

数据科学与工程数学基础

作业提交规范及第 2 次作业

教师：黄定江

助教：陈诺、刘文辉

2022 年 9 月 30 日

作业提交规范

1. 作业提交形式：使用 Word 或 \LaTeX 编写所得到的电子文档。若使用 Word 编写，将其另存为 PDF 形式，然后提交 PDF 文档。若使用 \LaTeX 编写，将其编译成 PDF 形式，然后提交 Tex 和 PDF 两个文档。
2. 作业命名规范：提交的电子文档必须命名为：“学号_姓名”。命名示例：52200000000_刘某某。
3. 作业提交途径：点击打开每次作业的传送门网址：[第 2 次作业提交传送门](#)，无需注册和登录，直接上传作业文档即可。注意：传送门将会在截至时间点到达后自动关闭。
4. 作业更改说明：如果需要修改已经提交的作业，只要在截至日期前，再次上传更改后的作业（切记保持同名），即可覆盖已有作业。
5. 作业评分说明：正常提交作业的按照实际评分记录；逾期补交作业的根据逾期情况在实际评分基础上酌情扣分；未交作业的当次作业记为 0 分。

第 2 次作业



提交截至时间：**2022/10/14 周五 12:00 (中午)**

理论部分 (范数与二次型)

习题 1. 假设 $M, P \in \mathbb{R}^{n \times n}$ 为对称阵, P 为正交阵,

$$A = \left(\begin{array}{c|c} M & PM \\ \hline MP & PMP \end{array} \right) \in \mathbb{R}^{2n \times 2n}.$$

(i) 证明 $A^T = A$.

(ii) 假设 $U \in \mathbb{R}^{m \times m}, V \in \mathbb{R}^{n \times n}$ 是正交矩阵, $D \in \mathbb{R}^{m \times n}$, 证明 $\|UDV\|_2 = \|D\|_2, \|UDV\|_F = \|D\|_F$.

(iii) 证明 $\|A\|_F = 2\|M\|_F, \|A\|_2 \leq 2\|M\|_2$. (提示: 将 A 分解, 并利用 (ii) 结论)

(iv) 假设 $n = 4, M = \text{diag}_{4 \times 4}(-2, 1, 0, 0), P = (e_4 | e_3 | e_2 | e_1)$. 证明 $\|A\|_p = 2 \forall p \in [1, \infty)$.

解. (i) 解法 1: 分解 A

$$A = \begin{pmatrix} I \\ P \end{pmatrix} M \begin{pmatrix} I & P \end{pmatrix} = \begin{pmatrix} I \\ P \end{pmatrix} \begin{pmatrix} M & M \\ M & M \end{pmatrix} \begin{pmatrix} I & P \end{pmatrix} = A^T$$

解法 2: 不分解 A

$$A^T = \left(\begin{array}{c|c} M^T & (MP)^T \\ \hline (PM)^T & (PMP)^T \end{array} \right) = \left(\begin{array}{c|c} M^T & P^T M^T \\ \hline M^T P^T & P^T M^T P^T \end{array} \right) = \left(\begin{array}{c|c} M & PM \\ \hline MP & PMP \end{array} \right) = A$$

(ii) 即证明 2 范数与 F 范数满足正交不变性

对 2 范数, 即证 $\|UD\|_2 = \|D\|_2, \|DV\|_2 = \|D\|_2$

$$\|UD\|_2 = \sqrt{\lambda_{\max}(D^T U^T U D)} = \sqrt{\lambda_{\max}(D^T D)} = \|D\|_2$$

$$\text{又因为 } \|DV\|_2 = x^T V^T V x = x^T x = \|x\|_2$$

$$\|DV\|_2 = \sup_{\|x\|_2=1} \|DVx\|_2 = \sup_{\|Vx\|_2=1} \|DVx\|_2 = \|D\|_2$$

对 F 范数,

$$\|UDV\|_F = \sqrt{\text{tr}(V^T D^T U^T U D V)} = \sqrt{\text{tr}(V V^T D^T D)} = \sqrt{\text{tr}(V^T V D^T D)} = \sqrt{\text{tr}(D^T D)} = \|D\|_F$$

