线性代数 I 习题课

https://tinyurl.com/la1-pku23f

T.-Y. Li

目录

0 代数系统

为了进行数学操作, 往往要指明运算的环境和规则. 譬如求解方程 $2x = 3$ 中的未知量 x 时, 人们无法在整数集 $\mathbb{Z}$ 中得到解, 只好扩充数系并在有理数集 $\mathbb{Q}$ 中声称有解 $x = \frac{3}{2}$.

除非特殊说明, 我们总是考虑能够进行加减乘除的集合 $\mathbb{F}$, 例如实数集 $\mathbb{R}$ 和复数集 $\mathbb{C}$.

下面罗列一些严格定义, 初学者不妨略过. 注意 s.t. 是 such that 的缩写, id 是恒同映射.

如果集合 G 上的二元 (封闭) 运算 $\star$ 满足

- 结合律: $(g_1 \star g_2) \star g_3 = g_1 \star (g_2 \star g_3),\ \forall g_1, g_2, g_3 \in G,$
- 存在单位元: $\exists e \in G$ s.t. $e \star (\cdot) = (\cdot) \star e = \mathrm{id}_G,$
- 存在逆元: $\forall g \in G,\ \exists h \in G$ s.t. $g \star h = h \star g = e,$

则称 $(G, \star)$ 是群; 易得单位元唯一性, 从而任意元素的逆元唯一.

如果群 $(G, \boxplus)$ 满足

- “加法” 交换律: $g \boxplus h = h \boxplus g,\ \forall g, h \in G,$

则称 $(G, \boxplus)$ 是交换群, 其单位元记为 $0_\boxplus$.

如果交换群 $(R, \boxplus)$ 上的二元 (封闭) 运算 $\boxtimes$ 满足 $\forall r_1, r_2, r_3 \in R,$

- 结合律: $(r_1 \boxtimes r_2) \boxtimes r_3 = r_1 \boxtimes (r_2 \boxtimes r_3),$
- 分配律: $r_1 \boxtimes (r_2 \boxplus r_3) = (r_1 \boxtimes r_2) \boxplus (r_1 \boxtimes r_3),\ (r_1 \boxplus r_2) \boxtimes r_3 = (r_1 \boxtimes r_3) \boxplus (r_2 \boxtimes r_3),$

则称 $(R, \boxplus, \boxtimes)$ 是环.

如果环 $(R, \boxplus, \boxtimes)$ 满足

- 存在 “乘法” 单位元: $\exists 1_\boxtimes \in R$ s.t. $1_\boxtimes \boxtimes (\cdot) = (\cdot) \boxtimes 1_\boxtimes = \mathrm{id}_R,$

则称 $(R, \boxplus, \boxtimes)$ 是幺环; 易得 “乘法” 单位元 $1_\boxtimes$ 唯一, 且当 $R \neq \{0_\boxplus\}$ 时 $1_\boxtimes \neq 0_\boxplus$.

如果环 $(R, \boxplus, \boxdot)$ 满足

- “乘法” 交换律: $r \boxdot s = s \boxdot r,\ \forall r, s \in R,$

则称 $(R, \boxplus, \boxdot)$ 是交换环.

如果交换环 $(F, \boxplus, \boxdot)$ 中 $(F \setminus \{0_\boxplus\}, \boxdot)$ 是 (交换) 群 (存在 “乘法” 逆), 则称 $(F, \boxplus, \boxdot)$ 是域; 常用 $(\mathbb{F}, +, \cdot)$ 代表域. 如果 $1 \in \mathbb{K} \subset \mathbb{C}$, 且 $\mathbb{K}$ 配备通常的加法和乘法后成为域, 则称 $\mathbb{K}$ 是数域.

1 线性方程组

给定 a_{ij} ($i=1,\dots,m$; $j=1,\dots,n$) 和 b_i ($i=1,\dots,m$) 后, 形如

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ & & & & \vdots & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{cases}$$

的关于未定元 x_j ($j=1,\dots,n$) 的方程组称为线性方程组 (system of linear equations).

具体求解时只需处理相应的增广矩阵 (augmented matrix)

$$\begin{pmatrix} A & b \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

可用的操作是下述初等行变换 (elementary row operations):

1. (swap) 两行交换;
2. (scale) 一行乘上非零常数;
3. (shear) 一行加上另一行的非零常数倍.

执行 Gauss–Jordan 消元 (elimination) 后, 系数矩阵 A 变为简化行阶梯形 (reduced row echelon form), 满足

1. 只含 0 的行在最下方;
2. 每一行从左到右第一个非零元 (称为主元, pivot) 取值为 1;
3. 主元所在列的其他元素均为 0;
4. 主元的列指标随行指标递增.

由此即可分析线性方程组解的情况.

考虑线性方程组 $Ax=b$ 中 b 为零向量的情形, 此时 $Ax=0$ 称为齐次 (homogeneous) 线性方程组. 由于零向量在初等行变换后仍为零向量, 增广矩阵中只有系数矩阵 A 是本质的. 如果 A 的主元个数恰为 A 的列数, 则方程只有零解 $x=0$, 否则存在自由变量从而方程有无穷多个解.

题 1. *(完全图的二部分解[i]) 记 $[n]=\{1,2,\dots,n\}$. 求满足下述性质的最小正整数 m: 存在 $[n]$ 的子集 $A_1,B_1,\dots,A_m,B_m$ 使得*

$$([n]\times[n])\setminus\{(i,i): i\in[n]\} = \bigsqcup_{k=1}^{m}\Big((A_k\times B_k)\sqcup(B_k\times A_k)\Big),$$

其中 $\sqcup$ 表示无交并.

[i] Jukna, S. (2011). *Extremal combinatorics: with applications in computer science*, 2nd ed., Theorem 13.7.

解. 断言所求为 $n-1$. 注意 $(\tilde{A}_k=\{k\},\ \tilde{B}_k=\{k+1,\ldots,n\})_{1\le k\le n-1}$ 满足条件. 若 $(A_k,B_k)_{1\le k\le m}$ 满足条件, 则 $\forall x=(x_i)_{i\in[n]}\in\mathbb{R}^n$,

$$\left(e_{[n]}x\right)^2-\sum_i x_i^2=\sum_{i\neq j}x_ix_j=\sum_k\left(\sum_{i\in A_k}\sum_{j\in B_k}+\sum_{i\in B_k}\sum_{j\in A_k}\right)x_ix_j=2\sum_k(e_{A_k}x)(e_{B_k}x),\qquad(\clubsuit)$$

其中 $e_A=(\mathbb{1}_{[j\in A]})_{j\in[n]}$ 表示 $A\subset[n]$ 的示性向量, 从而 $e_Ax=\sum_{j\in A}x_j$. 当 $m\le n-2$ 时, 齐次线性方程组 $\begin{cases}e_{[n]}x=0\\ e_{A_k}x=0\ (k\in[m])\end{cases}$ 有非零解, 代入($\clubsuit$)可得左端 <0 而右端 $=0$, 显然不可能. □

2　方阵的行列式

2.1　Cramer 法则的启发

我们借鉴 Cramer 的方法[ii]来导出行列式 (determinant). 考察 n 个 n 元线性方程

$$L_n:\begin{pmatrix}a_{11}&a_{12}&\ldots&a_{1n}\\a_{21}&a_{22}&\ldots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\ldots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}.$$

通过具体计算 $n=1,2$ 的例子, 拟设如下 Cramer 法则 (Cramer's rule): 方程 L_n 的解形如

$$x_j=\frac{\det_n(a_1,\ldots,a_{j-1},b,a_{j+1},\ldots,a_n)}{\det_n(a_1,\ldots,a_n)},\quad j=1,\ldots,n,$$

其中 $a_j=(a_{ij})_{1\le i\le n}$ 表示系数矩阵第 j 列.

