数据科学的数学基础
第 6 次理论作业

陈万祺　3220102895

2025 年 11 月 10 日

Problem 1

Given the following two problems:

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|_2^2 + s\|x\|_2^2 \quad (1)$$

and

$$\min_{x\in\mathbb{R}^m, \|x\|_2^2\le t} \|Ax-b\|_2^2 \quad (2).$$

Show that the two problems are equivalent in the sense that for any $s>0$, there exists $t>0$, such that the solution of problem (1) is also the solution of problem (2).

解： 设 x^* 为问题 (1) 的最优解。由一阶必要条件可得

$$\nabla\Big(\|Ax-b\|_2^2 + s\|x\|_2^2\Big)\Big|_{x=x^*} = 2A^\top(Ax^*-b) + 2sx^* = 0,$$

从而

$$(A^\top A + sI)x^* = A^\top b.$$

令 $t=\|x^*\|_2^2$，则 x^* 满足问题 (2) 的约束。为证 x^* 同时最优，假设存在 x' 满足 $\|x'\|_2^2\le t$ 且 $\|Ax'-b\|_2^2 < \|Ax^*-b\|_2^2$。
利用 $\|x'\|_2^2 \le t$ 得

$$\|Ax'-b\|_2^2 + s\|x'\|_2^2 \le \|Ax'-b\|_2^2 + st < \|Ax^*-b\|_2^2 + st = \|Ax^*-b\|_2^2 + s\|x^*\|_2^2,$$

这与 x^* 是问题 (1) 的最优解矛盾。因此，对任意 $s>0$ 取 $t=\|x^*\|_2^2$，问题 (1) 与问题 (2) 的解相同。

Problem 2

Apply the matching pursuit algorithm to solve the LASSO problem (only do two iterations):

$$\min_{x\in\mathbb{R}^3} \|Ax-b\|_2^2 + s\|x\|_1 \quad (3),$$

with $s=0.5$,

$$A = \begin{bmatrix} 2 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix}$$

and

$$b = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

解： 考虑 LASSO 目标

$$\min_{x\in\mathbb{R}^3} \|Ax-b\|_2^2 + s\|x\|_1, \quad s=0.5, \quad A = \begin{bmatrix} 2 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

记各列为 A_1, A_2, A_3。匹配追踪开启前，计算列范数

$$\|A_1\|_2^2 = 5, \qquad \|A_2\|_2^2 = 2, \qquad \|A_3\|_2^2 = 5,$$

以及软阈值参数

$$\lambda_j = \frac{s}{2\|A_j\|_2^2} \quad (j=1,2,3) \Rightarrow \lambda_1 = 0.05,\ \lambda_2 = 0.125,\ \lambda_3 = 0.05.$$

初始化 $x^{(0)}=0$，残差 $r^{(0)}=b$。以下给出两轮迭代：

1. 相关性

$$c = A^\top r^{(0)} = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}.$$

取绝对值最大的 c_1，得到临时解

$$a = \frac{c_1}{\|A_1\|_2^2} = 1, \qquad x_1^{(1)} = S_{\lambda_1}(a) = 0.95,$$

其中软阈值算子 $S_\lambda(a) = \operatorname{sign}(a)\max(|a|-\lambda, 0)$。更新向量与残差

$$x^{(1)} = \begin{bmatrix} 0.95 \\ 0 \\ 0 \end{bmatrix}, \qquad r^{(1)} = r^{(0)} - A_1(x_1^{(1)} - 0) = \begin{bmatrix} -0.90 \\ 2.05 \end{bmatrix}.$$

2. 再次计算相关性

$$c = A^\top r^{(1)} = \begin{bmatrix} 0.25 \\ -2.95 \\ 3.85 \end{bmatrix},$$

最大分量对应列 A_3。于是

$$a = \frac{c_3}{\|A_3\|_2^2} = 0.77, \qquad x_3^{(2)} = S_{\lambda_3}(a) = 0.72,$$

并更新

$$x^{(2)} = \begin{bmatrix} 0.95 \\ 0 \\ 0.72 \end{bmatrix}, \qquad r^{(2)} = r^{(1)} - A_3(x_3^{(2)} - 0) = \begin{bmatrix} 0.54 \\ 1.33 \end{bmatrix}.$$

完成两次迭代后，匹配追踪得到的系数向量为

$$x^{(2)} = [0.95,\ 0,\ 0.72]^\top.$$