http://www.ii.uni.wroc.pl/~sle/teaching/anp-1.pdf

Pracownia z ANALIZY NUMERYCZNEJ

Lista nr 1

Początek zapisów: **12 października 2009 r.**

Termin realizacji: **8 listopada 2009 r.**

Punktacja (podana przy każdym zadaniu): ***8, 10, albo 12 punktów***

Każde z zadań może być wybrane najwyżej przez cztery osoby (cztery zespoły dwuosobowe — w wypadku zadań P1.21 i P1.22) spośród wszystkich zapisanych na pracownię.

P1.1. 10 punktów Ciąg $x_k := 2^k \sin \frac{\pi}{2^k}$ $(k = 1, 2, \ldots)$ jest zbieżny do π. Wykazać, że ciąg ten spełnia każdy z następujących trzech związków rekurencyjnych:

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

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

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

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

```
s:=x[1];  c:=0;

  for i from 2 to n
    do
      y:=c+x[i];
      t:=s+y;
      c:=(s-t)+y;
      s:=t
    end
```

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

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

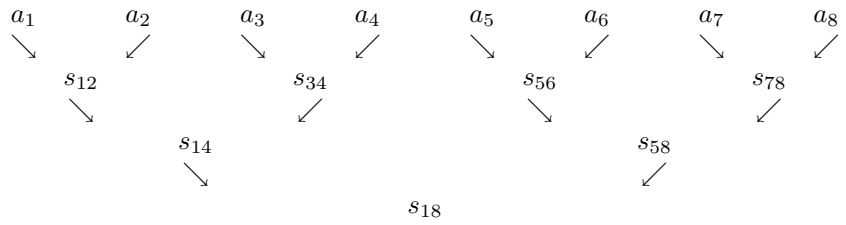
stosując

(a) algorytm sumowania składników w naturalnej kolejności (i) z pojedynczą precyzją, następnie (ii) z podwójną precyzją, a także

(b) algorytm sumowania z poprawkami, w arytmetyce z pojedynczą precyzją.

Porównaj wyniki. Podaj wnioski.

P1.3. 10 punktów Stosując strategię *dziel i zwyciężaj* wartość sumy $\sum_{k=1}^{8} a_k$ można wyznaczyć wykonując obliczenia zgodnie z następującym diagramem:



gdzie $s_{ij} := a_i + a_{i+1} + \ldots + a_j$. Zaproponuj podobny sposób wyznaczania wartości wyrażenia $\sum_{k=1}^{n} a_k$, gdzie $n := 2^m$ dla pewnego $m \in \mathbb{N}$ i porównaj go pod względem jakości numerycznej z tradycyjną metodą sumowania. Przeprowadź eksperymenty wykonując obliczenia z pojedynczą i podwójną precyzją dla odpowiednio dobranych elementów $a_1, a_2, \ldots, a_n$. Czy inne uporządkowanie elementów a_k coś zmienia? Co możemy zaobserwować dla dużych wartości n? Jak to wytłumaczyć?

P1.4. 8 punktów Niech $x := 1 + \pi/10^6$. Oblicz x^n dla $n = k \times 10^5$, $k = 1, 2, \ldots, 10$, stosując arytmetyki z pojedynczą i podwójną precyzją; oblicz błąd względny ρ_n wyniku obliczonego w arytmetyce z pojedynczą precyzją w odniesieniu do wyniku obliczonego w arytmetyce z podwójną precyzją. (Jaka powinna być, w przybliżeniu, wartość x^n dla $n = 10^6$?) Dla podanych wyżej wartości n wydrukuj: n, x^n, ρ_n, $\rho_n/(n \times \varepsilon)$, gdzie ε jest błędem reprezentacji maszynowej.

P1.5. 10 punktów Wartość skończonego *ułamka łańcuchowego*

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

można obliczyć metodą *„schodami w górę”*, używając wzorów

$$\begin{aligned} &U_n = b_n, \\ &U_k = \frac{a_{k+1}}{U_{k+1}} + b_k \qquad (k = n-1, n-2, \ldots, 0), \\ &C_n = U_0. \end{aligned}$$

Metoda *„schodami w dół”* jest następująca: obliczamy pomocnicze wielkości P_k i Q_k wg wzorów

$$\begin{aligned} &P_{-1} = 1, \quad P_0 = b_0, \qquad P_k = b_k P_{k-1} + a_k P_{k-2} \qquad (k = 1, 2, \ldots, n); \\ &Q_{-1} = 0, \quad Q_0 = 1, \qquad Q_k = b_k Q_{k-1} + a_k Q_{k-2} \qquad (k = 1, 2, \ldots, n). \end{aligned}$$

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. $a_1 = 1,\ a_2 = 1,\ a_3 = 9,\ a_4 = 25,\ a_5 = 49, \ldots,\ b_0 = 0,\ b_1 = 1,\ b_k = 2 \quad (k \geq 2)$; 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 $C_n \approx e$.

P1.6. 8 punktów

(a) Napisz program obliczający wartość sumy

$$S_N := \sum_{n=1}^{N} [1/n - 1/(n+1)]$$

bez żadnych przekształceń wstępnych! Jaka jest dokładna wartość sumy S_N? Wykonaj obliczenia dla $N = 10, 100, 1000, \ldots, 1000000$ i wydrukuj $|fl(S_N) - S_N|$. Omów wyniki.

(b) Powtórz czynności części (a) w stosunku do iloczynu $p_N := \prod_{n=1}^{N} [n/(n+1)]$.

