第四次上机作业

陈晓宇 121110008

考虑三对角对称矩阵 Tn，见 (6.0.2)。矩阵阶数分别取为 n = 100 和 n = 101，要求精确计算到准确特征值的小数点后第 6 位。

第一题

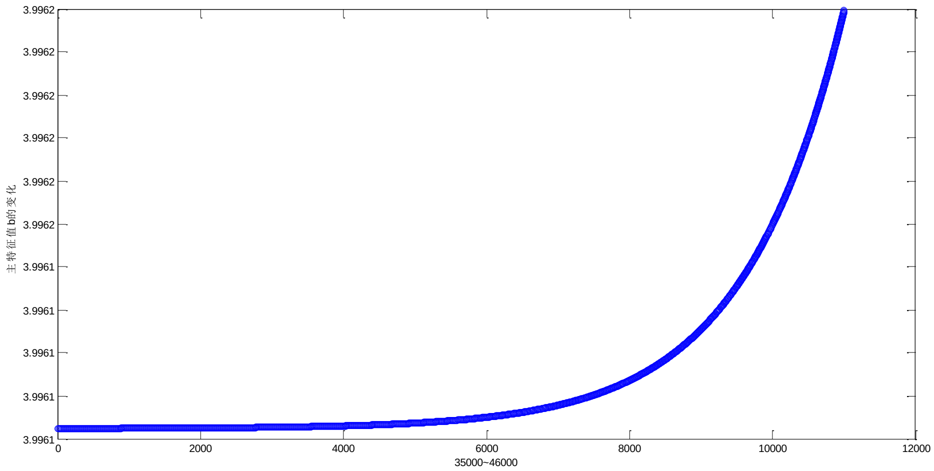
题目：取初始向量为 v0 = (1, 1, 1, . . . , 1)，用乘幂法计算主特征值及其相应的特征向量。请绘制主特征值误差的下降曲线，以及特征子空间距离的下降曲线。然后，请采用 Atiken 加速技巧和 Rayleigh 商技术分别对算法进行加速，并完成类似的工作。

目标：

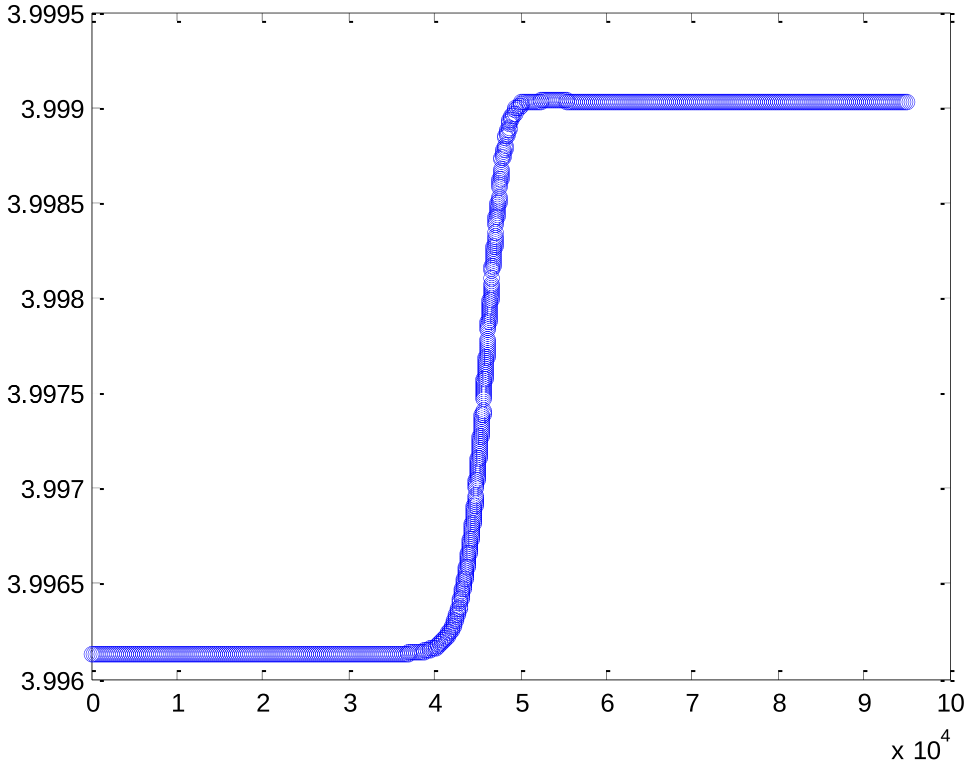
1. 观察在 n = 100 和 101 的时候，乘幂法是否有不同（在关掉误差容限的情况下），验证课上讲的：要在幂法的迭代过程收集到主特征信息，理论上需要初始向量在主特征子空间 span(X1)上的分量必须非 0，若这个分量接近 0 的时候，其数值表现为一个非常缓慢的假收敛过程，此时，舍入误差的积累可能起到积极的作用。
2. 绘制主特征值误差的下降曲线
3. 分别采用 Atiken 加速和 Rayleigh 加速，并绘制主特征值的误差下降曲线和特征子空间距离的下降曲线。

