林子开 21307110161

1

1. Let $A \in \mathbb{C}^{m\times n}$ and $b \in \mathbb{C}^m$. It is known that $x_* = A^\dagger b$ is always a minimizer of the least squares problem $\min_x \|Ax-b\|_2$ even if $\operatorname{rank}(A) < n$. Show that for any other minimizer y, we have

$$\|y\|_2 \geq \|x_*\|_2.$$

解: $A = U\Sigma_* V^*$ 则 $A^+ = V\Sigma_*^{-1}U^*$

$$\Sigma_* = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{m\times n} \qquad \Sigma_*^{-1} = \begin{bmatrix} \sigma_1^{-1} & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_n^{-1} & 0 & \cdots & 0 \end{bmatrix}_{n\times m}$$

满足 $\min_x \|Ax-b\|_2$ 的 x' 都可表示为:

$$x' = x^* + x_0$$

其中, $Ax_0 = 0$ (齐次方程组的解)

当 $\operatorname{Rank}(A) < n$ 时, x_0 可以不等于 0

而 x^* 满足: $\min_x \|Ax^* - b\|_2$

相当于一个"特解"

现在再加一个限定, 让所有属于零空间的部分都包含到 x_0 去, 则剩下的 x^* 就属于 A 的行空间

由行空间与零空间正交, 这时的 $\|x^*\|_2$ 就是最小的.

下面说明 X^* 就是 $X_* = A^+ b$

上面已证明, X^* 属于A的行空间, 下面需说明

① X_* 满足 $\|AX_* - b\|_2$ 最小 or $A^*AX_* = A^*b$

② X_* 在A的行空间中, 没有零空间的分量

$$A^*AX_* = A^* \cdot A \cdot A^+ b \qquad (A = U\Sigma_* V^*)$$

$$= V\Sigma_*^* U^* U \Sigma_* V^* V \Sigma_*^{-1} U^* b$$

$$= V \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix} \cdot \left(\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_n^{-1} & \end{array} \right) U^* b$$

$$= V \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_n & \end{array} \right) U^* b \text{ —— (1)}$$

而 $A^*b = (U\Sigma V^*)^* b$

$$= V \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_n & \end{array} \right) U^* b \text{ —— (2)}$$

由 (1) = (2) 可知

X_* 满足法方程 $A^*AX_* = A^*b$

又由 $A = \begin{pmatrix} \text{—} v_1^T \text{—} \\ \vdots \\ \text{—} v_m^T \text{—} \end{pmatrix} = U \cdot \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline & 0 & \end{pmatrix} V^*$

$$A^+ = V\Sigma_*^{-1}U^* \quad 其中: \Sigma_*^{-1} = \begin{bmatrix} \sigma_1^{-1} & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_n^{-1} & 0 & \cdots & 0 \end{bmatrix}_{n\times m}$$

则 $A^+b = V\cdot(\Sigma_*^{-1}U^*b)$

$X_* = A^+b$ 是 V 的列向量的线性组合

也即是 V^* 的行向量的线性组合

由 $A = U\Sigma_* V^*$ 可知，X_* 必落在 A 的行空间内

故 $X_* = X^*$

则 $\forall y \in \{x' : \min \|Ax' - b\|_2\}$

有 $\|y\|_2 \geq \|X_*\|_2$

2

2. Let $A \in \mathbb{C}^{m\times n}$ and $X \in \mathbb{C}^{n\times m}$. Suppose that for any $b \in \mathbb{C}^m$, $x = Xb$ is always a minimizer of the least squares problem $\min_x \|Ax - b\|_2$. Show that $AXA = A$ and $(AX)^* = AX$.

解：由(1)题可知，X 可分解成两个部分

$X = A^+ + T_0$

其中 T_0 会把 b 投影到零空间里，相当于给 x 加上了一个齐次方程的非零解 x_0，而 $x = x^* + x_0$ 仍有：

$A^*A(x^* + x_0) - A^*b = 0$

① $AXA = A(A^+ + T_0)A$

$$= AA^+A + AT_0A$$

而 T_0 把 b 投影到零空间里，T_0 与 A 的行空间正交，因此 $AT_0 = 0$

又由伪逆定义： $A^+A = I$

$$\therefore AXA = A$$

② $A(A^+ + T_0)$

$$= I + AT_0$$

$$= I + 0$$

$$= I$$

$$\therefore (AX)^* = I^* = AX$$

3 **3.** Generate a few least squares problems with condition numbers varying from 10^0 to 10^{15}. Compare the accuracy of the solutions produced by the following methods:
(a) solve the normal equation $A^*Ax = A^*b$ through the Cholesky factorization of A^*A;
(b) solve the equation $Rx = Q^*b$ through the (Householder) QR factorization $A = QR$.