利用递归的想法, 假设 $\det_n$ 是已知函数, 我们尝试求解方程 L_{n+1}, 进而合理地定义出 $\det_{n+1}$. 通过行变换消元, 设 $y_1,\ldots,y_n$ 满足

$$y_1a_{1j}+\cdots+y_na_{nj}=-a_{n+1,j},\quad j=1,\ldots,n,$$

则

$$(y_1a_{1,n+1}+\cdots+y_na_{n,n+1}+a_{n+1,n+1})x_{n+1}=y_1b_1+\cdots+y_nb_n+b_{n+1}.$$

根据归纳假设, (注意行指标和列指标)

$$y_i=\frac{\det_n\big((a_{1j})_{1\le j\le n},\ldots,(a_{i-1,j})_{1\le j\le n},(-a_{n+1,j})_{1\le j\le n},(a_{i+1,j})_{1\le j\le n},\ldots,(a_{nj})_{1\le j\le n}\big)}{\det_n\big((a_{1j})_{1\le j\le n},\ldots,(a_{nj})_{1\le j\le n}\big)}.$$

引入 $y_{n+1}=1$, 记 $A_{i,n+1}=y_i\det_n\big((a_{1j})_{1\le j\le n},\ldots,(a_{nj})_{1\le j\le n}\big)$. 于是

$$x_{n+1}=\frac{\sum_{i=1}^{n+1}b_iA_{i,n+1}}{\sum_{i=1}^{n+1}a_{i,n+1}A_{i,n+1}}.$$

为了得到

$$x_{n+1}=\frac{\det_{n+1}((a_{ij})_{1\le i\le n+1;\,1\le j\le n},(b_i)_{1\le i\le n+1})}{\det_{n+1}((a_{ij})_{1\le i,j\le n+1})},$$

定义

$$\det_{n+1}((a_{ij})_{1\le i,j\le n+1})=\sum_{i=1}^{n+1}a_{i,n+1}A_{i,n+1}\qquad(\diamond)$$

即可 (并不依赖 Cramer 法则, 稍后可直接验证此性质).

[ii] What is the origin of the determinant in linear algebra? — https://math.stackexchange.com/q/194579

2.2 置换的符号

约定 $\det_1(a) = a$. 注意到 $A_{i,n+1}$ 是关于 a_{kj} $(k \neq i;\ j \neq n+1)$ 的函数, 基于($\diamond$)归纳可得

$$\det_n((a_{ij})_{1\le i,j\le n}) = \sum_{\sigma\in S_n} s_n(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

其中 S_n 是全体从 $\{1,\ldots,n\}$ 到自身的一一对应 (称为置换, permutation) 构成的集合, s_n 是 S_n 上的函数. 显然 $s_1 \equiv 1$, 下面借助 det 的递归定义来计算 s_n. 取 $\tau \in S_{n+1}$, 令 $a_{ij} = \mathbb{1}_{[\tau(i)=j]}$, 则有

$$\begin{aligned}
s_{n+1}(\tau) &= \det_{n+1}((a_{ij})_{1\le i,j\le n+1}) \\
&= A_{\tau^{-1}(n+1),n+1} \\
&= \det_n\big((a_{ij})_{i<\tau^{-1}(n+1);\, j\le n}, (-a_{n+1,j})_{j\le n}, (a_{ij})_{\tau^{-1}(n+1)<i\le n;\, j\le n}\big) \\
&= (-1)^{\mathbb{1}_{[\tau(n+1)\neq n+1]}} \det_n\big((a_{ij})_{i<\tau^{-1}(n+1);\, j\le n}, (a_{n+1,j})_{j\le n}, (a_{ij})_{\tau^{-1}(n+1)<i\le n;\, j\le n}\big) \\
&= (-1)^{\mathbb{1}_{[\tau(n+1)\neq n+1]}} s_n(\tau_1),
\end{aligned}$$

其中 $\tau_1 \in S_n$ 定义为

$$\tau_1(i) = \begin{cases} \tau(i), & i \neq \tau^{-1}(n+1); \\ \tau(n+1), & i = \tau^{-1}(n+1). \end{cases}$$

通过归纳, 易见 s_n 值域为 $\{\pm 1\}$. 按照 $\tau(n+1)$ 的取值情况, 我们有 (不妨将置换等同于其取值序列)

- $s_{n+1}(\sigma(1),\ldots,\sigma(n),n+1) = s_n(\sigma),\ \forall \sigma \in S_n$;
- 若 $\tau \in S_{n+1}$ 满足 $\tau(n+1) \neq n+1$, 则

$$\begin{aligned}
s_{n+1}(\tau) &= s_{n+1}\big((\tau(i))_{i<\tau^{-1}(n+1)}, n+1, (\tau(i))_{\tau^{-1}(n+1)<i\le n}, \tau(n+1)\big) \\
= -s_n(\tau_1) &= -s_n\big((\tau(i))_{i<\tau^{-1}(n+1)}, \tau(n+1), (\tau(i))_{\tau^{-1}(n+1)<i\le n}\big) \\
&= -s_{n+1}\big((\tau(i))_{i<\tau^{-1}(n+1)}, \tau(n+1), (\tau(i))_{\tau^{-1}(n+1)<i\le n}, n+1\big).
\end{aligned}$$

  换言之, 交换 τ 在 $\tau^{-1}(n+1)$ 和 $n+1$ 处的取值将改变 s_{n+1} 的符号.

这两个性质完全决定了 s_n, 称为置换的符号函数, 它与置换的奇偶性一一对应. 对 $\sigma \in S_n$ 定义

$$\operatorname{sgn}(\sigma) = (-1)^{\operatorname{inv}(\sigma)}, \quad \text{其中 } \operatorname{inv}(\sigma) = |\{(i,j) : i<j \ \&\ \sigma(i) > \sigma(j)\}| \text{ 表示 } (\sigma(j))_{1\le j\le n} \text{ 的逆序数.}$$

在 S_n 上定义乘法为映射复合, 即 $\sigma\tau = \big(j \mapsto \sigma(\tau(j))\big) \in S_n,\ \forall \sigma,\tau \in S_n$.

题 2. *如果置换 σ 满足 $|\{i : \sigma(i) \neq i\}| = 2$, 则称 σ 为对换 (transposition). 证明:*

1. *若 $\sigma \in S_n$ 是对换, 则有 $\operatorname{sgn}(\sigma) = -1$ 以及 $\operatorname{sgn}(\sigma\tau) = -\operatorname{sgn}(\tau),\ \forall \tau \in S_n$; 进一步地, $s_n = \operatorname{sgn}$;*
2. *交错群 $A_n = \{\sigma \in S_n : \operatorname{sgn}(\sigma) = 1\}$ 基数为 $n!/2$;*
3. *(乘法群) 同态性质 $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau),\ \forall \sigma,\tau \in S_n$; 特别地, $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)^{-1} = \operatorname{sgn}(\sigma)$;*
4. *任给 $\sigma \in S_n$, 存在 $s \in \{\pm 1\}$ 使得 $\prod_{1\le i<j\le n}(x_{\sigma(i)} - x_{\sigma(j)}) = s\prod_{1\le i<j\le n}(x_i - x_j),\ \forall x_1,\ldots,x_n \in \mathbb{F}$, 并且 $s = \operatorname{sgn}(\sigma)$; 利用此刻画重新得到 sgn 的同态性质.*

证明. 留给读者. 注意任意置换都是一些对换的乘积. □

题 3. *(对称群的生成元) 试问 S_n $(n \geq 3)$ 可由其中最少几个置换通过反复作乘法得到? 说明理由.*

解. 断言至少需要 2 个. 假设某个 n-置换 σ 能生成全部 n-置换, 则考察 1 与 $j \in [n]$ 的对换可知轨道 $\{\sigma^\ell(1)\}_{\ell\in\mathbb{N}}$ 必为 $[n]$, 从而 $\sigma^n = \mathrm{id}_{[n]}$, 但是这表明 $|\{\sigma^\ell\}_{\ell\in\mathbb{N}}| = n < n! = |S_n|$, 矛盾. 往证生成元可取

$$S = \mathrm{swap}(1,2) = (2,1,3,\ldots,n) \quad \text{and} \quad C = (n,1,2,\ldots,n-1).$$

由于任意置换都是一些对换 swap 的乘积, 而 $\mathrm{swap}(k,\ell) = \mathrm{swap}(1,k)\,\mathrm{swap}(1,\ell)\,\mathrm{swap}(1,k)$, 所以只需用 S 和 C 的乘积表出 $\mathrm{swap}(1,k)$, $k\in[n]$. 易见

$$C^{n-k+1}SC^{k-1} = \mathrm{swap}(k,k+1),$$

且 $\mathrm{swap}(1,k+1) = \mathrm{swap}(1,k)\,\mathrm{swap}(k,k+1)\,\mathrm{swap}(1,k)$, 归纳即证. □

2.3　行列式的性质

现已推出下述 Leibniz 公式 (Leibniz formula for determinants)

$$\left.\begin{aligned}\left|(a_{ij})_{1\leq i,j\leq n}\right| &= \det((a_{ij})_{1\leq i,j\leq n}) = \textstyle\sum_{\sigma\in S_n} \mathrm{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} \\ & \qquad\qquad\qquad\qquad\quad \| \\ \left|(a_{ji})_{1\leq i,j\leq n}\right| &= \det((a_{ji})_{1\leq i,j\leq n}) = \textstyle\sum_{\tau\in S_n} \mathrm{sgn}(\tau)a_{\tau(1)1}a_{\tau(2)2}\cdots a_{\tau(n)n}\end{aligned}\right)\tau=\sigma^{-1} \quad .$$

不难验证:

- 多重线性 (multilinear)

$$\left|(a_{ij})_{j<k}^{i\leq n} \quad (ca_{ik}+b_i)^{i\leq n} \quad (a_{ij})_{j>k}^{i\leq n}\right| = c\left|(a_{ij})_{j<k}^{i\leq n} \quad (a_{ik})^{i\leq n} \quad (a_{ij})_{j>k}^{i\leq n}\right| + \left|(a_{ij})_{j<k}^{i\leq n} \quad (b_i)^{i\leq n} \quad (a_{ij})_{j>k}^{i\leq n}\right|;$$

- 反对称 (anti-symmetric)

$$\left|(a_{ij})_{j<k}^{i\leq n} \quad (b_i)^{i\leq n} \quad (a_{ij})_{k<j<\ell}^{i\leq n} \quad (c_i)^{i\leq n} \quad (a_{ij})_{j>\ell}^{i\leq n}\right| = -\left|(a_{ij})_{j<k}^{i\leq n} \quad (c_i)^{i\leq n} \quad (a_{ij})_{k<j<\ell}^{i\leq n} \quad (b_i)^{i\leq n} \quad (a_{ij})_{j>\ell}^{i\leq n}\right|.$$

题 4. *设 $f : (\mathbb{F}^n)^n \to \mathbb{F}$ 是 n 重线性反对称函数, 即 $\forall 1\leq j\leq n$, $\forall c\in\mathbb{F}$, $\forall \alpha_1,\ldots,\alpha_{n-1},\alpha_n,\beta\in\mathbb{F}^n$,*

- $f(\alpha_1,\ldots,\alpha_{j-1},c\alpha_j+\beta,\alpha_{j+1},\ldots,\alpha_n) = \begin{cases} cf(\alpha_1,\ldots,\alpha_{j-1},\alpha_j,\alpha_{j+1},\ldots,\alpha_n) \\ +f(\alpha_1,\ldots,\alpha_{j-1},\beta,\alpha_{j+1},\ldots,\alpha_n)\end{cases}$;
- $f(\ldots,\alpha_i,\ldots,\alpha_j,\ldots) = -f(\ldots,\alpha_j,\ldots,\alpha_i,\ldots)$, $\forall i<j$.

证明 $f = f(I_n)\det$, 其中 $I_n = (\mathbb{1}_{[i=j]})_{1\leq i,j\leq n}$ 是单位矩阵 (identity matrix).

证明. 记 I_n 第 j 列为 e_j. 任取 $\alpha_j = (a_{ij})_{1\leq i\leq n}\in\mathbb{F}^n$, $j=1,\ldots,n$, 有

$$\begin{aligned} f(\alpha_1,\alpha_2,\ldots,\alpha_n) &= f\big(\textstyle\sum_{i_1=1}^n a_{i_11}e_{i_1}, \sum_{i_2=1}^n a_{i_22}e_{i_2},\ldots,\sum_{i_n=1}^n a_{i_nn}e_{i_n}\big) && \Big)\ f\ \text{多重线性} \\ &= \textstyle\sum_{i_1=1}^n\sum_{i_2=1}^n\cdots\sum_{i_n=1}^n a_{i_11}a_{i_22}\cdots a_{i_nn} f(e_{i_1},e_{i_2},\ldots,e_{i_n}) && \Big)\ f\ \text{反对称} \\ &= \underbrace{\sum_{(i_1,i_2,\ldots,i_n)\in S_n} a_{i_11}a_{i_22}\cdots a_{i_nn}\,\mathrm{sgn}(i_1,i_2,\ldots,i_n)}_{=\det(\alpha_1,\alpha_2,\ldots,\alpha_n)}\ f(\underbrace{e_1,e_2,\ldots,e_n}_{=I_n}). \end{aligned}$$

注意 f 的反对称性蕴涵 $f(\ldots,e,\ldots,e,\ldots) = 0$. □

给定 $A \in \mathbb{F}^{p\times p}$ 和 $M \in \mathbb{F}^{q\times p}$, 考察 $X \in \mathbb{F}^{q\times q} \mapsto \begin{vmatrix} A & 0 \\ M & X \end{vmatrix} \in \mathbb{F}$, 则有

$$\begin{vmatrix} A & 0 \\ M & X \end{vmatrix} \xlongequal{\text{题4}} \begin{vmatrix} A & 0 \\ M & I_q \end{vmatrix} \left|X\right| \xlongequal{\text{Leibniz 公式}} \left|A\right| \left|X\right|.$$

具体计算时还经常用到 (注意行和列的对称性)

$$\left|(a_{ij})_{1\le i,j\le n}\right| = \left|(a_{ij})_{j<k}^{i\le n} \quad (a_{ik} + \textstyle\sum_{j\neq k} c_j a_{ij})^{i\le n} \quad (a_{ij})_{j>k}^{i\le n}\right|.$$

三种初等变换对行列式的影响即已明晰. 注意到初等变换不改变行列式的非零性, 对于 n 阶方阵 A,

$$\det A \neq 0 \iff A \text{ 的简化行阶梯形的主元有 } n \text{ 个} \iff \text{线性方程组 } Ax = b \text{ 存在唯一解}.$$

题 5. 设 $a, b, c \in \mathbb{R}$ 至多其一为零. 证明在 Oxy 平面上三条直线 $ax + by + c = 0$, $bx + cy + a = 0$, $cx + ay + b = 0$ 恰好交于一点 (x, y) 的充要条件是 $a + b + c = 0$.

证明. 题中三条直线交于一点当且仅当方程组

$$\begin{cases} ax + by = -c \\ bx + cy = -a \\ cx + ay = -b \end{cases}$$

存在唯一解, 即对应的增广矩阵

$$\begin{pmatrix} a & b & -c \\ b & c & -a \\ c & a & -b \end{pmatrix}$$

恰有 2 个主元且都在系数列. 换言之, $\begin{vmatrix} a & b & -c \\ b & c & -a \\ c & a & -b \end{vmatrix} = 0$ 且 $\begin{vmatrix} a & b \\ b & c \end{vmatrix}, \begin{vmatrix} a & b \\ c & a \end{vmatrix}, \begin{vmatrix} b & c \\ c & a \end{vmatrix}$ 不全为零. 计算可得

$$\begin{aligned}
\begin{vmatrix} a & b & -c \\ b & c & -a \\ c & a & -b \end{vmatrix} &= -\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} && \text{第 1 列加第 3 列} \\
&= -\begin{vmatrix} a+b+c & b & c \\ a+b+c & c & a \\ a+b+c & a & b \end{vmatrix} && \text{第 1 列加第 2 列} \\
&= -(a+b+c)\begin{vmatrix} 1 & b & c \\ 1 & c & a \\ 1 & a & b \end{vmatrix} && \text{第 3 行减第 1 行} \\
&= -(a+b+c)\begin{vmatrix} 1 & b & c \\ 0 & c-b & a-c \\ 0 & a-b & b-c \end{vmatrix} && \text{第 2 行减第 1 行} \\
&= -(a+b+c)[(c-b)(b-c) - (a-c)(a-b)] \\
&= (a+b+c)(a^2+b^2+c^2-ab-bc-ca) \\
&= (a+b+c)[(a-b)^2+(b-c)^2+(c-a)^2]/2.
\end{aligned}$$

若 $a=b=c$, 则三条直线重合, 不符合题意, 故必有 $a+b+c=0$. 反之, 当 $a+b+c=0$ 时, 三条直线有唯一交点 $(1,1)$. □

题 6. [iii] 设 $A,B\in\mathbb{R}^{n\times n}$ 满足 $|\det(A+zB)|\le 1,\ \forall z\in\mathbb{C}:|z|=1$. 证明 $(\det A)^2+(\det B)^2\le 1$.

证明. 记 $\det(A+zB)=f(z)=\sum_{k=0}^n c_k z^k$, 则 $c_0=\det A$ 且 $c_n=\det B$. 我们有

$$\int_0^{2\pi}\mathrm{e}^{\mathrm{i}kt}\,\mathrm{d}t=0\ (\forall k\in\mathbb{Z}\setminus\{0\})\implies 1\ge\frac{1}{2\pi}\int_0^{2\pi}f(\mathrm{e}^{\mathrm{i}t})\,\underbrace{\overline{f(\mathrm{e}^{\mathrm{i}t})}}_{=f(\mathrm{e}^{-\mathrm{i}t})}\,\mathrm{d}t=\sum_{k=0}^n c_k^2\ge(\det A)^2+(\det B)^2.$$

□

2.4 Laplace 展开式

在(◇)中, $n+1$ 阶行列式表为若干 n 阶行列式的加权和. 此观点可继续延伸.

