Pracownia z ANALIZY NUMERYCZNEJ

Lista nr 1

Początek zapisów: 16 października 2007 r.

Termin realizacji: 12 listopada 2007 r.

Punktacja: maksymalnie 8 punktów za każde zadanie

Każde zadanie może być wybrane najwyżej czterokrotnie.

P1.1. Ciąg $x_k := 2^k \sin \frac{\pi}{2^k}$ (k = 1, 2, ...) jest zbieżny do π . Wykazać, że ciąg ten spełnia każdy z następujących trzech związków rekurencyjnych:

(1)
$$x_{k+1} = 2^k \sqrt{2\left(1 - \sqrt{1 - (x_k/2^k)^2}\right)} (k = 1, 2, \dots; x_1 = 2);$$

(3)
$$x_{k+1} = x_k \sqrt{\frac{2x_k}{x_k + x_{k-1}}} (k = 2, 3, ...; x_1 = 2, x_2 = 2\sqrt{2}).$$

Stosując oddzielnie wzory (1)–(3) obliczać – z pojedynczą, a następnie z podwójną precyzją – kolejne wyrazy ciągu $\{x_k\}$ do chwili, gdy dwa kolejne wyrazy mają 5 identycznych początkowych cyfr. Powtórzyć obliczenia żądając stabilizacji 8 cyfr. Wyciągnąć wnioski.

P1.2. Następujący algorytm sumowania z poprawkami pozwala obliczyć z dużą dokładnością sumę s $\sum_{i=1}^{n} x_i$, w standardowej arytmetyce f:

Dowodzi się, że $f(s) = \sum_{i=1}^{n} (1+\xi_i)x_i$, gdzie $|\xi_i| \leq 2 \cdot 2^{-t} + O(n2^{-2t})$. Oblicz sumę

$$\sum_{k=1}^{10000} k^{-2}$$

stosując

- (a) algorytm sumowania składników w naturalnej kolejności (i) z pojedynczą precyzją, następnie (ii) z podwójna precyzją, a także
- (b) algorytm sumowania z poprawkami, w arytmetyce z pojedynczą precyzją.

Porównaj wyniki. Podaj wnioski.

- **P1.3.** (a) Oblicz przybliżenia wartości pochodnej funkcji e^x w punkcie x=0 jako wartości ilorazu różnicowego $(e^{x+h}-e^x)/h$ dla malejących wartości h. Użyj (a) $h=2^{-i}$, $i=1,2,\ldots,20$; (b) $h=(2.2)^{-i}$, $i=1,2,\ldots,20$; wydrukuj wyniki z 8 cyframi dziesiętnymi po kropce. Skomentuj wyniki.
 - (b) Powtórz obliczenia, używając $(e^{x+h}-e^{x-h})/(2h)$ jako przybliżonego wyrażenia dla pochodnej funkcji e^x .
- P1.4. Wartość skończonego ułamka łańcuchowego

$$C_n := b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \ldots + \frac{a_n}{b_n}}}$$

można obliczyć metodą "schodami w górę", używając wzorów

$$U_n = b_n,$$

 $U_k = \frac{a_{k+1}}{U_{k+1}} + b_k \qquad (k = n - 1, n - 2, \dots, 0),$
 $C_n = U_0.$

Metoda "schodami w dół" jest następująca: obliczamy pomocnicze wielkości P_k i Q_k wg wzorów

$$P_{-1} = 1, \quad P_0 = b_0, \qquad P_k = b_k P_{k-1} + a_k P_{k-2} \qquad (k = 1, 2, \dots, n);$$

 $Q_{-1} = 0, \quad Q_0 = 1, \qquad Q_k = b_k Q_{k-1} + a_k Q_{k-2} \qquad (k = 1, 2, \dots, n).$

Wówczas $C_n = P_n/Q_n$.

