Cramera stosuje się tylko dla $n \leq 4$.

2.3. Metoda eliminacji Gaussa

Rozwiązywanie układu równań liniowych metodą eliminacji Gaussa przebiega w dwóch etapach: pierwszy etap jest nazywany *etapem postępowania prostego (etapem eliminacji niewiadomych)*, drugi *etapem postępowania odwrotnego*.

Na etapie postępowania prostego wyjściowy układ równań zostaje przekształcony do postaci równoważnej (tzn. takiej, która ma dokładnie takie same rozwiązania, co układ wyjściowy) z trójkątną górną macierzą główną układu. Przekształcenie to jest realizowane w n krokach.

Krok 1 (eliminacja niewiadomej x_1 z równań 2, 3, ..., n).

Zakładamy, że w układzie wyjściowym, który zapiszemy jako

$$\begin{aligned} a_{11}^{(0)}x_1 + a_{12}^{(0)}x_2 + \dots + a_{1n}^{(0)}x_n &= a_{10}^{(0)}, \\ a_{21}^{(0)}x_1 + a_{22}^{(0)}x_2 + \dots + a_{2n}^{(0)}x_n &= a_{20}^{(0)}, \\ \dots & \\ a_{n1}^{(0)}x_1 + a_{n2}^{(0)}x_2 + \dots + a_{nn}^{(0)}x_n &= a_{n0}^{(0)}, \end{aligned} \tag{2.6}$$

element $a_{11}^{(0)} \neq 0$. Pierwsze równanie układu (2.6) dzielimy przez $a_{11}^{(0)}$, otrzymując

$$x_1 + a_{12}^{(1)}x_2 + \dots + a_{1n}^{(1)}x_n = a_{10}^{(1)}, \tag{2.7}$$

gdzie

$$a_{1j}^{(1)} = \frac{a_{1j}^{(0)}}{a_{11}^{(0)}} \quad \text{dla } j = 1, 2, \dots, n, 0.$$

Następnie od i -tego ($i = 2, 3, \dots, n$) równania układu (2.6) odejmujemy równanie (2.7) pomnożone przez a_{ii} , otrzymując

gdzie

$$a_{ij}^{(1)} = a_{ij}^{(0)} - a_{il}^{(0)} \cdot a_{lj}^{(1)} \quad \text{dla } i = 2, 3, \dots, n; \ j = 1, 2, \dots, n, 0.$$

Krok 2 (eliminacja niewiadomej x_2 z równań 3, 4, ..., n).

Schemat obliczeń kroku 1 powtarzamy w odniesieniu do układu równań (2.8). A więc zakładamy, że $a_{22}^{(1)} \neq 0$. Pierwsze równanie układu (2.8) dzielimy przez $a_{22}^{(1)}$, otrzymując

$$x_2 + a_{23}^{(2)} x_3 + \dots + a_{2n}^{(2)} x_n = a_{20}^{(2)}, \quad (2.9)$$

gdzie

$$a_{2j}^{(2)} = \frac{a_{2j}^{(1)}}{a_{22}^{(1)}} \quad \text{dla } j = 2, 3, \dots, n, 0.$$

Następnie od i -tego ($i = 3, 4, \dots, n$) równania układu (2.8) odejmujemy równanie (2.9) pomnożone przez $a_{i2}^{(1)}$, otrzymując

gdzie

$$a_{ij}^{(2)} = a_{ij}^{(1)} - a_{i2}^{(1)} \cdot a_{2j}^{(2)} \quad \text{dla } i = 3, 4, \dots, n; \ j = 2, 3, \dots, n, 0.$$

Kroki 3, 4, ..., $n - 1$.

Algorytm obliczeń w kolejnych krokach jest analogiczny do tych z kroków 1 i 2. Po wykonaniu obliczeń w $n - 1$ kroku, otrzymujemy dwa równania

$$x_{n-1} + a_{n-1,n}^{(n-1)} x_n = a_{n-1,0}^{(n-1)}, \quad (2.11)$$

gdzie

$$a_{n-1,j}^{(n-1)} = \frac{a_{n-1,j}^{(n-2)}}{a_{n-1,n-1}^{(n-2)}} \quad \text{dla } j = n, 0,$$

oraz

$$a_{nn}^{(n-1)} x_n = a_{n0}^{(n-1)}. \quad (2.12)$$

Krok n.

Zakładamy, że $a_{nn}^{(n-1)} \neq 0$. Równanie układu (2.12) dzielimy przez $a_{nn}^{(n-1)}$, otrzymując

$$x_n = a_{n0}^{(n)}. \quad (2.13)$$

Bardziej formalnie algorytm tego etapu został przedstawiamy jako algorytm 2.1.

Na etapie postępowania odwrotnego rozpatrujemy układ składający się z równań (2.7), (2.9), ..., (2.11), (2.13), tj.

$$\begin{aligned}
 x_1 + a_{12}^{(1)}x_2 + \dots + a_{1n}^{(1)}x_n &= a_{10}^{(1)}, \\
 x_2 + \dots + a_{2n}^{(2)}x_n &= a_{20}^{(2)}, \\
 &\dots, \\
 x_{n-1} + a_{n-1,n}^{(n-1)}x_n &= a_{n-1,0}^{(n-1)}, \\
 x_n &= a_{n0}^{(n)}.
 \end{aligned} \tag{2.14}$$

Algorytm rozwiązania tego układu jest oczywisty. Z ostatniego równania układu (2.14) otrzymujemy bezpośrednio wartość niewiadomej x_n . Wartość tę podstawiamy do równań $n-1, n-2, \dots, 1$. Teraz z równania $n-1$ możemy wyznaczyć wartość x_{n-1} itd. Symbolicznie algorytm postępowania odwrotnego został zapisany jako algorytm 2.2.

Algorytm 2.1. Metoda eliminacji Gaussa – etap eliminacji niewiadomych.

Etap postępowania prostego(n, A, A_0)

```

1   for k := 1 to n
2     do begin
3       for j := k to n
4         do  $a_{kj}^{(k)} := \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}}$ 
5          $a_{k0}^{(k)} := \frac{a_{k0}^{(k-1)}}{a_{kk}^{(k-1)}}$ 
6       for i := k + 1 to n
7         do begin
8           for j := k to n
9             do  $a_{ij}^{(k)} := a_{ij}^{(k-1)} - a_{ik}^{(k-1)}a_{kj}^{(k)}$ 
10             $a_{i0}^{(k)} := a_{i0}^{(k-1)} - a_{ik}^{(k-1)}a_{k0}^{(k)}$ 
11         end
12     end

```

Algorytm 2.2. Metoda eliminacji Gaussa – etap postępowania odwrotnego.

Etap postępowania odwrotnego(n, A, A_0, X)

```

1    $x_n := a_{n0}^{(n)}$ 
2   for i := n - 1 downto 1
3     do  $x_i := a_{i0}^{(i)} - \sum_{j=i+1}^n a_{ij}^{(i)}x_j$ 

```

Jak wynika z opisu metody, układ równań (2.1) można rozwiązać metodą eliminacji Gaussa, jeśli elementy $a_{kk}^{(k-1)}$ ($k = 1, 2, \dots, n$) są różne od zera. Między elementami $a_{kk}^{(k-1)}$ ($k = 1, 2, \dots, n$) oraz minorami głównymi macierzy głównej układu zachodzą związki wyrażone w twierdzeniu 2.2.

Twierdzenie 2.2. Zachodzą równości

$$M_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} = a_{11}^{(0)} \cdot a_{22}^{(1)} \cdot \dots \cdot a_{kk}^{(k-1)} \quad \text{dla } k = 1, 2, \dots, n . \quad (2.15)$$

Stąd prawdziwe jest

Twierdzenie 2.3. Układ równań liniowych (2.1) można rozwiązać metodą eliminacji Gaussa, jeśli wszystkie minory główne macierzy głównej układu są różne od zera.

Przykład 2.3. Rozwiązywanie układu równań liniowych

$$\begin{array}{lclcl} 1.2x_1 & +2.6x_2 & -0.1x_3 & 1.5x_4 & = 13.15, \\ 4.5x_1 & +9.8x_2 & -0.4x_3 & 5.7x_4 & = 49.84, \\ 0.1x_1 & -0.1x_2 & -0.3x_3 & -3.5x_4 & = -14.08, \\ 4.5x_1 & -5.2x_2 & +4.2x_3 & -3.4x_4 & = -46.51 \end{array}$$

metodą eliminacji Gaussa.

Rozwiązanie. Postępowanie proste

Krok 0

x_1	x_2	x_3	x_4	ai_0
1.20000	2.60000	-0.10000	1.50000	13.15000
4.50000	9.80000	-0.40000	5.70000	49.84000
0.10000	-0.10000	-0.30000	-3.50000	-14.08000
4.50000	-5.20000	4.20000	-3.40000	-46.51000

Krok 1 (eliminacja niewiadomej x_1 z równań 2, 3, 4)

x_1	x_2	x_3	x_4	ai_0
1.00000	2.16667	-0.08333	1.25000	10.95833
0.00000	0.04999	-0.02502	0.07500	0.52752
0.00000	-0.31667	-0.29167	-3.62500	-15.17583
0.00000	-14.95002	4.57499	-9.02500	-95.82249

Krok 2 (eliminacja niewiadomej x_2 z równań 3, 4)

x_1	x_2	x_3	x_4	ai_0
1.00000	2.16667	-0.08333	1.25000	10.95833
0.00000	1.00000	-0.50050	1.50030	10.55251
0.00000	0.00000	-0.45016	-3.14990	-11.83417
0.00000	0.00000	-2.90750	13.40452	61.93775

Krok 3 (eliminacja niewiadomej x_3 z równania 4)

x_1	x_2	x_3	x_4	ai_0
1.00000	2.16667	-0.08333	1.25000	10.95833
0.00000	1.00000	-0.50050	1.50030	10.55251
0.00000	0.00000	1.00000	6.99729	26.28881
0.00000	0.00000	0.00000	33.74914	138.37247

Krok 4 (wyznaczenie wartości niewiadomej x_4)

x_1	x_2	x_3	x_4	ai_0
1.00000	2.16667	-0.08333	1.25000	10.95833
0.00000	1.00000	-0.50050	1.50030	10.55251
0.00000	0.00000	1.00000	6.99729	26.28881
0.00000	0.00000	0.00000	1.00000	4.10003

Przykład 2.3 cd.

Postępowanie odwrotne

$$\begin{aligned}x_4 &= 4.10003 \\x_3 &= -2.40029 \\x_2 &= 3.19989 \\x_1 &= -1.29983\end{aligned}$$

Ostateczne rozwiązanie

$$\begin{aligned}x_1 &= -1.29983 \\x_2 &= 3.19989 \\x_3 &= -2.40029 \\x_4 &= 4.10003\end{aligned}$$

2.4. Metoda eliminacji Gaussa z pełnym wyborem elementu podstawowego

Warunki stosowalności metody eliminacji Gaussa, wyrażone w twierdzeniu 2.3, ograniczają istotnie wykorzystanie tej metody. Metodę tę obecnie zmodyfikujemy tak, aby mogła być stosowana do układu Cramera, tzn. do układu z $\det(A) \neq 0$.

Założymy, że układ równań (2.1) rozwiązuje metodą eliminacji Gaussa i zostało już wykonanych $k - 1$ kroków etapu postępowania prostego. Wyjściowy układ równań został przekształcony do układu postaci

$$\begin{aligned}x_1 + a_{12}^{(1)} x_2 + \dots + a_{1n}^{(1)} x_n &= a_{10}^{(1)}, \\x_2 + \dots + a_{2n}^{(2)} x_n &= a_{20}^{(2)}, \\&\dots, \\a_{kk}^{(k-1)} x_k + \dots + a_{kn}^{(k-1)} x_n &= a_{k0}^{(k-1)}, \\&\dots, \\a_{nk}^{(k-1)} x_k + \dots + a_{nn}^{(k-1)} x_n &= a_{n0}^{(k-1)}. \quad (2.16)\end{aligned}$$

Algorytm z pełnym wyborem elementu podstawowego jest następujący:

- wyszukujemy element $a_{rs}^{(k-1)}$ spełniający warunek:

$$\left| a_{rs}^{(k-1)} \right| = \max_{\substack{k \leq i \leq n \\ k \leq j \leq n}} \{ |a_{ij}^{(k-1)}| \}, \quad (2.17)$$

- 2) przestawiamy w układzie (2.16) równanie r z równaniem k oraz kolumnę s z kolumną k ,
 3) eliminujemy niewiadomą x_k z równań $k+1, k+2, \dots, n$, zgodnie z algorytmem k -tego kroku prostej eliminacji Gaussa.

Jeżeli $\det(A) \neq 0$, to żaden element podstawowy w tej metodzie nie będzie równy zeru.

Przykład 2.4. Rozwiązywanie układu równań liniowych

$$\begin{array}{l} 1.2 x_1 + 2.6 x_2 - 0.1 x_3 + 1.5 x_4 = 13.15 \\ 4.5 x_1 + 9.8 x_2 - 0.4 x_3 + 5.7 x_4 = 49.84 \\ 0.1 x_1 - 0.1 x_2 - 0.3 x_3 - 3.5 x_4 = -14.08 \\ 4.5 x_1 - 5.2 x_2 + 4.2 x_3 - 3.4 x_4 = -46.51 \end{array}$$

metodą eliminacji Gaussa z pełnym wyborem elementu podstawowego.

Rozwiązanie. Postępowanie proste.

Krok 0

x_1	x_2	x_3	x_4	ai_0
1.20000	2.60000	-0.10000	1.50000	13.15000
4.50000	9.80000	-0.40000	5.70000	49.84000
0.10000	-0.10000	-0.30000	-3.50000	-14.08000
4.50000	-5.20000	4.20000	-3.40000	-46.51000

Krok 1 ($a_{rs} = a_{22} = 9.80000$. Przestawiamy wiersz 1 z 2 oraz kolumnę 1 z 2)

x_2	x_1	x_3	x_4	ai_0
9.80000	4.50000	-0.40000	5.70000	49.84000
2.60000	1.20000	-0.10000	1.50000	13.15000
-0.10000	0.10000	-0.30000	-3.50000	-14.08000
-5.20000	4.50000	4.20000	-3.40000	-46.51000

Eliminacja niewiadomej x_2 z równań 2, 3, 4)

x_2	x_1	x_3	x_4	ai_0
1.00000	0.45918	-0.04082	0.58163	5.08571
0.00000	0.00613	0.00613	-0.01224	-0.07285
0.00000	0.14592	-0.30408	-3.44184	-13.57143
0.00000	6.88774	3.98774	-0.37552	-20.06431

Krok 2 ($a_{rs} = a_{42} = 6.88774$. Przestawiamy wiersz 2 z 4)

x_2	x_1	x_3	x_4	ai_0
1.00000	0.45918	-0.04082	0.58163	5.08571
0.00000	6.88774	3.98774	-0.37552	-20.06431
0.00000	0.14592	-0.30408	-3.44184	-13.57143
0.00000	0.00613	0.00613	-0.01224	-0.07285

Eliminacja niewiadomej x_3 z równań 3, 4

x_2	x_1	x_3	x_4	ai_0
1.00000	0.45918	-0.04082	0.58163	5.08571
0.00000	1.00000	0.57896	-0.05452	-2.91305
0.00000	0.00000	-0.38856	-3.43388	-13.14636
0.00000	0.00000	0.00258	-0.01191	-0.05499

Krok 3 ($a_{rs} = a_{34} = -3.43388$. Przestawiamy kolumnę 3 z 4)

x_2	x_1	x_4	x_3	ai_0
1.00000	0.45918	0.58163	-0.04082	5.08571
0.00000	1.00000	-0.05452	0.57896	-2.91305
0.00000	0.00000	-3.43388	-0.38856	-13.14636
0.00000	0.00000	-0.01191	0.00258	-0.05499

Eliminacja niewiadomej x_4 z równania 4

x_2	x_1	x_4	x_3	ai_0
1.00000	0.45918	0.58163	-0.04082	5.08571
0.00000	1.00000	-0.05452	0.57896	-2.91305
0.00000	0.00000	1.00000	0.11315	3.82843
0.00000	0.00000	0.00000	0.00393	-0.00939

Krok 4 (wyznaczenie wartości niewiadomej x_2)

x_2	x_1	x_4	x_3	ai_0
1.00000	0.45918	0.58163	-0.04082	5.08571
0.00000	1.00000	-0.05452	0.57896	-2.91305
0.00000	0.00000	1.00000	0.11315	3.82843
0.00000	0.00000	0.00000	1.00000	-2.38931

Postępowanie odwrotne

$$\begin{aligned}x_3 &= -2.38931, \\x_4 &= 4.09878, \\x_1 &= -1.30627, \\x_2 &= 3.20402.\end{aligned}$$

Rozwiązanie

$$\begin{aligned}x_1 &= -1.30627, \\x_2 &= 3.20402, \\x_3 &= -2.38931, \\x_4 &= 4.09878\end{aligned}$$

2.5. Metoda eliminacji Jordana

W tej metodzie rozwiązanie układu równań liniowych uzyskujemy w jednym etapie. Podobnie jak w metodzie eliminacji Gaussa, obliczenia przebiegają w n krokach.

Krok 1 (eliminacja niewiadomej x_1 z równań 2, 3, ..., n).

Zakładamy, że w układzie wyjściowym

element $a_{11}^{(0)} \neq 0$. Pierwsze równanie układu (2.18) dzielimy przez $a_{11}^{(0)}$. Następnie od i -tego równania ($i = 2, 3, \dots, n$) odejmujemy przekształcony wiersz pierwszy pomnożony przez $a_{ii}^{(0)}$. W efekcie otrzymujemy układ równań

gdzie:

$$a_{1j}^{(1)} = \frac{a_{1j}^{(0)}}{a_{11}^{(0)}} \quad \text{dla } j = 1, 2, \dots, n, 0,$$

$$a_{ij}^{(1)} = a_{ij}^{(0)} - a_{il}^{(0)} \cdot a_{lj}^{(1)} \quad \text{dla } i = 2, 3, \dots, n; \ j = 1, 2, \dots, n, 0.$$

Krok 2 (eliminacja niewiadomej x_2 z równań 1, 3, ..., n).

Zakładamy, że $a_{22}^{(1)} \neq 0$. Drugie równanie układu (2.19) dzielimy przez $a_{22}^{(1)}$. Następnie od i -tego ($i = 1, 3, \dots, n$) równania tego układu odejmujemy przekształcony wiersz drugi pomnożony przez $a_{i2}^{(1)}$. Otrzymujemy układ równań

$$\begin{aligned} x_1 + a_{13}^{(2)}x_3 + \dots + a_{1n}^{(2)}x_n &= a_{10}^{(2)}, \\ x_2 + a_{23}^{(2)}x_3 + \dots + a_{2n}^{(2)}x_n &= a_{20}^{(2)}, \\ a_{33}^{(2)}x_3 + \dots + a_{3n}^{(2)}x_n &= a_{30}^{(2)}, \\ \dots, \\ a_{n3}^{(2)}x_3 + \dots + a_{nn}^{(2)}x_n &= a_{n0}^{(2)}, \end{aligned}$$

gdzie

$$a_{2j}^{(2)} = \frac{a_{2j}^{(1)}}{a_{22}^{(1)}} \quad \text{dla } j = 2, 3, \dots, n, 0,$$

$$a_{ij}^{(2)} = a_{ij}^{(1)} - a_{i2j}^{(1)} \cdot a_{22j}^{(2)} \quad \text{dla } i = 1, 3, \dots, n; \ j = 2, 3, \dots, n, 0.$$

Kroki 3, 4, ..., n

W k -tym ($k = 3, 4, \dots, n$) kroku algorytmu eliminujemy niewiadomą x_k z równań $1, 2, \dots, k-1, k+1, \dots, n$, postępując podobnie jak w krokach 1 i 2 metody. W konsekwencji po n -tym kroku otrzymujemy układ równań

$$\begin{aligned}
 x_1 &= a_{10}^{(n)}, \\
 x_2 &= a_{20}^{(n)}, \\
 &\dots, \\
 x_{n-1} &= a_{n-1,0}^{(n)}, \\
 x_n &= a_{n0}^{(n)},
 \end{aligned} \tag{2.20}$$

reprezentujący gotowe rozwiązańe układu (2.1).

Przykład 2.5. Rozwiązywanie układu równań liniowych

$$\begin{array}{ccccc}
 1.2 x_1 & +2.6 x_2 & -0.1 x_3 & 1.5 x_4 & = 13.15 \\
 4.5 x_1 & +9.8 x_2 & -0.4 x_3 & 5.7 x_4 & = 49.84 \\
 0.1 x_1 & -0.1 x_2 & -0.3 x_3 & -3.5 x_4 & = -14.08 \\
 4.5 x_1 & -5.2 x_2 & +4.2 x_3 & -3.4 x_4 & = -46.51
 \end{array}$$

metodą eliminacji Jordana (bez wyboru elementów podstawowych).

Rozwiązanie

Krok 0

x_1	x_2	x_3	x_4	a_{i0}
1.20000	2.60000	-0.10000	1.50000	13.15000
4.50000	9.80000	-0.40000	5.70000	49.84000
0.10000	-0.10000	-0.30000	-3.50000	-14.08000
4.50000	-5.20000	4.20000	-3.40000	-46.51000

Krok 1 (eliminacja niewiadomej x_1 z równań 2,3,4)

x_1	x_2	x_3	x_4	a_{i0}
1.00000	2.16667	-0.08333	1.25000	10.95833
0.00000	0.04999	-0.02502	0.07500	0.52752
0.00000	-0.31667	-0.29167	-3.62500	-15.17583
0.00000	-14.95002	4.57499	-9.02500	-95.82249

Krok 2 (eliminacja niewiadomej x_2 z równań 1, 3, 4)

x_1	x_2	x_3	x_4	a_{i0}
1.00000	0.00000	1.00109	-2.00066	-11.90548
0.00000	1.00000	-0.50050	1.50030	10.55251
0.00000	0.00000	-0.45016	-3.14990	-11.83417
0.00000	0.00000	-2.90750	13.40452	61.93775

Krok 3 (eliminacja niewiadomej x_3 z równań 1, 2, 4)

x_1	x_2	x_3	x_4	a_{i0}
1.00000	0.00000	0.00000	-9.00558	-38.22294
0.00000	1.00000	0.00000	5.00244	23.71006
0.00000	0.00000	1.00000	6.99729	26.28881
0.00000	0.00000	0.00000	33.74914	138.37247

Krok 4 (eliminacja niewiadomej x_4 z równań 1, 2, 3)

x_1	x_2	x_3	x_4	ai_0
1.00000	0.00000	0.00000	0.00000	-1.29979
0.00000	1.00000	0.00000	0.00000	3.19991
0.00000	0.00000	1.00000	0.00000	-2.40029
0.00000	0.00000	0.00000	1.00000	4.10003

Rozwiązanie

$$x_1 = -1.29979,$$

$$x_2 = 3.19991,$$

$$x_3 = -2.40029,$$

$$x_4 = 4.10003.$$

2.6. Metoda LU

W tej metodzie rozwiązywanie układu równań (2.1) przebiega w trzech etapach.

- Macierz główną układu A przedstawiamy w postaci iloczynu

$$A = LU \quad (2.21)$$

macierzy trójkątnej dolnej L i macierzy trójkątnej górnej U

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & \dots & u_{1n} \\ 0 & 1 & \dots & u_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (2.22)$$

To przedstawienie umożliwia zapis układu (2.1) w postaci $AX = (LU)X = L(UX) = A_0$.

- Rozwiążujemy układ równań liniowych

$$LY = A_0 \quad (2.23)$$

lub w postaci rozwiniętej

$$\begin{aligned} l_{11}y_1 &= a_{10}, \\ l_{21}y_1 + l_{22}y_2 &= a_{20}, \\ \dots & \\ l_{n1}y_1 + l_{n2}y_2 + \dots + l_{nn}y_n &= a_{n0}. \end{aligned} \quad (2.24)$$

- Rozwiążujemy układ równań liniowych

$$UX = Y. \quad (2.25)$$

lub w postaci rozwiniętej

$$\begin{aligned} x_1 + u_{12}x_2 + \dots + u_{1n}x_n &= y_1, \\ x_2 + \dots + u_{2n}x_n &= y_2, \\ \dots & \\ x_n &= y_n. \end{aligned} \quad (2.26)$$

Poniżej prezentujemy algorytmy trzech etapów metody.

Algorytm 2.3. Metoda LU – rozkład trójkątno-trójkątny.

Rozkład trójkątno-trójkątny (n, A, L, U)

```

1      for  $k := 1$  to  $n$ 
2          do begin
3              for  $i := k$  to  $n$ 
4                  do  $l_{ik} := a_{ik} - \sum_{p=1}^{k-1} l_{ip} u_{pk}$ 
5                  for  $j := k+1$  to  $n$ 
6                       $u_{kj} := \frac{a_{ik} - \sum_{p=1}^{k-1} l_{kp} u_{pj}}{l_{kk}}$ 
7          end

```

Algorytm 2.4. Metoda LU – rozwiązywanie układu równań $LY = A_0$.

Wyznaczanie wektora $Y (n, L, A_0, Y)$

```

1      for  $k := 1$  to  $n$ 
2          do  $y_k := \frac{a_{k0} - \sum_{p=1}^{k-1} l_{kp} y_p}{l_{kk}}$ 

```

Algorytm 2.5. Metoda LU – rozwiązywanie układu równań $UX = Y$.

Wyznaczanie wektora $X (n, U, Y, X)$

```

1      for  $k := n$  downto 1
2          do  $x_k := y_k - \sum_{j=k+1}^n u_{kj} x_j$ 

```

W metodzie LU założyliśmy możliwość rozkładu trójkątno-trójkątnego macierzy głównej układu. W twierdzeniu 2.3 zostały sformułowane możliwości wykonania takiego rozkładu.

Twierdzenie 2.3. Jeżeli wszystkie minory główne macierzy A

$$M_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} \neq 0 \quad \text{dla } k = 1, 2, \dots, n,$$

to macierz A możemy przedstawić w postaci iloczynu macierzy trójkątnej dolnej L i macierzy trójkątnej górnej U . Rozkład ten jest jednoznaczny pod warunkiem, że ustalimy elementy głównej przekątnej jednej z macierzy trójkątnych.

Przykład 2.6. Rozwiązywanie układu równań liniowych

$$\begin{array}{lllll}
 1.2 x_1 & +2.6 x_2 & -0.1 x_3 & 1.5 x_4 & = 13.15 \\
 4.5 x_1 & +9.8 x_2 & -0.4 x_3 & 5.7 x_4 & = 49.84 \\
 0.1 x_1 & -0.1 x_2 & -0.3 x_3 & -3.5 x_4 & = -14.08 \\
 4.5 x_1 & -5.2 x_2 & +4.2 x_3 & -3.4 x_4 & = -46.51
 \end{array}$$

metodą LU .

Rozwiązanie

1. Wykonanie rozkładu $A = LU$, według algorytmu 2.3.

$k = 1$	$l_{11} = 1.20000$ $u_{11} = 1.00000$	$l_{21} = 4.50000$ $u_{12} = 2.16667$	$l_{31} = 0.10000$ $u_{13} = -0.08333$	$l_{41} = 4.50000$ $u_{14} = 1.25000$
$k = 2$		$l_{22} = 0.04999$ $u_{22} = 1.00000$	$l_{32} = -0.31667$ $u_{23} = -0.50040$	$l_{42} = -14.95002$ $u_{24} = 1.50030$
$k = 3$			$l_{33} = -0.45013$ $u_{33} = 1.00000$	$l_{43} = -2.90601$ $u_{34} = 6.99776$
$k = 4$				$l_{44} = 33.74008$ $u_{44} = 1.00000$

2. Rozwiązywanie układu $LY = A_0$:

$$\begin{array}{llll}
 1.20000 y_1 & & & = 13.15 \\
 4.50000 y_1 & +0.04999 y_2 & & = 49.84 \\
 0.10000 y_1 & -0.31667 y_2 & -0.45013 y_3 & = -14.08 \\
 4.50000 y_1 & -14.95002 y_2 & -2.90601 y_3 & +33.74008 y_4 = -46.51
 \end{array}$$

Obliczając według algorytmu 2.4, otrzymujemy

$$y_1 = 10.95833, \quad y_2 = 10.55211, \quad y_3 = 26.29084, \quad y_4 = 4.10008.$$

3. Rozwiązywanie układu $UX = Y$:

$$\begin{array}{llll}
 1.00000 x_1 & +2.16667 x_2 & -0.08333 x_3 & +1.25000 x_4 = 10.95833 \\
 & 1.00000 x_2 & -0.50040 x_3 & +1.50030 x_4 = 10.55211 \\
 & & 1.00000 x_3 & +6.99776 x_4 = 26.29084 \\
 & & & 1.00000 x_4 = 4.10008
 \end{array}$$

Wykonując obliczenia według algorytmu 2.5, mamy

$$x_4 = 4.09997, \quad x_3 = -2.39977, \quad x_2 = 3.20008, \quad x_1 = -1.30012.$$

2.7. Metoda QR

W tej metodzie rozwiązywanie układu równań (2.1) przebiega w trzech etapach.

1. Macierz główną układu A przedstawiamy w postaci iloczynu

$$A = QR \tag{2.27}$$

macierzy ortogonalnej Q i macierzy trójkątnej górnej R

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \dots & \dots & \dots & \dots \\ q_{n1} & q_{n2} & \dots & q_{nn} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & r_{nn} \end{bmatrix}$$

To przedstawienie umożliwia zapis układu (2.1) w postaci $AX = (QR)X = Q(RX) = A_0$.

2. Rozwiążujemy układ równań liniowych

$$QY = A_0. \quad (2.28)$$

Rozwiążanie ma prosty algorytm: $Y = Q^{-1}A_0 = Q^T A_0$.

3. Rozwiążujemy układ równań liniowych

$$RX = Y. \quad (2.29)$$

lub w postaci rozwiniętej

$$r_{11}x_1 + r_{12}x_2 + \dots + r_{1n}x_n = y_1,$$

$$r_{22}x_2 + \dots + r_{2n}x_n = y_2,$$

.....,

$$r_{nn}x_n = y_n.$$

Układ równań (2.2) możemy przekształcić do postaci (2.29) metodą obrotów Givensa. Metoda obrotów Givensa polega na wielokrotnym mnożeniu równania $AX = A_0$ przez macierze postaci

$$T_{ij}(\alpha) = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ & & & & & & \\ & & & \dots & & & \\ 0 & \dots & \cos \alpha & \dots & \sin \alpha & \dots & 0 \\ & \dots & \dots & \dots & \dots & \dots & \\ 0 & \dots & -\sin \alpha & \dots & \cos \alpha & \dots & 0 \\ & & & & & & \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \quad \begin{array}{l} i - \text{ta kolumna} \\ \downarrow \\ i - \text{ty wiersz} \\ \uparrow \\ j - \text{ty wiersz} \\ \dots \\ j - \text{ta kolumna} \end{array}$$

(różniącej się od macierzy jednostkowej wartościami elementów $t_{ii} = t_{jj} = \cos \alpha$, $t_{ij} = -t_{ji} = \sin \alpha$). Macierze $T_{ij}(\alpha)$ są macierzami ortogonalnymi dla dowolnego rzeczywistego α .

Zapiszmy równanie (2.1) jako $A^{(0)}X = A_0^{(0)}$. Jeśli to równanie pomnożymy lewostronnie przez macierz $T_{ij}(\alpha)$, to otrzymamy

$$A^{(1)}X = (T_{ij}(\alpha)A^{(0)})X = T_{ij}(\alpha)A^{(0)}X = T_{ij}(\alpha)A_0^{(0)} = A_0^{(1)},$$

przy czym elementy macierzy głównej $A^{(1)}$ i wektora wyrazów wolnych $A_0^{(1)}$ wyrażają się wzorami:

$$a_{ik}^{(1)} = \cos \alpha \cdot a_{ik}^{(0)} + \sin \alpha \cdot a_{jk}^{(0)} \quad \text{dla } k = 1, 2, \dots, n,$$

$$a_{0i}^{(1)} = \cos \alpha \cdot a_{0i}^{(0)} + \sin \alpha \cdot a_{0j}^{(0)},$$

$$a_{jk}^{(1)} = -\sin \alpha \cdot a_{ik}^{(0)} + \cos \alpha \cdot a_{jk}^{(0)} \quad \text{dla } k = 1, 2, \dots, n,$$

$$a_{0j}^{(1)} = -\sin \alpha \cdot a_{0i}^{(0)} + \cos \alpha \cdot a_{0j}^{(0)},$$

$$a_{lk}^{(1)} = a_{lk}^{(0)} \quad \text{dla } l \neq i, j; k = 1, 2, \dots, n,$$

$$a_{0l}^{(1)} = a_{0l}^{(0)} \quad \text{dla } l \neq i, j.$$

Powyższe równości zachodzą dla dowolnego rzeczywistego α . Wyznaczymy teraz takie α , dla którego $a_{jk}^{(1)} = 0$. Zakładając, że istnieje $a_{ik}^{(0)} \neq 0$, mamy

$$0 = a_{jk}^{(1)} = -\sin \alpha \cdot a_{ik}^{(0)} + \cos \alpha \cdot a_{jk}^{(0)}.$$

Stąd

$$\operatorname{tg} \alpha = \frac{a_{jk}^{(0)}}{a_{ik}^{(0)}} \text{ lub } \alpha = \operatorname{arctg} \left(\frac{a_{jk}^{(0)}}{a_{ik}^{(0)}} \right) \text{ dla pewnego } k (1 \leq k \leq n).$$

Poniższy algorytm przekształca wyjściowy układ równań w układ równoważny o postaci (2.30) poprzez $n(n-1)/2$ -krotne lewostronne mnożenie równania macierzowego przez macierze obrotu, zamieniając na zera kolejno elementy leżące na pozycjach $(2, 1), (3, 1), \dots, (n, 1), (3, 2), (4, 2), \dots, (n, 2), \dots, (n, n-1)$.

Algorytm 2.6. Metoda QR – przekształcenie układu równań $AX = A_0$ do postaci $RX = Y$ metodą obrotów Givensa.

Metoda obrotów Givensa (n, A, A_0)

```

1       $p := 0$ 
2      for  $i := 1$  to  $n-1$ 
3          do for  $j := i+1$  to  $n$ 
4              do begin
5                   $p := p + 1$ 
6                   $\alpha := \operatorname{arctg} \left( \frac{a_{ji}^{(p-1)}}{a_{ii}^{(p-1)}} \right)$ 
7                   $A^{(p)} := T_{ij}(\alpha)A^{(p-1)}$ 
8                   $A_0^{(p)} := T_{ij}(\alpha)A_0^{(p-1)}$ 
9              end
```

Metoda obrotów Givensa realizuje dwa początkowe etapy metody QR. Po wykonaniu algorytmu 2.6 macierz główna układu jest macierzą R , a kolumna wyrazów wolnych wektorem Y .

Układ równań (2.29) rozwiązuje według algorytmu 2.7.

Algorytm 2.7. Metoda QR – rozwiązywanie układu $RX = Y$

Rozwiązywanie układu $RX = Y$

1 **for** $k := n$ **downto** 1

2 **do** $x_k := \frac{y_k - \sum_{j=k+1}^n r_{kj}x_j}{r_{kk}}$

Przykład 2.7. Rozwiązywanie układu równań liniowych

$$\begin{array}{rrrrr} 1.2 x_1 & +2.6 x_2 & -0.1 x_3 & 1.5 x_4 & = 13.15, \\ 4.5 x_1 & +9.8 x_2 & -0.4 x_3 & 5.7 x_4 & = 49.84, \\ 0.1 x_1 & -0.1 x_2 & -0.3 x_3 & -3.5 x_4 & = -14.08, \\ 4.5 x_1 & -5.2 x_2 & +4.2 x_3 & -3.4 x_4 & = -58.51 \end{array}$$

metodą QR.

Rozwiązanie

1. Przekształcenie metodą obrotów Givensa układu $AX = A_0$ do postaci $RX = Y$.

Oznaczmy

$$A^{(0)} = \begin{bmatrix} 1.20000 & 2.60000 & -0.10000 & 1.50000 \\ 4.50000 & 9.80000 & -0.40000 & 5.70000 \\ 0.10000 & -0.10000 & -0.30000 & -3.50000 \\ 4.50000 & -5.20000 & 4.20000 & -3.40000 \end{bmatrix}, \quad A_0^{(0)} = \begin{bmatrix} 13.15000 \\ 49.84000 \\ -14.08000 \\ -58.51000 \end{bmatrix}.$$

Kładąc $p = 1$, $i = 1$, $j = 2$, mamy

$$\operatorname{tg}(\alpha) = \frac{a_{21}}{a_{11}} = 3.75000, \quad \alpha = 1.31019,$$

$$T_{12}(\alpha) = \begin{bmatrix} 0.25766 & 0.96623 & 0.00000 & 0.00000 \\ -0.96623 & 0.25766 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix},$$

$$A^{(1)} = \begin{bmatrix} 4.65725 & 10.13903 & -0.41226 & 5.89403 \\ 0.00000 & 0.01288 & -0.00644 & 0.01932 \\ 0.10000 & -0.10000 & -0.30000 & -3.50000 \\ 4.50000 & -5.20000 & 4.20000 & -3.40000 \end{bmatrix}, \quad A_0^{(1)} = \begin{bmatrix} 51.54541 \\ 0.13592 \\ -14.08000 \\ -58.51000 \end{bmatrix}.$$

Następnie kładąc $p = 2, i = 1, j = 3$, mamy

$$\begin{aligned} \operatorname{tg}(\alpha) &= \frac{a_{31}}{a_{11}} = 0.02147, \quad \alpha = 0.02147, \\ T_{13}(\alpha) &= \begin{bmatrix} 0.99977 & 0.00000 & 0.02147 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 & 0.00000 \\ -0.02147 & 0.00000 & 0.99977 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}, \\ A^{(2)} &= \begin{bmatrix} 4.65833 & 10.13454 & -0.41861 & 5.81754 \\ 0.00000 & 0.01288 & -0.00644 & 0.01932 \\ 0.00000 & -0.31763 & -0.20108 & -3.562572 \\ 4.50000 & -5.20000 & 4.20000 & -3.40000 \end{bmatrix}, \quad A_0^{(2)} = \begin{bmatrix} 51.23128 \\ 0.13592 \\ -15.18328 \\ -58.51000 \end{bmatrix}. \end{aligned}$$

Następnie kładąc $p = 3, i = 1, j = 4$, otrzymujemy

$$\begin{aligned} \operatorname{tg}(\alpha) &= \frac{a_{41}}{a_{11}} = -24.65478, \\ \alpha &= -1.53026, \quad T_{14}(\alpha) = \begin{bmatrix} 1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.04053 & -0.99918 & 0.00000 \\ 0.00000 & 0.99918 & 0.04053 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}, \\ A^{(3)} &= \begin{bmatrix} 6.47688 & 3.67615 & 2.61700 & 1.82186 \\ 0.00000 & 0.31789 & 0.29058 & 3.62352 \\ 0.00000 & 0.00000 & -0.01823 & -0.12763 \\ 0.00000 & -10.78123 & 3.31158 & -6.48726 \end{bmatrix}, \quad A_0^{(3)} = \begin{bmatrix} -3.80476 \\ 15.17631 \\ -0.47952 \\ -77.67617 \end{bmatrix}. \end{aligned}$$

Następnie kładąc $p = 4, i = 2, j = 3$, otrzymujemy

$$\begin{aligned} \operatorname{tg}(\alpha) &= \frac{a_{23}}{a_{22}} = -24.65478, \\ \alpha &= -1.53026, \quad T_{23}(\alpha) = \begin{bmatrix} 1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.04053 & -0.99918 & 0.00000 \\ 0.00000 & 0.99918 & 0.04053 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}, \\ A^{(4)} &= \begin{bmatrix} 6.47688 & 3.67615 & 2.61700 & 1.82186 \\ 0.00000 & 0.31789 & 0.29058 & 3.62352 \\ 0.00000 & 0.00000 & -0.01823 & -0.12763 \\ 0.00000 & -10.78123 & 3.31158 & -6.48726 \end{bmatrix}, \quad A_0^{(4)} = \begin{bmatrix} -3.80476 \\ 15.17631 \\ -0.47952 \\ -77.67617 \end{bmatrix}. \end{aligned}$$

Następnie kładąc $p = 5, i = 2, j = 4$, mamy

$$\operatorname{tg}(\alpha) = \frac{a_{24}}{a_{22}} = -33.91475, \alpha = -1.54132,$$

$$T_{24}(\alpha) = \begin{bmatrix} 1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.02947 & 0.00000 & -0.99957 \\ 0.00000 & 0.00000 & 1.00000 & 0.00000 \\ 0.00000 & 0.99957 & 0.00000 & 0.02947 \end{bmatrix},$$

$$A^{(5)} = \begin{bmatrix} 6.47688 & 3.67615 & 2.61700 & 1.82186 \\ 0.00000 & 10.78591 & -3.30157 & 6.59124 \\ 0.00000 & 0.00000 & -0.01823 & -0.12763 \\ 0.00000 & 0.00000 & 0.38806 & 3.43075 \end{bmatrix}, A_0^{(5)} = \begin{bmatrix} -3.80476 \\ 78.08972 \\ -0.47952 \\ 12.88038 \end{bmatrix}.$$

Następnie kładąc $p = 6, i = 3, j = 4$, otrzymujemy

$$\operatorname{tg}(\alpha) = \frac{a_{34}}{a_{33}} = -21.28336, \alpha = -1.52385,$$

$$T_{34}(\alpha) = \begin{bmatrix} 1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.04693 & -0.99890 \\ 0.00000 & 0.00000 & 0.99890 & 0.04693 \end{bmatrix},$$

$$A^{(6)} = \begin{bmatrix} 6.47688 & 3.67615 & 2.61700 & 1.82186 \\ 0.00000 & 10.78591 & -3.30157 & 6.59124 \\ 0.00000 & 0.00000 & -0.38848 & -3.43296 \\ 0.00000 & 0.00000 & 0.00000 & 0.03353 \end{bmatrix}, A_0^{(6)} = \begin{bmatrix} -3.80476 \\ 78.08972 \\ -12.88869 \\ 0.12552 \end{bmatrix}.$$

2. Rozwiązywanie układu równań $AX = Y$, czyli

$$\begin{bmatrix} 6.47688 & 3.67615 & 2.61700 & 1.82186 \\ 0.00000 & 10.78591 & -3.30157 & 6.59124 \\ 0.00000 & 0.00000 & -0.38848 & -3.43296 \\ 0.00000 & 0.00000 & 0.00000 & 0.03353 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3.80476 \\ 78.08972 \\ -12.88869 \\ 0.12552 \end{bmatrix}.$$

Postępując według algorytmu 2.7, otrzymujemy rozwiązanie

$$x_4 = 3.74392, x_3 = 0.09253, x_2 = 4.98042, x_1 = -4.50475.$$

2.8. Warianty metody iteracji prostej

Omawianie metod iteracyjnych rozpoczęmy od przedstawienia macierzy głównej układu równań (2.1) w postaci sumy

$$A = M + D + N, \quad (2.30)$$

gdzie

$$M = \begin{bmatrix} 0 & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix}, D = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}, N = \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Zakładając, że $a_{ii} \neq 0$ dla $i=1, 2, \dots, n$, układ równań

$$AX = (M + D + N)X = A_0$$

możemy zapisać w równoważnej postaci

$$X = -D^{-1}(M + N)X + D^{-1}A_0. \quad (2.31)$$

W dalszym ciągu układ ten będziemy zapisywać jako

$$X = HX + G \quad (2.32)$$

lub w postaci rozwiniętej

$$\begin{aligned} x_1 &= h_{12}x_2 + h_{13}x_3 + \dots + h_{1,n-1}x_{n-1} + h_{1n}x_n + g_1, \\ x_2 &= h_{21}x_1 + h_{23}x_3 + \dots + h_{2,n-1}x_{n-1} + h_{2n}x_n + g_2, \\ &\dots, \\ x_n &= h_{n1}x_1 + h_{n2}x_2 + h_{n3}x_3 + \dots + h_{n,n-1}x_{n-1} + g_n, \end{aligned} \quad (2.33)$$

gdzie:

$$\begin{aligned} h_{ii} &= 0 && \text{dla } i = 1, 2, \dots, n, \\ h_{ij} &= -\frac{a_{ij}}{a_{ii}} && \text{dla } i = 1, 2, \dots, n; j = 1, 2, \dots, i-1, i+1, \dots, n, \\ g_i &= \frac{a_{i0}}{a_{ii}} && \text{dla } i = 1, 2, \dots, n. \end{aligned}$$

2.8.1. Metoda iteracji prostej (metoda Jacobiego)

W metodzie Jacobiego wybieramy dowolny wektor $X^{(0)}$ i tworzymy ciąg kolejnych przybliżeń $X^{(k)}$ ($k = 1, 2, \dots$) według wzorów

$$x_i^{(k)} = \sum_{j=1}^{i-1} h_{ij} x_j^{(k-1)} + \sum_{j=i+1}^n h_{ij} x_j^{(k-1)} + g_i \quad \text{dla } i = 1, 2, \dots, n. \quad (2.34)$$

Twierdzenie 2.4. Jeżeli

$$\|D^{-1}(M + N)\|_1 \leq q < 1 \text{ lub } \|D^{-1}(M + N)\|_\infty \leq q < 1,$$

to:

- a) ciąg (2.34) jest zbieżny do rozwiązania układu równań (2.1) niezależnie od wyboru wektora $X^{(0)}$,
- b) zachodzą oszacowania

$$\|\bar{X} - X^{(k)}\| \leq \frac{q}{1-q} \cdot \|X^{(k)} - X^{(k-1)}\| \leq \frac{q^k}{1-q} \cdot \|X^{(1)} - X^{(0)}\|, \quad (2.35)$$

dla $k = 1, 2, \dots$

Opierając się na twierdzeniu 2.4, sformułujemy algorytm rozwiązywania układu równań (2.1) metodą Jacobiego (metodą iteracji prostej).

Algorytm 2.8. Rozwiązywanie układu równań liniowych $AX = A_0$ metodą Jacobiego.

Metoda Jacobiego ($n, A, A_0, \varepsilon, X$)

1. Wyznaczamy $q = \min\{\| -D^{-1}(M + N) \|_1; \| -D^{-1}(M + N) \|_\infty\}$.

Jeżeli $q < 1$, to przechodzimy do punktu 2 algorytmu.

2. *Układ równań (2.1) przekształcamy do postaci (2.33).*

3. *W charakterze $X^{(0)}$ wybieramy dowolny wektor.*

4. *Kolejne przybliżenie $X^{(k)}$ obliczamy ze wzorów (2.34).*

5. *Jeśli*

$$\frac{q}{1-q} \cdot \|X^{(k)} - X^{(k-1)}\| < \varepsilon,$$

to kończymy obliczenia przyjmując $\hat{X} = X^{(k)}$.

W przeciwnym przypadku powracamy do punktu 4 algorytmu.

Przykład 2.8. Rozwiązywanie układu równań liniowych

$$\begin{array}{rrrrr} 10.0 x_1 & -3.2 x_2 & +3.6 x_3 & +1.2 x_4 & = 38.36, \\ 0.9 x_1 & -5.0 x_2 & -0.4 x_3 & +0.7 x_4 & = 17.48, \\ 3.2 x_1 & -0.8 x_2 & -8.0 x_3 & -1.6 x_4 & = 38.48, \\ 1.5 x_1 & -3.0 x_2 & -4.5 x_3 & +10.0 x_4 & = 55.20, \end{array}$$

metodą Jacobiego (iteracji prostej) z dokładnością 10^{-5} .

Rozwiązanie. Wyjściowy układ równań $AX = A_0$ przekształcamy według wzorów (2.33) do postaci $X = HX + g$:

$$H = \begin{bmatrix} 0.00000 & 0.32000 & -0.36000 & -0.12000 \\ 0.18000 & 0.00000 & -0.08000 & 0.14000 \\ 0.40000 & -0.10000 & 0.00000 & -0.20000 \\ -0.15000 & 0.30000 & 0.45000 & 0.00000 \end{bmatrix}, \quad g = \begin{bmatrix} 3.83600 \\ -3.49600 \\ -4.81000 \\ 5.52000 \end{bmatrix}.$$

Następnie obliczamy

$$q = 0.89000, \quad r = \frac{q}{1-q} = 8.09091.$$

Przymajemy $X^{(0)} = [0.0 \ 0.0 \ 0.0 \ 0.0]^T$ i wykonujemy iteracje według wzorów (2.34). Obliczenia, aż do uzyskania dokładności $\varepsilon = 10^{-5}$, są przedstawione w tabeli 2.2.

Tabela 2.2. Iteracje metody Jacobiego dla rozwiązyania układu równań z przykładu 2.8

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$	$r \ X^{(k)} - X^{(k-1)} \ $
0	0.0000000000	0.0000000000	0.0000000000	0.0000000000	
1	3.8360000000	-3.4960000000	-4.8100000000	5.5200000000	1.4E+02
2	3.7864800000	-1.6479200000	-4.0300000000	1.7313000000	5.2E+01
3	4.5517096000	-2.2496516000	-3.4768760000	2.6441520000	2.3E+01
4	4.0504886080	-2.0283609120	-3.2931814000	2.5977538800	7.7E+00
5	4.0607393466	-2.1397719954	-3.5065192416	2.8219868052	4.5E+00
6	4.0749814718	-2.0894672256	-3.5361244229	2.6910238407	1.8E+00
7	4.1174524192	-2.1028700435	-3.5092654569	2.6906566212	6.7E-01
8	4.1035383560	-2.0974254010	-3.4908633522	2.6923516685	3.2E-01
9	4.0984524783	-2.1011647942	-3.4973124512	2.7043531178	2.2E-01
10	4.0981373742	-2.0998841213	-3.5013731528	2.7010920870	7.2E-02
11	4.1004003658	-2.1000725282	-3.5009750556	2.6996962387	3.4E-02
12	4.1003642623	-2.0998924563	3.4997718486	2.6994794117	1.3E-02
13	4.1000147501	-2.1000255673	-3.4997609318	2.7000802919	8.9E-03
14	4.0998961189	-2.1000052296	-3.5000076016	2.7000976980	3.3E-03
15	4.0999893393	-2.1000044128	-3.5000605691	2.7000105926	1.9E-03
16	4.1000191217	-2.0999955904	-3.5000059415	2.6999730192	1.1E-03
17	4.1000067877	-2.0999998601	-3.4999873961	2.6999957810	4.7E-04
18	4.0999960136	-2.1000003772	-3.4999964551	2.7000046956	2.4E-04
19	4.0999980397	-2.1000003438	-3.5000024960	2.7000020800	8.7E-05
20	4.1000005389	-2.0999998620	-3.5000011657	2.6999990677	5.9E-05
21	4.1000005757	-2.0999999403	-3.4999996118	2.6999994360	1.6E-05
22	4.0999999470	-2.1000000064	-3.4999996629	2.7000001062	1.1E-05
23	4.0999998639	-2.1000000216	-3.5000000418	2.7000001577	4.3E-06

2.8.2. Metoda Gaussa-Seidela

W metodzie Gaussa-Seidela wybieramy dowolny wektor $X^{(0)}$ i tworzymy ciąg kolejnych przybliżeń $X^{(k)}$ ($k = 1, 2, \dots$) według wzorów

$$x_i^{(k)} = \sum_{j=1}^{i-1} h_{ij} x_j^{(k)} + \sum_{j=i+1}^n h_{ij} x_j^{(k-1)} + g_i \quad \text{dla } i = 1, 2, \dots, n. \quad (2.36)$$

Twierdzenie 2.5. Jeżeli

$$\| -(M + D)^{-1} N \|_1 \leq p < 1 \text{ lub } \| -(M + D)^{-1} N \|_\infty \leq p < 1, \quad (2.37)$$

to:

- a) ciąg (2.36) jest zbieżny do rozwiązania układu równań (2.1) niezależnie od wyboru wektora $X^{(0)}$,
- b) zachodzą oszacowania

$$\| \bar{X} - X^{(k)} \| \leq \frac{p}{1-p} \cdot \| X^{(k)} - X^{(k-1)} \| \leq \frac{p^k}{1-p} \cdot \| X^{(1)} - X^{(0)} \| \quad (2.38)$$

dla $k = 1, 2, \dots$

Algorytm 2.9. Rozwiązywanie układu równań liniowych $AX = A_0$ metodą Gaussa-Seidela.

Metoda Gaussa-Seidela ($n, A, A_0, \varepsilon, X$)

1. Wyznaczamy $p = \min\{\|-(M+D)^{-1}N\|_1; \|-(M+D)^{-1}N\|_\infty\}$.
Jeżeli $p < 1$, to przechodzimy do punktu 2 algorytmu.
2. Układ równań (2.1) przekształcamy do postaci (2.33).
3. W charakterze $X^{(0)}$ wybieramy dowolny wektor.
4. Kolejne przybliżenie $X^{(k)}$ obliczamy ze wzorów (2.36).
5. Jeśli

$$\frac{p}{1-p} \cdot \|X^{(k)} - X^{(k-1)}\| < \varepsilon,$$

to kończymy obliczenia przyjmując $\hat{X} = X^{(k)}$.

W przeciwnym przypadku powracamy do punktu 4 algorytmu.

Przykład 2.9. Rozwiązywanie układu równań liniowych

$$\begin{array}{rrrrr} 10.0 x_1 & -3.2 x_2 & +3.6 x_3 & +1.2 x_4 & = 38.36, \\ 0.9 x_1 & -5.0 x_2 & -0.4 x_3 & +0.7 x_4 & = 17.48, \\ 3.2 x_1 & -0.8 x_2 & -8.0 x_3 & -1.6 x_4 & = 38.48, \\ 1.5 x_1 & -3.0 x_2 & -4.5 x_3 & +10.0 x_4 & = 55.20 \end{array}$$

metodą Gaussa-Seidela z dokładnością 10^{-5} .

Rozwiązanie. Wyjściowy układ równań $AX = A_0$ przekształcamy według wzorów (2.33) do postaci $X = HX + g$:

$$H = \begin{bmatrix} 0.00000 & 0.32000 & -0.36000 & -0.12000 \\ 0.18000 & 0.00000 & -0.08000 & 0.14000 \\ 0.40000 & -0.10000 & 0.00000 & -0.20000 \\ -0.15000 & 0.30000 & 0.45000 & 0.00000 \end{bmatrix}, \quad g = \begin{bmatrix} 3.83600 \\ -3.49600 \\ -4.81000 \\ 5.52000 \end{bmatrix}.$$

Następnie obliczamy:

$$p = 0.68204, \quad r = \frac{p}{1-p} = 2.14505.$$

Przymajemy $X^{(0)} = [0.0 \ 0.0 \ 0.0 \ 0.0]^T$ i wykonujemy iteracje według wzorów (2.38). Obliczenia, aż do uzyskania dokładności $\varepsilon=10^{-5}$, są przedstawione w tabeli 2.3.

Tabela 2.3. Iteracje metody Gaussa-Seidela dla rozwiązywania układu równań z przykładów 2.8 i 2.9

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$	$r \ X^{(k)} - X^{(k-1)} \ _1$
0	0.00000000000	0.00000000000	0.00000000000	0.00000000000	
1	3.83600000000	-2.80552000000	-2.99504800000	2.75517240000	2.7E+01
2	3.6858301920	-2.2072225894	-3.6659801443	2.6552676294	3.3E+00
3	4.1308095078	-2.0874384089	-3.4799858819	2.7081534043	1.7E+00
4	4.0958362181	-2.1012091336	-3.5031752803	2.6988329511	1.7E-01
5	4.1008962240	-2.0997480441	-3.4994332962	2.7001961699	2.5E-02
6	4.0998530721	-2.1000443195	-3.5000935732	2.6999666354	4.8E-03
7	4.1000235079	-2.0999929538	-3.4999846285	2.7000055049	7.9E-04
8	4.0999960605	-2.1000011681	-3.5000025600	2.6999990885	1.3E-04
9	4.1000006572	-2.0999998045	-3.4999995744	2.7000001516	2.1E-05
10	4.0999998912	-2.1000000324	-3.5000000706	2.6999999748	3.6E-06

2.9. Metoda Czebyszewa

Metodę Czebyszewa można zastosować do układów równań liniowych, których macierz główna A układu jest symetryczna i dodatnio określona. Ponadto metoda wymaga znajomości przedziału $[\alpha; \beta]$, w którym leżą wszystkie wartości własne $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ macierzy A . Wiadomo, że wartości własne macierzy symetrycznej i dodatnio określonej są dodatnie. Znalezienie β jest łatwe – wystarczy bowiem przyjąć

$$\beta = \|A\|_E = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2}.$$

Znacznie trudniej jest wyznaczyć dobre ograniczenie wartości własnych od dołu. Metodę Czebyszewa można jednak stosować, nie znając minimalnej wartości własne; wystarczy wziąć dowolne wartości α, β , spełniające warunki: $0 < \alpha < \beta$ i $\beta \geq \lambda_n$.

Metoda Czebyszewa została sformułowana w postaci algorytmu 2.10. Warunki zbieżności tej metody zostały sformułowane w twierdzeniu 2.6.

Twierdzenie 2.6. Jeżeli A jest macierzą symetryczną i dodatnio określona, której wartości własne leżą w przedziale $[\alpha; \beta]$, $0 < \alpha < \beta$, to dla dowolnego $s \geq 1$ i dla dowolnego początkowego wektora Y_0 , ciąg $X^{(1)}, X^{(2)}, \dots$ generowany przez algorytm 2.10 przy braku błędów zaokrągleń dąży do rozwiązywania układu równań (2.1).

Parametr s oznacza liczbę iteracji, po której należy dokonać odnowy algorytmu ze względu na narastanie błędów zaokrągleń. Doświadczenie podpowiada, że błędy zaokrągleń w każdej iteracji nie są na ogół duże, więc zaleca się stosować duże wartości s .

Algorytm 2.10. Rozwiązywanie układu równań liniowych $AX = A_0$ metodą Czebyszewa.