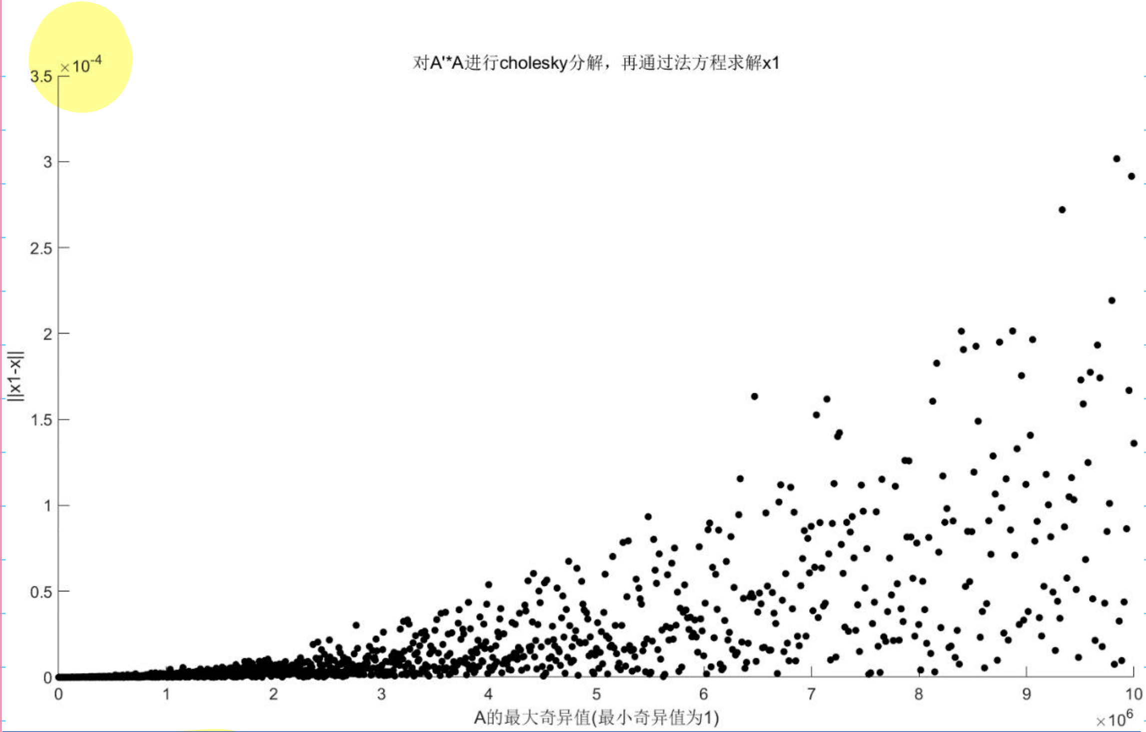
Matlab 代码实现如下：（后有图）

```
m=10;
n=4;
b=rand(m,1);
U =orth(rand(m,m));%生成一组标准正交基 u
V =orth(rand(n,n));%生成另一组标准正交基 v
sig=zeros(m,n);
sig(2,2)=10;%
sig(3,3)=5;
sig(4,4)=1;%最小的奇异值是 1
error1=zeros(1,((7-1)/0.001+1));
error2=zeros(1,((7-1)/0.001+1));
maxsigma=zeros(1,((7-1)/0.001+1));%初始化
%%
count=1;
for k= 1:0.001:7
    sig(1,1)=10^k;%最大的奇异值从 1 增加到 10^8，按照指数的方式增长
    A = U*sig*V';
    maxsigma(1,count)=10^k;
    %直接用 matlab 内部自带的方法计算伪逆，并认为这个解相对精准
    x=pinv(A)*b;

    %方法一，直接解法方程 normal  equation(Cholesky)
    R=chol(A'*A);
    x1=R\(R'\(A'*b));
    %[L,U]=lu(A'*A);
    %x1=U\(L\(A'*b));
    %x1=(A'*A)\(A'*b);
    error1(1,count)=norm(x1-x);

    %方法二，通过 Householder 方法进行 QR 分解
    %[Q,R]=qr(A);
    A2=A;
    Q_inv=eye(m)
    for j =1:n %使用 householder 方法
        v=A2(j:m,j);
        len=norm(v);
        if v(1,1)>=0%此处相当于对 v 与一个长度为||v||且只有第一个分量不为零的向量
作差
            v(1,1)=v(1,1)+len;%避免舍入误差
        else
            v(1,1)=v(1,1)-len;%避免舍入误差
        end
```