- (a) Porównaj te metody pod względem kosztu realizacji i własności numerycznych.
- (b) Wykonaj obliczenia dla następujących danych:
 - i. $a_1 = 1$, $a_2 = 1$, $a_3 = 9$, $a_4 = 25$, $a_5 = 49$,..., $b_0 = 0$, $b_1 = 1$, $b_k = 2$ $(k \ge 2)$; sprawdź, czy dla dostatecznie dużego n jest $C_n \approx \pi/4$;
 - ii. $b_0=2,\ a_k=b_k=k+1\quad (k\geq 1);$ sprawdź, czy dla dostatecznie dużego n jest $C_n\approx e.$
- **P1.5.** Napisz podprogram obliczający wartość logarytmu naturalnego w
g następującej metody. Jeśli x=1, to sprawa jest oczywista. W przeciwnym wypadku należy wyznaczyć takie $n\in \mathbf{Z}$ i $r\in [\frac{1}{2},1)$, że $x=r\times 2^n$. Następnie połóż $u:=(r-\sqrt{2}/2)/(r+\sqrt{2}/2)$ i oblicz przybliżoną wartość l
n $\frac{1+u}{1-u}$ ze wzoru

$$\ln \frac{1+u}{1-u} \approx u \cdot \frac{20790 - 21545.27u^2 + 4223.9187u^4}{10395 - 14237.635u^2 + 4778.8377u^4 - 230.41913u^6}.$$

Wreszcie, przyjmij, że $\ln x \approx \left(n - \frac{1}{2}\right) \ln 2 + \ln \frac{1+u}{1-u}$. Porównaj wartości obliczone w ten sposób z podawanymi przez podprogram biblioteczny (funkcję standardową) dla np. 100 wartości argumentu. Jaki jest największy błąd względny? Skomentuj wyniki.

P1.6. Funkcja cosinus ma następujące rozwinięcie w szereg potęgowy, zbieżne dla każdej wartości x:

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}.$$

Przybliżoną wartość $\cos x$ można otrzymać jako wartość wielomianu

$$c_{2n}(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!}.$$

- (a) Wykonaj obliczenia w arytmetyce z pojedynczą precyzją dla $n=0,1,2,\ldots,12$ oraz dla wybranych wartości x z przedziału [0,10].
- (b) Sporządź wykresy wielomianów c_2, c_4, \ldots, c_{24} w tym przedziale.
- (c) Skomentuj wyniki.

- **P1.7.** Stałą Eulera $\gamma = 0.577215664901532286\ldots$ definujemy jako granicę $\gamma := \lim_{n \to \infty} \gamma_n$, gdzie $\gamma_n := 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} \ln n$. Zakładając, że dla dostatecznie dużych wartości n jest $\gamma_n \gamma \approx c n^{-d}$, gdzie c i d>0 są pewnymi stałymi, spróbuj przy pomocy komputera wyznaczyć doświadczalnie wartości tych stałych.
- **P1.8.** Wiadomo, że $\pi = 4 \lim_{n \to \infty} s_n$, gdzie $s_n := \sum_{k=0}^n (-1)^k (2k+1)^{-1}$. Oblicz w arytmetyce z podwyższoną precyzją wartości s_n dla $n = 10^7$, sumując składniki w porządku (a) naturalnym i (b) odwrotnym. Oblicz błędy $|4.0 \times fl(s_n) \pi|$.
- **P1.9.** Wartości funkcji $f(x) = (x-1)^8$ można obliczać na różne sposoby, np:
 - (a) $x^8 8x^7 + 28x^6 56x^5 + 70x^4 56x^3 + 28x^2 8x + 1$,
 - (b) (((((((x-8)x+28)x-56)x+70)x-56)x+28)x-8)x+1,
 - (c) (x-1)(x-1)(x-1)(x-1)(x-1)(x-1)(x-1),
 - (d) $\{[(x-1)^2]^2\}^2$,
 - (e) $\begin{cases} e^{8 \ln |x-1|} & : & (x \neq 1), \\ 0 & : & (x = 1). \end{cases}$

Porównaj podane wyżej metody obliczania f(x) dla $x \in [0.99, 1.01]$ (np. w N równoodległych punktach tego przedziału). Wyniki eksperymentu przedstaw na odpowiednich wykresach, prze-analizuj je i skomentuj.