Metoda Czebyszewa ($\alpha, \beta, Y_0, s, \varepsilon$)

```

1      i := 0
2       $\omega_0 = \frac{\beta - \alpha}{\beta + \alpha}$ 
3       $X^{(-1)} := 0$ 
4      repeat
5       $X^{(i)} := Y_0$ 
6       $\omega_{i-1} := 0$ 
7       $\omega_i := \omega_0$ 
8      for k := 1 to s
9      do begin
10      $X^{(i+1)} := X^{(i)} + \omega_i \omega_{i-1} (X^{(i)} - X^{(i-1)}) - \frac{2}{\beta + \alpha} \cdot (1 + \omega_i \omega_{i-1}) (AX^{(i)} - A_0)$ 
11      $\omega_{i+1} := \frac{1}{2 \frac{\beta + \alpha}{\beta - \alpha} - \omega_i}$ 
12     i := i + 1
13   end
14    $Y_0 := X^{(i)}$ 
15 until  $\|X^{(i)} - X^{(i-1)}\| \leq \varepsilon$ 
```

Przykład 2.10. Rozwiązywanie układu równań liniowych

$$3.2x_1 + 1.8x_2 + 1.4x_3 + 1.2x_4 = 32.78,$$

$$1.8x_1 + 3.8x_2 + 3.2x_3 - 1.8x_4 = 32.40,$$

$$1.4x_1 + 3.2x_2 + 4.2x_3 - 1.5x_4 = 33.50,$$

$$1.2x_1 - 1.8x_2 - 1.5x_3 + 4.6x_4 = 5.89$$

metodą Czebyszewa.

Rozwiązanie. Mamy

$$A = \begin{bmatrix} 3.2 & 1.8 & 1.4 & 1.2 \\ 1.8 & 3.8 & 3.2 & -1.8 \\ 1.4 & 3.2 & 4.2 & -1.5 \\ 1.2 & -1.8 & -1.5 & 4.6 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 32.78 \\ 32.40 \\ 33.50 \\ 5.89 \end{bmatrix}, \quad \beta = \|A\|_E = 10.402884.$$

Przyjmujemy:

$$\alpha = 0.1 \cdot \beta = 1.040288, \quad Y_0 = A_0.$$

Wybrane iteracje metody zostały podane w tabeli 2.4.

Tabela 2.4. Iteracje metody Czebyszewa dla rozwiązyania układu równań z przykładów 2.10

i	ω_i	$x_1^{(i)}$	$x_2^{(i)}$	$x_3^{(i)}$	$x_4^{(i)}$	$\ X^{(i)} - X^{(i-1)} \ _1$
-1	0.000000	0.000000000	0.000000000	0.000000000	0.000000000	
0	0.818182	32.780000000	32.400000000	33.500000000	5.890000000	1.0E+02
1	0.614907	0.550476624	-10.651347352	-9.833611729	14.284525107	1.3E+02
2	0.546586	-3.365308186	1.595394691	2.396600364	-6.292348505	4.9E+01
3	0.526910	5.797265711	8.840606756	9.683478419	-2.620522462	2.7E+01
4	0.521503	5.713006918	0.271878700	1.087569071	7.762485578	2.8E+01
5	0.520037	6.396454152	4.952473023	5.790107738	2.917232423	1.5E+01
6	0.519640	4.757752965	3.731181477	4.566944243	2.135219733	4.9E+00
7	0.519533	4.732979913	3.153995470	3.996867195	2.795322977	1.8E+00
8	0.519505	5.393789543	3.916439258	4.772356036	2.756992882	2.2E+00
9	0.519497	5.234911250	3.567296659	4.428493346	2.956036432	1.1E+00
10	0.519495	5.160964113	3.599827158	4.466674240	2.804280117	3.0E-01
11	0.519494	5.178015937	3.649725610	4.521720548	2.758187804	1.7E-01
12	0.519494	5.182752901	3.597906137	4.473923688	2.821388206	1.7E-01
13	0.519494	5.199523149	3.619642122	4.499298673	2.811683928	7.4E-02
14	0.519494	5.193646745	3.614298482	4.496888147	2.804876902	2.0E-02
15	0.519494	5.190528877	3.607493749	4.492580975	2.804429156	1.5E-02
16	0.519494	5.194545917	3.610159501	4.497429549	2.802832711	1.3E-02
17	0.519494	5.195489153	3.607446674	4.496556400	2.804165409	5.9E-03
18	0.519494	5.195933081	3.606345997	4.497030330	2.803302322	2.9E-03
19	0.519494	5.196449300	3.605749947	4.497782471	2.802327640	2.8E-03
20	0.519494	5.196882404	3.604662983	4.497847012	2.802192729	1.7E-03
30	0.519494	5.199366204	3.600997805	4.499569919	2.800460148	4.3E-04
40	0.519494	5.199867197	3.600209006	4.499909875	2.800096392	8.9E-05
50	0.519494	5.199972179	3.600043785	4.499981119	2.800020193	1.9E-05
60	0.519494	5.199994172	3.600009173	4.499996045	2.800004230	3.9E-06
70	0.519494	5.199998779	3.600001922	4.499999171	2.800000886	8.2E-07
80	0.519494	5.199999744	3.600000403	4.499999826	2.800000186	1.7E-07
90	0.519494	5.199999946	3.600000084	4.499999964	2.800000039	3.6E-08
100	0.519494	5.199999989	3.600000018	4.499999992	2.800000008	7.6E-09
110	0.519494	5.199999998	3.600000004	4.499999998	2.800000002	1.6E-09
120	0.519494	5.200000000	3.600000001	4.500000000	2.800000000	3.3E-10
130	0.519494	5.200000000	3.600000000	4.500000000	2.800000000	6.9E-11

2.10. Metoda najszybszego spadku

Dla układu równań

$$AX = A_0, \quad (2.39)$$

którego macierz główna jest macierzą symetryczną i dodatnio określona, konstruujemy funkcjonał

$$F(X) = (AX, X) - 2(A_0, X) \quad (2.40)$$

(w powyższym wzorze (AX, X) oznacza iloczyn skalarny wektorów AX i X ; a (A_0, X) iloczyn skalarny wektorów A_0 i X).

Twierdzenie 2.7. Jeżeli A jest macierzą symetryczną i dodatnio określona, to funkcjonał (2.40) jest wypukły.

Twierdzenie 2.8. Funkcjonał (2.40) osiąga minimum dla $X = \bar{X}$ wtedy i tylko wtedy, gdy \bar{X} jest rozwiązaniem układu równań (2.39).

Na mocy twierdzenia 2.8, znalezienie rozwiązania układu (2.39) jest równoważne znalezieniu minimum funkcjonału (2.40).

Idea metody najszybszego spadku jest następująca:

- a) w charakterze $X^{(0)}$ wybieramy dowolny wektor z przestrzeni R^n ,
- b) kolejne przybliżenia obliczamy ze wzoru

$$X^{(k+1)} = X^{(k)} - \text{grad}(F(X^{(k)})) \quad \text{dla } k = 0, 1, \dots$$

przy czym współczynnik liczbowy μ_k jest dobierany w taki sposób, aby przy przejściu z $X^{(k)}$ do $X^{(k+1)}$ następowało największe zmniejszenie wartości funkcjonału (2.40).

Dla funkcjonału (2.40) mamy

$$F(X^{(k)}) = -(A_0 - AX^{(k)}) = -R^{(k)}, \quad (2.41)$$

$$\mu_k = \frac{(R^{(k)}, R^{(k)})}{(AR^{(k)}, R^{(k)})}. \quad (2.42)$$

Ostatecznie otrzymujemy

Algorytm 2.11. Rozwiązywanie układu równań liniowych $AX = A_0$ metodą najszybszego spadku.

1. W charakterze $X^{(0)}$ wybieramy dowolny wektor.

2. Kolejne przybliżenia obliczamy ze wzoru

$$R^{(k)} = A_0 - AX^{(k)}$$

$$X^{(k+1)} = X^{(k)} + \mu_k R^{(k)},$$

przy czym

$$\mu_k = \frac{(R^{(k)}, R^{(k)})}{(AR^{(k)}, R^{(k)})}.$$

3. Obliczenia kończymy, gdy

$$\|R^{(k)}\| < \varepsilon. \quad (2.43)$$

Wówczas przyjmujemy $\bar{X} = X^{(k)}$.

Do algorytmu 2.11 należy dołączyć kilka uwag. Po pierwsze kryterium zakończenia procesu iteracyjnego (2.43) jest heurystyczne i nie gwarantuje, że $\| \bar{X} - X^{(k)} \| < \varepsilon$. Po drugie, metodę najszybszego spadku można zastosować do dowolnego układu równań liniowych z nieosobliwą macierzą główną. Wynika to z następującego twierdzenia.

Twierdzenie 2.9. Jeżeli macierz A jest nieosobliwa, to AA^T oraz A^TA są macierzami symetrycznymi i dodatnio określonymi.

Zatem jeśli mamy układ równań liniowych z nieosobliwą macierzą główną, to możemy go przekształcić do postaci z macierzą główną symetryczną i dodatnio określona na jeden z dwóch sposobów:

I transformacja Gaussa

$$(A^T A)X = A^T(AX) = A^T A_0 ,$$

II transformacja Gaussa

$$\begin{aligned} X &= A^T Y, \\ (AA^T)Y &= A(A^T Y) = AX = A_0. \end{aligned}$$

Przykład 2.11. Rozwiązywanie układu równań liniowych

$$\begin{array}{lclll} 3.2 x_1 & +1.8 x_2 & +1.4 x_3 & +1.2 x_4 & = 32.78 \\ 1.8 x_1 & +3.8 x_2 & +3.2 x_3 & -1.8 x_4 & = 32.40 \\ 1.4 x_1 & +3.2 x_2 & +4.2 x_3 & -1.5 x_4 & = 33.50 \\ 1.2 x_1 & -1.8 x_2 & -1.5 x_3 & +4.6 x_4 & = 5.89 \end{array}$$

metodą najszybszego spadku.

Rozwiązanie. Macierz główna układu jest macierzą symetryczną i dodatnio określona. Wybrane iteracje wykonane według algorytmu 2.11 zostały zapisane w tabeli 2.5.

Tabela 2.5. Wybrane iteracje metody najszybszego spadku dla rozwiązywania układu równań

i	$x_1^{(i)}$ $r_1^{(i)}$	$x_2^{(i)}$ $r_2^{(i)}$	$x_3^{(i)}$ $r_3^{(i)}$	$x_4^{(i)}$ $r_4^{(i)}$	μ_i	$\ X^{(i)} - X^{(i-1)} \ _1$
0	0.000000000	0.000000000	0.000000000	0.000000000		
	32.780000000	32.400000000	33.500000000	5.890000000	0.129540	
1	4.246328600	4.197103314	4.339597563	0.762991930		1.35E+01
	4.645935610	-3.705710802	-2.957412515	11.348825113	0.157321	
2	4.977230696	3.614118095	3.874335194	2.548397660		3.56E+00
	1.865302738	1.896579153	2.517087796	0.511609292	0.131498	
3	5.222513848	3.863514026	4.205326628	2.615673141		8.93E-01
	0.087365392	-0.430711780	0.086373604	0.853202121	0.184325	
4	5.238617474	3.784123075	4.221247443	2.772939626		2.69E-01
	-0.032271424	0.074120371	0.486911874	-0.008570549	0.200712	

cd. tab. 2.5

5	5.232140198	3.798999955	4.318976712	2.771219410		1.21E-01
	-0.173079242	-0.286582727	0.035330790	0.180487463	0.194083	
6	5.198548452	3.743379108	4.325833820	2.806248966		1.31E-01
	-0.017103548	0.026352092	0.284090428	-0.030170262	0.201853	
7	5.195096056	3.748698348	4.383178227	2.800159018		7.22E-02
	-0.088604744	-0.182110714	0.021920784	0.097577615	0.195565	
8	5.177768089	3.713083907	4.387465160	2.819241763		7.63E-02
	0.002049744	0.005045255	0.170765146	-0.027085043	0.203946	
9	5.178186125	3.714112865	4.422291972	2.813717885		4.18E-02
	-0.043268981	-0.121006028	0.012328803	0.051915491	0.198014	
10	5.169618255	3.690151958	4.424733250	2.823997888		4.53E-02
	0.011524047	-0.003840499	0.106165361	-0.024558790	0.207079	
11	5.172004645	3.689356670	4.446717891	2.818912272		3.03E-02
	-0.019356646	-0.084618871	0.005405416	0.027516817	0.201253	
12	5.168109067	3.672326890	4.447805746	2.824450107		2.76E-02
	0.015594409	-0.006406700	0.069092283	-0.022304351	0.210669	
13	5.171394333	3.670977195	4.462361379	2.819751262		2.39E-02
	-0.007228259	-0.062227278	0.000630014	0.014772014	0.204418	
14	5.169916747	3.658256824	4.462490165	2.822770926		1.73E-02
	0.016592785	-0.006206934	0.047392416	-0.020048828	0.213661	
15	5.173461982	3.656930642	4.472616086	2.818487269		1.93E-02
	-0.001400741	-0.047662399	-0.002281488	0.008203469	0.206682	
16	5.173172473	3.647079677	4.472144543	2.820182780		1.23E-02
	0.015882971	-0.005146760	0.034170659	-0.017687521	0.215537	
17	5.176595844	3.645970359	4.479509590	2.816370462		1.57E-02
	0.001188675	-0.037523739	-0.003723914	0.004791895	0.207954	
18	5.176843034	3.638167136	4.478735186	2.817366957		9.82E-03
	0.014331839	-0.004044650	0.025647572	-0.015296018	0.216507	
19	5.179945975	3.637291442	4.484288061	2.814055264		1.28E-02
	0.002178684	-0.030032551	-0.004183936	0.002967300	0.208572	
20	5.180400387	3.631027491	4.483415409	2.814674160		8.21E-03
	0.012478712	-0.003140984	0.019818010	-0.013009004	0.216958	
30	5.192399430	3.611269103	4.494398922	2.805677765		3.16E-03
	0.005065627	-0.000998141	0.006620841	-0.005114268	0.217301	

40	5.197186422	3.604140033	4.497960488	2.802101331		1.17E-03
	0.001885108	-0.000358852	0.002408850	-0.001897039	0.217306	
50	5.198963934	3.601523164	4.499250426	2.800773770		4.30E-04
	0.000694596	-0.000131679	0.000885236	-0.000698731	0.217306	
60	5.199618715	3.600560484	4.499724210	2.800284756		1.58E-04
	0.000255638	-0.000048439	0.000325700	-0.000257149	0.217306	
70	5.199859693	3.600206248	4.499898516	2.800104786		5.83E-05
	0.000094072	-0.000017824	0.000119850	-0.000094627	0.217306	
80	5.199948369	3.600075895	4.499962656	2.800038559		2.14E-05
	0.000034617	-0.000006559	0.000044102	-0.000034821	0.217306	
90	5.199981001	3.600027928	4.499986258	2.800014189		7.89E-06
	0.000012738	-0.000002414	0.000016229	-0.000012814	0.217306	
100	5.199993009	3.600010277	4.499994943	2.800005221		2.90E-06
	0.000004688	-0.000000888	0.000005972	-0.000004715	0.217306	
110	5.199997427	3.600003782	4.499998139	2.800001921		1.07E-06
	0.000001725	-0.000000327	0.000002198	-0.000001735	0.217306	
120	5.199999053	3.600001392	4.499999315	2.800000707		3.93E-07
	0.000000635	-0.000000120	0.000000809	-0.000000638	0.217306	
130	5.199999652	3.600000512	4.499999748	2.800000260		1.45E-07
	0.000000234	-0.000000044	0.000000298	-0.000000235	0.217306	

2.11. Metody rozwiązywania wielkich układów równań z macierzami rozrzedzonymi

Metody skończone rozwiązywania układu równań (2.1) wymagają $O(n^2)$ komórek pamięci komputera oraz wykonania $O(n^3)$ działań arytmetycznych zmiennoprzecinkowych, a ich algorytmy są numerycznie poprawne. Zatem jeśli n nie jest duże i układ równań nie jest źle uwarunkowany, to rozwiązanie takiego układu nie nastręcza trudności.

W praktyce często pojawiają się układy równań liniowych, których wymiar n jest rzędu 10^3 , a nawet 10^4 . Zadania takie pojawiają się podczas numerycznego rozwiązywania równań różniczkowych cząstkowych. Układy takie mają szczególne właściwości: macierz główna układu ma dużo zer – często powyżej 70%, elementy niezerowe tworzą specjalne struktury, np. wstęgi, bloki. Takie układy należy rozwiązywać specjalnymi metodami, wykorzystującymi specyficzne własności macierzy głównej układu.

Istnieją dwie podstawowe strategie rozwiązywania wielkich układów równań liniowych z rozrzedzoną macierzą główną układu. Pierwsza strategia polega na zastosowaniu metod iteracyjnych. W metodach iteracyjnych operacją dominującą jest obliczanie iloczynu skalarnego wektorów rozmiaru n . Jeśli obliczamy

$$y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n,$$

to składniki $a_{ij}x_j$, których element macierzy a_{ij} jest równy zeru, możemy pominąć przy obliczaniu sumy, ponieważ są równe zeru bez względu na wartość x_j . W metodach iteracyjnych wystarczy zatem pamiętać tylko elementy niezerowe macierzy głównej układu (wraz z informacją o położeniu tych elementów), a podczas wykonywania obliczeń uwzględniać wyżej wzmiankowaną własność obliczania iloczynu skalarnego.

Druga strategia rozpoczyna się przekształceniem wyjściowego układu równań – przez przestawienie wierszy i kolumn – do postaci, w której macierz główna układu ma specjalną strukturę, np. wstęgową lub blokową. Z kolei stosujemy odpowiednią metodę skończoną, która zachowuje tę szczególną strukturę macierzy głównej układu, nie wykonując działań na elementach zerowych.

Kwadratową macierz A nazywamy *macierzą pasmową o szerokości pasma p* , jeśli $a_{ij} = 0$ dla wszystkich $|i - j| > p$.

Przykład 2.12. Rozwiązać układ równań

$$\begin{array}{lllllll}
 0.80 x_1 & +0.96 x_2 & & & & & = 1.28 \\
 3.40 x_1 & +4.28 x_2 & +0.42 x_3 & & & & = 5.46 \\
 0.90 x_1 & +3.18 x_2 & +5.81 x_3 & +0.42 x_4 & & & = 1.79 \\
 & 2.20 x_2 & +6.42 x_3 & +1.34 x_4 & +0.64 x_5 & & = 1.20 \\
 & & +0.80 x_3 & +2.44 x_4 & +1.86 x_5 & +0.23 x_6 & = 3.24 \\
 & & & 1.60 x_4 & +4.08 x_5 & +8.84 x_6 & +4.32 x_7 & = 28.96 \\
 & & & & 0.50 x_5 & +4.35 x_6 & +6.26 x_7 & +0.70 x_8 & = 7.04 \\
 & & & & & 4.00 x_6 & +8.10 x_7 & +4.06 x_8 & = 12.08
 \end{array}$$

Rozwiązanie. Powyższy układ rozwiążemy metodą *LU*. Rozkładu macierzy głównej układu A na iloczyn *LU* dokonujemy według algorytmu 2.3, otrzymując

$$L = \begin{bmatrix} 0.80 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 3.40 & 0.20 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.90 & 2.10 & 1.40 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 2.20 & 1.80 & 0.80 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.80 & 2.20 & 0.10 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.60 & 2.80 & 2.40 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.50 & 3.20 & 0.50 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 4.00 & 0.90 & 2.80 \end{bmatrix},$$

$$U = \begin{bmatrix} 1.00 & 1.20 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 2.10 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 & 0.30 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 & 0.80 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 2.30 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 1.80 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 1.40 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 \end{bmatrix}$$

Zwróćmy uwagę, że macierz L ma dwie podprzekątne (czyli tyle, ile ma ich macierz A), a macierz U ma jedną nadprzekątną (też tyle, ile ma ich macierz A).

Z kolei rozwiązujemy układ równań

$$\begin{array}{lll} 0.8 y_1 & & = 1.28 \\ 3.4 y_1 + 0.2 y_2 & & = 5.46 \\ 0.9 y_1 + 2.1 y_2 + 1.4 y_3 & & = 1.79 \\ 2.2 y_2 + 1.8 y_3 + 0.8 y_4 & & = 1.20 \\ 0.8 y_3 + 2.2 y_4 + 0.1 y_5 & & = 3.24 \\ 1.6 y_4 + 2.8 y_5 + 2.4 y_6 & & = 28.96 \\ 0.5 y_5 + 3.2 y_6 + 0.5 y_7 & & = 7.04 \\ 4.0 y_6 + 0.9 y_7 + 2.8 y_8 & & = 12.08 \end{array}$$

według algorytmu 2.4 i otrzymujemy wartości składowych wektora Y :

$$y_1 = 1.60, y_2 = 0.10, y_3 = 0.10, y_4 = 1.00, y_5 = 9.60, y_6 = 0.20, y_7 = 3.20, y_8 = 3.00.$$

Ostatecznie rozwiązujemy układ równań

$$\begin{array}{lll} x_1 + 1.2 x_2 & & = 1.60 \\ x_2 + 2.1 x_3 & & = 0.10 \\ x_3 + 0.3 x_4 & & = 0.10 \\ x_4 + 0.8 x_5 & & = 1.00 \\ x_5 + 2.3 x_6 & & = 9.60 \\ x_6 + 1.8 x_7 & & = 0.20 \\ x_7 + 1.4 x_8 & = 3.20 \\ x_8 & = 3.00 \end{array}$$

według algorytmu 2.5 i otrzymujemy wartości składowych wektora X :

$$x_8 = 3.0, x_7 = -1.0, x_6 = 2.0, x_5 = 5.0, x_4 = -3.0, x_3 = 1.0, x_2 = -2.0, x_1 = 4.0.$$

Na żadnym etapie metody nie korzystaliśmy z elementów leżących na pozycjach (i,j) gdy $j - i > 1$ oraz $i - j > 3$.

Układem równań liniowych o szczególnej strukturze jest układ z *trójdiamondową blokową macierzą główną*:

$$\begin{bmatrix} B_1 & C_1 & & 0 \\ A_2 & B_2 & C_2 & \\ \dots & \dots & \dots & \\ \dots & \dots & \dots & \dots & \dots \\ & \dots & \dots & \dots & \\ 0 & & & A_{n-1} & B_{n-1} & C_{n-1} \\ & & & A_n & B_n & \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_{n-1} \\ X_n \end{bmatrix} = \begin{bmatrix} D_1 \\ D_2 \\ \dots \\ D_{n-1} \\ D_n \end{bmatrix}, \quad (2.41)$$

gdzie A_i, B_i, C_i są podmacierzami kwadratowymi, wszystkie tego samego stopnia m . Rozwiązań takiego układu możemy otrzymać wykonując operacje na blokach w dwóch etapach.

Algorytm 2.13. Rozwiązywanie układu równań liniowych (2.41) metodą blokową.

Etap I:

```

 $Q_0 = 0$ 
 $U_0 = 0$ 
for  $k := 1$  to  $n - 1$ 
    do begin
         $P_k = (A_k Q_{k-1} + B_k)^{-1}$ 
         $Q_k = -P_k C_k$ 
         $U_k = P_k (D_k - A_k U_{k-1})$ 
    end
     $P_n = (A_n Q_{n-1} + B_n)^{-1}$ 
     $U_n = P_n (D_n - A_n U_{n-1})$ 

```

Etap II:

```

 $X_n := U_n$ 
for  $k := n - 1$  downto 1
    do  $X_k := Q_k X_{k+1} + U_k$ 

```

Przykład 2.13. Rozwiązać układ równań.

$5x_1$	$+4x_2$	$+8x_3$	$+9x_4$					$= 62.5$
$6x_1$	$+5x_2$	$+6x_3$	$+7x_4$					$= 55.0$
$9x_1$	$+4x_2$	$+4x_3$	$+3x_4$	$+6x_5$	$+4x_6$			$= 77.5$
$8x_1$	$+4x_2$	$+5x_3$	$+4x_4$	$+7x_5$	$+5x_6$			$= 89.0$
		$8x_3$	$+9x_4$	$+3x_5$	$+2x_6$	$+4x_7$	$+2x_8$	$= 93.5$
		$5x_3$	$+6x_4$	$+4x_5$	$+3x_6$	$+5x_7$	$+3x_8$	$= 94.0$
				$9x_5$	$+5x_6$	$+2x_7$	$+x_8$	$= 65.5$
				$3x_5$	$+2x_6$	$+3x_7$	$+2x_8$	$= 42.0$

Rozwiązanie. Strukturę blokową macierzy głównej układu i wektora wyrazów wolnych zaznaczono w treści zadania. Układ równań rozwiążemy według algorytmu 2.13.

$$P_1 = B_1^{-1} = \begin{bmatrix} 5.00000 & 4.00000 \\ 6.00000 & 5.00000 \end{bmatrix}^{-1} = \begin{bmatrix} 5.00000 & -4.00000 \\ -6.00000 & 5.00000 \end{bmatrix},$$

$$Q_1 = -P_1 C_1 = -\begin{bmatrix} 5.00000 & -4.00000 \\ -6.00000 & 5.00000 \end{bmatrix} \begin{bmatrix} 8.00000 & 9.00000 \\ 6.00000 & 7.00000 \end{bmatrix} = \begin{bmatrix} -16.00000 & -17.00000 \\ 18.00000 & 19.00000 \end{bmatrix},$$

$$U_1 = P_1 D_1 = \begin{bmatrix} 5.00000 & -4.00000 \\ -6.00000 & 5.00000 \end{bmatrix} \begin{bmatrix} 62.5 \\ 55.0 \end{bmatrix} = \begin{bmatrix} 92.50000 \\ -100.00000 \end{bmatrix},$$

$$P_2 = (A_2 Q_1 + B_2)^{-1} =$$

$$= \left(\begin{bmatrix} 9.00000 & 4.00000 \\ 8.00000 & 4.00000 \end{bmatrix} \begin{bmatrix} -16.00000 & -17.00000 \\ 18.00000 & 19.00000 \end{bmatrix} + \begin{bmatrix} 4.00000 & 3.00000 \\ 5.00000 & 4.00000 \end{bmatrix} \right)^{-1} =$$

$$= \begin{bmatrix} -1.64706 & 2.17647 \\ 1.50000 & -2.00000 \end{bmatrix},$$

$$Q_2 = -P_2 C_2 =$$

$$= -\begin{bmatrix} -1.64706 & 2.17647 \\ 1.50000 & -2.00000 \end{bmatrix} \begin{bmatrix} 6.00000 & 4.00000 \\ 7.00000 & 5.00000 \end{bmatrix} = \begin{bmatrix} -5.35294 & -4.29412 \\ 5.00000 & 4.0000 \end{bmatrix},$$

$$U_2 = P_2 (D_2 - A_2 U_1) =$$

$$= \begin{bmatrix} -1.64706 & 2.17647 \\ 1.50000 & -2.00000 \end{bmatrix} \left(\begin{bmatrix} 77.50000 \\ 89.00000 \end{bmatrix} - \begin{bmatrix} 9.00000 & 4.00000 \\ 8.00000 & 4.00000 \end{bmatrix} \begin{bmatrix} 92.50000 \\ -100.00000 \end{bmatrix} \right) =$$

$$= \begin{bmatrix} 38.41176 \\ -30.50000 \end{bmatrix},$$

$$P_3 = (A_3 Q_2 + B_3)^{-1} =$$

$$= \left(\begin{bmatrix} 8.00000 & 9.00000 \\ 5.00000 & 6.00000 \end{bmatrix} \begin{bmatrix} -5.35294 & -4.29412 \\ 5.00000 & 4.0000 \end{bmatrix} + \begin{bmatrix} 3.00000 & 2.00000 \\ 4.00000 & 3.00000 \end{bmatrix} \right)^{-1} =$$

$$= \begin{bmatrix} 2.47368 & -1.63158 \\ -3.236842 & 2.31579 \end{bmatrix},$$

$$Q_3 = -P_3 C_3 =$$

$$= -\begin{bmatrix} 2.47368 & -1.63158 \\ -3.236842 & 2.31579 \end{bmatrix} \begin{bmatrix} 4.00000 & 2.00000 \\ 5.00000 & 3.00000 \end{bmatrix} = \begin{bmatrix} -1.73684 & -0.05263 \\ 1.36842 & -0.47368 \end{bmatrix}$$

$$\begin{aligned}
 U_3 &= P_3(D_3 - A_3 U_2) = \\
 &= \begin{bmatrix} 2.47368 & -1.63158 \\ -3.236842 & 2.31579 \end{bmatrix} \left[\begin{bmatrix} 93.50000 \\ 94.00000 \end{bmatrix} - \begin{bmatrix} 8.00000 & 9.00000 \\ 5.00000 & 6.00000 \end{bmatrix} \begin{bmatrix} 38.41176 \\ -30.50000 \end{bmatrix} \right] = \\
 &= \begin{bmatrix} 11.57895 \\ 0.21053 \end{bmatrix},
 \end{aligned}$$

$$\begin{aligned}
 P_4 &= (A_4 Q_3 + B_4)^{-1} = \\
 &= \left(\begin{bmatrix} 9.00000 & 5.00000 \\ 3.00000 & 2.00000 \end{bmatrix} \begin{bmatrix} -1.73684 & -0.05263 \\ 1.36842 & -0.47368 \end{bmatrix} + \begin{bmatrix} 2.00000 & 1.00000 \\ 3.00000 & 2.00000 \end{bmatrix} \right)^{-1} = \\
 &= \begin{bmatrix} -0.17526 & -0.36082 \\ 0.10309 & 1.32990 \end{bmatrix},
 \end{aligned}$$

$$\begin{aligned}
 U_4 &= P_4(D_4 - A_4 U_3) = \\
 &= \begin{bmatrix} -0.17526 & -0.36082 \\ 0.10309 & 1.32990 \end{bmatrix} \left[\begin{bmatrix} 65.50000 \\ 42.00000 \end{bmatrix} - \begin{bmatrix} 9.00000 & 5.00000 \\ 3.00000 & 2.00000 \end{bmatrix} \begin{bmatrix} 11.57895 \\ 0.21053 \end{bmatrix} \right] = \\
 &= \begin{bmatrix} 4.50000 \\ 5.00000 \end{bmatrix},
 \end{aligned}$$

$$X_4 = U_4 = \begin{bmatrix} 4.50000 \\ 5.00000 \end{bmatrix},$$

$$X_3 = Q_3 X_4 + U_3 = \begin{bmatrix} -1.73684 & -0.05263 \\ 1.36842 & -0.47368 \end{bmatrix} \begin{bmatrix} 4.50000 \\ 5.00000 \end{bmatrix} + \begin{bmatrix} 11.57895 \\ 0.21053 \end{bmatrix} = \begin{bmatrix} 3.50000 \\ 4.00000 \end{bmatrix},$$

$$X_2 = Q_2 X_3 + U_2 = \begin{bmatrix} -5.35294 & -4.29412 \\ 5.00000 & 4.0000 \end{bmatrix} \begin{bmatrix} 3.50000 \\ 4.00000 \end{bmatrix} + \begin{bmatrix} 38.41176 \\ -30.50000 \end{bmatrix} = \begin{bmatrix} 2.50000 \\ 3.00000 \end{bmatrix},$$

$$X_1 = Q_1 X_2 + U_1 = \begin{bmatrix} -16.00000 & -17.00000 \\ 18.00000 & 19.00000 \end{bmatrix} \begin{bmatrix} 2.50000 \\ 3.00000 \end{bmatrix} + \begin{bmatrix} 92.50000 \\ -100.00000 \end{bmatrix} = \begin{bmatrix} 1.50000 \\ 2.00000 \end{bmatrix}.$$

Ostatecznym rozwiązaniem układu jest

$$x_1 = 1.5, x_2 = 2.0, x_3 = 2.5, x_4 = 3.0, x_5 = 3.5, x_6 = 4.0, x_7 = 4.5, x_8 = 5.0$$

2.12. Uwarunkowanie zadania rozwiązywania układu równań liniowych

Rozróźnimy dwa przypadki. W pierwszym badamy wpływ zaburzeń ΔA_0 wektora wyrazów wolnych układu równań (2.2) na zaburzenie ΔX rozwiązania tego układu, zakładając, że elementy macierzy głównej A układu nie są zaburzone. Mamy więc

$$A(X + \Delta X) = A_0 + \Delta A_0.$$

Dla tego przypadku prawdziwa jest nierówność

$$\frac{\|\Delta X\|}{\|X\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|\Delta A_0\|}{\|A_0\|}. \quad (2.42)$$

Wskaźnikiem uwarunkowania zadania ze względu na zaburzenie kolumny wyrazów wolnych jest wielkość

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|. \quad (2.43)$$

W drugim przypadku badamy wrażliwość wektora rozwiązania na zaburzenia ΔA macierzy głównej układu

$$(A + \Delta A)(X + \Delta X) = A_0,$$

zakładając, że wektor wyrazów wolnych nie jest zaburzony. Dla tego przypadku mamy nierówność

$$\frac{\|\Delta X\|}{\|X\|} \leq \frac{\|A\| \cdot \|A^{-1}\| \cdot \frac{\|\Delta A\|}{\|A\|}}{1 - \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\Delta A\|}{\|A\|}} = \frac{\text{cond}(A) \cdot \frac{\|\Delta A\|}{\|A\|}}{1 - \text{cond}(A) \cdot \frac{\|\Delta A\|}{\|A\|}}. \quad (2.44)$$

Również dla tego przypadku wskaźnik określony równością (2.43) charakteryzuje uwarunkowanie zadania.

Klasycznym przykładem źle uwarunkowanych układów równań liniowych są układy, w których macierzą główną jest macierz Hilberta

$$H_n = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \dots & \frac{1}{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \dots & \frac{1}{2n-1} \end{bmatrix}.$$

Przykładowo $\text{cond}(H_6) \approx 1.5 \cdot 10^7$, a $\text{cond}(H_{10}) \approx 1.6 \cdot 10^{13}$.

Przykładem układów dobrze uwarunkowanych są układy, których macierz główna jest macierzą ortogonalną. Dowodzi się, że $\text{cond}(Q) = 1$, jeśli Q jest macierzą ortogonalną.

Przykład 2.14. Zbadajmy uwarunkowanie zadania rozwiązania następującego układu równań liniowych

$$\begin{aligned} 5x_1 + 7x_2 &= 0.7, \\ 7x_1 + 10x_2 &= 1.0. \end{aligned} \quad (2.45)$$

Rozwiązanie. Obliczamy $\text{cond}(A)$:

$$A = \begin{bmatrix} 5 & 7 \\ 7 & 10 \end{bmatrix}, \text{ stąd } \|A\|_1 = \max\{5+7; 7+10\} = 17.$$

$$A^{-1} = \begin{bmatrix} 10 & -7 \\ -7 & 5 \end{bmatrix}, \text{ stąd } \|A^{-1}\|_1 = \max\{10+7; 7+5\} = 17.$$

Zatem na mocy wzoru (2.38) mamy

$$\text{cond}(A) = 17 \cdot 17 = 289.$$

Oznacza to, że błąd względny rozwiązania układu (2.45) może być 289 razy większy niż błąd względny wektora wyrazów wolnych tego układu. Nie dziwi zatem fakt, że rozwiązaniem układu (2.45) jest $x_1 = 0, x_2 = 0.1$, a rozwiązaniem układu równań

$$\begin{aligned} 5\hat{x}_1 + 7\hat{x}_2 &= 0.69, \\ 7\hat{x}_1 + 10\hat{x}_2 &= 1.01 \end{aligned}$$

jest $\hat{x}_1 = -0.17, \hat{x}_2 = 0.22$.

2.13. Zadania

Zad. 2.1. Złożoność obliczeniowa metody eliminacji Gaussa.

Wykazać, że dla wyznaczenia rozwiązania układu równań liniowych (2.1) metodą eliminacji Gaussa należy wykonać:

- $\frac{1}{3}n^3 + n^2 - \frac{1}{3}n$ mnożeń i dzieleni,
- $\frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n$ dodawań i odejmowań.

Łącznie więc trzeba wykonać

$$\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{1}{6}n \quad \text{działan arytmetycznych zmennoprzecinkowych.}$$

Złożoność obliczeniowa metody eliminacji Gaussa jest rzędu $O\left(\frac{2}{3}n^3\right)$.

Zad. 2.2. Warunki stosowności metody eliminacji Gaussa.

Oznaczmy przez M_k ($k = 1, 2, \dots, n$) minor główny stopnia k macierzy głównej układu (2.1). Wykazać, że

$$M_k = a_{11}^{(0)} \cdot a_{22}^{(1)} \cdot \dots \cdot a_{kk}^{(k-1)} \quad \text{dla } k = 1, 2, \dots, n.$$

Z tych równości wynika, że nierówność $a_{kk}^{(k-1)} \neq 0$ jest równoważna nierówności $M_k \neq 0$ dla $k = 1, 2, \dots, n$.

Zad. 2.3. Obliczanie wyznacznika macierzy.

Zastosuj przekształcenia równoważnościowe z rozdziału 2.3 i opracuj podprogram obliczania wyznacznika macierzy, wiedząc, że

- zamiana miejscami dwóch wierszy (lub kolumn) powoduje zmianę znaku wyznacznika,
- wartość wyznacznika nie zmieni się, jeżeli do dowolnego wiersza (kolumny) dodamy liniową kombinację innych wierszy (kolumn),
- jeżeli dowolny wiersz (kolumna) macierzy jest pomnożony przez stałą c , to tę stałą można wynieść przed symbol wyznacznika.

Zad. 2.4. Rozwiązywanie systemu układów równań liniowych z tą samą macierzą główną.

Metodę eliminacji Jordana (rozdział 2.5) można uogólnić na rozwiązywanie systemu układów równań liniowych

$$AX_1 = B_1,$$

$$AX_2 = B_2,$$

.....

$$AX_m = B_m$$

z tą samą macierzą główną układu A . Wystarczy powyższy system zapisać w postaci równania macierzowego

$$A \cdot [X_1, X_2, \dots, X_m] = [B_1, B_2, \dots, B_m]$$

i zastosować algorytm z rozdziału 2.5 przekształcając w każdym kroku metody elementy m kolumn wyrazów wolnych (zamiast jednej).

Opracować podprogram rozwiązywania takiego systemu układu równań liniowych. Jakie warunki musi spełniać macierz A , aby było można zastosować tę metodę?

Zad. 2.5. Odwracanie macierzy.

Załóżmy, że wszystkie minory główne macierzy A są różne od zera i oznaczmy

$$E_1 = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 \\ 1 \\ \dots \\ 0 \end{bmatrix}, \quad \dots \quad E_n = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 1 \end{bmatrix}.$$

Odwrócenie macierzy A jest równoważne rozwiązyaniu następującego systemu równań liniowych

$$AX_1 = E_1,$$

$$AX_2 = E_2,$$

.....

$$AX_n = E_n.$$

Opracuj podprogram odwracania macierzy A .

Zad. 2.6. Rozwiązywanie układu m równań liniowych z n niewiadomymi.

Algorytm z rozdziału 2.4 można zastosować do układu równań liniowych, w którym liczba równań jest różna od liczby niewiadomych. W tym przypadku zostanie wykonanych $r = \text{rzqd}(A) \leq \min\{m, n\}$ kroków algorytmu. Z równań o numerach $r + 1, r + 2, \dots, m$ możemy odczytać, czy wyjściowy układ równań jest sprzeczny czy nie (w jaki sposób?). Z równań o numerach $1, 2, \dots, r$ możemy odczytać rozwiązanie układu.

Zad. 2.7. Rozwiązywanie układów równań liniowych z macierzą pasmową.

Wykazać, że metoda eliminacji Gaussa zachowuje szerokość pasma, tzn. po etapie postępowania prostego macierz główna układu jest macierzą trójkątną górną, w której elementy leżące nad pasmem są równe zeru.

Zad. 2.8. Rozwiązywanie układów równań liniowych z trójdiamondalną macierzą blokową.

Wykazać, że w trakcie przekształceń metody eliminacji Gaussa nie pojawiają się elementy niezerowe poza blokami A_i, B_i, C_i .

Zad. 2.9. Złożoność obliczeniowa metody LU.

Wykazać, że dla wyznaczenia rozwiązania układu równań liniowych (2.1) metodą LU należy wykonać:

$$-\frac{1}{2}n^3 + n^2 - \frac{1}{3}n \quad \text{mnożeń i dzieleni},$$

$$-\frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n \quad \text{dodawań i odejmowań}.$$

Łącznie więc trzeba wykonać

$$\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n \quad \text{działan arytmetycznych zmiennoprzecinkowych.}$$

Złożoność obliczeniowa metody LU jest rzędu $O\left(\frac{2}{3}n^3\right)$.

Zad. 2.10. Własności rozkładu trójkątno-trójkątnego.

Wykazać, że dla rozkładu trójkątno-trójkątnego zachodzą równości

$$M_k = l_{11} \cdot l_{22} \cdot \dots \cdot l_{nn} \text{ dla } k = 1, 2, \dots, n,$$

Gdzie M_k ($k = 1, 2, \dots, n$) oznacza minor główny stopnia k macierzy A .

Zad. 2.11. Obliczanie wyznacznika macierzy na podstawie rozkładu trójkątno-trójkątnego.

Korzystając z własności wykazanych w zad. 2.10, opracować:

- a) podprogram obliczania $\det(A)$,
- b) podprogram sprawdzający czy macierz A jest dodatnio (ujemnie) określona.

Zad. 2.12. Rozkład trójkątno-trójkątny macierzy pasmowej.

Załóżmy, że macierz główna układu równań (2.1) jest macierzą pasmową o szerokości pasma p . Wykazać, że rozkład trójkątno-trójkątny zachowuje szerokość pasma, tzn. po etapie rozkładu $A = LU$ mamy:

$$l_{ij} = 0 \quad \text{dla } i - j > p,$$

$$u_{ij} = 0 \quad \text{dla } j - i > p.$$

Zad. 2.13. Rozkład trójkątno-trójkątny trójprzekątniowej macierzy blokowej.

Załóżmy, że macierz główna układu równań (2.1) jest trójprzekątnią macierzą blokową. Wykazać, że rozkład trójkątno-trójkątny zachowuje strukturę tej macierzy, tj. nie pojawiają się elementy niezerowe w macierzach L , U na pozycjach poza A_i , B_i , C_i .

Zad. 2.14. Rozwiązywanie układu równań liniowych z trójprzekątnią macierzą główną.

Dany jest układ równań liniowych

$$b_1 x_1 + c_1 x_2 = d_1,$$

$$a_2 x_1 + b_2 x_2 + c_2 x_3 = d_2,$$

$\dots \dots \dots \dots \dots \dots \dots \dots$

$$a_{n-1} x_{n-2} + b_{n-1} x_{n-1} + c_{n-1} x_n = d_{n-1},$$

$$a_n x_n + b_n x_n = d_n.$$

Wykazać, że algorytm

Etap I:

```

1       $q_0 := 0$ 
2       $u_0 := 0$ 
3      for  $k := 1$  to  $n$ 
4          do begin
5               $p_k := a_k q_{k-1} + b_k$ 
6               $q_k := -\frac{c_k}{p_k}$ 
7               $u_k := \frac{d_k - a_k u_{k-1}}{p_k}$ 
8      end

```

Etap II:

```

1       $x_n := u_n$ 
2      for  $k := n-1$  downto 1
3          do  $x_k := q_k x_{k+1} + u_k$ 

```

wyznacza rozwiązanie tego układu równań.

Zad. 2.15. Odwracanie macierzy trójkątnej górnej.

Niech A będzie macierzą trójkątną górną. Wiadomo, że macierz odwrotna do trójkątnej górnej jest również macierzą trójkątną górną. Oznaczmy $W = A^{-1}$. Korzystając z równości macierzowej

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ 0 & w_{22} & \dots & w_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix},$$

opracować algorytm obliczania elementów macierzy W .

Zad. 2.16. Odwracanie macierzy z wykorzystaniem rozkładu trójkątno-trójkątnego.

Opracować algorytm odwracania macierzy kwadratowej A według następującej koncepcji:

- przedstawić macierz A w postaci $A = LU$,
- wyznaczyć L^{-1} oraz U^{-1} ,
- wyznaczyć A^{-1} z równości $A^{-1} = U^{-1}L^{-1}$.

Zad. 2.17. Odwracanie macierzy z wykorzystaniem rozkładu ortogonalno-trójkątnego.

Opracować algorytm odwracania macierzy kwadratowej A według następującej koncepcji:

- przedstawić macierz A w postaci $A = QR$,
- wyznaczyć R^{-1} ,
- wyznaczyć A^{-1} z równości $A^{-1} = R^{-1}Q^T$.

Zad. 2.18. *Odwracanie macierzy przez podział na bloki.*

Opracować algorytm odwracania macierzy kwadratowej A według poniższej koncepcji. Macierz kwadratową A stopnia n możemy przedstawić w postaci

$$A = \begin{bmatrix} P & Q \\ R & S \end{bmatrix},$$

gdzie P i S są macierzami kwadratowymi odpowiednio stopnia p i s , Q jest macierzą o wymiarach $p \times s$, a R – macierzą o wymiarach $s \times p$. Podobnie zapiszemy

$$A^{-1} = \begin{bmatrix} K & L \\ M & N \end{bmatrix},$$

gdzie K, L, M, N mają takie same wymiary, jak odpowiednio macierze P, Q, R, S . Ponieważ $AA^{-1} = E$, więc możemy napisać dwa układy równań:

$$\begin{cases} PK + QM = E_{p \times p}, \\ RK + SM = 0_{s \times p} \end{cases}$$

$$\begin{cases} PL + QN = 0_{p \times s}, \\ RL + SN = E_{s \times s} \end{cases}$$

Z pierwszego układu równań wyznaczamy macierze K i M :

$$K = (P - QS^{-1}R)^{-1},$$

$$M = -S^{-1}RK,$$

z drugiego układu – macierze N i L :

$$N = (S - RP^{-1}Q)^{-1},$$

$$L = -P^{-1}QN.$$

Zad. 2.19. Obliczyć wskaźniki uwarunkowania następujących układów równań liniowych. Skomentuj otrzymane wyniki.

a) $9x_1 + 0.01x_2 = 18.05,$
 $8x_1 + 0.01x_2 = 16.05,$

b) $5x_1 + 551x_2 = 591,$
 $3x_1 + 331x_2 = 355,$

c) $6x_1 + 15x_2 = 78,$
 $7x_1 + 12x_2 = 80,$

d) $x_1 + \frac{1}{2}x_2 = 5,$

$\frac{1}{2}x_1 + \frac{1}{3}x_2 = 3,$

e) $(\sin \alpha)x_1 - (\cos \alpha)x_2 = 8$
 $(\cos \alpha)x_1 + (\sin \alpha)x_2 = 7$ (α jest parametrem – dowolną liczbą rzeczywistą).

Zad. 2.20. Rozwiązać poniższe układy równań liniowych za pomocą wybranej metody skończonej. W trakcie obliczeń wyniki zaokrągać na piątym miejscu po kropce dziesiętnej.

$$4.5x_1 + 2.8x_2 + 4.8x_3 + 2.2x_4 = 21.86,$$

$$9.3x_1 + 5.8x_2 + 9.9x_3 + 4.5x_4 = 45.48,$$

a) $0.1x_1 - 3.1x_2 + 4.2x_3 - 3.5x_4 = 8.38,$

$$1.4x_1 - 1.2x_2 + 4.2x_3 - 3.4x_4 = 21.68.$$

$$7.8x_1 + 8.4x_2 + 9.2x_3 + 1.8x_4 = 92.4,$$

$$0.5x_1 + 0.5x_2 + 4.3x_3 + 4.5x_4 = 34.4,$$

b) $6.9x_1 + 7.4x_2 + 8.0x_3 - 3.5x_4 = 60.8,$

$$5.3x_1 + 5.7x_2 + 6.1x_3 - 0.7x_4 = 54.6.$$

$$0.5x_1 + 0.3x_2 + 0.9x_3 + 0.4x_4 = 0.8,$$

$$9.5x_1 + 5.6x_2 + 17.2x_3 + 7.6x_4 = 14.9,$$

c) $0.2x_1 + 8.3x_2 - 16.0x_3 - 2.6x_4 = 49.1,$

$$9.9x_1 + 0.3x_2 + 30.0x_3 + 10.1x_4 = -20.7.$$

$$0.6x_1 + 0.7x_2 + 0.2x_3 + 0.2x_4 = 38.83,$$

$$12.0x_1 + 14.1x_2 + 0.9x_3 + 0.2x_4 = 602.21$$

d) $0.3x_1 + 0.4x_2 + 15.0x_3 + 4.2x_4 = 703.79,$

$$15.0x_1 + 17.1x_2 + 0.6x_3 + 14.0x_4 = 879.47.$$

$$24.0x_1 + 0.3x_2 + 16.0x_3 - 15.1x_4 = 13.82,$$

e) $0.2x_1 + 0.1x_3 = 0.21,$

$$12.5x_1 + 0.9x_2 + 16.0x_3 - 36.2x_4 = -13.95,$$

$$0.5x_1 + 24.0x_2 + 0.2x_3 + 86.8x_4 = 95.42.$$

$$1.0x_1 + 6.0x_2 - 9.0x_3 + 4.0x_4 = 4.0,$$

$$3.0x_1 + 18.1x_2 - 27.2x_3 + 12.1x_4 = 12.0,$$

f) $15.0x_1 + 2.0x_2 + 41.1x_3 + 7.2x_4 = 236.4,$

$$0.7x_1 + 3.0x_2 - 3.9x_3 + 2.1x_4 = 5.3.$$

$$\begin{aligned} & 20.0x_1 + 10.2x_2 + 39.6x_3 + 24.1x_4 = 38.8, \\ \textcircled{1} \quad & 12.0x_1 + 6.1x_2 + 23.8x_3 + 14.4x_4 = 22.8, \\ & 8.0x_1 + 25.0x_2 - 25.9x_3 + 12.3x_4 = 277.0, \\ & 6.0x_1 + 9.0x_2 + 0.1x_3 - 35.3x_4 = -87.4. \end{aligned}$$

$$\begin{aligned} & 0.6x_1 + 1.1x_2 - 0.4x_3 + 1.2x_4 = 8.32, \\ \textcircled{2} \quad & 0.8x_1 + 1.4x_2 - 0.4x_3 + 1.3x_4 = 10.42, \\ & 0.2x_1 + 0.3x_2 + 0.1x_3 + 1.9x_4 = 9.50, \\ & 0.6x_1 + 1.2x_2 - 0.6x_3 + 2.2x_4 = 11.42. \end{aligned}$$

$$\begin{aligned} & 0.1x_1 + 0.2x_2 - 0.1x_3 + 1.2x_4 = 3.45, \\ \textcircled{3} \quad & 0.3x_1 + 0.4x_2 + 0.1x_3 + 1.3x_4 = 4.23, \\ & 0.2x_1 + 0.3x_2 + 0.2x_3 + 0.6x_4 = 4.47, \\ & 0.1x_1 + 0.8x_2 - 1.3x_3 + 1.2x_4 = 0.03. \end{aligned}$$

$$\begin{aligned} & 0.2x_1 + 0.7x_2 - 0.8x_3 + 1.2x_4 = 34.8, \\ \textcircled{4} \quad & 0.6x_1 + 2.4x_2 - 3.1x_3 + 0.3x_4 = 64.1, \\ & 0.3x_1 + 0.7x_2 - 0.5x_3 + 0.6x_4 = 24.4, \\ & 0.4x_1 + 0.2x_2 + 0.6x_3 + 1.2x_4 = 19.6. \end{aligned}$$

$$\begin{aligned} & 0.3x_1 + 0.6x_2 - 0.5x_3 + 0.5x_4 = 23.6, \\ \textcircled{5} \quad & 0.7x_1 + 1.2x_2 - 0.3x_3 + 1.1x_4 = 44.8, \\ & 0.2x_1 + 0.8x_2 - 1.0x_3 + 0.5x_4 = 28.4, \\ & 1.3x_1 + 0.2x_2 + 3.5x_3 + 0.9x_4 = 7.6. \end{aligned}$$

$$\begin{aligned} & 4.2x_1 + 4.8x_2 + 2.9x_3 + 4.0x_4 = 155.2, \\ \textcircled{6} \quad & 5.5x_1 + 6.3x_2 + 3.8x_3 + 5.3x_4 = 204.6, \\ & 8.4x_1 + 9.3x_2 + 6.6x_3 + 9.5x_4 = 331.0, \\ & 7.2x_1 + 4.8x_2 + 8.6x_3 + 5.5x_4 = 175.6. \end{aligned}$$

$$\begin{aligned} & 8.1x_1 + 5.6x_2 + 13.1x_3 + 4.5x_4 = 64.90, \\ \textcircled{7} \quad & 8.8x_1 + 6.1x_2 + 14.2x_3 + 4.9x_4 = 70.52, \\ & 4.8x_1 + 7.6x_2 - 0.9x_3 + 9.2x_4 = 51.36, \\ & 2.7x_1 + 5.8x_2 - 3.5x_3 + 4.8x_4 = 26.09. \end{aligned}$$

$$\begin{aligned} & 9.1x_1 + 6.1x_2 + 15.1x_3 + 7.3x_4 = -40.0, \\ \textcircled{8} \quad & 12.1x_1 + 8.1x_2 + 20.1x_3 + 9.7x_4 = -53.2, \\ & 15.5x_1 + 4.8x_2 + 36.8x_3 + 9.1x_4 = -74.5, \\ & 5.6x_1 + 16.0x_2 - 15.2x_3 + 12.5x_4 = -11.9. \end{aligned}$$

$$\begin{aligned} & 15.1x_1 + 8.9x_2 + 27.5x_3 + 11.4x_4 = 199.59, \\ \text{o) } & 20.1x_1 + 11.9x_2 + 36.5x_3 + 15.2x_4 = 265.45, \\ & 15.5x_1 + 4.8x_2 + 36.8x_3 + 9.2x_4 = 222.69, \\ & 5.6x_1 + 16.0x_2 - 15.2x_3 + 9.7x_4 = 18.36. \end{aligned}$$

$$\begin{aligned} \text{p) } & 15.1x_1 + 0.3x_2 + 21.0x_3 + 1.2x_4 = 20.97, \\ & 20.0x_1 + 0.4x_2 + 28.0x_3 + 1.5x_4 = 27.80, \\ & 12.0x_1 + 0.2x_2 + 15.0x_3 + 1.6x_4 = 16.02, \\ & 16.0x_1 + 0.5x_2 + 27.0x_3 + 1.4x_4 = 27.48. \end{aligned}$$

$$\begin{aligned} \text{q) } & 0.3x_1 + 18.1x_2 + 2.0x_3 + 28.0x_4 = 35.83, \\ & 0.2x_1 + 12.1x_2 + 1.3x_3 + 16.0x_4 = 22.33, \\ & 0.5x_1 + 30.0x_2 + 2.8x_3 + 13.0x_4 = 39.10, \\ & 0.1x_1 + 17.0x_2 + 3.1x_3 + 30.0x_4 = 42.60. \end{aligned}$$

$$\begin{aligned} \text{r) } & 2.1x_1 + 20.1x_2 + 2.0x_3 + 0.9x_4 = 45.22, \\ & 4.2x_1 + 40.1x_2 + 3.0x_3 + 1.6x_4 = 81.02, \\ & 3.0x_1 + 28.7x_2 + 2.5x_3 + 1.2x_4 = 61.14, \\ & 5.0x_1 + 45.0x_2 + 2.0x_3 + 2.9x_4 = 90.30. \end{aligned}$$

$$\begin{aligned} \text{s) } & 0.1x_1 + 0.3x_2 + 0.4x_3 + 0.2x_4 = 23.5, \\ & 0.4x_1 + 0.8x_2 + 0.6x_3 + 0.3x_4 = 46.0, \\ & 0.3x_1 + 0.1x_2 + 0.7x_3 + 0.2x_4 = 30.0, \\ & 0.4x_1 + 0.8x_2 + 0.3x_3 + 0.3x_4 = 38.5. \end{aligned}$$

$$\begin{aligned} \text{t) } & 0.4x_1 + 1.7x_2 + 0.7x_3 + 0.9x_4 = 50.50, \\ & 0.2x_1 + 0.8x_2 + 0.4x_3 + 0.6x_4 = 27.18, \\ & 0.3x_1 + 0.3x_2 + 0.8x_3 + 0.8x_4 = 29.94, \\ & 0.4x_1 + 0.8x_2 + 0.3x_3 + 0.7x_4 = 29.13. \end{aligned}$$

Zad. 2.21. Rozwiązać poniższe układy równań liniowych za pomocą odpowiedniej metody skończonej. W trakcie obliczeń wyniki należy zaokrągać na piątym miejscu po kropce dziesiętnej.

$$\begin{aligned} \text{a) } & 4.8x_1 + 7.8x_2 - 7.2x_3 + 5.5x_4 = 67.19, \\ & 5.4x_1 + 8.7x_2 - 7.8x_3 + 6.2x_4 = 75.04, \\ & 5.6x_1 + 6.1x_2 + 3.6x_3 + 5.2x_4 = 49.40, \\ & 8.3x_1 + 5.2x_2 + 8.7x_3 + 6.2x_4 = 50.71. \end{aligned}$$

$$2.2x_1 + 3.9x_2 - 9.0x_3 + 1.4x_4 = -37.0,$$

b) $1.4x_1 + 4.1x_2 - 12.2x_3 + 1.2x_4 = -53.6,$

$$6.1x_1 + 3.7x_2 + 3.5x_3 + 2.6x_4 = 29.3,$$

$$3.6x_1 + 4.3x_2 - 6.3x_3 + 2.1x_4 = -22.1.$$

$$0.2x_1 + 0.9x_2 - 3.0x_3 + 0.8x_4 = 70.8,$$

c) $0.7x_1 + 0.3x_2 + 0.9x_3 + 0.9x_4 = 14.1,$

$$0.4x_1 + 0.6x_2 - 1.2x_3 + 0.7x_4 = 41.4,$$

$$0.5x_1 + 0.4x_2 + 0.1x_3 + 0.8x_4 = 22.9.$$

$$0.3x_1 + 0.9x_2 - 2.7x_3 + 0.2x_4 = 14.9,$$

d) $0.8x_1 + 2.4x_2 - 7.2x_3 + 0.3x_4 = 40.9,$

$$0.6x_1 + 1.3x_2 - 3.4x_3 - 0.4x_4 = 25.8,$$

$$0.7x_1 + 0.8x_2 + 0.6x_3 - 0.7x_4 = 16.4.$$

$$2.4x_1 + 4.8x_2 - 2.4x_3 - 9.6x_4 = 7.20,$$

e) $4.2x_1 + 8.4x_2 - 4.2x_3 - 9.9x_4 = 17.43,$

$$5.6x_1 + 6.1x_2 + 9.7x_3 + 3.1x_4 = 51.99,$$

$$4.9x_1 + 6.8x_2 + 4.2x_3 - 4.6x_4 = 35.59.$$

$$1.6x_1 + 3.2x_2 - 1.6x_3 - 4.6x_4 = 28.2,$$

f) $0.8x_1 + 1.6x_2 - 0.8x_3 - 2.4x_4 = 14.4,$

$$1.3x_1 + 2.1x_2 + 0.2x_3 - 1.9x_4 = 6.4,$$

$$0.6x_1 + 0.4x_2 - 4.2x_3 - 1.4x_4 = 28.0.$$

$$4.2x_1 + 8.4x_2 + 7.6x_3 - 4.2x_4 = 2.88,$$

g) $4.5x_1 + 9.0x_2 + 8.2x_3 - 4.5x_4 = 3.12,$

$$6.2x_1 + 2.1x_2 + 4.8x_3 + 9.8x_4 = 4.41,$$

$$7.6x_1 + 7.8x_2 + 6.9x_3 + 4.1x_4 = 3.94.$$

$$1.3x_1 + 3.9x_2 + 3.1x_3 - 3.9x_4 = 23.96,$$

h) $2.1x_1 + 6.3x_2 + 3.7x_3 - 6.2x_4 = 34.94,$

$$1.8x_1 + 2.1x_2 + 4.1x_3 + 1.2x_4 = 40.15,$$

$$0.9x_1 + 1.8x_2 + 3.5x_3 - 0.9x_4 = 22.99.$$

$$0.3x_1 + 0.4x_2 + 0.3x_3 + 0.5x_4 = 0.19,$$

i) $12.9x_1 + 17.2x_2 + 12.9x_3 + 14.5x_4 = 4.67,$

$$0.1x_1 - 0.2x_2 + 1.1x_3 - 0.3x_4 = -0.67,$$

$$15.7x_1 + 18.9x_2 + 21.7x_3 + 16.9x_4 = 2.30.$$

$$\begin{aligned} & 0.6x_1 + 15.6x_2 + 8.4x_3 + 0.5x_4 = 42.23, \\ \text{j)} \quad & 0.8x_1 + 20.8x_2 + 11.2x_3 + 12.3x_4 = 64.45, \\ & 0.7x_1 + 18.4x_2 + 9.9x_3 + 0.3x_4 = 49.53, \\ & 0.3x_1 + 25.5x_2 + 13.1x_3 + 27.3x_4 = 80.96. \end{aligned}$$

$$\begin{aligned} & 20.0x_1 + 30.0x_2 + 12.5x_3 + 16.3x_4 = 349.1, \\ \text{k)} \quad & 16.0x_1 + 24.0x_2 + 10.0x_3 + 13.3x_4 = 281.1, \\ & 32.0x_1 + 22.0x_2 + 13.5x_3 + 24.5x_4 = 404.5, \\ & 28.0x_1 + 36.7x_2 + 16.2x_3 + 11.3x_4 = 379.0. \end{aligned}$$

$$\begin{aligned} \text{l)} \quad & 15.0x_1 + 5.5x_2 + 20.0x_3 + 0.8x_4 = 35.20, \\ & 21.0x_1 + 7.7x_2 + 28.0x_3 + 1.1x_4 = 49.10, \\ & 12.9x_1 + 6.1x_2 + 17.2x_3 + 0.5x_4 = 32.69, \\ & 18.4x_1 + 4.8x_2 + 24.5x_3 + 0.7x_4 = 34.80. \end{aligned}$$

$$\begin{aligned} \text{m)} \quad & 10.0x_1 + 25.5x_2 + 17.5x_3 + 17.4x_4 = 51.73, \\ & 0.4x_1 + 1.0x_2 + 0.7x_3 + 0.7x_4 = 2.09, \\ & 28.8x_1 + 21.4x_2 + 25.1x_3 + 23.5x_4 = -17.23, \\ & 3.7x_1 + 4.2x_2 + 3.9x_3 + 4.8x_4 = 8.05. \end{aligned}$$

$$\begin{aligned} \text{n)} \quad & 20.0x_1 + 30.0x_2 + 25.0x_3 + 2.8x_4 = -73.94, \\ & 0.8x_1 + 1.2x_2 + 1.0x_3 + 0.1x_4 = -3.02, \\ & 18.6x_1 + 22.4x_2 + 20.5x_3 + 0.6x_4 = -82.21, \\ & 0.3x_1 + 0.8x_2 + 0.6x_3 + 15.0x_4 = 76.66. \end{aligned}$$

$$\begin{aligned} \text{o)} \quad & 0.3x_1 + 0.2x_2 + 0.3x_3 + 2.8x_4 = 45.39, \\ & 0.6x_1 + 0.4x_2 + 0.5x_3 + 1.4x_4 = 37.14, \\ & 0.5x_1 + 0.1x_2 + 0.3x_3 + 2.1x_4 = 37.39, \\ & 0.4x_1 + 0.7x_2 + 0.5x_3 - 0.8x_4 = 12.02. \end{aligned}$$

$$\begin{aligned} \text{p)} \quad & 0.8x_1 + 0.6x_2 - 0.1x_3 + 1.8x_4 = 39.88, \\ & 1.2x_1 + 0.9x_2 + 0.4x_3 + 2.4x_4 = 62.38, \\ & 0.4x_1 + 0.1x_2 + 0.3x_3 + 1.6x_4 = 27.24, \\ & 0.4x_1 + 0.7x_2 + 0.1x_3 - 0.8x_4 = 11.08. \end{aligned}$$

$$\begin{aligned} \text{q)} \quad & 0.8x_1 + 4.5x_2 + 6.2x_3 + 5.5x_4 = 29.20, \\ & 6.4x_1 + 3.6x_2 + 5.0x_3 + 1.6x_4 = 31.70, \\ & 7.2x_1 + 6.2x_2 + 6.7x_3 - 1.5x_4 = 41.90, \\ & 5.6x_1 + 8.2x_2 + 6.7x_3 + 7.6x_4 = 0.83. \end{aligned}$$

$$\begin{aligned} 5.0x_1 + 7.5x_2 + 6.1x_3 + 2.6x_4 &= 11.77, \\ \textcircled{1} \quad 6.0x_1 + 9.0x_2 + 7.3x_3 + 3.1x_4 &= 14.10, \\ 4.8x_1 + 3.2x_2 + 3.6x_3 + 2.4x_4 &= 8.44, \\ 2.4x_1 + 2.8x_2 + 2.5x_3 + 1.2x_4 &= 5.09. \end{aligned}$$

$$\begin{aligned} \textcircled{2} \quad 0.5x_1 + 1.4x_2 + 1.4x_3 + 3.2x_4 &= 0.74, \\ 2.5x_1 + 7.0x_2 + 7.1x_3 + 18.0x_4 &= 3.19, \\ 0.9x_1 - 0.8x_2 - 0.5x_3 - 4.2x_4 &= 2.93, \\ 1.2x_1 - 0.6x_2 + 0.2x_3 + 4.6x_4 &= 1.44. \end{aligned}$$

$$\begin{aligned} \textcircled{3} \quad 2.4x_1 + 3.8x_2 + 2.9x_3 + 6.6x_4 &= 3.91, \\ 4.8x_1 + 7.6x_2 + 5.9x_3 + 15.3x_4 &= 7.28, \\ 8.1x_1 - 2.7x_2 + 4.2x_3 + 5.3x_4 &= 19.47, \\ 1.2x_1 + 6.3x_2 + 3.4x_3 + 15.8x_4 &= -1.80. \end{aligned}$$

Zad. 2.22. Dla poniższych układów równań liniowych przyjąć

$$\mathbb{X}^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

a następnie obliczyć $X^{(1)}, X^{(2)}$ według:

- a) metody iteracji prostej,
- b) metody Gaussa-Seidela.