实验结果：

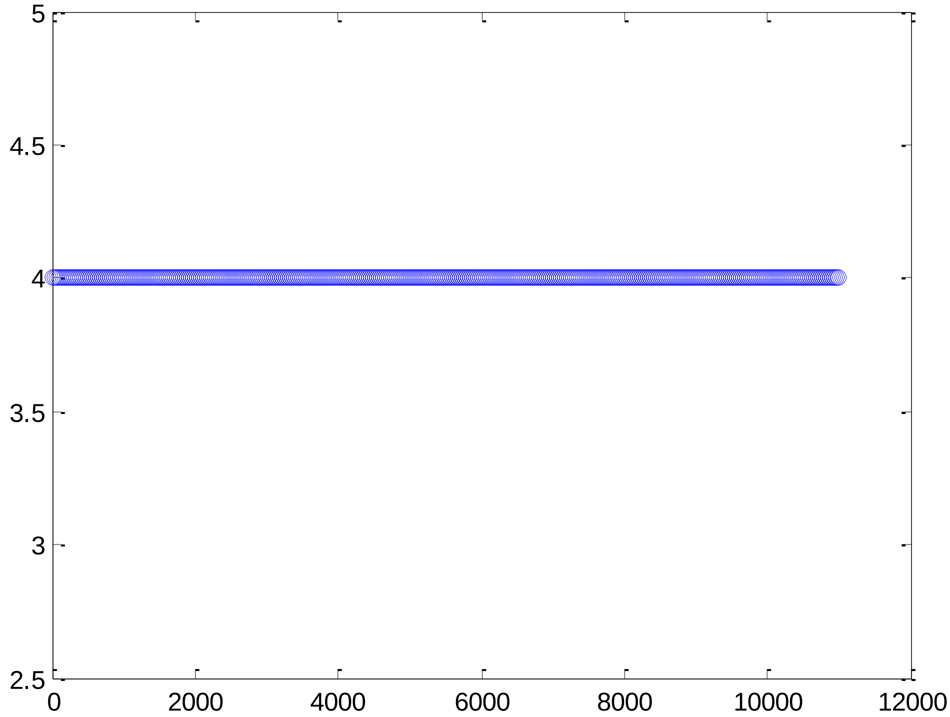
目标 1：



上图是 N=100 时，这是迭代步数在 35000~46000 之间主特征值的变化曲线，



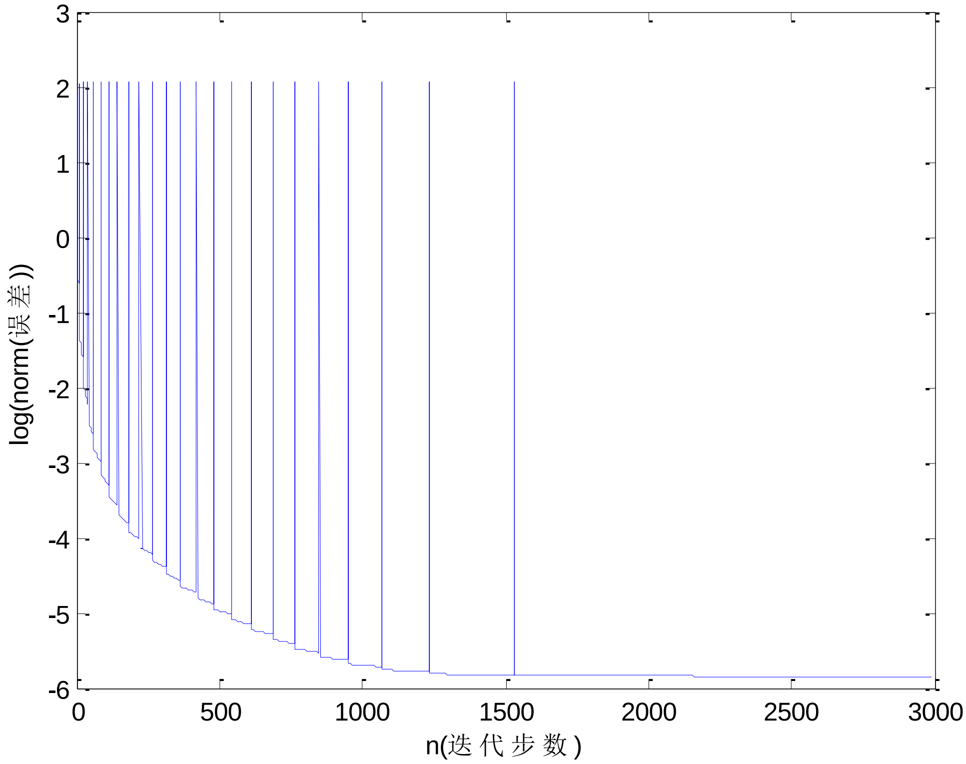
而上图是主特征值随着迭代步数的变化曲线，发现有两个平台，验证了书上所说的假收敛现象。



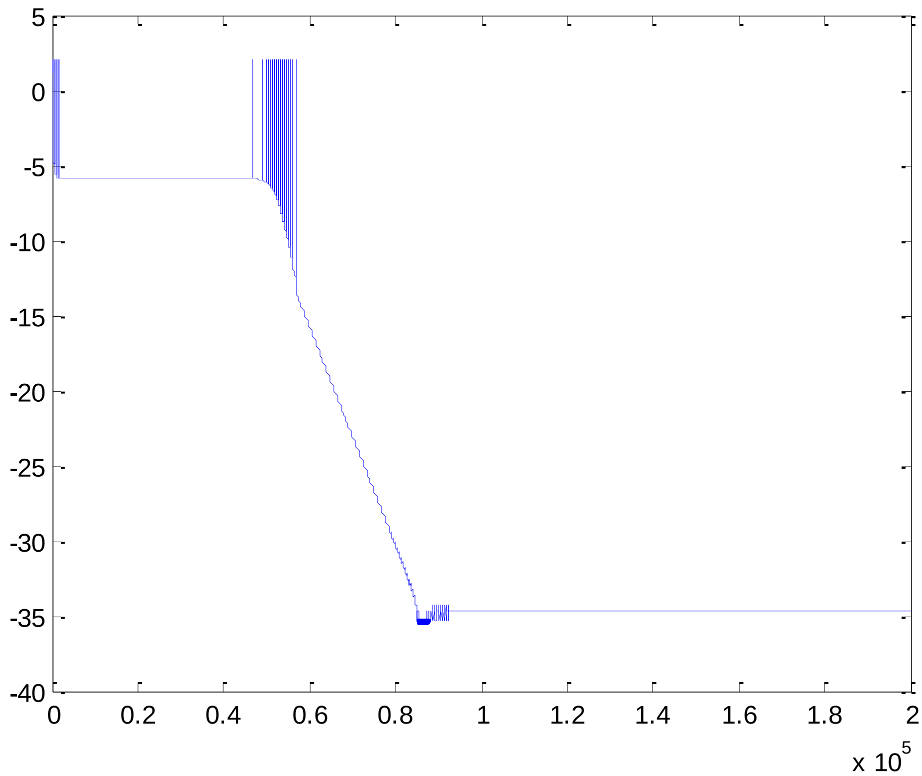
而上图则是 N=101 时，主特征值 b 随着迭代步数的变化曲线，发现，只有一个平台。

