Problemy i algorytmy analizy numerycznej

Stanisław Lewanowicz

4 października 2007 r.

Część I: Aproksymacja funkcji

Definicja 1. Szereg (nieskończony) $\sum_{k=1}^{\infty} c_k$ jest ***zbieżny*** i ma ***sumę*** s, jeśli ciąg sum częściowych

$$s_n := \sum_{k=1}^{n} c_k \qquad (n = 1, 2, \ldots)$$

jest zbieżny do granicy s.

Lemat 1. Jeśli $a_k \geq 0$ $(k \geq 0)$, $a_1 \geq a_2 \geq \ldots \geq a_k \geq \ldots$ i $\lim_{k\to\infty} a_k = 0$, to szereg przemienny

$$\sum_{k=1}^{\infty} (-1)^{k-1} a_k$$

jest zbieżny. Ponadto, jeśli s oznacza sumę szeregu, a s_n – n-tą sumę częściową

$$s_n := \sum_{k=1}^{n} (-1)^{k-1} a_k \qquad (n \geq 1),$$

to dla każdego n zachodzi nierówność

$$|s - s_n| \leq a_{n+1}.$$

Funkcja błędu

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

nie da się wyrazić prościej, przy użyciu bardziej elementarnych funkcji.

Zadanie 1: Podać metodę obliczenia $\operatorname{erf}(x)$ dla dowolnego $x \in [0, 1]$, z błędem mniejszym niż 10^{-12}.

Rozwiązanie: Korzystając z rozwinięcia $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$, dostajemy

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{k!}\, dt = \frac{2}{\sqrt{\pi}} \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{k!(2k+1)}.$$

Mamy tu szereg przemienny. Niech p_n będzie wielomianem

$$p_n(x) := \frac{2}{\sqrt{\pi}} \sum_{k=0}^{n-1} (-1)^k a_k, \quad \text{gdzie} \quad a_k = \frac{x^{2k+1}}{k!(2k+1)}.$$

Na mocy lematu 1 mamy

$$|\operatorname{erf}(x) - p_n(x)| < \frac{2\, a_n}{\sqrt{\pi}} = \frac{2x^{2n+1}}{\sqrt{\pi}\, n!(2n+1)} \leq \frac{2}{\sqrt{\pi}\, n!(2n+1)} < 10^{-12} \ \text{dla } n \geq 14.$$

Konkluzja: Dla każdego $x \in [0, 1]$ wartość wielomianu $p_{14}(x)$ przybliża $\operatorname{erf}(x)$ z błędem mniejszym niż 10^{-12}.

Część II: Całkowanie numeryczne

Zadanie 2: Obliczyć wartość całki oznaczonej

$$I := \int_a^b f(x)\,dx.$$

Sprawa jest prosta, jeśli znamy postać analityczną funkcji pierwotnej F funkcji podcałkowej, czyli takiej, że $F' = f$:

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Wiadomo, że całki wielu funkcji elementarnych, np. postaci

$$\int \frac{\sin x}{x}\,dx, \quad \int \frac{\cos x}{x}\,dx, \quad \int e^{-x^2}\,dx, \quad \int \sin x^2\,dx, \quad \int \frac{1}{\ln x}\,dx$$

nie dadzą się wyrazić *w sposób analityczny*, tj. przez funkcje elementarne za pomocą skończonej liczby działań arytmetycznych.

Często udaje się obliczyć dobre *przybliżenie* całki

$$\int_a^b f(x)\,dx$$

za pomocą sum, zwanych ***kwadraturami***, postaci

$$\sum_{k=0}^{n} A_k f(x_k),$$

gdzie $n, A_0, \ldots, A_n, x_0, \ldots, x_n$ są pewnymi odpowiednio dobranymi **stałymi**. Oto ***przykłady kwadratur***.