Porównać otrzymane wyniki.

$$\begin{aligned} \textcircled{1} \quad 10x_1 + 2x_2 + x_3 &= 48, \\ 3x_1 + 10x_2 + 2x_3 &= 46, \\ 2x_1 + 3x_2 + 10x_3 &= 37. \end{aligned}$$

$$\begin{aligned} \textcircled{2} \quad 10x_1 - 4x_2 - 3x_3 &= -7, \\ 5x_1 + 10x_2 - 4x_3 &= 3, \\ 4x_1 + 5x_2 + 10x_3 &= 44. \end{aligned}$$

$$\begin{aligned} \textcircled{3} \quad 10x_1 + 3x_2 - x_3 &= 22, \\ 2x_1 - 5x_2 + 3x_3 &= 27, \\ x_1 + 2x_2 + 10x_3 &= 51. \end{aligned}$$

$$\begin{aligned} \textcircled{4} \quad 2x_1 + 0.5x_2 &= 48, \\ 0.5x_1 + 2x_2 + 0.5x_3 &= 46, \\ + 0.5x_2 + 2x_3 &= 37. \end{aligned}$$

- $10x_1 - 2x_2 + 7x_3 = 62,$
e) $2x_1 + 20x_2 - 3x_3 = 24,$
 $7x_1 - 3x_2 + 10x_3 = 42.$

Zad. 2.23. Dla jakich wartości parametru α zbieżny jest proces Gaussa-Seidela dla układu równań liniowych:

$$\begin{aligned}\alpha x_1 + x_2 &= 13, \\ x_1 + \alpha x_2 &= 17.\end{aligned}$$

Zad. 2.24. Dla jakich wartości parametrów β, γ zbieżny jest proces Jacobiego dla układu równań liniowych:

$$\begin{aligned}x_1 + \beta x_2 &= 1, \\ \beta x_1 + \gamma x_2 &= 1.\end{aligned}$$

Zad. 2.25. Rozwiązać za pomocą metody Czebyszewa następujące układy równań liniowych:

- $5x_1 + 4x_2 + 2x_3 = 66,$
a) $4x_1 + 6x_2 + 3x_3 = 71,$
 $2x_1 + 3x_2 + 7x_3 = 52.$

- $10x_1 - 2x_2 + 4x_3 = 36,$
b) $-2x_1 + 10x_2 - 3x_3 = 18,$
 $4x_1 - 3x_2 + 10x_3 = 56.$

- $4x_1 - 5x_2 - 3x_3 = -1,$
c) $-5x_1 + 7x_2 + 2x_3 = 1,$
 $-3x_1 + 2x_2 + 8x_3 = 3.$

- $2x_1 + 4x_2 = 14,$
d) $4x_1 + 9x_2 + x_3 = 38,$
 $x_2 + 2x_3 = 19.$

- $3x_1 + 6x_2 + 3x_3 = 48,$
e) $8x_1 + 4x_2 - 6x_3 = 10,$
 $4x_1 + 7x_2 + 5x_3 = 65.$

Zad. 2.26. Dla jakich wartości parametru α układ równań liniowych

$$\begin{aligned}4x_1 + 2x_2 &= 1, \\ 2x_1 + \alpha x_2 + x_3 &= 1, \\ x_2 + x_3 &= 1\end{aligned}$$

można rozwiązać metodą Czebyszewa?

Zad. 2.27. Rozwiązać za pomocą metody najszybszego spadku następujące układy równań liniowych:

$$8x_1 + 4x_2 + 2x_3 = 44,$$

a) $4x_1 + 8x_2 + 4x_3 = 64,$

$$2x_1 + 4x_2 + 8x_3 = 68.$$

$$4x_1 + 3x_2 + 2x_3 = 41,$$

b) $3x_1 + 3x_2 + 2x_3 = 38,$

$$2x_1 + 2x_2 + 2x_3 = 30.$$

$$5x_1 + 3x_2 + 2x_3 = 27,$$

c) $3x_1 + 4x_2 + 3x_3 = 30,$

$$2x_1 + 3x_2 + 3x_3 = 25.$$

$$4x_1 + 7x_2 + 4x_3 = 45,$$

d) $5x_1 + 5x_2 + 3x_3 = 43,$

$$3x_1 + 7x_2 + 6x_3 = 42.$$

$$x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 = 22.0,$$

e) $\frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 = 14.5,$

$$\frac{1}{3}x_1 + \frac{1}{4}x_2 + \frac{1}{5}x_3 = 11.0.$$

Zad. 2.28. Wyznaczyć przybliżone z dokładnością ε rozwiązania poniższych układów równań liniowych za pomocą odpowiedniej metody iteracyjnej.

$$20x_1 - x_2 - 2x_3 - 3x_4 = 85,$$

a) $5x_1 + 20x_2 - 3x_3 - 4x_4 = 25,$

$$4x_1 + 3x_2 + 20x_3 - 5x_4 = 93,$$

$$3x_1 + 2x_2 + x_3 + 20x_4 = 61.$$

$$10x_1 - 3x_2 - 2x_3 - x_4 = 31,$$

b) $3x_1 + 10x_2 - 3x_3 - 2x_4 = -17,$

$$2x_1 + 3x_2 + 10x_3 - 3x_4 = 49,$$

$$x_1 + 2x_2 + 3x_3 + 10x_4 = -19.$$

$$15x_1 + 3x_2 + 5x_3 - 6x_4 = 15,$$

c) $4x_1 - 25x_2 + 2x_3 + 4x_4 = -16,$

$$2x_1 - 5x_2 + 20x_3 + x_4 = -56,$$

$$x_1 + 3x_2 + 6x_3 - 12x_4 = -80.$$

$$30x_1 - 10x_2 + 8x_3 - 5x_4 = 88.1,$$

d) $7x_1 + 30x_2 - 10x_3 + 8x_4 = 76.3,$
 $5x_1 + 7x_2 + 30x_3 - 10x_4 = 140.6,$
 $2x_1 + 5x_2 + 7x_3 + 30x_4 = 195.1.$

$$25x_1 - 4x_2 - 8x_3 + 12x_4 = 11.0,$$

e) $12x_1 - 25x_2 - 4x_3 - 8x_4 = 4.8,$
 $8x_1 + 12x_2 + 25x_3 - 4x_4 = 6.5,$
 $3x_1 + 8x_2 + 12x_3 - 25x_4 = -14.1.$

$$50x_1 + 20x_2 + 8x_3 + 5x_4 = 20.1,$$

f) $10x_1 - 50x_2 + 20x_3 + 8x_4 = 18.4,$
 $15x_1 + 10x_2 - 50x_3 + 20x_4 = 51.5,$
 $22x_1 + 15x_2 + 10x_3 + 50x_4 = 14.9.$

$$2.0x_1 + 0.8x_2 + 0.6x_3 + 0.5x_4 = 2.22,$$

g) $0.8x_1 - 2.0x_2 + 0.6x_3 + 0.5x_4 = 8.30,$
 $0.6x_1 + 0.6x_2 + 2.0x_3 + 0.7x_4 = 2.16,$
 $0.5x_1 + 0.5x_2 + 0.7x_3 - 2.0x_4 = 6.81.$

$$1.8x_1 - 0.6x_2 + 0.6x_3 - 0.5x_4 = 20.82,$$

h) $0.6x_1 - 1.6x_2 + 0.2x_3 - 0.6x_4 = 21.68,$
 $0.5x_1 - 0.4x_2 + 2.2x_3 - 0.8x_4 = 22.24,$
 $0.4x_1 - 0.5x_2 + 1.1x_3 - 2.0x_4 = 25.91.$

$$1.2x_1 + 0.3x_2 + 0.5x_3 + 0.3x_4 = 10.0,$$

i) $0.4x_1 + 0.7x_2 + 0.2x_3 + 0.1x_4 = 5.3,$
 $0.3x_1 + 0.2x_2 + 0.9x_3 + 0.3x_4 = 8.1,$
 $0.2x_1 + 0.1x_2 + 0.1x_3 + 0.8x_4 = 6.4.$

$$3.5x_1 + 1.2x_2 + 0.9x_3 + 1.3x_4 = 7.69,$$

j) $0.8x_1 + 3.8x_2 + 0.2x_3 + 1.7x_4 = 9.56,$
 $1.8x_1 + 0.9x_2 + 3.2x_3 + 0.3x_4 = 9.19,$
 $0.6x_1 + 1.5x_2 + 1.4x_3 + 3.6x_4 = 11.83.$

$$9.5x_1 + 2.1x_2 + 1.9x_3 + 2.8x_4 = 58.7,$$

k) $2.6x_1 + 9.2x_2 + 2.2x_3 + 1.8x_4 = -39.4,$
 $3.2x_1 + 2.4x_2 + 9.9x_3 + 2.7x_4 = 54.7,$
 $0.8x_1 + 1.8x_2 + 2.6x_3 + 8.6x_4 = -33.6.$

D) $25.2x_1 + 0.7x_2 + 7.7x_3 + 0.3x_4 = 28.64,$
 $0.6x_1 + 20.8x_2 + 0.2x_3 + 9.8x_4 = -22.66,$
 $8.2x_1 + 9.2x_2 + 28.7x_3 + 4.8x_4 = 30.24,$
 $2.4x_1 + 0.3x_2 + 0.8x_3 + 32.8x_4 = -26.85.$

III) $45.0x_1 + 0.1x_2 + 8.7x_3 + 7.9x_4 = 179.29,$
 $7.3x_1 + 36.0x_2 + 5.8x_3 + 0.1x_4 = 177.20,$
 $0.2x_1 + 9.2x_2 + 42.0x_3 + 4.8x_4 = 234.48,$
 $8.0x_1 + 8.7x_2 + 0.2x_3 + 32.0x_4 = 230.68.$

II) $10.8x_1 - 0.4x_2 - 0.3x_3 - 9.8x_4 = 73.38,$
 $0.3x_1 + 10.0x_2 - 8.9x_3 - 0.2x_4 = 33.48,$
 $8.2x_1 + 0.2x_2 + 10.2x_3 - 0.5x_4 = 76.59,$
 $0.1x_1 + 9.2x_2 + 0.1x_3 + 11.4x_4 = 49.80.$

II) $12.0x_1 + 0.4x_2 + 0.2x_3 + 7.8x_4 = 77.88,$
 $7.8x_1 + 10.0x_2 + 0.3x_3 + 0.4x_4 = 83.82,$
 $0.2x_1 + 8.6x_2 + 15.0x_3 + 0.2x_4 = 106.88,$
 $0.1x_1 + 0.3x_2 + 9.5x_3 + 11.0x_4 = 60.52.$

D) $20.0x_1 + 6.9x_2 + 0.7x_3 + 4.8x_4 = 87.67,$
 $4.7x_1 + 14.0x_2 + 8.9x_3 + 0.4x_4 = 53.29,$
 $0.7x_1 + 6.8x_2 + 25.0x_3 + 9.8x_4 = 83.37,$
 $7.1x_1 + 0.3x_2 + 8.3x_3 + 16.0x_4 = 60.74.$

II) $15.0x_1 + 8.1x_2 - 0.8x_3 + 6.1x_4 = 40.33,$
 $8.3x_1 + 25.0x_2 + 6.1x_3 - 0.3x_4 = 81.47,$
 $0.3x_1 - 5.9x_2 + 16.0x_3 + 9.8x_4 = 34.77,$
 $6.4x_1 + 7.3x_2 - 9.1x_3 + 32.0x_4 = 55.46.$

E) $24.0x_1 - 8.9x_2 + 8.7x_3 + 6.4x_4 = 74.25,$
 $8.5x_1 + 20.0x_2 - 6.5x_3 + 0.3x_4 = 48.55,$
 $9.7x_1 - 5.5x_2 + 25.0x_3 - 9.8x_4 = 46.35,$
 $5.8x_1 + 1.3x_2 - 9.8x_3 + 30.0x_4 = 90.45.$

5) $2.0x_1 - 0.5x_2 - 0.7x_3 - 0.8x_4 = 10.04,$
 $0.9x_1 + 2.0x_2 - 0.1x_3 - 0.9x_4 = 22.09,$
 $0.6x_1 + 0.3x_2 + 2.0x_3 - 0.3x_4 = 14.06,$
 $0.5x_1 + 0.8x_2 + 0.7x_3 + 2.0x_4 = 18.79.$

$$\begin{aligned} t) \quad & 1.5x_1 - 0.6x_2 + 0.2x_3 + 0.1x_4 = 6.06, \\ & 0.2x_1 - 1.4x_2 + 0.4x_3 + 0.8x_4 = 2.74, \\ & 0.4x_1 - 0.6x_2 + 1.3x_3 + 0.3x_4 = 4.94, \\ & 0.1x_1 - 0.2x_2 + 0.7x_3 + 1.6x_4 = 0.38. \end{aligned}$$

Zad. 2.29. Rozwiązać następujące układy równań liniowych.

$$\begin{array}{lll} x_1 & +0.5x_2 & = 7.05 \\ 0.5x_1 & +x_2 & +0.5x_3 & = 8.25 \\ & 0.5x_2 & +x_3 & +0.5x_4 & = 7.15 \\ a) & & 0.5x_3 & +x_4 & +0.5x_5 & = 6.10 \\ & & 0.5x_4 & +x_5 & 0.5x_6 & = 6.30 \\ & & & 0.5x_5 & +x_6 & +0.5x_7 & = 7.65 \\ & & & & 0.5x_6 & +x_7 & +0.5x_8 & = 9.00 \\ & & & & & 0.5x_7 & x_8 & = 7.45 \end{array}$$

$$\begin{array}{lll} 5x_1 & +4x_3 & = 72.2 \\ +5x_2 & & +4x_4 & = 53.6 \\ 2x_1 & +5x_3 & +4x_5 & = 74.6 \\ b) \quad 2x_2 & +5x_4 & 4x_6 & = 64.0 \\ & 2x_3 & +5x_5 & 4x_7 & = 66.8 \\ & 2x_4 & +5x_6 & 4x_8 & = 67.6 \\ & & 2x_5 & +5x_7 & = 43.6 \\ & & & 2x_6 & 5x_8 & = 46.4 \end{array}$$

$$\begin{array}{lll} 4x_1 & +8x_2 & +4x_3 & +2x_4 & = 40.0 \\ 3x_1 & +2x_2 & +5x_3 & +3x_4 & +6x_5 & = 31.8 \\ 6x_2 & +6x_3 & +4x_4 & +4x_5 & +3x_6 & = 38.5 \\ c) \quad 2x_3 & +2x_4 & +7x_5 & +4x_6 & +5x_7 & = 29.0 \\ & 5x_4 & +8x_5 & +8x_6 & +4x_7 & +2x_8 & = 39.0 \\ & 3x_5 & +5x_6 & +2x_7 & +3x_8 & & = 22.0 \\ & & 4x_6 & +3x_7 & +4x_8 & & = 22.0 \\ & & & 5x_7 & 6x_8 & & = 25.0 \end{array}$$

$$\begin{array}{lll} 8x_1 & +4x_2 & +3x_3 & +2x_4 & = 60.5 \\ 5x_1 & +8x_2 & +4x_3 & +3x_4 & = 65.0 \\ 4x_1 & +5x_2 & +8x_3 & +4x_4 & = 61.5 \\ d) \quad 3x_1 & +4x_2 & +5x_3 & +8x_4 & +4x_5 & +3x_6 & +2x_7 & +x_8 & = 82.0 \\ 2x_1 & +3x_2 & +4x_3 & +5x_4 & +8x_5 & +4x_6 & +3x_7 & +2x_8 & = 87.0 \\ x_1 & +2x_2 & +3x_3 & +4x_4 & +5x_5 & +8x_6 & +4x_7 & +3x_8 & = 90.0 \\ & & & & 4x_5 & +5x_6 & +8x_7 & +4x_8 & = 75.0 \\ & & & & 3x_5 & +4x_6 & +5x_7 & 8x_8 & = 78.0 \end{array}$$

$$\begin{array}{ccccccc}
 2x_1 & +6x_3 & & +4x_5 & & +5x_7 & = 46.0 \\
 & 3x_2 & & +3x_4 & & +3x_6 & +7x_8 = 61.0 \\
 & 5x_1 & +2x_3 & & +7x_5 & +3x_7 & = 42.0 \\
 \textcircled{e}) & 4x_2 & & +3x_4 & & +3x_6 & +2x_8 = 33.0 \\
 & 3x_1 & +2x_3 & & +2x_5 & +3x_7 & = 25.0 \\
 & 2x_2 & & +4x_4 & & +3x_6 & +5x_8 = 46.0 \\
 & 2x_1 & & +5x_3 & & +7x_5 & +2x_7 = 41.0 \\
 & 5x_2 & & +3x_4 & & +4x_6 & +3x_8 = 42.5
 \end{array}$$

$$\begin{array}{ccccccc}
 x_1 & +\frac{1}{2}x_2 & +\frac{1}{3}x_3 & +\frac{1}{4}x_4 & +\frac{1}{5}x_5 & +\frac{1}{6}x_6 & = 1 \\
 \frac{1}{2}x_1 & +\frac{1}{3}x_2 & +\frac{1}{4}x_3 & +\frac{1}{5}x_4 & +\frac{1}{6}x_5 & +\frac{1}{7}x_6 & = 1 \\
 \frac{1}{3}x_1 & +\frac{1}{4}x_2 & +\frac{1}{5}x_3 & +\frac{1}{6}x_4 & +\frac{1}{7}x_5 & +\frac{1}{8}x_6 & = 1 \\
 \textcircled{f}) & \frac{1}{4}x_1 & +\frac{1}{5}x_2 & +\frac{1}{6}x_3 & +\frac{1}{7}x_4 & +\frac{1}{8}x_5 & +\frac{1}{9}x_6 = 1 \\
 & \frac{1}{5}x_1 & +\frac{1}{6}x_2 & +\frac{1}{7}x_3 & +\frac{1}{8}x_4 & +\frac{1}{9}x_5 & +\frac{1}{10}x_6 = 1 \\
 & \frac{1}{6}x_1 & +\frac{1}{7}x_2 & +\frac{1}{8}x_3 & +\frac{1}{9}x_4 & +\frac{1}{10}x_5 & +\frac{1}{11}x_6 = 1
 \end{array}$$

Równania nieliniowe

3.1. Wstęp

W niniejszym rozdziale omówimy algorytmy wyznaczania rzeczywistych i jednokrotnych pierwiastków równań nieliniowych. Niech $y = f(x)$ będzie funkcją rzeczywistą zmiennej rzeczywistej x . Pierwiastkiem (zerem) równania

$$f(x) = 0 \quad (3.1)$$

(lub pierwiastkiem (zerem) funkcji $y = f(x)$) nazywamy każdą wartość \bar{x} zmiennej niezależnej x , dla której $f(\bar{x}) = 0$.

Zagadnienie przyblizonego obliczania pierwiastków definiujemy następująco: dla zadanej dokładności ε należy znaleźć takie \tilde{x} , dla którego zachodzi nierówność

$$|\bar{x} - \tilde{x}| < \varepsilon. \quad (3.2)$$

Przyblzoną wartość pierwiastka funkcji wyznaczamy w dwóch etapach:

- 1) najpierw lokalizujemy pierwiastki równania (3.1), tj.
 - a) znajdujemy liczbę p pierwiastków równania,
 - b) dla każdego pierwiastka \bar{x}_i znajdujemy taki przedział $[a_i ; b_i]$, że $\bar{x}_i \in [a_i ; b_i]$ oraz $[a_i ; b_i] \cap [a_j ; b_j] = \emptyset$ dla $i, j = 1, 2, \dots, p ; i \neq j$,
- 2) potem uściślamy przyblzoną wartość pierwiastka, tj. dla zadanego \bar{x}_i i zadanej dokładności ε znajdujemy taką wartość \tilde{x}_i , że $|\bar{x}_i - \tilde{x}_i| < \varepsilon$.

3.2. Lokalizacja pierwiastków

Dotychczas nie opracowano algorytmów lokalizacji pierwiastków rzeczywistych dowolnej funkcji ciągłej (odstępstwem od powyższego stwierdzenia jest twierdzenie Sturma, na mocy którego można skonstruować algorytm lokalizacji wszystkich rzeczywistych pierwiastków wielomianu o współczynnikach rzeczywistych).

Na etapie lokalizacji korzystamy z podstawowej wiedzy o funkcjach elementarnych.

Przykład 3.1. Zlokalizować rzeczywiste pierwiastki funkcji

$$f(x) = \ln(x+2) - 2x^2 + 1. \quad (3.3)$$

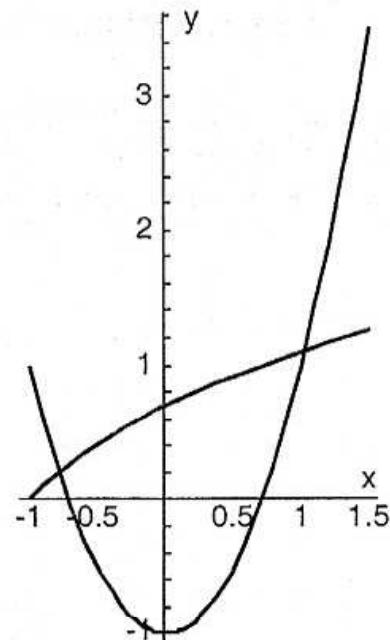
Rozwiązanie. Funkcję (3.3) przyrównujemy do zera, otrzymując równanie

$$\ln(x+2) - 2x^2 + 1 = 0,$$

które następnie przekształcamy do postaci

$$\ln(x+2) = 2x^2 - 1.$$

Z łatwością wykonujemy wykresy znanych funkcji $y_1 = \ln(x+2)$ oraz $y_2 = 2x^2 - 1$. Odcięte punktów przecięcia się wykresów są pierwiastkami funkcji (3.3). Z tabeli 3.1 oraz wykresu wnioskujemy, że istnieją dwa rzeczywiste pierwiastki funkcji (3.3) leżące w przedziałach $\bar{x}_1 \in [-0.8; -0.7]$, $\bar{x}_2 \in [1.0; 1.1]$.



Rys. 3.1. Wykresy funkcji $y_1 = \ln(x+2)$ i $y_2 = 2x^2 - 1$

Tabela 3.1. Lokalizacja pierwiastków funkcji $y = \ln(x+2) - 2x^2 + 1$

x	$y_1 = \ln(x+2)$	$y_2 = 2x^2 - 1$	$y = \ln(x+2) - 2x^2 + 1$
-1.0	0.00000	1.00000	-1.00000
-0.9	0.09531	0.62000	-0.52469
-0.8	0.18232	0.28000	-0.09768
-0.7	0.26236	-0.02000	0.28236
-0.6	0.33647	-0.28000	0.61647
-0.5	0.40547	-0.50000	0.90547
-0.4	0.47000	-0.68000	1.15000
...
0.7	0.99325	-0.02000	1.01325
0.8	1.02962	0.28000	0.74962
0.9	1.06471	0.62000	0.44471
1.0	1.09861	1.00000	0.09861
1.1	1.13140	1.42000	-0.28860
1.2	1.16315	1.88000	-0.71685
1.3	1.19392	2.38000	-1.18608
1.4	1.22378	2.92000	-1.69622

3.3. Metoda bisekcji (metoda połowienia przedziału)

Zakładamy, że wcześniej został zlokalizowany w przedziale $[a ; b]$ rzeczywisty, jednokrotny pierwiastek funkcji $y = f(x)$ – zatem $f(a)f(b) < 0$. Metoda bisekcji jest metodą uściślania przybliżonej wartości pierwiastka, która wymaga założenia:

$$f(x) \in C[a ; b]. \quad (3.4)$$

Metoda bisekcji

```

1      k := 0
2      a0 := a
3      b0 := b
4      while |bk - ak| ≥ ε
5          do      begin
6              ck :=  $\frac{a_k + b_k}{2}$ 
7              if f(ak)f(ck) ≤ 0
8                  then    begin
9                      ak+1 := ak
10                     bk+1 := ck
11                     end
12                 else    begin
13                     ak+1 := ck
14                     bk+1 := bk
15                     end
16                 k := k + 1
17             end
18             x̃ :=  $\frac{a_k + b_k}{2}$ 

```

Przykład 3.2. Wyznaczyć metodą bisekcji ujemny pierwiastek funkcji

$$f(x) = \ln(x + 2) - 2x^2 + 1$$

z dokładnością $\varepsilon = 10^{-5}$.

Rozwiązanie. W przykładzie 3.1 pokazano, że ujemny pierwiastek funkcji (3.3) leży w przedziale $[-0.8 ; -0.7]$. Obliczenia kolejnych kroków metody bisekcji są zestawione w tabeli 3.2.

Tabela 3.2. Wyznaczanie ujemnego pierwiastka funkcji
 $f(x) = \ln(x+2) - 2x^2 + 1$
z dokładnością $\varepsilon = 10^{-5}$ metodą bisekcji

$a =$	-0.800000	$b =$	-0.700000	$\varepsilon =$	1.0E-03
k	a_k	b_k	c_k	$f(a_k)f(c_k)$	$ b_k - a_k $
0	-0.800000	-0.700000	-0.750000	-9.6E-03	1.0E-01
1	-0.800000	-0.750000	-0.775000	-1.7E-04	5.0E-02
2	-0.800000	-0.775000	-0.787500	4.7E-03	2.5E-02
3	-0.787500	-0.775000	-0.781250	1.1E-03	1.3E-02
4	-0.781250	-0.775000	-0.778125	2.4E-04	6.2E-03
5	-0.778125	-0.775000	-0.776563	4.7E-05	3.1E-03
6	-0.776563	-0.775000	-0.775781	6.1E-06	1.6E-03
7	-0.775781	-0.775000	-0.775391	-2.2E-07	7.8E-04
8	-0.775781	-0.775391	-0.775586	8.3E-07	3.9E-04
9	-0.775586	-0.775391	-0.775488	1.3E-07	2.0E-04
10	-0.775488	-0.775391	-0.775439	6.8E-09	9.8E-05
11	-0.775439	-0.775391	-0.775415	-2.0E-09	4.9E-05
12	-0.775439	-0.775415	-0.775427	-5.3E-10	2.4E-05
13	-0.775439	-0.775427	-0.775433	2.1E-10	1.2E-05
14	-0.775433	-0.775427	-0.775430	-3.5E-11	6.1E-06

3.4. Metoda stycznych (metoda Newtona-Raphsona)

Metoda stycznych jest inną niż metoda bisekcji metodą uściślania przybliżonej wartości pierwiastka. Zakładamy zatem, że wcześniej został zlokalizowany w przedziale $[a ; b]$ niewyistny, jednokrotny pierwiastek funkcji $y = f(x)$; a zatem $f(a)f(b) < 0$. Metoda stycznych wymaga silniejszych niż metoda bisekcji założeń odnośnie do funkcji $y = f(x)$, a mianowicie:

D) $f(x) \in C^2[a ; b],$ (3.5)

D) dla każdego $x \in [a ; b]$ albo $f'(x) < 0$ albo $f'(x) > 0$ ($f'(x)$ nie zmienia znaku na przedziale $[a ; b]$), (3.6)

D) dla każdego $x \in [a ; b]$ albo $f''(x) < 0$ albo $f''(x) > 0$ ($f''(x)$ nie zmienia znaku na przedziale $[a ; b]$). (3.7)

Metoda Newtona-Raphsona

W charakterze x_0 wybieramy ten kraniec przedziału $[a ; b]$ (tzn. albo $x_0 = a$, albo $x_0 = b$), który spełnia warunek:

$$f(x_0)f''(x_0) > 0. \quad (3.8)$$

2. Wyznaczamy: $m = \min \{|f'(a)|; |f'(b)|\}$

3. **while** $\frac{|f(x_k)|}{m} \geq \varepsilon$

4. **do begin**

$$5. \quad x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}$$

6. $k := k + 1$

7. **end**

8. $\tilde{x}_k := x_k$

(3.9)

Ciąg wygenerowany według reguł (3.8)÷(3.9) nosi nazwę *procesu Newtona-Raphsona*. Jego własności zostały sformułowane w poniższym twierdzeniu.

Twierdzenie 3.1. Jeżeli funkcja $y = f(x)$ spełnia założenia (3.5)÷(3.7), to proces Newtona-Raphsona ma następujące własności:

$$1) \quad \lim_{k \rightarrow \infty} x_k = \bar{x} \quad (3.10)$$

- (proces jest zbieżny do pierwiastka funkcji $y = f(x)$ leżącego w przedziale $[a; b]$),
 2) ciąg (3.8)÷(3.9) jest ściśle monotoniczny, tzn. albo $x_0 < x_1 < \dots < x_k < \dots < \bar{x}$, albo

$$x_0 > x_1 > \dots > x_k > \dots > \bar{x}, \quad (3.11)$$

- 3) k -ty wyraz procesu spełnia nierówność:

$$|\bar{x} - x_k| \leq \frac{|f(x_k)|}{m}, \quad (3.12)$$

gdzie $m = \min \{|f'(a)|; |f'(b)|\}$.

Przykład 3.3. Obliczyć metodą Newtona-Raphsona ujemny pierwiastek funkcji

$$f(x) = \ln(x+2) - 2x^2 + 1$$

z dokładnością $\varepsilon = 10^{-5}$.

Rozwiążanie. W przykładzie 3.1 pokazano, że ujemny pierwiastek funkcji (3.3) leży w przedziale $[-0.8; -0.7]$. Na początku wykażemy, że funkcja (3.3) spełnia w przedziale $[-0.8; -0.7]$ warunki (3.5)÷(3.7).

— Dla $x > -2$ funkcja (3.3) jest ciągła i ma ciągle pochodne wszystkich rzędów.

$$— f'(x) = \frac{1}{x+2} - 4x = \frac{-4x^2 - 8x + 1}{x+2} \begin{cases} > 0 & \text{dla } -2 < x < \frac{\sqrt{5}}{2} - 1 \approx 0.118 \\ < 0 & \text{dla } x > \frac{\sqrt{5}}{2} - 1 \approx 0.118 \end{cases}$$

Zatem w przedziale $[-0.8; -0.7]$ spełniona jest nierówność $f'(x) > 0$.

$$— f''(x) = -\frac{1}{(x+2)^2} - 4 < 0 \quad \text{dla } x > -2.$$

Zatem funkcja (3.3) spełnia w przedziale $[-0.8 ; -0.7]$ warunki (3.5)÷(3.7) i w związku z tym możemy zastosować metodę stycznych do uściślenia przybliżonej wartości pierwiastka.

Wyznaczamy x_0 oraz m . Z tabeli 3.1 odczytujemy

$$f(-0.8) = -0.09768 < 0, \quad f(-0.7) = +0.28286 > 0.$$

Więc $x_0 = -0.8$, ponieważ $f(-0.8)f'(-0.8) > 0$. Następnie obliczamy

$$m = \min \left\{ \left| \frac{1}{-0.8+2} - 4(-0.8) \right|; \left| \frac{1}{-0.7+2} - 4(-0.7) \right| \right\} = \min \{4.033; 3.569\} \cong 3.569.$$

Wartości kolejnych iteracji metody stycznych są podane w tabeli 3.3.

Tabela 3.3. Wyznaczanie ujemnego pierwiastka funkcji

$$f(x) = \ln(x+2) - 2x^2 + 1 \\ z \text{ dokładnością } 10^{-5} \text{ metodą stycznych}$$

$$x_0 = -0.800000 \quad m = 3.569231 \quad \epsilon = 1.0E-03$$

k	x_k	$f(x_k)$	$f'(x_k)$	$\frac{f(x_k)}{f'(x_k)}$	$\frac{ f(x_k) }{m}$
0	-0.800000	-0.097678	4.033333	-0.024218	2.7E-02
1	-0.775782	-0.001374	3.919977	-0.000350	3.8E-04
2	-0.775432	0.000000	3.918341	0.000000	8.0E-08

3.5. Metoda iteracji prostej

Metoda iteracji prostej jest metodą uściślenia przybliżonej wartości pierwiastka. Założymy zatem, że wcześniej został zlokalizowany w przedziale $[a ; b]$ jednokrotny pierwiastek funkcji $y = f(x)$; zatem $f(a)f(b) < 0$.

Metoda iteracji prostej składa się z dwóch etapów.

I. Najpierw równanie

$$f(x) = 0, \quad (3.13)$$

przekształcamy do równoważnej postaci

$$x = \phi(x) \quad (3.14)$$

(takie przekształcenie jest zawsze wykonalne i zazwyczaj można tego dokonać kilkoma sposobami).

II. Wybieramy z przedziału $[a ; b]$ przybliżenie x_0 . Kolejne iteracje obliczamy ze wzoru

$$x_{k+1} = \phi(x_k) \text{ dla } k = 0, 1, \dots \quad (3.15)$$

W twierdzeniu 3.2 zostały sformułowane warunki wystarczające zbieżności ciągu (3.15) do pierwiastka funkcji (3.13) leżącego w przedziale $[a ; b]$.

Twierdzenie 3.2. Niech funkcja $y = \varphi(x)$ będzie określona, ciągła i różniczkowalna w przedziale domkniętym $[a ; b]$ oraz $\varphi'(x) \in [a ; b]$ dla każdego $x \in [a ; b]$. Jeśli nierówność

$$|\varphi'(x)| \leq q < 1$$

zachodzi dla każdego $x \in [a ; b]$, to:

- 1) proces (3.15) jest zbieżny niezależnie od wyboru $x_0 \in [a ; b]$ oraz $\lim_{k \rightarrow \infty} x_k = \bar{x} \in [a ; b]$,
- 2) zachodzą nierówności

$$|\bar{x} - x_k| \leq \frac{q}{1-q} |x_k - x_{k-1}| \leq \frac{q^k}{1-q} |x_1 - x_0| \text{ dla } k = 1, 2, \dots \quad (3.16)$$

Metoda iteracji prostej

1. Równanie (3.13) przekształcamy do takiej postaci równoważnej (3.14), która spełnia założenia twierdzenia 3.2.
2. W charakterze przybliżenia zerowego wybieramy dowolny $x_0 \in [a ; b]$, np.

$$x_0 = \frac{a+b}{2}.$$

3. $k := 0$
4. **repeat**
5. $k := k + 1$
6. $x_k := \varphi(x_{k-1})$
7. **until** $\frac{q}{1-q} |x_k - x_{k-1}| < \varepsilon$
8. $\tilde{x} := x_k$

Przykład 3.4. Obliczyć metodą iteracji prostej ujemny pierwiastek funkcji

$$f(x) = \ln(x+2) - 2x^2 + 1$$

z dokładnością $\varepsilon = 10^{-5}$.

Rozwiązanie. W przykładzie 3.1 pokazano, że ujemny pierwiastek funkcji (3.3) leży w przedziale $[-0.8 ; -0.7]$, a dodatni w przedziale $[1.0 ; 1.1]$. Najpierw musimy przekształcić równanie

$$\ln(x+2) - 2x^2 + 1 = 0$$

do równoważnej postaci (3.14), spełniającej założenia twierdzenia 3.2. Równanie możemy przekształcić w następujący sposób:

$$x = e^{2x^2-1} - 2 = \varphi_0(x).$$

Dla tego przedstawienia mamy

$$|\varphi'_0(x)| = \left| 4xe^{2x^2-1} \right| = 4|x|e^{2x^2-1}.$$

Ta funkcja jest:

— w przedziale $[-0.8 ; -0.7]$ monotonicznie malejąca, zatem dla $x \in [-0.8 ; -0.7]$ zachodzi

$$|\varphi'_0(x)| \geq |\varphi'_0(-0.7)| = 2.74456;$$

— w przedziale $[1.0 ; 1.1]$ monotonicznie rosnąca, zatem dla $x \in [1.0 ; 1.1]$ zachodzi

$$|\varphi'_0(x)| \geq |\varphi'_0(1.0)| = 10.87313.$$

W obydwu przedziałach $|\varphi'_0(x)| > 1$, więc utworzony proces (3.15) przez odwzorowanie $x = \varphi_0(x)$ jest rozbieżny.

Równanie $\ln(x+2) - 2x^2 + 1 = 0$ może być przedstawione w równoważnej postaci jako para równań:

$$x = -\sqrt{\frac{\ln(x+2)+1}{2}} = \varphi_1(x), \quad x = \sqrt{\frac{\ln(x+2)+1}{2}} = \varphi_2(x).$$

Dla obydwu funkcji mamy

$$|\varphi'_{1,2}(x)| = \left| \frac{1}{\sqrt{2}} \cdot \frac{1}{2} \cdot \frac{1}{\sqrt{\ln(x+2)+1}} \cdot \frac{1}{x+2} \right| = \frac{\sqrt{2}}{4} \cdot \frac{1}{\sqrt{\ln(x+2)+1}} \cdot \frac{1}{x+2}.$$

Dla $x \geq -1$ funkcja $y = |\varphi'_{1,2}(x)|$ jest monotonicznie malejąca, zatem

$$|\varphi'_{1,2}(x)| \leq |\varphi'_{1,2}(-1.0)| = \frac{\sqrt{2}}{4} = 0.35355 < 1.$$

W związku z tym, w procesie iteracyjnym użyjemy:

— funkcji

$$\varphi_1(x) = -\sqrt{\frac{\ln(x+2)+1}{2}}$$

dla wyznaczenia przybliżonej wartości pierwiastka z przedziału $[-0.8 ; -0.7]$,

— funkcji

$$\varphi_2(x) = \sqrt{\frac{\ln(x+2)+1}{2}}$$

dla wyznaczenia przybliżonej wartości pierwiastka z przedziału $x \in [1.0 ; 1.1]$.

Do przyjęto:

$$\max_{-0.8 \leq x \leq -0.7} |\varphi'_1(x)| = 0.27096 \quad \text{dla pierwiastka z przedziału } [-0.8 ; -0.7],$$

$$\max_{1.0 \leq x \leq 1.1} |\varphi'_2(x)| = 0.08135 \quad \text{dla pierwiastka z przedziału } x \in [1.0 ; 1.1].$$

Wartości kolejnych iteracji metody iteracji prostej są podane w tabeli 3.4.

Tabela 3.4. Wyznaczanie pierwiastków funkcji

$$f(x) = \ln(x+2) - 2x^2 + 1$$

metodą iteracji prostej, z dokładnością 10^{-5} .

$$\begin{aligned} x_0 &= -0.75000 \\ q &= 0.27096 \end{aligned}$$

$$\begin{aligned} x_0 &= 1.05000 \\ q &= 0.08135 \end{aligned}$$

k	x_k	$\frac{q}{1-q} x_k - x_{k-1} $	k	x_k	$\frac{q}{1-q} x_k - x_{k-1} $
0	-0.75000		0	1.05000	
1	-0.78203	8.6E-02	1	1.02838	5.8E-02
2	-0.77369	2.2E-02	2	1.02665	4.7E-03
3	-0.77589	5.9E-03	3	1.02651	3.7E-04
4	-0.77531	1.6E-03	4	1.02650	3.0E-05
5	-0.77546	4.1E-04	5	1.02650	2.4E-06
6	-0.77542	1.1E-04			
7	-0.77543	2.8E-05			
8	-0.77543	7.5E-06			

3.6. Efektywność metod przybliżonego obliczania pierwiastków funkcji

Obecnie naszym celem jest porównanie metod uściślania przybliżonej wartości jednokrotnego pierwiastka funkcji $y = f(x)$. Zakładamy, że metoda iteracyjna generuje ciąg $x_0, x_1, \dots, x_k, \dots$ kolejnych przybliżeń zbieżny do pierwiastka \bar{x} funkcji $y = f(x)$. Każdej metodzie można przyporządkować liczbę wykładnikiem p ($p \geq 1$) zwaną *wykładnikiem zbieżności metody* (lub *rządem metody*), oraz stałą K , zwaną *stałą asymptotyczną błędu metody*, które spełniają warunek

$$\limsup_{k \rightarrow \infty} \frac{|x_{k+1} - \bar{x}|}{|x_k - \bar{x}|^p} = K < \infty. \quad (3.17)$$

Rząd metody p i stała K charakteryzuje szybkość zbieżności metody iteracyjnej: ciąg kolejnych przybliżeń $x_0, x_1, \dots, x_k, \dots$ jest tym szybciej zbieżny do pierwiastka, im większy jest rząd metody i im mniejsza jest stała asymptotyczna błędu. Spośród tych dwóch wielkości istotniejszą rolę odgrywa wykładnik zbieżności p . Przykładowo,

— dla metody iteracji prostej: $p = 1$ oraz $K = |\varphi(\bar{x})|$,

— dla metody stycznych: $p = 1$ oraz $K = \frac{1}{2} \frac{|f''(\bar{x})|}{|f'(\bar{x})|}$.

3.7. Zadania

Zad. 3.1. Liczba iteracji w metodzie bisekcji

Wykazać, że dla wyznaczenia z dokładnością ε przybliżonej wartości pierwiastka ~~zaznaczonego~~ w przedziale $[a ; b]$ trzeba w metodzie bisekcji wykonać

$$\text{Entier}\left(\lg \frac{|b-a|}{\varepsilon}\right)$$

iteracji.

Zad. 3.2. Kryterium zakończenia obliczeń w metodzie Newtona-Raphsona.

Udowodnić, że dla procesu Newtona-Raphsona zachodzi nierówność

$$|\bar{x} - x_k| \leq \frac{M_2}{2m_1} (x_k - x_{k-1})^2,$$

gdzie:

$$m_1 = \min_{a \leq x \leq b} |f'(x)|,$$

$$M_2 = \max_{a \leq x \leq b} |f''(x)|.$$

Nierówność tę można wykorzystać jako kryterium zakończenia obliczeń w metodzie Newtona-Raphsona.

Zad. 3.3. Kryterium zakończenia obliczeń w metodzie iteracji prostej.

Udowodnić, że jeśli proces iteracji prostej jest zbieżny, to zachodzi nierówność

$$|\bar{x} - x_k| \leq \frac{q}{1-q} |x_k - x_{k-1}|.$$

Nierówność tę można wykorzystać jako kryterium zakończenia obliczeń w metodzie iteracji prostej.

Zad. 3.4. Podprogram obliczania pierwiastka .

Zakładamy, że dysponujemy narzędziem obliczeniowym wykonującym 5 działań ~~uni~~miennoprzecinkowych: dodawanie, odejmowanie, mnożenie, dzielenie i potęgowanie. Opracować algorytm obliczania $\sqrt[m]{a}$ dla dowolnej dodatniej liczby rzeczywistej a i dowolnego naturalnego m ($m \geq 2$).

~~Wskazówka.~~ Oznaczyć $x = \sqrt[m]{a}$. Równość ta jest równoważna równości $x^m = a$. Oznacza to, że należy znaleźć pierwiastek funkcji $f(x) = x^m - a$. Dowieść, że do znalezienia pierwiastka tej funkcji można zastosować metodę Newtona-Raphsona.

Zad. 3.5. Stosując metodę bisekcji, obliczyć z dokładnością ε rzeczywiste pierwiastki ~~zaznaczonych~~ funkcji:

$$f(x) = 4 \sin x - x + 1,$$

$$f(x) = e^{-2x} + x - 1,$$

$$f(x) = x \cos x - \ln x,$$

- d) $f(x) = \sin^3 x + \cos^3 x - 3x + 7$,
e) $f(x) = \sqrt[3]{x^2} - \sqrt[3]{x^2 - 1} - 3(x-1)^2$.

Zad. 3.6. Stosując metodę Newtona-Raphsona, obliczyć z dokładnością ε rzeczywiste pierwiastki poniższych funkcji:

- a) $f(x) = x^3 - x + 5$,
b) $f(x) = e^{-x} - x$,
c) $f(x) = \cos x - 2x^2$,
d) $f(x) = \ln x - x^2 + 2x$,
e) $f(x) = \sin^2 x - 2(x-1)^2$.

Zad. 3.7. Stosując metodę iteracji prostej, obliczyć z dokładnością ε rzeczywiste pierwiastki poniższych funkcji:

- a) $f(x) = x^5 + x - 3$,
b) $f(x) = \ln(x-1) + 2x$,
c) $f(x) = x - \sin x - 0.25$,
d) $f(x) = x^6 - \sqrt{x} - 1$,
e) $f(x) = \operatorname{arctg} x + 2x - 4$.

Zad. 3.8. Obliczyć z dokładnością ε rzeczywiste pierwiastki poniższych funkcji:

- a) $f(x) = \sin x - x + 1$,
b) $f(x) = \ln x + x - 2$,
c) $f(x) = x^5 + x - 5$,
d) $f(x) = x^2 \ln x - 1$,
e) $f(x) = e^{-x^2} - x^3 + 1$,
f) $f(x) = \cos x - \sqrt{x}$,
g) $f(x) = \operatorname{arctg} x + x - 2$,
h) $f(x) = x^5 + 3x - 2$,
i) $f(x) = e^x + x + 1$,
j) $f(x) = \sqrt[4]{x} e^x - 1$,
k) $f(x) = \operatorname{arcctg} x + 5x + 1$,
l) $f(x) = x \lg(x-1) - 5$,
m) $f(x) = \sin hx - \sqrt{x+2}$,
n) $f(x) = 2^x - x^2 + 4x - 4$,
o) $f(x) = x^2 \sqrt{x} - 2$,
p) $f(x) = e^{2x} - 5x - 3$,

- $f(x) = \ln x - x^2 + 2,$
 $f(x) = \cos x - 3x^2 + 3x,$
 $f(x) = 6 \sin x - x^2 - 2,$
 $f(x) = \arctg x - 2x^2 + 2.$

Zad. 3.9. Obliczyć z dokładnością ε rzeczywiste pierwiastki poniższych funkcji:

- $f(x) = 2 \operatorname{arcctg} x - x^2,$
 $f(x) = x^6 - 4x + 2,$
 $f(x) = x^3 - 5\sqrt[3]{x^2} + 1,$
 $f(x) = 2^x - 3x,$
 $f(x) = 2x^3 - 2 \ln(x+1) - 1,$
 $f(x) = \cosh x - 2\sqrt[3]{x+1},$
 $f(x) = 4x^6 + 4x^2 - 8x - 5,$
 $f(x) = 5 \cos(x-1) - 4x^6,$
 $f(x) = \log x - x^4 + 1,$
 $f(x) = 5x \log(x+2) - 1,$
 $f(x) = x^4 - 4x^3 + 2x^2 - 8,$
 $f(x) = (x+1)e^{x-1} - 1,$
 $f(x) = 2^x + 2x^2 - 3,$
 $f(x) = \ln(x+3) - \cosh x + 1,$
 $f(x) = \sinh x - x^4 + 2,$
 $f(x) = 2e^x(x^3 - 3x^2 - x + 3) - 1,$
 $f(x) = 5x^3 - 10x^2 + \ln(x+2),$
 $f(x) = \sin^3 x + \cos^3 x - x^2 + 1,$
 $f(x) = x^2 \ln(x+3) - 1,$
 $f(x) = \cos 6x - x^2.$

Rozwiązywanie układów równań nieliniowych

4.1. Wstęp

W niniejszym rozdziale omówimy wybrane metody rozwiązywania układów równań

$$\begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0, \\ f_2(x_1, x_2, \dots, x_n) &= 0, \\ \dots & \\ f_n(x_1, x_2, \dots, x_n) &= 0, \end{aligned} \tag{4.1}$$

w których $f_i(x_1, x_2, \dots, x_n)$ ($i=1, 2, \dots, n$) są funkcjami rzeczywistymi, zmiennych rzeczywistych x_1, x_2, \dots, x_n i nie wszystkie są liniowe. Oznaczając

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \quad O = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \quad F = \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{bmatrix}, \quad J(X) = \begin{bmatrix} \frac{\partial f_1(X)}{\partial x_1} & \frac{\partial f_1(X)}{\partial x_2} & \dots & \frac{\partial f_1(X)}{\partial x_n} \\ \frac{\partial f_2(X)}{\partial x_1} & \frac{\partial f_2(X)}{\partial x_2} & \dots & \frac{\partial f_2(X)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(X)}{\partial x_1} & \frac{\partial f_n(X)}{\partial x_2} & \dots & \frac{\partial f_n(X)}{\partial x_n} \end{bmatrix},$$

można układ (4.1) zapisać w zwięzzej postaci

$$F(X) = O. \tag{4.2}$$

Rozwiązaniem układu równań (4.1) nazywamy każdy wektor

$$\bar{X} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \dots \\ \bar{x}_n \end{bmatrix},$$

spełniający równość $F(\bar{X}) = O$.

Metodologia numerycznego rozwiązywania układów równań nieliniowych jest uogólnieniem na przypadek przestrzeni euklidesowej n -wymiarowej metodologii numerycznego obliczania pierwiastków funkcji jednej zmiennej. Zatem wyznaczenie przyblizonego rozwiązania układu równań (4.1) jest realizowane w dwóch etapach:

- 1) najpierw *lokalizujemy rozwiązania układu* (4.1), tj.

- a) znajdujemy liczbę p rozwiązań układu równań,
 b) dla każdego rozwiązania \bar{X}_i znajdujemy taki obszar domknięty $D_i \subset R^n$, że $X_i \in \text{Int}(D_i)$ oraz $D_i \cap D_j = \emptyset$ dla $i, j = 1, 2, \dots, p; i \neq j$,
 c) następnie uściślamy przybliżone rozwiązania, tj. dla każdego rozwiązania \bar{X} układu i zadanej dokładności ε znajdujemy taki wektor

$$\tilde{X} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{bmatrix}, \text{że } \|\bar{X} - \tilde{X}\| < \varepsilon.$$

4.2 Lokalizacja rozwiązań

Niestety, nie opracowano dotychczas algorytmów lokalizacji rozwiązań układu równań ogólnej postaci. Na tym etapie wykorzystujemy wiedzę z analizy matematycznej, geometrii, algebry.

Przykład 4.1. Zlokalizować rozwiązania układu równań

$$\begin{aligned} \arctg x_1 - x_2^3 &= 0, \\ 4x_1^2 + 9x_2^2 &= 36. \end{aligned} \tag{4.3}$$

Rozwiązanie. Pierwsze z równań układu (4.3) przekształcamy do postaci funkcji zmiennej zależnej x_1 :

$$x_2 = \sqrt[3]{\arctg x_1},$$

$$y = f_a(x_1) = \sqrt[3]{\arctg x_1}.$$

Drugi równanie układu (4.3) generuje dwie funkcje zmiennej niezależnej x_1 :

$$y = f_d(x_1) = -\frac{1}{3}\sqrt{36 - 4x_1^2},$$

$$y = f_g(x_1) = \frac{1}{3}\sqrt{36 - 4x_1^2}.$$

W tabeli 4.1 zostały stabilowane funkcje $y = f_a(x)$, $y = f_d(x)$, $y = f_g(x)$, a ich wykresy są pokazane na rys. 4.1.

