计算物理作业 1

季宇桐 2100017820

2023 年 10 月 28 日

1 数值误差的避免

为便于观察变化趋势，本大题的呈现结果为一表格，代表了 x 从 0 到 100 以步长为 10 增加时，从 0 到 n 求和的结果，每一行代表某一求和上限下不同 x 值的结果，可以用如下示意表表示，其中 result= $\sum_{n=0}^{N} f_n(x)$，生成结果应在表格内容中（具体数据运行程序）。

result \ x / N	0	10	20	30	···	90	100
0	1	1	1	1	···	1	1
1	1						
2	1						
3	1			···	···		
4	1						
···	···						
100	1						

表 1: 示意表

1.1 直接展开法

本题采用直接展开，分别计算每一项并相加。从程序运行结果可以看到，此方法得到的结果不收敛，并且摆动越来越大，在 $n = 34$ 附近就溢出了，用此方法计算不可行，在计算其中的大项时会导致溢出。

1.2 递归法

本题编写递归函数并调用。从运行结果来看，由于使用递归规避了大数的阶乘和幂，递归法可以收敛到稳定的值，但是由于存在大数相消效应，导致计算得到的结果误差很大，如对 e^{-10} 的计算收敛到了 $-6.25618\text{e}-05$，明显偏离实际值 4.54×10^{-5}。

1.3 取倒数法

由于计算 e^x 时全为正数，结果是稳定的，所以可以稳定获得正确的值，取倒数后获得稳定收敛的 e^{-x}，经验证，所得结果与真实值相符，这是有效可行的算法。

2 矩阵的模与条件数

(a)

依题意，

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

由于这是个上三角矩阵，因此在计算矩阵的行列式时每一列只有对角元有贡献，即

$$\det(A)=1^n=1$$

因此该矩阵是非奇异的。

(b)

使用初等行变换法求逆矩阵，$(A,I)\to(I,A^{-1})$. 将 A 变为单位矩阵时可以使用如下的步骤：先将第 n 行加到第 $n-1$ 行，再把变换后的第 $n-1$ 和第 n 行加到第 $n-2$ 行，以此类推。将相同的初等行变换作用到单位矩阵上，于是得到

$$A^{-1}=\begin{pmatrix}1 & 1 & 2 & \cdots & 2^{n-3} & 2^{n-2}\\ & 1 & 1 & 2 & \cdots & 2^{n-3}\\ & & \ddots & \ddots & \vdots & \vdots\\ & & & 1 & 1 & 2\\ & & & & 1 & 1\\ & & & & & 1\end{pmatrix}$$

(c)

由定义

$||\vec{x}||_\infty=\max_{1\le i\le n}|x_i|$,

$$\begin{aligned}||A||_\infty &= \sup_{\vec{x}\neq 0}\frac{||A\vec{x}||_\infty}{||\vec{x}||_\infty}\\ &= \sup_{\vec{x}\neq 0}\frac{\max_{1\le i\le n}\{\sum_{j=1}^m|a_{ij}x_j|\}}{\max_{1\le i\le n}\{|x_i|\}}\\ &= \sup_{\vec{x}\neq 0}\max_{1\le i\le n}\{\sum_{j=1}^m|a_{ij}|\frac{|x_j|}{|x|_{\max}}\}\end{aligned}$$

要求上确界，在 $\vec{x}$ 的每一个分量都是 $|x|_{\max}$ 时可以取到，即

$$||A||_\infty=\max_{1\le i\le n}\{\sum_{j=1}^m|a_{ij}|\}$$

(d)

对幺正矩阵有 $U^\dagger U = 1$，因此

$$\frac{||U\vec{x}||_2}{||\vec{x}||_2} = \frac{(\vec{x}^\dagger U^\dagger U \vec{x})^{\frac{1}{2}}}{(\vec{x}^\dagger \vec{x})^{\frac{1}{2}}} = \frac{(\vec{x}^\dagger \vec{x})^{\frac{1}{2}}}{(\vec{x}^\dagger \vec{x})^{\frac{1}{2}}} = 1$$

同理

$$\frac{||U^\dagger\vec{x}||_2}{||\vec{x}||_2} = \frac{(\vec{x}^\dagger U U^\dagger \vec{x})^{\frac{1}{2}}}{(\vec{x}^\dagger \vec{x})^{\frac{1}{2}}} = \frac{(\vec{x}^\dagger \vec{x})^{\frac{1}{2}}}{(\vec{x}^\dagger \vec{x})^{\frac{1}{2}}} = 1$$