记 $[n]=\{1,\dots,n\}$ 和 $\binom{[n]}{r}=\{\mathcal{I}\subset[n]:|\mathcal{I}|=r\}$, 并约定任意 $\mathcal{I}\in\binom{[n]}{r}$ 均没有逆序对. 任给 $\mathcal{I}=(i_1,\dots,i_r)\in\binom{[n]}{r}$, 令 $\mathcal{I}^\complement=[n]\setminus\mathcal{I}\in\binom{[n]}{n-r}$, 定义

$$\mathscr{S}(\mathcal{I})=\operatorname{inv}(\mathcal{I},\mathcal{I}^\complement)=(i_1-1)+\cdots+(i_r-r)=\left(\textstyle\sum_{i\in\mathcal{I}}i\right)-\tfrac{r(r+1)}{2}.$$

对于 $A=(a_{ij})_{i\in[n],\,j\in[n]}$, 记 $A[\mathcal{I},\mathcal{J}]=(a_{ij})_{i\in\mathcal{I},\,j\in\mathcal{J}}$, 并定义 A 的 $[\mathcal{I},\mathcal{J}]$-代数余子式 (cofactor) 为

$$A_{\mathcal{I}\mathcal{J}}=(-1)^{\mathscr{S}(\mathcal{I})+\mathscr{S}(\mathcal{J})}\det A[\mathcal{I}^\complement,\mathcal{J}^\complement],\quad \mathcal{I},\mathcal{J}\in\tbinom{[n]}{r}.$$

从 Leibniz 公式出发, Laplace 给出了如下按行展开式 (Laplace expansion): $\forall\mathcal{I}\in\binom{[n]}{r}$,

$$\det A=\sum_{\mathcal{J}\in\binom{[n]}{r}}(-1)^{\mathscr{S}(\mathcal{I})+\mathscr{S}(\mathcal{J})}(\det A[\mathcal{I},\mathcal{J}])(\det A[\mathcal{I}^\complement,\mathcal{J}^\complement])=\sum_{\mathcal{J}\in\binom{[n]}{r}}(\det A[\mathcal{I},\mathcal{J}])A_{\mathcal{I}\mathcal{J}}.$$

当 $r=1$ 时,

$$\det A=\sum_{j\in[n]}(-1)^{i+j}A[i,j]\det A[\{i\}^\complement,\{j\}^\complement]=\sum_{j\in[n]}a_{ij}A_{ij},\quad\forall i\in[n].\tag{◇'}$$

一个颇有意义的特例是

$$\sum_{j\in[n]}a_{kj}A_{ij}=\sum_{j\in[n]}(-1)^{i+j}A[k,j]\det A[\{i\}^\complement,\{j\}^\complement]=0,\quad\forall k\ne i.$$

这些公式应用于 $A^\top=(a_{ji})_{i\in[n],\,j\in[n]}$ 即得 $\det A$ 的按列展开式. 至此, 通过叠加原理 (superposition principle) 容易验证 Cramer 法则: 既然 $x=(A_{ij}/\det A)_{j\in[n]}$ 是 $Ax=(\mathbb{1}_{[k=i]})_{k\in[n]}$ 的解, 那么 $x=\sum_i b_i(A_{ij}/\det A)_{j\in[n]}$ 就是 $Ax=\sum_i b_i(\mathbb{1}_{[k=i]})_{k\in[n]}=(b_i)_{i\in[n]}$ 的解.

题 7. 记 $A=(a_{ij})_{1\le i,j\le n}$ 的 (i,j)-代数余子式为 A_{ij}. 证明 $\begin{vmatrix} & A & & x_1\\ & & & \vdots\\ & & & x_n\\ z_1 & \dots & z_n & y\end{vmatrix}=\Big|A\Big|y-\sum_{i,j=1}^n A_{ij}x_iz_j$.

证明. 原式左端 $\xlongequal{\text{按最后一行展开}}\Big|A\Big|y+\sum_{j=1}^n(-1)^{n+1+j}z_j\begin{vmatrix}a_{11}&\dots&a_{1,j-1}&a_{1,j+1}&\dots&a_{1n}&x_1\\ \vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\ a_{n1}&\dots&a_{n,j-1}&a_{n,j+1}&\dots&a_{nn}&x_n\end{vmatrix}$

$\xlongequal{\text{按最后一列展开}}\Big|A\Big|y+\sum_{j=1}^n(-1)^{n+1+j}z_j\sum_{i=1}^n(-1)^{i+n}x_i(-1)^{i+j}A_{ij}=$ 原式右端. □

[iii] Andreescu, T.; Mortici, C.; Tetiva, M. (2017). *Mathematical bridges*, §3.

记 $\mathbf{1}=(1,\cdots,1)^\top$ 和 $\boldsymbol{t}=(t_1,\ldots,t_n)^\top$, 则

$$\left|(a_{ij}+t_j)_{1\le i,j\le n}\right| = \begin{vmatrix} A+\mathbf{1}\boldsymbol{t}^\top & \mathbf{1} \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} A & \mathbf{1} \\ -\boldsymbol{t}^\top & 1 \end{vmatrix} = \left|A\right| + \sum_{i=1}^n \sum_{j=1}^n A_{ij} t_j.$$

题 8. *证明对角阵的非零子式 (子矩阵的行列式, minor) 一定是主子式 (行列指标相同的子式, principal minor), 从而是某些对角元的乘积.*

证明. 只需对角阵不是主子式的子式均为零. 任取 $A=(a_{ij})_{1\le i,j\le n}$, 其中 $a_{ij}=0$ $(\forall i\ne j)$. 考察子式

$$\begin{vmatrix} a_{i_1j_1} & \cdots & a_{i_1j_r} \\ \vdots & \ddots & \vdots \\ a_{i_rj_1} & \cdots & a_{i_rj_r} \end{vmatrix}, \quad \begin{array}{l} 1\le i_1<i_2<\cdots<i_r\le n; \\ 1\le j_1<j_2<\cdots<j_r\le n. \end{array}$$

若存在 l 使 $i_l\ne j_l$, 记 l_0 为最小的那个, 不妨(否则转置)设 $i_{l_0}<j_{l_0}$, 则上式第 l_0 行全为零, 按这一行展开即证. □

结合题7, $$\begin{vmatrix} 1+t & t & \cdots & t \\ t & 2+t & \ddots & \vdots \\ \vdots & \ddots & \ddots & t \\ t & \cdots & t & n+t \end{vmatrix} = n!\Big(1+t\sum_{j=1}^n \frac{1}{j}\Big).$$

题 9. *(连项式[iv]) 令*

$$K_n(x_1,x_2,\ldots,x_n) = \begin{vmatrix} x_1 & 1 & 0 & \cdots & 0 \\ -1 & x_2 & 1 & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & -1 & x_n \end{vmatrix}.$$

证明

$$\frac{K_n(x_1,\ldots,x_n)}{K_{n-1}(x_2,\ldots,x_n)} = [x_1;x_2,\ldots,x_n] = x_1 + \cfrac{1}{x_2+\cfrac{1}{\ddots+\cfrac{1}{x_n}}}.$$

证明. 按第一行展开可得 $K_n(x_1,\ldots,x_n)=x_1K_{n-1}(x_2,\ldots,x_n)+K_{n-2}(x_3,\ldots,x_n)$. (特例是 Fibonacci 数列.) 重写为 $\dfrac{K_n(x_1,\ldots,x_n)}{K_{n-1}(x_2,\ldots,x_n)} = x_1 + 1\Big/\dfrac{K_{n-1}(x_2,\ldots,x_n)}{K_{n-2}(x_3,\ldots,x_n)}$, 归纳即得. □

题 10. *定义 n 阶方阵 A_i 满足第 i 行第 j 列为 x_ja_{ij} 且第 $k\ne i$ 行第 j 列为 a_{kj}. 求 $\sum_{i=1}^n \det A_i$.*

解. 令 $A=(a_{ij})_{1\le i,j\le n}$ 并记其 (i,j)-代数余子式为 A_{ij}, 则

$$\begin{aligned} \text{原式} &= \sum_{i=1}^n \sum_{j=1}^n x_j a_{ij} A_{ij} \\ &= \sum_{j=1}^n x_j \sum_{i=1}^n a_{ij} A_{ij} = \sum_{j=1}^n x_j \det A. \end{aligned}$$

□

[iv] Graham, R.L.; Knuth, D.E.; Patashnik, O. (1994). *Concrete mathematics: a foundation for computer science*, 2nd ed., §6.7.

