2月20日作业

1

1. In the lecture we briefly discussed the convergence of the Gregory–Leibniz series

$$\frac{\pi}{4} = \arctan 1 = \cancel{\sum_{k=1}^{\infty} \frac{(-1)^k}{2k-1}}.$$

应该是: $\arctan x = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{2n-1} x^{2n-1}$

If we compute $\pi/4$ through the partial sum

$$S_{2n} = \sum_{k=1}^{2n} \frac{(-1)^{k-1}}{2k-1},$$

the convergence is slow. Estimate the truncation error, and propose a simple correction to S_{2n} based on your error estimate.

(optional) Estimate the truncation error again for the corrected scheme.

解: 记 $r_{2n} = \sum_{k=1}^{\infty} \frac{(-1)^k}{2k-1} - \sum_{k=1}^{2n} \frac{(-1)^k}{2k-1}$

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

注意到 $\left\{\frac{1}{2k-1}\right\}$ 单调递减且 $\lim_{k\to+\infty} \frac{1}{2k-1} = 0$

则 $\sum_{k=1}^{\infty} \frac{(-1)^k}{2k-1}$ 是 Leibniz 级数, 收敛

注意到 $\{S_{2n}\}$ 是关于 n 单调递增的数列:

$$S_{2(n+1)} = S_{2n} + \left(\frac{(-1)^{2n}}{2(2n+1)-1} + \frac{(-1)^{2n+1}}{2(2n+2)-1}\right)$$

$$= S_{2n} + \left(\frac{1}{4n+1} - \frac{1}{4n+3}\right) > S_{2n}$$

那么: $\left(\frac{1}{4n+1} - \frac{1}{4n+3}\right) < r_{2n}$

另一方面，由于：

$$r_{2n} = \frac{1}{4n+1} - \left(\frac{1}{4n+3} - \frac{1}{4n+5}\right) - \left(\frac{1}{4n+5} - \frac{1}{4n+7}\right) \cdots$$

$$\Rightarrow \quad r_{2n} < \frac{1}{4n+1}$$

故截断误差的范围为：

$$\frac{1}{4n+1} - \frac{1}{4n+3} < r_{2n} < \frac{1}{4n+1}，\text{或有 } |r_{2n}| < \frac{1}{4n+1}$$

对于 S_{2n}，可以通过 $\hat{S} = S_{2n} + \frac{1}{4n+1}$

进行修正。$\square$

(Optional) 正确的计算方法是：

$$\arctan\frac{1}{2} + \arctan\frac{1}{3} = \frac{\pi}{4}$$

同样计算：$S_{2n} = \sum_{k=1}^{2n} \frac{(-1)^{k-1}}{2k-1} \cdot \left(\left(\frac{1}{2}\right)^{2k-1} + \left(\frac{1}{3}\right)^{2k-1}\right)$

同样定义 $r_{2n}' = S_\infty - S_{2n}$，则类似的有：

$$|r_{2n}'| < \frac{1}{2n+1} \cdot \left(\left(\frac{1}{2}\right)^{4n+1} + \left(\frac{1}{3}\right)^{4n+1}\right) \quad \square$$

2

2. In principle, the sine function can be evaluated through the Taylor series expansion

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \qquad (x \in (-\infty, +\infty)).$$

Let us consider two computational schemes to evaluate the sine function.
(a) Directly truncate the Taylor series. Make sure that the truncation error is less than the rounding error bound for any input x.
(b) First shift x to the interval $(-\pi/2, \pi/2]$, and then apply scheme (a).
Scheme (b) is in general more accurate. Can you provide a theoretical analysis?

Sample at least 1000 points in $[-10, 10]$ (e.g., using the MATLAB/Octave statement `linspace(-10, 10, 1000)`) and plot the error of scheme (a) relative to scheme (b) (e.g., using the MATLAB/Octave function `semilogy`). Can you explain the result?