1. Zastąpmy pod znakiem całki funkcję $f(x)$ funkcją liniową

$$w(x) = \frac{1}{b-a}\left[(b-x)\,f(a) + (x-a)\,f(b)\right].$$

Zauważmy, $w(a) = f(a)$ oraz $w(ba) = f(b)$. Mamy

$$\int_a^b f(x)\,dx \approx \int_a^b w(x)\,dx = \frac{b-a}{2}\left[f(a) + f(b)\right].$$

Otrzymaliśmy kwadraturę zwaną ***wzorem trapezów***. Np.

$$\int_0^{0.8} \frac{\sin x}{x}\,dx \approx \frac{0.8-0}{2}\left[\frac{\sin 0.8}{0.8} + 1\right] = 0.758.$$

Ile jest wart ten wynik? Cierpliwości...

2. Dla $n > 0$ określamy $h := (b - a)/n$. Zapisujemy całkę w postaci

$$\int_a^b f(x)\,dx = \int_a^{a+h} f(x)\,dx + \int_{a+h}^{a+2h} f(x)\,dx + \ldots + \int_{b-h}^{b} f(x)\,dx,$$

a następnie do każdej z całek po prawej stronie wzoru stosujemy wzór trapezów. Dostajemy

$$\int_a^b f(x)\,dx \approx \frac{h}{2}\left[f(a) + f(a+h)\right] + \frac{h}{2}\left[f(a+h) + f(a+2h)\right] + \ldots$$

$$\ldots + \frac{h}{2}\left[f(b-h) + f(b)\right] =: T_n,$$

gdzie

$$T_n := h\left[\frac{f(a)}{2} + f(a+h) + f(a+2h) + \ldots + f(a+(n-1)h) + \frac{f(b)}{2}\right]$$

jest kwadraturą, którą nazywamy ***złożonym wzorem trapezów***.

Lemat 2. Dla dowolnej funkcji f ciągłej w przedziale $[a, b]$ jest

$$\lim_{n\to\infty} T_n = \int_a^b f(x)\, dx.$$

Przykład. Weźmy całkę

$$\int_0^{0.8} \frac{\sin x}{x}\, dx = 0.7720957855.$$

Oto jakie wyniki uzyskano:

n	T_n	n	T_n
1	0.7586780454	11	0.7719855932
2	0.7687573650	12	0.7720031942
3	0.7706133484	13	0.7720168916
4	0.7712621711	14	0.7720277600
5	0.7715623482	15	0.7720365280
6	0.7717253715	16	0.7720437039
7	0.7718236573	17	0.7720496510
8	0.7718874436	18	0.7720546348
9	0.7719311731	19	0.7720588525
10	0.7719624514	20	0.7720624536

wart. dokł. = **0.7720957855**

Jak widać ciąg T_n jest ***bardzo wolno zbieżny***.

Oto sposób przyspieszenia (niemal) za darmo:

Można wykazać, że ciąg

$$T'_n := \frac{4T_n - T_{n/2}}{3} \qquad (n = 2, 4, \ldots)$$

jest znacznie ***szybciej zbieżny***! Np.

$$T'_{10} = \frac{4T_{10} - T_5}{3} = \mathbf{0.7720958192.}$$

Porównajmy ***błąd tego przybliżenia***, czyli

$$|I - T'_{10}| = 3.4 \cdot 10^{-8},$$

z błędem

$$|I - T_{10}| = 1.3 \cdot 10^{-4}!$$

n	T_n	T'_n
1	0.7586780454	
2	0.7687573650	0.7721171382
3	0.7706133484	
4	0.7712621711	0.7720971069
5	0.7715623482	
6	0.7717253715	0.7720960448
7	0.7718236573	
8	0.7718874436	0.7720958676
9	0.7719311731	
10	0.7719624514	0.7720958196
12	0.7720031942	0.7720958025
14	0.7720277600	0.7720957945
16	0.7720437039	0.7720957904
18	0.7720546348	0.7720957879
20	0.7720624536	0.7720957879

wart. dokł. $= \mathbf{0.7720957855}$

Część III: Związki rekurencyjne

Zadanie: Obliczyć 20 początkowych wyrazów ciągu $\{y_k\}$, określonego wzorem

$$y_k := \int_0^1 x^k e^x \, dx \qquad (k = 0, 1, \ldots).$$

Własności ciągu:

- $\{y_k\}$ monotonicznie maleje do zera;
- $\frac{1}{k+1} \le y_k \le \frac{e}{k+1}$, tak więc zbieżność jest bardzo wolna;
- ciąg $\{y_k\}$ spełnia związek rekurencyjny

(1) $$y_{k+1} = e - (k+1)y_k \qquad (k = 0, 1, \ldots).$$

Oczywiście, $y_0 = e - 1$, skąd $y_1 = e - y_0 = 1$.

Algorytm I: Stosując (1) dla wartości k rosnących od 1 do 19 obliczymy potrzebne wyrazy ciągu:

$$(2) \qquad y_{k+1} = e - (k+1)y_k \qquad (k = 1, 2, \ldots 19).$$

Ocena wyników. Druga kolumna tabeli zawiera wyniki obliczeń na IBM PC w arytmetyce Single. Wyraz y_{11} jest większy od y_{10} (!); następne wyrazy są zmiennych znaków, a ich **wartości bezwzględne rosną bardzo szybko**!!!

- Błąd wielkości y_2 równy $\delta = 10^{-7}$ jest mnożony przez 3, gdy obliczamy y_3.
- Błąd wielkości y_3 równy 3δ jest mnożony przez 4, gdy obliczamy y_4.
- Wynikający stąd błąd 12δ jest mnożony przez 5, gdy obliczamy y_5 itd.
- Dlatego przy obliczaniu y_{10} błąd może być rzędu $\frac{1}{2}10!\delta \approx 2 \cdot 10^6 \cdot \delta \approx 2 \cdot 10^{-1}$. Dla y_{20} analogiczna wielkość jest już rzędu $10^{18}\delta = 10^{11}$.

Mamy tu przykład nieprzyjemnej wady algorytmu obliczeniowego, zwanej ***niestabilnością***.

	Algorytm I	**wart. dokładne**
1	1.0000000	1.00000000000
2	0.7182817	0.71828182846
3	0.5634365	0.56343634308
4	0.4645357	0.46453645613
5	0.3956032	0.39559954780
6	0.3446627	0.34468454165
7	0.3056431	0.30549003693
8	0.2731371	0.27436153302
9	0.2600479	0.24902803130
10	0.1178026	0.22800151549
11	1.4224529	0.21026515811
12	$-1.4351153 \cdot 10^{+1}$	0.19509993116
13	$1.8928328 \cdot 10^{+2}$	0.18198272334
14	$-2.6472476 \cdot 10^{+3}$	0.17052370130
15	$3.9711434 \cdot 10^{+4}$	0.16042630893
16	$-6.3538025 \cdot 10^{+5}$	0.15146088554
17	$1.0801467 \cdot 10^{+7}$	0.14344677430
18	$-1.9442640 \cdot 10^{+8}$	0.13623989100
19	$3.6941015 \cdot 10^{+9}$	0.12972389989
20	$-7.3882026 \cdot 10^{+10}$	0.12380383076

Algorytm II: Stosujemy równanie (1) dla k **malejącego** od 20 do 1:

$$y_k = (e - y_{k+1}/(k+1) \qquad (k = 20, 19, \ldots, 1). \tag{3}$$

Niezbędną do rozpoczęcia obliczeń wartość y_{21} obliczymy z równania

$$y_{21} = e - 21y_{20},$$

próbując przyjąć, że $y_{21} \approx y_{20}$ (założenie naturalne wobec poprzedniej uwagi, że $\{y_k\}$ maleje wolno). Stąd $y_{21} = \frac{e}{22}$.