Z tabeli 4.1 oraz z rys. 4.1 wnioskujemy, że istnieją dwa rozwiązania układu równań (4.3), które leżą w przedziałach:

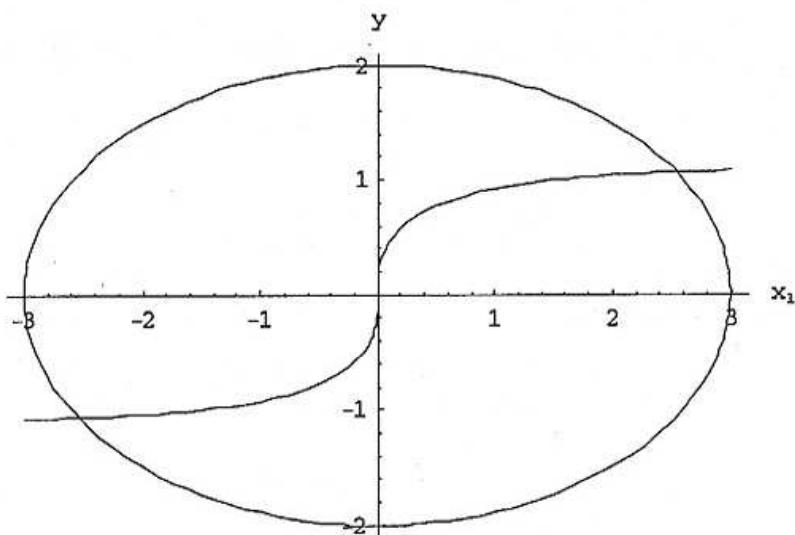
- rozwiązanie 1: $-3.0 < \bar{x}_{11} < -2.5, \quad -1.5 < \bar{x}_{21} < -1.0,$
- rozwiązanie 2: $2.5 < \bar{x}_{12} < 3.0, \quad 1.0 < \bar{x}_{22} < 1.5.$

Umawiamy, że składowe rozwiązań spełniają zależności:

$$\begin{aligned} \bar{x}_{12} &= -\bar{x}_{11}, \\ \bar{x}_{22} &= -\bar{x}_{21}. \end{aligned} \tag{4.4}$$

Tabela 4.1. Lokalizacja rozwiązań układu (4.3)

x	$y = f_a(x)$	$y = f_d(x)$	$y = f_g(x)$
-3.0	-1.07694	0.00000	0.00000
-2.8	-1.07079	-0.71802	0.71802
-2.6	-1.06373	-0.99778	0.99778
-2.4	-1.05553	-1.20000	1.20000
-2.2	-1.04592	-1.35974	1.35974
-2.0	-1.03451	-1.49071	1.49071
-1.8	-1.02080	-1.60000	1.60000
...
2.2	1.04592	-1.35974	1.35974
2.4	1.05553	-1.20000	1.20000
2.6	1.06373	-0.99778	0.99778
2.8	1.07079	-0.71802	0.71802
3.0	1.07694	0.00000	0.00000



Rys. 4.1. Krzywe układu równań (4.3)

4.3. Metoda najszybszego spadku

Metoda najszybszego spadku jest metodą wyznaczania minimum lokalnego funkcjonału

$$U(x_1, x_2, \dots, x_n) = \sum_{i=1}^n f_i^2(x_1, x_2, \dots, x_n) \quad (4.5)$$

(x_1, x_2, \dots, x_n) oznacza lewą stronę i -tego równania układu (4.1)). Związek tej metody z rozwiązywaniem układu równań (4.1) jest sformułowany w twierdzeniu 4.1.

Twierdzenie 4.1. Wektor $X \in R^n$ jest rozwiązaniem układu równań (4.1) wtedy i tylko wtedy, gdy $U(X) = 0$.

Biorąc pod uwagę, że każdego $U(X) \geq 0$ dla każdego $X \in R^n$, problem znalezienia rozwiązania układu równań (4.1) jest równoważny znalezieniu wektorów $X \in R^n$, w których funkcjonał (4.5) osiąga minimum globalne – równe zeru.

Metoda najszybszego spadku jest iteracyjną metodą gradientową wyznaczania minimum lokalnego funkcjonału (4.5).

Algorytm 4.1. Rozwiązywanie układu równań nieliniowych $F(X) = O$ metodą najszybszego spadku.

Metoda najszybszego spadku ($X^{(0)}, \varepsilon$)

```

1      W charakterze  $X^{(0)}$  wybieramy dowolny wektor z przestrzeni  $R^n$ 
2       $k := 0$ 
3      repeat
4           $\mu_k = \frac{F^T(X^{(k)})J(X^{(k)})J^T(X^{(k)})F(X^{(k)})}{(J(X^{(k)})J^T(X^{(k)})F(X^{(k)}))^T(J(X^{(k)})J^T(X^{(k)})F(X^{(k)}))}$ 
5           $X^{(k+1)} = X^{(k)} - \mu_k J^T(X^{(k)})F(X^{(k)})$ 
6      until  $\|X^{(k+1)} - X^{(k)}\| < \varepsilon$ 
7       $\tilde{X} := X^{(k+1)}$ 

```

Warto podkreślić kilka własności metody najszybszego spadku.

Kryterium $\|X^{(k+1)} - X^{(k)}\| < \varepsilon$ – zakończenia procesu iteracyjnego – jest heurystyczne i nie gwarantuje spełnienia nierówności $\|\bar{X} - X^{(k+1)}\| < \varepsilon$.

Ciąg wektorów $X^{(0)}, X^{(1)}, \dots, X^{(k)}, \dots$ wyznaczony metodą najszybszego spadku jest zbiegły do wektora, w którym funkcjonał (4.5) przyjmuje minimum lokalne. Od wyboru wektora $X^{(0)}$ zależy, czy będzie to minimum lokalne czy minimum globalne funkcjonału.

Proces iteracyjny generowany metodą najszybszego spadku jest na ogół wolno zbieżny.

Przykład 4.2. Uściślić metodą najszybszego spadku rozwiązanie układu równań

$$\operatorname{arctg} x_1 - x_2^3 = 0,$$

$$4x_1^2 + 9x_2^2 = 36,$$

leżące w prostokącie $D_2 = \{(x_1, x_2); 2.5 < x_1 < 3.0, 1.0 < x_2 < 1.5\}$. W charakterze $X^{(0)}$ przyjąć środek prostokąta D_2 , tj.

$$D_2, \text{ tj. } X^{(0)} = \begin{bmatrix} 2.75 \\ 1.25 \end{bmatrix}.$$

Rozwiązanie. Początkowe iteracje wygenerowane metodą najszybszego spadku zostały zestawione w tabeli 4.2.

Tabela 4.2. Początkowe iteracje metody najszybszego spadku dla układu równań (4.3)

k	$X^{(k)}$	$F(X^{(k)})$	$J(X^{(k)})$		$J^T(X^{(k)})F(X^{(k)})$	μ_k	$\ X^{(k)} - X^{(k-1)}\ _1$
0	2.750000	-0.731100	0.116788	-4.687500	182.789616	0.000999	
	1.250000	8.312500	22.000000	22.500000	190.458280		
1	2.567404	0.009224	0.131725	-3.369170	9.732014	0.001265	3.7E-01
	1.059744	0.473766	20.539233	19.075386	9.006201		
2	2.555093	0.045568	0.132829	-3.297120	0.115817	0.007464	2.4E-02
	1.048351	0.005370	20.440748	18.870319	-0.048910		
3	2.554229	0.044249	0.132907	-3.299417	-0.104611	0.001537	1.2E-03
	1.048716	-0.005407	20.433832	18.876890	-0.248067		
4	2.554390	0.043012	0.132892	-3.301816	0.109433	0.007467	5.4E-04
	1.049097	0.005075	20.435118	18.883752	-0.046174		
5	2.553573	0.041764	0.132966	-3.303986	-0.098804	0.001536	1.2E-03
	1.049442	-0.005108	20.428581	18.889958	-0.234486		
6	2.553724	0.040594	0.132952	-3.306255	0.103382	0.007470	5.1E-04
	1.049802	0.004796	20.429795	18.896441	-0.043584		
7	2.552952	0.039415	0.133022	-3.308306	-0.093305	0.001535	1.1E-03
	1.050128	-0.004825	20.423617	18.902301	-0.221603		
8	2.553095	0.038308	0.133009	-3.310449	0.097648	0.007472	4.8E-04
	1.050468	0.004531	20.424763	18.908425	-0.041135		
9	2.552366	0.037193	0.133075	-3.312387	-0.088099	0.001534	1.0E-03
	1.050775	-0.004557	20.418926	18.913957	-0.209388		
10	2.552501	0.036147	0.133063	-3.314413	0.092218	0.007475	4.6E-04
	1.051097	0.004281	20.420007	18.919740	-0.038818		
11	2.551812	0.035093	0.133125	-3.316243	-0.083171	0.001534	9.8E-04
	1.051387	-0.004303	20.414492	18.924963	-0.197809		
12	2.551939	0.034103	0.133113	-3.318157	0.087076	0.007477	4.3E-04
	1.051690	0.004043	20.415513	18.930423	-0.036628		

Przykład 4.3. Uściślić metodą najszybszego spadku rozwiązanie układu równań (4.3). W charakterze $X^{(0)}$ przyjąć wektor

$$X^{(0)} = \begin{bmatrix} -3.00 \\ -3.00 \end{bmatrix}.$$

Rozwiązanie. Wybrane iteracje metody zostały zestawione w tabeli 4.3.

Tabela 4.3. Uściślanie metodą najszybszego spadku rozwiązania układu równań (4.3)

$X^{(k)}$	$F(X^{(k)})$	$J(X^{(k)})$		$J^T(X^{(k)})F(X^{(k)})$	μ_k	$\ X^{(k)} - X^{(k-1)}\ _1$
-3.000000	25.750954	0.100000	-27.000000	-1941.424905	0.000243	
-3.000000	81.000000	-24.000000	-54.000000	-5069.275764		
-2.528362	4.336983	0.135270	-9.382777	-357.809265	0.000673	1.7E+00
-1.768500	17.718783	-20.226893	-31.833001	-604.734969		
-2.287434	1.364056	0.160453	-5.559474	-29.203781	0.001053	6.5E-01
-1.361307	1.607841	-18.299473	-24.503534	-46.981223		
-2.256694	1.103975	0.164132	-5.162881	2.721219	0.061886	8.0E-02
-1.311854	-0.140694	-18.053548	-23.613368	-2.377439		
-2.425098	0.400358	0.145326	-4.069746	5.225616	0.001256	3.2E-01
-1.164724	-0.266352	-19.400788	-20.965032	3.954716		
-2.431661	0.419705	0.144656	-4.104527	0.732296	0.024347	1.2E-02
-1.169690	-0.034523	-19.453286	-21.054427	-0.995832		
-2.449490	0.319675	0.142857	-3.936134	3.800202	0.001276	4.2E-02
-1.145445	-0.191598	-19.595918	-20.618013	2.692080		
-2.454340	0.332548	0.142373	-3.959782	0.548406	0.022487	8.3E-03
-1.148881	-0.025519	-19.634718	-20.679856	-0.789086		
-2.466672	0.261616	0.141154	-3.838411	2.935537	0.001291	3.0E-02
-1.131137	-0.146889	-19.733374	-20.360461	1.986530		
-2.470462	0.270950	0.140782	-3.855840	0.428319	0.021346	6.4E-03
-1.133702	-0.019742	-19.763699	-20.406634	-0.641874		
-2.479605	0.217472	0.139891	-3.763202	2.343841	0.001303	2.3E-02
-1.120000	-0.116622	-19.836844	-20.160004	1.532716		
...
-2.539902	0.009294	0.134208	-3.397158	0.084988	0.001360	7.0E-04
-1.064136	-0.004121	-20.319214	-19.154453	0.047366		
-2.540017	0.009498	0.134198	-3.397569	0.012226	0.017844	1.8E-04
-1.064201	-0.000539	-20.320139	-19.155613	-0.021945		
-2.540236	0.008138	0.134178	-3.395070	0.074356	0.001360	6.1E-04
-1.063809	-0.003605	-20.321884	-19.148564	0.041404		
-2.540337	0.008316	0.134169	-3.395429	0.010694	0.017832	1.6E-04
-1.063865	-0.000471	-20.322693	-19.149578	-0.019211		
...
-2.542585	0.000002	0.133963	-3.380349	0.000016	0.001363	1.3E-07
-1.061500	-0.000001	-20.340680	-19.107006	0.000009		
-2.542585	0.000002	0.133963	-3.380349	0.000002	0.017752	3.5E-08
-1.061500	0.000000	-20.340680	-19.107006	-0.000004		
-2.542585	0.000002	0.133963	-3.380348	0.000014	0.001363	1.2E-07
-1.061500	-0.000001	-20.340680	-19.107005	0.000008		

4.4. Metoda Newtona-Raphsona

Metoda Newtona-Raphsona jest uogólnieniem metody stycznych poszukiwania przybliżonej wartości pierwiastka funkcji jednej zmiennej. W tej metodzie kolejne przybliżenia obliczamy ze wzoru

$$X^{(k+1)} = X^{(k)} - J^{-1}(X^{(k)})F(X^{(k)}) \quad \text{dla } k = 0, 1, \dots \quad (4.6)$$

W poniższym twierdzeniu zostały sformułowane warunki wystarczające zbieżności ciągu (4.6) do rozwiązywania układu równań (4.1).

Twierdzenie 4.2 (Kantorowicz).

Niech funkcje $y = f_i(x_1, x_2, \dots, x_n)$ ($i = 1, 2, \dots, n$) będą określone i ciągle wraz ze swymi pochodnymi cząstkowymi rzędu pierwszego i drugiego w pewnym otoczeniu D rozwiązania \bar{X} układu równań (4.1). Założmy, że $X^{(0)} \in Int(D)$ oraz zachodzą warunki:

$$1) \text{ istnieje } J^{-1}(X^{(0)}) \text{ oraz } \|J^{-1}(X^{(0)})\| \leq A_0, \quad (4.7)$$

$$2) \|J^{-1}(X^{(0)})F(X^{(0)})\| \leq B_0, \quad (4.8)$$

3) istnieje taka stała rzeczywista, nieujemna C , że dla każdego $X \in D$ zachodzi

$$\sum_{k=1}^n \left| \frac{\partial^2 f_i(X)}{\partial x_j \partial x_k} \right| \leq C \quad \text{dla } i, j = 1, 2, \dots, n, \quad (4.9)$$

4) stałe A_0, B_0, C spełniają warunek

$$\mu_0 = 2 \cdot n \cdot A_0 B_0 C < 1. \quad (4.10)$$

Wówczas:

1) proces Newtona-Raphsona (4.6) startujący z przybliżenia $X^{(0)}$ jest zbieżny oraz $\lim_{k \rightarrow \infty} X^{(k)} = \bar{X} \in D$,

2) zachodzą nierówności

$$\|X - X^{(k)}\| \leq 2B_k \quad \text{dla } k = 0, 1, \dots, \quad (4.11)$$

3) dla kolejnych przybliżeń $X^{(k)}$ spełniona jest nierówność

$$\|\bar{X} - X^{(k)}\| \leq \frac{\mu_0^{2^k-1} B_0}{2^{k-1}} \quad \text{dla } k = 0, 1, \dots. \quad (4.12)$$

Algorytm 4.2. Rozwiązywanie układu równań nieliniowych $F(X) = O$ metodą Newtona-Raphsona.

Metoda Newtona-Raphsona($X^{(0)}, \varepsilon$)

- 1 Znaleźć $X^{(0)}$ spełniające warunki (4.7)-(4.10) twierdzenia Kantorowicza.
- 2 $k := -1$
- 3 repeat
- 4 $k := k + 1$
- 5 $X^{(k+1)} := X^{(k)} - J^{-1}(X^{(k)})F(X^{(k)})$
- 6 until $\|X^{(k+1)} - X^{(k)}\| < \frac{\varepsilon}{2}$
- 7 $\tilde{X} := X^{(k+1)}$

zamiast obliczania kolejnego przybliżenia $X^{(k+1)}$ ze wzoru (4.6), wektor ten wyznacza się przez rozwiązywanie układu równań liniowych

$$J(X^{(k)})X^{(k+1)} = J(X^{(k)})X^{(k)} - F(X^{(k)}).$$

Przykład 4.4. Uściślić metodą Newtona-Raphsona rozwiązanie układu równań (4.3) leżące w prostokącie $D_2 = \{(x_1 ; x_2); 2.5 < x_1 < 3.0, 1.0 < x_2 < 1.5\}$.

Rozwiązanie. W charakterze $X^{(0)}$ przyjmiemy środek prostokąta D_2 , tj.

$$X^{(0)} = \begin{bmatrix} 2.75 \\ 1.25 \end{bmatrix}. \quad (4.13)$$

Na początku sprawdzamy, czy dla tego przybliżenia są spełnione założenia twierdzenia Kantorowicza. Dla układu (4.3) mamy

$$F(X) = \begin{bmatrix} \arctgx_1 - x_2^3 \\ 4x_1^2 + 9x_2^2 - 36 \end{bmatrix}, \quad J(X) = \begin{bmatrix} 1 & -3x_2^2 \\ 1+x_1^2 & 18x_2 \\ 8x_1 & \end{bmatrix}.$$

Zatem

$$A_0 = \|J^{-1}(X^{(0)})\|_1 = \left\| \begin{bmatrix} 0.21276 & 0.04433 \\ -0.20803 & 0.00110 \end{bmatrix} \right\|_1 = 0.42079,$$

$$B_0 = \|J^{-1}(X^{(0)} F(X^{(0)})\|_1 = \left\| \begin{bmatrix} 0.21276 & 0.04433 \\ -0.20803 & 0.00110 \end{bmatrix} \begin{bmatrix} -0.73110 \\ 8.31250 \end{bmatrix} \right\|_1 = 0.21290,$$

$$\max_{\substack{2.5 \leq x_1 \leq 3.0 \\ 1.0 \leq x_2 \leq 1.5}} \left\{ \left| \frac{\partial^2 f_1(X)}{\partial x_1^2} \right| + \left| \frac{\partial^2 f_1(X)}{\partial x_1 \partial x_2} \right| \right\} = \max_{\substack{2.5 \leq x_1 \leq 3.0 \\ 1.0 \leq x_2 \leq 1.5}} \left\{ -\frac{2x_1}{(1+x_1^2)^2} + |0| \right\} = \left| \frac{2 \cdot 2.5}{(1+2.5^2)^2} \right| \equiv 0.095,$$

$$\max_{\substack{2.5 \leq x_1 \leq 3.0 \\ 1.0 \leq x_2 \leq 1.5}} \left\{ \left| \frac{\partial^2 f_1(X)}{\partial x_2 \partial x_1} \right| + \left| \frac{\partial^2 f_1(X)}{\partial x_2^2} \right| \right\} = \max_{\substack{2.5 \leq x_1 \leq 3.0 \\ 1.0 \leq x_2 \leq 1.5}} \{0| + |-6x_2| \} = |6 \cdot 1.5| = 9.000,$$

$$\max_{\substack{2.5 \leq x_1 \leq 3.0 \\ 1.0 \leq x_2 \leq 1.5}} \left\{ \left| \frac{\partial^2 f_2(X)}{\partial x_1^2} \right| + \left| \frac{\partial^2 f_2(X)}{\partial x_1 \partial x_2} \right| \right\} = \max_{\substack{2.5 \leq x_1 \leq 3.0 \\ 1.0 \leq x_2 \leq 1.5}} \{8| + |0| \} = 8.000,$$

$$\max_{\substack{2.5 \leq x_1 \leq 3.0 \\ 1.0 \leq x_2 \leq 1.5}} \left\{ \left| \frac{\partial^2 f_2(X)}{\partial x_2 \partial x_1} \right| + \left| \frac{\partial^2 f_2(X)}{\partial x_2^2} \right| \right\} = \max_{\substack{2.5 \leq x_1 \leq 3.0 \\ 1.0 \leq x_2 \leq 1.5}} \{0| + |18| \} = 18.000.$$

Zatem

$$C = \max\{0.095 ; 9.000 ; 8.000 ; 18.000\} = 18.000$$

$$\mu_0 = 2 \cdot n \cdot A_0 B_0 C = 2 \cdot 2 \cdot 0.42079 \cdot 0.21290 \cdot 18.000 = 6.45035.$$

Zatem nie są spełnione warunki wystarczające twierdzenia Kantorowicza ($\mu_0 > 1$). Nie oznacza to, że proces Newtona-Raphsona nie jest zbieżny. W tabeli 4.4 zostały podane

wartości kolejnych iteracji procesu Newtona-Raphsona startującego z przybliżenia zerowego (4.13). Z tabeli widać, że proces jest zbieżny do rozwiązania leżącego w prostokącie D_2 .

Tabela 4.4. Iteracje metody Newtona-Raphsona dla układu równań (4.3)

k	$X^{(k)}$	$F(X^{(k)})$	$J(X^{(k)})$		$J^T(X^{(k)})F(X^{(k)})$	$2\ X^{(k)} - X^{(k-1)}\ _1$
0	2.75000000	-0.731100	0.116788	-4.687500	0.212903	
	1.25000000	8.312500	22.000000	22.500000	0.161272	
1	2.53709675	-0.095157	0.134465	-3.555983	-0.005182	7.48E-01
	1.08872762	0.415390	20.296774	19.597097	0.026564	
2	2.54227896	-0.002287	0.133991	-3.384576	-0.000306	6.35E-02
	1.06216387	0.006458	20.338232	19.118950	0.000664	
3	2.54258529	-0.000001	0.133963	-3.380348	0.000000	1.94E-03
	1.06150022	0.000004	20.340682	19.107004	0.000000	
4	2.54258546	0.000000	0.133963	-3.380346	0.000000	1.14E-06
	1.06149982	0.000000	20.340684	19.106997	0.000000	

Czytelnik zechce sprawdzić, że dla przybliżenia zerowego

$$X^{(0)} = \begin{bmatrix} 2.55 \\ 1.05 \end{bmatrix}$$

współczynnik $\mu_0 = 0.46997$ i już w trzeciej iteracji otrzymujemy rozwiązanie układu równań (4.3) z dokładnością $\varepsilon = 1.5 \cdot 10^{-8}$.

4.5. Metoda symulowanego wyżarzania

Metoda symulowanego wyżarzania (ang. *simulated annealing method*) jest metodą losowego poszukiwania minimum globalnego funkcji.

Wyżarzanie jest procesem technologicznym stosowanym głównie w produkcji stali i szkła. Ma ono na celu usunięcie naprężeń własnych w materiale, nadanie materiałowi struktury drobnoziarnistej i polepszenie jego plastyczności. Wyżarzanie polega na nagrzaniu materiału do określonej temperatury, wygrzaniu go w tej temperaturze przez czas dostatecznie długi dla wyrównania temperatury w całej masie materiału, a następnie powolnym chłodzeniu. Jeżeli chłodzenie przebiega wystarczająco wolno, to zostaną zredukowane naprężenia własne materiału, jego struktura zostaje przekrystalizowana i zwiększą się jego właściwości plastyczne.

Algorytm symulowanego wyżarzania jest matematycznym opisem ostatniego etapu wyżarzania – etapu powolnego, kontrolowanego chłodzenia materiału. Metoda symuluje stany ciała $X^{(0)}, X^{(1)}, \dots$ podczas jego chłodzenia, w sposób kontrolowany. Sterowanie procesem chłodzenia polega na regulowaniu szybkości obniżania temperatury i jest w algorytmie opisane parametrem t_k . W każdym stanie $X^{(k)}$ ciało ma energię $E(X^{(k)})$ naprężen wewnętrznych. Algorytm dobrego wyżarzania powinien generować ciąg $X^{(0)}, X^{(1)}, \dots$ zbieżny do stanu \bar{X} , w którym funkcja $E(X)$ osiąga minimum (globalne).

Algorytm 4.3. Rozwiązywanie układu równań nieliniowych $F(X) = O$ metodą symulowanego wyżarzania.

Metoda symulowanego wyżarzania(liczba-losowa)

```

1       $k := 0$ 
2      Wylosuj  $X^{(b)}$ 
3      repeat
4          Ustal parametr kontrolny  $t_k$ 
5          Ustal liczbę powtórzeń  $L_k$ 
6          for  $p := 1$  to  $L_k$ 
7              do begin { $p$ }
8                  Wylosuj  $X^{(n)}$ 
9                  if  $E(X^{(n)}) \leq E(X^{(b)})$ 
10                     then  $X^{(b)} := X^{(n)}$ 
11                      $\frac{E(X^{(b)}) - E(X^{(n)})}{t_k} > \text{liczba-losowa} \in [0;1]$ 
12                     then  $X^{(b)} := X^{(n)}$ 
13                 end { $p$ }
14              $k := k + 1$ 
15         until Kryterium zakończenia obliczeń

```

Odpowiednikiem ilości energii chłodzonego ciała w stanie $X^{(k)}$ może być wartość funkcji (4.5), tj.

$$E(X^{(k)}) = U(X^{(k)}) = \sum_{i=1}^n f_i^2(X^{(k)}).$$

Typową cechą algorytmu symulowanego wyżarzania jest to, że oprócz akceptacji rozwiązań o lepszym (mniejszym) koszcie również akceptuje rozwiązania o gorszym koszcie w ograniczonym zakresie. Dowodzi się, że metodą symulowanego wyżarzania można znaleźć minimum globalne, jeśli temperatura jest obniżana nie szybciej niż $t_0/\ln k$.

4.5. Zadania

Zad. 4.1. Rozwiązać z dokładnością ε poniższe układy dwóch równań z dwiema niewiadomymi x, y :

$$e^x - 2y = 0,$$

$$x - y^2 + 1 = 0,$$

$$\sin x - y = 0,$$

$$x - y^2 - 1 = 0,$$

c) $x^3y - 3 = 0,$
 $\ln(x+2) - y = 0.$

d) $4x^2 + 9y^2 - 1 = 0,$
 $x^3 - y = 0,$

e) $2x^4 - x^4y + 1 = 0,$
 $x^3 - y^2 = 0.$

Zad. 4.2. Rozwiązać z dokładnością ε poniższe układy dwóch równań z dwiema niewiadomymi x, y :

a) $4x^2 + 10y^2 - 40 = 0,$
 $\ln(x-1) + y = 0,$

b) $\sin(x+1) - y - 1.5 = 0,$
 $2x + \cos y - 2 = 0,$

c) $16x^2 - 9y^2 - 144 = 0,$
 $e^{x-5} - 10y = 0,$

d) $\cos(x-2) + y + 1 = 0,$
 $2x + \sin(y+2) - 4 = 0,$

e) $-x^2 + 4y^2 - 4 = 0,$
 $\cosh x - y - 5 = 0,$

f) $10x^2 + y^2 - 6y - 1 = 0,$
 $\sin x - y + 2 = 0,$

g) $4x^2 - y^2 - 4 = 0,$
 $\operatorname{arctg} x - y^3 = 0,$

h) $\cos x - y + 1 = 0,$
 $x^2 + 2y - 4.2 = 0,$

i) $x^3 + y^3 - 4 = 0,$
 $x^4 + y^2 - 3 = 0,$

j) $\cosh x - y^5 = 0,$
 $25x^2 - 100x + 4y^2 = 0,$

$$x^2 + 5y^2 - 40 = 0,$$

$$\sinh x - y^5 = 0,$$

$$\sin(3x) - y = 0,$$

$$\cos(x + \frac{\pi}{2}) - 5y + 5 = 0,$$

$$x - \cosh y = 0,$$

$$10x - 3y^2 - 11 = 0,$$

$$x^3 y - 1 = 0,$$

$$(x-1)^3 - y + 1 = 0,$$

$$\log_5 x - \sqrt[3]{y} = 0,$$

$$x + y^4 - 4 = 0,$$

$$e^{x+y} + y = 0,$$

$$e^{x-y} + x = 0,$$

$$x + 3 \ln x - y^2 = 0,$$

$$2x^2 - 5x - xy + 1 = 0,$$

$$x^2 + y^2 - 16 = 0,$$

$$2 \cosh \frac{x}{2} - y = 0,$$

$$x^2 - 5x + 6 - y(x^2 + 1) = 0,$$

$$e^{-x} - y^3 = 0$$

$$x - y^2 + 1 = 0,$$

$$x^5(y^3 + y^2 - 2y) - 3 = 0.$$

Zad. 4.3. Rozwiązać z dokładnością ε poniższe układy dwóch równań z dwiema niewiadomymi x, y .

$$x^{\frac{2}{3}} + y^{\frac{2}{3}} - 3^{\frac{2}{3}} = 0,$$

$$3 \sinh x - y = 0,$$

$$x^2 y - 4xy + 3y - 10 = 0,$$

$$xy^2 - 15xy + 50x - 5 = 0,$$

c) $x^3 + y^3 - 6xy = 0,$
 $\cos x - y^2 + 1 = 0,$

d) $\operatorname{tg}(xy + 0.3) - x^2 = 0,$
 $\frac{1}{2}x^2 + 2y^2 - 1 = 0,$

e) $e^{x+2y} + 3y = 0,$
 $x - \cosh(y + 2) - 1 = 0,$

f) $x^{2/3} - (x^2 - 1)^{1/3} - y + 1 = 0,$
 $e^{-x} - y^2 = 0,$

g) $5x^2 + 30x - 2y^2 + 4y + 63 = 0,$
 $\ln(x + 5) - y^3 + 20 = 0,$

h) $x^3 - y^2(x - 1) = 0,$
 $\sin^5 x + \cos^5 x - y + 10 = 0,$

i) $(x + 2)^{2/3}(x - 1) - y^{1/3} = 0,$
 $\arctg x - y^3 = 0,$

j) $x^{2/3} - y^{2/3} - 3 = 0,$
 $\ln(x - 1) - y^{1/3} = 0,$

k) $(x^2 + y^2)^2 + 32(y^2 - x^2) = 0,$
 $x - \sin^3 y - \cos^3 y + 2 = 0,$

l) $\sin x + 2\sin y - 1 = 0,$
 $2\sin(3x) + 3\sin(3y) - \frac{3}{10} = 0,$

m) $x - \sin x \cosh y = 0,$
 $y - \cos x \sinh y = 0,$

n) $x^2y^4 - y^4 + 16 = 0,$
 $x^3y^2 - 8x^3y + 15x^3 - 1 = 0,$

o) $4 + x + y - x^2 + 2xy + 3y^2 = 0,$
 $1 + 2x - 3y + x^2 + xy - 2y^2 = 0,$

$$\arctg \frac{x}{y} - \frac{5}{2} = 0,$$

$$x + y - \left(\sin x + \sin y + \frac{3}{4} \right) = 0,$$

$$(2x^2 - y^2) + 16(x + y) - 48 = 0,$$

$$x^2 + y^2 - 16 = 0,$$

$$\sin(x + y) - y^2 - 1.4 = 0,$$

$$\cosh(x^2 - y^2) + 3x^2 y + 2.6 = 0,$$

$$\sin(y^3 - x^2) - x = 0,$$

$$(x^2 - 0.25)y^5 - 5 = 0,$$

$$e^{-2.1x+1} - xy + y = 0,$$

$$x^2 + y^2 - 50 = 0.$$

4.4. Rozwiązać z dokładnością ε następujące układy równań:

$$x^2 + 25y^2 + 4z^2 - 100 = 0,$$

$$x - \sin \frac{y}{10} - \sin \frac{z}{5} - 5 = 0,$$

$$e^{-x} + e^{-y} - 2z = 0,$$

$$\operatorname{arctan} x + \operatorname{arctan} y - z = 0,$$

$$25x^2 + 16y^2 + 16z^2 - 400 = 0,$$

$$x - 2^{-y} - 3^{-z} + 2 = 0,$$

$$x^2 + y^4 - 1 = 0,$$

$$x^2 - y + 8z^6 = 0,$$

$$x^2 + y^2 - 10z = 0,$$

$$\sin 2x + \cos 3y - z + 2 = 0,$$

$$\operatorname{arctan} x - 2y + \operatorname{arctan} z + 10 = 0,$$

$$x^2 - 10y + z^2 = 0,$$

$$3^{-x} + 2^y - z + 2 = 0,$$

$$25x^2 + 3^{-y} - 3z + 1 = 0,$$

$$\cos^2(x + z) - y + 3 = 0,$$

- $x(y+z)-2=0,$
- f) $x-y^2-z^2=0,$
 $e^y - e^{-y} - 2x = 0,$
 $2x^2 + y^2 - 10z - 30 = 0,$
- g) $\log x + \log y - z + 2 = 0,$
 $\sqrt[3]{x} - y = 0,$
 $16x^2 + 36y^2 - 9z^2 + 144 = 0,$
- h) $x - \ln(y+z+7) - 2 = 0,$
 $10x + 10y + 3z + 30 = 0,$
 $(x-1)^2 - (y-2)^2 - 10z = 0,$
- i) $\sin(x+y) + \cos(x+y) + z + 4 = 0,$
 $x - e^{-y^2-z^2} - 2 = 0,$
 $(x+1)^2 - (y+1)^2 + 12z = 0,$
- j) $\sin(xz) + \cos(xz) - y + 2 = 0,$
 $2^{-x^2-y^2} + z + 1 = 0,$
 $\arctan(xy) - z + 1 = 0 = 0,$
- k) $\cosh(x+y) - z - 5 = 0,$
 $\sin \frac{x+z}{10} - y = 0,$
 $2 \cdot 3^{-y-z} - x = 0,$
- l) $2 \cdot 8^{-y-z+3} - x - 0.6 = 0,$
 $10x(y-2z)^2 - 1 = 0,$
 $\cos^3 x - y + \cos^3 z + 4 = 0,$
- m) $4x^2 + 2(y-4)^2 + z^2 - 20 = 0,$
 $18x^4 + 3y^6 + x^4y^6 - x^4y^6z = 0,$
 $3e^x + 2e^y - 6z^2 - 24 = 0,$
- n) $2\sqrt{x} + \sqrt[4]{y} - z^3 = 0,$
 $x^5 - 2y = 0,$

- $e^{x+z} - e^{-x-z} - 2y = 0,$
- $(x+z)^2 - y + 0.5 = 0,$
- $2x^4 - z = 0,$
- $x^2 + y^2 + z^2 - 4 = 0,$
- $\cos 3x - 5y + \cos 2z = 0,$
- $e^{-2x} + e^{-2y} - z - 1 = 0,$
- $\arctan 3x + \arctan 3y - 3z + 1 = 0,$
- $x^2 + y^2 + z^2 - 64 = 0,$
- $z(x^2 + y^2) - 10 = 0,$
- $e^{-2y^2-3z^2} - x + 5 = 0,$
- $20x - 2y^4 - z^6 = 0,$
- $2(x-5)^4 - y + z^4 = 0,$
- $\sin^5 5x + \cos^3 3y - z + 2 = 0,$
- $\arctan 2x - 3y + \arctan 2z + 15 = 0,$
- $x^2 - 20y + z^2 = 0,$
- $2^{-x} + 2^{-y} - z + 2 = 0,$
- $3^x + 3^y - z + 2 = 0,$
- $\sin \frac{x}{5} + \cos \frac{y}{5} + 2y + 3 = 0.$

Zad. 4.5. Rozwiązać z dokładnością ε następujące układy równań:

- $x^2 + 4y^2 + 25z^2 - 100 = 0,$
- $20x - 4\sin \frac{y}{5} - 5\sin \frac{z}{4} - 20 = 0,$
- $e^{-2x} + e^{-2y} - 10z = 0,$
- $\arctan x + \arctan y - 2z = 0,$
- $4x^2 + 4y^2 + 9z^2 - 36 = 0,$
- $x - 5^{-y} - 4^{-z} + 2 = 0,$
- $4z(x^2 + y^2) + 5 = 0,$
- $3x^4 + y + 2z^6 = 0,$
- $x^2 - y + z^2 - 6 = 0,$

$$\sin 5x + \cos 3y - z + 2 = 0,$$

d) $3\arctan 2x - 6y + 2\arctan 3z + 30 = 0,$

$$x^4 - 20y + z^2 = 0,$$

$$1.7^{-x} + 2.2^{-y} - z + 1.5 = 0,$$

e) $1.5^x + 2.5^y - 3z + 2.5 = 0,$

$$\sin^2(y+z) - x + 2 = 0,$$

$$x(2y+z) - 1 = 0,$$

f) $x^2 + y^2 + z^2 - 4 = 0,$

$$y^3 + z = 0,$$

$$x - 2(y-1)^2 - 3(z-1)^2 = 0,$$

g) $\log(y+2) - x + \log(z+3) + 2 = 0,$

$$y^3 - z = 0,$$

$$25x^2 + 400y^2 - 16z^2 + 400 = 0,$$

h) $5x - 3^{z+y-10} + 10 = 0,$

$$9x - 16y + 2z - 20 = 0,$$

$$20x - 2y^2 + 5z^2 = 0,$$

i) $(x+2)\left(3\sqrt{y} + \sqrt[4]{z}\right) + 1 = 0,$

$$\sin 2x - y + \sin 2z + 5 = 0,$$

$$3x^2 + 12y - 2z^2 = 0,$$

j) $(z-5)(3x^4 + 2y^4) - 1 = 0,$

$$\cos \frac{x}{2} - y + \sin \frac{z}{2} - 5 = 0,$$

$$\arctan(x+y) - z - 2 = 0,$$

k) $\cosh(x^2 + y^2) + z - 5 = 0,$

$$\cos \frac{x+z}{10} - y - 1 = 0,$$

$$3x^2 - 15y + z^2 = 0,$$

l) $9x^2 - 72y + 4z^2 + 144 = 0,$

$$y(x-z)^3 - 1 = 0,$$

$$x^2 + 2y^2 - x^2 y^2 z + 2x^2 y^2 = 0,$$

$$x^2 z - y + \sin^3 z + 6 = 0,$$

$$x^2 + 15(y-6)^2 + 12z^2 - 300 = 0,$$

$$x^2 + 4\sqrt{y} - z^2 = 0,$$

$$x - y^2 = 0,$$

$$x^2 + y^2 - z - 4 = 0,$$

$$x^2 - e^{-x-z} - 2y = 0,$$

$$x^2 + z^3 - y - 1 = 0,$$

$$x^2 + e^{-x} - 2z - 2 = 0,$$

$$x^2 + 2\sqrt[4]{y} - z + 1 = 0,$$

$$x^2 + y^2 + z^2 - 16 = 0,$$

$$x^2 - 2y^2 = 0,$$

$$x^2 + z^2 - 5 = 0,$$

$$-\log(y+4) - \log(z+5) = 0,$$

$$x - z = 0,$$

$$(x-2)^2 - 900y^2 - 225z^2 - 8100 = 0,$$

$$(x-2)^2 + 2(y+2)z^2 + 5 = 0,$$

$$x + y + z - 14 = 0,$$

$$x^2 + y^2 - 18z = 0,$$

$$\frac{y}{2} + \cos \frac{z}{2} = 0,$$

$$x - y - z + 1 = 0,$$

$$x^2 - y^2 + 8z = 0,$$

$$x^2 + \cos^2 3y - z - 4 = 0,$$

$$x + 3y - z - 3 = 0.$$

Rozwiązać z dokładnością ε następujące układy równań:

$$x^2 + y^2 + 4z^2 - 100 = 0,$$

$$\frac{x}{2} - y + \cos \frac{z}{2} - 4 = 0,$$

$$(x^2 + 3y^2) - 1 = 0,$$

$$2 \arctan \frac{x}{5} + 3 \arctan \frac{y}{5} - z = 0,$$

b) $4x^2 + 4y^2 + 9z^2 - 36 = 0,$

$$e^{-x^2-y^2} - z = 0,$$

$$2x^2y + 3z^2y - 1 = 0,$$

c) $8x^4 + y + 10z^6 - 10 = 0,$

$$2x^6 + 6y^4 + z = 0,$$

$$\sin^3 x + \cos^3 y - z = 0,$$

d) $\arctan \frac{x}{10} - y + \arctan \frac{z}{10} + 6 = 0,$

$$x^2 - 5y + z^4 = 0,$$

$$4^x + 4^{-y} - z + 1 = 0,$$

e) $5^{-x} + 5^y - 3z + 1.5 = 0,$

$$\sin(2y + 3z) + x + 2 = 0,$$

$$3^{-x} - y + 2^{-z} = 0,$$

f) $\log(x+1) - y + \log(z+1) + 1 = 0,$

$$x^2 + y^2 + z^2 - 100 = 0,$$

$$4x^2 - 9y^2 + 4z^2 + 36 = 0,$$

g) $z - \log(x+y+3) - 3 = 0,$

$$4x - 2y + 4z - 8 = 0,$$

$$5(x-3)^2 + 40y - 4(z+2)^2 = 0,$$

h) $4(x-3)^2 - y + 2(z+2)^2 - 5 = 0,$

$$\sin x^2 + \sin y^2 - z = 0,$$

$$5(x-1)^2 + 60y - 6(z+1)^2 = 0,$$

i) $\sin^3(x+y) - z + 5 = 0,$

$$x + 5y + z - 10 = 0,$$

$$\arctan(x^2 + y^2) - z = 0 = 0,$$

j) $x^6 + y^6 - 5z - 30 = 0,$

$$\sin \frac{xz}{10} - y = 0,$$

$$\sqrt[10]{x} + \sqrt[8]{y} - z = 0,$$

□) $\sqrt[3]{x+6} - y + \sqrt[8]{y+6} - z = 0,$

$$25x(3y-2z)^2 - 1 = 0,$$

$$xyz - xz - 10y - 8z = 0,$$

□) $\arctan x + 2 \arctan y - z = 0,$

$$x - (y-6)^2 - (z-5)^2 - 100 = 0,$$

$$\sqrt[4]{x} - \sqrt{y} - \sqrt[6]{z} = 0,$$

■) $y^5 - z^2 = 0,$

$$2x - y^4 - 2z^4 + 6 = 0,$$

$$\sqrt{x+z-1} - y^2 = 0,$$

■) $\ln(x+z+1) - y - 0.5 = 0,$

$$x^4 - 10z = 0,$$

$$20x - 5y^2 - 4z^2 = 0,$$

□) $7x^2 - 84y + 6z^2 = 0,$

$$10x(2y+3z)^4 - 1 = 0,$$

$$10x(2y-3z^3)^2 + 1 = 0,$$

□) $x - \arctan z + 5 = 0,$

$$(x-5)^2 - y^2 - z^2 - 40 = 0,$$

$$2^x + 3^y - z^3 = 0,$$

□) $\sqrt{x} + \sqrt[4]{y} - z^2 + 4 = 0,$

$$100(x^2 - y) + (z-4)^2 = 0,$$

$$e^{x+z} + e^{-x-z} - 2y^4 - 2 = 0,$$

□) $\sqrt[3]{x+z+1} - y^2 = 0,$

$$x^3 - z = 0.$$

Zad. 4.7. Rozwiązać z dokładnością ε następujące układy równań:

$$25x^2 + 4y^2 + z^2 - 100 = 0,$$

■) $\sin 3x + \cos 2y - z + 3 = 0,$

$$y(x^4 + 2z^2) - 4 = 0,$$

$$2 \arctan 3x + 3 \arctan 3y - z = 0,$$

b) $x^2 + y^2 + z^2 - 100 = 0,$

$$e^{-2x^2-y^2} + z = 0,$$

$$20(x^4 + z^4)(y - 2) + 1 = 0,$$

c) $10x^6 - y + 10z^2 - 2 = 0,$

$$x^2 + y^2 - 20z = 0,$$

$$\sin^2 x + \cos^2 y - z + 3 = 0,$$

d) $\arctan \frac{x}{4} - y + \arctan \frac{z}{3} + 150 = 0,$

$$x^4 - 10y + z^4 = 0,$$

$$4^x + 3^y - z + 1 = 0,$$

e) $6^x + 5^y - z + 1 = 0,$

$$\sin \frac{x}{3} + \cos \frac{y}{3} - 2z + 4 = 0,$$

$$3\sqrt[3]{x} + 5\sqrt[4]{y} - z - 2 = 0,$$

f) $4x^2 + 25y^2 + z^2 - 100 = 0,$

$$3x - \sqrt{y+1} = 0,$$

$$(x^2 + z^2)y - 2 = 0,$$

g) $2 \log(x+3) - y + 3 \log(z+4) = 0,$

$$x - \sqrt[3]{z} = 0;$$

$$36x^2 - 100y^2 + 225z^2 = 0,$$

h) $2^{x+y-6} - 4z - 4 = 0,$

$$6x - y - 6z + 6 = 0,$$

$$36x^2 - 100y^2 + 225z^2 + 400 = 0,$$

i) $2^{x+y-6} - 4z - 4 = 0,$

$$6x - y - 6z + 6 = 0,$$

$$5x^2 + 30y - 3z^2 = 0,$$

j) $\cos^2(x+y) - z - 5 = 0,$

$$5x + 10y - z - 10 = 0,$$

$$\operatorname{arcctg}(x+z) - y = 0,$$

■) $x^4 + y + z^4 - 6 = 0,$

$$\sin \frac{y+z}{10} - x = 0,$$

$$2^{-x-y} - z = 0,$$

■) $3^{-x-y+2} - z - 0.5 = 0,$

$$20z(2x-y)^4 - 1 = 0,$$

$$xy^2z^4 - 8y^2 - 2z^4 - 3y^2z^4 = 0,$$

■) $\sin^5 x + \sin^7 y - z + 5 = 0,$

$$x^2 + 6y^2 + 5z^2 - 300 = 0,$$

$$2x^4 - \sqrt[3]{z} = 0,$$

■) $4x^2 + y^2 + 25z^2 - 100 = 0,$

$$\sqrt[4]{x} - 8\sqrt{y} + \sqrt[6]{z} = 0,$$

$$e^{x+z} - e^{-x-z} - 2y - 4 = 0,$$

■) $\ln(x+z) - y = 0,$

$$x - z^3 = 0,$$

$$e^{x+y} - e^{z-x} - 4.5 = 0,$$

■) $\cosh(x-y+z^2) - xyz - 1.8 = 0,$

$$\sinh(x+y+z) + x^2 - z^2 - 6.0 = 0,$$

$$x + x^2 - 2yz - 0.1 = 0,$$

■) $y - y^2 + 3xz + 0.2 = 0,$

$$z + z^2 + 2xy - 0.3 = 0,$$

$$\sin(x+y^2+z) + xyz + 0.80 = 0,$$

■) $x^y + y^2 + z^3 - 0.5z = 0,$

$$x \tanh\left(\frac{y-x}{z}\right) + yz + 1.25 = 0,$$

$$(z-1)(3x^4 + y^6) - 2 = 0,$$

■) $\operatorname{arcctg}(x+z) - y^3 = 0,$

$$25x^2 + 4(y-3)^2 + 25z^2 - 100 = 0,$$

$$4x^2 + y^2 + 4z^2 - 36 = 0,$$

t) $x^4 - y + 3z^4 - 1 = 0,$

$$9x^3 - z^2 = 0.$$

Zad. 4.8. Rozwiązać z dokładnością ε następujące układy równań:

$$25x^2 + 4y^2 + 4z^2 - 100 = 0,$$

a) $\cos \frac{x}{10} - 10y + \cos \frac{z}{10} + 2 = 0,$

$$z(3x^2 + 5y^2) - 4 = 0,$$

$$2\arctan \frac{x}{10} + 2\arctan \frac{y}{10} - z = 0,$$

b) $x^2 + y^2 + z^2 - 100 = 0,$

$$z(x^4 + y^4) - 2 = 0,$$

$$(x+3)(y^2 + z^2) - 1 = 0,$$

c) $x + 2y^4 + 3z^6 = 0,$

$$x^4 - y + z^4 = 0,$$

$$\sin \frac{x}{5} + \cos \frac{y}{5} - 5z + 15 = 0,$$

d) $\arctan 5x - 3y + \arctan 5z + 15 = 0,$

$$x^4 - 20y + z^4 = 0,$$

$$5^{-x} + 5^{-y} - z + 1 = 0,$$

e) $4^{-x} + 4^{-y} - z + 2 = 0,$

$$\sin \frac{y}{10} + \cos \frac{z}{10} + 2x + 3 = 0,$$

$$3\sqrt{x} + 2\sqrt{y} - z + 1 = 0,$$

f) $x^2 + y^2 + z^2 - 25 = 0,$

$$x^3 - y - 1 = 0,$$

$$(x-2)(2y^2 + 3z^2 + 1) - 4 = 0,$$

g) $2x - \ln y - \ln z = 0,$

$$\sqrt[4]{y} - 2z = 0,$$

- $4x^2 + 8x - 36y^2 - 9z^2 - 32 = 0,$
- $y(x+1)^2 + 3yz^2 - 2 = 0,$
- $2x + y + z + 8 = 0,$
- $9x^2 - 4y^2 - 72z = 0,$
- $x - \sin 2y - \cos 3z - 4 = 0,$
- $x + 4y + 4z - 4 = 0,$
- $x^2 - 4y^2 + 8z = 0,$
- $\sin^3 x + \cos^3 y - z - 4 = 0,$
- $x + y + z - 1 = 0,$
- $\arcctg(x^2 + z^2) - y = 0,$
- $1.5^x + 2.5^{-xz} + 2y = 0,$
- $e^x - e^{-x} - 5z = 0,$
- $\sqrt[4]{x} - y + \sqrt[8]{z} + 0.5 = 0,$
- $\sqrt[4]{x+5} - z + \sqrt[8]{y+5} = 0,$
- $10z(x-y)^4 - 1 = 0,$
- $x^3 - y^2 - z^4 + 1 = 0,$
- $3x + 2y^3 - 20z + 507.3 = 0,$
- $5\ln(x-3) - y^2 + z + 9.003 = 0,$
- $\log(xy) - \sqrt{z} = 0,$
- $e^{-x} + e^{-y} - z^2 = 0,$
- $x^2 - y + z^4 = 0,$
- $2^{x+z} - y - 1 = 0,$
- $\ln(x+z+2) - y^5 = 0,$
- $10x^2 + 10(y-1)^2 - z = 0.$

Czytelnik: CHOJNACKA, Aleksandra
 Tytuł: Metody numeryczne i przykłady zastosowania
 Autor: (P) Ratajczak, Tadeusz i Przykłady zastosowania
 Sygnatura: BG P 714873-000005
 Kod kreskowy: 714873000005 (informacja z 2005)

Interpolacja

5.1. Interpolacja w sensie Lagrange'a

Dana jest funkcja $y = f(x)$, określona i ciągła w przedziale $[a ; b]$, ale znamy jej wartości y_0, y_1, \dots, y_n , które ona przyjmuje dla skończonego zbioru argumentów x_0, x_1, \dots, x_n , tzn.

$$y_i = f(x_i) \quad \text{dla } i = 0, 1, \dots, n \quad (5.1)$$

Argumenty będziemy nazywali *węzłami interpolacji* i będziemy zakładali, że

$$a \leq x_0 < x_1 < \dots < x_n \leq b.$$

Taką funkcję nazywamy *funkcją tabelaryczną*, ponieważ można ją określić, podając tabelę (tab. 5.1).

Tabela 5.1. Tabelaryczna postać funkcji interpolowanej

x_i	x_0	x_1	...	x_n
$y_i = f(x_i)$	y_0	y_1	...	y_n

Zagadnienie *interpolacji w sensie Lagrange'a* polega na znalezieniu funkcji $y = F(x)$, pewnej, określonej klasy, która w węzłach interpolacji przyjmuje te same wartości co funkcja $y = f(x)$, tj.

$$F(x_i) = y_i \quad \text{dla } i = 0, 1, \dots, n.$$

Funkcję $y = F(x)$ będziemy nazywać *funkcją interpolującą*, a $y = f(x)$ – *funkcją interpolowaną*.

Zagadnienie interpolacji wykorzystuje się zazwyczaj do obliczenia przybliżonej wartości funkcji $y = f(x)$ dla argumentów x różnych od węzłów interpolacji. W tym celu dobiera się funkcję interpolującą $y = F(x)$, możliwie dobrze przybliżającą funkcję $y = f(x)$, oblicza się wartość $F(x)$ i przyjmuje się, że

$$f(x) \approx F(x).$$

Ścisłe rzecz biorąc, obliczona wartość $F(x)$ jest obarczona błędem

$$R(x) = f(x) - F(x), \quad (5.2)$$

zwanym *błędem interpolacji* (*błędem metody*), który powinniśmy oszacować, gdy obliczamy wartość $F(x)$.

Odróżnia się pojęcie *interpolacji w węższym sensie*, oznaczające zagadnienie poszukiwania wartości funkcji $y = f(x)$ dla $x \in (x_0 ; x_n)$, od *ekstrapolacji* oznaczającej poszukiwanie wartości funkcji $y = f(x)$ dla $x < x_0$ lub $x > x_n$. Zastosowanie zagadnienia interpolacji ilustruje przykład 5.1.

Przykład 5.1. Odczytane wartości temperatury zestawiono w tabeli.

Tabela 5.2. Temperatura T w funkcji czasu t

Czas t [godz.]	12	13	14	15	16
Temperatura T [$^{\circ}$ C]	24	25	23	20	16

Była temperatura o godzinie 14.30?

Sądzę, że Czytelnik zgodzi się z założeniem, że temperatura T jest funkcją ciągłą zmiennej t (czasu). Mamy zatem wyznaczyć wartość funkcji $T = T(t)$ dla argumentu $t = 14.5$ (czas zapisany w systemie dziesiętnym).

W zależności od klasy funkcji interpolującej, zagadnienie interpolacji może mieć dość jedno rozwiązanie, wiele rozwiązań (nawet nieskończenie wiele) lub nie mieć rozwiązań wcale. W dalszym ciągu przyjmiemy, że funkcja interpolująca jest *wielomianem interpolującym*, tzn. funkcją postaci

$$F(x) = a_0 \varphi_0(x) + a_1 \varphi_1(x) + \dots + a_n \varphi_n(x), \quad (5.3)$$

$$\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x) \quad (5.4)$$

funkcjami określonymi na przedziale $[a ; b]$ i tworzą na tym przedziale układ liniowo niezależny. Zbiór funkcji (5.4) będziemy nazywali *bazą*. Aby wielomian uogólniony (5.3) do funkcji interpolującej, muszą zachodzić równości

$$\begin{aligned} F(x_0) &= a_0 \varphi_0(x_0) + a_1 \varphi_1(x_0) + \dots + a_n \varphi_n(x_0) = y_0, \\ F(x_1) &= a_0 \varphi_0(x_1) + a_1 \varphi_1(x_1) + \dots + a_n \varphi_n(x_1) = y_1, \\ &\dots, \\ F(x_n) &= a_0 \varphi_0(x_n) + a_1 \varphi_1(x_n) + \dots + a_n \varphi_n(x_n) = y_n. \end{aligned} \quad (5.5)$$

Równania (5.5) tworzą układ $n+1$ równań liniowych z niewiadomymi a_0, a_1, \dots, a_n . Jeśli

$$\begin{vmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \dots & \varphi_n(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \dots & \varphi_n(x_1) \\ \dots & \dots & \dots & \dots \\ \varphi_0(x_n) & \varphi_1(x_n) & \dots & \varphi_n(x_n) \end{vmatrix} \neq 0,$$

układ (5.5) ma dokładnie jedno rozwiązanie. W tym przypadku, rozwiązując ten układ, oznaczamy współczynniki a_0, a_1, \dots, a_n , które z funkcji (5.3) czynią funkcję interpolującą.

5.2. Wzór Lagrange'a

Niech bazę tworzą funkcje

$$\varphi_0(x) = 1, \varphi_1(x) = x, \dots, \varphi_n(x) = x^n$$

(funkcje te tworzą układ liniowo niezależny dla $-\infty < x < +\infty$). W tym przypadku funkcja (5.3) przyjmuje postać

$$F(x) = a_0 + a_1 x + \dots + a_n x^n, \quad (5.5)$$

a więc jest zwykłym wielomianem. Układ (5.5) przyjmuje postać

$$\begin{aligned} a_0 + a_1 x_0 + a_2 x_0^2 + \dots + a_n x_0^n &= y_0, \\ a_0 + a_1 x_1 + a_2 x_1^2 + \dots + a_n x_1^n &= y_1, \\ \dots & \\ a_0 + a_1 x_n + a_2 x_n^2 + \dots + a_n x_n^n &= y_n. \end{aligned} \quad (5.7)$$

Wyznacznik macierzy głównej tego układu jest wyznacznikiem Vandermonda

$$\begin{vmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^n \end{vmatrix} = \prod_{i>j} (x_i - x_j) \neq 0,$$

a więc układ równań (5.7) ma dokładnie jedno rozwiązanie. Rozwiązujeć układ (5.7) otrzymujemy współczynniki szukanego wielomianu interpolacyjnego (5.6). Praktycznie ten sposób nie wyznacza się wielomianów interpolujących, ponieważ zadanie rozwiązywania układu równań (5.7) jest na ogół zadaniem *źle uwarunkowanym*.

Jedną z częściej stosownych funkcji interpolacyjnych jest *wielomian interpolacyjny Lagrange'a*

$$L_n(x) = \sum_{i=0}^n y_i \frac{(x - x_0)(x - x_1) \cdot \dots \cdot (x - x_{i-1})(x - x_{i+1}) \cdot \dots \cdot (x - x_n)}{(x_i - x_0)(x_i - x_1) \cdot \dots \cdot (x_i - x_{i-1})(x_i - x_{i+1}) \cdot \dots \cdot (x_i - x_n)}. \quad (5.8)$$

Dla obliczenia wartości $L_n(x)$ dla zadanego x wykorzystuje się inną postać wielomianu Lagrange'a, wymagającą wykonania mniejszej ilości działań arytmetycznych

$$L_n(x) = \omega_{n+1}(x) \sum_{i=0}^n \frac{y_i}{(x - x_i)(x_i - x_0)(x_i - x_1) \cdot \dots \cdot (x_i - x_{i-1})(x_i - x_{i+1}) \cdot \dots \cdot (x_i - x_n)}, \quad (5.9)$$

gdzie

$$\omega_{n+1}(x) = (x - x_0)(x - x_1) \cdot \dots \cdot (x - x_n). \quad (5.10)$$

Dowodzi się, że (5.8) jest wielomianem stopnia co najwyżej n i że błąd interpolacji z użyciem wielomianu interpolacyjnego Lagrange'a spełnia nierówność

$$|f(x) - L_n(x)| \leq |\omega_{n+1}(x)| \frac{M_{n+1}}{(n+1)!}, \quad (5.11)$$

$$M_{n+1} = \max_{a \leq x \leq b} |f^{(n+1)}(x)|.$$

Przykład 5.2. Zastosować wielomian interpolacyjny Lagrange'a do rozwiązania zadania zadanego w przykładzie 5.1.