- **P1.10.** Wykorzystując jedynie podstawowe działania arytmetyczne (+, -, *, /), zaproponuj efektywny sposób wyznaczania wartości funkcji sinus i cosinus z dokładnością bliską dokładności maszynowej. Opracowany algorytm porównaj z funkcjami bibliotecznymi.
- **P1.11.** Niech $\{s_n\}$ będzie ciągiem zbieżnym do granicy s. Ciąg Δ^2 Aitkena

$$t_n = \frac{s_n s_{n+2} - s_{n+1}^2}{s_{n+2} - 2s_{n+1} + s_n} \qquad (n = 0, 1, \ldots)$$

jest w wielu wypadkach zbieżny do s szybciej niż $\{s_n\}$, tzn.

$$\lim_{n \to \infty} \frac{t_n - s}{s_n - s} = 0.$$

(a) Oblicz 20 początkowych wyrazów ciągów $\{s_n\}$ i $\{t_n\}$ oraz $\{e_n:=s_n-s\}$ i $\{d_n:=t_n-s\}$ w wypadku

i.
$$s_n = \sum_{j=1}^n (-1)^j (2j+1)^{-1}, \ s = \pi/4 \approx 0.7853981634;$$

ii.
$$s_n = \sum_{k=1}^n k^{-3/2}, \ s \approx 2.612375348685488.$$

Czy mamy do czynienia z istotnym przyspieszeniem zbieżności? Powtórz doświadczenie dla innych danych.

(b) Zauważ, że zbieżność ciągu $\{t_n\}$ można przyspieszyć w analogiczny sposób, definiując ciąg $\{u_n\}$ wzorem

$$u_n = \frac{t_n t_{n+2} - t_{n+1}^2}{t_{n+2} - 2t_{n+1} + t_n} \qquad (n = 0, 1, \ldots).$$

Korzystając z tej obserwacji wykonaj kilka doświadczeń obliczeniowych, m. in. dla danych z punktu 1.11a.

(c) Uogólniając metodę, zaproponuj sposób przyspieszenia ciqgu $\{u_n\}!$ Sprawdź eksperymentalnie jego skuteczność.

- **P1.12.** Ciąg $\{y_n\}$ jest określony wzorem $y_n := \int_0^1 t^n e^t dt$ $(n = 0, 1, \ldots).$
 - (a) Sprawdź, że ciąg $\{y_n\}$ monotonicznie maleje do zera.
 - (b) Sprawdź, że zachodzi związek $y_{n+1} = e (n+1)y_n$ (n=0,1,...) i wyznaczyć wartość początkową y_0 . Korzystając z tego wyniku oblicz w standardowej arytmetyce wyrazy $y_0, y_1,..., y_N$ dla N=20. Czy otrzymane wyniki są wiarygodne?
 - (c) Oto inny sposób realizacji tego samego zadania. Zauważ, że wobec nierówności $\frac{1}{n+1} \leq y_n \leq \frac{e}{n+1}$ (sprawdzić!) ciąg jest wolno zbieżny, więc y_N i y_{N-1} są prawie sobie równe; z równania $y_N = e Ny_{N-1}$ wynika wówczas, że w przybliżeniu jest $y_N = e/(N+1)$. Następnie za pomocą podanego wcześniej związku rekurencyjnego oblicz $y_{N-1}, y_{N-2}, \ldots, y_0$. Czy wartość y_0 jest dokładna? A inne wyrazy ciągu? Podsumuj wyniki doświadczeń.
- P1.13. Funkcja Bessla ma następujące rozwinięcie w szereg potęgowy, zbieżne dla każdej wartości x:

$$J_0(x) = \sum_{k=0}^{\infty} (-1)^k \frac{(x/2)^{2k}}{(k!)^2}.$$

Załóżmy, że chcemy wykorzystać to rozwinięcie dla obliczenia przybliżonej wartości $J_0(x)$ z błędem $\leq \frac{1}{2}10^{-6}$. Dla wybranych, umiarkowanych oraz dużych wartości argumentu x, np. x=20, odpowiedz na następujące pytania:

- (a) Ilu składników szeregu trzeba użyć?
- (b) Jaki jest największy co do modułu użyty składnik?
- (c) Jakie zjawisko obserwujemy?
- **P1.14.** Ciąg funkcji Bessla J_n określamy wzorem

$$J_n(x) := \frac{1}{\pi} \int_0^{\pi} \cos(x \sin t - nt) dt$$
 $(n = 0, 1, ...).$

Łatwo zauważyć, że $|J_n(x)| \leq 1$. Wiadomo, że zachodzi związek

$$J_{n+1}(x) = \frac{2n}{x} J_n(x) - J_{n-1}(x)$$
 $(n = 1, 2, ...).$

(a) i. Wykorzystać ten związek oraz znane wartości $J_0(1)\approx 0.765197865,$ $J_1(1)\approx 0.4400505857$ do obliczenia w arytmetyce z pojedynczą precyzją wartości

$$J_0(1), J_1(1), \ldots, J_{20}(1).$$

Co można powiedzieć o wiarygodności wyników?

- ii. Rozważyć następujący algorytm.
 - \bullet Wybrać N>20 i określić pomocnicze wartości

$$\begin{split} c_{N+1}^{(N)} &:= 0; \qquad c_N^{(N)} := 1.0, \\ c_{k-1}^{(N)} &:= \frac{2k}{r} c_k^{(N)} - c_{k+1}^{(N)} \quad (k = N, N-1, \dots, 1). \end{split}$$

• Następnie obliczyć stałą $\lambda := J_0(x)/c_0^{(N)}$ oraz wielkości

$$j_k^{(N)} := \lambda c_k^{(N)}$$
 $(k = 0, 1, \dots, N).$

• Wówczas jest $j_k^{(N)} \approx J_k(x)$ dla $k = 0, 1, \dots, N$.

Wykonać obliczenia w arytmetyce z pojedynczą precyzją dla x=1 oraz dla N=25 i N=30. Przedyskutować wyniki.

(b) Powtórzyć obliczenia z punktu 1.14
a w arytmetykach z podwyższoną precyzją. Przedyskutować wyniki.

P1.15. Wiadomo, że suma szeregu

(4)
$$S_2 := \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2 + 1}$$

wynosi 0.36398547250893341852488170816398... Spróbuj wyznaczyć wartość tej liczby z dokładnością 10 i 16 cyfr za pomocą sum częściowych szeregu (4). Następnie zauważ, że

$$\frac{\pi^2}{12} - S_2 = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2(k^2+1)}, \qquad S_2 - \frac{\pi^2}{12} + \frac{7\pi^4}{720} = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^4(k^2+1)}$$

i wykorzystaj te związki, aby przyspieszyć obliczenia. Postępując podobnie, zaproponuj efektywny sposób wyznaczania wartości szeregu

$$S_n := \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^n + 1}$$
 $(n = 2, 4, 6, \ldots).$

P1.16. Dla danych x > 0 i $\varepsilon > 0$ przybliżoną wartość e^x można otrzymać jako

$$s_n := \sum_{k=0}^n \frac{x^k}{k!},$$

gdzie n jest najmniejszą liczbą naturalną, dla której zachodzi nierówność $|x^{n+1}/(n+1)!| < \varepsilon$. (Jeśli x < 0, można zastosować wzór $e^x = 1/e^{-x}$.)

Obliczenie e^x można sprowadzić do obliczenia e^u dla $|u|<\frac{1}{2}\ln 2$, co oznacza sumowanie <u>bardzo</u> szybko zbieżnego szeregu. Mianowicie, dla danego x wyznaczamy takie stałe m oraz u, że

$$e^x = 2^m e^u$$
.

Algorytm jest następujący: $z := x/\ln 2$; $m := \lfloor z + \sigma_{\frac{1}{2}} \rfloor$; w := z - m; $u := w \ln 2$, gdzie $\sigma := \operatorname{sgn} x$. Wykonać obliczenia sprawdzające powyższy algorytm dla kilku serii wartości x (małych, średnich i dużych), podając w każdym wypadku wartości n, błędu bezwzględnego i względnego.