(iii) 解法 1: 分解 A , 由于 $\begin{pmatrix} I \\ P \end{pmatrix}$ 为正交阵, 由 (ii) 得

$$\|A\|_F = \left\| \begin{pmatrix} M & M \\ M & M \end{pmatrix} \right\|_F, \quad \|A\|_2 = \left\| \begin{pmatrix} M & M \\ M & M \end{pmatrix} \right\|_2$$

$$\|A\|_F^2 = \|M\|_F^2 + \|M\|_F^2 + \|M\|_F^2 + \|M\|_F^2 = 4\|M\|_F^2$$

或

$$\begin{aligned}
\|A\|_F &= \left\| \begin{pmatrix} M & M \\ M & M \end{pmatrix} \right\|_F = \left\| \begin{pmatrix} I \\ I \end{pmatrix} M \begin{pmatrix} I & I \end{pmatrix} \right\|_F \\
&= \sqrt{\text{tr} \left(\begin{pmatrix} I \\ I \end{pmatrix} M \begin{pmatrix} I & I \end{pmatrix} \begin{pmatrix} I \\ I \end{pmatrix} M \begin{pmatrix} I & I \end{pmatrix} \right)} \\
&= \sqrt{2 \text{tr} \left(M \begin{pmatrix} I & I \end{pmatrix} \begin{pmatrix} I \\ I \end{pmatrix} M \right)} \\
&= 2\sqrt{\text{tr}(M^2)} = 2\|M\|_F
\end{aligned}$$

$$\text{记 } B = \begin{pmatrix} M & M \\ M & M \end{pmatrix}, w = \begin{pmatrix} x \\ y \end{pmatrix}, \|A\|_2 = \|B\|_2 = \sup_{\|w\|_2 \neq 0} \frac{\|Bw\|_2}{\|w\|_2} = \sup_{\|w\|_2=1} \|Bw\|_2$$

$$\begin{aligned}
\left\| B \begin{pmatrix} x \\ y \end{pmatrix} \right\|_2^2 &= \left\| \begin{pmatrix} Mx + My \\ Mx + My \end{pmatrix} \right\|_2^2 \\
&= 2\|Mx + My\|_2^2 \\
&\leq 2(\|Mx\|_2 + \|My\|_2)^2 \\
&\leq 4\|Mx\|_2^2 + 4\|My\|_2^2 \\
&\leq 4\|M\|_2^2\|x\|_2^2 + 4\|M\|_2^2\|y\|_2^2 \text{ (相容性)} \\
&= 4\|M\|_2^2\|w\|_2^2 = 4\|M\|_2^2
\end{aligned}$$

因此 $\|A\|_2 = \sup_{\|w\|_2=1} \|Bw\|_2 = 2\|M\|_2$

解法 2: 不分解 A

$$\begin{aligned}
\|A\|_F^2 &= \|M\|_F^2 + \|MP\|_F^2 + \|PM\|_F^2 + \|PMP\|_F^2 = 4\|M\|_F^2 \\
\|Aw\|_2^2 &= \left\| \begin{pmatrix} Mx + PMy \\ MPx + PMPy \end{pmatrix} \right\|_2^2 \\
&= \|Mx + PMy\|_2^2 + \|MPx + PMPy\|_2^2 \\
&\leq (\|Mx\|_2 + \|PMY\|_2)^2 + (\|MPx\|_2 + \|PMPy\|_2)^2 \\
&\leq 2\|Mx\|_2^2 + 2\|PMY\|_2^2 + 2\|MPx\|_2^2 + 2\|PMPy\|_2^2 \\
&\leq 2\|M\|_2^2\|x\|_2^2 + 2\|PM\|_2^2\|y\|_2^2 + 2\|MP\|_2^2\|x\|_2^2 + 2\|PMP\|_2^2\|y\|_2^2 \\
&= 2\|M\|_2^2\|x\|_2^2 + 2\|M\|_2^2\|y\|_2^2 + 2\|M\|_2^2\|x\|_2^2 + 2\|M\|_2^2\|y\|_2^2 \\
&= 4\|M\|_2^2(\|x\|_2^2 + \|y\|_2^2) \\
&= 4\|M\|_2^2\|w\|_2^2
\end{aligned}$$

(iv)

$$M = \begin{pmatrix} -2 & & & & & & & \\ & 1 & & & & & & \\ & & 0 & & & & & \\ & & & 0 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 0 \end{pmatrix} \quad P = \begin{pmatrix} & & & & & & 1 & \\ & & & & & & & 1 \\ & & & & & & & & 1 \\ & & & & & & & & & 1 \\ & & & & & & & & & & 1 \\ & & & & & & & & & & & 1 \\ & & & & & & & & & & & & 1 \\ 1 & & & & & & & & & & & & & \end{pmatrix} \quad A = \left(\begin{array}{cccc|cccc} -2 & & & & & & & 0 \\ & 1 & & & & & & \\ & & 0 & & & & 1 & \\ & & & 0 & & & & -2 \\ \hline & & & & -2 & & 0 & \\ & & & & & 1 & & 0 \\ & & & & & & 0 & 1 \\ & & & & & & & & 1 & -2 \end{array} \right)$$

$$\begin{aligned} \|Ax\|_p^p &= \left\| (-2x_1, x_2, x_6, -2x_5, -2x_4, x_3, x_7, -2x_8)^T \right\|_p^p \\ &= |-2x_1|^p + |x_2|^p + |x_6|^p + |-2x_5|^p + |-2x_4|^p + |x_3|^p + |x_7|^p + |-2x_8|^p \\ &= 2^p |x_1|^p + |x_2|^p + |x_3|^p + 2^p |x_4|^p + 2^p |x_5|^p + |x_6|^p + |x_7|^p + 2^p |x_8|^p \\ &\leq 2^p (|x_1|^p + |x_2|^p + |x_3|^p + |x_4|^p + |x_5|^p + |x_6|^p + |x_7|^p + |x_8|^p) \\ &= 2^p \|x\|_p^p. \end{aligned}$$