Rozwiązanie. Przypomnijmy, że dana jest tabela wartości temperatur

Czas t [godz.]	12	13	14	15	16
Temperatura T [$^{\circ}$ C]	24	25	23	20	16

Wależy określić temperaturę o godzinie 14.30. Do obliczenia $T(14.5)$ zastosujemy wzór Lagrange'a (5.9).

$$L_4(t) = (t - t_0)(t - t_1)(t - t_2)(t - t_3)(t - t_4) \left\{ \frac{T_0}{(t - t_0)(t_0 - t_1)(t_0 - t_2)(t_0 - t_3)(t_0 - t_4)} + \right. \\ \left. + \frac{T_1}{(t - t_1)(t_1 - t_0)(t_1 - t_2)(t_1 - t_3)(t_1 - t_4)} + \frac{T_2}{(t - t_2)(t_2 - t_0)(t_2 - t_1)(t_2 - t_3)(t_2 - t_4)} \right. \\ \left. + \frac{T_3}{(t - t_3)(t_3 - t_0)(t_3 - t_1)(t_3 - t_2)(t_3 - t_4)} + \frac{T_4}{(t - t_4)(t_4 - t_0)(t_4 - t_1)(t_4 - t_2)(t_4 - t_3)} \right\}.$$

$$L_4(14.5) = (14.5 - 12)(14.5 - 13)(14.5 - 14)(14.5 - 15)(14.5 - 16) \cdot$$

$$\left\{ \frac{24}{(14.5 - 12)(12 - 13)(12 - 14)(12 - 15)(12 - 16)} + \right.$$

$$\left. \frac{25}{(14.5 - 13)(13 - 12)(13 - 14)(13 - 15)(13 - 16)} + \right.$$

$$\left. \frac{23}{(14.5 - 14)(14 - 12)(14 - 13)(14 - 15)(14 - 16)} + \right.$$

$$\left. \frac{20}{(14.5 - 15)(15 - 12)(15 - 13)(15 - 14)(15 - 16)} + \right.$$

$$\left. \frac{16}{(14.5 - 16)(16 - 12)(16 - 13)(16 - 14)(16 - 15)} \right\} =$$

$$= (2.5)(1.5)(0.5)(-0.5)(-1.5)$$

$$\left\{ \frac{24}{(2.5)(-1.0)(-2.0)(-3.0)(-4.0)} + \frac{25}{(1.5)(1.0)(-1.0)(-2.0)(-3.0)} + \right.$$

$$\left. + \frac{23}{(0.5)(2.0)(1.0)(-1.0)(-2.0)} + \frac{20}{(-0.5)(3.0)(2.0)(1.0)(-1.0)} + \right.$$

$$\left. + \frac{16}{(-1.5)(4.0)(3.0)(2.0)(1.0)} \right\} =$$

$$\begin{aligned}
 &= (1.40625) \cdot \left\{ \frac{24}{60} - \frac{25}{9} + \frac{23}{2} + \frac{20}{3} - \frac{16}{36} \right\} = \\
 &= 1.40625 \cdot \{0.40000 - 2.77778 + 11.50000 + 6.66667 - 0.44444\} = \\
 &= 1.40625 \cdot 15.34444 = 21.57813 \approx 21.6.
 \end{aligned}$$

Odpowiedź: O godz. 14.30 temperatura była (w przybliżeniu) równa 21.6°C . Nie potrafimy oszacować błędu obliczonego wyniku, ponieważ nie mamy informacji o $T^{(5)}(t)$.

Gdy planujemy eksperyment, możemy postawić pytanie: jak w przedziale $[a ; b]$ wybrać węzły x_0, x_1, \dots, x_n , aby błąd interpolacji (5.11) był najmniejszy? To zagadnienie rozwiązał Czebyszew, a w niniejszym skrypcie rozwiązanie jest sformułowane w postaci twierdzeń 6.12 oraz 6.9. Z tych twierdzeń wynika, że optymalne węzły są zerami wielomianu Czebyszewa pierwszego rodzaju i w związku z tym wyrażają się wzorami

$$x_k = \frac{b-a}{2} \cdot \cos \frac{2k+1}{2n+2} \pi + \frac{b+a}{2} \quad k = 0, 1, \dots, n.$$

W tym przypadku nierówność (5.11) przybiera postać

$$|f(x) - L_n(x)| \leq \frac{M_{n+1}}{(n+1)!} \cdot \frac{(b-a)^{n+1}}{2^{2n+1}}.$$

5.3. Interpolacja Czebyszewa

Jeżeli za bazę przyjąć wielomiany Czebyszewa pierwszego rodzaju, określone wzorami

$$\begin{aligned}
 T_0(x) &= 1, \\
 T_1(x) &= x, \\
 T_k(x) &= 2 \cdot x \cdot T_{k-1}(x) - T_{k-2}(x) \quad \text{dla } k = 2, 3, \dots,
 \end{aligned} \tag{5.12}$$

to funkcja interpolująca przyjmie postać

$$F(x) = a_0 T_0(x) + a_1 T_1(x) + \dots + a_n T_n(x). \tag{5.13}$$

Współczynniki a_0, a_1, \dots, a_n wyznaczamy, rozwiązuając układ równań

$$\begin{aligned}
 F(x_0) &= T_0(x_0)a_0 + T_1(x_0)a_1 + \dots + T_n(x_0)a_n = y_0, \\
 F(x_1) &= T_0(x_1)a_0 + T_1(x_1)a_1 + \dots + T_n(x_1)a_n = y_1, \\
 &\dots \\
 F(x_n) &= T_0(x_n)a_0 + T_1(x_n)a_1 + \dots + T_n(x_n)a_n = y_n.
 \end{aligned} \tag{5.14}$$

Wielomian (5.13) ze współczynnikami otrzymanymi przez rozwiązywanie układu równań (5.14) nosi nazwę wielomianu interpolacyjnego Czebyszewa.

Wielomiany Czebyszewa pierwszego rodzaju są określone w przedziale $[-1 ; 1]$. Jeżeli interpolujemy funkcję $y = f(t)$, przy czym $a \leq t \leq b$, to trzeba wykonać podstawienie

$$t = \frac{a+b}{2} + \frac{b-a}{2} x \tag{5.15}$$

$a \leq t \leq b$ wtedy i tylko wtedy, gdy $-1 \leq x \leq 1$). Interpolacja Czebyszewa jest mniej wrażliwa na błędy zaokrągleń od interpolacji z użyciem wzoru Lagrange'a.

Druga ważna własność interpolacji Czebyszewa ujawnia się, gdy węzły interpolacji zbiegają się z zerami wielomianu Czebyszewa stopnia $n+1$

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

Wówczas macierz główna układu równań (5.14) pomnożona przez $\sqrt{2/n+1}$ jest macierzą diagonalną. W tym przypadku obliczenia upraszczają się – możemy podać bezpośrednie wzory na obliczenie współczynników a_0, a_1, \dots, a_n

$$a_k = \frac{2}{n+1} \cdot \sum_{i=0}^n T_k(x_i) y_i \quad \text{dla } k = 0, 1, 2, \dots, n. \quad (5.16)$$

Przykład 5.3. Zastosować wielomian interpolacyjny Czebyszewa do rozwiązania zadania formułowanego w przykładzie 5.1.

Rozwiązanie. Przypomnijmy, że dana jest tabela wartości temperatur

Czas t [godz.]	12	13	14	15	16
Temperatura T [$^{\circ}\text{C}$]	24	25	23	20	16

Należy określić temperaturę o godzinie 14.30. Do obliczenia $T(14.5)$ zastosujemy wzór interpolacyjny (5.13). Ze wzorów (5.12) mamy:

$$T_0(x) = 1,$$

$$T_1(x) = x,$$

$$T_2(x) = 2x^2 - 1,$$

$$T_3(x) = 4x^3 - 3x,$$

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

czym $-1 \leq x \leq 1$. Ponieważ funkcja interpolowana jest określona w przedziale $[a; b] = [12; 16]$, należy zastosować przekształcenie argumentów odwrotne do (5.15)

$$x = \frac{2}{b-a} \cdot t - \frac{a+b}{b-a} = \frac{2}{16-12} \cdot t - \frac{12+16}{16-12} = \frac{1}{2} \cdot t - 7.$$

Zatem funkcję $T = T(t)$ będziemy interpolowali wielomianem

$$F(x(t)) = a_0 T_0\left(\frac{t}{2} - 7\right) + a_1 T_1\left(\frac{t}{2} - 7\right) + a_2 T_2\left(\frac{t}{2} - 7\right) + a_3 T_3\left(\frac{t}{2} - 7\right) + a_4 T_4\left(\frac{t}{2} - 7\right).$$

Współczynniki a_0, a_1, a_2, a_3, a_4 obliczamy, rozwiązuając układ równań liniowych

$$\begin{bmatrix} T_0(t_0) & T_1(t_0) & T_2(t_0) & T_3(t_0) & T_4(t_0) \\ T_0(t_1) & T_1(t_1) & T_2(t_1) & T_3(t_1) & T_4(t_1) \\ T_0(t_2) & T_1(t_2) & T_2(t_2) & T_3(t_2) & T_4(t_2) \\ T_0(t_3) & T_1(t_3) & T_2(t_3) & T_3(t_3) & T_4(t_3) \\ T_0(t_4) & T_1(t_4) & T_2(t_4) & T_3(t_4) & T_4(t_4) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} T_0 \\ T_1 \\ T_2 \\ T_3 \\ T_4 \end{bmatrix},$$

czyli

$$\begin{bmatrix} 1.0 & -1.0 & 1.0 & -1.0 & 1.0 \\ 1.0 & -0.5 & -0.5 & 1.0 & -0.5 \\ 1.0 & 0.0 & -1.0 & 0.0 & 1.0 \\ 1.0 & 0.5 & -0.5 & -1.0 & -0.5 \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 24.0 \\ 25.0 \\ 23.0 \\ 20.0 \\ 16.0 \end{bmatrix}.$$

Rozwiązaniem tego układu jest:

$$a_0 = 21.66667, a_1 = -4.33333, a_2 = -1.50000, a_3 = 0.33333, a_4 = -0.16667.$$

Ponieważ dla $t = 14.5$, $x = \frac{14.5}{2} - 7 = 0.25$, więc

$$F(0.25) = 21.66667 \cdot (1) - 4.33333 \cdot (0.25) - 1.50000 \cdot (2 \cdot 0.25^2 - 1) + \\ + 0.33333 \cdot (4 \cdot 0.25^3 - 3 \cdot 0.25) - 0.16667 \cdot (8 \cdot 0.25^4 - 0.25^2 + 1) = 21.57813$$

Odpowiedź: O godz. 14.30 temperatura była (w przybliżeniu) równa 21.6°C .

5.4. Interpolacja trygonometryczna

W wielu przypadkach zachodzi potrzeba interpolowania funkcji okresowej. Założymy w dalszym ciągu, że funkcja interpolowana $y = f(t)$ jest funkcją okresową, o okresie głównym $[0 ; 2\pi]$. W przypadku ogólniejszym, gdy okresem głównym funkcji interpolowanej $y = f(t)$ jest przedział $[a ; b]$, stosujemy przekształcenie

$$t = a + \frac{b-a}{2\pi}x$$

($a \leq t \leq b$ wtedy i tylko wtedy, gdy $0 \leq x \leq 2\pi$).

W przypadku funkcji okresowej za bazę przyjmiemy *układ trygonometryczny*

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos mx, \sin mx$$

Wówczas funkcja interpolująca przyjmie postać *wielomianu trygonometrycznego*

$$F(x) = a_0 + \sum_{k=1}^m (a_k \cos kx + b_k \sin kx), \quad (5.17)$$

przy czym $n = 2m$. Współczynniki $a_0, a_1, b_1, \dots, a_n, b_n$ wyznaczamy, rozwiązujeć układ równań

$$\begin{aligned} F(x_0) &= a_0 + a_1 \cos x_0 + b_1 \sin x_0 + \dots + a_m \cos mx_0 + b_m \sin mx_0 & = y_0, \\ F(x_1) &= a_0 + a_1 \cos x_1 + b_1 \sin x_1 + \dots + a_m \cos mx_1 + b_m \sin mx_1 & = y_1, \end{aligned} \quad (5.18)$$

.....

$$F(x_{2m}) = a_0 + a_1 \cos x_{2m} + b_1 \sin x_{2m} + \dots + a_m \cos mx_{2m} + b_m \sin mx_{2m} = y_{2m}.$$

W tym najważniejszy jest przypadek, gdy węzły interpolacji są równoodległe, tj.

$$x_i = \frac{2i}{2m+1} \pi, \quad \text{dla } i = 0, 1, \dots, 2m. \quad (5.19)$$

W tym przypadku macierz główna układu równań (5.18) pomnożona przez $\sqrt{2/(2m+1)}$ jest ortogonalną. W konsekwencji możemy podać bezpośrednie wzory na obliczenie współczynników $a_0, a_1, b_1, \dots, a_m, b_m$:

$$a_0 = \frac{2}{2m+1} \cdot \sum_{i=0}^{2m} y_i$$

$$a_k = \frac{2}{2m+1} \cdot \sum_{i=0}^{2m} y_i \cos kx_i = \frac{2}{2m+1} \cdot \sum_{i=0}^{2m} y_i \cos k \frac{2i}{2m+1} \pi \quad (k = 1, 2, \dots, m), \quad (5.20)$$

$$b_k = \frac{2}{2m+1} \cdot \sum_{i=0}^{2m} y_i \sin kx_i = \frac{2}{2m+1} \cdot \sum_{i=0}^{2m} y_i \sin k \frac{2i}{2m+1} \pi \quad (k = 1, 2, \dots, m). \quad (5.21)$$

Interpolacja w przypadku węzłów równoodległych

Analizujemy obecnie przypadek, gdy argumenty funkcji dyskretnej

x_i	x_0	x_1	...	x_n
$y_i = f(x_i)$	y_0	y_1	...	y_n

rozłożone w równych odstępach, tzn.

$$x_{i+1} - x_i = h = \text{const dla } i = 0, 1, \dots, n-1.$$

Wzory interpolacyjne wykorzystują pojęcie różnicy skończonej. Różnicą skońzoną dwukrotną rzędu pierwszego dla argumentu x_i nazywamy różnicę

$$\Delta y_i = f(x_i+h) - f(x_i) = y_{i+1} - y_i \quad \text{dla } i = 0, 1, \dots, n-1 \quad (5.22)$$

Różnice skończone progresywne rzędu $k+1$ definiujemy rekurencyjnie

$$\Delta^{k+1} y_i = \Delta(\Delta^k y_i) = \Delta^k y_{i+1} - \Delta^k y_i \quad \text{dla } i = 0, 1, \dots, n-k. \quad (5.23)$$

1. Wzór Stirlinga

Węzły interpolacji ponumerujemy następująco: ..., $x_{-2}, x_{-1}, x_0, x_1, x_2, \dots$. Ponadto

$$q = \frac{x - x_0}{h}. \quad (5.24)$$

Wzór interpolacyjny Stirlinga rozpięty na węzłach $x_{-m}, \dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots, x_m$ przedstawia się następująco:

$$\begin{aligned} S_{2m}(x) = & y_0 + q \cdot \frac{\Delta y_{-1} + \Delta y_0}{2} + \frac{q^2}{2!} \cdot \Delta^2 y_{-1} + \frac{q(q^2 - 1^2)}{3!} \cdot \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \dots + \\ & + \frac{q(q^2 - 1^2)(q^2 - 2^2) \cdot \dots \cdot (q^2 - (m-1)^2)}{(2m-1)!} \cdot \frac{\Delta^{2m-1} y_{-m} + \Delta^{2m-1} y_{-m+1}}{2} \\ & + \frac{q^2(q^2 - 1^2)(q^2 - 2^2) \cdot \dots \cdot (q^2 - (m-1)^2)}{(2m)!} \cdot \Delta^{2m} y_{-m}. \end{aligned} \quad (5.25)$$

Błąd interpolacji z użyciem wzoru Stirlinga jest w przybliżeniu równy

$$\begin{aligned} R_{2m}(x) = & f(x) - S_{2m}(x) \approx \\ & \approx \frac{q(q^2 - 1^2)(q^2 - 2^2) \cdot \dots \cdot (q^2 - m^2)}{(2m+1)!} \cdot \frac{\Delta^{2m+1} y_{-m-1} + \Delta^{2m+1} y_{-m}}{2}. \end{aligned} \quad (5.26)$$

Przykład 5.4. Znane są następujące wartości funkcji $y = f(x)$.

x_i	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6
y_i	1.00000	1.14870	1.31951	1.51572	1.74110	2.00000	2.29740	2.63902	3.03143

Obliczyć $f(0.5)$ z dokładnością $\varepsilon = 10^{-5}$.

Tabela 5.3. Różnice skończone funkcji $y = f(x)$ z przykładu 5.6

i	x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$	$\Delta^5 y_i$	$\Delta^6 y_i$
-3	0.0	1.00000						
			0.14870					
-2	0.2	1.14870		0.02211				
			0.17081		0.00329			
-1	0.4	1.31951		0.02540		0.00048		
			0.19621		0.00377		0.00010	
0	0.6	1.51572		0.02917		0.00058		-0.00005
			0.22538		0.00435		0.00005	
1	0.8	1.74110		0.03352		0.00063		0.00006
			0.25890		0.00498		0.00011	
2	1.0	2.00000		0.03850		0.00074		0.00000
			0.29740		0.00572		0.00011	
3	1.2	2.29740		0.04422		0.00085		
			0.34162		0.00657			
4	1.4	2.63902		0.05079				
			0.39241					
5	1.6	3.03143						

Rozwiązanie. Do obliczenia przybliżonej wartości $f(0.5)$ zastosujemy wzór Stirlinga. Obliczenia rozpoczęmy od odpowiedniego ponumerowania węzłów funkcji interpolowanej. Zauważmy, że jeśli argument $x = 0.5$ leży w pobliżu x_0 , to q obliczone według wzoru (5.24) jest małe i w konsekwencji kolejne składniki wzoru Stirlinga szybko maleją. Oznacza to, że obliczenia wartości $S_{2m}(x)$ zadaną dokładnością ε trzeba we wzorze (5.25) wziąć pod uwagę tylko kilka początkowych składników. W naszym przykładzie przyjmiemy $x_0 = 0.6$, co daje

$$q = \frac{0.5 - 0.6}{0.2} = -0.5.$$

W tabeli 5.3 zestawiono różnice skończone aż do rzędu szóstego. Tłustym drukiem zostały zaznaczone liczby, które wystąpią we wzorze Stirlinga. Wartość $S_{2k}(x)$ obliczamy iteracyjnie (tab. 5.4), sprawdzając w każdej iteracji czy zachodzi warunek $|R_{2k}(x)| < \varepsilon$.

Tabela 5.4. Iteracyjne obliczanie $S_{2k}(x)$

k	$S_{2k}(x)$	$R_{2k}(x)$
0	1.515720	-0.105398
1	1.413968	0.000254
2	1.414220	-0.000001

5.2. Pierwszy wzór interpolacyjny Newtona

Węzły interpolacji ponumerujemy następująco x_0, x_1, \dots, x_n . Podobnie jak we wzorze Stirlinga oznaczamy

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

Wzór interpolacyjny Newtona rozpięty na węzłach x_0, x_1, \dots, x_n przedstawia się następująco:

$$\begin{aligned} P_n^I(x) = y_0 + q \cdot \Delta y_0 + \frac{q(q-1)}{2!} \cdot \Delta^2 y_0 + \frac{q(q-1)(q-2)}{3!} \cdot \Delta^3 y_0 + \dots \\ + \frac{q(q-1) \cdots (q-n+1)}{n!} \cdot \Delta^n y_0 \end{aligned} \quad (5.27)$$

Wzór interpolacji z użyciem pierwszego wzoru Newtona jest w przybliżeniu równy

$$R_n^I(x) = f(x) - P_n^I(x) \approx \frac{q(q-1)(q-2) \cdots (q-n)}{(n+1)!} \cdot \Delta^{n+1} y_0. \quad (5.28)$$

Przykład 5.5. Znane są następujące wartości funkcji $y = f(x)$

0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6
1.00000	1.14870	1.31951	1.51572	1.74110	2.00000	2.29740	2.63902	3.03143

Obliczyć $f(0.1)$ z dokładnością $\varepsilon = 10^{-5}$.

Rozwiązanie. Do obliczenia przybliżonej wartości $f(0.1)$ zastosujemy pierwszy wzór interpolacyjny Newtona. Węzły interpolacji należy tak ponumerować, aby argument $x = 0.1$ leżał w pobliżu węzła x_0 . Zatem przyjmiemy $x_0 = 0.0, x_1 = 0.2, x_2 = 0.4, \dots$ Daje to:

$$q = \frac{0.0 - 0.1}{0.2} = -0.5 .$$

W tabeli 5.5 zestawiono różnice skończone (tłustym drukiem zostały wyróżnione liczby, które wystąpią w pierwszym wzorze interpolacyjnym Newtona).

Tabela 5.5. Różnice skończone funkcji $y = f(x)$ z przykładu 5.7

i	x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$	$\Delta^5 y_i$	$\Delta^6 y_i$
0	0.0	1.00000						
			0.14870					
1	0.2	1.14870		0.02211				
			0.17081		0.00329			
2	0.4	1.31951		0.02540		0.00048		
			0.19621		0.00377		0.00010	
3	0.6	1.51572		0.02917		0.00058		-0.00005
			0.22538		0.00435		0.00005	
4	0.8	1.74110		0.03352		0.00063		0.00006
			0.25890		0.00498		0.00011	
5	1.0	2.00000		0.03850		0.00074		0.00000
			0.29740		0.00572		0.00011	
6	1.2	2.29740		0.04422		0.00085		
			0.34162		0.00657			
7	1.4	2.63902		0.05079				
			0.39241					
8	1.6	3.03143						

Tabela 5.6. Iteracyjne obliczanie $P_k^I(0.1)$

k	$P_k^I(0.1)$	$R_k^I(0.1)$
0	1.000000	0.074350
1	1.074350	-0.002764
2	1.071586	0.000206
3	1.071792	-0.000019
4	1.071773	0.000003

Wartość $P_k^I(0.1)$ obliczamy iteracyjnie (tabela 5.6), sprawdzając w każdej iteracji, czy spełnia się nierówność $|R_k^I(0.1)| < \varepsilon$.

5.3. Drugi wzór interpolacyjny Newtona

Węzły interpolacji ponumerujemy podobnie, jak w pierwszym wzorze interpolacyjnym Newtona x_0, x_1, \dots, x_n , ale przez q oznaczymy

$$q = \frac{x - x_n}{h}.$$

Drugie wzór interpolacyjny Newtona, rozpięty na węzłach x_0, x_1, \dots, x_n , przedstawia się następująco:

$$\begin{aligned} P_n^{II}(x) = y_n + q \cdot \Delta y_{n-1} + \frac{q(q+1)}{2!} \cdot \Delta^2 y_{n-2} + \frac{q(q+1)(q+2)}{3!} \cdot \Delta^3 y_{n-3} + \dots \\ + \frac{q(q+1) \cdots (q+n-1)}{n!} \cdot \Delta^n y_0. \end{aligned} \quad (5.29)$$

Blad interpolacji z użyciem drugiego wzoru Newtona jest w przybliżeniu równy

$$R_k^{II}(x) = f(x) - P_k^{II}(x) \approx \frac{q(q+1)(q+2) \cdots (q+k)}{(k+1)!} \cdot \Delta^{n+1} y_{n-(k+1)}. \quad (5.30)$$

Przykład 5.6. Znane są następujące wartości funkcji $y = f(x)$

0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6
1.00000	1.14870	1.31951	1.51572	1.74110	2.00000	2.29740	2.63902	3.03143

Obliczyć $f(1.5)$ z dokładnością $\varepsilon = 10^{-5}$.

Rozwiązanie. Do obliczenia przybliżonej wartości $f(1.5)$ zastosujemy drugi wzór interpolacyjny Newtona. Węzły interpolacyjne numerujemy w taki sposób, aby $x_n = 1.6$. Zatem

$$q = \frac{1.5 - 1.6}{0.2} = -0.5.$$

Tabela 5.7. Różnice skończone funkcji $y = f(x)$ z przykładu 5.8

x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$	$\Delta^5 y_i$	$\Delta^6 y_i$
0.0	1.00000						
		0.14870					
0.2	1.14870		0.02211				
		0.17081		0.00329			
0.4	1.31951		0.02540		0.00048		
		0.19621		0.00377		0.00010	

3	0.6	1.51572		0.02917		0.00058		-0.00005
			0.22538		0.00435		0.00005	
4	0.8	1.74110		0.03352		0.00063		0.00006
			0.25890		0.00498		0.00011	
5	1.0	2.00000		0.03850		0.00074		0.00000
			0.29740		0.00572		0.00011	
6	1.2	2.29740		0.04422		0.00085		
			0.34162		0.00657			
7	1.4	2.63902		0.05079				
			0.39241					
8	1.6	3.03143						

W tabeli 5.7 zestawiono różnice skończone funkcji $y = f(x)$, a tłustym drukiem wyróżnione zostały liczby, które wystąpią w drugim wzorze interpolacyjnym Newtona.

Wartość $P_k^{II}(1.5)$ obliczamy iteracyjnie (tabela 5.6), sprawdzając w każdej iteracji, czy zachodzi nierówność $|R_k^{II}(1.5)| < \varepsilon$.

Tabela 5.8. Iteracyjne obliczanie $P_k^{II}(1.5)$

k	$P_k^{II}(1.5)$	$R_k^{II}(1.5)$
0	3.031430	-0.196205
1	2.835225	-0.006349
2	2.828876	-0.000411
3	2.828465	-0.000033
4	2.828432	-0.000003

5.6. Interpolacja funkcjami sklejonymi

Będziemy interpolowali funkcję

x_i	x_0	x_1	...	x_n
$y_i = f(x_i)$	y_0	y_1	...	y_n

której węzły nie muszą być rozmieszczone ze stałym krokiem h . Oznaczmy

$$\Delta = \{a = x_0 < x_1 < \dots < x_n = b\}.$$

Funkcję sklejaną stopnia m odpowiadającą podziałowi Δ nazywamy funkcję rzeczywistą $S_m(\Delta) : [a ; b] \rightarrow R$, interpolującą funkcję, $y = f(x)$ która:

- na każdym podprzedziale $[x_j; x_{j+1}]$ jest wielomianem stopnia co najwyżej m (na każdym podprzedziale $[x_j; x_{j+1}]$ może być innym wielomianem),
- na całym przedziale $[a; b]$ (również w węzłach interpolacji) ma ciągłe pochodne do stopnia $m-1$ włącznie.

W praktyce wystarczy się ograniczyć się do *funkcji sklejanych stopnia trzeciego*. Podstawowe własności tych funkcji zostały sformułowane w twierdzeniu 5.1.

Twierdzenie 5.1. (Jednoznaczność funkcji sklejanej i minimalizacja krzywizny całkowitej).

Wśród funkcji interpolujących funkcję $y = f(x)$, mających ciągłe pochodne do drugiego stopnia, włącznie, funkcja sklejana stopnia trzeciego $y = S_3(\Delta)$, spełniająca jeden z trzech poniższych warunków:

$$S_3''(\Delta, a) = S_3''(\Delta, b) = 0, \quad (5.31)$$

$$f'(a) = S_3'(\Delta, a), \quad f'(b) = S_3'(\Delta, b), \quad (5.32)$$

$$y = f(x) \text{ jest funkcją okresową,} \quad (5.33)$$

minimalizuje krzywiznę całkowitą

$$\int_a^b |f''(x)|^2 dx .$$

Zauważmy:

$$h_{j+1} = x_{j+1} - x_j \quad \text{dla } j = 0, 1, \dots, n-1,$$

$$M_j = S_3''(\Delta, x_j) \quad \text{dla } j = 0, 1, \dots, n.$$

Konstrukcja funkcji $y = S_3(\Delta, x)$ na przedziale $[x_j; x_{j+1}]$, dla ustalonego j ($0 \leq j \leq n-1$), poczyna się od spostrzeżenia, że $y = S_3''(\Delta, x)$ powinna być funkcją liniową (ponieważ $S_3(\Delta, x)$ ma być wielomianem trzeciego stopnia). Budując na węzłach x_j oraz x_{j+1} wielomian interpolacyjny Lagrange'a, mamy

$$S_3''(\Delta, x) = M_j \frac{(x - x_{j+1})}{(x_j - x_{j+1})} + M_{j+1} \frac{(x - x_j)}{(x_{j+1} - x_j)} = M_j \frac{(x_{j+1} - x)}{h_{j+1}} + M_{j+1} \frac{(x - x_j)}{h_{j+1}} .$$

Całkując powyższą równość dwukrotnie, otrzymujemy

$$S_3(\Delta, x) = M_j \frac{(x_{j+1} - x)^3}{6h_{j+1}} + M_{j+1} \frac{(x - x_j)^3}{6h_{j+1}} + A_j(x - x_j) + B_j .$$

Stąd A_j, B_j wyznaczamy, rozwiązuając układ równań:

$$S_3(\Delta, x_j) = M_j \frac{h_{j+1}^2}{6} + B_j = y_j ,$$

$$S_3(\Delta, x_{j+1}) = M_{j+1} \frac{h_{j+1}^2}{6} + A_j h_{j+1} + B_j = y_{j+1} ,$$

uwierciedlający żądanie, aby funkcja $y = S_3(\Delta, x)$ była na przedziale $[x_j; x_{j+1}]$ funkcją interpolacyjną. Ostatecznie otrzymujemy wzory wyrażające $y = S_3(\Delta, x)$ oraz jej pochodną jako funkcję nieznanego parametrów M_0, M_1, \dots, M_n :

$$S(\Delta, x) = M_j \frac{(x_{j+1} - x)^3}{6h_{j+1}} + M_{j+1} \frac{(x - x_j)^3}{6h_{j+1}} + \\ + \left(\frac{y_{j+1} - y_j}{h_{j+1}} - \frac{h_{j+1}}{6}(M_{j+1} - M_j) \right)(x - x_j) + y_j - M_j \frac{h_{j+1}^2}{6},$$

$$S'_3(\Delta, x) = -M_j \frac{(x_{j+1} - x)^2}{2h_{j+1}} + M_{j+1} \frac{(x - x_j)^2}{2h_{j+1}} + \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{h_{j+1}}{6}(M_{j+1} - M_j).$$

Dla wyznaczenia wartości parametrów M_1, M_2, \dots, M_{n-1} , formułujemy warunek ciągłości pierwszych pochodnych $y = S'_3(\Delta, x)$ w węzłach x_1, x_2, \dots, x_{n-1} , tj.

$$S'_3(\Delta, x_{j-}) = S'_3(\Delta, x_{j+}) \text{ dla } j = 1, 2, \dots, n-1.$$

Wartości parametrów M_0, M_n określają warunki (5.31) lub (5.32). Jeśli zakładamy spełnienie warunków (5.31), to mamy

$$M_0 = M_n = 0.$$

Jeśli zakładamy spełnienie warunków (5.32), to mamy związki:

$$f'(a) = S'_3(\Delta, a) = S'_3(\Delta, x_{0+}) = \frac{y_1 - y_0}{h_1} - \frac{h_1}{3}M_0 - \frac{h_1}{6}M_1,$$

$$f'(b) = S'_3(\Delta, b) = S'_3(\Delta, x_{n-}) = \frac{y_n - y_{n-1}}{h_n} + \frac{h_n}{3}M_n + \frac{h_n}{6}M_{n-1},$$

a z nich otrzymujemy wzory na wartości poszukiwanych parametrów:

$$M_0 = \frac{3}{h_1^2}(y_1 - y_0) - \frac{3}{h_1}f'(a) - \frac{M_1}{2},$$

$$M_n = \frac{3}{h_n^2}f'(b) - \frac{3}{h_n}(y_n - y_{n-1}) - \frac{M_{n-1}}{2}.$$

Algorytm interpolowania funkcjami sklejonymi trzeciego stopnia

1. Wyznaczamy wartości parametrów M_0, M_1, \dots, M_n przez rozwiązywanie układu równań

$$\begin{aligned} 2M_0 + \lambda_0 M_1 &= d_0, \\ \mu_1 M_0 + 2M_1 + \lambda_1 M_2 &= d_1, \\ \mu_2 M_1 + 2M_2 + \lambda_2 M_3 &= d_2, \\ \dots & \\ \mu_{n-1} M_{n-2} + 2M_{n-1} + \lambda_{n-1} M_n &= d_{n-1}, \\ \mu_n M_{n-1} + 2M_n &= d_n, \end{aligned} \tag{5.34}$$

gdzie dla $j = 1, 2, \dots, n-1$,

$$\lambda_j = \frac{h_j}{h_j + h_{j+1}}, \quad \lambda_{j+1} = \frac{h_{j+1}}{h_j + h_{j+1}}, \quad d_j = \frac{6}{h_j + h_{j+1}} \left(\frac{y_{j+1} - y_j}{h_{j+1}} - \frac{y_j - y_{j-1}}{h_j} \right)$$

— gdy zachodzą warunki (5.31), to:

$$\lambda_0 = 0, d_0 = 0,$$

$$\lambda_n = 0, d_n = 0,$$

— zachodzą warunki (5.32), to

$$\lambda_0 = 1, \quad d_0 = \frac{6}{h_1} \left(\frac{y_1 - y_0}{h_1} - f'(a) \right),$$

$$\lambda_n = 1, \quad d_n = \frac{6}{h_n} \left(f'(b) - \frac{y_n - y_{n-1}}{h_n} \right).$$

■ Dla $x \in [x_j; x_{j+1}]$ funkcję $y = S_3(\Delta, x)$ wyznaczamy ze wzoru

$$S_3(\Delta, x) = \alpha_j + \beta_j(x - x_j) + \gamma_j(x - x_j)^2 + \delta_j(x - x_j)^3, \quad (5.35)$$

$$\alpha_j = S_3(\Delta, x_j) = y_j,$$

$$\beta_j = \frac{S'_3(\Delta, x_j)}{1!} = \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{2M_j + M_{j+1}}{6} h_{j+1},$$

$$\gamma_j = \frac{S''_3(\Delta, x_j)}{2!} = \frac{M_j}{2},$$

$$\delta_j = \frac{S'''_3(\Delta, x_j)}{3!} = \frac{M_{j+1} - M_j}{6h_{j+1}}.$$

Zwróćmy uwagę, że układ równań liniowych (5.34) jest układem dobrze uwarunkowanym i że dla każdego równania j zachodzi warunek $2 > \lambda_j + \mu_j = 1$.

Jakość przybliżenia funkcji $y = f(x)$ przez $y = S_3(\Delta, x)$ można ocenić na podstawie nierówności

$$|f(x) - S_3(\Delta, x)| \leq \frac{7}{8} KM_4 \|\Delta\|^4, \quad (5.36)$$

$$\|\Delta\| = \max_{0 \leq j \leq n-1} |x_{j+1} - x_j|, \quad K = \max_{0 \leq j \leq n-1} \frac{\|\Delta\|}{|x_{j+1} - x_j|}, \quad M_4 = \max_{a \leq x \leq b} |f^{(IV)}(x)|.$$

Ze nierówności (5.36) wynika, że błąd bezwzględny przybliżenia możemy zmniejszać przez zmniejszanie wielkości $\|\Delta\|$.

Przykład 5.7. Skonstruować $y = S_3(\Delta, x)$ dla funkcji

x_j	$\frac{1}{2}$	1	$\frac{3}{2}$	2
$y_j = f(x_j)$	2	1	$\frac{2}{3}$	$\frac{1}{2}$

wiedząc ponadto, że $f\left(\frac{1}{2}\right) = -4$, $f'(2) = -\frac{1}{4}$.

Rozwiązanie. Węzły interpolacji są rozmieszczone ze stałym krokiem $h_1 = h_2 = h_3 = 0.5$. Dane do obliczeń zostały zestawione w tabeli 5.9.

Tabela 5.9. Podstawowe wielkości potrzebne do wyznaczenia $y = S_3(\Delta, x)$

j	0	1	2	3
x_j	$\frac{1}{2}$	1	$\frac{3}{2}$	2
$y_j = f(x_j)$	2	1	$\frac{2}{3}$	$\frac{1}{2}$
$f'(x_j)$	-4			$-\frac{1}{4}$

Najpierw wyznaczamy wartości parametrów M_0, M_1, M_2, M_3 z układu równań (5.3-), który w tym przypadku przybiera postać

$$2M_0 + M_1 = \frac{6}{1} \cdot \left(\frac{\frac{1}{2} - 2}{\frac{1}{2}} - (-4) \right) = 24,$$

$$\frac{1}{2}M_0 + 2M_1 + \frac{1}{2}M_2 = \frac{6}{1} \cdot \left(\frac{\frac{2}{3} - 1}{\frac{1}{2}} - \frac{1 - 2}{\frac{1}{2}} \right) = 8,$$

$$\frac{1}{2}M_1 + 2M_2 + \frac{1}{2}M_3 = \frac{6}{1} \cdot \left(\frac{\frac{1}{2} - \frac{2}{3}}{\frac{1}{2}} - \frac{\frac{2}{3} - 1}{\frac{1}{2}} \right) = 2,$$

$$M_2 + 2M_3 = \frac{6}{1} \cdot \left(-\frac{1}{4} - \frac{\frac{1}{2} - \frac{2}{3}}{\frac{1}{2}} \right) = 1.$$

Rozwiązaniem tego układu jest:

$$M_0 = \frac{173}{15}, \quad M_1 = \frac{14}{15}, \quad M_2 = \frac{11}{15}, \quad M_3 = \frac{2}{15}.$$

Zatem wyznaczamy $y = S_3(\Delta, x)$ dla poszczególnych podprzedziałów.

■ Dla $x \in \left[\frac{1}{2}; 1\right]$ kolejno mamy:

$$\alpha_0 = 2,$$

$$\beta_0 = \frac{\frac{1-2}{2} - \left(2 \cdot \frac{173}{15} + \frac{14}{15}\right)}{6} \cdot \frac{1}{2} = -4,$$

$$\gamma_0 = \frac{1}{2} \cdot \frac{173}{15} = \frac{173}{30},$$

$$\delta_0 = \frac{\frac{14}{15} - \frac{173}{15}}{6 \cdot \frac{1}{2}} = -\frac{53}{15}.$$

Zatem dla $x \in \left[\frac{1}{2}; 1\right]$ mamy

$$S_3(\Delta, x) = 2 - 4 \cdot \left(x - \frac{1}{2}\right) + \frac{173}{30} \left(x - \frac{1}{2}\right)^2 - \frac{53}{15} \left(x - \frac{1}{2}\right)^3.$$

■ Dla $x \in \left[1; \frac{3}{2}\right]$ kolejno mamy

$$\alpha_1 = 1,$$

$$\beta_1 = \frac{\frac{3}{2}-1}{\frac{1}{2}} - \frac{\left(2 \cdot \frac{14}{15} + \frac{11}{15}\right)}{6} \cdot \frac{1}{2} = -\frac{53}{60},$$

$$\gamma_1 = \frac{1}{2} \cdot \frac{14}{15} = \frac{7}{15},$$

$$\delta_1 = \frac{\frac{11}{15} - \frac{14}{15}}{6 \cdot \frac{1}{2}} = -\frac{1}{15}.$$

Zatem dla $x \in \left[1; \frac{3}{2}\right]$ mamy

$$S_3(\Delta, x) = 1 - \frac{53}{60}(x-1) + \frac{7}{15}(x-1)^2 - \frac{1}{15}(x-1)^3.$$

3. Dla $x \in \left[\frac{3}{2}; 2\right]$ kolejno mamy

$$\alpha_2 = \frac{2}{3},$$

$$\beta_2 = \frac{\frac{1}{2} - \frac{2}{3}}{\frac{1}{2}} - \frac{\left(2 \cdot \frac{11}{15} + \frac{2}{15}\right)}{6} \cdot \frac{1}{2} = -\frac{7}{15},$$

$$\gamma_2 = \frac{1}{2} \cdot \frac{11}{15} = \frac{11}{30},$$

$$\delta_2 = \frac{\frac{2}{15} - \frac{11}{15}}{6 \cdot \frac{1}{2}} = -\frac{1}{5}.$$

Zatem dla $x \in \left[\frac{3}{2}; 2\right]$ mamy

$$S_3(\Delta, x) = \frac{2}{3} - \frac{7}{15} \left(x - \frac{3}{2}\right) + \frac{11}{30} \left(x - \frac{3}{2}\right)^2 - \frac{1}{5} \left(x - \frac{3}{2}\right)^3.$$

Poszukiwaną funkcją jest

$$S_3(\Delta, x) = \begin{cases} 2 - 4 \cdot \left(x - \frac{1}{2}\right) + \frac{173}{30} \left(x - \frac{1}{2}\right)^2 - \frac{53}{15} \left(x - \frac{1}{2}\right)^3 & \text{dla } \frac{1}{2} \leq x \leq 1 \\ 1 - \frac{53}{60} (x-1) + \frac{7}{15} (x-1)^2 - \frac{1}{15} (x-1)^3 & \text{dla } 1 \leq x \leq \frac{3}{2}, \\ \frac{2}{3} - \frac{7}{15} \left(x - \frac{3}{2}\right) + \frac{11}{30} \left(x - \frac{3}{2}\right)^2 - \frac{1}{5} \left(x - \frac{3}{2}\right)^3 & \text{dla } \frac{3}{2} \leq x \leq 2 \end{cases}$$

Obecnie spróbujemy ocenić przybliżenie funkcji interpolowanej $f(x) = 1/x$ na odcinku $[0.5; 2]$ za pomocą funkcji (5.36). Mamy

$$\|\Delta\| = \frac{1}{2}, \quad K = 1, \quad M_4 = \max_{\frac{1}{2} \leq x \leq 2} \left| \frac{24}{x^5} \right| = 24 \cdot 2^5,$$

$$|f(x) - S_3(\Delta, x)| \leq \frac{7}{8} \cdot 1 \cdot 24 \cdot 2^5 \cdot \frac{1}{2^4} = 42 \quad \text{dla } \frac{1}{2} \leq x \leq 2.$$

Obliczone szacowanie (5.58) jest bardzo silne. Przekonuje o tym tablica 5.10 wartości funkcji interpolowanej i interpolującej.

Tabela 5.10. Tablica wartości funkcji interpolowanej i interpolującej

x	$f(x) = 1/x$	$S_3(\Delta, x)$	$f(x) - S_3(\Delta, x)$
0,5	2,0000	2,0000	0,0000
0,6	1,6667	1,6541	0,0125
0,7	1,4286	1,4024	0,0262
0,8	1,2500	1,2236	0,0264
0,9	1,1111	1,0965	0,0146
1,0	1,0000	1,0000	0,0000
1,1	0,9091	0,9163	-0,0072
1,2	0,8333	0,8415	-0,0081
1,3	0,7692	0,7752	-0,0060
1,4	0,7143	0,7171	-0,0028
1,5	0,6667	0,6667	0,0000
1,6	0,6250	0,6235	0,0015
1,7	0,5882	0,5864	0,0018
1,8	0,5556	0,5543	0,0013
1,9	0,5263	0,5259	0,0004
2,0	0,5000	0,5000	0,0000

5.7. Inne zagadnienia interpolacyjne

5.7.1. Zagadnienie interpolacyjne Taylora

Zakładamy, że funkcja interpolowana $y = f(x)$ jest określona w pewnym otoczeniu argumentu $x = x_0$ oraz ma pochodne rzędów aż do $n+1$ włącznie w tym otoczeniu. Znane są wartości

$$f^{(k)}(x_0) = y_0^{(k)} \quad \text{dla } k = 0, 1, \dots, n \quad (f^{(0)}(x_0) = f(x_0))$$

funkcji $y = f(x)$ i jej n pochodnych dla $x = x_0$. Należy znaleźć funkcję $y = F(x)$, ciągłą i mającą w otoczeniu argumentu $x = x_0$ ciągłe pochodne do rzędu n włącznie, która spełnia $n+1$ równości

$$F^{(k)}(x_0) = y_0^{(k)} \quad \text{dla } k = 0, 1, \dots, n \quad (F^{(0)}(x_0) = F(x_0)).$$

Rozwiązaniem tego zagadnienia jest znany z analizy matematycznej wzór Taylora

$$F(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \quad (5.37)$$

Reszta wzoru interpolacyjnego Taylora jest równa

$$R(x) = f(x) - F(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!}(x - x_0)^{n+1}, \quad (5.38)$$

gdzie $x_0 < \theta < x$ lub $x < \theta < x_0$.

Jeżeli $x_0 = 0$, to wzór Taylora nosi nazwę *wzoru Maclaurina*. Zatem wzór Maclaurina ma postać

$$F(x) = f(x_0) + \frac{f'(0)}{1!} \cdot x + \frac{f''(0)}{2!} \cdot x^2 + \dots + \frac{f^{(n)}(0)}{n!} \cdot x^n \quad (5.39)$$

Reszta wzoru Maclaurina jest równa

$$R(x) = f(x) - F(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!} \cdot x^{n+1}, \quad (5.40)$$

gdzie $0 < \theta < x$ lub $x < \theta < 0$. W dodatku E zostały podane rozwinięcia w szereg Maclaurina często stosowanych funkcji elementarnych.

Wzór Taylora (Maclaurina) jest dobrym narzędziem do obliczania wartości funkcji $y = f(x)$ dla argumentów bliskich x_0 .

5.7.2. Zagadnienie interpolacyjne Hermite'a

Zakładamy, że funkcja interpolowana $y = f(x)$ jest określona i p -krotnie ($p \geq 0$) różniczkowalna w przedziale $[a ; b]$ zawierającym węzły interpolacji x_0, x_1, \dots, x_n .

Dla każdego k ($k = 0, 1, \dots, n$) znane są wartości

$$f^{(s)}(x_k) = y_k^{(s)} \quad (s = 0, 1, \dots, d_k)$$

(funkcji $y = f(x)$ i jej d_k początkowych pochodnych dla argumentu $x = x_k$; d_k są liczbami całkowitymi nieujemnymi). Należy znaleźć wielomian $y = H(x)$ spełniający $d_0 + d_1 + \dots + d_n + n + 1$ równości

$$H^{(s)}(x_k) = y_k^{(s)} \quad (k = 0, 1, \dots, n ; s = 0, 1, \dots, d_k).$$

Rozwiązaniem powyższego zagadnienia jest *wzór interpolacyjny Hermite'a*

$$H_n(x) = \sum_{k=0}^n \sum_{s=0}^{d_k} y_k^{(s)} \frac{B(x)}{s!(x - x_k)^{d_k+1}} \sum_{i=0}^{d_k-s} c_{ki} (x - x_k)^{i+s}, \quad (5.41)$$

gdzie

$$B(x) = (x - x_0)^{d_0+1} (x - x_1)^{d_1+1} \cdot \dots \cdot (x - x_n)^{d_n+1}.$$

Współczynniki c_{ki} we wzorze (5.41) wyznaczamy z tożsamości

$$\frac{(x - x_k)^{d_k+1}}{B(x)} = \sum_{i=0}^{\infty} c_{ki} (x - x_k)^i \quad \text{dla } k = 0, 1, \dots, n. \quad (5.42)$$

Reszta wzoru interpolacyjnego Hermite'a jest równa

$$R(x) = f(x) - H(x) = (x - x_0)^{d_0+1} (x - x_1)^{d_1+1} \cdot \dots \cdot (x - x_n)^{d_n+1} \frac{f^{(n+1)}(\theta)}{(n+1)!}, \quad (5.43)$$

dla pewnego $\theta \in (a ; b)$. Zagadnienie interpolacyjne Hermite'a jest naturalnym uogólnieniem zagadnień interpolacyjnych Lagrange'a i Taylora.

Rozpatrzmy dwa szczególne przypadki wzoru interpolacyjnego Hermita. W pierwszym przypadku założmy, że $d_0 = d_1 = \dots = d_n = 0$ (chcemy interpolować w sensie Lagrange'a). Wykorzystując oznaczenie (5.10), mamy

$$B(x) = (x - x_0)(x - x_1) \cdot \dots \cdot (x - x_n) = \omega_{n+1}(x) = \omega(x).$$

Współczynniki c_{k0} ($k = 0, 1, \dots, n$) obliczamy ze wzoru (5.42), przyjmując $d_k = 0$, a następnie podstawiając $x = x_k$. Ponieważ lewa strona wzoru (5.42) jest wyrażeniem nieoznaczonym, stosujemy jednokrotnie regułę l'Hospitala, otrzymując

$$\frac{1}{\omega'(x)} = \sum_{i=0}^{\infty} c_{ki} (x - x_k)^i, \quad k = 0, 1, \dots, n.$$

Mamy

$$c_{k0} = \frac{1}{\omega'(x_k)}, \quad k = 0, 1, \dots, n.$$

Wstawiając obliczone współczynniki do wzoru (5.41), ostatecznie otrzymujemy

$$H_n(x) = \sum_{k=0}^n f(x_k) \cdot \frac{\omega(x)}{(x - x_k)\omega'(x_k)}.$$

Jak należało oczekiwać, otrzymaliśmy wzór interpolacyjny Lagrange'a.

Rozpatrzmy drugi przypadek, kiedy $d_0 = d_1 = \dots = d_n = 1$. Współczynniki c_{ki} ($k = 0, 1, \dots, n$) wyznaczamy z tożsamości

$$\frac{(x - x_k)^2}{\omega^2(x)} = \sum_{i=0}^{\infty} c_{ki} (x - x_k)^i \quad \text{dla } k = 0, 1, \dots, n. \quad (5.44)$$

Pierw dwukrotnie stosujemy regułę l'Hospitala do lewej strony tożsamości (5.44), a następnie podstawiamy $x = x_k$. W konsekwencji otrzymujemy

$$c_{k0} = \frac{1}{\omega'^2(x_k)}, \quad k = 0, 1, \dots, n.$$

Następnie różniczkujemy obie strony tożsamości (5.44), trzykrotnie stosujemy regułę l'Hospitala do lewej strony i podstawiamy $x = x_k$. Ostatecznie otrzymujemy

$$c_{k1} = -\frac{\omega''(x_k)}{\omega'^2(x_k)}, \quad k = 0, 1, \dots, n.$$

Podstawiając obliczone współczynniki do wzoru (5.44), otrzymujemy wzór

$$\begin{aligned} H(x) &= \sum_{k=0}^n f(x_k) \cdot \frac{\omega^2(x)}{(x - x_k)^2 \omega'^2(x_k)} \cdot \left(1 - \frac{\omega''(x_k)}{\omega'(x_k)} (x - x_k) \right) + \\ &\quad + \sum_{k=0}^n f'(x_k) \cdot \frac{\omega^2(x)}{(x - x_k)^2 \omega'^2(x_k)} \cdot (x - x_k) \end{aligned} \quad (5.45)$$

nazywany wzorem interpolacyjnym Hermite'a o węzłach dwukrotnych.

Przykład 5.8. Znane są następujące wartości funkcji $y = f(x)$

x_i	0.000000	0.500000	1.000000	1.500000
$y_i = f(x_i)$	1.000000	0.877583	0.540302	0.070737
$y'_i = f'(x_i)$	0.000000	-0.479426	-0.841471	-0.997495

Dla powyższej funkcji wyznaczyć wielomian interpolacyjny Lagrange'a $L_3(x)$ oraz wielomian interpolacyjny Hermite'a o węzłach dwukrotnych. Pokazać różnice w przybliżeniu funkcji.

Rozwiązanie. Wielomian $y = L_3(x)$ konstrujemy według wzoru (5.9). Wielomian Hermite'a $y = H_3(x)$ według wzoru (5.45). Dla zobrazowania jakości przybliżania funkcji $y = f(x)$ wielomianami interpolacyjnymi stablicowano wielomiany $L_3(x)$, $H_3(x)$ oraz różnice $f(x) - L_3(x)$, $f(x) - H_3(x)$ w przedziale $[0.0 ; 1.5]$, a wyniki umieszczone w tabeli 5.11.

Tabela 5.11. Tablica wartości funkcji $y = L_3(x)$ i $y = H_3(x)$ oraz odchyлеń

i	x_i	$L_3(x_i)$	$f(x_i) - L_3(x_i)$	$H_3(x_i)$	$f(x_i) - H_3(x_i)$
0	0.000	1.000000	0.0E+00	1.000000	0.0E+00
1	0.125	0.994055	-1.9E-03	0.992198	6.0E-08
2	0.250	0.970810	-1.9E-03	0.968912	6.4E-08
3	0.375	0.931556	-1.0E-03	0.930508	2.0E-08
4	0.500	0.877583	0.0E+00	0.877583	0.0E+00
5	0.625	0.810180	7.8E-04	0.810963	1.2E-08
6	0.750	0.730639	1.0E-03	0.731689	2.2E-08
7	0.875	0.640250	7.5E-04	0.640997	1.2E-08
8	1.000	0.540302	0.0E+00	0.540302	0.0E+00
9	1.125	0.432087	-9.1E-04	0.431176	1.9E-08
10	1.250	0.316894	-1.6E-03	0.315322	5.8E-08
11	1.375	0.196014	-1.5E-03	0.194548	5.3E-08
12	1.500	0.070737	0.0E+00	0.070737	0.0E+00

5.7.3. Zagadnienie interpolacyjne Abela-Gonczarowa

Zakładamy, że funkcja interpolowana $y = f(x)$ jest określona i ma pochodne aż do rzędu $n+1$ w przedziale $[a ; b]$, zawierającym węzły interpolacji x_0, x_1, \dots, x_n . Znane są wartości

$$f^{(k)}(x_k) = y_k^{(k)} \quad \text{dla } k = 0, 1, \dots, n.$$

Należy znaleźć wielomian $P(x)$ spełniający następujące warunki:

$$P^{(k)}(x_k) = y_k^{(k)} \quad \text{dla } k = 0, 1, \dots, n.$$

Rozwiązańiem tego zagadnienia jest wzór interpolacyjny Abela-Gonczarowa

$$P_n(x) = \sum_{k=0}^n y_k^{(k)} A_k(x), \quad (5.46)$$

$$A_0(x) = 1,$$

$$A_k(x) = \int_{x_0}^x dt_1 \int_{x_1}^{t_1} dt_2 \cdots \int_{x_{k-1}}^{t_{k-1}} dt_k \quad \text{dla } k = 1, 2, \dots, n. \quad (5.47)$$

Wielomian interpolacji z użyciem wzoru Abela-Gonczarowa jest równy

$$R(x) = f(x) - P(x) = \int_{x_0}^x dt_1 \int_{x_1}^{t_1} dt_2 \cdots \int_{x_n}^{t_n} f^{(n+1)}(t_{k+1}) dt_{k+1}. \quad (5.48)$$

Zadanie 5.9. Na podstawie danych:

$$f(0.0) = 1.000000,$$

$$f(0.5) = -0.479426,$$

$$f'(1.0) = -0.540302,$$

$$f''(1.5) = 0.997495$$

wyznaczyć wielomian interpolacyjny Abela-Gonczarowa.

Rozwiązanie. Na podstawie powyższych danych możemy wyznaczyć wielomian interpolacyjny Abela-Gonczarowa stopnia 3

$$P_3(x) = f(0.0) A_0(x) + f'(0.5) A_1(x) + f''(1.0) A_2(x) + f'''(1.5) A_3(x).$$

Wielomiany $A_k(x)$ wyznaczamy ze wzoru (5.45):

$$A_0(x) = 1,$$

$$A_1(x) = x - x_0, \quad A_1(x) = x - x_0,$$

$$A_2(x) = \frac{(x^2 - x_0^2)}{2!} - x_1(x - x_0),$$

$$A_3(x) = \frac{(x^3 - x_0^3)}{3!} - \frac{x_2}{2!}(x^2 - x_0^2) + \left(x_1 x_2 - \frac{x_1^2}{2} \right) (x - x_0).$$

$$P_3(x) = 1.000000 - 0.479426 \cdot x - 0.540302 \cdot \frac{x^2 - x}{2} + 0.997495 \cdot \left(\frac{x^3}{3!} - \frac{x^2}{2!} + \frac{3}{8} x \right).$$

W tabeli 5.12 podano wybrane wartości wielomianu $y = P_3(x)$ w przedziale $[0.0 ; 1.5]$.

Tabela 5.12. Wartości wielomianu interpolacyjnego Abela-Gonczarowa $y = P_3(x)$

i	x_i	$A_0(x_i)$	$A_1(x_i)$	$A_2(x_i)$	$A_3(x_i)$	$P_3(x_i)$	$f(x_i) - P_3(x_i)$
0	0.000	1.000000	0.000000	0.000000	0.000000	1.000000	0.0E+00
1	0.125	1.000000	0.125000	-0.054688	0.039388	1.008909	-1.7E-02
2	0.250	1.000000	0.250000	-0.093750	0.065104	0.995738	-2.7E-02
3	0.375	1.000000	0.375000	-0.117188	0.079102	0.962436	-3.2E-02
4	0.500	1.000000	0.500000	-0.125000	0.083333	0.910950	-3.3E-02
5	0.625	1.000000	0.625000	-0.117188	0.079753	0.843229	-3.2E-02
6	0.750	1.000000	0.750000	-0.093750	0.070313	0.761221	-3.0E-02
7	0.875	1.000000	0.875000	-0.054688	0.056966	0.666874	-2.6E-02
8	1.000	1.000000	1.000000	0.000000	0.041667	0.562137	-2.2E-02
9	1.125	1.000000	1.125000	0.070313	0.026367	0.448957	-1.8E-02
10	1.250	1.000000	1.250000	0.156250	0.013021	0.329284	-1.4E-02
11	1.375	1.000000	1.375000	0.257813	0.003581	0.205065	-1.1E-02
12	1.500	1.000000	1.500000	0.375000	0.000000	0.078248	-7.5E-03

5.8. Zadania

Zad. 5.1. Zbudować wielomiany interpolacyjne dla następujących funkcji:

x_i	1	2	3
$y_i = f_1(x_i)$	3	1	2

x_i	0	0.50	1	2
$y_i = f_2(x_i)$	0	0.25	1	-1

x_i	0.0	0.5	1.0	2.0	3.0
$y_i = f_3(x_i)$	0.0	4.5	4.0	0.0	12.0

x_i	0.0	0.2	0.4	0.6	0.8
$y_i = f_4(x_i)$	0.50	0.32	0.18	0.08	0.02

x_i	0.0	0.5	1.0	1.5	2.0
$y_i = f_5(x_i)$	0.75	1.00	2.75	7.50	16.75

Zad. 5.2. W poniższej tabeli podane są wartości funkcji $y = \sin x$ dla wybranych argumentów

x_i	0.0	0.3	0.6	0.9	1.2	1.5
$y_i = \sin x_i$	0.0000	0.2955	0.5646	0.7833	0.9320	0.9975

- Zbudować wielomian interpolacyjny opierając się na wartościach zawartych w tabeli.
- Oceneć dokładność, z jaką można obliczyć przybliżoną wartość funkcji $y = \sin x$ z wykorzystaniem wielomianu interpolacyjnego.
- Korzystając z wielomianu interpolacyjnego, obliczyć: $\sin 0.1$, $\sin 0.5$, $\sin 1.0$, $\sin(\pi/2)$.