P1.17. (a) Sprawdzić, że ciąg Fibonacciego

(5)
$$F_k = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^k - \left(\frac{1-\sqrt{5}}{2} \right)^k \right]$$

spełnia związek rekurencyjny

(6)
$$z_k = z_{k-1} + z_{k-2} (k = 2, 3, ...).$$

Sprawdzić, że $F_0=1$ i $F_1=1$ są początkowymi dwoma wyrazami ciągu.

(b) Sprawdzić, że ciąg

(7)
$$G_k = \left(\frac{1 - \sqrt{5}}{2}\right)^k.$$

również spełnia związek rekurencyjny (6) dla pewnych (jakich?) danych G_0 i G_1 .

- (c) Obliczyć w arytmetykach z pojedynczą i podwyższoną precyzją 50 początkowych wyrazów ciągów $\{F_k\}$ i $\{G_k\}$ następującymi dwoma sposobami, a następnie porównać wyniki otrzymane za pomocą:
 - i. związku rekurencyjnego (6),
 - ii. wzorów (5), (7).

Objaśnić wyniki.

P1.18. Napisz podprogram obliczający dwa pierwiastki x_1 i x_2 trójmianu kwadratowego $f(x) = ax^2 + bx + c$ o rzeczywistych współczynnikach a, b i c, jak również wartości $f(x_1)$ i $f(x_2)$. Użyj wzorów redukujących błędy zaokrągleń. Sprawdź działanie podprogramu m. in. dla

$$\begin{array}{lll} (a,\,b,\,c) & = & (0,1,0),\; (0,0,1),\; (0,0,0),\; (1,1,0),\; (2,10,1),\; (1,-4,3.99999),\\ & & (1,-8.01,16.004),\; (2\times10^{17},10^{18},10^{17}),\; (10^{-17},-10^{17},10^{17}). \end{array}$$

P1.19. Zaproponuj efektywny numerycznie program wyznaczający rozwiązania równania algebraicznego postaci

$$ax^3 + bx^2 + cx + d = 0.$$

gdzie a, b, c i d są liczbami rzeczywistymi.

P1.20. Opracować i sprawdzić na przykładach procedury funkcyjne, obliczające z dokładnością bliską dokładności maszynowej wartości następujących funkcji matematycznych:

$$g_1(x) := x + e^x - e^{3x}, \quad g_2(x) := \log x - 1, \quad g_3(x) := \sqrt{x^2 + 1} - 1.$$

W każdym wypadku zbadać, czy istnieje groźba utraty cyfr znaczących wyniku i – w razie potrzeby – zaproponować sposób unikniecia groźby.

 $\mathbf{P1.21.}$ Sprawdzić, że dla dowolnych stałych A i B ciąg

$$x_k = A(1+\sqrt{3})^k + B(1-\sqrt{3})^k$$

spełnia związek rekurencyjny

$$x_k = 2(x_{k-1} + x_{k-2})$$
 $(k = 3, 4, ...).$

Sprawdzić, że jeśli $x_1 = 1$ i $x_2 = 1 - \sqrt{3}$ są początkowymi dwoma wyrazami ciągu, to A = 0 i $B = (1 - \sqrt{3})^{-1}$. Obliczyć (w arytmetykach z pojedynczą i podwyższoną precyzją) 100 początkowych wyrazów ciągu następującymi trzema sposobami, a następnie porównać wyniki:

- (a) za pomocą związku rekurencyjnego,
- (b) za pomocą wzoru $x_k = B(1 \sqrt{3})^k$,
- (c) za pomocą wzoru $x_k = \tilde{A}(1+\sqrt{3})^k + B(1-\sqrt{3})^k$, gdzie $\tilde{A} := 2^{-t}$ (błąd reprezentacji liczb rzeczywistych w arytmetyce maszynowej).
- P1.22. Wykazać, że ciąg

(8)
$$I_n := \int_0^1 \frac{x^n}{x+10} dx \qquad (n=0,1,\ldots)$$

ma następujące własności:

- (a) $I_0 = \ln \frac{11}{10}$;
- (b) $I_n + 10I_{n-1} = \frac{1}{n}$ (n = 1, 2, ...);
- (c) $\frac{1}{11(n+1)} \le I_n \le \frac{1}{10(n+1)}$.

