Numerical Algebra：Homework 10

HUANG Wenchong
3200100006

May 7, 2022

2. 证：将 $\varphi(x)=x^TAx-2b^Tx$ 在 x_{k-1} 处 Taylor展开得：

$$\varphi(x_k)=\varphi(x_{k-1})+\nabla^T\varphi(x_{k-1})(x_k-x_{k-1})+\frac{1}{2}(x_k-x_{k-1})^T\nabla^2\varphi(x_{k-1})(x_k-x_{k-1})$$
$$=\varphi(x_{k-1})-2r_{k-1}^T(\alpha_{k-1}r_{k-1})+\alpha_{k-1}^2(r_{k-1}^TAr_{k-1})$$
$$=\varphi(x_{k-1})-\frac{(r_{k-1}^Tr_{k-1})^2}{r_{k-1}^TAr_{k-1}}$$

又：
$$r_{k-1}^TA^{-1}r_{k-1}=(b-Ax_{k-1})^TA^{-1}(b-Ax_{k-1})$$
$$=b^TA^{-1}b+x_{k-1}^TAx_{k-1}-2b^Tx_{k-1}$$
$$=b^TA^{-1}b+\varphi(x_{k-1})$$

由A正定 $\Rightarrow$ A^{-1}正定 $\Rightarrow$ $b^TA^{-1}b>0$ $\Rightarrow$ $\varphi(x_{k-1})<r_{k-1}^TA^{-1}r_{k-1}$

于是：
$$\varphi(x_k)=\varphi(x_{k-1})\left[1-\frac{(r_{k-1}^Tr_{k-1})^2}{\varphi(x_{k-1})\cdot(r_{k-1}^TAr_{k-1})}\right]$$
$$<\varphi(x_{k-1})\left[1-\frac{r_{k-1}^Tr_{k-1}}{r_{k-1}^TAr_{k-1}}\cdot\frac{r_{k-1}^Tr_{k-1}}{r_{k-1}^TA^{-1}r_{k-1}}\right]$$
$$<\varphi(x_{k-1})\left(1-\frac{1}{\|A\|_2\cdot\|A^{-1}\|_2}\right)=\left(1-\frac{1}{\kappa_2(A)}\right)\varphi(x_{k-1})$$ □

3. 证：设 $n+1$ 步终止，即 $x_{n+1}=x_*$，$Ax_*-b=0$.

由 $x_{n+1}=x_n+\alpha_nr_n$，得 $A(x_n+\alpha_nr_n)-b=0$.

而 $r_n=b-Ax_n$，故 $b=Ax_n+r_n$，从而：
$$A(x_n+\alpha_nr_n)-(Ax_n+r_n)=0$$
$$\Rightarrow\alpha_n(Ar_n)=r_n\quad\Rightarrow\quad Ar_n=\alpha_n^{-1}r_n$$ □

即 r_n 是A的特征向量，且 α_n^{-1} 为对应的特征值.

6. 证：记 $\Phi(t)=\varphi(y_{i-1}+te_i)$，由凸函数的性质，$\Phi(t)$取min $\Leftrightarrow$ $\Phi'(t)=0$.

$$\Phi'(t)=e_i^T\nabla\varphi(y_{i-1}+te_i)=2e_i^T(A(y_{i-1}+te_i)-b)=2(e_i^TAy_{i-1}+t-e_i^Tb)$$
$$\Phi'(t)=0\Leftrightarrow t=e_i^Tb-e_i^TAy_{i-1}=b_i-\sum_{j=1}^{i}a_{ij}y_j^{(k+1)}-\sum_{j=i+1}^{n}a_{ij}y_j^{(k)}$$

如此更新 $y_i=y_{i-1}+te_i$

与 G-S 迭代法中更新第k步的第i个分量格式相同.

故 x_k 就是G-S迭代第k步的结果，此方法与G-S法相同. □

7. 证：由实对称性，知存在以A的特征向量构成的 $\mathbb{R}^n$ 空间基底.

设 A 有特征值 $\lambda_1,\dots,\lambda_k$，代数重数为 $m_1,\dots,m_k$，λ_i 对应特征向量 $x_1^{(i)}\dots x_{m_i}^{(i)}$（单位正交）

则，$\forall r\in\mathbb{R}^n$，$\exists\ \alpha_{ij}$ $(1\le i\le k,\ 1\le j\le m_i)$，s.t.
$$r=\sum_{i=1}^{k}\sum_{j=1}^{m_i}\alpha_{ij}x_j^{(i)}$$

从而：
$$A^sr=\sum_{i=1}^{k}\sum_{j=1}^{m_i}\alpha_{ij}\lambda_i^sx_j^{(i)},\quad s=1,2,\dots,n-1$$

用归纳法证明 $\dim\operatorname{span}(r,Ar,\dots,A^{n-1}r)\le k$.