Zad. 5.3. Zbudować wielomiany interpolacyjne $y = s_3(x)$ dla następujących funkcji

x_i	1	2	3
$y_i = f_1(x_i)$	2	1	5

x_i	1	2	3
$y_i = f_2(x_i)$	2	1	5

x_i	1.0	1.5	2.0	3.0
$y_i = f_3(x_i)$	1	3	0	-2

x_i	0.2	0.4	0.6	0.8
$y_i = f_4(x_i)$	0.68	1.52	2.52	3.68

Zad. 5.4. Czy poniższa funkcja

$$s(x) = \begin{cases} (x-1)^3 & ; 0 \leq x \leq 1 \\ 2(x-1)^3 & ; 1 \leq x \leq 2 \end{cases}$$

jest funkcją sklejaną trzeciego stopnia na przedziale $[0 ; 2]$?

Zad. 5.5. Czy poniższa funkcja

$$s(x) = \begin{cases} x^3 - 3x^2 + 2x + 1 & ; 1 \leq x \leq 2 \\ -x^3 + 9x^2 - 22x + 17 & ; 2 \leq x \leq 3 \end{cases}$$

jest funkcją sklejaną trzeciego stopnia na przedziale $[1 ; 3]$?

Zad. 5.6. Dany jest przedział $[a ; b]$ oraz liczba c spełniająca warunek $a < c < b$. Wykazać, że funkcja

$$s(x) = \begin{cases} 0 & ; a \leq x \leq c \\ (x-c)^3 & ; c \leq x \leq b \end{cases}$$

jest funkcją sklejaną trzeciego stopnia na przedziale $[a ; b]$.

Zad. 5.7. Dysponujemy kalkulatorem umożliwiającym wykonywanie tylko czterech podstawowych działań arytmetycznych. Zastosuj odpowiednie wzory interpolacyjne do obliczenia wartości poniższych funkcji. Obliczenie wartości funkcji uzupełnij obliczeniem o takim jakim jest ona obarczona.

Wiedząc, że $\sqrt{4} = 2$, $\sqrt{9} = 3$, $\sqrt{16} = 4$

wyznaczyć $\sqrt{10}$.

5.7.2. Wiedząc, że: $\sqrt{100} = 10$, $\sqrt{121} = 11$, $\sqrt{144} = 12$, $\sqrt{169} = 13$

obliczyć: a) $\sqrt{130}$, b) $\sqrt{160}$.

5.1.3. Wiedząc, że $\frac{1}{\sqrt{4}} = \frac{1}{2}$, $\frac{1}{\sqrt{9}} = \frac{1}{3}$, $\frac{1}{\sqrt{16}} = \frac{1}{4}$, $\frac{1}{\sqrt{25}} = \frac{1}{5}$

obliczyć: a) $\frac{1}{\sqrt{5}}$, b) $\frac{1}{\sqrt{20}}$.

5.1.4. Wiedząc, że $\sqrt[3]{\frac{1}{8}} = \frac{1}{2}$, $\sqrt[3]{\frac{8}{27}} = \frac{2}{3}$, $\sqrt[3]{1} = 1$, $\sqrt[3]{\frac{27}{8}} = \frac{3}{2}$

obliczyć: a) $\sqrt[3]{2}$, b) $\sqrt[3]{3}$.

5.1.5. Wiedząc, że $\lg \frac{1}{2} = -1$, $\lg 1 = 0$, $\lg 2 = 1$, $\lg 4 = 2$

obliczyć: a) $\lg 3$, b) $\lg 5$.

5.1.6. Wiedząc, że $\ln \frac{1}{e} = -1$, $\lg 1 = 0$, $\ln e = 1$, $\ln e^2 = 2$

obliczyć: a) $\ln 2$, b) $\ln 8$.

5.1.7. Wiedząc, że $\log 10 = 1$, $\log 100 = 2$, $\log 1000 = 3$

obliczyć: a) $\log 30$, b) $\log 85$.

5.1.8. Wiedząc, że $\log \frac{1}{10} = -1$, $\log 1 = 0$, $\log 10 = 1$, $\log 100 = 2$

obliczyć: a) $\log 2$, b) $\log 75$.

5.1.9. Wiedząc, że $2^0 = 1$, $2^1 = 2$, $2^2 = 4$, $2^3 = 8$

obliczyć: a) $\sqrt{2}$, b) $\sqrt[3]{4}$.

5.1.10. Wiedząc, że $3^{-1} = \frac{1}{3}$, $3^0 = 1$, $3^1 = 3$, $3^2 = 9$,

obliczyć: a) $\sqrt[4]{3}$, b) $\sqrt{3}$.

5.1.11. Znając e^{-1} , e^0 , e , e^2 ,

obliczyć \sqrt{e} .

5.1.12. Wiedząc, że

$10^{-1} = \frac{1}{10}$, $10^0 = 1$, $10^1 = 10$, $10^2 = 100$,

obliczyć: a) $\sqrt{10}$, b) $\frac{1}{\sqrt{10}}$.

E1.13. Wiedząc, że $\sin \frac{\pi}{12} = \frac{\sqrt{6}-\sqrt{2}}{4}$, $\sin \frac{\pi}{8} = \frac{\sqrt{2}-\sqrt{2}}{2}$, $\sin \frac{\pi}{6} = \frac{1}{2}$

wiećzyć $\sin \frac{\pi}{9}$.

E1.14. Wiedząc, że $\cos \frac{\pi}{12} = \frac{\sqrt{6}+\sqrt{2}}{4}$, $\cos \frac{\pi}{8} = \frac{\sqrt{2}+\sqrt{2}}{2}$, $\cos \frac{\pi}{6} = \frac{\sqrt{3}}{2}$

wiećzyć $\cos \frac{\pi}{9}$.

E1.15. Wiedząc, że $\operatorname{tg} \frac{\pi}{12} = 2 - \sqrt{3}$, $\operatorname{tg} \frac{\pi}{8} = \sqrt{2} - 1$, $\operatorname{tg} \frac{\pi}{6} = \frac{\sqrt{3}}{3}$

wiećzyć $\operatorname{tg} \frac{\pi}{9}$.

E1.16. Wiedząc, że $\operatorname{ctg} \frac{\pi}{12} = 2 + \sqrt{3}$, $\operatorname{ctg} \frac{\pi}{8} = \sqrt{2} + 1$, $\operatorname{ctg} \frac{\pi}{6} = \sqrt{3}$

wiećzyć $\operatorname{ctg} \frac{\pi}{9}$.

E1.17. Wiedząc, że $\sin \frac{\pi}{8} = \frac{\sqrt{2}-\sqrt{2}}{2}$, $\sin \frac{\pi}{6} = \frac{1}{2}$, $\sin \frac{\pi}{4} = \frac{\sqrt{2}}{2}$, $\sin \frac{\pi}{3} = \frac{\sqrt{3}}{2}$

wieć: a) $\sin \frac{5\pi}{36}$, b) $\sin \frac{5\pi}{18}$, c) $\sin \frac{7\pi}{18}$.

E1.18. Wiedząc, że $\sin \frac{\pi}{8} = \frac{\sqrt{2}+\sqrt{2}}{2}$, $\sin \frac{\pi}{6} = \frac{\sqrt{3}}{2}$, $\sin \frac{\pi}{4} = \frac{\sqrt{2}}{2}$, $\sin \frac{\pi}{3} = \frac{1}{2}$

wieć: a) $\cos \frac{5\pi}{36}$, b) $\cos \frac{5\pi}{18}$, c) $\cos \frac{7}{18}$.

E1.19. Wiedząc, że $\operatorname{tg} \frac{\pi}{8} = \sqrt{2} - 1$, $\operatorname{tg} \frac{\pi}{6} = \frac{\sqrt{3}}{3}$, $\operatorname{tg} \frac{\pi}{4} = 1$, $\operatorname{tg} \frac{\pi}{3} = \sqrt{3}$

wieć: a) $\operatorname{tg} \frac{5\pi}{36}$, b) $\operatorname{tg} \frac{5\pi}{18}$, c) $\operatorname{tg} \frac{7}{18}$.

E1.20. Wiedząc, że $\operatorname{ctg} \frac{\pi}{8} = \sqrt{2} + 1$, $\operatorname{ctg} \frac{\pi}{6} = \sqrt{3}$, $\operatorname{ctg} \frac{\pi}{4} = 1$, $\operatorname{ctg} \frac{\pi}{3} = \frac{\sqrt{3}}{3}$

wieć: a) $\operatorname{ctg} \frac{5\pi}{36}$, b) $\operatorname{ctg} \frac{5\pi}{18}$, c) $\operatorname{ctg} \frac{7}{18}$.

5.1.21. Wiedząc, że $\arctg(2-\sqrt{3})=\frac{\pi}{12}$, $\arctg(\sqrt{2}-1)=\frac{\pi}{8}$, $\arctg\left(\frac{\sqrt{3}}{3}\right)=\frac{\pi}{6}$

obliczyć: a) $\arctg\left(\frac{1}{4}\right)$, b) $\arctg\left(\frac{1}{2}\right)$.

5.1.22. Wiedząc, że $\arcsin\left(\frac{1}{2}\right)=\frac{\pi}{6}$, $\arcsin\left(\frac{\sqrt{2}}{2}\right)=\frac{\pi}{4}$, $\arcsin\left(\frac{\sqrt{3}}{2}\right)=\frac{\pi}{3}$

obliczyć: a) $\arcsin\left(\frac{2}{3}\right)$, b) $\arcsin\left(\frac{3}{4}\right)$.

5.1.23. Wiedząc, że $\arctg(\sqrt{3}-2)=-\frac{\pi}{12}$, $\arctg 0=0$, $\arctg(2-\sqrt{3})=\frac{\pi}{12}$, $\arctg(\sqrt{2}-1)=\frac{\pi}{8}$

obliczyć: a) $\arctg\left(\frac{1}{5}\right)$, b) $\arctg\left(\frac{1}{3}\right)$, c) $\arctg\left(\frac{1}{2}\right)$.

5.1.24. Wiedząc, że $\arcsin\left(-\frac{1}{2}\right)=-\frac{\pi}{6}$, $\arcsin 0=0$, $\arcsin\left(\frac{1}{2}\right)=\frac{\pi}{6}$, $\arcsin\left(\frac{\sqrt{2}}{2}\right)=\frac{\pi}{4}$

obliczyć: a) $\arcsin\left(\frac{1}{4}\right)$, b) $\arcsin\left(\frac{2}{3}\right)$, c) $\arcsin\left(\frac{3}{4}\right)$.

Zad. 5.8. Dana jest funkcja $y=f(x)$ określona i ciągła w przedziale $[a ; b]$. Należy:

- 1) funkcję $y=f(x)$ stablicować w przedziale $[a ; b]$ z krokiem $h_1 = \frac{b-a}{10}$,
- 2) skonstruować wielomian interpolacyjny Stirlinga $S_{10}(x)$ dla utworzonej funkcji tabelarycznej,
- 3) stablicować funkcje $y=f(x)$, $y=S_{10}(x)$, $y=f(x) - S_{10}(x)$ w przedziale $[a ; b]$ z krokiem $h_2 = \frac{b-a}{40}$,
- 4) wykonać – w tym samym układzie współrzędnych – wykresy funkcji $y=f(x)$, $y=S_{10}(x)$.
 - a) $f(x)=|x|$ na $[-1; 1]$,
 - b) $f(x)=1-|x|$ na $[-1; 1]$,
 - c) $f(x)=e^{-x^2}$ na $[-1; 1]$,
 - d) $f(x)=x+1$ na $[-1; 1]$,
 - e) $f(x)=\cos x$ na $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$,

1) $f(x) = \sin x$ na $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$,

2) $f(x) = \operatorname{tg} x$ na $\left[-\frac{\pi}{4}; \frac{\pi}{4}\right]$,

3) $f(x) = \operatorname{ctg} x$ na $\left[\frac{\pi}{4}; \frac{3\pi}{4}\right]$,

4) $f(x) = \arcsin x$ na $[-1; 1]$,

5) $f(x) = \arctg x$ na $[-1; 1]$,

6) $f(x) = 2^x$ na $[-1; 1]$,

7) $f(x) = 5^x$ na $[-1; 1]$,

8) $f(x) = \log x$ na $\left[\frac{1}{2}; 4\right]$,

9) $f(x) = \log_2 x$ na $\left[\frac{1}{2}; 4\right]$,

10) $f(x) = \sqrt{x}$ na $[0; 4]$,

11) $f(x) = \sqrt[6]{x}$ na $[0; 4]$,

12) $f(x) = shx$ na $[-1; 1]$,

13) $f(x) = chx$ na $[-1; 1]$,

14) $f(x) = thx$ na $[-1; 1]$,

15) $f(x) = cthx$ na $\left[\frac{1}{2}; 2\right]$.

Zad. 5.9. Dana jest funkcja $y = f(x)$ określona i ciągła w przedziale $[a; b]$. Należy:

■ funkcję $y = f(x)$ stabilicować w przedziale $[a; b]$ z krokiem $h_1 = \frac{b-a}{10}$,

■ skonstruować wielomian interpolacyjny Stirlinga $S_{10}(x)$ dla utworzonej funkcji tabelarycznej,

■ stabilicować funkcje $y = f(x)$, $y = S_{10}(x)$, $y = f(x) - S_{10}(x)$ w przedziale $[a; b]$ z krokiem $h_2 = \frac{b-a}{40}$,

■ wykonać – w tym samym układzie współrzędnych wykresy funkcji $y = f(x)$, $y = S_{10}(x)$.

- a) $f(x) = \frac{1}{x}$ na $\left[\frac{1}{2}; 2\right]$,
- b) $f(x) = \frac{1}{x^4}$ na $\left[\frac{1}{2}; 2\right]$,
- c) $f(x) = \sqrt{1-x^2}$ na $[-1; 1]$,
- d) $f(x) = \sqrt{1-\frac{x^2}{4}}$ na $[-2; 2]$,
- e) $f(x) = 3\sqrt{1-x^2}$ na $[-1; 1]$,
- f) $f(x) = \frac{(x-1)^3}{(x+1)^2}$ na $\left[\frac{1}{2}; 4\right]$,
- g) $f(x) = \sqrt{\frac{8-x^3}{3x}}$ na $\left[\frac{1}{2}; 2\right]$,
- h) $f(x) = \sin \frac{1}{x}$ na $\left[\frac{1}{\pi}; \frac{2}{\pi}\right]$,
- i) $f(x) = \sin^3 x + \cos^3 x$ na $\left[\frac{\pi}{4}; \frac{5\pi}{4}\right]$,
- j) $f(x) = \sin^3 x + \cos^3 x$ na $\left[\frac{3\pi}{4}; \frac{7\pi}{4}\right]$,
- k) $f(x) = (x+2)^2(x-1)^3$ na $[-2; 1]$,
- l) $f(x) = \frac{(x-1)^3}{(x+1)^2}$ na $[0; 4]$.

Aproksymacja

6.1. Wprowadzenie

Aproksymacja oznacza przybliżanie. Zadaniem aproksymacji jest znalezienie prostej funkcji $y = F(x)$ określonej klasy, przybliżającej zadaną funkcję $y = f(x)$. Przyczyną formułowania takiego zadania może być fakt, że funkcja $y = f(x)$ jest określona niewygodnym w ~~analitycznym~~ wzorem analitycznym. Inną przyczyną może być znajomość tylko skońzonego ~~analitycznego~~ wartości tej funkcji, np. wartości odczytanych w trakcie pomiaru. Funkcję $y = f(x)$ ~~analityczną~~ nazywali *funkcją aproksymowaną*, a funkcję $y = F(x)$ – *funkcją aproksymującą* lub ~~analityczną~~ *przybliżeniem funkcji* $y = f(x)$. Funkcji $y = F(x)$ poszukujemy zawsze z określonej rozmaitości funkcji, np. spośród wielomianów. Przybliżenie jednej funkcji przez inną powoduje pojawienie się błędów, zwanych *błędami aproksymacji (przybliżenia)*. W zależności od sposobu mierzenia błędu aproksymacji rozróżniamy dwa jej zasadnicze rodzaje: *aproksymację jednostajną i aproksymację średniokwadratową*.

W *aproksymacji jednostajnej* zakładamy, że funkcja aproksymowana $y = f(x)$ oraz funkcja aproksymująca $y = F(x)$ są określone i ciągłe na przedziale $[a ; b]$. Błąd aproksymacji jest mierzony za pomocą *normy Czebyszewa*

$$\|f - F\| = \sup_{a \leq x \leq b} \|f(x) - F(x)\|. \quad (6.1)$$

W *aproksymacji średniokwadratowej* wyróżniamy dwa przypadki. Jeśli funkcja aproksymowana $y = f(x)$ jest określona i ciągła na przedziale $[a ; b]$, to mamy do czynienia z *aproksymacją średniokwadratową ciągłą*. W tym przypadku błąd aproksymacji określony jest wzorem

$$\|f - F\| = \sqrt{\int_a^b w(x)[f(x) - F(x)]^2 dx}. \quad (6.2)$$

Jeśli funkcja aproksymowana $y = f(x)$ jest *funkcją dyskretną*, tzn. jej znane wartości ~~można~~ przedstawić za pomocą tabeli.

Tabela 6.1. Znane wartości funkcji dyskretnej

x_i	x_0	x_1	...	x_n
$y_i = f(x_i)$	y_0	y_1	...	y_n

to mamy przypadek *aproksymacji średniokwadratowej dyskretnej*, dla której błąd aproksymacji zdefiniowany jest wzorem

$$\| f - F \| = \sum_{i=0}^n w(x_i) [f(x_i) - F(x_i)]^2. \quad (6.3)$$

We wzorach (6.2) i (6.3) $y = w(x)$ jest funkcją rzeczywistą, o wartościach nieujemnych i nosi nazwę *funkcji wagowej*.

Najczęściej poszukujemy funkcji aproksymującej w postaci *wielomianu uogólnionego*

$$F(x) = a_0 \varphi_0(x) + a_1 \varphi_1(x) + \dots + a_m \varphi_m(x) \quad (6.4)$$

w którym funkcje

$$\varphi_0(x), \varphi_1(x), \dots, \varphi_m(x)$$

są wybranymi a priori funkcjami bazowymi $m+1$ wymiarowej przestrzeni liniowej. W tym przypadku zadanie aproksymacji sprowadza się do określenia współczynników a_0, a_1, \dots, a_m . Najczęściej stosowanymi bazami są:

— *baza jednomianów*: $1, x, x^2, \dots, x^m$,

— *baza trygonometryczna*: $1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos mx, \sin mx$.

Wynika to z następujących właściwości wielomianów oraz wielomianów trygonometrycznych.

Twierdzenie 6.1 (Weierstrass).

Dla każdej funkcji $y = f(x)$ określonej i ciągłej na domkniętym i ograniczonym odcinku $[a ; b]$ istnieje taki wielomian $w_n(x)$, który przybliża jednostajnie funkcję $y = f(x)$ na odcinku $[a ; b]$.

Twierdzenie 6.2 (Weierstrass).

Dla każdej funkcji $y = f(x)$ określonej i ciągłej na R oraz okresowej o okresie 2π istnieje taki wielomian trygonometryczny

$$S_m(x) = \frac{a_0}{2} + \sum_{k=1}^m (a_k \cos kx + b_k \sin kx),$$

który przybliża jednostajnie funkcję $y = f(x)$.

6.2. Aproksymacja średniokwadratowa dyskretna

Zakładamy, że znane wartości funkcji aproksymowanej $y = f(x)$ zostały zestawione w tabeli 6.1. Funkcję $y = f(x)$ będziemy aproksymowali wielomianem uogólnionym (6.4). Błąd aproksymacji jest obliczany według wzoru (6.3). Zadaniem aproksymacji średniokwadratowej dyskretnej jest wyznaczenie takich współczynników a_0, a_1, \dots, a_m wielomianu (6.4), przy których błąd aproksymacji (6.3) jest najmniejszy. Powyższe zagadnienie rozwiązujemy metodą najmniejszych kwadratów. W tym celu odległość $\|f - F\|$ rozpatrujemy jako funkcję $m+1$ zmiennych niezależnych a_0, a_1, \dots, a_m

$$D_m(a_0, a_1, \dots, a_m) = \|f - F\| = \sum_{i=1}^n w(x_i) \left[y_i - \sum_{k=0}^m a_k \varphi_k(x_i) \right]^2.$$

Wystając z warunku koniecznego istnienia minimum funkcji wielu zmiennych mamy

$$\frac{\partial \mathcal{D}_m(a_0, a_1, \dots, a_m)}{\partial a_j} = \sum_{i=1}^n 2w(x_i) \left[y_i - \sum_{k=0}^m a_k \varphi_k(x_i) \right] \varphi_j(x_i) = 0 \quad \text{dla } j = 0, 1, \dots, m$$

po uporządkowaniu składników względem a_k

$$\sum_{k=0}^m \left(\sum_{i=0}^n w(x_i) \varphi_k(x_i) \varphi_j(x_i) \right) a_k = \sum_{i=0}^n w(x_i) \varphi_j(x_i) y_i \quad \text{dla } j = 0, 1, \dots, m.$$

Macząc przez

$$(\varphi_j, \varphi_k) = \sum_{i=0}^n w(x_i) \varphi_j(x_i) \varphi_k(x_i) \quad \text{dla } j, k = 0, 1, \dots, m,$$

$$(\varphi_j, f) = \sum_{i=0}^n w(x_i) \varphi_j(x_i) f(x_i) \quad \text{dla } j = 0, 1, \dots, m,$$

wiadzamy zagadnienie znalezienia optymalnych współczynników a_0, a_1, \dots, a_m do rozwiązania następującego układu równań liniowych

$$\begin{aligned} (\varphi_0, \varphi_0) a_0 + (\varphi_0, \varphi_1) a_1 + \dots + (\varphi_0, \varphi_m) a_m &= (\varphi_0, f), \\ (\varphi_1, \varphi_0) a_0 + (\varphi_1, \varphi_1) a_1 + \dots + (\varphi_1, \varphi_m) a_m &= (\varphi_1, f), \\ \dots, \\ (\varphi_m, \varphi_0) a_0 + (\varphi_m, \varphi_1) a_1 + \dots + (\varphi_m, \varphi_m) a_m &= (\varphi_m, f), \end{aligned} \tag{6.5}$$

z tego nazwę *układu normalnego*.

Aproksymacja średniokwadratowa dyskretna za pomocą wielomianów

Jeżeli w charakterze funkcji bazowych przyjmiemy ciąg jednomianów

$$1, x, x^2, \dots, x^m,$$

funkcja aproksymująca przybierze postać wielomianu

$$F_m(x) = a_0 + a_1 x + \dots + a_m x_m. \tag{6.6}$$

W normalny przybierze postać

$$\begin{aligned} s_0 a_0 + s_1 a_1 + \dots + s_m a_m &= t_0, \\ s_1 a_0 + s_2 a_1 + \dots + s_{m+1} a_m &= t_1, \\ \dots, \\ s_m a_0 + s_{m+1} a_1 + \dots + s_{2m} a_m &= t_m, \end{aligned} \tag{6.7}$$

czytym

$$s_k = \sum_{i=0}^n x_i^k \quad \text{dla } k = 0, 1, \dots,$$

$$t_k = \sum_{i=0}^n y_i x_i^k \quad \text{dla } k = 0, 1, \dots.$$

Można wykazać, że jeżeli argumenty x_0, x_1, \dots, x_n są różne i $m \leq n$, to wyznacznik układu (6.7) jest różny od zera, a więc układ ten ma jednoznaczne rozwiązanie. Rozwiązuje powyższy układ równań liniowych, wyznaczamy współczynniki a_0, a_1, \dots, a_m .

Przykład 6.1. Przebieg niektórych zjawisk dobrze opisuje funkcja

$$f(x) = e^{-(\alpha x^2 + \beta \cdot x + \gamma)}, \quad (6.8)$$

w której parametry α, β, γ są liczbami rzeczywistymi. W tabeli 6.2 zostały odnotowane wyniki doświadczenia tego rodzaju zjawiska.

Tabela 6.2. Wartości funkcji (6.8)

x_i	1.00	1.25	1.50	1.75	2.00	2.25	2.50
$f(x_i)$	0.160	0.990	3.095	4.485	3.075	1.010	0.145

Wyznaczyć wartości parametrów α, β, γ dla przeprowadzonego doświadczenia.

Rozwiązanie. Logarytmując równość (6.8), otrzymujemy

$$y = g(x) = \ln(f(x)) = -\alpha \cdot x^2 - \beta \cdot x - \gamma. \quad (6.9)$$

Ostatnia równość sugeruje, że problem sprowadza do wyznaczenia współczynników wielomianu drugiego stopnia

$$F(x) = a_2 x^2 + a_1 x + a_0 = -\alpha \cdot x^2 - \beta \cdot x - \gamma,$$

najlepiej przybliżającego – w sensie najmniejszych kwadratów – funkcję z tabeli 6.3.

Tabela 6.3. Wartości funkcji (6.9)

x_i	1.00	1.25	1.50	1.75	2.00	2.25	2.50
$g(x_i) = \ln(f(x_i))$	-1.833	-0.010	1.130	1.501	1.123	0.010	-1.931

Współczynniki a_0, a_1, a_2 wyznaczymy, rozwiązujeając układ równań liniowych

$$\begin{aligned} s_0 a_0 + s_1 a_1 + s_2 a_2 &= t_0, \\ s_1 a_0 + s_2 a_1 + s_3 a_2 &= t_1, \\ s_2 a_0 + s_3 a_1 + s_4 a_2 &= t_2. \end{aligned} \quad (6.10)$$

W tabeli 6.4 zestawiono obliczenia prowadzące do wyznaczenia współczynników s_i, t_i .

Tabela 6.4. Wyznaczanie współczynników s_i, t_i

x_i	$f(x_i)$	$g(x_i)$	x_i^0	x_i^1	x_i^2	x_i^3	x_i^4	$y_i x_i^0$	$y_i x_i^1$	$y_i x_i^2$
1,00	0,160	-1,833	1,000	1,000	1,000	1,000	1,000	-1,833	-1,833	-1,833
1,25	0,990	-0,010	1,000	1,250	1,563	1,953	2,441	-0,010	-0,013	-0,016
1,50	3,095	1,130	1,000	1,500	2,250	3,375	5,063	1,130	1,695	2,542
1,75	4,485	1,501	1,000	1,750	3,063	5,359	9,379	1,501	2,626	4,596
2,00	3,075	1,123	1,000	2,000	4,000	8,000	16,000	1,123	2,247	4,493
2,25	1,010	0,010	1,000	2,250	5,063	11,391	25,629	0,010	0,022	0,050
2,50	0,145	-1,931	1,000	2,500	6,250	15,625	39,063	-1,931	-4,828	-12,069
			s_0	s_1	s_2	s_3	s_4	t_0	t_1	t_2
			7,000	12,250	23,188	46,703	98,574	-0,010	-0,083	-2,236

W konsekwencji układ (6.10) przybiera postać

$$\begin{aligned} 7.000a_0 + 12.250a_1 + 23.188a_2 &= -0.010, \\ 12.250a_0 + 23.188a_1 + 46.703a_2 &= -0.083, \\ 23.188a_0 + 46.703a_1 + 98.574a_2 &= -2.236. \end{aligned}$$

Rozwiązaniem tego układu jest

$$a_0 = -16.854, a_1 = 21.016, a_2 = -6.015.$$

Ostatecznie poszukiwanymi wartościami parametrów są

$$\alpha = 6.015, \beta = -21.016, \gamma = 16.854.$$

Różnice między funkcją aproksymowaną i aproksymującą można odczytać z tabeli 6.5.

Tablica 6.5. Różnice między funkcjami aproksymowaną i aproksymującą

x	$f(x)$	$F(x)$	$f(x)-F(x)$
1,00	0,160	0,157	0,003
1,25	0,990	1,017	-0,027
1,50	3,095	3,114	-0,019
1,75	4,485	4,493	-0,008
2,00	3,075	3,056	0,019
2,25	1,010	0,980	0,030
2,50	0,145	0,148	-0,003

Przykład 6.2. Tor dysku, kuli czy oszczepu rzuconego przez sportowca jest zbliżony do kuli. Opierając się na danych zamieszczonych w tabeli 6.6, wyznaczyć współczynniki kuli najlepiej, w sensie najmniejszych kwadratów, przybliżającej tor kuli.

Tabela 6.6. Współrzędne toru kuli pchniętej przez lekkoatletę

x_i [m]	0.00	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00	4.50	5.00
y_i [m]	1.60	2.01	2.40	2.76	3.11	3.43	3.73	4.01	4.27	4.51	4.73
x_i [m]	5.50	6.00	6.50	7.00	7.50	8.00	8.50	9.00	9.50	10.00	10.50
y_i [m]	4.93	5.10	5.25	5.38	5.49	5.58	5.65	5.70	5.72	5.73	5.71
x_i [m]	11.00	11.50	12.00	12.50	13.00	13.50	14.00	14.50	15.00	15.50	16.00
y_i [m]	5.67	5.61	5.53	5.43	5.27	5.08	4.88	4.68	4.42	4.17	3.88
x_i [m]	16.50	17.00	17.50	18.00	18.50	19.00	19.50	20.00			
y_i [m]	3.55	3.18	2.77	2.34	1.78	1.28	0.73	0.18			

Rozwiązanie. Współczynniki wielomianu aproksymującego

$$F_2(x) = a_0 + a_1x + a_2x^2$$

obliczymy, rozwiązując układ równań (6.10). Proces obliczania współczynników s_i, t_i został przedstawiony w tabeli 6.7. Układ (6.10) przybiera w tym przypadku postać

$$\begin{bmatrix} 41.00 & 410.00 & 5535.00 \\ 410.00 & 5535.00 & 84050.00 \\ 5535.00 & 84050.00 & 1361333.25 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 167.24 \\ 1600.55 \\ 19207.89 \end{bmatrix}.$$

Czytelnik zechce zwrócić uwagę na specyficzną strukturę macierzy głównej układu. Czym ona jest spowodowana? Rozwiązaniem tego układu jest

$$a_0 = 1.45, a_1 = 0.91, a_2 = -0.05.$$

Układ równań liniowych (6.7) jest dla $m \geq 6$ na ogół układem źle uwarunkowanym. Oznacza to, że współczynniki a_0, a_1, \dots, a_m otrzymane przez rozwiązanie tego układu równań mogą być tak bardzo zaburzone, że nie nadają się do praktycznego wykorzystania. W przypadku gdy $m \geq 6$ w charakterze funkcji bazowych, przyjmujemy układ wielomianów ortogonalnych.

Tabela 6.7. Proces obliczania współczynników s_i i t_i

x_i	$f(x_i)$	x_i^0	x_i^1	x_i^2	x_i^3	x_i^4	$y_i x_i^0$	$y_i x_i^1$	$y_i x_i^2$
1	2	3	4	5	6	7	8	9	10
0.00	1.60	1.00	0.00	0.00	0.00	0.00	1.60	0.00	0.00
0.50	2.01	1.00	0.50	0.25	0.13	0.06	2.01	1.00	0.50
1.00	2.40	1.00	1.00	1.00	1.00	1.00	2.40	2.40	2.40
1.50	2.76	1.00	1.50	2.25	3.38	5.06	2.76	4.14	6.22
2.00	3.11	1.00	2.00	4.00	8.00	16.00	3.11	6.22	12.43
2.50	3.43	1.00	2.50	6.25	15.63	39.06	3.43	8.58	21.45
3.00	3.73	1.00	3.00	9.00	27.00	81.00	3.73	11.20	33.60
3.50	4.01	1.00	3.50	12.25	42.88	150.06	4.01	14.05	49.18

cd. tab. 6.7

1	2	3	4	5	6	7	8	9	10	
4.00	4.27	1.00	4.00	16.00	64.00	256.00	4.27	17.10	68.39	
4.50	4.51	1.00	4.50	20.25	91.13	410.06	4.51	20.31	91.38	
5.00	4.73	1.00	5.00	25.00	125.00	625.00	4.73	23.65	118.23	
5.50	4.93	1.00	5.50	30.25	166.38	915.06	4.93	27.09	148.98	
6.00	5.10	1.00	6.00	36.00	216.00	1296.00	5.10	30.60	183.58	
6.50	5.25	1.00	6.50	42.25	274.63	1785.06	5.25	34.14	221.91	
7.00	5.38	1.00	7.00	49.00	343.00	2401.00	5.38	37.69	263.82	
7.50	5.49	1.00	7.50	56.25	421.88	3164.06	5.49	41.21	309.06	
8.00	5.58	1.00	8.00	64.00	512.00	4096.00	5.58	44.67	357.34	
8.50	5.65	1.00	8.50	72.25	614.13	5220.06	5.65	48.04	408.30	
9.00	5.70	1.00	9.00	81.00	729.00	6561.00	5.70	51.28	461.51	
9.50	5.72	1.00	9.50	90.25	857.38	8145.06	5.72	54.37	516.47	
10.00	5.73	1.00	10.00	100.00	1000.00	10000.00	5.73	57.26	572.64	
10.50	5.71	1.00	10.50	110.25	1157.63	12155.06	5.71	59.94	629.40	
11.00	5.67	1.00	11.00	121.00	1331.00	14641.00	5.67	62.37	686.07	
11.50	5.61	1.00	11.50	132.25	1520.88	17490.06	5.61	64.51	741.89	
12.00	5.53	1.00	12.00	144.00	1728.00	20736.00	5.53	66.34	796.06	
12.50	5.43	1.00	12.50	156.25	1953.13	24414.06	5.43	67.82	847.71	
13.00	5.27	1.00	13.00	169.00	2197.00	28561.00	5.27	68.51	890.63	
13.50	5.08	1.00	13.50	182.25	2460.38	33215.06	5.08	68.58	925.83	
14.00	4.88	1.00	14.00	196.00	2744.00	38416.00	4.88	68.32	956.48	
14.50	4.68	1.00	14.50	210.25	3048.63	44205.06	4.68	67.86	983.97	
15.00	4.42	1.00	15.00	225.00	3375.00	50625.00	4.42	66.30	994.50	
15.50	4.17	1.00	15.50	240.25	3723.88	57720.06	4.17	64.64	1001.84	
16.00	3.88	1.00	16.00	256.00	4096.00	65536.00	3.88	62.08	993.28	
16.50	3.55	1.00	16.50	272.25	4492.13	74120.06	3.55	58.58	966.49	
17.00	3.18	1.00	17.00	289.00	4913.00	83521.00	3.18	54.06	919.02	
17.50	2.77	1.00	17.50	306.25	5359.38	93789.06	2.77	48.48	848.31	
18.00	2.34	1.00	18.00	324.00	5832.00	104976.00	2.34	42.12	758.16	
18.50	1.78	1.00	18.50	342.25	6331.63	117135.06	1.78	32.93	609.21	
19.00	1.28	1.00	19.00	361.00	6859.00	130321.00	1.28	24.32	462.08	
19.50	0.73	1.00	19.50	380.25	7414.88	144590.06	0.73	14.24	277.58	
20.00	0.18	1.00	20.00	400.00	8000.00	160000.00	0.18	3.60	72.00	
			s_0	s_1	s_2	s_3	s_4	t_0	t_1	t_2
			41.00	410.00	5535.00	84050.00	1361333.25	167.24	1600.55	19207.89

Przykład 6.3. W poniższej tabeli zostały odnotowane wyniki przeprowadzonego doświadczenia

x_i	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5
$y_i = f(x_i)$	1.02	0.62	0.50	0.60	0.98	1.55	3.12	5.08

Przeprowadzający doświadczenie stwierdził, że badana funkcja jest zbliżona do funkcji kwadratowej, a wartość $f(2.5)$ jest obarczona błędem grubym. Wyznacz wartość funkcji $y = f(x)$ dla $x = 2.5$.

Rozwiązanie. Ponieważ wartość $f(2.5) = 1.55$ jest obarczona błędem grubym, więc nie możemy uwzględniać jej w obliczeniach. Obliczenia oprzemy na funkcji tabelarycznej

x_i	0.0	0.5	1.0	1.5	2.0	3.0	3.5
$y_i = g(x_i)$	1.02	0.62	0.50	0.60	0.98	3.12	5.08

która – zgodnie z uwagą uczynioną przez przeprowadzającego eksperyment – powinniśmy przybliżać parabolą

$$F_2(x) = a_0 + a_1x + a_2x^2.$$

Współczynniki a_0, a_1, a_2 wyznaczmy, rozwiązując układ równań liniowych (6.10). W tabeli 6.8 zestawiono obliczenia prowadzące do wyznaczenia współczynników s_i, t_i .

Rozwiązujeając układ równań

$$\begin{bmatrix} 7.000 & 11.500 & 28.750 \\ 11.500 & 28.750 & 82.375 \\ 28.750 & 82.375 & 253.188 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1.124 \\ -1.495 \\ 0.739 \end{bmatrix}.$$

Otrzymujemy: $a_0 = 1.124, a_1 = -1.495, a_2 = 0.739$.

Tabela 6.8. Proces obliczania współczynników s_i i t_i

x	$g(x)$	x_i^0	x_i^1	x_i^2	x_i^3	x_i^4	$y_i x_i^0$	$y_i x_i^1$	$y_i x_i^2$
0.0	1.02	1.000	0.000	0.000	0.000	0.000	1.020	0.000	0.000
0.5	0.62	1.000	0.500	0.250	0.125	0.063	0.620	0.310	0.155
1.0	0.50	1.000	1.000	1.000	1.000	1.000	0.500	0.500	0.500
1.5	0.60	1.000	1.500	2.250	3.375	5.063	0.600	0.900	1.350
2.0	0.98	1.000	2.000	4.000	8.000	16.000	0.980	1.960	3.920
3.0	3.12	1.000	3.000	9.000	27.000	81.000	3.120	9.360	28.080
3.5	5.08	1.000	3.500	12.250	42.875	150.063	5.080	17.780	62.230
		s_0	s_1	s_2	s_3	s_4	t_0	t_1	t_2
		7.000	11.500	28.750	82.375	253.188	11.920	30.810	96.235

Za rozwiązanie zadania przyjmujemy

$$F_2(2.5) = 1.124 - 1.495 \cdot 2.5 + 0.739 \cdot 2.5^2 = 2.00.$$

■ Aproksymacja średniokwadratowa dyskretna wielomianami ortogonalnymi

Układ wielomianów $\varphi_0(x), \varphi_1(x), \dots, \varphi_k(x), \dots$ nazywamy *układem ortogonalnym* funkcji wagowej $y = w(x)$ (z wagą $w(x)$) na zbiorze $X = \{x_0, x_1, \dots, x_n\}$ wtedy i tylko wtedy gdy

$$(\varphi_j, \varphi_k) = \sum_{i=0}^n w(x_i) \varphi_j(x_i) \varphi_k(x_i) \begin{cases} = 0 & \text{dla } j \neq k, \\ > 0 & \text{dla } j = k. \end{cases}$$

Układ ortogonalny wielomianów jest układem funkcji liniowo niezależnych. Można ten układ potraktować jako bazę pewnej podprzestrzeni funkcyjnej i aproksymować funkcję $y = f(x)$ za pomocą liniowej kombinacji wielomianów tej bazy.

W przypadku gdy funkcje bazowe tworzą układ ortogonalny, układ normalny (6.5) ma postać

$$\begin{aligned} (\varphi_0, \varphi_0) a_0 &= (\varphi_0, f), \\ (\varphi_1, \varphi_1) a_1 &= (\varphi_1, f), \\ \dots &\dots \\ (\varphi_m, \varphi_m) a_m &= (\varphi_m, f), \end{aligned} \tag{6.11}$$

zauważ, że $(\varphi_j, \varphi_k) = 0$ dla $j, k = 0, 1, \dots, m; j \neq k$ na mocy definicji ortogonalności. Rozwiążemy układ (6.11) określone jest wzorami

$$a_j = \frac{(\varphi_j, f)}{(\varphi_j, \varphi_j)} = \frac{\sum_{i=0}^n w(x_i) \varphi_j(x_i) f(x_i)}{\sum_{i=0}^n w(x_i) \varphi_j^2(x_i)} \quad \text{dla } j = 0, 1, \dots, m. \tag{6.12}$$

Wzory te określają sposób obliczania współczynników wielomianu aproksymującego.

Błąd aproksymacji można obliczyć ze wzoru

$$D_m^{\min} = \|f - F\| = \sqrt{\sum_{i=0}^n w(x_i) f^2(x_i) - \sum_{k=0}^m \frac{\left(\sum_{i=0}^n w(x_i) \varphi_k(x_i) f(x_i) \right)^2}{\sum_{i=0}^n w(x_i) \varphi_k^2(x_i)}}. \tag{6.13}$$

Z powyższego wzoru wynika, że

$$D_0^{\min} \geq D_1^{\min} \geq D_2^{\min} \geq \dots$$

Co znaczy to, że odchylenie średniokwadratowe monotonicznie maleje wraz ze wzrostem stopnia wielomianu aproksymującego. Dalej podajemy przykłady układów wielomianów ortogonalnych.

6.4.1. Aproksymacja za pomocą wielomianów p_k

Zakładamy, że znane wartości funkcji aproksymowanej $y = f(x)$ zostały zestawione w tabeli 6.1, w której węzły x_0, x_1, \dots, x_n nie muszą być równoodległe.

Wielomiany p_k określone są następującymi wzorami rekurencyjnymi

$$\begin{aligned} p_{-1}(x) &= 0, \\ p_0(x) &= 1, \\ p_{k+1}(x) &= (x - \alpha_{k+1})p_k(x) - \beta_k p_{k-1}(x) \text{ dla } k = 0, 1, \dots, \end{aligned} \quad (6.14)$$

gdzie β_0 jest dowolne,

$$\beta_k = \frac{\sum_{i=0}^n p_k^2(x_i)}{\sum_{i=0}^n p_{k-1}^2(x_i)} \quad \text{dla } k = 1, 2, \dots, \quad (6.15)$$

$$\alpha_{k+1} = \frac{\sum_{i=0}^n x_i p_k^2(x_i)}{\sum_{i=0}^n p_k^2(x_i)} \quad \text{dla } k = 0, 1, \dots. \quad (6.16)$$

Wielomiany $p_k(x)$ ($k = 0, 1, \dots$) tworzą na zbiorze $X = \{x_0, x_1, \dots, x_n\}$ układ ortogonalny z funkcją wagową $y = w(x) \equiv 1$. Zatem funkcję $y = f(x)$ możemy przybliżać za pomocą wielomianu

$$F_m(x) = a_0 p_0(x) + a_1 p_1(x) + \dots + a_m p_m(x). \quad (6.17)$$

Przykład 6.4. Zadanie sformułowane w przykładzie 6.1 rozwiązać z wykorzystaniem wielomianów p_k .

Rozwiązanie. Logarytmując równość (6.8) sprowadzamy problem do aproksymacji funkcji określonej w tabeli 6.9 wielomianem drugiego stopnia.

Tabela 6.9. Wartości funkcji $y = g(x)$

x_i	1.00	1.25	1.50	1.75	2.00	2.25	2.50
$g(x_i) = \ln(f(x_i))$	-1.833	-0.010	1.130	1.501	1.123	0.010	-1.931

W niniejszym przykładzie w charakterze funkcji aproksymującej przyjmujemy wielomian aproksymacyjny (6.17) przy $m = 2$:

$$\begin{aligned} F_2(x) &= a_0 p_0(x) + a_1 p_1(x) + a_2 p_2(x) = \\ &= a_0 \cdot 1 + a_1 \cdot (x - \alpha_1) + a_2 ((x - \alpha_2)(x - \alpha_1) - \beta_1). \end{aligned} \quad (6.18)$$

Obliczanie współczynników $a_0, a_1, a_2, \alpha_1, \alpha_2, \beta_1$ zostało przedstawione w tabeli 6.10.

Tabela 6.10. Proces aproksymacji funkcji $y = g(x)$ za pomocą wielomianów p_k

x_i	y_i	$p_{-1}(x_i)$	$F_{-1}(x_i)$	$p_0(x_i)$	$F_0(x_i)$	$p_1(x_i)$	$F_1(x_i)$	$p_2(x_i)$	$F_2(x_i)$
-1.00	-1.833	0.000	0.000	1.000	-0.001	-0.750	0.027	0.313	-1.853
-0.75	-0.010	0.000	0.000	1.000	-0.001	-0.500	0.017	0.000	0.017
-0.50	1.130	0.000	0.000	1.000	-0.001	-0.250	0.008	-0.188	1.136
-0.25	1.501	0.000	0.000	1.000	-0.001	0.000	-0.001	-0.250	1.503
0.00	1.123	0.000	0.000	1.000	-0.001	0.250	-0.011	-0.188	1.117
0.25	0.010	0.000	0.000	1.000	-0.001	0.500	-0.020	0.000	-0.020
0.50	-1.931	0.000	0.000	1.000	-0.001	0.750	-0.029	0.313	-1.909
		$\alpha_1 =$	1.750	$\alpha_2 =$	1.750				
		$\beta_0 =$	1.000	$\beta_1 =$	0.250				
		$\alpha_0 =$	-0.001	$\alpha_1 =$	-0.037	$\alpha_2 =$	-6.016		

Współczynniki α, β, γ wyznaczamy poprzez przerównanie wyrażenia (6.18) do prawej strony funkcji $\ln(f(x)) = -(\alpha + \beta x + \gamma)$:

$$F_2(x) = \alpha_0 \cdot 1 + \alpha_1 \cdot (x - \alpha_1) + \alpha_2 \cdot ((x - \alpha_2)(x - \alpha_1) - \beta_1) = -(\alpha x^2 + \beta x + \gamma).$$

Przez przerównanie współczynników przy jednakowych potęgach x otrzymujemy:

$$\alpha = -\alpha_2 = 6.016,$$

$$\beta = -(\alpha_1 - \alpha_2(\alpha_1 + \alpha_2)) = -21.018,$$

$$\gamma = -(\alpha_2(\alpha_1\alpha_2 - \beta_1) - \alpha_1\alpha_1) = 16.854.$$

6.2. Aproksymacja za pomocą wielomianów Grama

Zakładamy, że węzły funkcji aproksymowanej, oznaczone tu przez u_0, u_1, \dots, u_n , są punktami podziału odcinka $[-1; 1]$ na n równych części, tj.

$$u_i = \frac{2i}{n} - 1 \quad \text{dla } i = 0, 1, \dots, n.$$

Tabela 6.11. Znane wartości funkcji aproksymowanej

u_i	u_0	u_1	...	u_n
$y_i = f(u_i)$	y_0	y_1	...	y_n

Wielomiany Grama są określone następującymi związkami rekurencyjnymi

$$G_{-1}(u) = 0,$$

$$G_0(u) = \alpha_0 = \frac{1}{\sqrt{n+1}},$$

$$G_k(u) = \alpha_k \cdot u \cdot G_{k-1}(u) - \gamma_k G_{k-2}(u) \quad \text{dla } k = 1, 2, \dots, n-1, \quad (6.19)$$

gdzie

$$\alpha_k = \frac{n}{k} \sqrt{\frac{4k^2 - 1}{(n+1)^2 - k^2}} \quad \text{dla } k = 1, 2, \dots, n-1, \quad (6.20)$$

$$\gamma_1 \text{ jest dowolne, } \gamma_k = \frac{\alpha_k}{\alpha_{k-1}} \quad \text{dla } k = 2, 3, \dots, n-1. \quad (6.21)$$

Wielomiany Grama $G_k(u)$ ($k = 0, 1, \dots, n-1$) tworzą układ ortogonalny na zbiorze

$$X = \{u_0, u_1, \dots, u_n\} = \left\{-1, \frac{2}{n}-1, \frac{4}{n}-1, \dots, 1\right\}$$

z funkcją wagową $y = w(x) \equiv 1$. Funkcję $y = f(u)$ określoną na odcinku $[-1; 1]$ możemy przybliżać za pomocą wielomianu

$$Q_m(u) = a_0 G_0(u) + a_1 G_1(u) + \dots + a_m G_m(u) \quad (m \leq n-1). \quad (6.22)$$

Przykład 6.5. Zadanie sformułowane w przykładzie 6.1 rozwiązać z wykorzystaniem wielomianów Grama.

Rozwiązanie. Do rozwiązania powyższego zadania możemy użyć wielomianów Grama, ponieważ węzły są rozmiieszczone ze stałym krokiem $h = 0.25$. Ponieważ wielomiany Grama są ortogonalne na odcinku $[-1; 1]$, musimy znaleźć przekształcenie odcinka $[1.0; 2.5]$ na $[-1; 1]$. Dla dwóch odcinków skończonej długości istnieje przekształcenie liniowe przekształcające jeden odcinek w drugi. Łatwo stwierdzić, że w naszym przypadku takie przekształcenie ma postać

$$x = \frac{3}{4}u + \frac{7}{4}, \quad u = \frac{4}{3}x - \frac{7}{3}, \quad (6.23)$$

gdzie $-1 \leq u \leq 1$ oraz $1 \leq x \leq 2.5$. Zatem funkcję $g(x) = \ln(f(x))$ będziemy przybliżali wielomianem

$$Q_2\left(\frac{4}{3}x - \frac{7}{3}\right) = a_0 G_0\left(\frac{4}{3}x - \frac{7}{3}\right) + a_1 G_1\left(\frac{4}{3}x - \frac{7}{3}\right) + a_2 G_2\left(\frac{4}{3}x - \frac{7}{3}\right). \quad (6.24)$$

Obliczenia współczynników $a_0, a_1, a_2, \alpha_1, \gamma_1, \alpha_2, \gamma_2$ są przedstawione w tabeli 6.12.

Pozostają do wyznaczenia wartości parametrów α, β, γ . Odpowiednie wzory na obliczenie ich wartości otrzymamy, przerównując współczynniki przy jednakowych potęgach x w równości

$$\begin{aligned} Q_2\left(\frac{4}{3}x - \frac{7}{3}\right) &= a_0 \cdot \alpha_0 + a_1 \left(\alpha_0 \alpha_1 \left(\frac{4}{3}x - \frac{7}{3} \right) \right) + a_2 \left(\alpha_0 \alpha_1 \alpha_2 \left(\frac{4}{3}x - \frac{7}{3} \right)^2 - \alpha_0 \gamma_2 \right) = \\ &= -(\alpha \cdot x^2 + \beta x + \gamma). \end{aligned}$$

Zatem mamy

$$\gamma = -\left(a_0 \alpha_0 - \frac{7}{3} \cdot a_1 \alpha_0 \alpha_1 + \frac{49}{9} a_2 \alpha_0 \alpha_1 \alpha_2 - a_2 \alpha_0 \gamma_2\right) = 16.856,$$

$$\alpha = -\left(\frac{16}{9}a_2\alpha_0\alpha_1\alpha_2\right) = 6.015.$$

Tabela 6.12. Proces aproksymacji funkcji $y = g(x)$ za pomocą wielomianów Grama

x_i	u_i	$f(x_i)$	$G_{-1}(x_i)$	$G_0(x_i)$	$Q_0(x_i)$	$G_1(x_i)$	$Q_1(x_i)$	$G_2(x_i)$	$Q_2(x_i)$
1.00	-1.00	-1.833	0.000	0.378	-0.001	-0.567	0.027	0.546	-1.853
1.25	-0.67	-0.010	0.000	0.378	-0.001	-0.378	0.017	0.000	0.017
1.50	-0.33	1.130	0.000	0.378	-0.001	-0.189	0.008	-0.327	1.136
1.75	0.00	1.501	0.000	0.378	-0.001	0.000	-0.001	-0.436	1.503
2.00	0.33	1.123	0.000	0.378	-0.001	0.189	-0.011	-0.327	1.117
2.25	0.67	0.010	0.000	0.378	-0.001	0.378	-0.020	0.000	-0.020
2.50	1.00	-1.931	0.000	0.378	-0.001	0.567	-0.029	0.546	-1.909
						$k =$	1	$k =$	2
						$\alpha_1 =$	1.500	$\alpha_2 =$	1.732
						$\gamma_1 =$	1.000	$\gamma_2 =$	1.155
				$a_0 =$	-0.004	$a_1 =$	-0.049	$a_2 =$	-3.446

4.3. Aproxymacja za pomocą wielomianów trygonometrycznych

Zakładamy, że funkcja aproksymowana $y = f(v)$ jest funkcją ciągłą, okresową o okresie okresem głównym $[0 ; 2\pi]$, a znane są jej wartości zestawione w tabeli 6.13, w której v_0, v_1, \dots, v_n są punktami odcinka $[0 ; 2\pi]$ określonymi wzorami

$$v_i = \frac{2\pi}{n+1} i \quad (i = 0, 1, \dots, n).$$

Tabela 6.13. Znane wartości aproksymowanej funkcji okresowej

v_i	v_0	v_1	...	v_n
$y_i = f(v_i)$	y_0	y_1	...	y_n

One funkcji

$$1, \cos v, \sin v, \cos 2v, \sin 2v, \dots, \cos(mv), \sin(mv), \dots,$$

Ważnym układem trygonometrycznym, jest układem funkcji ortogonalnych na zbiorze

$$V = \{v_0, v_1, \dots, v_n\} = \left\{0, \frac{2\pi}{n+1}, \frac{4\pi}{n+1}, \dots, \frac{2n\pi}{n+1}\right\}. \quad (6.24)$$

funkcją wagową $y = w(v) \equiv 1$.

Funkcję $y = f(v)$ ciągłą i okresową o okresie 2π z okresem głównym $[0 ; 2\pi]$ można przybliżać wielomianami trygonometrycznymi rzędu m

$$S_m(v) = \frac{a_0}{2} + \sum_{k=1}^m (a_k \cos kv + b_k \sin kv). \quad (6.25)$$

Współczynniki optymalne (w sensie aproksymacji średniokwadratowej) wielomianu trygonometrycznego wyrażają się wzorami

$$a_k = \frac{2}{n+1} \sum_{i=0}^n f(v_i) \cos kv_i = \frac{2}{n+1} \sum_{i=0}^n f\left(\frac{2\pi}{n}i\right) \cos \frac{2k\pi}{n}i \quad \text{dla } k = 0, 1, \dots, \quad (6.26)$$

$$b_k = \frac{2}{n+1} \sum_{i=0}^n f(v_i) \sin kv_i = \frac{2}{n+1} \sum_{i=0}^n f\left(\frac{2\pi}{n}i\right) \sin \frac{2k\pi}{n}i \quad \text{dla } k = 1, 2, \dots. \quad (6.27)$$

i noszą nazwę *współczynników Fouriera*. Wielomian trygonometryczny ze współczynnikami Fouriera nosi nazwę *wielomianu Fouriera*.

Dla funkcji $y = f(v)$ ciągłej, parzystej i okresowej, o okresie 2π z okresem głównym $[0 ; 2\pi]$ współczynniki Fouriera $b_k = 0$, dla $k = 1, 2, \dots$. Zatem taką funkcję można przybliżać wielomianami postaci

$$S_m(v) = \frac{a_0}{2} + \sum_{k=1}^m a_k \cos kv. \quad (6.28)$$

Dla funkcji $y = f(v)$ ciągłej, nieparzystej i okresowej, o okresie 2π z okresem głównym $[0 ; 2\pi]$ współczynniki Fouriera $a_k = 0$, dla $k = 0, 1, \dots$. Zatem taką funkcję można przybliżać wielomianami postaci

$$S_m(v) = \sum_{k=1}^m b_k \sin kv. \quad (6.29)$$

Przykład 6.6. W tabeli 6.14 podano wybrane wartości funkcji $y = f(t)$ ciągłej dla $-\infty < t < \infty$, okresowej o okresie $[-1 ; 1]$.

Tablica 6.14. Wartości funkcji okresowej

t_i	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8
$y_i = f(t_i)$	0.0	0.4	0.8	1.2	1.6	2.0	1.6	1.2	0.8	0.4

Zbudować dla tej funkcji wielomiany Fouriera aż do stopnia 5.

Rozwiązanie. Funkcja $y = f(t)$ jest ciągła i okresowa, a węzły są rozmieszczone ze stałym krokiem $h = 0.2$ ($n = 9$). Ponieważ wielomiany trygonometryczne są ortogonalne na odcinku $[0 ; 2\pi]$, musimy znaleźć przekształcenie odcinka $[-1 ; 1]$ na $[0 ; 2\pi]$. Takie przekształcenie ma postać

$$v = \pi \cdot t + \pi, \quad t = \frac{v}{\pi} - 1,$$

gdzie $-1 \leq t \leq 1$ oraz $0 \leq v \leq 2\pi$. Zauważmy ponadto, że $y = f(t)$ jest funkcją parzystą, za-

możemy ją aproksymować za pomocą wielomianów (6.28). Obliczanie współczynników a_k zostało przedstawione w tabeli 6.15.

Tabela 6.15. Wyznaczanie optymalnych współczynników a_k

t_i	v_i	$f(v_i)$	Wartości $f(v_i) \cos(kv_i)$ dla					
			$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
-1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-0.8	0.6283	0.4000	0.4000	0.3236	0.1236	-0.1236	-0.3236	-0.4000
-0.6	1.2566	0.8000	0.8000	0.2472	-0.6472	-0.6472	0.2472	0.8000
-0.4	1.8850	1.2000	1.2000	-0.3708	-0.9708	0.9708	0.3708	-1.2000
-0.2	2.5133	1.6000	1.6000	-1.2944	0.4944	0.4944	-1.2944	1.6000
0.0	3.1416	2.0000	2.0000	-2.0000	2.0000	-2.0000	2.0000	-2.0000
0.2	3.7699	1.6000	1.6000	-1.2944	0.4944	0.4944	-1.2944	1.6000
0.4	4.3982	1.2000	1.2000	-0.3708	-0.9708	0.9708	0.3708	-1.2000
0.6	5.0265	0.8000	0.8000	0.2472	-0.6472	-0.6472	0.2472	0.8000
0.8	5.6549	0.4000	0.4000	0.3236	0.1236	-0.1236	-0.3236	-0.4000
			a_0	a_1	a_2	a_3	a_4	a_5
			2.0000	-0.8378	0.0000	-0.1222	0.0000	-0.0800

Wielomiany (6.28) stopni od 0 do 5 zostały stablicowane w tabeli 6.16.