借助($\diamond$')和归纳法, 可以计算出 Vandermonde 矩阵

$$V_n(x_1,\dots,x_n)=(x_i^{j-1})_{1\le i,j\le n}=\begin{pmatrix}1 & x_1 & x_1^2 & \cdots & x_1^{n-1}\\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1}\end{pmatrix}$$

的行列式

$$\det V_n(x_1,\dots,x_n)=\prod_{i>j}(x_i-x_j).$$

题 11. *(Lagrange 插值多项式[v]) 任给 $y_0,y_1,\dots,y_n$ 和互异的 $x_0,x_1,\dots,x_n$, 求不超过 n 次的多项式*

$$p(x)=\textstyle\sum_{j=0}^n a_jx^j\quad \text{s.t.}\quad p(x_i)=y_i\ (\forall i).$$

解. 记 Vandermonde 矩阵 $V_{n+1}=(x_{i-1}^{j-1})_{1\le i,j\le n+1}$ 删去第 k 行第 ℓ 列所得子矩阵为 $V_{n;k,\ell}=(x_{i-1}^{j-1})_{j\neq\ell}^{i\neq k}$. 根据 Cramer 法则, 由 $V_{n+1}(a_{j-1})_{1\le j\le n+1}=(y_{i-1})_{1\le i\le n+1}$ 可解得

$$a_{j-1}=\frac{1}{\det V_{n+1}}\sum_{i=1}^{n+1}(-1)^{i+j}y_{i-1}\det V_{n;i,j},\qquad j=1,2,\dots,n+1.$$

将 $\det V_{n+1}=\prod_{1\le i<j\le n+1}(x_{j-1}-x_{i-1})$ 视作关于 x_{k-1} 的多项式, 则有

$$\sum_{\ell=1}^{n+1}(-1)^{k+\ell}(\det V_{n;k,\ell})x_{k-1}^{\ell-1}=\Bigg(\prod_{\substack{1\le i<j\le n+1\\ i,j\neq k}}(x_{j-1}-x_{i-1})\Bigg)(-1)^{n+1-k}\underbrace{\prod_{d\neq k}(x_{k-1}-x_{d-1})}_{=\sum_{\ell=1}^{n+1}(-1)^{n+1-\ell}\sigma_{n+1-\ell}\left((x_{d-1})_{d\neq k}\right)x_{k-1}^{\ell-1}},$$

其中 $\sigma_0=1$ 且 $\sigma_m(z_1,\dots,z_n)=\sum_{1\le i_1<\cdots<i_m\le n}z_{i_1}\cdots z_{i_m}$, $m=1,\dots,n$. 于是

$$\det V_{n;k,\ell}=\sigma_{n+1-\ell}\big((x_{d-1})_{d\neq k}\big)\prod_{\substack{1\le i<j\le n+1\\ i,j\neq k}}(x_{j-1}-x_{i-1})$$

$$\Longrightarrow\ \frac{\det V_{n;k,\ell}}{\det V_{n+1}}=\frac{\sigma_{n+1-\ell}\big((x_{d-1})_{d\neq k}\big)}{(-1)^{n+1-k}\prod_{d\neq k}(x_{k-1}-x_{d-1})},$$

从而

$$a_{\ell-1}=(-1)^{n+\ell-1}\sum_{k=1}^{n+1}y_{k-1}\frac{\sigma_{n+1-\ell}\big((x_{d-1})_{d\neq k}\big)}{\prod_{d\neq k}(x_{k-1}-x_{d-1})},\qquad \ell=1,2,\dots,n+1.$$

$\square$

展开 $p(x)=\sum_{i=0}^n y_i\prod_{j\neq i}\frac{x-x_j}{x_i-x_j}$ 更为简捷.

题 12. 求 $\left\{(z_1,\dots,z_n)\in\mathbb{C}^n:\sum_{j=1}^n z_j^k=0,\ \forall k=1,2,\dots,n\right\}$.

解. 利用对称多项式理论和 Vieta 公式即可一眼看穿只有零解. 假设该集合有非零元, 记其不同的非零分量值为 $w_1,\dots,w_r\in\mathbb{C}$, 相应的出现次数为 $n_1,\dots,n_r\in\mathbb{N}\subset\mathbb{C}$, 则

$$w_1^kn_1+w_2^kn_2+\cdots+w_r^kn_r=0\quad(k=1,2,\dots,r).$$

系数矩阵为 $(w_j^k)_{1\le k,j\le r}$, 其行列式 $w_1\cdots w_r\det V_r(w_1,\dots,w_r)$ 非零. 这蕴涵 $n_1=\cdots=n_r=0$, 矛盾! 是故所求为 $\{0\}$. $\square$

[v] Hoffman, K.; Kunze, R. (1971). *Linear algebra*, 2nd ed., §4.3.

题 13. 计算行列式

$$\begin{vmatrix} 2 & a & 6 & 3 & c \\ 0 & 1 & 2 & 0 & 1 \\ 3 & 7 & b & 4 & 5 \\ 0 & 4 & 1 & 0 & 2 \\ 0 & 1 & 2 & 0 & 3 \end{vmatrix}.$$

解. 对第 $2,4,5$ 行作 Laplace 展开, 原式 $= (-1)^{(2+4+5)+(2+3+5)} \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} \begin{vmatrix} 1 & 2 & 1 \\ 4 & 1 & 2 \\ 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 \\ 4 & 1 & 2 \\ 0 & 0 & 2 \end{vmatrix} = -14.$ □

题 14. *(Putnam 全美大学生数学竞赛 2018 年 A2)* 用 $2^{[n]} \setminus \{\emptyset\}$ 表示 $[n] = \{1, \dots, n\}$ 的全体非空子集的集合. 求 $M_n = \left(\mathbb{1}_{[A \cap B \neq \emptyset]}\right)_{A, B \in 2^{[n]} \setminus \{\emptyset\}}$ 的行列式.

解. 对 n 归纳: 当 $n = 1$ 时显然有 $\det M_1 = 1$; 利用

$$\begin{array}{c c} & \begin{array}{ccc} n \notin B & & n \in B \end{array} \\ \begin{array}{c} n \notin A \\ n \in A \end{array} & \begin{pmatrix} M_{n-1} & 0 & M_{n-1} \\ 0 & 1 & \mathbf{1}^\top \\ M_{n-1} & \mathbf{1} & \mathbf{1}\mathbf{1}^\top \end{pmatrix} \end{array} = M_n$$

可得

$$\det M_n \xlongequal[\text{接着按第 } 2^{n-1} \text{ 行展开}]{\text{后 } (2^{n-1}-1) \text{ 列减去第 } 2^{n-1} \text{ 列}} \begin{vmatrix} M_{n-1} & M_{n-1} \\ M_{n-1} & 0 \end{vmatrix} \xlongequal{\text{按后 } (2^{n-1}-1) \text{ 行展开}} -(\det M_{n-1})^2, \quad n \geq 2,$$

从而 $\det M_n = -1,\ n \geq 2.$ □

3 向量空间 $\mathbb{F}^n$

集合配备加法和数乘 (标量乘法, scalar multiplication) 即成为线性空间 (linear space), 其中的元素常称为向量 (vector). 域 $\mathbb{F}$ 上的 $\mathbb{F}^n$ 是最典型的例子, 并且约定向量按列写出. 线性空间的子集继承加法和数乘后, 如果本身能成为线性空间, 即满足运算的封闭性, 则称其为原空间的线性子空间 (subspace). 设 W 是 $\mathbb{F}$-线性空间 V 的子集, 则 W 是线性子空间当且仅当

$$cw_1 + w_2 \in W \ (\forall c \in \mathbb{F},\ w_1, w_2 \in W).$$

3.1 线性组合

设 V 是 $\mathbb{F}$-线性空间. 任取有限个向量 $v_1, \dots, v_p \in V$ 和相同多个标量 $c_1, \dots, c_p \in \mathbb{F}$, 则 $\sum_{j=1}^p c_j v_j$ 称为 $v_1, \dots, v_p$ 的一个线性组合 (linear combination). 向量 $v_1, \dots, v_p$ 的所有线性组合构成一个线性子空间

$$\operatorname{span}\{v_1, \dots, v_p\} = \{\textstyle\sum_{j=1}^p c_j v_j : c_1, \dots, c_p \in \mathbb{F}\},$$

称为 $v_1, \dots, v_p$ 张成 (span) 的子空间. 若 $w \in \operatorname{span}\{v_1, \dots, v_p\}$, 则称 $v_1, \dots, v_p$ 能线性表示 (express) 出 w. 求解线性方程组 $Ax = b$ 就相当于使 A 的列向量线性表示出 b.

题 15. 当 a,b 取何值时, $\alpha_1=\begin{pmatrix}1\\1\\4\end{pmatrix}$, $\alpha_2=\begin{pmatrix}-2\\1\\5\end{pmatrix}$, $\alpha_3=\begin{pmatrix}a\\2\\10\end{pmatrix}$ 可以线性表出 $\beta=\begin{pmatrix}1\\b\\-1\end{pmatrix}$? 何时表示方法唯一?