首先，若 $k=1$. 假定有 $\beta_1,\dots,\beta_n\in\mathbb{R}$，s.t.
$$\beta_1r+\beta_2Ar+\dots+\beta_nA^{n-1}r=0$$

i.e. $0=\sum_{s=1}^{n}\beta_s A^{s-1}r = \sum_{s=1}^{n}\beta_s\sum_{j=1}^{m_1}\alpha_{ij}\lambda_1^{s-1}x_j^{(1)}$

取 $j=1,\dots,m_1=n$，上式两边同乘 $x_j^{(1)T}$ 得：

$$0=\sum_{s=1}^{n}\beta_s\alpha_{ij}\lambda_1^{s-1}x_j^{(1)T}x_j^{(1)}=\sum_{s=1}^{n}\beta_s\alpha_{ij}\lambda_1^{s-1}$$

$$\Rightarrow \begin{pmatrix}\alpha_{11} & \alpha_{11}\lambda_1 & \cdots & \alpha_{11}\lambda_1^{n-1}\\ \alpha_{12} & \alpha_{12}\lambda_1 & \cdots & \alpha_{12}\lambda_1^{n-1}\\ \vdots & \vdots & & \vdots\\ \alpha_{1n} & \alpha_{1n}\lambda_1 & \cdots & \alpha_{1n}\lambda_1^{n-1}\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_n\end{pmatrix}=\begin{pmatrix}0\\0\\ \vdots\\0\end{pmatrix}$$

系数矩阵显然秩不大于1，故解空间维数不小于 $n-1$

这表明 $\dim\operatorname{span}(r, Ar, \dots, A^{n-1}r)\leq 1$.

现假设结论对 $1,2,\dots,k-1$ 已成立，考虑 k $(k\geq 2)$.

同样假定 $\beta_1,\dots,\beta_n\in\mathbb{R}$，使：

$$\beta_1 r+\beta_2 Ar+\cdots+\beta_n A^{n-1}r=0$$

由归纳假设，

$$\begin{pmatrix}A_1\\ \vdots\\ A_{k-1}\end{pmatrix}\begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}，其中 A_i=\begin{pmatrix}\alpha_{i1} & \alpha_{i1}\lambda_i & \cdots & \alpha_{i1}\lambda_i^{m_i-1}\\ \alpha_{i2} & \alpha_{i2}\lambda_i & \cdots & \alpha_{i2}\lambda_i^{m_i-1}\\ \vdots & \vdots & & \vdots\\ \alpha_{im_i} & \alpha_{im_i}\lambda_i & \cdots & \alpha_{im_i}\lambda_i^{m_i-1}\end{pmatrix}$$

总有 $\operatorname{rank}(A_i)\leq 1$，故解空间维数不小于 $n-(k-1)$. 而 $\operatorname{rank}(A_k)\leq 1$ 也成立

$$\Rightarrow \begin{pmatrix}A_1\\ \vdots\\ A_k\end{pmatrix}\begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$$ 解空间维数不小于 $n-k$

这表明 $\dim\operatorname{span}(r, Ar, \cdots, A^{n-1}r)\leq k$.

由归纳法即得证. □

8. 证：共轭梯度法产生的向量组 $\{r_0,\dots,r_k\}$ 是 Krylov 子空间：

$$\operatorname{span}\{r_0, Ar_0, \dots, A^{n-1}r_0\}$$

的正交基，而由 Ex7. 知 Krylov 子空间维数至多为 l（互异特征值的个数）

故 $k\leq l-1$，即最多迭代 l 步. □

1 上机题报告

1.1 CG 方法通用子程序设计

```
function [x,step] = CG(A,b,err)
    x = zeros(size(A,1),1);
    r = A*x-b;
    p = -r;
    step = 0;
    while vecnorm(r) >= err
        step = step + 1;
        alpha = - (r.'*p)/(p.'*A*p);
        x = x + alpha*p;
        tmp = r.'*r;
        r = r + alpha*A*p;
        beta = (r.'*r)/tmp;
        p = -r + beta*p;
    end
end
```

1.2 数值实验 - 上机习题 5.2

用 Hilbert 矩阵测试共轭梯度法：

$$a_{ij} = (i+j+1)^{-1}, \quad b_i = \frac{1}{3}\sum_{j=1}^{n} a_{ij} \qquad (1 \le i, j \le n)$$

我们给定 $N = 100$，对 $n = 1, 2, ..., N$ 阶的 Hilbert 矩阵，用 CG 方法求解线性方程组，停机准则设为：

$$||r_k||_2 < err = 10^{-14}$$

求解后将结果与精确解进行比较，绘制残差关于阶数的图线。同时，观察迭代次数，绘制迭代次数关于阶数的图线。代码如下（附件 ex_5_2.m）：

```
N = 100;
index = zeros(N,1);
err = zeros(N,1);
step = zeros(N,1);
for n = 1:N
    A = hilb(n);
    b = zeros(n,1);
    for i = 1:n
        b(i) = sum(A(i,1:n))/3;
    end
    index(i) = i;
    [x,step(i)] = CG(A,b,1e-14);
```

```
    err(i) = vecnorm(x-ones(n,1)/3);
end
subplot(1,2,1)
plot(index, err);
xlabel('(1) residual of n','fontname', 'Times new roman','interpreter', 'latex');
box off
subplot(1,2,2)
plot(index, step);
xlabel('(2) iterations of n','fontname', 'Times new roman','interpreter', 'latex');
box off
```