实验分析：因为 n=100 时，初始向量与主特征值所对应的特征向量之间的内积为 0，也就是初始向量在主特征子空间的分量为 0，而 n=101 时，二者的内积不为 0.

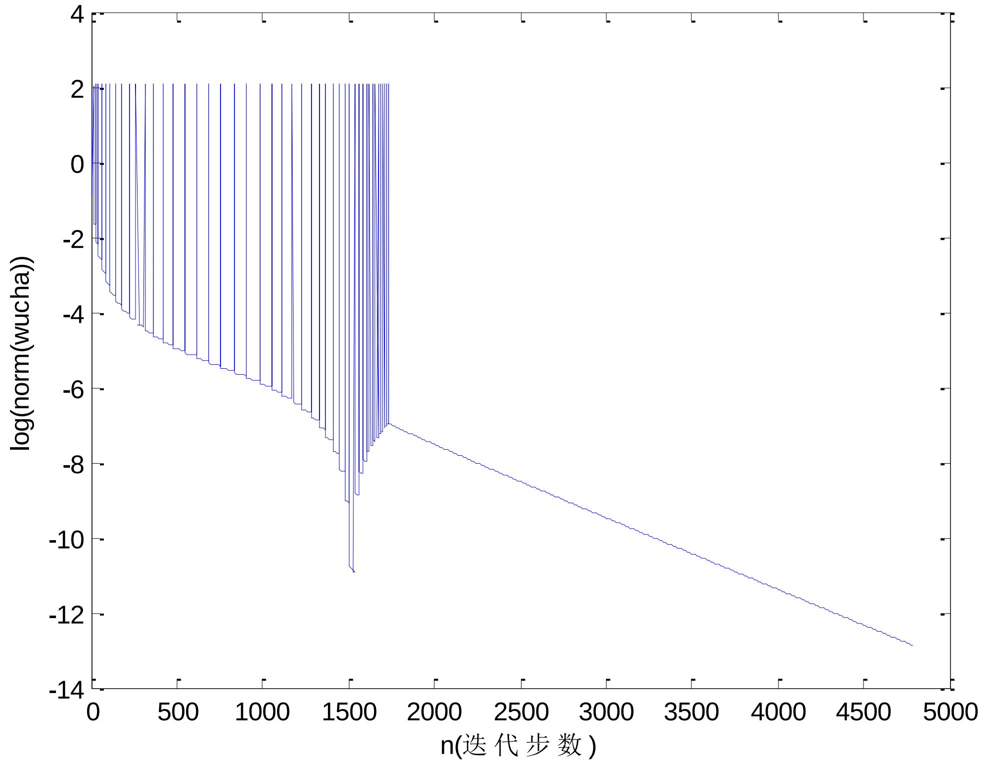
目标 2:



上图是 n = 100 时，主特征值误差随着迭代步数的变化曲线（采用了停机标准），下图是不采用停机标准所作出的图像。



发现出现了两个平台，更能说明假收敛现象。



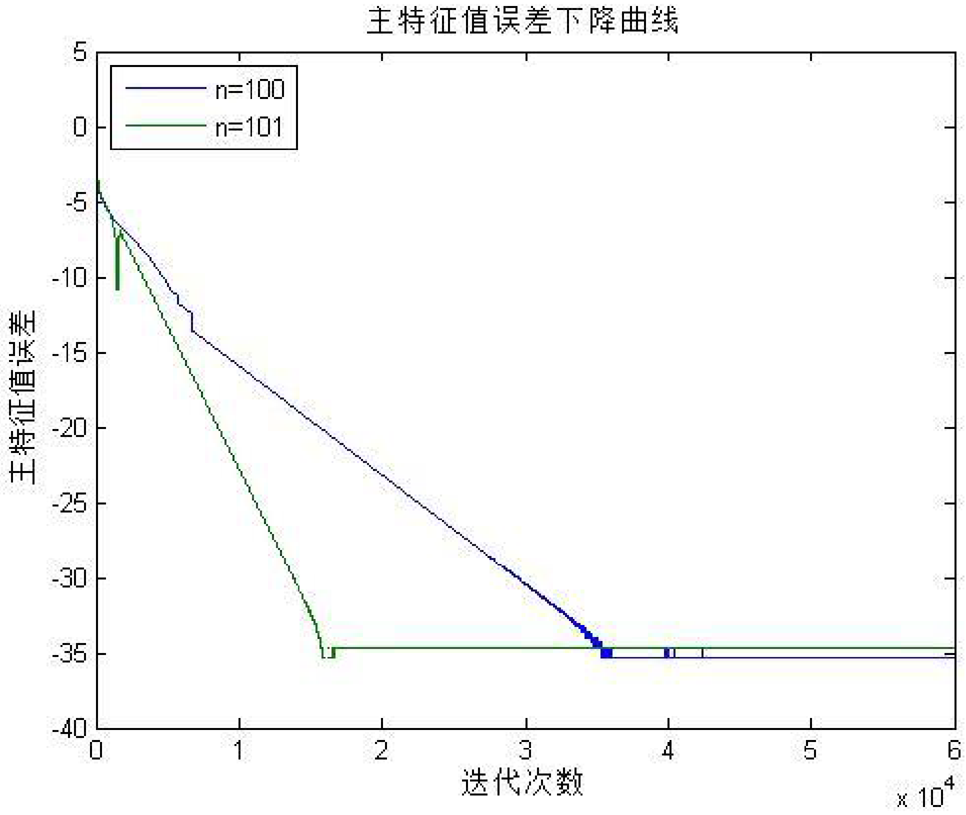
上图是 n = 101 时候的主特征值误差下降曲线，没有出现假收敛现象。

目标 3：分别使用 Aitken 和 Rayleigh 加速技巧对乘幂法进行加速，下面简述两种加速技巧。

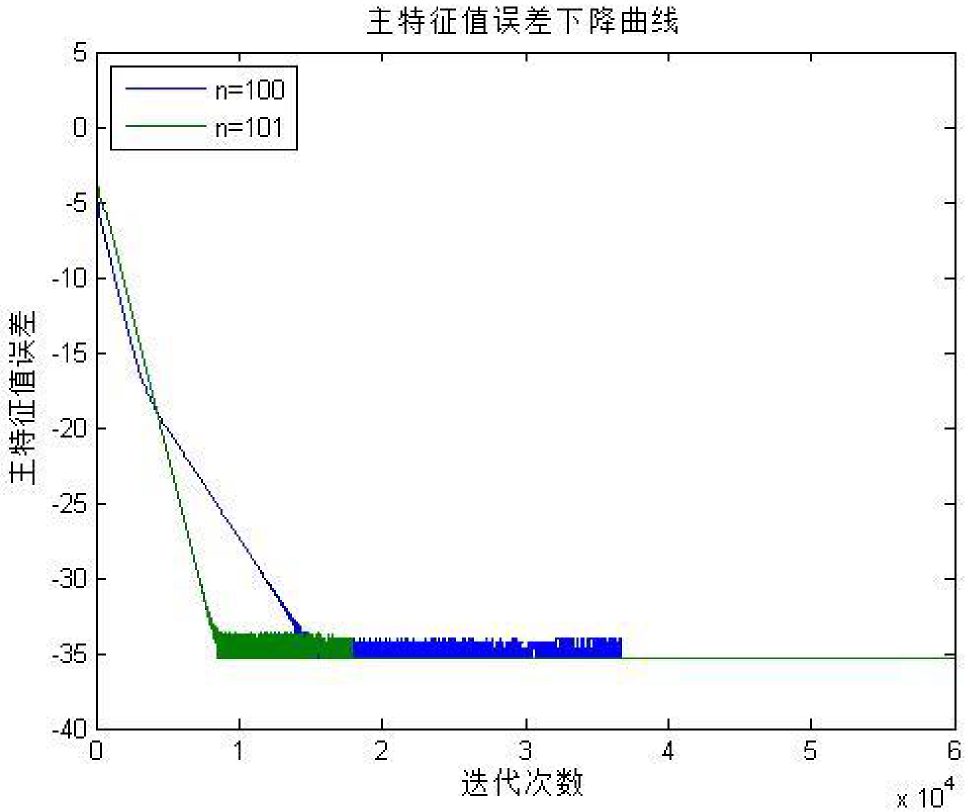
Aitken：利用 $m_k = m_k - \frac{(\triangle m_k)^2}{\triangle^2 m_k}$ 代替原来的 m_k 计算

Rayleigh：利用 $\mathrm{R}(\mathrm{v}_k) = \frac{v_k{}^T A v_k}{v_k{}^T v_k} = \lambda_1 + \mathrm{O}(|\frac{\lambda_2}{\lambda_1}|^{2k})$ 代替原来的 m_k 计算，发现 Rayleigh 商加速法是二阶收敛的。

实验结果和分析：



上图是 Aitken 加速。



上图是 Rayleigh 加速，发现二阶收敛。

第二题

题目：用反幂法求解最靠近 2 的特征值，及其对应的特征向量。观察是否有所谓的“一次迭代”特性。

目标：利用正幂法对于逆矩阵求解按模最小的特征值，注意的是对非奇异矩阵才能使用，对于奇异矩阵没有逆矩阵。即取 n=101 时，不能用反幂法解。

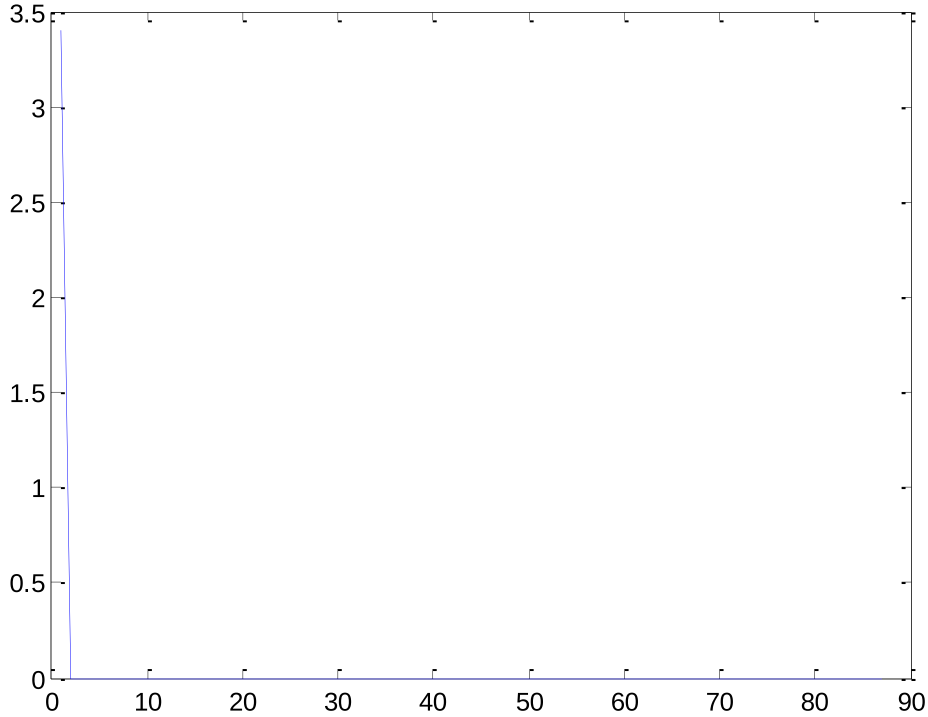
实验过程：原理是$\begin{cases} \mathrm{A}u_k = v_{k-1} \\ v_k = \dfrac{u_k}{m_k} \end{cases}$，其中$m_k = \max(u_k)$，则