解. 求线性表出的系数就是解线性方程组. 对增广矩阵作 Gauss 消元:

$$\begin{pmatrix}\alpha_1 & \alpha_2 & \alpha_3 & \beta\end{pmatrix}=\begin{pmatrix}1&-2&a&1\\1&1&2&b\\4&5&10&-1\end{pmatrix}\xrightarrow[\text{第 3 行减第 2 行的 4 倍}]{\text{第 1 行减第 2 行}}\begin{pmatrix}0&-3&a-2&1-b\\1&1&2&b\\0&1&2&-1-4b\end{pmatrix}\xrightarrow{\text{第 2 行减第 3 行}}$$
$$\to\begin{pmatrix}0&-3&a-2&1-b\\1&0&0&1+5b\\0&1&2&-1-4b\end{pmatrix}\xrightarrow{\text{第 1 行加第 3 行的 3 倍}}\begin{pmatrix}0&0&a+4&-2-13b\\1&0&0&1+5b\\0&1&2&-1-4b\end{pmatrix}\to\begin{pmatrix}1&0&0&1+5b\\0&1&2&-1-4b\\0&0&a+4&-2-13b\end{pmatrix},$$

有解当且仅当主元不出现在常数列, 即不发生情形 $\begin{cases}a=-4\\b\neq-\frac{2}{13}\end{cases}$; 有唯一解当且仅当 $a\neq-4$. □

3.2 线性无关向量组

下面考虑线性表出的唯一性. 如果一组向量 $v_1,\dots,v_p$ 线性表示出零向量的方式 (线性组合系数) 不唯一, 即存在不全为零的标量 $c_1,\dots,c_p$ 使得 $\sum_{j=1}^p c_jv_j=0$, 则称 $v_1,\dots,v_p$ 线性相关 (linearly dependent); 否则称之为线性无关 (linearly independent).

题 16. *证明: 若一组向量线性无关, 则这组向量线性表出任何向量的方式至多有一个.*

证明. 若向量 w 可被 $v_1,\dots,v_p$ 线性表出为 $w=\sum_{j=1}^p c_jv_j=\sum_{j=1}^p \tilde{c}_jv_j$, 则 $\sum_{j=1}^p(c_j-\tilde{c}_j)v_j=w-w=0$. 根据线性无关性, 必有 $c_j-\tilde{c}_j=0$ $(\forall j)$, 即 $c_j=\tilde{c}_j$ $(\forall j)$. □

题 17. *证明: 若一组向量线性表出某一个向量的方式存在且唯一, 则这组向量线性无关.*

证明. 设向量 $v_1,\dots,v_p$ 线性表出向量 w 的方式恰为 $w=\sum_{j=1}^p c_jv_j$. 任取标量 $d_1,\dots,d_p$ 适合 $\sum_{j=1}^p d_jv_j=0$, 则 $\sum_{j=1}^p(c_j+d_j)v_j=w+0=w$. 根据唯一性, $c_j+d_j=c_j\implies d_j=0$ $(\forall j)$. □

设向量 $v_1,\dots,v_p$ 线性无关, 则

$$\operatorname{span}\{v_1,\dots,v_p\}=\{w: v_1,\dots,v_p,w \text{ 线性相关}\}.$$

我们自然地要关注线性无关向量组. 线性空间 V 的基 (basis) 指的是极大 (maximal) 的线性无关向量组(存在性归于 Zorn 引理, 可以使其包含任意线性无关向量组), 即一组线性无关的向量 $v_1,\dots,v_n$, 满足对任意 $w\in V$ 均有 $v_1,\dots,v_n,w$ 线性相关, 亦即

$$V=\operatorname{span}\{v_1,\dots,v_n\}.$$

题18表明线性空间 V 的任意两组基中向量的个数是相同的, 记为 $\dim V$, 称作 V 的维数 (dimension). 根据极大性, 线性空间 V 中超过 $\dim V$ 个向量必然线性相关.

题 18. *(Steinitz 替换引理[vi]) 设 $V=\operatorname{span}\{v_1,\dots,v_t\}$. 任取线性无关的 $w_1,\dots,w_s\in V$, 证明: $s\le t$, 并且存在 $\{1,\dots,t\}$ 的 $(t-s)$ 元子集 $\mathcal{I}$ 使得*

$$\operatorname{span}(\{w_1,\dots,w_s\}\cup\{v_i\}_{i\in\mathcal{I}})=V.$$

[vi] Jänich, K. (1994). *Linear algebra*, §3.6.

证明. 当 $s=0$ 时命题显然. 下面进行归纳: 若

$$w_{s+1} \in \operatorname{span}(\{w_1,\dots,w_s\} \cup \{v_i\}_{i\in\mathcal{I}}) \setminus \operatorname{span}\{w_1,\dots,w_s\},$$

则存在 $|\mathcal{I}|$ 个标量 c_i 使得 $w_{s+1} - \sum_{i\in\mathcal{I}} c_i v_i \in \operatorname{span}\{w_1,\dots,w_s\}$, 其中必存在 $j \in \mathcal{I}$ 适合 $c_j \neq 0$, 于是 $t = s + |\mathcal{I}| \geq s+1$ 且

$$v_j \in \frac{1}{c_j} w_{s+1} - \sum_{i\in\mathcal{I}\setminus\{j\}} \frac{c_i}{c_j} v_i + \operatorname{span}\{w_1,\dots,w_s\} \subset \operatorname{span}(\{w_1,\dots,w_s,w_{s+1}\} \cup \{v_i\}_{i\in\mathcal{I}\setminus\{j\}}),$$

即得所求. □

如果 W 是 V 的子空间, 那么**题18**蕴涵

$$\dim W \leq \dim V,$$

且等号成立当且仅当 $W=V$.

题 19. 设 $U = \operatorname{span}\{v_1,\dots,v_p\} \neq \operatorname{span}\{v_1,\dots,v_p,v_{p+1}\} = V$. 证明: $v_{p+1} \notin U$, 但是

$$v_{p+1} \in W_w = \operatorname{span}\{v_1,\dots,v_p,w\}, \ \forall w \in V \setminus U.$$

证明. 若 $v_{p+1} \in U$, 则 $V \subset U \subset V$, 这与 $U \neq V$ 矛盾! 任取 $w \in V\setminus U$, 有 $U \subsetneq W_w \subset V$. 我们得到 $\dim U < \dim W_w \leq \dim V \leq \dim U + 1$, 从而 $\dim W_w = \dim V$, 故 $W_w = V \ni v_{p+1}$. □

3.3 矩阵的秩

对于一个矩阵, 其列向量张成子空间的维数称为列秩 (column rank), 其行向量张成子空间的维数称为行秩 (row rank). **题20**将说明行秩等于列秩, 从而统称为矩阵的秩 (rank).

题 20. *(CR 分解[vii]) 设 $A \in \mathbb{F}^{m\times n}$, 各列依次记为 $\alpha_1,\alpha_2,\dots,\alpha_n$. 定义 A 的列空间 (column space) 为 $\operatorname{col} A = \operatorname{span}\{\alpha_1,\alpha_2,\dots,\alpha_n\}$. 通过**算法1**构造 $C \in \mathbb{F}^{m\times r}$. 证明:*

Algorithm 1: Column-Row Rank-Revealing Factorization (Heuristic, Incomplete)

Input: $A = \begin{pmatrix}\alpha_1 & \alpha_2 & \dots & \alpha_n\end{pmatrix}$

Output: $C = \begin{pmatrix}\gamma_1 & \gamma_2 & \dots & \gamma_r\end{pmatrix} = \begin{pmatrix}\alpha_{j_1} & \alpha_{j_2} & \dots & \alpha_{j_r}\end{pmatrix}$

```
begin
    C ← ∅;
    for j = 1, 2, ..., n do
        if α_j ∉ col C then
            C ← (C  α_j);
```

1. *C 的 r 列在 $\mathbb{F}^m$ 中线性无关;*
2. *存在唯一的 $R = (x_{ij})_{1\leq j\leq n}^{1\leq i\leq r} \in \mathbb{F}^{r\times n}$ 使得 $A = \begin{pmatrix}Cx_1 & Cx_2 & \dots & Cx_n\end{pmatrix}$, 其中 $x_j = (x_{ij})_{1\leq i\leq r}$;*
3. *A 的列秩为 r;*

[vii] Strang, G.; Moler, C. (2022). LU and CR elimination.