测试结果如下图所示：

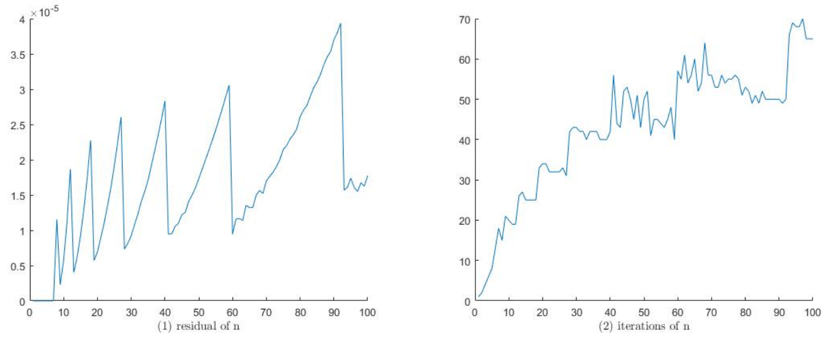

图 1: 用 CG 法解 Hilbert 矩阵方程，残差与迭代次数关于阶数 n 的图线

1.3 数值实验 - 上机习题 5.3

使用 Jacobi 迭代法、G-S 迭代法和共轭梯度法求解下述方程：

$$\begin{bmatrix} 10 & 1 & 2 & 3 & 4 \\ 1 & 9 & -1 & 2 & -3 \\ 2 & -1 & 7 & 3 & -5 \\ 3 & 2 & 3 & 12 & -1 \\ 4 & -3 & -5 & -1 & 15 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 12 \\ -27 \\ 14 \\ -17 \\ 12 \end{bmatrix}$$

容易知道上述方程的精确解为 $(1, -2, 3, -2, 1)^T$。现分别用三种迭代法求解，与精确解做误差比较，并输出迭代次数。测试用代码如下（附件 ex_5_3.m）：

```
A = [[10 1 2 3 4]; [1 9 -1 2 -3]; [2 -1 7 3 -5]; [3 2 3 12 -1]; [4 -3 -5 -1 15]];
b = [12; -27; 14; -17; 12];
s = [1;-2;3;-2;1];
eps = 1e-12;
[x1,step1] = jacobi(A,b,eps);
[x2,step2] = gauss_seidel(A,b,eps);
[x3,step3] = CG(A,b,eps);
solutions = [x1 x2 x3]
residures = [vecnorm(x1-s) vecnorm(x2-s) vecnorm(x3-s)]
iterations = [step1 step2 step3]
```

输出结果如下（从左到右分别为 Jacobi 迭代、G-S 迭代、CG 方法的结果）：

```
solutions =


          1.00000000000183          0.999999999999123                           1
         -2.00000000000168          -1.99999999999928                          -2
          2.99999999999692           3.00000000000125                           3
         -1.99999999999951          -2.00000000000012                          -2
         0.999999999997802           1.00000000000079                           1



residures =


      4.55356206907407e-12       1.86230250377141e-12       1.11022302462516e-15



iterations =


   138    80     5
```

可以看到，Jacobi 迭代和 G-S 迭代所花费的迭代次数远远高于共轭梯度法，而且所得解的精度也不如共轭梯度法。事实上，我们可以计算 Jacobi 迭代法和 G-S 迭代法的渐进收敛速度：

$$R_\infty(M^{(Jacobi)}) = -\ln\rho(M^{(Jacobi)}) = -\ln 0.826786 = 0.19$$

$$R_\infty(M^{(G-S)}) = -\ln\rho(M^{(G-S)}) = -\ln 0.684558 = 0.38$$

对于共轭梯度法我们有收敛性估计：

$$||x_* - x_k||_A \leq 2\left(\frac{\sqrt{\kappa_2}-1}{\sqrt{\kappa_2}+1}\right)^k ||x_* - x_0||_A$$

类似地我们可以计算：

$$-\ln\left(\frac{\sqrt{\kappa_2}-1}{\sqrt{\kappa_2}+1}\right) = -\ln 0.545826 = 0.605$$

这个值高于 Jacobi 迭代法和 G-S 迭代法的渐进收敛速度。而且由于上式对共轭梯度法的收敛速度给出的估计是十分粗糙的，共轭梯度法实际计算时，收敛速度还要比它快得多。所以共轭梯度法在三者的比较中占据绝对优势完全是预料之中的。