当$k \to \infty$时，$m_k \to \dfrac{1}{\lambda_n}$，$v_k \to \dfrac{x_n}{\max(x_n)}$。下面考虑一个更一般且有用的形式，用矩阵$\mathrm{A} - \mathrm{qI}$代替 A 应用反幂法，迭代公式如下

$$\begin{cases} (A - qI)u_k = v_{k-1} \\ v_k = \dfrac{u_k}{m_k} \end{cases}$$

其中$m_k = \max(u_k)$,则 $q + \dfrac{1}{m_k} \to \lambda (k \to \infty)$。由于题目要求靠近 2 的特征值，则取$q = 2$即可。

结果与评价：



上图是主特征值误差随着迭代次数的变化，可以很明显的看到“一次迭代”的效果。容易看出，在初始的几步具有很好的性价比，一下次就很快的逼近了主特征值，后来的只有小幅变化来调整。

第三题

题目：用幂法求解第二主特征值及其特征向量；用同时迭代方法求解T_n的前两个主特征值。比较两者的计算效果。

目标：分别运用降阶法和子空间同时迭代法求解T_n的前两个特征值。

实验过程：

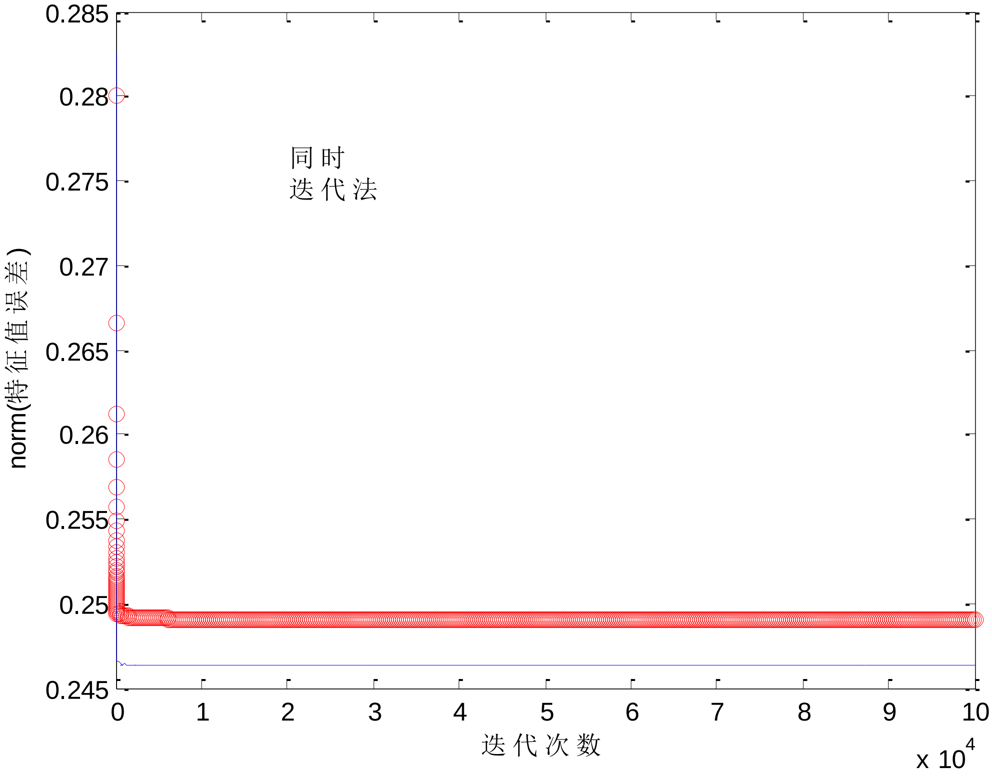
一．降阶法

取用 Householder 变换矩阵$S_1 = I - b^{-1}uu^T$，其中$\|u\|_2 \neq 0$，且$b = \frac{1}{2}\|u\|_2$，在S_1的作用下，$A = S_1 A S_1^{-1} = \begin{vmatrix} \lambda_1 & \omega^T \\ 0 & B_2 \end{vmatrix}$，而其中$B_2$的主特征值就是$A$的第二大特征值。

