第九次作业-solution

Lecturer: 杨启哲　　Last modified: 2024 年 5 月 14 日

1. 假设现在有 4 个点：$(0,0),(8,1),(8,3),(20,4)$:
   (1) 用最小二乘法拟合出一个线性函数 $f(x)=ax+b$，列出对应的方程组 $A\mathbf{x}=\mathbf{b}$ 并给出每个点上的误差。
   (2) 用最小二乘法拟合出一个线性函数 $f(x)=ax^2+bx+c$，列出对应的方程组 $A\mathbf{x}=\mathbf{b}$ 并且给出每个点上的误差。

解答.

- 对应的方程为：

$$\begin{bmatrix} 0 & 1 \\ 8 & 1 \\ 8 & 1 \\ 20 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix} (\text{记为} A\mathbf{x}=\mathbf{b})$$

注意到 $rank(A)=2$，从而 A^TA 是可逆的，从而我们有其最小二乘解为：

$$\begin{bmatrix} a \\ b \end{bmatrix} = (A^TA)^{-1}A^T\mathbf{b} = \frac{1}{51}\begin{bmatrix} 10 \\ 12 \end{bmatrix}$$

即对应的线性函数为 $f(x)=\frac{10}{51}x+\frac{12}{51}$，每个点上的误差为：

$$\begin{aligned} e_1 &= 0-\frac{12}{51} = -\frac{12}{51} \\ e_2 &= 1-\frac{10\cdot 8+12}{51} = -\frac{41}{51} \\ e_3 &= 3-\frac{10\cdot 8+12}{51} = \frac{61}{51} \\ e_4 &= 4-\frac{10\cdot 20+12}{51} = -\frac{8}{51} \end{aligned}$$

- 对应的方程为：

$$\begin{bmatrix} 0 & 0 & 1 \\ 64 & 8 & 1 \\ 64 & 8 & 1 \\ 400 & 20 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix} (\text{记为} A\mathbf{x}=\mathbf{b})$$

意到 $rank(A)=3$，从而 A^TA 是可逆的，从而我们有其最小二乘解为：

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = (A^TA)^{-1}A^T\mathbf{b} = \frac{1}{240}\begin{bmatrix} -1 \\ 68 \\ 0 \end{bmatrix}$$

即对应的二次函数为 $f(x)=-\frac{1}{240}x^2+\frac{68}{240}x$，每个点上的误差为：

$$
\begin{aligned}
e_1 &= 0\\
e_2 &= 1-\frac{-8^2+68*8}{240}=-1\\
e_3 &= 3-\frac{-8^2+68*8}{240}=1\\
e_4 &= 4-\frac{-20^2+68*20}{240}=0
\end{aligned}
$$

□

Remark 0.1

抱歉，这道题不小心设置了较大的计算量，大家可以借助一些计算矩阵的工具进行计算，比如一个在线的矩阵计算器 https://matrix.reshish.com/zh/。但是可以看到最后的结果确实是误差比较小的。关键是要能够理解最小二乘法的原理。

2. 通过 *Gram−Schmidt* 正交化得到下列向量张成空间的对应标准正交基：

$$
\mathbf{a}=\begin{bmatrix}1\\1\\2\end{bmatrix},\ \mathbf{b}=\begin{bmatrix}1\\-1\\0\end{bmatrix},\ \mathbf{c}=\begin{bmatrix}1\\0\\4\end{bmatrix}
$$

解答. 我们先得到一组正交基 $\mathbf{q}_1,\cdots,\mathbf{q}_3$：

(1) $\mathbf{q}_1=\mathbf{a}=\begin{bmatrix}1\\1\\2\end{bmatrix}$.

(2) $\mathbf{q}_2=\mathbf{b}-\frac{\mathbf{q}_1\mathbf{q}_1^T}{\mathbf{q}_1^T\mathbf{q}}\mathbf{b}=\begin{bmatrix}1\\-1\\0\end{bmatrix}$（因为 $\mathbf{a}\perp\mathbf{b}$）.

(3) $\mathbf{q}_3=\mathbf{c}-\frac{\mathbf{q}_1\mathbf{q}_1^T}{\mathbf{q}_1^T\mathbf{q}_1}\mathbf{c}-\frac{\mathbf{q}_2\mathbf{q}_2^T}{\mathbf{q}_2^T\mathbf{q}_2}\mathbf{c}=\begin{bmatrix}1\\0\\4\end{bmatrix}-\frac{3}{2}\begin{bmatrix}1\\1\\2\end{bmatrix}-\frac{1}{2}\begin{bmatrix}1\\-1\\0\end{bmatrix}=\begin{bmatrix}-1\\-1\\1\end{bmatrix}$.