4. A 的简化行阶梯形矩阵 R_0 的非零行恰为 R;

5. R 的 r 行在 $\mathbb{F}^n$ 中线性无关, 从而 A 的行秩为 r.

证明. 1. 设 $c_1, c_2, \dots, c_r \in \mathbb{F}$ 满足 $\sum_{i=1}^r c_i \gamma_i = 0$. 由 $\gamma_r \notin \operatorname{span}\{\gamma_1, \gamma_2, \dots, \gamma_{r-1}\}$ 可得 $c_r = 0$, 类似递推则有 $0 = c_{r-1} = c_{r-2} = \cdots = c_1$.

2. 相当于求解 $\alpha_j = C x_j = \sum_{i=1}^r x_{ij} \gamma_i$, 其中 $x_j = (x_{ij})_{1 \le i \le r}$ 是未知量. 分类讨论如下:

   (a) 若 $j = j_k$ 对某个 $1 \le k \le r$ 成立, 则方程化为 $\sum_{i=1}^r \lambda_i \gamma_i = 0$, 其中 $\lambda_k = x_{kj} - 1$ 且 $\lambda_i = x_{ij}$ $(\forall i \ne k)$, 所以存在唯一解 $x_j = (\mathbb{1}_{[i=k]})_{1 \le i \le r}$;

   (b) 若 $j < j_1$, 则 $\alpha_j = 0$, 此时必有 $x_j = 0 \in \mathbb{F}^r$;

   (c) 若 $j_k < j < j_{k+1}$ 对某个 $1 \le k \le r$ 成立 (其中约定 $j_{r+1} = n+1$), 则 $\alpha_j \in \operatorname{span}\{\gamma_1, \gamma_2, \dots, \gamma_k\}$, 从而存在 $\mu_1, \dots, \mu_k \in \mathbb{F}$ 使得 $\alpha_j = \sum_{i=1}^k \mu_i \gamma_i$. 此时方程化为 $\sum_{i=1}^r \nu_i \gamma_i = 0$, 其中 $\nu_i = x_{ij} - \mu_i$ $(\forall i \le k)$ 且 $\nu_i = x_{ij}$ $(\forall i > k)$. 如前可得 $\nu_i = 0$ $(\forall i)$, 所以 $x_j \in \mathbb{F}^r$ 存在且唯一.

3. 只需 A 和 C 有相同的列空间. 根据定义有 $\operatorname{col} C \subset \operatorname{col} A$, 而2表明 $\operatorname{col} A \subset \operatorname{col} C$.

4. 记行向量 $R_{ik} = (x_{ij})_{j_k < j < j_{k+1}}$. 对于列向量 v 和行向量 $\rho = \begin{pmatrix} \rho_1 & \dots & \rho_s \end{pmatrix}$, 记 $v\rho = \begin{pmatrix} \rho_1 v & \dots & \rho_s v \end{pmatrix}$. 由2可得

$$R = \begin{pmatrix} 1 & R_{11} & 0 & R_{12} & \dots & 0 & R_{1r} \\ & & 1 & R_{22} & \dots & 0 & R_{2r} \\ & & & & \ddots & \vdots & \vdots \\ 0 & & & & & 1 & R_{rr} \end{pmatrix},$$

$$A = \begin{pmatrix} 0 & \gamma_1 & \gamma_1 R_{11} & \gamma_2 & \gamma_1 R_{12} + \gamma_2 R_{22} & \dots & \gamma_r & \sum_{j=1}^r \gamma_j R_{jr} \end{pmatrix}.$$

因为行变换不改变列向量之间的线性相关性($\gamma_j^{(j-1)} \notin \operatorname{span}\{e_1, \dots, e_{j-1}\}$), 所以 Gauss 消元给出

$$\begin{aligned} A &\to \begin{pmatrix} 0 & e_1 & e_1 R_{11} & \gamma_2^{(1)} & e_1 R_{12} + \gamma_2^{(1)} R_{22} & \dots & \gamma_r^{(1)} & e_1 R_{1r} + \sum_{j>1} \gamma_j^{(1)} R_{jr} \end{pmatrix} \\ &\to \begin{pmatrix} 0 & e_1 & e_1 R_{11} & e_2 & e_1 R_{12} + e_2 R_{22} & \dots & \gamma_r^{(2)} & e_1 R_{1r} + e_2 R_{2r} + \sum_{j>2} \gamma_j^{(2)} R_{jr} \end{pmatrix} \\ &\to \cdots \\ &\to \begin{pmatrix} 0 & e_1 & e_1 R_{11} & e_2 & e_1 R_{12} + e_2 R_{22} & \dots & e_r & \sum_{j=1}^r e_j R_{jr} \end{pmatrix} = \begin{pmatrix} R \\ 0 \end{pmatrix} = R_0, \end{aligned}$$

其中 $e_j = (\mathbb{1}_{[i=j]})_{1 \le i \le m}$. 注意 A 的主元所在列构成 C, 执行**算法**1之前不妨作 Gauss 消元.

5. 因为 R 有子矩阵 I_r, 所以 R 各行线性无关. 由4可知 A 和 R 有相同的行空间 (row space, 由行向量张成), 故 A 的行秩与 R 一样也是 r.

$\square$

应用列秩的定义, (注意转置后也有类似结论)

$$\max\{\operatorname{rank} A, \operatorname{rank} B\} \le \operatorname{rank} \begin{pmatrix} A & B \end{pmatrix} \le \operatorname{rank} A + \operatorname{rank} B;$$

$$\operatorname{rank} \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \ge \operatorname{rank} A + \operatorname{rank} C, \quad \text{当 } B = 0 \text{ 时等号成立.}$$

题 21. [viii] 记 $[n]=\{1,2,\dots,n\}$ 和 $\binom{[n]}{k}=\{K\subset[n]:|K|=k\}$, $\forall k\in[n]\cup\{0\}$. 任给 $0\le a\le b\le n$, 证明矩阵 $M_{n;a,b}=(\mathbb{1}_{[A\subset B]})_{A\in\binom{[n]}{a},\,B\in\binom{[n]}{b}}$ 的秩为 $\min\{\binom{n}{a},\binom{n}{b}\}$.

证明. 由 $\mathbb{1}_{[A\subset B]}=\mathbb{1}_{[[n]\setminus B\subset[n]\setminus A]}$ 可知 $M_{n;a,b}^{\top}=M_{n;n-b,n-a}$. 注意到 $\binom{n}{n-k}=\binom{n}{k}$, 不妨设 $\binom{n}{a}\le\binom{n}{b}$, 即 $a+b\le n$. 考虑对 n 归纳: 当 $n=1$ 时命题显然; 利用

$$\begin{array}{c|cc} & n\notin B & n\in B \\ \hline n\notin A & & \\ n\in A & & \end{array}\quad \begin{pmatrix} M_{n-1;a,b} & M_{n-1;a,b-1} \\ 0 & M_{n-1;a-1,b-1}\end{pmatrix}=M_{n;a,b}$$

易见 $\operatorname{rank}M_{n-1;a,b}+\operatorname{rank}M_{n-1;a-1,b-1}\le\operatorname{rank}M_{n;a,b}$, 从而 (行) 满秩性质得以赓续(杨辉三角). □

矩阵的秩也可以用行列式刻画. 方阵 A 满秩当且仅当

$$A\text{ 各列线性无关}\iff\text{线性方程组 }Ax=0\text{ 存在唯一解}\iff\det A\ne0.$$

一般地, 任给矩阵 $A\in\mathbb{F}^{m\times n}$, 有

$$\operatorname{rank}A=\max\left\{r:\exists\mathcal{I}\in\tbinom{[m]}{r},\ \mathcal{J}\in\tbinom{[n]}{r}\text{ s.t. }\det A[\mathcal{I},\mathcal{J}]\ne0\right\}.$$

线性方程组存在解的条件也有秩的描述. 对于 $\beta\in\mathbb{F}^m$, 显然 $\operatorname{col}A\subset\operatorname{col}\begin{pmatrix}A&\beta\end{pmatrix}$, 考察取等条件可知

$$\beta\in\operatorname{col}A\iff\operatorname{rank}A=\operatorname{rank}\begin{pmatrix}A&\beta\end{pmatrix}.$$

题 22. *(Plücker 坐标)* 设 $A\in\mathbb{F}^{n\times r}$ 各列线性无关. 证明 $\operatorname{col}A$ 取决于

$$p_{\mathcal{I}}=\det A[\mathcal{I},[r]],\quad \mathcal{I}\in\tbinom{[n]}{r}.$$

证明. 对于 $\beta=(b_i)_{i\in[n]}\in\mathbb{F}^n$, 有 $\operatorname{rank}\begin{pmatrix}A&\beta\end{pmatrix}\le r$

$$\iff\textstyle\sum_{k=1}^{r+1}(-1)^{k+r+1}p_{i_1,\dots,i_{k-1},i_{k+1},\dots,i_{r+1}}b_{i_k}=0,\ \forall(i_1,\dots,i_{r+1})\in\binom{[n]}{r+1}.$$ □