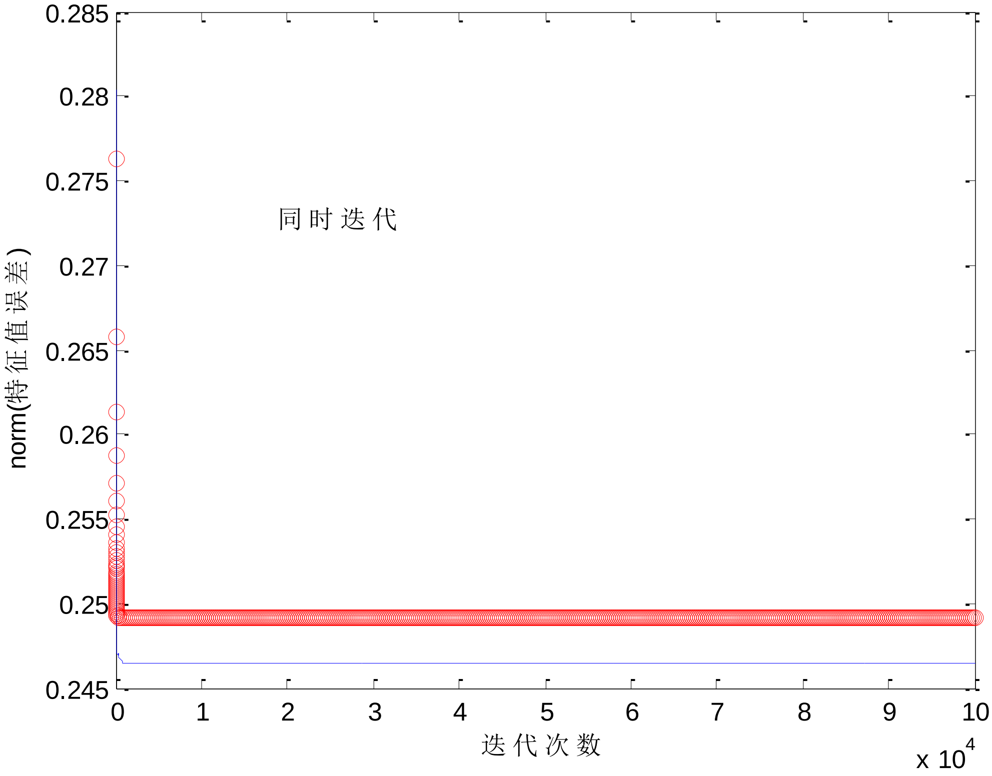
二．同时迭代法

实矩阵 A 有完备的特征向量系，取 m(m<n)个初始近似向量组成一个$m*n$阶列直交阵V_0，对 k=1,2……作下述操作：计算$U_k = AV_{k-1}$；计算$m*n$阶实对称矩阵$B_k = V_{k-1}^T U_k$；计算B_k的特征值的排列顺序为$|\mu_1^{(k)}| \geq |\mu_2^{(k)}| \geq \cdots\cdots \geq |\mu_m^{(k)}|$，并计算其特征向量矩阵$W_k$，取它为一个直交矩阵；计算$U_k W_k$；对作$U_k W_k$直交三角分解：$U_k W_k = V_k R_k$，其中$R_k$为上三角矩阵，$V_k$为$m*n$阶列直交阵；检验相邻两次迭代得到的$\mu_j^{(k)}$和$\mu_j^{(k-1)}$之间的差是否满足精度要求，如果满足则取$\mu_j^{(k)}$作为$\lambda_j$的近似值，$V_k$的第 j 列向量作为与之相应的特征向量，否则，对 k 进行+1。

结果与评价:

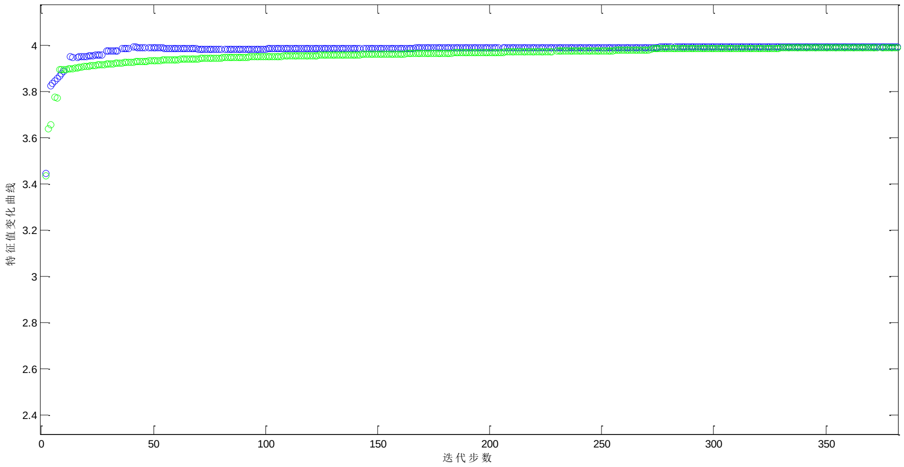


上图是 n = 100 时，主特征值，第二大特征值误差随着迭代步数的变化曲线，红色的是主特征值，蓝色的是第二主特征值。

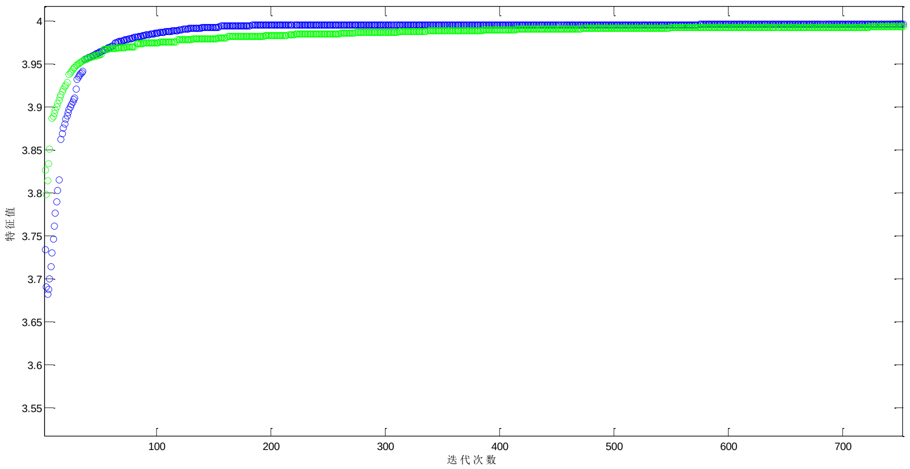


上图是 n = 101 时，主特征值，第二大特征值误差随着迭代步数的变化曲线，红色的是主特征值，蓝色的是第二主特征值。

书上给出的是子空间投影法和空间同时迭代法的综合，也就是上述的实现过程。如果只是空间同时迭代法，只需要任取 m 个列直交向量，组成初始矩阵 V_0，然后，循环执行 $U_k = AV_{k-1}$，$U_k = V_k R_k$。