最后将 $\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3$ 单位化便得到了一组标准正交基：

$$
\frac{\sqrt{6}}{6}\begin{bmatrix}1\\1\\2\end{bmatrix},\quad \frac{\sqrt{2}}{2}\begin{bmatrix}1\\-1\\0\end{bmatrix},\quad \frac{\sqrt{3}}{3}\begin{bmatrix}-1\\-1\\1\end{bmatrix}
$$

□

3. 考虑下列矩阵 Q:

$$
Q=c\begin{bmatrix}1&-1&-1&-1\\-1&1&-1&-1\\-1&-1&1&-1\\-1&-1&-1&1\end{bmatrix}
$$

(1) 求一个合适的 c 使得 Q 是一个正交矩阵。

(2) 求 $\mathbf{b}=(1,1,1,1)$ 到 Q 中每一个列向量上的投影 $\mathbf{p}_1,\cdots,\mathbf{p}_4$

(3) 求 $\mathbf{b}=(1,1,1,1)$ 到 Q 的列空间的投影 $\mathbf{p}$。

解答.

(1) 我们有:

$$Q^TQ=c^2\begin{bmatrix}4&0&0&0\\0&4&0&0\\0&0&4&0\\0&0&0&4\end{bmatrix}$$

从而 $c^2=\frac{1}{4}$，即令 $c=\pm\frac{1}{2}$，Q 是一个正交矩阵。

(2) 若 $c=0$，则我们有 $\mathbf{p}_1=\mathbf{p}_2=\mathbf{p}_3=\mathbf{p}_4=\mathbf{0}$，否则记 $\frac{1}{c}Q$ 的列向量分别为 $\mathbf{q}_1,\cdots,\mathbf{q}_4$，我们有:

$$\mathbf{p}_1=\frac{1}{4}\mathbf{q}_1\mathbf{q}_1^T\mathbf{b}=\frac{1}{2}\begin{bmatrix}-1\\1\\1\\1\end{bmatrix}$$

$$\mathbf{p}_2=\frac{1}{4}\mathbf{q}_2\mathbf{q}_2^T\mathbf{b}=\frac{1}{2}\begin{bmatrix}1\\-1\\1\\1\end{bmatrix}$$

$$\mathbf{p}_3=\frac{1}{4}\mathbf{q}_3\mathbf{q}_3^T\mathbf{b}=\frac{1}{2}\begin{bmatrix}1\\1\\-1\\1\end{bmatrix}$$

$$\mathbf{p}_4=\frac{1}{4}\mathbf{q}_4\mathbf{q}_4^T\mathbf{b}=\frac{1}{2}\begin{bmatrix}1\\1\\1\\-1\end{bmatrix}$$

(3) 若 $c=0$ 则我们有 $\mathbf{p}=\mathbf{0}$，否则注意到 Q 的列向量都是正交的，所以我们有:

$$\mathbf{p}=\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3+\mathbf{p}_4=\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

另一个解释在于 $\mathbf{b}\in C(Q)$。

□

4. 求出下列矩阵的 QR 分解:

$$A=\begin{bmatrix}1&2&4\\0&0&5\\0&3&6\end{bmatrix}$$

(hint: 先通过 $Gram\text{–}Schmidt$ 正交化得到一组标准正交的向量)

解答. 我们将其列向量 $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}2\\0\\3\end{bmatrix}, \begin{bmatrix}4\\5\\6\end{bmatrix}$ 进行 $Gram-Schmidt$ 正交化:

- $\mathbf{q}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$.
- $\mathbf{q}_2 = \begin{bmatrix}2\\0\\3\end{bmatrix} - \frac{\mathbf{q}_1\mathbf{q}_1^T}{\mathbf{q}_1^T\mathbf{q}_1}\begin{bmatrix}2\\0\\3\end{bmatrix} = \begin{bmatrix}0\\0\\3\end{bmatrix}$.
- $\mathbf{q}_3 = \begin{bmatrix}4\\5\\6\end{bmatrix} - \frac{\mathbf{q}_1\mathbf{q}_1^T}{\mathbf{q}_1^T\mathbf{q}_1}\begin{bmatrix}4\\5\\6\end{bmatrix} - \frac{\mathbf{q}_2\mathbf{q}_2^T}{\mathbf{q}_2^T\mathbf{q}_2}\begin{bmatrix}4\\5\\6\end{bmatrix} = \begin{bmatrix}0\\5\\0\end{bmatrix}$.