Tabela 6.16. Tablica wartości wielomianów Fouriera

t_i	v_i	$f(v_i)$	$S_0(v_i)$	$S_1(v_i)$	$S_2(v_i)$	$S_3(v_i)$	$S_4(v_i)$	$S_5(v_i)$
-1.0	0.0000	0.0000	1.0000	0.1622	0.1622	0.0400	0.0400	-0.0400
-0.8	0.6283	0.4000	1.0000	0.3222	0.3222	0.3600	0.3600	0.4400
-0.6	1.2566	0.8000	1.0000	0.7411	0.7411	0.8400	0.8400	0.7600
-0.4	1.8850	1.2000	1.0000	1.2589	1.2589	1.1600	1.1600	1.2400
-0.2	2.5133	1.6000	1.0000	1.6778	1.6778	1.6400	1.6400	1.5600
0.0	3.1416	2.0000	1.0000	1.8378	1.8378	1.9600	1.9600	2.0400
0.2	3.7699	1.6000	1.0000	1.6778	1.6778	1.6400	1.6400	1.5600
0.4	4.3982	1.2000	1.0000	1.2589	1.2589	1.1600	1.1600	1.2400
0.6	5.0265	0.8000	1.0000	0.7411	0.7411	0.8400	0.8400	0.7600
0.8	5.6549	0.4000	1.0000	0.3222	0.3222	0.3600	0.3600	0.4400
1.0	6.2832	0.0000	1.0000	0.1622	0.1622	0.0400	0.0400	-0.0400
			D_0	D_1	D_2	D_3	D_4	D_5
			4.6000	0.1170	0.1170	0.0176	0.0176	0.0176

6.5. Aproksymacja średniokwadratowa całkowa

Zakładamy, że funkcja aproksymowana $y = f(x)$ jest określona i ciągła na przedziale $[a; b]$. Funkcję $y = f(x)$ będziemy aproksymowali za pomocą wielomianu uogólnionego

$$F(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \dots + a_m\varphi_m(x)$$

Zagadnienie aproksymacji średniokwadratowej całkowej polega na znalezieniu takich współczynników c_0, c_1, \dots, c_m , dla których odległość $\|f - \varphi\|$ określona wzorem (6.2) jest najmniejsza. Powyższe zagadnienie rozwiązuje się metodą najmniejszych kwadratów. W tym celu odległość $\|f - \varphi\|$ rozpatrujemy jako funkcję

$$I_m(c_0, c_1, \dots, c_m) = \|f - \varphi\|^2 = \int_a^b w(x) \left[f(x) - \sum_{k=0}^m c_k \varphi_k(x) \right]^2 dx$$

$m+1$ zmiennych niezależnych c_0, c_1, \dots, c_m . Korzystając z warunku koniecznego istnienia minimum funkcji wielu zmiennych, mamy

$$\frac{\partial I_m(c_0, c_1, \dots, c_m)}{\partial c_j} = 2 \int_a^b w(x) \left[f(x) - \sum_{k=0}^m c_k \varphi_k(x) \right] \varphi_j(x) dx = 0 \quad \text{dla } j = 0, 1, \dots, m$$

lub po uporządkowaniu względem c_k

$$\sum_{k=0}^m \left(\int_a^b w(x) \varphi_k(x) \varphi_j(x) dx \right) c_k = \int_a^b w(x) \varphi_j(x) f(x) dx \quad \text{dla } j = 0, 1, \dots, m.$$

Oznaczając przez

$$g_{jk} = \int_a^b w(x) \varphi_j(x) \varphi_k(x) dx \quad \text{dla } j, k = 0, 1, \dots, m,$$

$$h_j = \int_a^b w(x) \varphi_j(x) f(x) dx \quad \text{dla } j = 0, 1, \dots, m$$

sprowadzamy zagadnienie znalezienia optymalnych współczynników c_0, c_1, \dots, c_m do rozwiązania następującego układu normalnego równań liniowych

$$\begin{aligned} g_{00}c_0 + g_{01}c_1 + \dots + g_{0m}c_m &= h_0, \\ g_{10}c_0 + g_{11}c_1 + \dots + g_{1m}c_m &= h_1, \\ \dots & \\ g_{m0}c_0 + g_{m1}c_1 + \dots + g_{mm}c_m &= h_m. \end{aligned} \tag{6.26}$$

Jednak przy tak ogólnych założeniach odnośnie bazowych funkcji aproksymujących nie można nawet orzec, czy ten układ równań liniowych jest układem kramersowskim.

6. Aproksymacja średniokwadratowa całkowa za pomocą wielomianów

Jeżeli w charakterze funkcji bazowych przyjmiemy ciąg jednomianów

$$1, x, x^2, \dots, x^m,$$

to funkcja aproksymująca przybierze postać wielomianu

$$F_m(x) = a_0 + a_1x + \dots + a_mx_m. \quad (6.27)$$

Układ normalny przybierze postać

$$\begin{aligned} s_0a_0 + s_1a_1 + \dots + s_ma_m &= t_0, \\ s_1a_0 + s_2a_1 + \dots + s_{m+1}a_m &= t_1, \\ \dots &, \\ s_ma_0 + s_{m+1}a_1 + \dots + s_{2m}a_m &= t_m, \end{aligned} \quad (6.28)$$

czytajemy czym

$$s_k = \int_a^b x^k dx = \frac{b^{k+1} - a^{k+1}}{k+1} \quad \text{dla } k = 0, 1, \dots, \quad (6.29)$$

$$t_k = \int_a^b x^k f(x) dx \quad \text{dla } k = 0, 1, \dots. \quad (6.30)$$

Można wykazać, że jeżeli $m \leq n$, to wyznacznik układu (6.28) jest różny od zera, a więc układ ten ma jednoznaczne rozwiązanie. Rozwiązujeając powyższy układ równań liniowych, wyznaczamy współczynniki a_0, a_1, \dots, a_m wielomianu aproksymującego.

Przykład 6.7. Znaleźć wielomian drugiego stopnia najlepiej przybliżający, w sensie metryki średniokwadratowej, funkcję

$$f(x) = \sqrt{x}$$

w przedziale $[0 ; 2]$.

Rozwiązanie. Poszukujemy współczynników a_0, a_1, a_2 wielomianu

$$F_2(x) = a_0 + a_1x + a_2x^2.$$

Dla ich znalezienia formułujemy układ równań (6.28), gdy $m = 2$.

$$s_0 = \int_0^2 dx = 2, \quad s_1 = \int_0^2 x dx = 2, \quad s_2 = \int_0^2 x^2 dx = \frac{8}{3},$$

$$s_3 = \int_0^2 x^3 dx = 4, \quad s_4 = \int_0^2 x^4 dx = \frac{32}{5},$$

$$t_0 = \int_0^2 \sqrt{x} dx = \frac{4}{3}\sqrt{2}, \quad t_1 = \int_0^2 x\sqrt{x} dx = \frac{8}{5}\sqrt{2}, \quad t_2 = \int_0^2 x^2\sqrt{x} dx = \frac{16}{7}\sqrt{2}.$$

Otrzymujemy układ równań

$$\begin{bmatrix} 2 & 2 & \frac{8}{3} \\ 2 & \frac{8}{3} & 4 \\ \frac{8}{3} & 4 & \frac{32}{5} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}\sqrt{2} \\ \frac{8}{5}\sqrt{2} \\ \frac{16}{7}\sqrt{2} \end{bmatrix},$$

którego rozwiązaniem jest

$$a_0 = \frac{6}{35}\sqrt{2}, \quad a_1 = \frac{24}{35}\sqrt{2}, \quad a_2 = -\frac{\sqrt{2}}{7}.$$

Aby móc przekonać o jakości wyznaczonego przybliżenia, w tabeli 6.17 zostały stablicowane funkcja aproksymowana i aproksymująca w przedziale $[0 ; 2]$.

Tabela 6.17. Tablica wartości funkcji $y = \sqrt{x}$ i $y = F_2(x)$

x	$f(x) = \sqrt{x}$	$y = F_2(x)$	$\sqrt{x} - F_2(x)$
0.0	0.000000	0.242437	-0.242437
0.2	0.447214	0.428305	0.018909
0.4	0.632456	0.598010	0.034445
0.6	0.774597	0.751553	0.023043
0.8	0.894427	0.888934	0.005493
1.0	1.000000	1.010153	-0.010153
1.2	1.095445	1.115208	-0.019763
1.4	1.183216	1.204102	-0.020886
1.6	1.264911	1.276833	-0.011922
1.8	1.341641	1.333401	0.008239
2.0	1.414214	1.373807	0.040406

6.7. Aproksymacja średniokwadratowa całkowa wielomianami ortogonalnymi

Układ wielomianów $\varphi_0(x), \varphi_1(x), \dots, \varphi_k(x), \dots$ nazywamy *układem ortogonalnym z funkcją wagową $y = w(x)$ (z wagą $w(x)$)* na odcinku $[a ; b]$ wtedy i tylko wtedy, gdy

$$(\varphi_j, \varphi_k) = \int_a^b w(x) \varphi_j(x) \varphi_k(x) dx \begin{cases} = 0 & \text{dla } j \neq k, \\ > 0 & \text{dla } j = k. \end{cases}$$

Układ ortogonalny funkcji (w tym również wielomianów) jest układem funkcji liniowo niezależnych. Można zatem układ $\varphi_0(x), \varphi_1(x), \dots, \varphi_k(x), \dots$ potraktować jako bazę pewnej

przestrzeni funkcyjnej i aproksymować funkcję $y = f(x)$ za pomocą liniowej kombinacji funkcji bazy (za pomocą wielomianów uogólnionych).

W przypadku gdy funkcje $\varphi_0(x), \varphi_1(x), \dots, \varphi_k(x), \dots$ tworzą układ ortogonalny, na mocy definicji ortogonalności, mamy

$$(\varphi_j, \varphi_k) = 0 \quad \text{dla } j, k = 0, 1, \dots, m; j \neq k.$$

Układ równań (6.26) upraszcza się do postaci

$$\begin{aligned} g_{00}c_0 &= h_0, \\ g_{11}c_1 &= h_1, \\ &\dots, \\ g_{mm}c_m &= h_m. \end{aligned}$$

Ponieważ – na mocy definicji ortogonalności – mamy $g_{jj} > 0$ dla $j = 0, 1, \dots, m$, więc powyższy układ ma jednoznaczne rozwiązanie określone wzorami

$$c_j = \frac{h_j}{g_{jj}} = \frac{\int_a^b w(x) \varphi_j f(x) dx}{\int_a^b w(x) \varphi_j^2(x) dx} \quad \text{dla } j = 0, 1, \dots, m. \quad (6.31)$$

Powyższe wzory określają sposób obliczania współczynników wielomianu aproksymującego. Odchylenie średniokwadratowe funkcji aproksymowanej $y = f(x)$ od funkcji aproksymującej $y = \varphi(x)$ wyraża się wzorem

$$I_m^{\min} = \|f - \varphi\|^2 = \int_a^b w(x) f^2(x) dx - \sum_{k=0}^m \frac{\left(\int_a^b w(x) f(x) \varphi_k(x) dx \right)^2}{\int_a^b w(x) \varphi_k^2(x) dx}. \quad (6.32)$$

Wzór (6.32) wynika, że

$$I_0^{\min} \geq I_1^{\min} \geq I_2^{\min} \geq \dots,$$

wraz ze wzrostem stopnia wielomianu aproksymującego maleje monotonicznie odchylenie średniokwadratowe.

6.7.1. Aproksymacja za pomocą wielomianów Legendre'a

Wielomiany Legendre'a są określone wzorem Rodriguesa

$$P_k(x) = \frac{1}{2^k k!} \frac{d^k}{dx^k} (x^2 - 1)^k \quad \text{dla } k = 0, 1, \dots \quad (6.33)$$

W szczególności mamy:

$$\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x, \\ P_2(x) &= \frac{1}{2}(2x^2 - 1), \\ P_3(x) &= \frac{1}{2}(5x^3 - 3x), \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3), \\ P_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15). \end{aligned} \quad (6.34)$$

Ze wzoru (6.33) widać, że $P_k(x)$ są funkcjami parzystymi dla k parzystych i nieparzystymi dla k nieparzystych. Dalsze własności tych wielomianów są sformułowane w poniższych twierdzeniach.

Twierdzenie 6.3. Wielomiany Legendre'a tworzą układ ortogonalny w przedziale $[-1 ; 1]$, tj.

$$\int_{-1}^1 P_k(x) P_m(x) dx = 0 \quad \text{dla } k, m = 0, 1, 2, \dots, \quad k \neq m.$$

Twierdzenie 6.4. Kwadrat normy wielomianu Legendre'a jest równy

$$\|P_k\|^2 = \int_{-1}^1 P_k^2(x) dx = \frac{2}{2k+1} \quad \text{dla } k = 0, 1, \dots$$

Twierdzenie 6.5. Wielomian Legendre'a $P_k(x)$ dla $k \geq 1$ ma k różnych pierwiastków rzeczywistych, leżących w przedziale otwartym $(-1 ; 1)$.

Twierdzenie 6.6. Wielomian Legendre'a $P_k(x)$

a) wyraża się wzorem

$$P_k(x) = \frac{1}{2^k} \sum_{i=0}^{\left[\frac{k}{2}\right]} (-1)^i \binom{k}{i} \binom{2k-2i}{k} x^{k-2i} \quad \text{dla } k = 0, 1, \dots,$$

b) można wyznaczyć go ze wzoru rekurencyjnego

$$P_0(x) = 1,$$

$$P_1(x) = x,$$

$$P_{k+1}(x) = \frac{1}{k+1} ((2k+1) \cdot x \cdot P_k(x) - k \cdot P_{k-1}(x)) \quad \text{dla } k = 1, 2, \dots.$$

Z własności tych wynika, że funkcję $y=f(x)$ określoną i ciągłą na odcinku $[-1 ; 1]$ można przybliżać za pomocą wielomianów postaci

$$f(x) \approx Q_m(x) = \sum_{k=0}^m c_k P_k(x),$$

gdzie

$$c_k = \frac{2k+1}{2} \int_{-1}^1 f(x) P_k(x) dx.$$

Przykład 6.8. Stosując wielomiany Legendre'a, znaleźć wielomian drugiego stopnia najlepiej przybliżający – w sensie metryki średniokwadratowej – funkcję wymierną

$$f(u) = \frac{1}{u}$$

na odcinku $[0.5 ; 2.5]$.

Rozwiązanie. Ponieważ wielomiany Legendre'a tworzą układ ortogonalny na odcinku $[-1 ; 1]$ rozpoczęmy rozwiązanie zadania od znalezienia przekształcenia tego odcinka na odcinek $[0.5 ; 2.5]$. Łatwo znaleźć, że takim przekształceniem jest

$$u = x + \frac{3}{2} \quad \left(x = u - \frac{3}{2}, \quad -1 \leq x \leq 1, \quad \frac{1}{2} \leq u \leq \frac{5}{2} \right).$$

Poszukiwane przybliżenie będzie miało postać

$$\begin{aligned} f(u) &= f\left(x + \frac{3}{2}\right) = F_2^L(x) = c_0 P_0(x) + c_1 P_1(x) + c_2 P_2(x) = \\ &= c_0 P_0\left(u - \frac{3}{2}\right) + c_1 P_1\left(u - \frac{3}{2}\right) + c_2 P_2\left(u - \frac{3}{2}\right) = \\ &= c_0 + c_1 \left(u - \frac{3}{2}\right) + c_2 \cdot \frac{1}{2} \left(3\left(u - \frac{3}{2}\right)^2 - 1\right). \end{aligned}$$

Kolejno obliczamy:

$$c_0 = \frac{1}{2} \int_{-1}^1 \frac{dx}{x + 1.5} = \frac{1}{2} \ln 5 = 0.804719,$$

$$c_1 = \frac{3}{2} \int_{-1}^1 \frac{x dx}{x + 1.5} = 3 - \frac{9}{4} \ln 5 = -0.621235,$$

$$c_2 = \frac{5}{2} \int_{-1}^1 \frac{0.5(3x^2 - 1) dx}{x + 1.5} = \frac{115}{16} \ln 5 - \frac{45}{4} = 0.317835.$$

Ostatecznie mamy

$$F_2^L(u) = 0.804719 - 0.621235 \cdot \left(u - \frac{3}{2}\right) + 0.317835 \cdot \frac{1}{2} \left(3\left(u - \frac{3}{2}\right)^2 - 1\right).$$

W tabeli 6.18 zostały stabclicowane funkcje aproksymowana i aproksymująca w przedziale [0.5 ; 2.5].

Tabela 6.18. Tablica wartości funkcji $y = \frac{1}{x}$ i $y = F_2^L(x)$

x	$f(x) = \frac{1}{x}$	$y = F_2^L(x)$	$\frac{1}{x} - F_2^L(x)$
0.5	2.000000	1.743789	0.256211
0.7	1.428571	1.447911	-0.019340
0.9	1.111111	1.190174	-0.079062
1.1	0.909091	0.970576	-0.061485
1.3	0.769231	0.789119	-0.019888
1.5	0.666667	0.645801	0.020865
1.7	0.588235	0.540624	0.047611
1.9	0.526316	0.473588	0.052728
2.1	0.476190	0.444691	0.031499
2.3	0.434783	0.453935	-0.019152
2.5	0.400000	0.501319	-0.101319

6.7.2. Aproksymacja za pomocą wielomianów Czebyszewa pierwszego rodzaju

Wielomiany Czebyszewa pierwszego rodzaju są określone wzorami

$$T_k(x) = \frac{(x + \sqrt{x^2 - 1})^k + (x - \sqrt{x^2 - 1})^k}{2} \quad \text{dla } k = 0, 1, \dots . \quad (6.35)$$

Dla $|x| \leq 1$ wielomiany Czebyszewa możemy zapisać w tzw. postaci trygonometrycznej

$$T_k(x) = \cos(k \cdot \arccos x) \quad \text{dla } k = 0, 1, \dots \quad (6.36)$$

W szczególności mamy

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_2(x) &= 2x^2 - 1, \\ T_3(x) &= 4x^3 - 3x, \\ T_4(x) &= 8x^4 - 8x^2 + 1, \\ T_5(x) &= 16x^5 - 20x^3 + 5x. \end{aligned} \quad (6.37)$$

Ze wzoru (6.35) widać, że $T_k(x)$ są funkcjami parzystymi dla k parzystymi i nieparzystymi dla k nieparzystymi. Dalsze własności tych wielomianów są sformułowane w przedstawionych twierdzeniach.

Twierdzenie 6.7. Wielomiany Czebyszewa pierwszego rodzaju $T_k(x)$ można wyznaczyć ze wzoru rekurencyjnego:

$$T_0(x) = 1,$$

$$\begin{aligned} T_1(x) &= x, \\ T_{k+1}(x) &= 2 \cdot x \cdot T_k(x) - T_{k-1}(x) \quad \text{dla } k = 1, 2. \end{aligned}$$

Twierdzenie 6.8 (Wartości szczególne wielomianów Czebyszewa $T_k(x)$).

- a) $T_k(1) = 1$,
- b) $T_k(-1) = (-1)^k$,
- c) $T_{2k}(0) = (-1)^k$,
- d) $T_{2k+1}(0) = 0$.

Twierdzenie 6.9 (Pierwiastki wielomianu Czebyszewa $T_n(x)$).

Wielomian Czebyszewa $T_n(x)$ ($n \geq 1$) ma n pierwiastków rzeczywistych. Wszystkie pierwiastki są różne, leżą w przedziale otwartym $(-1 ; 1)$ i wyrażają się wzorami

$$x_k = \cos \frac{2k-1}{2n} \pi, \quad k = 1, 2, \dots, n.$$

Twierdzenie 6.10 (Własności ekstremalne wielomianu Czebyszewa $T_n(x)$).

Wielomian Czebyszewa $T_n(x)$ ($n \geq 1$) przyjmuje w przedziale domkniętym $[-1 ; 1]$ $n+1$ wartości ekstremalnych. Wartości ekstremalne osiągane są dla argumentów

$$y_k = \cos \frac{k\pi}{n}, \quad k = 0, 1, \dots, n$$

i są równe

$$T_n(y_k) = (-1)^k.$$

Twierdzenie 6.11 (Ortogonalność wielomianów Czebyszewa).

Wielomiany Czebyszewa pierwszego rodzaju tworzą na odcinku $[-1 ; 1]$ układ ortogonalny z funkcją wagową

$$w(x) = \frac{1}{\sqrt{1-x^2}}.$$

Ścisiej, zachodzą równości

$$\int_{-1}^1 \frac{T_k(x)T_m(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0 & \text{dla } k \neq m, \\ \pi & \text{dla } k = m = 0, \\ \frac{\pi}{2} & \text{dla } k = m \neq 0. \end{cases}$$

Funkcję ciągłą $y = f(x)$ w przedziale $[-1 ; 1]$ możemy aproksymować za pomocą sum wielciennych szeregu Czebyszewa

$$T(x) = \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k T_k(x).$$

Współczynniki optymalne – w sensie aproksymacji średniokwadratowej – wyrażają się wzorami

$$c_k = \frac{2}{\pi} \int_{-1}^1 \frac{f(x)T_k(x)}{\sqrt{1-x^2}} dx = \frac{2}{\pi} \int_0^\pi f(\cos x) \cos(kx) dx \quad \text{dla } k = 0, 1, \dots .$$

Wielomiany Czebyszewa pierwszego rodzaju mają szczególną własność, wyrażoną w poniższym twierdzeniu.

Twierdzenie 6.12. Wśród wszystkich wielomianów $w_n(x)$ stopnia n o najwyższym współczynniku równym 1 znormalizowany wielomian Czebyszewa pierwszego rodzaju stopnia n , tj.

$$\tilde{T}_n(x) = \frac{T_n(x)}{2^{n-1}},$$

najmniej odchyla się od zera w przedziale $[-1; 1]$ w sensie normy jednostajnej, tzn.

$$\max_{-1 \leq x \leq 1} |w_n(x)| \geq \max_{-1 \leq x \leq 1} |\tilde{T}_n(x)| = \frac{1}{2^{n-1}} .$$

Przykład 6.9. Stosując wielomiany Czebyszewa pierwszego rodzaju, znaleźć wielomian drugiego stopnia najlepiej przybliżający – w sensie metryki średniokwadratowej – funkcję wymierną

$$f(u) = 1/u$$

na odcinku $[0.5; 2.5]$.

Rozwiążanie. Ponieważ wielomiany Czebyszewa tworzą układ ortogonalny na odcinku $[-1; 1]$ rozpoczęmy rozwiązywanie zadania od znalezienia przekształcenia tego odcinka na odcinek $[0.5; 2.5]$. Łatwo znaleźć, że takim przekształceniem jest

$$u = x + \frac{3}{2} \quad \left(x = u - \frac{3}{2}, \quad -1 \leq x \leq 1, \quad \frac{1}{2} \leq u \leq \frac{5}{2} \right).$$

Poszukiwane przybliżenie będzie miało postać

$$\begin{aligned} f(u) &= f\left(x + \frac{3}{2}\right) \approx F_2^C(x) = c_0 T_0(x) + c_1 T_1(x) + c_2 T_2(x) = \\ &= c_0 T_0\left(u - \frac{3}{2}\right) + c_1 T_1\left(u - \frac{3}{2}\right) + c_2 T_2\left(u - \frac{3}{2}\right) = \\ &= \frac{c_0}{2} + c_1 \left(u - \frac{3}{2}\right) + c_2 \left(2\left(u - \frac{3}{2}\right)^2 - 1\right), \end{aligned}$$

gdzie

$$c_k = \frac{2}{\pi} \int_0^\pi \frac{\cos kt \cdot dt}{\cos t + 1.5} = \frac{4\sqrt{5}}{5} \left(\frac{\sqrt{5}-3}{2}\right)^k \quad \text{dla } k = 0, 1, 2.$$

Stąd

$$c_0 = \frac{4\sqrt{5}}{5} = 1.788854,$$

$$c_1 = \frac{2\sqrt{5}}{5}(\sqrt{5} - 3) = -0.683282,$$

$$c_2 = \frac{\sqrt{5}}{5}(\sqrt{5} - 3)^2 = 0.260990.$$

Ostatecznie mamy

$$F_2^C(u) = 0.894427 - 0.683282\left(u - \frac{3}{2}\right) + 0.260990\left(2\left(u - \frac{3}{2}\right)^2 - 1\right).$$

W tabeli 6.19 zostały stablicowane w przedziale [0.5 ; 2.5] funkcja aproksymowana $y = 1/x$ i aproksymująca $y = F_2^C(x)$. Do tabeli 6.19 dołączono kolumnę $\frac{1}{x} - F_2^L(x)$ z tabeli 6.18 dla porównania aproksymacji wielomianami Czebyszewa i wielomianami Lagrange'a.

Tabela 6.19. Tablica wartości funkcji $y = 1/x$ i $y = F_2^C(x)$

x	$f(x) = 1/x$	$y = F_2^C(x)$	$\frac{1}{x} - F_2^C(x)$	$\frac{1}{x} - F_2^L(x)$
0.5	2.000000	1.838699	0.161301	0.256211
0.7	1.428571	1.514130	-0.085558	-0.019340
0.9	1.111111	1.231319	-0.120208	-0.079062
1.1	0.909091	0.990266	-0.081175	-0.061485
1.3	0.769231	0.790972	-0.021742	-0.019888
1.5	0.666667	0.633437	0.033230	0.020865
1.7	0.588235	0.517660	0.070576	0.047611
1.9	0.526316	0.443641	0.082675	0.052728
2.1	0.476190	0.411381	0.064810	0.031499
2.3	0.434783	0.420879	0.013903	-0.019152
2.5	0.400000	0.472136	-0.072136	-0.101319

6.7.3. Aproksymacja za pomocą wielomianów Hermite'a

Wielomiany Hermite'a są określone wzorami

$$H_k(x) = (-1)^k e^{x^2} \cdot \frac{d^k}{dx^k}(e^{-x^2}) \quad \text{dla } k = 0, 1, \dots .$$

W szczególności mamy:

$$\begin{aligned}
 H_0(x) &= 1, \\
 H_1(x) &= 2x, \\
 H_2(x) &= 4x^2 - 2, \\
 H_3(x) &= 8x^3 - 12x, \\
 H_4(x) &= 16x^4 - 48x^2 + 12, \\
 H_5(x) &= 32x^5 - 160x^3 + 12x, \\
 H_6(x) &= 64x^6 - 480x^4 + 504x^2 - 120.
 \end{aligned}$$

Ze wzoru (6.33) widać, że $H_k(x)$ są funkcjami parzystymi dla k parzystych i nieparzystymi dla k nieparzystych. Dalsze własności tych wielomianów są sformułowane w dalszych twierdzeniach.

Twierdzenie 6.13. Wielomiany Hermite'a

a) wyrażają się wzorami

$$H_k(x) = k! \sum_{i=0}^{\text{entier}\left(\frac{k}{2}\right)} (-1)^i \frac{(2x)^{k-2i}}{i!(k-2i)!} \quad \text{dla } k = 0, 1, \dots,$$

b) można wyznaczyć ze wzoru rekurencyjnego

$$\begin{aligned}
 H_0(x) &= 1, \\
 H_1(x) &= 2x, \\
 H_{k+1}(x) &= 2 \cdot x \cdot H_k(x) - 2 \cdot k \cdot H_{k-1}(x) \quad k = 1, 2, \dots,
 \end{aligned}$$

c) spełniają równość

$$e^{-t^2 + 2tx} = \sum_{k=0}^{\infty} \frac{t^k}{k!} H_k(x).$$

Twierdzenie 6.14 (Wartości szczególne wielomianów Hermite'a).

- a) $H_{2n}(0) = (-1)^n \cdot 2^n \cdot (2n-1)!!$
- b) $H_{2n+1}(0) = 0$.

Twierdzenie 6.15. Wielomiany Hermite'a tworzą układ wielomianów ortogonalnych na prostej $(-\infty \leq x \leq +\infty)$ z funkcją wagową $w(x) = e^{-x^2}$. Dokładniej

$$\int_{-\infty}^{+\infty} e^{-x^2} H_k(x) H_m(x) dx = \begin{cases} 0 & \text{dla } k \neq m, \\ 2^k k! \sqrt{\pi} & \text{dla } k = m. \end{cases}$$

Zatem funkcję ciągłą $y = f(x)$ można na prostej $(-\infty \leq x \leq +\infty)$ przybliżać za pomocą wielomianów postaci

$$Q_m(x) = \sum_{k=0}^m c_k H_k(x),$$

gdzie

$$c_k = \frac{1}{2^k k! \sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} f(x) H_k(x) dx \quad \text{dla } k = 0, 1, \dots.$$

6.7.4. Aproksymacja za pomocą wielomianów trygonometrycznych

Układ trygonometryczny funkcji

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos(mx), \sin(mx), \dots$$

jest układem funkcji ortogonalnych na odcinku $[0 ; 2\pi]$ z funkcją wagową $y = w(x) \equiv 1$, ponieważ dla $k = 1, 2, \dots$ zachodzą równości

$$\int_0^{2\pi} \cos(kx) dx = 0, \quad \int_0^{2\pi} \sin(kx) dx = 0,$$

$$\int_0^{2\pi} \sin(kx) \cos(mx) dx = 0, \quad \int_0^{2\pi} \sin(kx) \sin(mx) dx = 0, \quad \int_0^{2\pi} \cos(kx) \cos(mx) dx = 0, \quad \text{dla } k \neq m$$

$$\int_0^{2\pi} 1^2 dx = 2\pi, \quad \int_0^{2\pi} \sin^2(kx) dx = \pi, \quad \int_0^{2\pi} \cos^2(mx) dx = \pi.$$

Zatem funkcję $y = f(x)$ ciągłą i okresową o okresie 2π z okresem głównym $[0 ; 2\pi]$ można przybliżać wielomianami trygonometrycznymi rzędu m :

$$S_m(x) = \frac{a_0}{2} + \sum_{k=1}^m (a_k \cos(kx) + b_k \sin(kx)).$$

Współczynniki optymalne – w sensie aproksymacji średniokwadratowej – wielomianu trygonometrycznego wyrażają się wzorami:

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx \quad \text{dla } k = 0, 1, \dots,$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx \quad \text{dla } k = 1, 2, \dots$$

i noszą nazwę *współczynników Fouriera* funkcji $y = f(x)$. Wielomian trygonometryczny ze współczynnikami Fouriera nosi nazwę *wielomianu Fouriera*.

Układ funkcji

$$1, \cos x, \cos 2x, \dots, \cos(mx), \dots$$

jest również układem ortogonalnym z wagą $w(x) = 1$ na odcinku $[0 ; 2\pi]$. Jeśli więc funkcja aproksymowana $y = f(x)$ jest *funkcją parzystą*, ciągłą i okresową o okresie 2π z okresem głównym $[0 ; 2\pi]$, to możemy ją przybliżać za pomocą funkcji postaci

$$Q_m(x) = \frac{a_0}{2} + \sum_{k=1}^m a_k \cos(kx).$$

Układ funkcji

$$\sin x, \sin 2x, \dots, \sin(mx), \dots$$

jest także układem ortogonalnym z wagą $w(x) = 1$ na odcinku $[0 ; 2\pi]$. Jeśli funkcja aproksymowana $y = f(x)$ jest *funkcją nieparzystą*, ciągłą i okresową o okresie 2π z okresem głównym $[0 ; 2\pi]$, to możemy ją przybliżać za pomocą funkcji postaci

$$Q_m(x) = \sum_{k=1}^m a_k \sin(kx).$$

Przykład 6.10. Zbudować wielomian Fouriera rzędu 4 dla funkcji

$$f(t) = 2(1 - |t|), \quad -1 \leq t \leq 1$$

okresowej, o okresie głównym $[-1 ; 1]$.

Rozwiążanie. Funkcja $y = f(x)$ jest ciągła i okresowa. Ponieważ wielomiany trygonometryczne są ortogonalne na odcinku $[0 ; 2\pi]$, znajdujemy przekształcenie odcinka $[-1 ; 1]$ na $[0 ; 2\pi]$. Takie przekształcenie ma postać

$$x = \pi \cdot t + \pi, \quad t = \frac{x}{\pi} - 1,$$

gdzie $-1 \leq t \leq 1$ oraz $0 \leq x \leq 2\pi$. Zauważmy ponadto, że $y = f(x)$ jest funkcją parzystą, zatem możemy ją aproksymować za pomocą wielomianów (6.28). Zgodnie ze wzorami (6.11) mamy

$$a_k = \frac{1}{\pi} \int_0^{2\pi} 2 \left(1 - \left|\frac{x}{\pi} - 1\right|\right) \cos(kx) dx = \int_{-1}^1 2(1 - |t|) \cos(k\pi t + k\pi) dt.$$

Stąd mamy

$$a_0 = \int_{-1}^1 2(1 - |t|) dt = 2,$$

$$a_1 = \int_{-1}^1 2(1 - |t|) \cos(\pi t + \pi) dt = -\frac{8}{\pi^2} = -0.810569,$$

$$a_2 = \int_{-1}^1 2(1 - |t|) \cos(2\pi t + 2\pi) dt = 0.$$

Ostatecznie mamy

$$Q_2(t) = \frac{a_0}{2} + a_1 \cos(\pi \cdot t + \pi) = 1 - 0.810569 \cdot \cos(\pi \cdot t + \pi) = 1 + 0.810569 \cdot \cos(\pi \cdot t).$$

W tabeli 6.20 zostały podane wartości funkcji aproksymowanej $f(t) = 2(1 - |t|)$ oraz wielomianu Fouriera $Q_2(t)$, stablicowanych w przedziale $[-1 ; 1]$ z krokiem $h = 0.2$.

Tabela 6.20. Stablicowane funkcje $y = 2(1 - |t|)$ i $y = Q_2(t)$

t	$f(t) = 2(1 - t)$	$y = Q_2(t)$	$f(t) - y = Q_2(t)$
-1.0	0.000000	0.189431	-0.189431
-0.8	0.400000	0.344236	0.055764
-0.6	0.800000	0.749520	0.050480
-0.4	1.200000	1.250480	-0.050480
-0.2	1.600000	1.655764	-0.055764
0.0	2.000000	1.810569	0.189431
0.2	1.600000	1.655764	-0.055764
0.4	1.200000	1.250480	-0.050480
0.6	0.800000	0.749520	0.050480
0.8	0.400000	0.344236	0.055764
1.0	0.000000	0.189431	-0.189431

6.8. Aproksymacja za pomocą wzorów empirycznych

Czasami mamy pewne informacje o postaci analitycznej funkcji aproksymującej. Bywa również tak, że funkcję aproksymującą dobieramy na podstawie podobieństwa jej wykresu do wykresu funkcji analizowanej. Takimi często stosowanymi wzorcami są:

- 1) $f_1(x) = ax^b + c$,
- 2) $f_2(x) = e^{ax^2 + bx + c}$,
- 3) $f_3(x) = ae^{bx} + c$,
- 4) $f_4(x) = ax + b$,
- 5) $f_5(x) = ax^2 + bx + c$,
- 6) $f_6(x) = \frac{ax + b}{cx + d}$,
- 7) $f_7(x) = \frac{1}{ax^2 + bx + c}$,
- 8) $f_8(x) = \frac{x}{ax^2 + bx + c}$,

$$9) \quad f_9(x) = a + \frac{b}{x} + \frac{c}{x^2},$$

$$10) \quad f_{10}(x) = ax^b e^{cx},$$

$$11) \quad f_{11}(x) = ae^{bx} + ce^{dx}.$$

Funkcje te mają mało parametrów oraz łatwą interpretację w modelu fizycznym. Znalezienie optymalnych współczynników a, b, c, d nie zawsze musi być łatwe. Spróbujemy to pokazać na przykładzie aproksymacji funkcji zadanej tabelą 6.1 za pomocą funkcji $y = f_3(x) = ae^{bx} + c$. Zgodnie z metodą najmniejszych kwadratów, należy wyznaczyć takie współczynniki a, b, c , dla których wyrażenie

$$D(a, b, c) = \sum_{i=0}^n (y_i - (ae^{bx_i} + c))^2$$

osiaga wartość minimalną. To zagadnienie prowadzi do zadania rozwiązania następującego układu równań:

$$\begin{aligned} \frac{\partial D(a, b, c)}{\partial a} &= \sum_{i=0}^n 2(y_i - (ae^{bx_i} + c)) \cdot e^{bx_i} = 0, \\ \frac{\partial D(a, b, c)}{\partial b} &= \sum_{i=0}^n 2(y_i - (ae^{bx_i} + c)) \cdot a \cdot e^{bx_i} \cdot x_i = 0, \\ \frac{\partial D(a, b, c)}{\partial c} &= \sum_{i=0}^n 2(y_i - (ae^{bx_i} + c)) = 0. \end{aligned}$$

Po wprowadzeniu oznaczeń $u_i = e^{x_i}$ ($i = 0, 1, \dots, n$), otrzymujemy do rozwiązania układ

$$\begin{aligned} \sum_{i=0}^n (y_i - (a \cdot u_i^b + c)) \cdot u_i^b &= 0, \\ \sum_{i=0}^n (y_i - (a \cdot u_i^b + c)) \cdot a \cdot u_i^b \cdot x_i &= 0, \\ \sum_{i=0}^n (y_i - (a \cdot u_i^b + c)) &= 0 \end{aligned}$$

trzech równań nieliniowych z trzema niewiadomymi a, b, c . Rozwiązanie tego układu równań jest problemem samym w sobie.

6.9. Zadania

Zad. 6.1. Dana jest funkcja dyskretna

x_i	x_0	x_1	...	x_n
$y_i = f(x_i)$	y_0	y_1	...	y_n

która chcemy przybliżyć za pomocą jednej z poniższych funkcji:

- a) $f_4(x) = ax + b$,
- b) $f_5(x) = ax^2 + bx + c$,
- c) $f_9(x) = a + \frac{b}{x} + \frac{c}{x^2}$,
- d) $f_{10}(x) = ax^b e^{cx}$.

Znaleźć wzory na optymalne – w sensie aproksymacji średniokwadratowej – współczynniki a, b, c .

Zad. 6.2. Dowieść, że wielomiany p_k (patrz rozdział 6.4.1) najniższych stopni p_0, p_1, p_2 tworzą układ ortogonalny z wagą $w(x) = 1$ na zbiorze $X = \{x_0, x_1, \dots, x_n\}$.

Zad. 6.3. Dowieść, że wielomiany Grama (patrz rozdział 6.4.2) najniższych stopni G_0, G_1, G_2 tworzą układ ortogonalny z wagą $w(x) = 1$ na zbiorze

$$X = \left\{ -1, \frac{2}{n}-1, \frac{4}{n}-1, \dots, 1 \right\}.$$

Zad. 6.4. Dowieść, że wielomiany Czebyszewa T_k (patrz rozdział 6.7.2) najniższych stopni T_0, T_1, T_2 tworzą układ ortogonalny z wagą

$$w(x) = \frac{1}{\sqrt{1-x^2}}$$

na odcinku $[-1; 1]$.

Zad. 6.5. W poniższej tabeli spisano wyniki przeprowadzonego eksperymentu. Przeprowadzający doświadczenie stwierdził, że analizowane zjawisko w przedziale $[0.0 ; 2.0]$ można opisać funkcją monotonicznie rosnącą i wypukłą. Eksperymentator stwierdził ponadto, że w tabeli wartości y_i dla $t = 0.4$ oraz $t = 1.2$ są obarczone błędem grubym. Wyznacz takie nowe wartości $y_i(0.4), y_i(1.2)$, aby odzwierciedlały przebieg badanego zjawiska.

t_i	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8
y_i	1.50	1.70	2.10	2.15	2.45	3.15	3.30	5.20	7.65	15.25

Zad. 6.6. Przebieg niektórych zjawisk dobrze opisuje funkcja

$$f(t) = \alpha_0 e^{-\beta t},$$

w której współczynniki α_0, β są dodatnimi liczbami rzeczywistymi. W poniższej tabeli zostały odnotowane wyniki doświadczenia tego rodzaju zjawiska.

t_i	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
y_i	4.10	2.65	1.75	1.25	0.80	0.55	0.40	0.30

Wyznaczyć wartości współczynników α_0, β dla przeprowadzonego doświadczenia.

Zad. 6.7. Dana jest funkcja $y = f(x)$ określona w przedziale $[a ; b]$. Należy:

- 1) stablicować funkcję $y = f(x)$ w przedziale $[a ; b]$ z krokiem $h = 0.2$;
- 2) skonstruować wielomian aproksymacyjny przybliżający funkcję dyskretną (otrzymaną w efekcie stablicowania funkcji $y = f(x)$) z zadaną dokładnością ε ;
- 3) wykonać – w tym samym układzie współrzędnych – wykresy funkcji $y = f(x)$ oraz wielomianu aproksymacyjnego. Wskazane jest dołączyć wykresy wielomianów aproksymacyjnych niższych stopni.

$$a) \quad f(x) = \begin{cases} x & \text{dla } 0 \leq x \leq 1 \\ 1 & \text{dla } 1 \leq x \leq 3, \\ \frac{-x+5}{2} & \text{dla } 3 \leq x \leq 5 \end{cases}$$

$$b) \quad f(x) = \begin{cases} x & \text{dla } 0 \leq x \leq 1 \\ \frac{1}{2}\sqrt{1-(x-2)^2} + 1 & \text{dla } 1 \leq x \leq 3, \\ \frac{-x+5}{2} & \text{dla } 3 \leq x \leq 5 \end{cases}$$

$$c) \quad f(x) = \begin{cases} x & \text{dla } 0 \leq x \leq 1 \\ \frac{1}{2}(x+1) & \text{dla } 1 \leq x \leq 3, \\ 5-x & \text{dla } 3 \leq x \leq 5 \end{cases}$$

$$d) \quad f(x) = \begin{cases} x & \text{dla } 0 \leq x \leq 2 \\ \frac{1}{3}(10-2x) & \text{dla } 2 \leq x \leq 5, \end{cases}$$

$$e) \quad f(x) = \sqrt{x} \quad \text{dla } 0 \leq x \leq 4,$$

$$f) \quad f(x) = \begin{cases} 1+\sqrt{1-(x-1)^2} & \text{dla } 0 \leq x \leq 2, \\ 1-\sqrt{1-(x-3)^2} & \text{dla } 2 \leq x \leq 5, \end{cases}$$

$$g) \quad f(x) = e^{-x^2} \quad \text{dla } -2 \leq x \leq 2,$$

$$h) \quad f(x) = \begin{cases} 0 & \text{dla } 0 \leq x \leq 1 \\ 1 & \text{dla } 1 < x < 3, \\ 0 & \text{dla } 3 \leq x \leq 4 \end{cases}$$

$$i) \quad f(x) = \begin{cases} \cosh x & \text{dla } -2 \leq x \leq 0 \\ -\cosh x & \text{dla } 0 \leq x \leq 2 \end{cases},$$

j) $f(x) = \begin{cases} 2x & \text{dla } 0 \leq x \leq 1 \\ -x+3 & \text{dla } 1 \leq x \leq 2 \\ x-1 & \text{dla } 2 \leq x \leq 3 \\ -2x+8 & \text{dla } 3 \leq x \leq 4 \end{cases}$

k) $f(x) = \begin{cases} x & \text{dla } 0 \leq x \leq 1 \\ -x+2 & \text{dla } 1 \leq x \leq 2 \\ x-2 & \text{dla } 2 \leq x \leq 3 \\ -x+4 & \text{dla } 3 \leq x \leq 4 \end{cases}$

l) $f(x) = \begin{cases} \sqrt{1-(x-1)^2} & \text{dla } 0 \leq x \leq 2 \\ \sqrt{1-(x-3)^2} & \text{dla } 2 \leq x \leq 4 \end{cases}$

m) $f(x) = \begin{cases} 1 & \text{dla } 0 \leq x \leq 1 \\ \frac{1}{2}(3-x) & \text{dla } 1 \leq x \leq 3 \end{cases}$

n) $f(x) = \begin{cases} 1 & \text{dla } 0 \leq x \leq 2 \\ (3-x) & \text{dla } 2 \leq x \leq 3 \end{cases}$

o) $f(x) = \begin{cases} 1 & \text{dla } 0 \leq x < 2 \\ 0 & \text{dla } 2 \leq x \leq 3 \end{cases}$

p) $f(x) = \begin{cases} 1 & \text{dla } 0 \leq x \leq 2 \\ \sqrt{1-(x-2)^2} & \text{dla } 2 \leq x \leq 3 \end{cases}$

q) $f(x) = \begin{cases} \frac{1}{2}x & \text{dla } 0 \leq x \leq 2 \\ \sqrt{1-(x-2)^2} & \text{dla } 2 \leq x \leq 3 \end{cases}$

r) $f(x) = \begin{cases} \sqrt{1-(x-1)^2} & \text{dla } 0 \leq x \leq 1 \\ 1 & \text{dla } 1 \leq x \leq 2 \\ \sqrt{1-(x-2)^2} & \text{dla } 2 \leq x \leq 3 \end{cases}$

s) $f(x) = \begin{cases} 1 & \text{dla } 0 \leq x \leq 1 \\ 0 & \text{dla } 1 < x < 2 \\ 1 & \text{dla } 2 \leq x \leq 3 \end{cases}$

t) $f(x) = \begin{cases} x & \text{dla } 0 \leq x \leq 1 \\ 1-\sqrt{1-(x-2)^2} & \text{dla } 1 \leq x \leq 3 \\ -x+4 & \text{dla } 3 \leq x \leq 4 \end{cases}$

Zad. 6.8. Funkcje zdefiniowane w zadaniu 6.7 przybliżyć z dokładnością ε według zasad aproksymacji średniokwadratowej całkowej. Wykonać – w tym samym układzie współrzędnych – wykresy funkcji aproksymowanej $y = f(x)$ oraz wielomianu aproksymacyjnego. Wskazane jest dołączyć wykresy wielomianów aproksymacyjnych niższych stopni.

Zad. 6.9. Dana jest funkcja okresowa $y = f(x)$ o okresie głównym $[a ; b]$. Znaleźć wielomian Fouriera najniższego stopnia przybliżający średniokwadratowo funkcję $y = f(x)$ z dokładnością ε .

a) $f(x) = 2\sqrt{1-x^2}$

z okresem głównym $[-1; 1]$,

b) $f(x) = \sqrt{1 - \frac{x^2}{4}}$

z okresem głównym $[-2; 2]$,

c) $f(x) = \begin{cases} 2-x^2 & \text{dla } -1 \leq x \leq 1 \\ (x-2)^2 & \text{dla } 1 \leq x \leq 3 \end{cases}$

z okresem głównym $[-1; 3]$,

d) $f(x) = \begin{cases} \sqrt{x} & \text{dla } 0 \leq x \leq 4 \\ 6-x & \text{dla } 4 \leq x \leq 6 \end{cases}$

z okresem głównym $[0; 6]$,

e) $f(x) = \cosh x$

z okresem głównym $[-1; 1]$,

f) $f(x) = \sin^3 x + \cos^3 x$

z okresem głównym $[0; 2\pi]$,

g) $f(x) = \sin^5 x + \cos^5 x$

z okresem głównym $[0; 2\pi]$,

h) $f(x) = |x|$

z okresem głównym $[-1; 1]$,

i) $f(x) = \begin{cases} 1+x & \text{dla } -1 \leq x \leq 0 \\ 1-x & \text{dla } 0 < x \leq 1 \end{cases}$

z okresem głównym $[-1; 1]$,

j) $f(x) = \begin{cases} 1+x & \text{dla } -1 \leq x \leq 0 \\ 1-x & \text{dla } 0 < x < 2 \\ x-3 & \text{dla } 2 \leq x \leq 3 \end{cases}$

z okresem głównym $[-1; 3]$,

k) $f(x) = \begin{cases} 1 & \text{dla } 0 \leq x \leq 1 \\ 2 & \text{dla } 1 < x < 3 \\ 1 & \text{dla } 3 \leq x \leq 4 \end{cases}$

z okresem głównym $[0; 4]$,

l) $f(x) = \begin{cases} 1 & \text{dla } 0 \leq x \leq 1 \\ (x-2)^2 & \text{dla } 1 < x \leq 3 \end{cases}$

z okresem głównym $[0; 3]$.

Różniczkowanie numeryczne

7.1. Sformułowanie zagadnienia

Różniczkowaniem numerycznym nazywamy obliczanie pochodnych funkcji tabelarycznej

Tabela 7.1. Tabela wartości funkcji dyskretniej $y = f(x)$

x_i	x_0	x_1	...	x_n
$y_i = f(x_i)$	y_0	y_1	...	y_n

Ponieważ funkcja dyskretna nie jest funkcją różniczkowalną (nie jest nawet funkcją ciągłą), zagadnienie różniczkowania funkcji zadanej tabelą 7.1 definiuje się następująco: funkcję $y = f(x)$ zastępuje się funkcją różniczkowalną $y = F(x)$, której pochodne dobrze przybliżają pochodne funkcji $y = f(x)$, a następnie przyjmuje się

$$f^{(k)}(x) \approx F^{(k)}(x) \text{ dla } k = 1, 2, \dots \quad (7.1)$$

Ścisłe rzecz biorąc obliczona wartość $F^{(k)}(x)$ jest obarczona błędem metody różniczkowania

$$r(k, x) = f^{(k)}(x) - F^{(k)}(x). \quad (7.2)$$

Najczęściej w charakterze funkcji $y = F(x)$ przyjmuje się wielomian, bowiem wielomiany mają ciągle pochodne wszystkich rzędów dla każdego rzeczywistego x oraz przybliżają jednostajnie funkcję ciągłą $y = f(x)$ na odcinku domkniętym $[a ; b]$.

7.2. Różniczkowanie z wykorzystaniem interpolacyjnych funkcji sklejanych

Dobre przybliżenia pochodnych pierwszego, drugiego i trzeciego rzędu otrzymuje się, wybierając w charakterze funkcji $y = F(x)$ funkcję sklejaną trzeciego stopnia $y = S_3(\Delta, x)$, określoną wzorem (5.35). Ogólny algorytm obliczania pochodnych jest następujący.

Algorytm różniczkowania z wykorzystaniem funkcji sklejanych trzeciego stopnia

1. Wyznaczamy wartości parametrów M_0, M_1, \dots, M_n poprzez rozwiązanie układu równań liniowych (5.34).
2. Dla $x \in [x_j ; x_{j+1}]$ wartości pochodnych obliczamy ze wzorów

$$S'_3(\Delta, x) = \beta_j + 2\gamma_j(x - x_j) + 3\delta_j(x - x_j)^2, \quad (7.3)$$

$$S_3''(\Delta, x) = 2\gamma_j + 6\delta_j(x - x_j), \quad (7.4)$$

$$S_3'''(\Delta, x) = 6\delta_j, \quad (7.5)$$

gdzie

$$\beta_j = \frac{S_3'(\Delta, x_j)}{1!} = \frac{y_{j+1} - y_j}{h_{j+1}} - \frac{2M_j + M_{j+1}}{6} h_{j+1},$$

$$\gamma_j = \frac{S_3''(\Delta, x_j)}{2!} = \frac{M_j}{2},$$

$$\delta_j = \frac{S_3'''(\Delta, x_j)}{3!} = \frac{M_{j+1} - M_j}{6h_{j+1}}.$$

Jakość przybliżenia pochodnych funkcji $y = f(x)$ przez pochodne funkcji sklejanej $y = S_3(\Delta, x)$ można ocenić, opierając się na nierównościach:

$$|f'(x) - S_3'(\Delta, x)| \leq \frac{7}{4} KM_4 \|\Delta\|^3, \quad (7.6)$$

$$|f''(x) - S_3''(\Delta, x)| \leq \frac{7}{4} KM_4 \|\Delta\|^2, \quad (7.7)$$

$$|f'''(x) - S_3'''(\Delta, x)| \leq 2KM_4 \|\Delta\|, \quad (7.8)$$

gdzie:

$$\|\Delta\| = \max_{0 \leq j \leq n-1} |x_{j+1} - x_j|, \quad K = \max_{0 \leq j \leq n-1} \frac{\|\Delta\|}{|x_{j+1} - x_j|}, \quad M_4 = \max_{a \leq x \leq b} |f^{(IV)}(x)|.$$

Z nierówności (7.6) – (7.8) wynika, że błąd bezwzględny różniczkowania możemy zmniejszać poprzez zmniejszanie wielkości $\|\Delta\|$.

Przykład 7.1. O funkcji $y = f(x)$ wiemy, że

Tabela 7.2. Tabela znanych wartości funkcji $y = f(x)$

x_j	$\frac{1}{2}$	1	$\frac{3}{2}$	2
$y_j = f(x_j)$	2	1	$\frac{2}{3}$	$\frac{1}{2}$
$y'_j = f'(x_j)$	-4			$-\frac{1}{4}$

Wyznaczyć $y = f'(x)$, $y = f''(x)$.

Rozwiązanie. Funkcje $y = S_3'(\Delta, x)$ i $y = S_3''(\Delta, x)$ – stanowiące przybliżenia funkcji $y = f'(x)$ oraz $y = f''(x)$ – wyznaczymy ze wzorów (7.3) i (7.4).

$$S'_3(\Delta, x) = \begin{cases} -4 + \frac{173}{15}\left(x - \frac{1}{2}\right) - \frac{53}{5}\left(x - \frac{1}{2}\right)^2 & \text{dla } \frac{1}{2} \leq x \leq 1 \\ -\frac{53}{60} + \frac{14}{15}(x-1) - \frac{1}{5}(x-1)^2 & \text{dla } 1 \leq x \leq \frac{3}{2}, \\ -\frac{7}{15} + \frac{11}{15}\left(x - \frac{3}{2}\right) - \frac{3}{5}\left(x - \frac{3}{2}\right)^2 & \text{dla } \frac{3}{2} \leq x \leq 2 \end{cases} \quad (7.9)$$

$$S''_3(\Delta, x) = \begin{cases} \frac{173}{15} - \frac{106}{5}\left(x - \frac{1}{2}\right) & \text{dla } \frac{1}{2} \leq x \leq 1 \\ \frac{14}{15} - \frac{2}{15}(x-1) & \text{dla } 1 \leq x \leq \frac{3}{2}. \\ \frac{11}{15} - \frac{6}{5}\left(x - \frac{3}{2}\right) & \text{dla } \frac{3}{2} \leq x \leq 2 \end{cases} \quad (7.10)$$

Błąd bezwzględny szacujemy na podstawie wzorów (7.6) i (7.7). Na odcinku [0.5 ; 2] mamy: $\|\Delta\| = 0.5$, $K = 1$, $M_4 = 24 \cdot 2^5$;

$$|f'(x) - S'_3(\Delta, x)| \leq \frac{7}{4} \cdot 1 \cdot 24 \cdot 2^5 \cdot \frac{1}{2^3} = 168,$$

$$|f''(x) - S''_3(\Delta, x)| \leq \frac{7}{4} \cdot 1 \cdot 24 \cdot 2^5 \cdot \frac{1}{2^2} = 336.$$

Tabela 7.3. Tablica wartości pierwszej i drugiej pochodnej funkcji z tabeli 7.2

x	$f'(x)$	$S'_3(x)$	$f'(x) - S'_3(x)$	$f''(x)$	$S''_3(x)$	$f''(x) - S''_3(x)$
0,5	-4,0000	-4,0000	0,0000	16,0000	11,5333	4,4667
0,6	-2,7778	-2,9527	0,1749	9,2593	9,4133	-0,1541
0,7	-2,0408	-2,1173	0,0765	5,8309	7,2933	-1,4624
0,8	-1,5625	-1,4940	-0,0685	3,9063	5,1733	-1,2671
0,9	-1,2346	-1,0827	-0,1519	2,7435	3,0533	-0,3098
1,0	-1,0000	-0,8833	-0,1167	2,0000	0,9333	1,0667
1,1	-0,8264	-0,7920	-0,0344	1,5026	0,8933	0,6093
1,2	-0,6944	-0,7047	0,0102	1,1574	0,8533	0,3041
1,3	-0,5917	-0,6213	0,0296	0,9103	0,8133	0,0970
1,4	-0,5102	-0,5420	0,0318	0,7289	0,7733	-0,0445
1,5	-0,4444	-0,4667	0,0222	0,5926	0,7333	-0,1407
1,6	-0,3906	-0,3993	0,0087	0,4883	0,6133	-0,1251
1,7	-0,3460	-0,3440	-0,0020	0,4071	0,4933	-0,0863
1,8	-0,3086	-0,3007	-0,0080	0,3429	0,3733	-0,0304
1,9	-0,2770	-0,2693	-0,0077	0,2916	0,2533	0,0383
2,0	-0,2500	-0,2500	0,0000	0,2500	0,1333	0,1167

Ostatnie oszacowania mówią o małej wartości otrzymanych wzorów (7.9) i (7.10). Jedną z interpretacji tego faktu może być duży krok ($h = 0.5$) stablicowania funkcji $y = f(x)$. Inną interpretacją mogą być bardzo ostre nierówności (7.6)÷(7.7). Precyzyjniejsze informacje o różniczkowaniu są zawarte w tabeli 7.3.