P1.7. 8 punktów Obliczaj wyrazy ciągów

$$\textbf{(a)} \quad s_n := \sum_{k=0}^{n} (-1)^k k!^{-2}, \qquad \textbf{(b)} \quad t_n := \sum_{k=0}^{n} k!^{-2}$$

do chwili, gdy dwa kolejne wyrazy są równe w wybranej arytmetyce maszynowej. (Wybierz odpowiedni sposób generowania składników sum!)

P1.8. 10 punktów 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.

P1.9. 8 punktów 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) $(x-1)(x-1)(x-1)(x-1)(x-1)(x-1)(x-1)(x-1)$,

(d) $\left\{\left[(x-1)^2\right]^2\right\}^2$,

(e) $\begin{cases} e^{8\ln|x-1|} & : \quad (x \neq 1), \\ 0 & : \quad (x = 1). \end{cases}$

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, przeanalizuj je i skomentuj.

P1.10. 10 punktów Wykorzystując jedynie podstawowe działania arytmetyczne $(+, -, *, /)$, zaproponuj efektywny sposób wyznaczania wartości funkcji $f(x) = \arctan x$ z dokładnością bliską dokładności maszynowej. Opracowany algorytm porównaj z funkcją biblioteczną.

P1.11. 10 punktów 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.

P1.12. 10 punktów 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. $s_n = \sum_{j=0}^{n} (-1)^j (2j+1)^{-1}$, $s = \pi/4 \approx 0.7853981634$;

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 1.12a.

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

P1.13. 10 punktów Ciąg $\{y_n\}$ jest określony wzorem $y_n := \int_0^1 t^n e^t \, dt \qquad (n = 0, 1, \ldots)$.

(a) Sprawdź, że ciąg $\{y_n\}$ monotonicznie maleje do zera.

(b) Sprawdź, że zachodzi związek $y_{n+1} = e - (n+1)y_n$ $(n = 0, 1, \ldots)$ i wyznaczyć wartość początkową y_0. Korzystając z tego wyniku oblicz w standardowej arytmetyce wyrazy $y_0, y_1, \ldots, y_N$ dla $N = 20$. Czy otrzymane wyniki są wiarygodne?

(c) Oto inny sposób realizacji tego samego zadania. Zauważ, że wobec nierówności $\frac{1}{n+1} \le y_n \le \frac{e}{n+1}$ (sprawdzić!) ciąg jest wolno zbieżny, więc y_N i y_{N-1} są prawie sobie równe; z równania $y_N = e - N y_{N-1}$ 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.

P1.14. 10 punktów *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 $\le \frac{1}{2}10^{-6}$. 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?

P1.15. 10 punktów Ciąg *funkcji Bessla* J_n określamy wzorem

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

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

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

(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.

- Wybrać $N > 20$ i określić pomocnicze wartości

$$c_{N+1}^{(N)} := 0; \qquad c_N^{(N)} := 1.0,$$
$$c_{k-1}^{(N)} := \frac{2k}{x} c_k^{(N)} - c_{k+1}^{(N)} \quad (k = N, N-1, \ldots, 1).$$

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

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

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

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

(b) Powtórzyć obliczenia z punktu 1.15a w arytmetykach z podwyższoną precyzją. Przedyskutować wyniki.

P1.16. 12 punktów Wiadomo, że suma szeregu

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

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

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

P1.17. 8 punktów Niech będzie $\Delta u_k = u_{k+1} - u_k$. Sprawdź, że

$$\sum_{k=1}^{n} \Delta u_k = u_{n+1} - u_1.$$

Oblicz wartość po lewej stronie powyższej równości z pojedynczą precyzją (**bez żadnych przekształceń wstępnych!**), a wartość po prawej stronie – z podwójną precyzją dla $u_k = \ln(1+k)$ oraz $n = 1000(1000)10000$. Wydrukuj tabelkę różnic między tymi wartościami. Oblicz $\sum_{k=1}^{n} u_k$ z pojedynczą i podwójną precyzją. Objaśnij wyniki.

P1.18. 12 punktów 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.

P1.19. 8 punktów 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 uniknięcia groźby.

P1.20. 10 punktów Sprawdzić, że dla dowolnych stałych A i B ciąg

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

spełnia związek rekurencyjny

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

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).

P1.21. Zadanie dla dwuosobowego zespołu. 12 punktów Wykorzystując jedynie działania arytmetyczne $+, -, *, /$, zaproponuj efektywny sposób wyznaczania wartości funkcji sinus w **dziedzinie liczb zespolonych** z dokładnością bliską dokładności maszynowej.

P1.22. Zadanie dla dwuosobowego zespołu. 12 punktów Niech X i Y będą macierzami kwadratowymi stopnia n, gdzie n jest liczbą parzystą. Iloczyn macierzy

$$Z = X\,Y$$

definiujemy następująco:

$$z_{ij} = \sum_{k=1}^{n} x_{ik}\,y_{kj} \qquad (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n), \tag{5}$$

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 (5) należy wykonać n^3 mnożeń. Macierze Z, X i Y możemy zapisać w tzw. *postaci blokowej*:

$$Z = \begin{bmatrix} R & S \\ T & U \end{bmatrix}, \qquad X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \qquad Y = \begin{bmatrix} E & G \\ F & H \end{bmatrix}.$$

Jeśli $Z = X\,Y$, to

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

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:

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

gdzie

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

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 (5)) 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.