并将其单位化可得一组标准正交基:

$$\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}$$

从而我们可以得到其 QR 分解:

$$A = \begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&2&4\\0&3&6\\0&0&5\end{bmatrix}$$

□

5. 假设 P 是一个投影矩阵, 即 $P = A(A^TA)^{-1}A^T$:

(1) 证明 $P^2 = P$, 并且解释为什么 $P(P\mathbf{b})$ 等于 $P\mathbf{b}$。

(2) 证明 P 是对称矩阵, 即 $P^T = P$。

(3) 证明, 对任一矩阵 C, 如果其满足 $C^2 = C$, $C^T = C$, 则 C 是一个投影矩阵, 即证明对任意的 $\mathbf{b}, \mathbf{c}$ 有 $(\mathbf{b} - C\mathbf{b}) \perp C\mathbf{c}$, 从而这可以成为投影矩阵的另一个定义。

(4) 证明, 给定投影矩阵 P_1, P_2, $P_1P_2 = P_2P_1$ 当且仅当 P_1P_2 是一个投影矩阵。(利用上一问的结论)

解答.

(1) 我们有:

$$P^2 = A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T = P$$

$P(P\mathbf{b}) = P\mathbf{b}$ 是显然的, 这是因为 $P(P\mathbf{b})$ 是 $P\mathbf{b}$ 在 $C(A)$ 上的投影, 而 $P\mathbf{b}$ 是 $\mathbf{b}$ 在 $C(A)$ 上的投影, 从而 $P\mathbf{b} \in C(A)$, 所以 $P(P\mathbf{b}) = P\mathbf{b}$。

(2) 我们有:

$$P^T = (A(A^TA)^{-1}A^T)^T = (A^T)^T((A^TA)^{-1})^TA^T = A(A^TA)^{-1}A^T = P$$

(3) 对任意的 $\mathbf{b}, \mathbf{c}$ 我们有:

$$(\mathbf{b} - C\mathbf{b}) \cdot C\mathbf{c} = (\mathbf{b} - C\mathbf{b})^TC\mathbf{c} = \mathbf{b}^TC\mathbf{c} - \mathbf{b}^TC^TC\mathbf{c} = \mathbf{b}^TC\mathbf{c} - \mathbf{b}^TC^2\mathbf{c} = \mathbf{b}^TC\mathbf{c} - \mathbf{b}^TC\mathbf{c} = 0$$

即 $(\mathbf{b} - C\mathbf{b}) \perp C\mathbf{c}$, 从而 C 是一个将任一向量 $\mathbf{b}$ 投影到 $C(C)$ 上的投影矩阵。

(4) 若 $P_1P_2 = P_2P_1$，我们有：

$$
\begin{aligned}
(P_1P_2)^T &= P_2^T P_1^T = P_2P_1 = P_1P_2 \\
(P_1P_2)^2 &= P_1P_2P_1P_2 = P_1P_1P_2P_2 = P_1P_2
\end{aligned}
$$

从而 P_1P_2 是一个投影矩阵。
若 P_1P_2 是一个投影矩阵，则：

$$P_1P_2 = (P_1P_2)^T = P_2^T P_1^T = P_2P_1$$

□

6. 证明可逆矩阵 A 的 QR 分解是唯一的，这里 Q 是一个正交矩阵，R 是一个正对角元的上三角矩阵。

解答. 反设其中存在两个 QR 分解 $A = Q_1R_1 = Q_2R_2$，其中 Q_1, Q_2 是正交矩阵，R_1, R_2 是上三角矩阵，我们有：

$$Q_2^{-1}Q_1 = R_2R_1^{-1}$$

这意味着 $R' = R_2R_1^{-1}$ 是一个正交矩阵并且是一个上三角矩阵。下用数学归纳法证明 R' 的第 i 列是 $\mathbf{e} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0. \end{bmatrix}$

- 当 $i = 1$ 时，由上三角矩阵的定义及 R 的每一列都是单位向量可得 R' 的第一列是 $\mathbf{e}_1$，显然成立。
- 假设当 $i = k$ 时成立，即 R' 的第 k 列是 $\mathbf{e}_k$，则考虑 R' 的第 $k+1$ 列 $\mathbf{r}_{k+1}$，则我们有：
  (1) 对于 $i > k+1$，$\mathbf{r}_{k+1}(i) = 0$.
  (2) 对于 $i < k+1$，注意到 R' 是正交矩阵，从而第 $k+1$ 列和第 i 列是正交的，从而 $\mathbf{r}_{k+1}(i) = 0$.
  (3) 对于 $i = k+1$，注意到 R' 是一个正对角元的上三角矩阵，从而 $\mathbf{r}_{k+1}(k+1) = 1$.

□

Remark 0.2

这里抱歉，课上一开始的版本有些许问题，可逆矩阵的 QR 分解的上三角矩阵要求对角元是正的。并且我们利用到了如下两个结论：

- 任意正交矩阵的乘积还是正交矩阵，
- 任意上三角矩阵的乘积还是上三角矩阵（包括其逆）。