作业6

(数值算法与案例分析)

李维杰

2024 年 10 月 27 日

题目1. 设有瘦高上双对角矩阵$A \in \mathbb{R}^{m\times n}$(即$a_{i,j} \neq 0$当且仅当$i-j \in \{0,-1\}$).设计一个基于Givens Rotations的算法解决岭回归问题

$$\min \|Ax-b\|_2^2 + \lambda\|x\|_2^2,$$

其中λ是一个给定的正数.

解答. 问题等价于标准最小二乘问题

$$\min_{Cx=d} \left\| \begin{bmatrix} A \\ \sqrt{\lambda} I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2.$$

此问题的法方程为

$$\begin{bmatrix} A \\ \sqrt{\lambda} I \end{bmatrix}^* \begin{bmatrix} A \\ \sqrt{\lambda} I \end{bmatrix} x = \begin{bmatrix} A \\ \sqrt{\lambda} I \end{bmatrix}^* \begin{bmatrix} b \\ 0 \end{bmatrix}.$$

即

$$(A^*A + \lambda I)x = A^*b.$$

其中,$A^*A + \lambda I$是一个三对角阵,故可以使用n次Givens Rotations,将$A^*A + \lambda I$变换为上双对角阵,也即上三角阵R. 同时,将这n次Givens Rotations作用于A^*b得到向量c,最后解线性方程组$Rx = c$即可.

题目2. 设计一个基于Gaussian elimination消去约束系统的算法解决带约束的最小二乘问题

$$\min_{Cx=d} \|Ax-b\|_2.$$

解答. 对$Cx=d$做Gaussian elimination,得

$$\begin{bmatrix} R & C_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = d'.$$

其中,$\begin{bmatrix} R & C_2 \end{bmatrix}$是一个梯形矩阵.于是解得

$$x_1 = R^{-1}(d' - C_2 x_2).$$

按$\begin{bmatrix} x_1 & x_2 \end{bmatrix}$的规模大小对$A$分块为$A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$,则

$$\begin{aligned} \min_{Cx=d} \|Ax-b\|_2 &= \min \|A_1 x_1 + A_2 x_2 - b\|_2 \\ &= \min \left\|(A_2 - A_1 R^{-1} C_2) x_2 - (b - A_1 R^{-1} d')\right\|_2. \end{aligned}$$

题目3. 最小二乘问题$\min \|Ax-b\|_2$等价于Hermite不定增广线性系统

$$\begin{bmatrix} I & A \\ A^* & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix},$$

对于带约束的最小二乘问题$\min\limits_{Cx=d} \|Ax-b\|_2$,同样存在类似的增广线性系统,请找出它.

解答. **算法1.** $\min\limits_{Cx=d} \|Ax-b\|_2$对应的增广线性系统为

$$\begin{bmatrix} A^*A & C^* \\ C & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ z \end{bmatrix} = \begin{bmatrix} A^*b \\ d \end{bmatrix}.$$

事实上,假设存在其它满足$Cx=d$的x,且有$x \neq \hat{x}$,下面证明

$$\|Ax-b\|_2 \geqslant \|A\hat{x}-b\|_2$$

即可.

由于$A^*A\hat{x} + C^*z = A^*b$,可知$A^*(A\hat{x}-b) = -C^*z$,故

$$\begin{aligned} 2(x-\hat{x})^* A^*(A\hat{x}-b) &= -2(x-\hat{x})^* C^* z \\ &= -2[C(x-\hat{x})]^* z \\ &= 0. \end{aligned}$$

于是

$$\begin{aligned}\|Ax-b\|_2^2 &= \|A(x-\hat{x})+A\hat{x}-b\|_2^2\\ &= \|A(x-\hat{x})\|_2^2+\|A\hat{x}-b\|_2^2+2(x-\hat{x})^*A^*(A\hat{x}-b)\\ &= \|A(x-\hat{x})\|_2^2+\|A\hat{x}-b\|_2^2\\ &\geqslant \|A\hat{x}-b\|_2^2\end{aligned}$$

当且仅当$x=\hat{x}$时取等.

算法2. 设Lagrange函数

$$L(x,\lambda)=\frac{1}{2}(\|Ax-b\|_2^2+\lambda\|Cx-d\|_2^2),$$

令

$$\begin{cases}\frac{\partial L}{\partial x}=A^*(A\hat{x}-b)+\lambda C^*(Cx-d)=0\\ C\hat{x}-d=0\end{cases},$$

则$\min\limits_{Cx=d}\|Ax-b\|_2$对应的增广线性系统为

$$\begin{bmatrix} I & A\\ A^* & 0\\ 0 & C\end{bmatrix}\begin{bmatrix} b-A\hat{x}\\ \hat{x}\end{bmatrix}=\begin{bmatrix} b\\ 0\\ d\end{bmatrix}.$$

题目4. 分别使用CGS/CGS2/MGS/MGS2算法进行Arnoldi过程并可视化正交性损失. 使用的样例可以是随机生成的1000×1000矩阵和30维的Krylov subspace.

解答. (代码见Problem4.m)

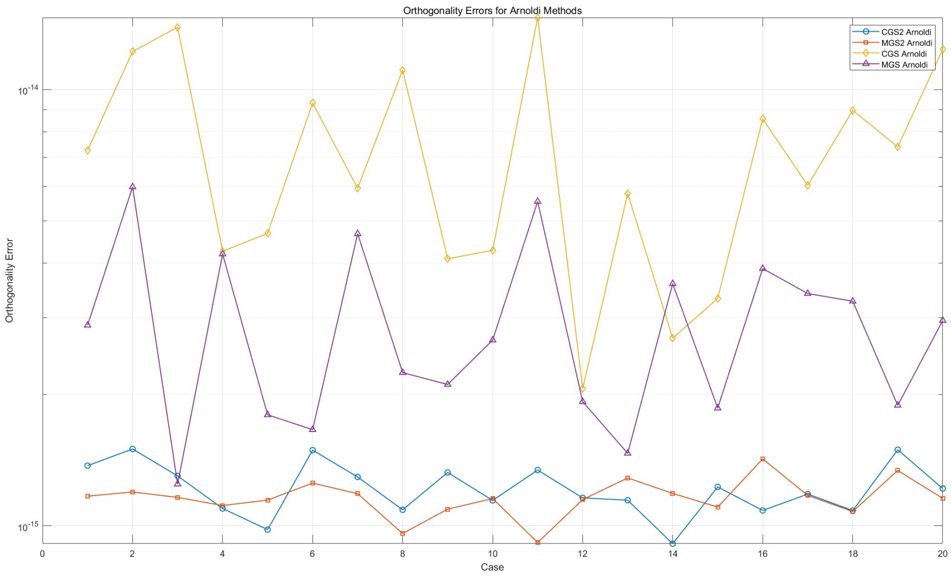

图 1: 四种算法在20种随机样例下的正交性损失$\|Q^*Q - I_n\|_\mathsf{F}$

题目5. 假设矩阵的秩已经给定,编写一个程序来解决亏秩最小二乘问题.

解答. 通过列选主元的QR分解求得亏秩条件下的最小二乘解为$\hat{x}$, 并设满秩条件下的最小二乘解为$\tilde{x} = (A^*A)^{-1}A^*b$. 对于随机生成的亏秩最小二乘问题$(A, b)$,两种解与观测值之间的差异度为

$$\|A\hat{x} - b\|_2 = \|A\tilde{x} - b\|_2 = 0.2723.$$

且$\hat{x}_i \neq 0$当且仅当$col(A)_i$是列空间$C(A)$的一个基.(代码见Problem5.m)