Computational Physics(A)
Assignment 1

Chon Hei Lo*(罗俊熙)

School of Physics, Peking University

November 5, 2023

注 *1:* 此作业的解答如无说明，统一使用爱因斯坦求和约定。

1 Problems & Solutions

1.1 数值误差的避免 (15pt)

对 x 从 0 到 100，以 10 为步长，编写程序，比较、讨论下列三种计算 e^{-x} 的方法：

(a) (5pt) 直接展开法

$$e^{-x} = \sum_{0}^{\infty} \frac{(-1)^n x^n}{n!}. \tag{1}$$

(b) (5pt) 递归法

$$e^{-x} = \sum_{0}^{\infty} s_n = \sum_{0}^{\infty} \frac{(-1)^n x^n}{n!}, \tag{2}$$

其中 $s_n = -s_{n-1}\frac{x}{n}$。

(c) (5pt) 先计算 e^x，然后求倒数。

Solution: 参考程序 `1-1.py`，由于 x 十分巨大，因此求和部份只取至第 10 项。其结果如??。由于展开只限于 $x = 0$ 附近时成立，对 x 取过大，由于阶乘项和指数项所带来的修正极为巨大，而且需要在 n 取至比 x 大很多才够收敛，因此无论是直接展开还是递归法，都不是个计算 e^x 的好方法。最好还是在对数域进行运算，这样能够保证相对误差是一致的。

*Email: see.looooo@stu.pku.edu.cn; StudentID: 2000012508

Table 1: 不同方法计算 e^{-x} 的结果

x	真实值	直接展开法	递归法	倒数展开法	倒数递归法
0	1.00e+0	1.00e+0	1.00e+0	1.00e+0	1.00e+0
10	4.54e-5	-2.77e+1	-2.77e+1	4.56e-5	4.56e-5
20	2.06e-9	-2.18e+7	-2.18e+7	4.38e-9	4.38e-9
30	9.36e-14	-5.80e+10	-5.80e+10	4.28e-12	4.28e-12
40	4.25e-18	-1.52e+13	-1.52e+13	2.41e-14	2.41e-14
50	1.93e-22	-1.13e+15	-1.13e+15	4.03e-16	4.03e-16
60	8.76e-27	-3.79e+16	-3.79e+16	1.38e-17	1.38e-17
70	3.98e-31	-7.35e+17	-7.35e+17	7.83e-19	7.83e-19
80	1.80e-35	-9.55e+18	-9.55e+18	6.47e-20	6.47e-20
90	8.19e-40	-9.15e+19	-9.15e+19	7.13e-21	7.13e-21

1.2 矩阵的模与条件数 (25pt)

考虑一个具体的上三角矩阵 $A \in \mathbb{R}^{n\times n}$，其所有对角元都为 1，而所有的上三角部分矩阵元都是 -1。

(a) (5pt) 计算矩阵 A 的行列式，说明 A 的确不是奇异矩阵。

(b) (5pt) 给出矩阵的逆矩阵 A^{-1} 的形式。

(c) (5pt) 如果我们采用矩阵 p 模的定义，

$$\|A\|_p = \sup_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} \tag{3}$$

其中等式右边的模函数 $\|\cdot\|_p$ 是标准定义的矢量 p 模，说明如果取 $p \to \infty$，得到的所谓 ∞ 模为：

$$\|A\|_\infty = \max_{i=1,\ldots,n} \sum_{j=1}^{N} |a_{ij}| \tag{4}$$

(d) (5pt) 矩阵的模有多种定义方法。一种常用的是 $p = 2$ 的欧氏模 $\|\cdot\|_2$。我们有一个幺正矩阵 $U \in \mathbb{C}^{n\times n}$，证明：

$$\|U\|_2 = \left\|U^\dagger\right\|_2 = 1. \tag{5}$$

和证明对于任意的 $A \in \mathbb{C}^{n\times n}$，有：

$$\|UA\|_2 = \|A\|_2. \tag{6}$$

因此，如果利用欧氏模来定义条件数，$K_2(A) = K_2(UA)$。

(e) (5pt) 利用这个定义计算上面给出的具体的矩阵的条件数 $K_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty$。

Solution:

(a) 由于 A 是上三角矩阵，其行列式可以表示为其对角元之积：

$$|A| = \prod_{i=1}^{n} A_{ii} = 1$$

(b) 用行变换的形式求解，若增广矩阵 $[I|A]$，可以通过行变换成为 $[B|I]$，则 $B = A^{-1}$。具体过程如下：

$$[I|A] = \begin{pmatrix} 1 & \dots & 0 & 0 & 0 & 1 & \dots & -1 & -1 & -1 \\ & \ddots & & \vdots & & & \ddots & & \vdots & \\ & & 1 & 0 & 0 & & & 1 & -1 & -1 \\ & & & 1 & 0 & & & & 1 & -1 \\ & & & & 1 & & & & & 1 \end{pmatrix}.$$

$$= \begin{pmatrix} 1 & \dots & 0 & 0 & 1 & 1 & \dots & -1 & -1 & 0 \\ & \ddots & & \vdots & & & \ddots & & \vdots & \\ & & 1 & 0 & 1 & & & 1 & -1 & 0 \\ & & & 1 & 1 & & & & 1 & 0 \\ & & & & 1 & & & & & 1 \end{pmatrix}.$$

$$= \begin{pmatrix} 1 & \dots & 0 & 1 & 2 & 1 & \dots & -1 & 0 & 0 \\ & \ddots & & \vdots & & & \ddots & & \vdots & \\ & & 0 & 1 & 2 & & & -1 & 0 & 0 \\ & & 1 & 1 & 2 & & & 1 & 0 & 0 \\ & & & 1 & 1 & & & & 1 & 0 \\ & & & & 1 & & & & & 1 \end{pmatrix}.$$

$$= \begin{pmatrix} 1 & \dots & 0 & 2 & 4 & 1 & \dots & 0 & 0 & 0 \\ & \ddots & & \vdots & & & \ddots & & \vdots & \\ & & 1 & 2 & 4 & & & 0 & 0 & 0 \\ & & 1 & 1 & 2 & & & 1 & 0 & 0 \\ & & & 1 & 1 & & & & 1 & 0 \\ & & & & 1 & & & & & 1 \end{pmatrix}.$$

$$
= \dots
$$

$$
= \begin{pmatrix}
1 & \dots & 2^{n-4} & 2^{n-3} & 2^{n-2} & 1 & \dots & 0 & 0 & 0 \\
 & \ddots & & \vdots & & & \ddots & & \vdots & \\
 & & 1 & 2 & 4 & & & 0 & 0 & 0 \\
 & & 1 & 1 & 2 & & & 1 & 0 & 0 \\
 & & & 1 & 1 & & & & 1 & 0 \\
 & & & & 1 & & & & & 1
\end{pmatrix}.
$$

故可知，A 的逆为：

$$
A^{-1} = \begin{pmatrix}
1 & 1 & 2 & 4 & & 2^{n-3} & 2^{n-2} \\
 & 1 & 1 & 2 & \dots & 2^{n-4} & 2^{n-3} \\
 & & 1 & 1 & & \vdots & \\
 & & & \ddots & 1 & 2 & 4 \\
 & & & & 1 & 1 & 2 \\
 & & & & & 1 & 1 \\
 & & & & & & 1
\end{pmatrix} \tag{7}
$$

(c) 参考教材的证明方法，在 $\|x\|_\infty = 1$ 的空间上进行证明：

$$
\begin{aligned}
\|A\|_\infty &= \max_{\|x\|_\infty=1} \|Ax\|_\infty \\
&= \max_{\|x\|_\infty=1} \left\| \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \alpha_3^T \\ \vdots \\ \alpha_n^T \end{pmatrix} x \right\|_\infty \\
&= \max_{\|x\|_\infty=1} \left\| \begin{pmatrix} \alpha_1^T x \\ \alpha_2^T x \\ \alpha_3^T x \\ \vdots \\ \alpha_n^T x \end{pmatrix} \right\|_\infty \\
&= \max_{\|x\|_\infty=1} \max_i (|a_i^T x|) \\
&= \max_{\|x\|_\infty=1} \max_i (|\sum_j^n a_{ij} x_j|)
\end{aligned}
$$

$$
\begin{aligned}
&\le \max_{\|x\|_\infty=1} \max_i (\sum_j^n |a_{ij}x_j|) \\
&= \max_{\|x\|_\infty=1} \max_i (\sum_j^n |a_{ij}||x_j|) \\
&\le \max_{\|x\|_\infty=1} \max_i (\sum_j^n [|a_{ij}| \max_k |x_k|]) \\
&= \max_{\|x\|_\infty=1} \max_i (\sum_j^n |a_{ij}|)\|x\|_\infty \\
&= \max_i (\sum_j^n |a_{ij}|) \\
&= \max_i (\|a_i\|_1)
\end{aligned}
$$