解：首先对(a)方法的截断误差进行估计

设以第N项开始满Leibniz级数的条件，即：

$$\frac{x^{2N+3}}{(2N+3)!} \Big/ \frac{x^{2N+1}}{(2N+1)!} < 1$$

$$\Rightarrow \quad x^2 < (2N+2)(2N+3)$$

则只要取 $N = \lceil \frac{x}{2} \rceil$ 即有

$$|r_n| < \frac{x^{2N+3}}{(2N+3)!} \quad \text{其中 } r_n \text{ 是截断误差}$$

另一方面，对于舍入误差，我的能力有限，只能估个大概：

对于第k项而言，不妨认为分母(2k+1)!用长整型存储，不会有舍入误差

那么：

$$\left[fl\left(\frac{x^{2k+1}}{(2k+1)!}\right)-\frac{x^{2k+1}}{(2k+1)!}\right]\le\frac{(x\cdot(1+\theta_{2k}))^{2k+1}}{(2k+1)!}$$

或者不严格一点：

$$fl\left(\frac{x^{2k+1}}{(2k+1)!}\right)\approx\frac{[x\cdot(1+\theta_{2k})]^{2k+1}}{(2k+1)!}$$

对于前 N 项和而言：

$$fl\left(\sum_{n=0}^{N}(-1)^n\frac{x^{2n+1}}{(2n+1)!}\right)$$

$$=fl\left(\sum_{n=0}^{N}(-1)^n\frac{(x(1+\theta_{2n}))^{2n+1}}{(2n+1)!}\right)$$

$$=\sum_{n=0}^{N}(-1)^n\frac{x(1+\theta_{2n+n})^{2n+1}}{(2n+1)!}$$

其中，对于 θ_{2n+N} 而言，$2n$ 来源于 x^{2n+1} 连乘，n 来源于 $(N+1)$ 个通项相加

于是可以得到：

$$|fl(S_N)-S_N|<\sum_{n=0}^{N}(-1)^n\frac{x^{2n+1}\cdot\theta_{3n(2n+1)}}{(2n+1)!}$$

$$< \Theta_{6N^2+3N} \sum_{n=0}^{N} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

$$\leq \Theta_{6N^2+3N} \cdot \max\{x, x^{2n+1}\} \cdot \sum_{n=0}^{N} (-1)^n \frac{1}{(2n+1)!}$$