上图是 n=100 时，特征值变化的曲线，蓝色的是主特征值，绿色的是第二主特征值，为了防止假收敛的现象，选取的初始向量是随机的，保证了在特征子空间上的投影不为 0。由于不会出现假收敛现象，于是不会出现第二主特征值大于主特征值的现象。



上图是 n=101 时，特征值变化的曲线，蓝色的是主特征值，绿色的是第二主特征值。

第四题

题目：分别用古典 Jacobi 方法、循环 Jacobi 和阈值 Jacobi 方法求解T_n的全部特征值；绘制相应的收敛过程。

实现过程：下面分别叙述三种 Jacobi 方法的实现思路

一． 古典 Jacobi 方法

古典 Jacobi 方法，令A_0=A，在第 k 步中，取$a_{pq}^k = \arg\max_{i<j} |a_{ij}|$，将其作为旋转主元，然后令$A = R(p,q)AR(p,q)^T$，其中$R(p,q)$为 Givens 平面旋转矩阵。当$k \to \infty$时，可以证明$diag(A) \to eig(A)$。

二． 循环 Jacobi 方法

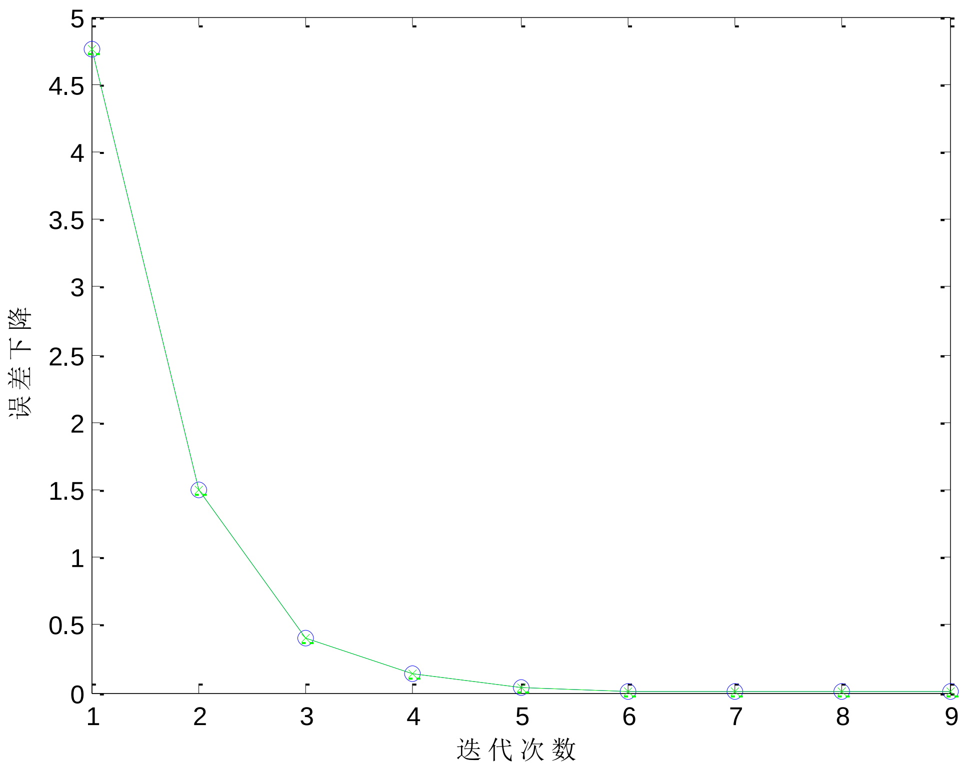
循环的 Jacobi 方法，是对古典 Jacobi 方法的一个改进，古典的 Jacobi 方法需要旋转主元的全局搜索，而循环的 Jacobi 方法直接对矩阵的所有非对角元素，依次进行$\frac{n(n-1)}{2}$次 Jacobi 旋转。