(d) 幺正矩阵的定义为 $UU^\dagger = I$，那么：

$$
\begin{aligned}
\|U\|_2 &= \sup_{x\neq 0} \frac{\|Ux\|_2}{\|x\|_2} \\
&= \sup_{x\neq 0} \frac{(Ux)^\dagger Ux}{x^\dagger x} \\
&= \sup_{x\neq 0} \frac{x^\dagger U^\dagger Ux}{x^\dagger x} \\
&= \sup_{x\neq 0} \frac{x^\dagger (U^\dagger U)x}{x^\dagger x} \\
&= 1
\end{aligned}
$$

同理可证 $\left\|U^\dagger\right\|_2 = 1$。类似地：

$$
\begin{aligned}
\|UA\|_2 &= \sup_{x\neq 0} \frac{\|UAx\|_2}{\|x\|_2} \\
&= \sup_{x\neq 0} \frac{(UAx)^\dagger UAx}{x^\dagger x} \\
&= \sup_{x\neq 0} \frac{x^\dagger A^\dagger U^\dagger UAx}{x^\dagger x} \\
&= \sup_{x\neq 0} \frac{x^\dagger A^\dagger (U^\dagger U)Ax}{x^\dagger x} \\
&= \sup_{x\neq 0} \frac{x^\dagger A^\dagger Ax}{x^\dagger x} \\
&= \|A\|_2
\end{aligned}
$$

(e) 由于我们已经知道 A 和 A^{-1} 的表达式，因此易知：

$$
K(A) = \|A\|_\infty \left\|A^{-1}\right\|_\infty = 2^{n-1}. \tag{8}
$$

1.3 Hilbert 矩阵 (30pt)

本题中我们将考虑一个着名的、接近奇异的矩阵，称为 Hilbert 矩阵。

(a) (5pt) 考虑区间 $[0,1]$ 上的任意函数 $f(x)$，我们试图用一个 $(n-1)$ 次的多项式 $P_n(x)=\sum_{i=1}^{n}c_i x^{i-1}$（从而有 n 个待定的系数 c_i) 来近似 $f(x)$。构建两者之间的差的平方的积：

$$D=\int_0^1[(\sum_{i=1}^{n}c_i x^{i-1})-f(x)]^2\mathrm{d}x \tag{9}$$

如果我们要求 D 取极小值，说明各个系数 c_i 所满足的方程为

$$\sum_{j=1}^{n}(H_n)_{ij}c_j=b_i \tag{10}$$

其中 $i,j=1,...,n$。或者简写为矩阵形式：$H_n\mathbf{c}=\mathbf{b}$，其中 $\mathbf{c},\mathbf{b}\in\mathbb{R}^n$，而 $H_n\in\mathbb{R}^{n\times n}$ 就称为 n 阶的 Hilbert 矩阵。给出矩阵 H_n 的矩阵元的表达式和矢量 $\mathbf{b}$ 的表达式 (用包含函数 $f(x)$ 的积分表达)。

(b) (10pt) 请证明矩阵 H_n 是对称正定的矩阵，即对于任意的 $c\in\mathbb{R}^n$，说明 $c^T\cdot H_n c\geq 0$ 其中等号只有当 $c=0$ 时会取得。进而运用线性代数的知识论证 Hilbert 矩阵 H_n 是非奇异的。

(c) (5pt) 虽然矩阵 H_n 是非奇异的，但是它的行列式随着 n 的增加会迅速地减小。事实上，它的行列式竟然有严格的表达式：

$$\begin{cases}\det(H_n)=\dfrac{c_n^4}{c_{2n}}\\ \qquad c_n=1!\times 2!\times...\times(n-1)!\end{cases} \tag{11}$$

因此 $\det(H_n)$ 会随着 n 的增加而迅速指数减小。结合上述 $\det(H_n)$ 的表达式，估计出 $\det(H_n),n\leq 10$ 的数值。(提示：取对数)

(d) (10pt) 由于 Hilbert 矩阵的近奇异性，它具有非常巨大的条件数。因此在求解它的线性方程时，误差会被放大。为了有所体会，请写两个程序，分别利用 GEM 和 Cholesky 分解来求解线性方程 $H_n x=b$，其中 $b=(1,1,...,1)^T\in\mathbb{R}^n$。从小的 n 开始并逐步增加 n（比如说一直到 $n=10$)，两种方法给出的解有差别吗？如果有，你认为哪一个更为精确呢？简单说明理由。

Solution:

(a) 若要求 $D(\mathbf{c})$ 取最小值，那么则有：

$$\frac{\partial D(\mathbf{c})}{\partial c_i} = 0. \tag{12}$$

代入积分式后，可以得到

$$\begin{aligned}\frac{\partial D(\mathbf{c})}{\partial c_i} &= \frac{\partial}{\partial c_i}\int_0^1 \left(\sum_{j=1}^{n} c_j x^{j-1} - f(x)\right)^2 \mathrm{d}x \\ &= \int_0^1 \left(\sum_{j=1}^{n} c_j x^{j-1} - f(x)\right) x^{i-1}\mathrm{d}x = 0\end{aligned}$$

整理后可得：

$$\begin{aligned}\int_0^1 \sum_{j=1}^{n} c_j x^{i+j-2}\mathrm{d}x &= \int_0^1 f(x)x^{i-1}dx \\ \sum_{j=1}^{n} \frac{c_j}{i+j-1} &= \int_0^1 f(x)x^{i-1}dx\end{aligned}$$

定义 $(H_n)_{ij} = \frac{1}{i+j-1}$ 和 $\mathbf{b} = \int_0^1 f(x)x^i dx$，则上式化简为：

$$(H_n)_{ij}c_j = b_i \tag{13}$$

(b) 即证 $\mathbf{c}^T H_n \mathbf{c} > 0$，其中左式为：

$$\begin{aligned}\sum_{i,j}^{n} \frac{c_i c_j}{i+j-1} &= \sum_{i,j}^{n} \int_0^1 c_i c_j t^{i+j-2}\mathrm{d}t \\ &= \int_0^1 \left(\sum_{i}^{n} c_i t^{i-1}\right)^2 \mathrm{d}t \geq 0\end{aligned}$$

当且仅当 $c_i = 0$ 时，等号成立。

(c) 先对 $\det(H_n)$ 取对数，直接计算不进行进一步近似是因为斯特林公式的误差项会在 c_n 中累积。

$$\begin{aligned}\ln\det(H_n) &= \ln\left(\frac{c_n^4}{c_{2n}}\right) \\ &= 4\ln(c_n) - \ln(c_{2n})\end{aligned}$$

使用 numpy 来进行求解，程式码如下：

```
import numpy as np
def c(n):
```

```python
    def factorial(n):
        return np.prod(np.arange(1, n+1, dtype=np.float64))
    out = 1
    for i in range(1,n):
        out *= factorial(i)
    return out
for n in range(1, 11):
    print(f"ln det(H_{n})={4*np.log(c(n))-np.log(c(2*n))}")
```

如??，可见到 H_n 的行列式值下降得十分快。

Table 2: H_n

$\ln\|H_1\|$	$\ln\|H_2\|$	$\ln\|H_3\|$	$\ln\|H_4\|$	$\ln\|H_5\|$	$\ln\|H_6\|$	$\ln\|H_7\|$	$\ln\|H_8\|$	$\ln\|H_9\|$	$\ln\|H_{10}\|$
0.00	−2.48	−7.68	−15.62	−26.31	−39.77	−55.99	−74.98	−96.74	−121.26

(d) 程序如 `1-4.py`，由于输出的内容太多，这里就不列出，`1-4.py` 文件的最后已经附上其输出。对于 $n \leq 10$ 两个方法的输出几乎是一样的，但对于 $n = 13$，时，两个方法的输出存在一定精度差异，其中 GEM 更为准确（通过计算 $H_n\mathbf{x} - 1$ 得到），原因我认为是因为进行 Cholesky 分解时，多解了次方程组，因此误差的累积更为严重，而对 GEM 分解，是直接进行变换解得的，因此误差更小。

1.4 矩阵与二次型 (15pt)

(a) (5pt) 已知矩阵 B 的形式为:

$$B = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

求解矩阵的特征值和特征向量

(b) (10pt) 将矩阵 B 进行对角化，$B = Q\Sigma Q^T$，其中 $Q^{-1} = Q^T$，写出矩阵 Q, Σ。对二次型 $\frac{1}{2}u^T Bu = 1$，进行作图，画出矩阵 B 的三个特征向量的方向。

Solution:

(a) 使用 λ 多项式求解：

$$\begin{aligned}\det(B-\lambda I) &= \begin{vmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)\begin{vmatrix} 2-\lambda & -1 \\ -1 & 1-\lambda \end{vmatrix} + \begin{vmatrix} -1 & 0 \\ -1 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)(\lambda^2-3\lambda) \\ &= \lambda(1-\lambda)(\lambda-3)\end{aligned}$$