$$< \Theta_{6N^2+3N} \cdot \max\{1, x^{2N+1}\}.$$

现在要求 $r_n < \Theta_{6N^2+3N} \cdot \max\{1, x^{2N+1}\}$

只需: $$\frac{x^{2N+3}}{(2N+3)!} < 6N^2 u \cdot \max\{1, x^{2N+1}\}$$

当 $|x| \geq 1$ 时, 要求

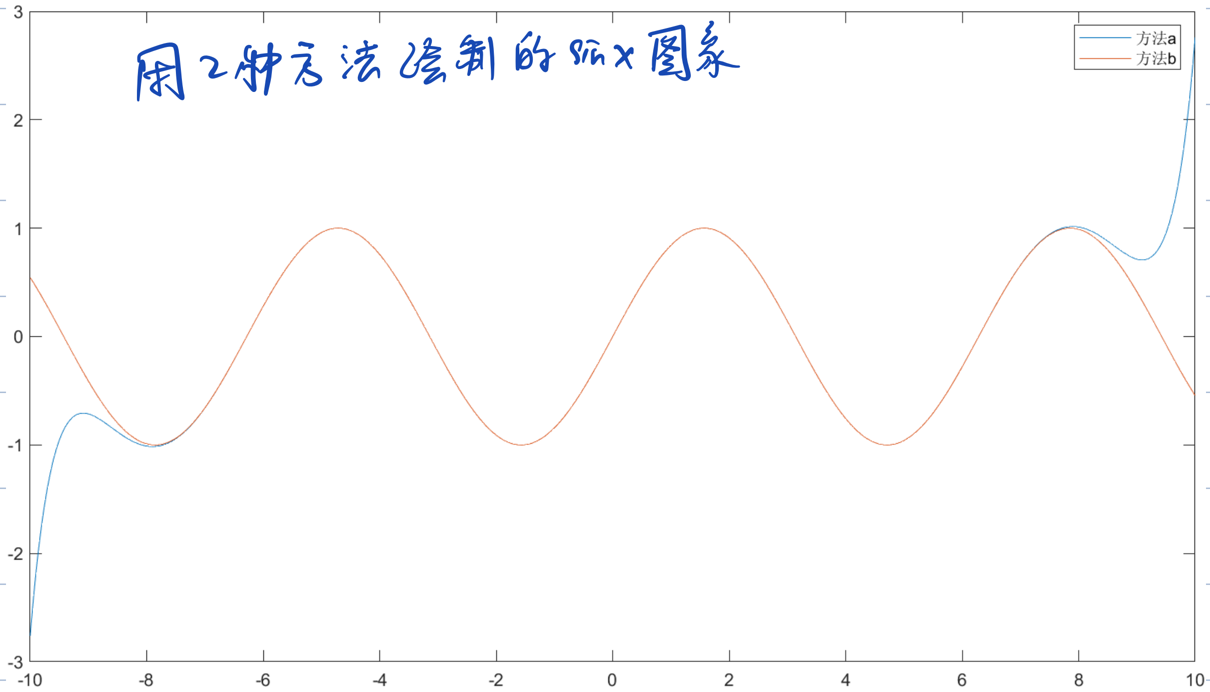
$$\frac{x^2}{(2N+3)!} < 6N^2 u$$

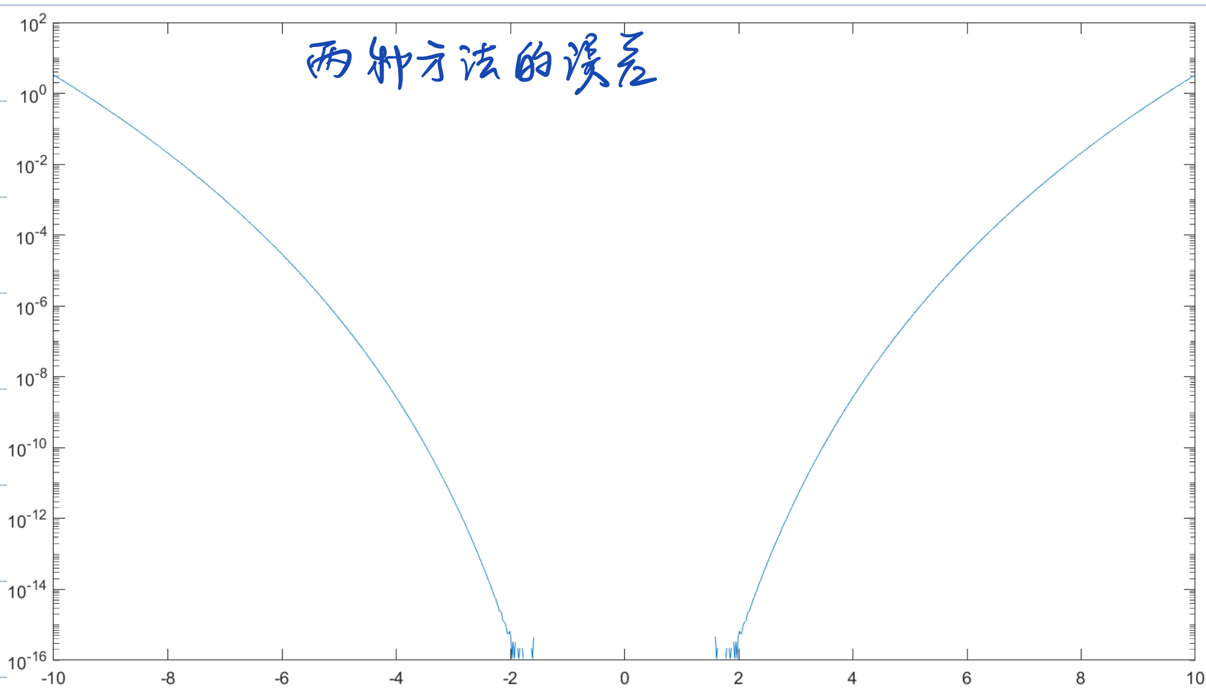
当 $|x| < 1$ 时要求:

$$\frac{1}{(2N+3)!} < 6N^2 u$$

对于 $x \in [-10, 10]$ 而言, 取 $N = 10$ 足够

有关图象如下:





首先说明为什么(b)方法比(a)更准确：

在N≥10时，由上面的讨论可知，当|x|<10，n>10时

$\left\{(-1)^n \frac{x^{2n+1}}{(2n+1)!}\right\}$ 是交错级数，则对于(a)有：

$$\left|r_N^{(a)}\right| < \frac{x^{21}}{(21)!}$$

对于(b)而言，注意到 $|\hat{x}| = |x - k\pi| \leq \frac{\pi}{2}$

则 $$|r_n^{(b)}| < \frac{\left(\frac{\pi}{2}\right)^{21}}{(21)!}$$

当 x 比较大时，由方法(a)产生的截断误差会明显大于由(b)方法产生的截断误差，则相较之下，(b)方法得到的结果更准确。

下面解释 (a) & (b) 的相对误差：

观察发现，|x| 较小时，误差也较小，|x| 增大时，误差会指数型增长。

设 $$E = \sum_{n=0}^{10} (-1)^n \frac{x^{2n+1} - (x-k\pi)^{2n+1}}{(2n+1)!}$$

其中 $|x - k\pi| < \frac{\pi}{2}$

当 |x| 较大时， $$E = \frac{x^{21}}{21!} - \frac{(x-k\pi)^{21}}{21!} + o(x^{21})$$

注意到 $$\lim_{x\to+\infty} \frac{(x-k\pi)^{21}}{21!} \Big/ x^{21} = 0$$

$$\Rightarrow \quad E = \frac{x^{21}}{21!} + o(x^{21})$$

$$\Rightarrow \quad |E| \geq \left| \frac{x^{21}}{21!} + o(x^{21}) \right|$$

$$\Rightarrow \quad \log|E| = O(\log|x|)$$

因此，当 $|x|$ 增大时 $|E|$ 也增大，并且整体上呈指数增长 □

3

3. A backward error can sometimes be interpreted as a forward error in a certain sense. Let us consider the following example.

You are given a nonsingular matrix $A \in \mathbb{R}^{n\times n}$ and a vector $b \in \mathbb{R}^n$. For an approximate solution $\hat{x} \in \mathbb{R}^n$ to the linear system $Ax = b$, try to find two vector norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ such that

$$\|r\|_\alpha = \|\hat{x} - x_*\|_\beta,$$

where $r = b - A\hat{x}$ is the residual vector and $x_* = A^{-1}b$ is the exact solution.

解：可取 $\|\cdot\|_\alpha = \|\cdot\|_2$，$\|\cdot\|_\beta = \|\cdot\|_A$

其中 $\|v\|_A = \sqrt{v^T A^T A v} = \|Av\|_2$

由于 $\det(A) \neq 0$，则当且仅当 $v=0$，才有 $\|Av\|_2 = 0$

故满足正定性

而 $\|\alpha v\|_A = \sqrt{\alpha v^T A^T A v \alpha} = |\alpha| \|Av\|_2 = |\alpha| \|v\|_A$

故满足齐次性

又有 $\|v+w\|_A = \|A(v+w)\|_2$

$\leq \|Av\|_2 + \|Aw\|_2 = \|v\|_A + \|w\|_A$

故满足三角不等式

$\Rightarrow \|\cdot\|_A$ 是一个范数

注意到：

$r = b - A\hat{x} = A(A^{-1}b - \hat{x}) = A(x_* - \hat{x})$

$$\Rightarrow \quad \|b - A\hat{x}\|_2 = \|A(x_* - \hat{x})\|_2$$

$$\Rightarrow \quad \|b - A\hat{x}\|_2 = \|x_* - \hat{x}\|_A$$

$$\Rightarrow \quad \|r\|_2 = \|\hat{x} - x_*\|_A \quad \square$$

4

4. (optional) Given a nonzero vector $x \in \mathbb{R}^n$ and a symmetric positive definite matrix $A \in \mathbb{R}^{n\times n}$, both already stored in the floating-point format. Estimate the rounding error for evaluating $x^\top Ax$.

You may assume that there is no overflow or (gradual) underflow.

正定性怎么用？

解：一方面：$fl(x^T y) = (x + \Delta x)^T y$

其中 $|\Delta x_i| \le \gamma_n |x_i|$

另一方面：$fl(Ax) = A \cdot (x + \Delta x)$

也同样有 $|\Delta x_i| \le \gamma_n |x_i|$

那么：$|fl(x^TAx) - x^TAx|$

$= |fl(x^T \cdot fl(Ax)) - x^TAx|$

$= |fl(x^T \cdot (A \cdot (x + \Delta x_2))) - x^TAx|$

$= |(x + \Delta x_1)^T A (x + \Delta x_2) - x^TAx|$

$= |\Delta x_1^T A x + x^T A \Delta x_2 + \Delta x_1^T A \Delta x_2|$

$$\leq |\Delta X_1^T A x| + |x^T A \Delta X_2| + |\Delta X_1^T A \Delta X_2|$$

$$\leq \gamma_n \sum_{i,j} x_i a_{ij} x_j + \gamma_n \sum_{i,j} x_i a_{ij} x_j + \gamma_n^2 \sum_{i,j} x_i a_{ij} x_j$$

$$= O(\gamma_n x^T A x) \quad \square$$

Question：以上过程并没有利用到 $A \succ 0$ 的好性质，但我印象里正定阵往往是数值稳定的。不知如何利用这个特性呢？