因此 $\|A\|_p = \sup_{\|x\|_p=1} \|Ax\|_p = 2$

(即如果一个变换只将某些维度倍乘并交换顺序, 不作维度间相加的操作, 那矩阵范数即最大拉伸倍数)

习题 2. 假设 $P \in \mathbb{R}^{n \times n} \setminus \{0\}$ 是一个投影矩阵.

(i) 证明 $P y = y \ \forall y \in \mathcal{R}(P)$. $P x - x \in \mathcal{N}(P) \ \forall x \in \mathbb{R}^n$.

(即证明投影 P 沿着零空间 $\mathcal{N}(P)$ 投影到列空间 $\mathcal{R}(P)$)

(ii) 证明 P 的特征值 $\lambda \in \Lambda(P) \subseteq \{0, 1\}$. 假设 $\mathcal{R}(P) = \text{span}(u_1, \dots, u_r)$, $\mathcal{N}(P) = \text{span}(v_{r+1}, \dots, v_n)$, 试找到 P 的特征分解 $P = X D X^{-1}$ 并证明 $\text{tr}(P) = \text{rank}(P)$. (提示: 利用 (i) 结论.)

(iii) 证明当 $P \neq I_n$, $\det(P) = 0$.

(iv) 证明当 P 是正交投影矩阵 ($P^2 = P = P^T$) 时, $I_n - 2P$ 是正交矩阵.

(v) 假设 $A \in \mathbb{R}^{n \times m}$, $m \leq n$, $\text{rank}(A) = m$. $P = A(A^T A)^{-1} A^T$ 证明 P 是正交投影矩阵, $\text{rank}(P) = m$. (提示: 利用 (ii) 结论.)

解. (i) $\forall y \in \mathcal{R}(P)$ 即对 $x \in \mathbb{R}^n$, $y = Px$, $P y = P^2 x = P x = y$.

$\forall x \in \mathbb{R}^n$, $P(Px - x) = P^2 x - Px = Px - Px = 0$.

(ii) 对 $\lambda \in \Lambda(P)$, $x \in \mathbb{R}^n \setminus \{0\}$, 有 $Px = \lambda x$. 由于 $P = P^2$, $\lambda x = Px = P(Px) = P(\lambda x) = \lambda Px = \lambda^2 x$. 因为 $x \neq 0 \in \mathbb{R}^n$, 故 $\lambda = \lambda^2$, $\lambda \in \{0, 1\}$. 因此 $\Lambda(P) \subseteq \{0, 1\}$.

由 (i) 可知, $\forall i = 1, \dots, r$, $u_i \in \mathcal{R}(P)$, $P u_i = u_i$. $\forall j = r+1, \dots, n$, $v_j \in \mathcal{N}(P)$, $P v_j = 0$.

故令 $X := (u_1 | \dots | u_r | v_{r+1} | \dots | v_n) \in \mathbb{R}^{n \times n}$, $D := \text{diag}_{n \times n}(\underbrace{1, \dots, 1}_{r \text{ times}}, 0, \dots, 0) \in \mathbb{R}^{n \times n}$, 此时

$$P = XDX^{-1}$$

(注：也可理解为 SVD (后续课程会讲)，即 $P = UDV^T, U \in \mathcal{R}(P), V \in \mathcal{N}(P)$)

$$\text{tr}(P) = \text{tr}(XDX^{-1}) = \text{tr}(D) = r.$$

(iii) 反证 $\det(P) \neq 0 \implies P = I_n$. 由于 $\det(P) \neq 0$, P 可逆. 故由 $P^2 = P$, 得 $P^{-1}P^2 = P^{-1}P$, $P = I_n$.

(iv) 由于 P 是正交投影矩阵, $P^2 = P = P^T$. 令 $Q := I_n - 2P, Q^T = I_n - 2P^T = Q, Q^2 = I_n - 4P + 4P^2 = I_n$. 因此, $Q^T Q = Q Q^T = I_n$.

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

$$P^T = A \left((A^T A)^{-1} \right)^T A^T = A \left((A^T A)^T \right)^{-1} A^T = A(A^T A)^{-1} A^T = P.$$

由 (ii), $\text{rank}(P) = \text{tr}(P) = \text{tr}(A(A^T A)^{-1} A^T) = \text{tr}((A^T A)^{-1} A^T A) = \text{tr}(I_m) = m$.