Ocena wyników. Wyniki obliczeń zamieszczono w trzeciej kolumnie następnej tabeli. Poza dwiema ostatnimi liczbami wszystkie pozostałe wyniki są poprawnie zaokrąglonymi wynikami dokładnymi!

	Algorytm I	Algorytm II	wart. dokładne
1	1.0000000	1.0000000	1.00000000000
2	0.7182817	0.7182818	0.71828182846
3	0.5634365	0.5634363	0.56343634308
4	0.4645357	0.4645364	0.46453645613
5	0.3956032	0.3955995	0.39559954780
6	0.3446627	0.3446845	0.34468454165
7	0.3056431	0.3054900	0.30549003693
8	0.2731371	0.2743615	0.27436153302
9	0.2600479	0.2490280	0.24902803130
10	0.1178026	0.2280015	0.22800151549
11	1.4224529	0.2102652	0.21026515811
12	$-1.4351153 \cdot 10^{+1}$	0.1950999	0.19509993116
13	$1.8928328 \cdot 10^{+2}$	0.1819827	0.18198272334
14	$-2.6472476 \cdot 10^{+3}$	0.1705237	0.17052370130
15	$3.9711434 \cdot 10^{+4}$	0.1604263	0.16042630893
16	$-6.3538025 \cdot 10^{+5}$	0.1514609	0.15146088554
17	$1.0801467 \cdot 10^{+7}$	0.1434468	0.14344677430
18	$-1.9442640 \cdot 10^{+8}$	0.1362393	0.13623989100
19	$3.6941015 \cdot 10^{+9}$	0.1297362	0.12972389989
20	$-7.3882026 \cdot 10^{+10}$	0.1235583	0.12380383076

Część IV: Rozwiązywanie układów równań liniowych

Układ równań liniowych

$$
(4) \qquad \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \frac{137}{60} \\ \frac{29}{20} \\ \frac{153}{140} \\ \frac{743}{840} \\ \frac{1879}{2520} \end{bmatrix}
$$

ma dokładne rozwiązanie

$$x_1 = x_2 = \ldots = x_5 = 1.$$

Macierz układu (4) to ciesząca się złą sławą ***macierz Hilberta***. Proszę popatrzeć, jaki ma trudny charakter.

Rozważmy układ powstający z układu (4) przez ***dodanie małej liczby δ do ostatniej składowej wektora prawych stron***:

$$(5)\qquad \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \\ x_5' \end{bmatrix} = \begin{bmatrix} \frac{137}{60} \\ \frac{29}{20} \\ \frac{153}{140} \\ \frac{743}{840} \\ \frac{1879}{2520} + \delta \end{bmatrix}$$

Dla $\delta = 10^{-8}$ układ (5) ma rozwiązanie

$$x_1' = \frac{10000063}{10000000} = 1.0000063, \quad x_2' = \frac{499937}{500000} = 0.9998740,$$
$$x_3' = \frac{1000567}{1000000} = 1.0005670, \; x_4' = \frac{499559}{500000} = 0.9991180,$$
$$x_5' = \frac{1000441}{1000000} = 1.0004410.$$

Natomiast dla $\delta = 10^{-5}$ układ (5) ma rozwiązanie

$$x_1' = \frac{10063}{10000} = 1.0063, \quad x_2' = \frac{437}{500} = 0.8740, \quad x_3' = \frac{1567}{1000} = 1.5670,$$

$$x_4' = \frac{59}{500} = 0.1180, \quad x_5' = \frac{1441}{1000} = 1.4410.$$

Mamy tu przykład **zadania źle uwarunkowanego:**

małe względne zaburzenia danych

mogą spowodować **bardzo duże odkształcenie wyniku**.

- Czy można rozpoznać zadanie źle uwarunkowane PRZED wykonaniem obliczeń? MOŻNA!
- **Jak postępować** z zadaniami źle uwarunkowanymi? UNIKAĆ ich, a jeśli to niemożliwe, traktować bardzo ostrożnie.