对于 $||UA||_2$，有

$$||UA||_2 = \frac{(\vec{x}^\dagger A^\dagger U^\dagger U A \vec{x})^{\frac{1}{2}}}{(\vec{x}^\dagger \vec{x})^{\frac{1}{2}}} = \frac{(\vec{x}^\dagger A^\dagger A \vec{x})^{\frac{1}{2}}}{(\vec{x}^\dagger \vec{x})^{\frac{1}{2}}} = ||A||_2$$

因此，

$$K_2(UA) = ||UA||_2 = ||A||_2 = K_2(A)$$

(e)

根据 (c)，最大显然在第一行取到，即

$$\begin{aligned}
&||A||_\infty = \max_{1\le i\le n} \sum_{j=1}^{m} |a_{ij}| = n \\
&||A^{-1}||_\infty = \max_{1\le i\le n} \sum_{j=1}^{m} |a'_{ij}| = 2^{n-1} \\
&K_\infty(A) = n2^{n-1}
\end{aligned}$$

3 Hilbert 矩阵

(a)

取极小值要求 $\frac{\partial D}{\partial c_k} = 0$，即

$$\begin{aligned}
\frac{\partial D}{\partial c_k} &= 0 \\
\int_0^1 2x^{k-1}\left(\sum_{i=1}^{n} c_i x^{i-1} - f(x)\right) dx &= 0 \\
\int_0^1 \left(\sum_{i=1}^{n} c_i x^{k+i-2} - f(x)x^{i-1}\right) dx &= 0 \\
\sum_{i=1}^{n} c_i \frac{1}{k+i-1} &= \int_0^1 f(x) x^{k-1}\, dx
\end{aligned}$$

左边是哑指标，$j \to i, i \to k$可以写作，

$$\sum_{j=1}^{n} c_j \frac{1}{i+j-1} = \int_0^1 f(x) x^{i-1}\, dx$$

对照

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

得,

$$(H_n)_{ij} = \frac{1}{i+j-1}$$
$$b_i = \int_0^1 f(x) x^{i-1}\, dx$$

(b)

由 $(H_n)_{ij} = \frac{1}{i+j-1} = \frac{1}{j+i-1} = (H_n)_{ji}$，故 H 是对称的。

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n} c_i (H_n)_{ij} c_j &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{c_i c_j}{i+j-1} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \int_0^1 x^{i+j-2}\, dx \\ &= \int_0^1 \left(\sum_{i=1}^{n} c_i x^{i-1}\right)^2 dx \geq 0\end{aligned}$$

当且仅当所有的 $c_i = 0$ 时取等，因此该矩阵是正定的。由线性代数知识，正定矩阵合同于单位矩阵，即存在可逆矩阵 C，使得 $H = C^{\mathbf{T}} I C$, 则 $\det(H) = \det(C^{\mathbf{T}} C) = \det(C^{\mathbf{T}}) \det(C) = \det(C)^2$，因此该矩阵是非奇异的。

(c)

取对数后 $\ln\det(H_n) = 4\ln c_n - \ln c_{2n}$，由 Stirling 公式，$\ln(n!) \approx n\ln n - n$，得 $\ln\det(H_n) = 4\sum_{i=1}^{n-1}(i\ln i - i) - \sum_{i=1}^{2n-1}(i\ln i - i)$，下面给出 $n \leq 10$ 的计算结果

n	2	3	4	5	6	7	8	9	10
$\ln\det(H_n)$	-2.68	-9.73	-19.92	-33.15	-49.36	-68.52	-90.60	-115.59	-143.46

由此可见以非常快的速度指数减小。

(d)

具体的计算结果见程序运行，两个结果在 n 增大时出现明显差异，我认为 GEM 方法更准确，因为 Hilbert 矩阵极度病态，在 n 较大时 Cholesky 分解中的开平方等运算的误差会被放大，而 GEM 不涉及开方，相对准确一些。尽管在 n 较小时 Cholesky 算法的结果更接近真实值，但是随着 n 增大，Cholesky 算法很快失效，这也可以从后几项 Cholesky 算法结果无法计算看出来。

4 矩阵与二次型

(a)

依题意

$$B - \lambda I = \begin{pmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{pmatrix} \quad \det(B - \lambda I) = \lambda(\lambda - 1)(3 - \lambda) = 0$$

得到 $\lambda_1=1,\lambda_2=3,\lambda_3=0$，对应特征向量计算如下：

$$(B-\lambda_1 I)\vec{x}_1=\begin{pmatrix}0 & -1 & 0\\ -1 & 1 & -1\\ 0 & -1 & 0\end{pmatrix}\begin{pmatrix}x_{11}\\ x_{12}\\ x_{13}\end{pmatrix}=0$$