Korzystając ze związku rekurencyjnego (b) obliczyć – z pojedynczą, a następnie z podwójną precyzją – kolejno I_0, I_1, \ldots, I_{20} . Następnie powtórzyć obliczenia, przyjmując $I_{21} := 0$ (dlaczego jest to sensowne przybliżenie?) i – korzystając z (b) – otrzymać $I_{20}, I_{19}, \ldots, I_{0}$. Porównać wyniki i wyciągnąć wnioski.

P1.23. Zadanie dla dwuosobowego zespołu Wykorzystując jedynie podstawowe działania arytmetyczne (+,-,*,/), zaproponuj efektywny sposób wyznaczania wartości funkcji sinus w **dziedzinie liczb** zespolonych z dokładnością bliską dokładności maszynowej.

P1.24. Zadanie dla dwuosobowego zespołu Niech X i Y będą macierzami kwadratowymi stopnia n, gdzie n jest liczbą parzystą. Iloczyn macierzy

$$Z = X Y$$

definiujemy następująco:

(9)
$$z_{ij} = \sum_{k=1}^{n} x_{ik} y_{kj} \qquad (i = 1, 2, \dots, n; \ j = 1, 2, \dots, n),$$

gdzie z_{ij} jest elementem *i*-tego wiersza i *j*-tej kolumny macierzy Z (wielkości x_{ij} i y_{ij} mają znaczenie analogiczne). Aby wyznaczyć iloczyn macierzy korzystając ze wzoru (9) należy wykonać n^3 mnożeń. Macierze Z, X i Y możemy zapisać w tzw. postaci blokowej:

$$Z = \left[egin{array}{cc} R & S \\ T & U \end{array}
ight], \qquad X = \left[egin{array}{cc} A & B \\ C & D \end{array}
ight], \qquad Y = \left[egin{array}{cc} E & G \\ F & H \end{array}
ight].$$

Jeśli Z = XY, to

$$R = AE + BF$$
, $S = AG + BH$, $T = CE + DF$, $U = CG + DH$.

W tym wypadku musimy obliczyć 8 iloczynów macierzy stopnia n/2, czyli ponownie wykonać $8(n/2)^3=n^3$ mnożeń. Sprawdź jednak, że prawdziwe są równości:

$$R = P_5 + P_4 - P_2 + P_6$$
, $S = P_1 + P_2$, $T = P_3 + P_4$, $U = P_5 + P_1 - P_3 - P_7$

gdzie

$$P_1 = A(G - H),$$
 $P_2 = (A + B)H,$ $P_3 = (C + D)E,$ $P_4 = D(F - E),$ $P_5 = (A + D)(E + H),$ $P_6 = (B - D)(F + H),$ $P_7 = (A - C)(E + G).$

Stosując powyższą procedurę obliczamy tylko 7 iloczynów macierzy stopnia n/2. Wykonujemy zatem $7/8 \cdot n^3$ mnożeń. Jeśli $n=2^k$, to iloczyny macierzy stopnia n/2 obliczamy podobnie (jeśli n nie jest potęgą dwójki możemy rozszerzyć macierze uzupełniając je zerami do odpowiedniego rozmiaru). Powyższe postępowanie nosi nazwę algorytmu Strassena mnożenia macierzy.

1. Porównaj pod względem szybkości i dokładności tradycyjny algorytm mnożenia macierzy (wzór (9)) z algorytmem Strassena. 2. Obliczenia przeprowadź dla macierzy o rozmiarach od 4 do 500. Dla macierzy X o znanej macierzy odwrotnej X^{-1} (konsultacje) oblicz wartości błędów

$$\Delta(XX^{-1}-I), \qquad \Delta(X^{-1}X-I),$$

gdzie I jest macierzą jednostkową, natomiast $\Delta(X) := \sum_{i=1}^n \sum_{j=1}^n x_{ij}^2$. 3. Dla danych macierzy X, Y i V oblicz $\Delta((XY)V - X(YV))$. 4. Skomentuj wyniki.