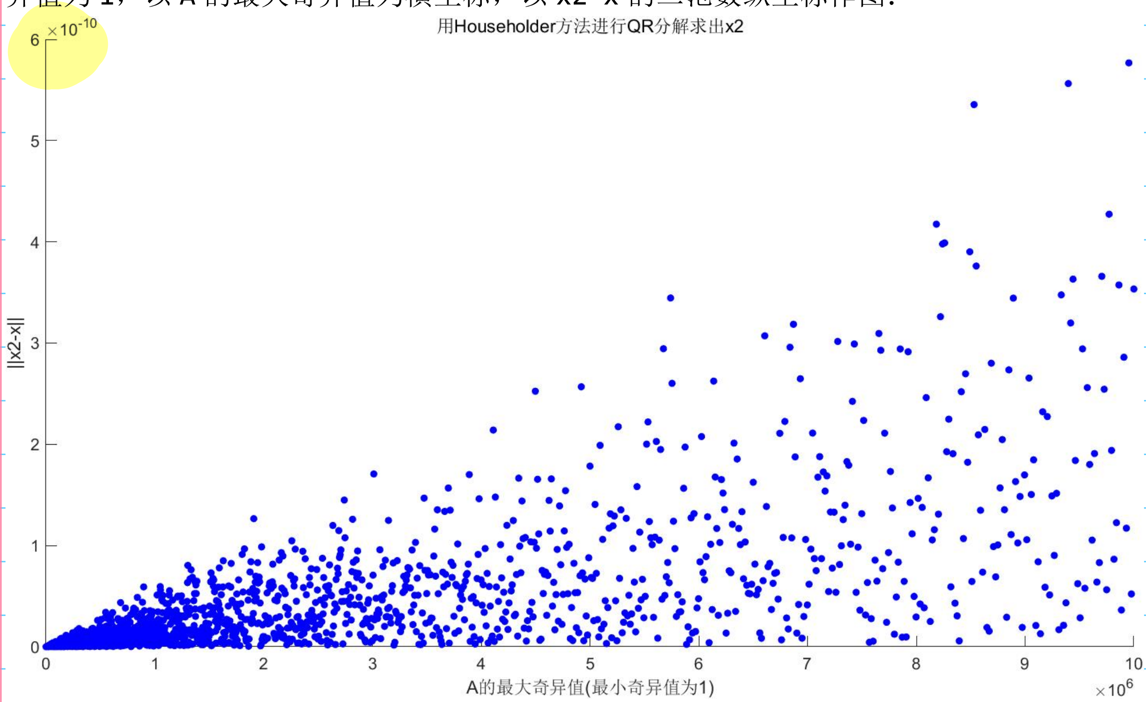
```
        H=eye(m-j+1)-2*(v*v')./(v'*v);
        A2(j:m,j:n)=H*A2(j:m,j:n);%Householder 反射子只需要作用于 schur 补即可
        H2=[eye(j-1),zeros(j-1,m-j+1);zeros(m-j+1,j-1),H];%分块矩阵，左上角为
eye，右下角为 H，其他为零
        Q_inv=H2*Q_inv;
    end
    R=A2;
    x2=R\(Q_inv*b);

    %计算 x2-x1 的范数 norm
    error2(1,count)=norm(x2-x);

    count=count+1;
end
figure(1);
scatter(maxsigma,error1,20,"black","filled");
xlabel("A 的最大奇异值(最小奇异值为 1)");
ylabel("||x1-x||");
title("对 A'*A 进行 cholesky 分解，再通过法方程求解 x1")
figure(2)
scatter(maxsigma,error2,20,"blue","filled");
xlabel("A 的最大奇异值(最小奇异值为 1)");
ylabel("||x2-x||");
title("用 Householder 方法进行 QR 分解求出 x2")
```

直接用 matlab 自带的方法计算伪逆左乘到 b 上，得到 x，认为这个 x 较为准确

通过 cholesky 分解的方法得到 x1,计算 x1-x 的二范数，并以此来衡量 x1 的误差。A 的最小奇异值为 1，以 A 的最大奇异值为横坐标，以 x1-x 的二范数纵坐标作图：