对 $\lambda=0$，解 $Bx=0$，得到：

$$x_{\lambda=0} = \begin{pmatrix}1,1,1\end{pmatrix}^T$$

同理可以得到：

$$x_{\lambda=1} = \begin{pmatrix}1,0,-1\end{pmatrix}^T, \quad x_{\lambda=3} = \begin{pmatrix}1,-2,1\end{pmatrix}^T$$

因此，B 的特征值为 $\lambda_1=0,\lambda_2=1,\lambda_3=3$，对应的特征向量为：

$$x_{\lambda=0} = \begin{pmatrix}1,1,1\end{pmatrix}^T, \quad x_{\lambda=1} = \begin{pmatrix}1,0,-1\end{pmatrix}^T, \quad x_{\lambda=3} = \begin{pmatrix}1,-2,1\end{pmatrix}^T \tag{14}$$

(b) 由于 B 是非简并的，因此三个特征向量自然就是正交的，因此可以直接进行归一化：

$$Q^T = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

因此，B 可以表示为：

$$B = Q\Sigma Q^T$$

求出二次型：

$$\begin{aligned}\frac{1}{2}u^TBu &= 1 \\ \frac{1}{2}u^TQ\Sigma Q^Tu &= 1 \\ \frac{1}{2}(Qu)^T\Sigma(Qu) &= 1 \\ 1 &= \frac{1}{2}(\frac{1}{\sqrt{3}}x - \frac{2}{\sqrt{6}}z)^2 + \frac{3}{2}(\frac{1}{\sqrt{3}}x - \frac{1}{\sqrt{2}}y + \frac{1}{\sqrt{6}}z)^2\end{aligned}$$

进行作图，画出矩阵 B 的三个特征向量的方向，程式码如 `1-4.py`，其结果如??。

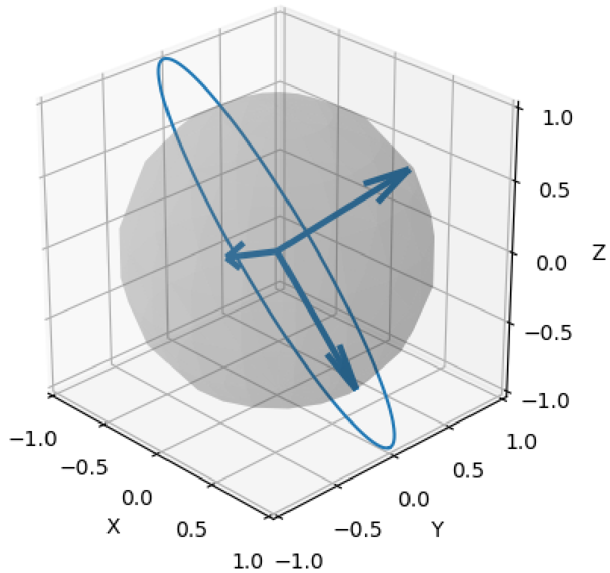

Figure 1: 矩阵 B 的三个特征向量的方向及其二次型

1.5 正定矩阵 (15pt)

对称矩阵 K 如果满足对于任意非零向量 u 都有 $u^TKu>0$，就被称为"正定"矩阵。

(a) (7pt) 我们根据矩阵

$$A=\begin{pmatrix}1 & & & \\ -1 & 1 & & \\ & -1 & 1 & \\ & & -1 & 1 \\ & & & -1\end{pmatrix}$$

构造

$$K_4=A^TA=\begin{pmatrix}2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2\end{pmatrix}$$

请证明对于每一个非零向量 u，都有 $u^TK_4u=u^TA^TAu>0$。请先说明为什么 $u^TA^TAu\geq 0$，然后说明只能取大于号。

(b) (8pt) 对于哪些 b 的取值，这个矩阵 S 是正定的？对于哪些 b，它是半正定的？它的主元是什么？

Solution:

(a) u^TK_4u 满足：

$$\begin{aligned}u^TK_4u&=u^TA^TAu\\&=(Au)^TAu\geq 0\end{aligned}$$

当且仅当 $Au=0$ 时，等号成立，即 $u\neq\mathbf{0}$，因此 $u^TK_4u>0$。

(b) 由正定和半正定的定义，显然 S 的一阶主子式为正，考虑二阶主子式：

$$\det(S)=8-b^2$$

当 $-2\sqrt{2}\geq b\geq 2\sqrt{2}$ 时，S 是半正定的；当 $-2\sqrt{2}<b<2\sqrt{2}$ 时，S 是正定的。