7.3. Różniczkowanie z wykorzystaniem wzoru interpolacyjnego Stirlinga

W przypadku węzłów równoodległych, można je ponumerować tak jak w tabeli 7.4

Tabela 7.4. Tabela wartości funkcji z węzłami ponumerowanymi do użycia wzoru Sterlinga

x_i	...	x_{-k}	...	x_{-1}	x_0	x_1	...	x_k	...
$y_i = f(x_i)$...	y_{-k}	...	x_{-1}	x_0	x_1	...	y_k	...

i w charakterze funkcji $y = F(x)$ przyjąć wielomian interpolacyjny Stirlinga

$$F(x) = S_{2n}(x) = y_0 + \sum_{k=1}^n \left\{ \frac{q \cdot \prod_{l=1}^{k-1} (q^2 - l^2)}{(2k-1)!} \cdot \frac{\Delta^{2k-1} y_{-k} + \Delta^{2k-1} y_{-k+1}}{2} + \frac{\prod_{l=0}^{k-1} (q^2 - l^2)}{(2k)!} \cdot \Delta^{2k} y_{-k} \right\}$$

Błąd interpolacji z zastosowaniem wzoru Stirlinga jest w przybliżeniu równy

$$R_{2n}(x) = f(x) - S_{2n}(x) \approx \frac{q(q^2 - 1^2) \cdot \dots \cdot (q^2 - n^2)}{(2n+1)!} \cdot \frac{\Delta^{2n+1} y_{-n} + \Delta^{2n+1} y_{-n-1}}{2}.$$

Różniczkując wzór Stirlinga mamy

$$\begin{aligned} \frac{dS_{2n}(x)}{dx} &= \frac{dS_{2n}(x)}{dq} \cdot \frac{dq}{dx} = \frac{1}{h} \cdot \frac{dS_{2n}(x)}{dq} = \\ &= \frac{1}{h} \cdot \sum_{k=1}^n \left\{ \frac{\prod_{l=1}^{k-1} (q^2 - l^2) + \sum_{p=1}^{k-1} 2q^2 \prod_{\substack{l=1 \\ l \neq p}}^{k-1} (q^2 - l^2)}{(2k-1)!} \cdot \frac{\Delta^{2k-1} y_{-k} + \Delta^{2k-1} y_{-k+1}}{2} + \right. \\ &\quad \left. + \frac{\sum_{p=0}^{k-1} 2q \prod_{\substack{l=0 \\ l \neq p}}^{k-1} (q^2 - l^2)}{(2k)!} \cdot \Delta^{2k} y_{-k} \right\}. \end{aligned}$$

Błąd różniczkowania możemy obliczyć z następującego wzoru:

$$\begin{aligned}
r_{2n}(1, x) &= \frac{df(x)}{dx} - \frac{dS_{2n}(x)}{dx} \approx \frac{dR_{2n}(x)}{dx} = \\
&= \frac{1}{h} \cdot \frac{\prod_{l=1}^n (q^2 - l^2) + \sum_{p=1}^n 2q^2 \prod_{\substack{l=1 \\ l \neq p}}^n (q^2 - l^2)}{(2n+1)!} \cdot \frac{\Delta^{2n+1} y_{-n} + \Delta^{2n+1} y_{-n-1}}{2}.
\end{aligned}$$

Powyższe wzory są mało czytelne. Stają się one bardziej użyteczne w przypadku, gdy obliczamy $S'_{2n}(x)$ dla $x = x_0$. W tym przypadku $q = 0$ oraz

$$\frac{dS_{2n}(x_0)}{dx} = \frac{1}{h} \cdot \sum_{k=1}^n (-1)^{k-1} \frac{(k-1)!}{k(k+1) \cdot \dots \cdot (2k-1)} \cdot \frac{\Delta^{2k-1} y_{-k} + \Delta^{2k-1} y_{-k+1}}{2}, \quad (7.11)$$

$$r_{2n}(1, x_0) \approx \frac{(-1)^n}{h} \cdot \frac{n!}{(n+1) \cdot \dots \cdot (2n+1)} \cdot \frac{\Delta^{2n+1} y_{-n} + \Delta^{2n+1} y_{-n-1}}{2}. \quad (7.12)$$

Przykład 7.2. W tabeli 6.6 podano współrzędne kuli pchniętej przez lekkoatletę. Obliczyć prędkość kuli w różnych punktach toru.

Rozwiązanie. Tabela 6.6 zawiera wartości funkcji, której zmienna niezależna x zmienia się ze stałym krokiem $h = 0.5$. Do obliczenia pochodnych tej funkcji zastosujemy wzór (7.11). Do oceny błędu różniczkowania zostanie wykorzystany wzór (7.12). Obliczenia zostały zestawione w tabeli 7.5.

Tabela 7.5. Obliczanie pochodnych ze wzoru (7.11)

i	x_i	$y_i = f(x_i)$	$\Delta^1 y_i$	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$	$\Delta^5 y_i$	$S'_{2}(x_i)$	$r_2(1, x_i)$	$S'_{4}(x_i)$	$r_4(1, x_i)$
1	2	3	4	5	6	7	8	9	10	11	12
0	0.00	1.60	0.41	-0.02	-0.01	0.03	-0.07				
1	0.50	2.01	0.39	-0.03	0.02	-0.04	0.07				
2	1.00	2.40	0.36	-0.01	-0.02	0.03	-0.04				
3	1.50	2.76	0.35	-0.03	0.01	-0.01	0.01	0.71	-0.0017		
4	2.00	3.11	0.32	-0.02	0.00	0.00	0.00	0.67	0.0000		
5	2.50	3.43	0.30	-0.02	0.00	0.00	0.00	0.62	0.0017		
6	3.00	3.73	0.28	-0.02	0.00	0.00	0.00	0.58	-0.0017	0.58	0.0000
7	3.50	4.01	0.26	-0.02	0.00	0.00	-0.01	0.54	0.0000	0.54	0.0010
8	4.00	4.27	0.24	-0.02	0.00	-0.01	0.03	0.50	0.0000	0.50	-0.0010
9	4.50	4.51	0.22	-0.02	-0.01	0.02	-0.03	0.46	0.0000	0.46	0.0003
10	5.00	4.73	0.20	-0.03	0.01	-0.01	0.01	0.42	0.0000	0.42	0.0000
11	5.50	4.93	0.17	-0.02	0.00	0.00	0.00	0.37	0.0017	0.37	0.0000
12	6.00	5.10	0.15	-0.02	0.00	0.00	-0.09	0.32	0.0000	0.32	-0.0003
13	6.50	5.25	0.13	-0.02	0.00	-0.09	0.45	0.28	-0.0017	0.28	0.0007

7.4. Zadania

Zad. 7.1. Pokazać, że jeśli w charakterze funkcji $y = F(x)$ przyjmiemy pierwszy wielomian interpolacyjny Newtona $P_n^I(x)$ (rozdział 5.5.2), to

$$\frac{dP_n^I(x_0)}{dx} = \frac{1}{h} \sum_{k=1}^n (-1)^{k-1} \frac{\Delta^k y_0}{k},$$

$$r_n(1, x_0) = \frac{df(x_0)}{dx} - \frac{dP_n^I(x_0)}{dx} \approx (-1)^{n+1} \frac{\Delta^{n+1} y_0}{n+1}.$$

Zad. 7.2. Pokazać, że jeśli w charakterze funkcji $y = F(x)$ przyjmiemy drugi wielomian interpolacyjny Newtona $P_n^{II}(x)$ (rozdział 5.5.3), to

$$\frac{dP_n^{II}(x_n)}{dx} = \frac{1}{h} \sum_{k=1}^n \frac{\Delta^k y_{n-k}}{k},$$

$$r_n(1, x_n) = \frac{df(x_n)}{dx} - \frac{dP_n^{II}(x_n)}{dx} \approx \frac{\Delta^{n+1} y_{-1}}{n+1}.$$

Zad. 7.3. Dana jest w postaci tabeli funkcja dyskretna $y = f(x)$. Obliczyć $y = f'(x_i)$, $f''(x_i)$ w każdym węźle interpolacji. Podane w tabelach nazwy stabilowanej funkcji mogą posłużyć do sprawdzenia jakości różniczkowania numerycznego.

a)

x_i	0.0	0.2	0.4	0.6	0.8	1.0	1.2
$y_i = e^{x_i}$	1.000000	1.221403	1.491825	1.822119	2.225541	2.718282	3.320117

x_i	1.4	1.6	1.8	2	2.2	2.4	2.6
$y_i = e^{x_i}$	4.055200	4.953032	6.049647	7.389056	9.025013	11.023176	13.463738

b)

x_i	0.0	0.2	0.4	0.6	0.8	1.0	1.2
$y_i = \sin x_i$	0.000000	0.198669	0.389418	0.564642	0.717356	0.841471	0.932039

x_i	1.4	1.6	1.8	2	2.2	2.4	2.6
$y_i = \sin x_i$	0.985450	0.999574	0.973848	0.909297	0.808496	0.675463	0.515501

c)

x_i	0.0	0.2	0.4	0.6	0.8	1.0	1.2
$y_i = \arctg x_i$	0.000000	0.197396	0.380506	0.540420	0.674741	0.785398	0.876058

x_i	1.4	1.6	1.8	2.0	2.2	2.4	2.6
$y_i = \operatorname{arctg} x_i$	0.950547	1.012197	1.063698	1.107149	1.144169	1.176005	1.203622

d)

x_i	1.0	1.2	1.4	1.6	1.8	2.0	2.2
$y_i = \ln x_i$	0.000000	0.182322	0.336472	0.470004	0.587787	0.693147	0.788457

x_i	2.4	2.6	2.8	3	3.2	3.4	3.6
$y_i = \ln x_i$	0.875469	0.955511	1.029619	1.098612	1.163151	1.223775	1.280934

e)

x_i	0.0	0.2	0.4	0.6	0.8	1.0	1.2
$y_i = \sqrt[4]{x_i}$	0.000000	0.668740	0.795271	0.880112	0.945742	1.000000	1.046635

x_i	2.4	2.6	2.8	3	3.2	3.4	3.6
$y_i = \sqrt[4]{x_i}$	1.244666	1.269823	1.293569	1.316074	1.337481	1.357906	1.377449

f)

x_i	0.2	0.4	0.6	0.8	1.0	1.2	1.4
$y_i = 1/x_i$	5.000000	2.500000	1.666667	1.250000	1.000000	0.833333	0.714286

x_i	2.6	2.8	3.0	3.2	3.4	3.6	3.8
$y_i = 1/x_i$	0.384615	0.357143	0.333333	0.312500	0.294118	0.277778	0.263158

g)

x_i	0.0	0.2	0.4	0.6	0.8	1.0
$y_i = \sin^3 x_i + \cos^3 x_i$	1.000000	0.949225	0.840439	0.742221	0.707333	0.753552

x_i	1.2	1.4	1.6	1.8	2	2.2
$y_i = \sin^3 x_i + \cos^3 x_i$	0.857238	0.961891	0.998696	0.911848	0.679759	0.324669

h)

x_i	0.0	0.2	0.4	0.6	0.8
$y_i = (2x_i - 1)(2x_i - 3)(2x_i - 7)$	-21.000000	-10.296000	-2.728000	2.088000	4.536000

x_i	1.0	1.2	1.4	1.6	1.8
$y_i = (2x_i - 1)(2x_i - 3)(2x_i - 7)$	5.000000	3.864000	1.512000	-1.672000	-5.304000

x_i	2.0	2.2	2.4	2.6	2.8
$y_i = (2x_i - 1)(2x_i - 3)(2x_i - 7)$	-9.000000	-12.376000	-15.048000	-16.632000	-16.744000

j)

x_i	0.0	0.2	0.4	0.6	0.8	1.0
$y_i = \sin 3x_i + \cos 3x_i$	1.000000	1.389978	1.294397	0.746646	-0.061931	-0.848872

x_i	1.2	1.4	1.6	1.8	2.0	2.2
$y_i = \sin 3x_i + \cos 3x_i$	-1.339279	-1.361837	-0.908666	-0.138072	0.680755	1.261774

k)

x_i	0.2	0.4	0.6	0.8	1.0	1.2
$y_i = \log x_i$	-0.698970	-0.397940	-0.221849	-0.096910	0.000000	0.079181

x_i	1.4	1.6	1.8	2.0	2.2	2.4
$y_i = \log x_i$	0.146128	0.204120	0.255273	0.301030	0.342423	0.380211

l)

x_i	0.1	0.3	0.5	0.7	0.9	1.1
$y_i = \operatorname{ctg} x_i$	9.966644	3.232728	1.830488	1.187242	0.793551	0.508968

x_i	1.3	1.5	1.7	1.9	2.1	2.3
$y_i = \operatorname{ctg} x_i$	0.277616	0.070915	-0.129927	-0.341635	-0.584848	-0.893484

m)

x_i	0.1	0.3	0.5	0.7	0.9	1.1
$y_i = \frac{2}{x_i^2}$	200.000000	22.222222	8.000000	4.081633	2.469136	1.652893

x_i	1.3	1.5	1.7	1.9	2.1	2.3
$y_i = \frac{2}{x_i^2}$	1.183432	0.888889	0.692042	0.554017	0.453515	0.378072

Całkowanie numeryczne

8.1. Sformułowanie zagadnienia

Całkowaniem numerycznym nazywamy obliczanie całki oznaczonej funkcji dyskretnej.

Tabela 8.1. Tabela wartości funkcji dyskretnej $y = f(x)$

x_i	x_0	x_1	...	x_n
$y_i = f(x_i)$	y_0	y_1	...	y_n

Obliczanie numeryczne całki jednokrotnej nosi nazwę *kwadratury*, a całki wielokrotnej *kubatury*. Odpowiednie wzory całkowania numerycznego nazywamy *wzorami kwadratur (kubatur)*.

Kwadratura (kubatura) polega na zastąpieniu w przedziale całkowania $[a ; b]$ funkcji podcałkowej $y = f(x)$ funkcją aproksymującą lub interpolującą $y = F(x)$ i przyjęciu równości

$$\int_a^b f(x)dx = \int_a^b F(x)dx + R_F . \quad (8.1)$$

R_F w powyższym wzorze oznacza *błąd metody całkowania*.

Całkowanie numeryczne umożliwia obliczenie całki oznaczonej funkcji empirycznej. W tym przypadku tabela 8.1 przedstawia spisane wyniki określonego doświadczenia. Całkowanie numeryczne umożliwia również obliczenie całki oznaczonej funkcji, której funkcja pierwotna nie daje się przedstawić w postaci superpozycji skończonej ilości funkcji elementarnych. Przykładami takich funkcji są: $f(x) = e^{-x^2}$, $g(x) = \sin x^2$. Aby obliczyć całkę oznaczoną takiej funkcji, należy najpierw utworzyć tablicę wartości tej funkcji, a następnie zastosować metodę całkowania numerycznego.

Jeśli w charakterze funkcji $y = F(x)$ przyjmiemy wielomian interpolacyjny, to wzór (8.1) przyjmie postać

$$\int_a^b f(x)dx = \sum_{j=0}^n A_j f(x_j) + R_F , \quad (8.2)$$

w którym x_j są wybranymi węzłami interpolacji, A_j stałymi współczynnikami zależnymi od sposobu wyboru węzłów i niezależnymi od postaci funkcji $y = f(x)$, a R_F błędem metody całkowania.

8.2. Kwadratury Newtona-Cotesa

Wzory całkowania numerycznego, w których przyjęto, że wszystkie węzły interpolacji są rozmieszczone w jednakowych odstępach, noszą nazwę wzorów *Newtona-Cotesa*. Praktycznie nie stosuje się wzorów Newtona-Cotesa opartych na wielomianach interpolacyjnych wysokich stopni. Dalej przedstawiamy najczęściej stosowane wzory: wzór trapezów i wzór parabol (wzór *Simpsona*).

8.2.1. Złożony wzór trapezów

Przedział całkowania $[a ; b]$ dzielimy na n podprzedziałów równej długości. Mamy zatem:

$$h_n = \frac{b-a}{n},$$

$$x_0 = a, x_j = x_0 + jh_n \quad \text{dla } j = 1, 2, \dots, n \quad (x_n = b).$$

W każdym podprzedziale $[x_j ; x_{j+1}]$ ($j = 0, 1, \dots, n - 1$) funkcję podcałkową $y = f(x)$ zastępujemy wielomianem pierwszego stopnia (wielomiany są na ogół różne w różnych podprzedziałach). Wzór (8.2) przyjmuje postać

$$\int_a^b f(x) dx = h_n \cdot \left(\frac{1}{2} y_0 + y_1 + y_2 + \dots + y_{n-1} + \frac{1}{2} y_n \right) - \frac{(b-a)h_n^2}{12} f''(c_n), \quad (8.3)$$

przy czym $a < c_n < b$. Wzór

$$I_n = h_n \cdot \left(\frac{1}{2} y_0 + y_1 + y_2 + \dots + y_{n-1} + \frac{1}{2} y_n \right) \quad (8.4)$$

określa algorytm obliczania przybliżonej wartości całki i nosi nazwę *ogólnego wzoru trapezów*. Wyrażenie

$$R_1 = -\frac{(b-a)h_n^2}{12} f''(c_n)$$

określa błąd całkowania metodą trapezów.

Przykład 8.1. Obliczyć

$$\int_0^2 e^{-\frac{x^2}{2}} dx \quad (8.5)$$

z dokładnością ε stosując wzór trapezów.

Rozwiązanie. Rozwiązywanie zadania rozpoczęmy od wyznaczenia liczby n podprzedziałów, na które należy podzielić przedział $[0 ; 2]$, aby całka obliczona według wzoru (8.4) była obarczona błędem nie większym niż ε . Liczba n jest najmniejszą liczbą całkowitą, spełniającą nierówność

$$\begin{aligned}|R_1| &= \left| -\frac{(b-a)h_n^2}{12} f''(c_n) \right| = \left| \frac{(b-a)^3}{12 \cdot n^2} f''(c_n) \right| \leq \frac{(b-a)^3}{12 \cdot n^2} \cdot \max_{0 \leq x \leq 2} |f''(x)| = \\ &= \frac{(b-a)^3 M_2}{12 \cdot n^2} \leq \varepsilon.\end{aligned}$$

Tak postawione zadanie wymaga obliczenia $M_2 = \max_{0 \leq x \leq 2} |f''(x)|$. Kolejno wyznaczamy

$$f(x) = e^{-\frac{x^2}{2}}, \quad f'(x) = -xe^{-\frac{x^2}{2}}, \quad f''(x) = (x^2 - 1)e^{-\frac{x^2}{2}},$$

stąd

$$M_2 = \max_{0 \leq x \leq 2} |f''(x)| = |f''(0)| = 1.$$

W tabeli 8.2 podano wyznaczone z nierówności

$$n \geq \sqrt{\frac{(b-a)^3 M_2}{12 \cdot \varepsilon}}$$

liczby podprzedziałów n dla kilku dokładności ε .

Tabela 8.2. Liczba podprzedziałów n w metodzie trapezów jako funkcja dokładności ε

Dokładność ε	Liczba podprzedziałów n dla całki (8.5)
10^{-2}	9
10^{-3}	26
10^{-4}	82

Przyjmując $n = 10$, mamy

$$\begin{aligned}\int_0^2 e^{-\frac{x^2}{2}} dx &= 0.2 \cdot \left(\frac{1}{2} f(0.0) + f(0.2) + f(0.4) + f(0.6) + f(0.8) + f(1.0) + \right. \\ &\quad \left. + f(1.2) + f(1.4) + f(1.6) + f(1.8) + \frac{1}{2} f(2.0) \right) + R_1 = \\ &= 0.2 \cdot \left(\frac{1}{2} (1.000 + 0.980 + 0.923 + 0.835 + 0.726 + 0.607 + \right. \\ &\quad \left. + 0.487 + 0.375 + 0.278 + 0.198 + \frac{1}{2} 0.135) \right) + R_1 = 1.195 + R_1.\end{aligned}$$

8.2.2. Złożony wzór parabol (wzór Simpsona)

Przedział całkowania $[a ; b]$ dzielimy na $n = 2m$ podprzedziałów równej długości. Mamy zatem:

$$h_m = \frac{b-a}{2m}, \quad x_0 = a, \quad x_j = x_0 + jh_m \text{ dla } j = 1, 2, \dots, n \quad (x_n = b).$$

W podprzedziałach podwójnej długości $[x_{2j} ; x_{2j+2}]$ ($j = 0, 1, \dots, m - 1$) funkcję podcałkową $y = f(x)$ zastępujemy wielomianem drugiego stopnia. Wzór (8.2) przyjmuje w tym przypadku postać

$$\int_a^b f(x) dx = \frac{h_m}{3} \cdot (y_0 + y_{2m} + 2(y_2 + y_4 + \dots + y_{2m-2}) + 4(y_1 + y_3 + \dots + y_{2m-1})) - \frac{(b-a)h_m^4}{180} f^{(IV)}(c_m), \quad (8.6)$$

gdzie $a < c_m < b$. Wyrażenie

$$I_{2m} = \frac{h_m}{3} \cdot (y_0 + y_{2m} + 2(y_2 + y_4 + \dots + y_{2m-2}) + 4(y_1 + y_3 + \dots + y_{2m-1})) \quad (8.7)$$

określa algorytm obliczania przybliżonej wartości całki i nosi nazwę *ogólnego wzoru Simpsona (wzoru parabol)*. Błąd całkowania metodą Simpsona wyraża się wzorem

$$R_2 = -\frac{(b-a)h_m^4}{180} f^{(IV)}(c_m).$$

Przykład 8.2. Stosując wzór Simpsona, obliczyć

$$\int_0^2 e^{-\frac{x^2}{2}} dx$$

z dokładnością ε .

Rozwiązanie. Podobnie jak w przykładzie 8.1, rozwiązywanie zadania rozpoczęliśmy od wyznaczenia liczby $n = 2m$ podprzedziałów, na które należy podzielić przedział $[0 ; 2]$, aby całka obliczona według wzoru (8.7) była obarczona błędem nie większym niż ε . Liczba n jest najmniejszą liczbą całkowitą parzystą, spełniającą nierówność

$$\begin{aligned} |R_2| &= \left| -\frac{(b-a)h_m^4}{180} f^{(IV)}(c_m) \right| = \left| \frac{(b-a)^5}{180 \cdot n^4} f^{(IV)}(c_m) \right| \leq \frac{(b-a)^5}{180 \cdot n^4} \cdot \max_{0 \leq x \leq 2} |f^{(IV)}(x)| = \\ &= \frac{(b-a)^5 M_4}{180 \cdot n^4} \leq \varepsilon. \end{aligned}$$

A więc trzeba obliczyć $M_4 = \max_{0 \leq x \leq 2} |f^{(IV)}(x)|$. Kolejno wyznaczamy

$$f'''(x) = (-x^3 + 3x) e^{-\frac{x^2}{2}}, \quad f^{(IV)}(x) = (x^4 - 6x^2 + 3) e^{-\frac{x^2}{2}}.$$

Nietrudno sprawdzić, że

$$M_4 = \max_{0 \leq x \leq 2} |f^{(IV)}(x)| = |f^{(IV)}(0)| = 3.$$

W tabeli 8.3 podano wyznaczone z nierówności

$$n \geq \sqrt{\frac{(b-a)^5 M_4}{180 \cdot \varepsilon}}$$

parzyste liczby podprzedziałów n dla trzech dokładności ε .

Tabela 8.3. Liczba podprzedziałów n w metodzie Simpsona jako funkcja dokładności ε

Dokładność ε	Liczba podprzedziałów n dla całki (8.5)
10^{-2}	4
10^{-3}	6
10^{-4}	10

Przyjmując $n = 6$, mamy

$$\begin{aligned} \int_0^2 e^{-\frac{x^2}{2}} dx &= \frac{1}{9} \cdot \left(f(0) + 2 \cdot \left(f\left(\frac{2}{3}\right) + f\left(\frac{4}{3}\right) \right) + 4 \cdot \left(f\left(\frac{1}{3}\right) + f(1) + f\left(\frac{5}{3}\right) \right) + f(2) \right) + R_2 = \\ &= \frac{1}{9} \cdot (1.0000 + 2 \cdot (0.8007 + 0.4111) + 4 \cdot (0.9460 + 0.6065 + 0.2494) + 0.1353) + \\ &+ R_2 = 1.1963 + R_2. \end{aligned}$$

8.3. Ekstrapolacja Richardsona

Stosowanie wzorów (8.3), (8.6) wymaga wyznaczenia odpowiednio

$$M_2 = \max_{a \leq x \leq b} |f''(x)|$$

lub

$$M_4 = \max_{a \leq x \leq b} |f^{(IV)}(x)|.$$

Tę niedogodność możemy pokonać, stosując algorytm zwany *ekstrapolacją Richardsona*. Polega on na utworzeniu ciągu przybliżeń wielkości

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

poprzez podział przedziału całkowania $[a ; b]$ na coraz większą liczbę podprzedziałów i zastosowanie wzoru trapezów (lub Simpsona) z przybliżoną oceną błędu całkowania. Jest prawdopodobne, że taki ciąg będzie monotonicznie zbieżny do wielkości I .

8.3.1. Ekstrapolacja Richardsoна z wykorzystaniem wzoru trapezów

Jeżeli przedział $[a ; b]$ podzielimy na n podprzedziałów równej długości i przybliżoną wartość całki

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

obliczymy ze wzoru trapezów, to

$$I = I_n - \frac{(b-a)^3}{12 \cdot n^2} \cdot f''(c_n). \quad (8.8)$$

Jeśli innym razem przedział $[a ; b]$ podzielimy na k ($k > n$) podprzedziałów równej długości i przybliżoną wartość całki policzymy ze wzoru trapezów, to

$$I = I_k - \frac{(b-a)^3}{12 \cdot k^2} \cdot f''(c_k), \quad (8.9)$$

gdzie $c_n, c_k \in (a ; b)$. Jeśli założymy, że $f''(c_n) \approx f''(c_k)$, to eliminując obydwie wartości z równań (8.8) i (8.9), otrzymujemy

$$I \approx \frac{k^2 I_k - n^2 I_n}{k^2 - n^2} = I_k + \frac{n^2}{k^2 - n^2} \cdot (I_k - I_n). \quad (8.10)$$

Porównując wzory (8.9) i (8.10) stwierdzamy, że znaleźliśmy przybliżony wzór na błąd całkowania metodą trapezów, tj.

$$R_1 = - \frac{(b-a)^3}{12 \cdot k^2} \cdot f''(c_k) \approx \frac{n^2}{k^2 - n^2} (I_k - I_n). \quad (8.11)$$

Wzór (8.11) jest tym dokładniejszy, im mniej różnią się pochodne $f''(c_n), f''(c_k)$. Wartość tego wzoru jest niebagatelna, gdyż umożliwia nam oszacowanie wartości całki policzonej metodą trapezów bez wyznaczania pochodnych funkcji podcałkowych.

Przykład 8.3. Obliczyć

$$\int_0^2 e^{-\frac{x^2}{2}} dx$$

z dokładnością ε , stosując ekstrapolację Richardsoна z wykorzystaniem wzoru trapezów.

Tabela 8.4. Ekstrapolacja Richardsoña z wykorzystaniem wzoru trapezów całki (8.5)

n	I_n	$(I_n - I_{n-1})/3$
2	1,17420	
4	1,19067	0,005490
8	1,19488	0,001403
16	1,19594	0,000353
32	1,19620	0,000087

Rozwiązanie. Ze względu na wygodę obliczeń przyjmiemy zasadę, że przedział dzielimy na połowy. W tabeli 8.4 zostały zestawione przybliżone wartości całki (8.5) policzone metodą trapezów, gdy przedział całkowania $[0 ; 2]$ dzielimy kolejno na 2, 4, 8, 32 podprzedziały.

8.3.2. Ekstrapolacja Richardsona z wykorzystaniem wzoru Simpsona

Jeśli przedział $[a ; b]$ podzielimy na $n = 2m$ podprzedziałów równej długości i wartość całki

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

obliczymy ze wzoru (8.7), to

$$I = I_m - \frac{(b-a)^5}{180 \cdot 2^4 m^4} \cdot f^{(IV)}(c_m). \quad (8.12)$$

Jeśli innym razem przedział $[a ; b]$ podzielimy na $2p$ ($p > m$) podprzedziałów równej długości i przybliżoną wartość całki policzymy ze wzoru Simpsona, to

$$I = I_p - \frac{(b-a)^5}{180 \cdot 2^4 p^4} \cdot f^{(IV)}(c_p), \quad (8.13)$$

gdzie $c_m, c_p \in (a ; b)$. Jeśli założymy, że $f^{(IV)}(c_m) \approx f^{(IV)}(c_p)$, to eliminując obydwie wartości z równań (8.12) i (8.13), otrzymujemy

$$I \approx \frac{p^4 I_p - m^4 I_m}{p^4 - m^4} = I_p + \frac{m^4}{p^4 - m^4} \cdot (I_p - I_m). \quad (8.14)$$

Porównując wzory (8.13) i (8.14) stwierdzamy, że ostatni jest wzorem całkowania metodą Simpsona z przybliżoną oceną błędu całkowania, tj.

$$R_2 = - \frac{(b-a)^5}{180 \cdot 2^4 p^4} \cdot f^{(IV)}(c_p) \approx \frac{m^4}{p^4 - m^4} \cdot (I_p - I_m). \quad (8.15)$$

Wzór (8.15) jest tym dokładniejszy, im mniej różnią się pochodne $f^{(IV)}(c_m), f^{(IV)}(c_p)$. Wzór (8.14) umożliwia obliczenie (przybliżone) całki metodą Simpsona bez wyznaczania

$$M_4 = \max_{a \leq x \leq b} |f^{(IV)}(x)|.$$

Przykład 8.4. Obliczyć

$$\int_0^2 e^{-\frac{x^2}{2}} dx$$

z dokładnością ε , stosując ekstrapolację Richardsona z wykorzystaniem wzoru Simpsona.

Rozwiązanie. Podobnie jak w przykładzie (8.3) przyjmiemy zasadę, że przedziały dzielimy na połowy. W tabeli 8.5 zostały zestawione przybliżone wartości całki (8.5) policzone metodą Simpsona, gdy przedział całkowania $[0 ; 2]$ dzielimy kolejno na $p = 2, 4, 8$ ($p = 2m$) podprzedziałów.

Tabela 8.5. Ekstrapolacja Richardsoна z wykorzystaniem wzoru Simpsona całki (8.8)

P	I_p	$(I_p - I_{p-1})/15$
2	1.18715	
4	1.19617	0.00060
8	1.19628	0.00001

8.4. Kwadratury Gaussa

Wzór kwadratur Gaussa ma postać

$$\int_a^b f(x)dx = \frac{b-a}{2} \cdot \sum_{i=1}^n A_i f(x_i) + R_n, \quad (8.16)$$

gdzie

$$x_i = \frac{b-a}{2} t_i + \frac{a+b}{2} \quad (i = 1, 2, \dots, n),$$

a t_1, t_2, \dots, t_n są pierwiastkami wielomianu Legendre'a $P_n(t)$ n -tego stopnia. Współczynniki A_1, A_2, \dots, A_n otrzymuje się przez rozwiązanie następującego układu równań liniowych

$$A_1 t_1^k + A_2 t_2^k + \dots + A_n t_n^k = \frac{1 - (-1)^{k+1}}{k+1} \quad (k = 0, 1, \dots, n-1).$$

Błąd R_n kwadratur Gaussa jest równy

$$R_n = \frac{(b-a)^{2n+1} (n!)^4 f^{(2n)}(c)}{(2n+1)[(2n)!]^3} \quad \text{dla pewnego } c \ (a < c < b). \quad (8.17)$$

Wzór kwadratur Gaussa został tak skonstruowany, że jest dokładny dla wielomianów stopni do $2n-1$ włącznie.

Przykład 8.5. Obliczyć

$$\int_0^2 e^{-\frac{x^2}{2}} dx,$$

stosując kwadratury Gaussa.

Rozwiązanie. Rozwiążanie zadania rozpoczęmy od dyskusji na temat wyboru stopnia wielomianu interpolacyjnego. Do oceny dokładności całkowania służy wzór (8.17). Z tego wzoru otrzymujemy

$$|R_n| \leq \frac{(b-a)^{2n+1} (n!)^4 M_{2n}}{(2n+1)[(2n)!]^3},$$

gdzie $M_{2n} = \max_{a \leq x \leq b} |f^{(2n)}(x)|$. Oszacujmy błąd dla $n = 3$;

$$f(x) = e^{-\frac{x^2}{2}}, \quad f^{(VI)}(x) = (x^6 - 15x^4 + 45x^2 - 15) \cdot e^{-\frac{x^2}{2}}.$$

Stąd $M_6 = |f^{(IV)}(0)| = 15$. A zatem

$$|R_3| \leq \frac{2^7 \cdot (3!)^4 \cdot 15}{7 \cdot (6!)^3} = \frac{1}{1050} = 0.00095.$$

Założymy, że ta dokładność jest satysfakcyjna. Przystępujemy do obliczania całki. Z twierdzenia 6.6 mamy

$$P_3(t) = \frac{1}{2} \cdot (5t^3 - 3t).$$

Wielomian ten ma 3 pierwiastki rzeczywiste

$$t_1 = -\sqrt{\frac{3}{5}}, \quad t_2 = 0, \quad t_3 = \sqrt{\frac{3}{5}}.$$

Ponieważ $x_i = t_i + 1$, więc

$$x_1 = 1 - \sqrt{\frac{3}{5}}, \quad x_2 = 1, \quad x_3 = 1 + \sqrt{\frac{3}{5}}.$$

Współczynniki A_1, A_2, A_3 otrzymamy, rozwiązuając układ równań

$$\begin{array}{lcl} +A_1 & +A_2 & +A_3 \\ -\sqrt{\frac{3}{5}} \cdot A_1 & +0 \cdot A_2 & +\sqrt{\frac{3}{5}} \cdot A_3 \\ \frac{3}{5} \cdot A_1 & +0 \cdot A_2 & +\frac{3}{5} \cdot A_3 \end{array} = 2, \quad = 0, \quad = \frac{2}{3}.$$

Rozwiązuając ten układ otrzymujemy

$$A_1 = A_3 = \frac{5}{9}, \quad A_2 = \frac{8}{9}.$$

Ostatecznie, na mocy wzoru (8.16), mamy

$$\int_0^2 e^{-\frac{x^2}{2}} dx = \frac{2}{2} \cdot \left(A_1 e^{-\frac{x_1^2}{2}} + A_1 e^{-\frac{x_2^2}{2}} + A_1 e^{-\frac{x_3^2}{2}} \right) + R_3 = 1.1958 + R_3.$$

8.5. Zadania

Zad. 8.1. Wzór Newtona (reguła trzech ósmych).

Przedział całkowania $[a; b]$ dzielimy na $n = 3m$ podprzedziałów równej długości. Mamy zatem

$$h_m = \frac{b-a}{3m}, \quad x_0 = a, \quad x_j = jh_m \quad \text{dla } j = 1, 2, \dots, n \quad (x_n = b).$$

Na podprzedziałach potrójnej długości $[x_{3j}; x_{3j+3}]$ ($j = 0, 1, \dots, m - 1$) funkcję podcałkową $y = f(x)$ zastępujemy wielomianem trzeciego stopnia. Pokazać, że w tym przypadku wzór (8.2) przyjmuje postać

$$\int_a^b f(x) dx = \frac{3}{8} \cdot h_m \cdot (y_0 + y_{3m} + 2(y_3 + y_6 + \dots + y_{3m-3}) + \\ + 3(y_1 + y_2 + y_4 + y_5 + \dots + y_{3m-2} + y_{3m-1})) - \frac{(b-a)h_m^4}{80} f^{(IV)}(c_m),$$

gdzie $a < c_m < b$.

Zad. 8.2. Wykazać, że w kwadraturach Gaussa opartych na wielomianach interpolacyjnych stopnia jeden i dwa węzły t_i i współczynniki A_i są równe

n	t_i	A_i
1	$t_1 = 0$	$A_1 = 2$
2	$-t_1 = t_2 = \frac{\sqrt{3}}{3}$	$A_1 = A_2 = 1$

Zad. 8.3. Oblicz z dokładnością ε następujące całki:

a) $\int_0^a e^{x^2} dx \quad (a > 0),$

b) $\int_0^a \frac{e^{bx}}{x} dx \quad (a > 0, b > 0),$

c) $\int_0^a \frac{e^{bx}}{x^2} dx \quad (a > 0, b > 0),$

d) $\int_0^a \frac{e^{bx}}{x^3} dx \quad (a > 0, b > 0),$

e) $\int_0^a \frac{dx}{\sin^3 x} \quad (a > 0),$

f) $\int_0^a \frac{dx}{\sin^5 x} \quad (a > 0),$

g) $\int_0^a \frac{dx}{\cos^3 x} \quad (a > 0),$

h) $\int_0^a \frac{dx}{\cos^5 x} \quad (a > 0),$

i) $\int_0^a \operatorname{tg}^3 x dx \quad (a > 0),$

j) $\int_1^a \operatorname{ctg}^3 x dx \quad (a > 1),$

k) $\int_0^a e^{-px^2} x^2 dx \quad (a > 0, p > 0),$

l) $\int_0^a \frac{e^{-px}}{x^2} dx \quad (a > 0, p > 0),$

m) $\int_0^a e^{-px} \sin^{2k} x dx \quad (a > 0, p > 0, k \geq 1),$

n) $\int_0^a e^{-px} \cos^{2k} x dx \quad (a > 0, p > 0, k \geq 1),$

o) $\int_0^a \frac{\ln(1+x)}{1+x^2} dx \quad (a > 0),$

p) $\int_0^a \ln(1+x) x^{p-1} dx \quad (a > 0, p > 0),$

q) $\int_0^a e^{-px} \ln x dx \quad (a > 0, p > 0),$

r) $\int_0^a \frac{\operatorname{arctg} x}{x} dx \quad (a > 0),$

s) $\int_0^a \frac{x \cdot \operatorname{arctg} x}{1-x^4} dx \quad (a > 0),$

t) $\int_0^a x^p \operatorname{arctg} x dx \quad (a > 0, p > 0).$

Zad. 8.4. Oblicz z dokładnością ε całki zdefiniowane w zadaniu 1.5.

Zad. 8.5. Oblicz z dokładnością ε następujące całki:

a) $\int_a^b \frac{dx}{\sqrt{1-k^2 \sin^2 x}} \quad (0 < k < 1),$

- b) $\int_0^{\frac{\pi}{4}} \frac{x}{\sin x + \cos x} \cdot \frac{dx}{\cos x},$
- c) $\int_0^1 e^{-px \ln x} dx,$
- d) $\int_0^a \sin^2 x \cdot \frac{dx}{x^2} \quad (a > 0),$
- e) $\int_0^a \operatorname{tg} x \cdot \frac{dx}{x} \quad (a > 0),$
- f) $\int_0^a (1 - \cos qx) e^{-px} \cdot \frac{dx}{x} \quad (a > 0, p > 0),$
- g) $\int_0^1 \frac{\ln x dx}{1 + x + x^2},$
- h) $\int_0^1 \frac{x \ln x dx}{1 + x + x^2},$
- i) $\int_0^1 \ln x \cdot \frac{x^{2n} dx}{1 + x} \quad (n \geq 1),$
- j) $\int_0^{\frac{\pi}{4}} \ln \cos x dx,$
- k) $\int_0^a \frac{\cos px dx}{\cosh qx} \quad (a > 0, q \neq 0),$
- l) $\int_0^a \frac{\sin px dx}{\sinh qx} \quad (a > 0, q \neq 0),$
- m) $\int_0^a x e^{-p^2 x^2} \cos qx dx \quad (a > 0, q \neq 0),$
- n) $\int_0^a x e^{-p^2 x^2} \operatorname{tg} qx dx \quad (a > 0, p \neq 0),$
- o) $\int_0^1 \frac{\operatorname{tg} x}{\sin x + \cos x} dx,$

p) $\int_a^b \frac{dx}{2 + \cos x^2},$

q) $\int_a^b e^{-4x^3 + 2x^2 + 1} dx,$

r) $\int_a^b e^{2\cos x} dx,$

s) $\int_a^b x^x dx \quad (b > a > 0),$

t) $\int_0^a \frac{\sin x dx}{1+x^2} \quad (a > 0).$



DODATKI

A. Wzory sumacyjne

$$\sum_{k=1}^n k = \frac{n(n+1)}{2} = s_1 \quad (\text{A.1})$$

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6} = s_2 \quad (\text{A.2})$$

$$\sum_{k=1}^n k^3 = \left(\frac{n(n+1)}{2} \right)^2 = s_1^2 = s_3 \quad (\text{A.3})$$

$$\sum_{k=1}^n k^4 = \frac{1}{5} \cdot s_2 \cdot (3n^2 + 3n - 1) = \binom{n+1}{5} + 11\binom{n+2}{5} + 11\binom{n+3}{5} + \binom{n+4}{5} \quad (\text{A.4})$$

$$\sum_{k=1}^n k(k+1) = \frac{1}{3} \cdot n(n+1)(n+2) \quad (\text{A.5})$$

$$\sum_{k=1}^n k(k+1)^2 = \frac{1}{12} \cdot n(n+1)(n+2)(3n+5) \quad (\text{A.6})$$

$$\sum_{k=1}^n k(k+1)^3 = \frac{1}{60} \cdot n(n+1)(n+2)(12n^2 + 39n + 29) \quad (\text{A.7})$$

$$\sum_{k=1}^n k(k+1)(k+2) = \frac{1}{4} \cdot n(n+1)(n+2)(n+3) \quad (\text{A.8})$$

$$\sum_{k=1}^n k(k+1)(k+2)^2 = \frac{1}{20} \cdot n(n+1)(n+2)(n+3)(4n+11) \quad (\text{A.9})$$

$$\sum_{k=1}^n k(k+1)(k+2)^3 = \frac{1}{60} \cdot n(n+1)(n+2)(n+3)(10n^2 + 54n + 71) \quad (\text{A.10})$$

B. Normy wektorów i macierzy

Niech

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}$$

oznaczają wektory przestrzeni R^n , a α niech oznacza liczbę rzeczywistą. Wyrażenie

$$\|X\|_p = \sqrt[p]{\sum_{i=1}^n |x_i|^p} \quad (\text{B.1})$$

jest dla każdego rzeczywistego p z przedziału $1 \leq p \leq \infty$ normą, tzn. dla każdych $X, Y \in R^n$ oraz dowolnej liczby rzeczywistej α są spełnione własności:

$$\|X\|_p \geq 0 \quad \text{oraz} \quad (\|X\|_p = 0) \Leftrightarrow X = O \quad (\text{B.2})$$

$$\|\alpha X\|_p = |\alpha| \cdot \|X\|_p \quad (\text{B.3})$$

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p \quad (\text{nierówność trójkąta}). \quad (\text{B.4})$$

Wyrażenie (B.1) nosi nazwę p -tej normy Höldera. W praktyce stosowane są trzy normy:

$$\|X\|_1 = \sum_{i=1}^n |x_i| \quad (\text{norma pierwsza}), \quad (\text{B.5})$$

$$\|X\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} \quad (\text{norma druga lub euklidesowa}), \quad (\text{B.6})$$

$$\|X\|_\infty = \max_{1 \leq i \leq n} |x_i| \quad (\text{norma nieskończoność}). \quad (\text{B.7})$$

Dla dowolnego $X \in R^n$ zachodzą nierówności

$$\|X\|_\infty \leq \|X\|_2 \leq \|X\|_1 \leq \sqrt{n} \|X\|_2 \leq n \|X\|_\infty. \quad (\text{B.8})$$

Niech

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

będzie macierzą o m wierszach i n kolumnach. Normą macierzy A indukowaną przez normę wektorową nazywamy liczbę

$$\|A\|_p = \sup_{\substack{X \in R^n \\ X \neq O}} \frac{\|AX\|_p}{\|X\|_p}. \quad (\text{B.9})$$

Odpowiednio dla $p = 1, 2, \infty$ normy indukowane przybierają postać

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|, \quad (\text{B.10})$$

$$\|A\|_2 = \max_{\lambda \in \text{Spec}(A^T A)} \sqrt{\lambda} \quad (\text{B.11})$$

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|. \quad (\text{B.12})$$

Przykład B.1. Niech

$$X = \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \quad A = \begin{bmatrix} 3 & 6 & -2 & 9 \\ 2 & -4 & 3 & -6 \\ 4 & 5 & -7 & 8 \end{bmatrix}.$$

Obliczyć $\|X\|_1, \|X\|_2, \|X\|_\infty, \|A\|_1, \|A\|_E, \|A\|_\infty$.

Rozwiązanie. Kolejno mamy

$$\|X\|_1 = |3| + |-4| + |5| = 12,$$

$$\|X\|_2 = \sqrt{3^2 + (-4)^2 + 5^2} = \sqrt{50},$$

$$\|X\|_\infty = \max\{|3|, |-4|, |5|\} = 5,$$

$$\begin{aligned} \|A\|_1 &= \max\{|3| + |2| + |4|, |6| + |-4| + |5|, |-2| + |3| + |-7|, |9| + |-6| + |8|\} = \\ &= \max\{9, 15, 12, 23\} = 23, \end{aligned}$$

$$\begin{aligned} \|A\|_E &= \sqrt{3^2 + 6^2 + (-2)^2 + 9^2 + 2^2 + (-4)^2 + 3^2 + (-6)^2 + 4^2 + 5^2 + (-7)^2 + 8^2} = \\ &= \sqrt{349}, \end{aligned}$$

$$\begin{aligned} \|A\|_\infty &= \max\{|3| + |6| + |-2| + |9|, |2| + |-4| + |3| + |-6|, |4| + |5| + |-7| + |8|\} = \\ &= \max\{20, 15, 24\} = 24. \end{aligned}$$

Wprowadźmy dalsze oznaczenia

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}, \quad C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1q} \\ c_{21} & c_{22} & \dots & c_{2q} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nq} \end{bmatrix}.$$

Normy macierzy mają podobne własności jak normy wektorów, a mianowicie:

$$\|A\|_p \geq 0 \quad \text{oraz} \quad \|A\|_p = 0 \Leftrightarrow A = O, \quad (\text{B.13})$$

$$\|\alpha A\|_p = |\alpha| \cdot \|A\|_p, \quad (\text{B.14})$$

$$\|A + B\|_p \leq \|A\|_p + \|B\|_p \quad (\text{nierówność trójkąta}), \quad (\text{B.15})$$

$$\|A + B\|_p \geq \|A\|_p - \|B\|_p, \quad (\text{B.16})$$

$$\|AC\|_p \leq \|A\|_p \cdot \|C\|_p. \quad (\text{B.17})$$

Z ostatniej nierówności wynika, że

$$\|AX\|_p \leq \|A\|_p \cdot \|X\|_p. \quad (\text{B.18})$$

Lemat B.1. Niech M będzie macierzą kwadratową stopnia n , a E niech oznacza macierz jednostkową stopnia n . Jeżeli $\|M\|_p < 1$, to macierz $E + M$ jest nieosobliwa oraz

$$\|(E + M)\|_p \leq \frac{1}{1 - \|M\|_p} \quad \text{dla } p = 1, \infty. \quad (\text{B.19})$$

C. Liczby i wielomiany Bernoulliego

Liczbami Bernoulliego B_k ($k = 0, 1, \dots$) nazywamy współczynniki przy $x^k/k!$ w rozwińciu funkcji $x/(e^x - 1)$ w szereg Maclaurina, tzn. w rozwińciu

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} B_k \frac{x^k}{k!}. \quad (\text{C.1})$$

Liczby Bernoulliego można obliczyć z następującego wzoru rekurencyjnego

$$B_0 = 1,$$

$$\sum_{i=0}^k \binom{k+1}{i} B_i = 0 \quad \text{dla } k = 1, 2, \dots \quad (\text{C.2})$$

Wszystkie liczby Bernoulliego są liczbami wymiernymi. Liczby o numerach nieparzystych są, z wyjątkiem $B_1 = -(1/2)$, równe zero. Liczby o numerach parzystych, począwszy od wyrazu $B_2 = 1/6$, są na przemian dodatnie i ujemne. Początkowe liczby Bernoulliego:

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0,$$

$$B_4 = -\frac{1}{30}, \quad B_5 = 0, \quad B_6 = \frac{1}{42}, \quad B_7 = 0,$$

$$B_8 = -\frac{1}{30}, \quad B_9 = 0, \quad B_{10} = \frac{5}{66}, \quad B_{11} = 0,$$

$$B_{12} = -\frac{691}{2730}, \quad B_{13} = 0, \quad B_{14} = \frac{7}{6}, \quad B_{15} = 0,$$

$$B_{16} = -\frac{3617}{510}, \quad B_{17} = 0, \quad B_{18} = \frac{43867}{798}, \quad B_{19} = 0.$$

Wielomianami Bernoulliego $B_k(x)$ ($k = 0, 1, \dots$) nazywamy współczynniki przy $t^k/k!$ w rozwińciu funkcji

$$f(t, x) = \frac{te^{tx}}{e^t - 1}$$

w szereg potęgowy względem zmiennej t , tzn. w rozwinięciu

$$f(t, x) = \frac{te^{tx}}{e^t - 1} = \sum_{k=0}^{\infty} B_k(x) \frac{t^k}{k!}. \quad (\text{C.3})$$

Wielomiany Bernoulliego można obliczyć z następującego wzoru

$$B_k(x) = \sum_{i=0}^k \binom{k}{i} B_i x^{k-i} \quad \text{dla } k = 0, 1, \dots, \quad (\text{C.4})$$

w którym B_i są liczbami Bernoulliego. Ze wzoru (C.4) otrzymujemy początkowe wielomiany Bernoulliego:

$$B_0(x) = 1,$$

$$B_1(x) = x - \frac{1}{2},$$

$$B_2(x) = x^2 - x + \frac{1}{6},$$

$$B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x,$$

$$B_4(x) = x^4 - 2x^3 + x^2 - x - \frac{1}{30}.$$

Niżej podajemy kilka własności liczb i wielomianów Bernoulliego.

$$B_k(0) = B_k \quad (k = 0, 1, \dots), \quad (\text{C.5})$$

$$B'_k(x) = k \cdot B_{k-1}(x) \quad (k = 1, 2, \dots), \quad (\text{C.6})$$

$$B_k(1-x) = (-1)^k \cdot B_k(x) \quad (k = 1, 2, \dots), \quad (\text{C.7})$$

$$B_k(mx) = m^{k-1} \sum_{p=0}^{m-1} B_k \cdot \left(x + \frac{p}{m} \right) \quad (k, m = 1, 2, \dots), \quad (\text{C.8})$$

$$B_k(x+1) - B_k(x) = k \cdot x^{k-1}, \quad (k = 1, 2, \dots) \quad (\text{C.9})$$

$$B_k(0) = B_k(1) = 0 \quad (k = 1, 2, \dots), \quad (\text{C.10})$$

$$B_k(x+1) = \sum_{i=0}^k \binom{k}{i} B_i(x) \quad (k = 1, 2, \dots). \quad (\text{C.11})$$

Powstanie liczb i wielomianów Bernoulliego związane jest z wyliczeniem sumy $1^k = 2^k + \dots + n^k$. Dziś wiadomo, że między sumami tej postaci a liczbami i wielomianami Bernoulliego istnieje związek

$$\sum_{p=1}^{n-1} p^k = \frac{B_{k+1}(n) - B_{k+1}}{k+1} \quad (\text{C.12})$$

D. Liczby Fibonacciego

Na początku XIII wieku Leonardo Fibonacci opublikował rozwiązańe zadania o szybkości rozmnażania się stada królików. Zadanie to można sformułować następująco.

Zadanie D.1. Na początku mamy parę nowo narodzonych królików. O każdej parze królików zakładamy, że:

- nowa para staje się płodna po miesiącu życia,
- każda płodna para rodzi jedną parę nowych królików w ciągu miesiąca,
- króliki nigdy nie umierają.

Pytanie brzmi: ile będzie par królików w k -tym miesiącu?

Rozwiązanie. Ze sposobu rozmnażania się królików wynika, że w kolejnym miesiącu liczba par królików jest równa liczbie par z poprzedniego miesiąca (króliki nie umierają) plus liczba par nowo narodzonych (a tych jest tyle, ile było par dwa miesiące wcześniej). Zatem liczba par królików w k -tym miesiącu jest równa:

$$\begin{aligned} F_1 &= 1, \\ F_2 &= 1, \\ F_k &= F_{k-1} + F_{k-2} \quad \text{dla } k > 2. \end{aligned} \quad (\text{D.1})$$

Liczby określone zależnością (D.1) noszą nazwę *liczb Fibonacciego*. Oto początkowe liczby Fibonacciego

$$\begin{array}{lllll} F_1 = 1, & F_2 = 1, & F_3 = 2, & F_4 = 3, & F_5 = 5, \\ F_6 = 8, & F_7 = 13, & F_8 = 21, & F_9 = 34, & F_{10} = 55, \\ F_{11} = 89, & F_{12} = 144, & F_{13} = 233, & F_{14} = 377, & F_{15} = 610, \\ F_{16} = 987, & F_{17} = 1597, & F_{18} = 2584, & F_{19} = 4181, & F_{20} = 6765. \end{array}$$

Liczby Fibonacciego można obliczyć ze wzoru

$$F_k = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^k - \left(\frac{1-\sqrt{5}}{2} \right)^k \right) \quad \text{dla } k = 1, 2, \dots . \quad (\text{D.2})$$

Ciekawostką jest fakt, że w *złotym podziale odcinka*, proporcja jest równa $(\sqrt{5}+1)/2$. Natomiast odwrotność tej proporcji jest równa $(\sqrt{5}-1)/2$.

E. Rozwinięcia funkcji elementarnych w szereg Maclaurina

Rozwinięcie funkcji w szereg	Przedział zbieżności	
$\frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k = 1 - x + x^2 - x^3 + \dots$	$ x < 1$	(E.1)
$\sqrt{x+1} = \sum_{k=0}^{\infty} \binom{1}{k} \cdot x^k = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \dots$	$ x \leq 1$	(E.2)
$\frac{1}{\sqrt{x+1}} = \sum_{k=0}^{\infty} \binom{-1}{k} \cdot x^k = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \dots$	$ x < 1$	(E.3)
$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \cdot x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$	$ x < \infty$	(E.4)
$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \cdot x^{2k} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$	$ x < \infty$	(E.5)
$\operatorname{tg} x = \sum_{k=1}^{\infty} \frac{2^{2k}(2^{2k}-1)B_k}{(2k)!} \cdot x^{2k-1} = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \dots$	$ x < \frac{\pi}{2}$	(E.6)
$x \cdot \operatorname{ctg} x = 1 - \sum_{k=1}^{\infty} \frac{2^{2k}B_k}{(2k)!} \cdot x^{2k} = 1 - \frac{1}{3}x^2 - \frac{1}{45}x^4 - \dots$	$ x < \pi$	(E.7)
$\arcsin x = \sum_{k=0}^{\infty} \binom{-\frac{1}{2}}{k} \cdot \frac{1}{2k+1} \cdot x^{2k+1} = x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \dots$	$ x < 1$	(E.8)
$\arccos x = \frac{\pi}{2} - \sum_{k=0}^{\infty} \binom{-\frac{1}{2}}{k} \cdot \frac{1}{2k+1} \cdot x^{2k+1} = \frac{\pi}{2} - \left(x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \dots \right)$	$ x < 1$	(E.9)
$\operatorname{arctg} x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \cdot x^{2k+1} = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \dots$	$ x < 1$	(E.10)
$\operatorname{arcctg} x = \frac{\pi}{2} + \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \cdot x^{2k+1} = \frac{\pi}{2} + \left(x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \dots \right)$	$ x < 1$	(E.11)

Rozwinięcie funkcji w szereg	Przedział zbieżności
$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots$	$ x < \infty$
$a^x = \sum_{k=0}^{\infty} \frac{(x \cdot \ln a)^k}{k!} = 1 + \frac{\ln a}{1!}x + \frac{\ln^2 a}{2!}x^2 + \frac{\ln^3 a}{3!}x^3 + \dots$	$ x < \infty$
$\ln(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k} = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \dots$	$-1 < x \leq 1$
$\sinh x = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} \cdot x^{2k+1} = x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \frac{1}{7!}x^7 + \dots$	$ x < \infty$
$\cosh x = \sum_{k=0}^{\infty} \frac{1}{(2k)!} \cdot x^{2k} = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \frac{1}{6!}x^6 + \dots$	$ x < \infty$
$\operatorname{tgh} x = \sum_{k=1}^{\infty} \frac{2^{2k}(2^{2k}-1)B_{2k}}{(2k)!} \cdot x^{2k-1} = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 - \frac{17}{315}x^7 + \dots$	$ x < \frac{\pi}{2}$
$\operatorname{ctgh} x = \frac{1}{x} + \sum_{k=1}^{\infty} \frac{2^{2k}B_{2k}}{(2k)!} \cdot x^{2k-1} = \frac{1}{x} + \frac{1}{3}x - \frac{1}{45}x^3 + \frac{2}{945}x^5 - \dots$	$ x < \pi$
$\operatorname{arctgh} x = \frac{1}{2} \ln \frac{1+x}{1-x} = \sum_{k=0}^{\infty} \frac{1}{2k+1} x^{2k+1} = x + \frac{1}{3}x^3 + \frac{1}{5}x^5 + \frac{1}{7}x^7 + \dots$	$ x < 1$

F. Rozwinięcia funkcji elementarnych w ułamki łańcuchowe

Rozwinięcie funkcji w ułamek łańcuchowy	Przedział zbieżności
$e^x = 1 + \frac{x}{1 - x + \frac{x}{2 - x + \frac{2x}{3 - x + \dots}}}$ \dots $\frac{nx}{n+1-x+\dots}$	$ x < \infty$ (F.1)
$\ln(x+1) = \frac{x}{1 + \frac{x}{2 + \frac{x}{3 + \frac{2x}{2 + \frac{2x}{5 + \dots}}}}}$ \dots $\frac{nx}{2 + \frac{nx}{2n+1+\dots}}$	$x > -1$ (F.2)
$\operatorname{tg} x = \frac{x}{1 - \frac{x^2}{3 - \frac{x^2}{5 - \dots}}}$ \dots $\frac{x^2}{-2n+1-\dots}$	(F.3)
$\operatorname{arctgx} = \frac{x}{1 + \frac{x^2}{3 + \frac{4x^2}{5 + \dots}}}$ \dots $\frac{n^2 x^2}{2n+1+\dots}$	$ x < \infty$ (F.4)

Rozwinięcie funkcji w ułamek łańcuchowy	Przedział zbieżności
$\operatorname{tgh} x = \frac{x}{1 + \frac{x^2}{3 + \frac{x^2}{5 + \dots}}}$ \dots $\frac{x^2}{2n+1+\dots}$	$ x < \infty$ (F.5)
$\operatorname{artgh} x = \frac{x}{1 - \frac{x^2}{3 - \frac{4x^2}{5 - \dots}}}$ \dots $\frac{n^2 x^2}{-2n+1-\dots}$	(F.6)