通过 QR 分解的方法得到 x2,计算 x2-x 的二范数，并以此来衡量 x2 的误差。A 的最小奇异值为 1，以 A 的最大奇异值为横坐标，以 x2-x 的二范数纵坐标作图：



当 A 的最大奇异值很大时，条件数也会相应地增大。但是，QR 分解的条件数增加的速度不如用法方程求解的条件数增加的速度快。从误差的数量级也可以看出，QR 分解的精确度更高，法方程求解的精确度较低。

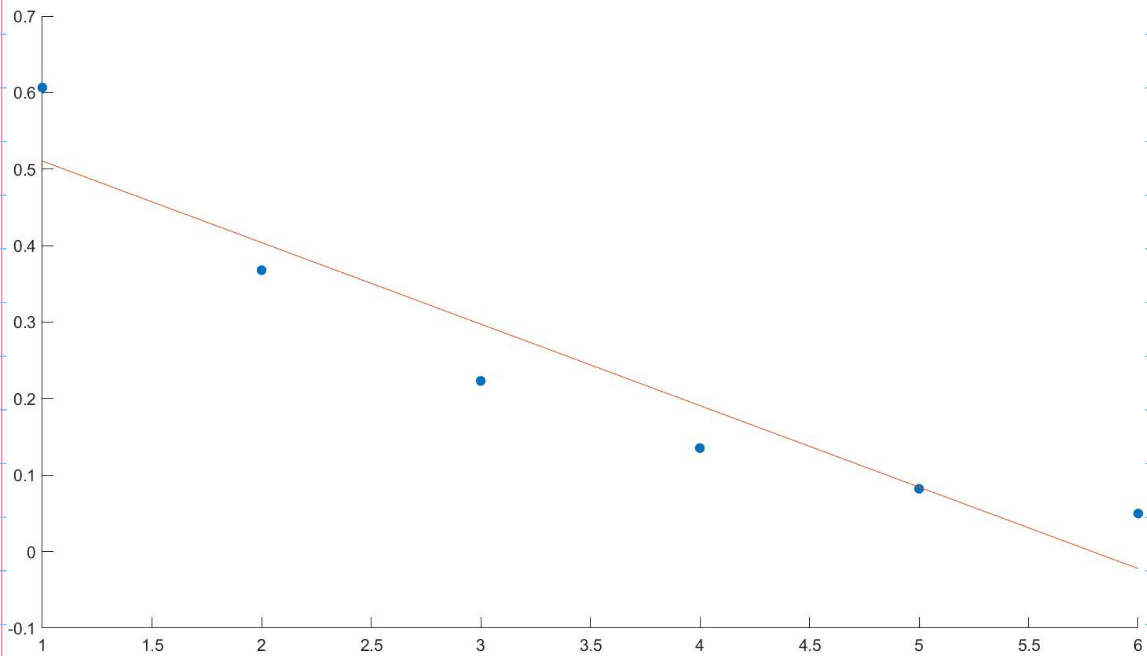
4

4. Find the "best" straight line that approximately pass through the data set $\{(n, e^{-n/2}) \in \mathbb{R}^2 : n \in \{1,2,3,4,5,6\}\}$. Visualize your result and clarify in what sense your solution is the best.

Matlab 代码实现如下:

```
n=1:6;
y=(exp(-n/2))';
%绘制散点图
scatter(n,y,"filled");
hold on
%用最小二乘法求斜率和截距
A=[n',ones(6,1)];
x=pinv(A)*y;
k=x(1,1);
b=x(2,1);
%绘制对应直线
Y=k*n+b;
plot(n,Y);
```

作图为:



说明: 对于 $(x_1, y_1) \cdots (x_6, y_6)$

要找到 k, b, 使得

$$\sum_{i=1}^{6} (y_i - (kx_i + b))^2 \text{ 最小}$$

问题可以转化为解超定方程组：

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \\ x_5 & 1 \\ x_6 & 1 \end{bmatrix} \begin{bmatrix} k \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix}$$

记为：$A\tilde{x} = \tilde{b}$

若 $\sum_{i=1}^{6}(y_i - (kx_i + b))^2$ 最小

$\Longleftrightarrow \quad \|A\tilde{x} - \tilde{b}\|_2$ 最小

则只要取：

$\tilde{x} = \begin{pmatrix} k \\ b \end{pmatrix} = A^+\tilde{b}$ 即可