令 $|x_1|=1$，得特征值 $\lambda=1$ 的特征向量是 $\vec{x}_1=(\frac{1}{\sqrt{2}},0,\frac{-1}{\sqrt{2}})^{\mathrm{T}}$. 同理，

$$(B-\lambda_2 I)\vec{x}_2=\begin{pmatrix}-2 & -1 & 0\\ -1 & -1 & -1\\ 0 & -1 & -2\end{pmatrix}\begin{pmatrix}x_{21}\\ x_{22}\\ x_{23}\end{pmatrix}=0$$

令 $|x_2|=1$，得特征值 $\lambda=3$ 的特征向量是 $\vec{x}_2=(\frac{1}{\sqrt{6}},\frac{-2}{\sqrt{6}},\frac{1}{\sqrt{6}})^{\mathrm{T}}$. 同理，

$$(B-\lambda_2 I)\vec{x}_2=\begin{pmatrix}1 & -1 & 0\\ -1 & 2 & -1\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix}x_{31}\\ x_{32}\\ x_{33}\end{pmatrix}=0$$

令 $|x_3|=1$，得特征值 $\lambda=0$ 的特征向量是 $\vec{x}_2=(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}})^{\mathrm{T}}$.

(b)

根据实对称矩阵对角化的知识有，

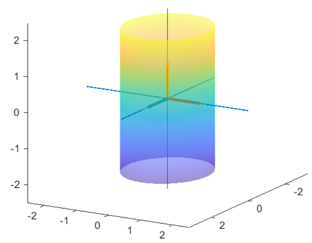
$$Q=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\end{pmatrix}$$
$$\Sigma=\begin{pmatrix}1 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 0\end{pmatrix}$$

验证了 $QQ^{\mathrm{T}}=I$,$B=Q\Sigma Q^{\mathbf{T}}$.

对二次型 $\frac{1}{2}u^{\mathbf{T}}Bu$ 做单位正交变换 $\tilde{u}=Q^{\mathbf{T}}u$，则变换后的二次型是 $\tilde{x}^2+3\tilde{y}^2-2=0$，变换后的特征向量为

$$\vec{x_1}=\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}\ \vec{x_2}=\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}\ \vec{x_3}=\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}$$

绘制图像如下



5 正定矩阵

(a)

$$A\vec{x} = \begin{pmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \\ x_4 - x_3 \\ -x_4 \end{pmatrix}$$

故 $u^{\mathbf{T}}K_4u = (Au)^{\mathbf{T}}(Au) = x_1^2 + (x_2 - x_1)^2 + (x_3 - x_2)^2 + (x_4 - x_3)^2 + (-x_4)^2 \geq 0$

若上式等于零，则 $x_1 = x_2 - x_1 = x_3 - x_2 = x_4 - x_3 = x_4 = 0$，进而 $x_1 = x_2 = x_3 = x_4 = 0$，与 $\vec{x}$ 是非零向量矛盾，故 $u^{\mathbf{T}}K_4u > 0$.

(b)

首先我们将 S 对角化，先计算其特征多项式，为

$$\det(S - \lambda I) = \lambda^2 - 6\lambda + 8 - b^2$$

由求根公式

$$\lambda = 3 \pm \sqrt{1 + b^2}$$

由于 S 是实对称矩阵，因此对应的特征向量在归一化后可以组成单位正交矩阵 U 使 S 对角化，即 $U^{\mathbf{T}}SU = \mathrm{diag}\{\lambda_1, \lambda_2\}$. 对于任意的非零向量 $\vec{u}$，总可以找到 $\tilde{u} = U^{-1}u$，则 $u^{\mathbf{T}}Su = \tilde{u}^{\mathbf{T}}U^{\mathbf{T}}SU\tilde{u} = \lambda_1\tilde{x_1}^2 + \lambda_2\tilde{x_2}^2$.

因此，当 $\lambda_1 > 0, \lambda_2 > 0$ 时矩阵是正定的，即 $-2\sqrt{2} < b < 2\sqrt{2}$；当有一个特征值为 0 时 S 是半正定的，此时 $b = \pm 2\sqrt{2}$.

Gauss 消元后有

$$S \to \begin{pmatrix} 2 & b \\ & 4\frac{b^2}{2} \end{pmatrix}$$

因此主元为 2 和 $4 - \frac{b^2}{2}$，可以看出当主元为正时原矩阵正定，主元 ≥ 0 时原矩阵半正定.