三． 阀值 Jacobi 方法

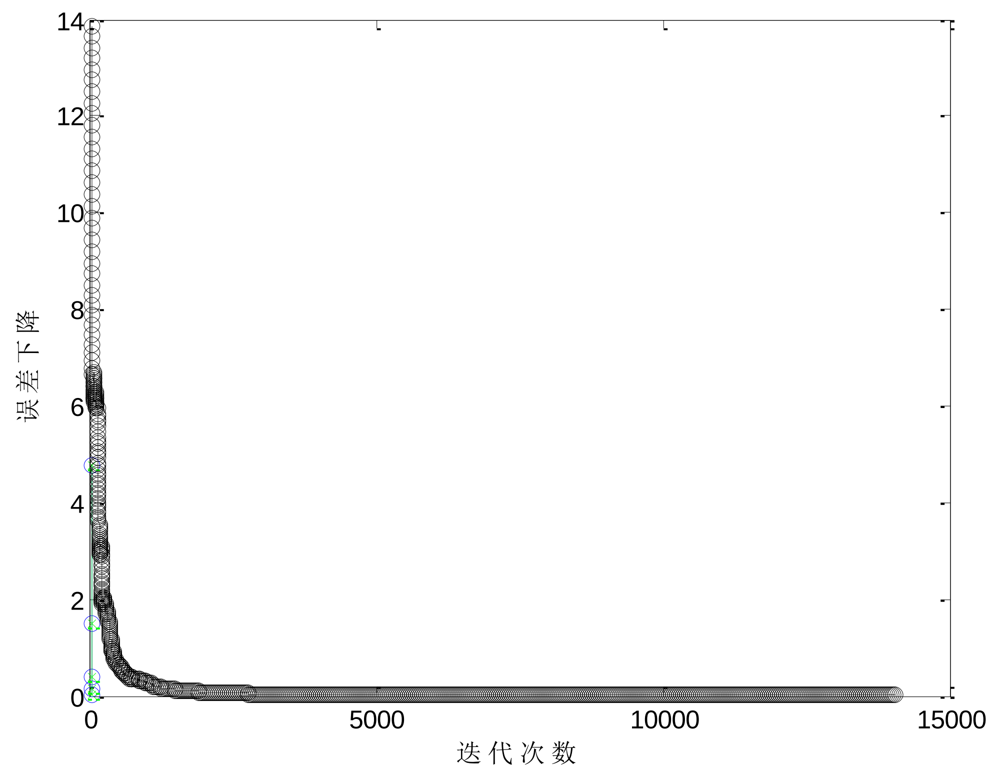
阀值 Jacobi 方法是在循环 Jacobi 方法的改进。设定一个初始的阀值σ，然后逐步设置阀值，直至达到机器精度或用户要求为止，即$\delta_k = \delta_{k-1}/\sigma$，$k = 1,2\cdots\cdots$，当$|a_{pq}|$小于设定的阀值时，跳过相应的 Jacobi 旋转操作，直至所有的非对角线元素均小于给定的阀值时，按照上述规则，重新设立阀值。

实验结果和分析：

会发现，由于阀值和循环的处理，迭代步数的大幅度下降。



这是 n=100 时阀值和循环的 Jacobi 的误差下降曲线



而这是 n=100 时古典 Jacobi 迭代产生的误差下降曲线，比较

两者的迭代次数，相差的比较巨大。书上说当循环 Jacobi 方法收敛时，则该方法具有渐进平方收敛速度，此方法的扫描次数和 logn 成正比，从第一幅图可以看出。

第五题

题目： 用二分法加原点位移反幂法求解在(1,2)间的特征值。

实现过程： 二分法是对实对称三角矩阵 $T=\begin{pmatrix} a_1 & b_1 & 0 \\ b_1 & \ddots & b_{n-1} \\ 0 & b_{n-1} & a_n \end{pmatrix}$ 进行特征值的计算。定义 $T-\lambda I$ 的 r 阶顺序主子式为 $p_r(\lambda)$，它们满足三项递推式 $p_r(\lambda)=(a_r-\lambda)p_{r-1}(\lambda)-b_{r-1}^2 p_{r-2}(\lambda)$，其中 $p_0(\lambda)$=1，用 $S_k(\alpha)$ 表示序列 $p_0(\alpha),p_1(\alpha),\cdots\cdots$，$p_k(\alpha)$ 中相邻两个数中符号相同的数目。对于矩阵 T 的任意一个特征值 λ_i 都满足 $|\lambda_1|\le\max|\lambda_i|=\rho(T)\le|T|_\infty$，于是据此初步锁定根的范围，然后在区间[$|T|_\infty$,$|T|_\infty$]中寻找矩阵 T 的特征值。根据如下定理：对于给定的实数 α，$p_r(\lambda)$ 恰有 $S_r(\alpha)$ 个根严格大于 α，设 a<b 为两个实数，那么由于 $S_n(a)\ge S_n(b)$，则 $S_n(a)-S_n(b)$ 恰是矩阵在区间[a,b]中的特征值个数。设 λ_k 位于区间(a,b]中，取 c=(a+b)/2，计算 $S_n(c)$，如果 $S_n(c)<S_n(a)$，则 $\lambda_k\in(a,c]$，反之则 $\lambda_k\in(c,b]$。再对区间二分，直到最后区间长度小于预先给定的精度控制量，便取最后区间的中点作为 λ_k 的近似值。

在本次实验中，我们先用二分法确定粗糙的特征值，然后再利用原点位移反幂法进行调整，使其更加精确。

实验结果和分析：

T_{100}的特征值=(

1.96889637615930

1.90671921922517

1.84463230542199

1.78269569982905

1.72096932211215

1.65951288855520

1.59838585428856

1.53764735577006

1.47735615357428

1.41757057554550

1.35834823244260

1.29974710161722

1.24182319231924

1.18463277011662

1.12823149087300

1.07267293602935

1.01801183805336)

T_{101}的特征值=(

1.93840988288766

1.87687818773211

1.81546328107340

1.75422341867057

1.69321669024263

1.63250096436686

1.57213383358701

1.51217255978325

1.45267401985583

1.39369465177391

1.33529040104068

1.27751666762569

1.22042825341464

1.16407931022643

1.10852328844692

1.05381288632798)

第六题

题目： 阈值Jacobi方法具有求解小特征值的优势。考虑对称正定矩阵 $A=\begin{pmatrix}10^{40} & 10^{29} & 10^{19}\\ 10^{29} & 10^{20} & 10^{9}\\ 10^{19} & 10^{9} & 1\end{pmatrix}$ 直接计算可知其特征值为 10^{40}，9.9×10^{19}，9.81818×10^{-1}.请用阈值Jacobi方法求解三个特征值。在Matlab中，我们可以利用eig()计算矩阵的特征值。请比较它们的数值结果。

实验结果和分析：

$$threshold(A)=\begin{pmatrix} 10^{40} \\ 9.9*10^{19} \\ 9.81818*10^{-1} \end{pmatrix},$$

$$eig(A)=\begin{pmatrix} 10^{40} \\ 3.9667878456105*10^{23} \\ -8.1*10^{19} \end{pmatrix}$$

可以看出内置函数eig的计算已经错误了。而按照阀值所进行的运算则是正确的。