3.4 线性映射与线性方程组

所谓映射 (mapping), 就是某种指派规则, 将定义域 (domain) 中的每个元素 (element) 都唯一地对应到陪域 (codomain) 中的某个元素. 如果映射 f 有定义域 X 和陪域 Y, 那么常写成 $f:X\to Y$, 其值域 (range) 为 $f(X)=\{f(x):x\in X\}$; 代数学中值域称为像 (image), 记作 $\operatorname{im}f$. 不难验证表1的

表 1: 映射 $f:X\to Y$ 的性质

	单 (injective, 符号 $\hookrightarrow$)	满 (surjective, 符号 $\twoheadrightarrow$)
集合论定义	$\forall y\in f(X),\ \exists!x\in X,\ \text{s.t. }f(x)=y$	$\forall y\in Y,\ \exists x\in X,\ \text{s.t. }f(x)=y$
范畴论定义	$\forall g,h:Z\to X$, 有 $(f\circ g=f\circ h\implies g=h)$	$\forall g,h:Y\to Z$, 有 $(g\circ f=h\circ f\implies g=h)$
代数性质	(左逆) $\exists\varphi:Y\to X,\ \text{s.t. }\varphi\circ f=\mathrm{id}_X$	(右逆) $\exists\psi:Y\to X,\ \text{s.t. }f\circ\psi=\mathrm{id}_Y$

同一列中各行等价. 双 (bijective) 射是既单又满的映射, 也称作一一对应, 其左逆和右逆存在且相等. 表1中的各项刻画将应用于线性映射, 之后会不加说明地援引. 尽管有限维空间性质优良, 在无穷维则需要多加小心, 建议记住无穷维空间上的推移算子 (shift operator) 作为典例: ($LR_0=\mathrm{id}$ 而 $R_0L\ne\mathrm{id}$)

$$\begin{aligned}L&:(x_0,x_1,x_2,\dots)\in\mathbb{F}^\infty\mapsto(x_1,x_2,\dots)\in\mathbb{F}^\infty,\\ R_{x_0}&:(x_1,x_2,\dots)\in\mathbb{F}^\infty\mapsto(x_0,x_1,x_2,\dots)\in\mathbb{F}^\infty\quad(x_0\in\mathbb{F}).\end{aligned}$$

[viii] Bollobás, B. (2006). *The art of mathematics: coffee time in Memphis*, Problem 130.

设 X 和 Y 是 $\mathbb{F}$-线性空间. 映射 $T: X \to Y$ 称为 $\mathbb{F}$-线性的 (linear), 若

$$T(cx_1 + x_2) = cT(x_1) + T(x_2), \quad \forall c \in \mathbb{F},\ x_1, x_2 \in X.$$

对于线性映射 $T: X \to Y$, 易得 $\operatorname{im} T$ 是 Y 的线性子空间; 定义核 (kernel)/零空间 (null space) 为

$$\ker T = T^{-1}(0) = \{x \in X : T(x) = 0\},$$

它是 X 的线性子空间.

任给矩阵 $A = \begin{pmatrix} \alpha_1 & \dots & \alpha_n \end{pmatrix} \in \mathbb{F}^{m\times n}$, 我们总可以将之等同于线性映射

$$A : x = (x_1, \dots, x_n)^\top \in \mathbb{F}^n \mapsto Ax = x_1\alpha_1 + \dots + x_n\alpha_n \in \mathbb{F}^m.$$

于是, 线性方程组 $Ax = b$ 有解当且仅当 $b \in \operatorname{col} A = \operatorname{im} A$, 且 x 是 $Ax = 0$ 的解当且仅当 $x \in \ker A$. 若 $Ax_0 = b$, 则(非齐次方程通解 = 非齐次方程特解 + 齐次方程通解)

$$\{x : Ax = b\} = \{x_0 + w : Aw = 0\} = x_0 + \ker A,$$

这里定义 $x_0 + W = \{x_0 + w : w \in W\}$. 特别地,

$$A \text{ 是单射} \iff \ker A = \{0\}.$$

利用 Gauss 消元,

$$\dim\ker A = \text{自由变量个数} = \text{所有变量个数} - \text{主元个数} = n - \operatorname{rank} A.$$

题 23. *设 $A \in \mathbb{F}^{m\times n}$, $\beta \in \operatorname{col} A \subset \mathbb{F}^m$, $\gamma \in \mathbb{F}^n$, $d \in \mathbb{F}$. 证明下述两个命题等价:*

1. 方程 $Ax = \beta$ 的任何解 $x \in \mathbb{F}^n$ 都满足 $\gamma^\top x = d$;　　2. 存在 $\alpha \in \mathbb{F}^m$ 使得 $\gamma = A^\top\alpha$ 且 $d = \beta^\top\alpha$.

证明. 1 $\iff \{x \in \mathbb{F}^n : Ax = c\beta\} \subset \{x \in \mathbb{F}^n : \gamma^\top x = cd\}, \forall c \in \mathbb{F}$　(分别讨论 $c = 0$ 和 $c \neq 0$ 情形)

$$\iff \ker\begin{pmatrix} A & \beta \end{pmatrix} = \ker\begin{pmatrix} A & \beta \\ \gamma^\top & d \end{pmatrix}$$

$$\iff \operatorname{rank}\begin{pmatrix} A & \beta \end{pmatrix} = \operatorname{rank}\begin{pmatrix} A & \beta \\ \gamma^\top & d \end{pmatrix} \iff \operatorname{rank}\begin{pmatrix} A^\top \\ \beta^\top \end{pmatrix} = \operatorname{rank}\begin{pmatrix} A^\top & \gamma \\ \beta^\top & d \end{pmatrix} \iff 2.$$

□

4 矩阵运算

4.1 基本概念

所有 m 行 n 列 $\mathbb{F}$-值矩阵构成的集合 $\mathbb{F}^{m\times n}$ 自然地成为 $\mathbb{F}$-线性空间, 并且等同于所有从 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的 $\mathbb{F}$-线性映射的集合. 下面定义矩阵乘法 (matrix multiplication): 对于矩阵 A 和 B, 它们的乘积 AB 应当理解为映射复合, 即满足

$$(AB)x = A(Bx), \quad \forall x.$$

这要求 A 的列数等于 B 的行数, 当 $A \in \mathbb{F}^{m\times n}$ 且 $B \in \mathbb{F}^{n\times p}$ 时 $AB \in \mathbb{F}^{m\times p}$. 记 p 阶单位矩阵 (identity matrix) 为

$$I_p = \begin{pmatrix} e_1 & \dots & e_p \end{pmatrix} = (\mathbb{1}_{[i=j]})_{1\le i,j\le p},$$

它等同于 $\mathbb{F}^p$ 上的恒同映射, 当阶数自明时简记为 I. 若 $B = \begin{pmatrix} \beta_1 & \dots & \beta_p \end{pmatrix}$, 则

$$(AB)e_j = A(Be_j) = A\beta_j\ (\forall j) \implies AB = \begin{pmatrix} A\beta_1 & \dots & A\beta_p \end{pmatrix} = \left(\sum_{k=1}^n A[i,k]B[k,j]\right)_{1\le j\le p}^{1\le i\le m}.$$

https://pytorch.org/blog/inside-the-matrix/

题 24. 设 $V=[n]=\{1,\dots,n\}$, 则 $E\subset V\times V$ 表示 V 中点之间的有向边. 定义

$$A=(\mathbb{1}_{[(i,j)\in E]})_{i,j\in V},$$

称为 (V,E) 的邻接矩阵. 不难将 A 推广为非负整数值矩阵.

1. 解释 A^ℓ 的 (i,j)-元 $A^\ell[i,j]$ 的含义;
2. 求 $i,j\in V$ 的图上距离, 即从 i 到 j 的最短路径的长度.

解. 1. 我们归纳地说明 $A^\ell[i,j]$ 表示从 i 到 j 的长度为 ℓ 的路径的数目. 当 $\ell=1$ 时, 这是显然的. 当 $\ell>1$ 时, $A^\ell[i,j]=\sum_k A^{\ell-1}[i,k]A[k,j]$, 其中 $A^{\ell-1}[i,k]$ 表示从 i 到 k 的长度为 $\ell-1$ 的路径的数目, $A[k,j]$ 表示从 k 到 j 的长度为 1 的路径的数目, 二者相乘后遍历 k 求和即得从 i 到 j 的长度为 ℓ 的路径的数目.

2. 从 i 到 j 存在长度为 ℓ 的路径当且仅当 $A^\ell[i,j]>0$, 于是所求为 $d(i,j)=\min\{\ell : A^\ell[i,j]>0\}$. □

题 25. *(Pauli 矩阵) 求下列矩阵的两两乘积:*

$$\sigma_0=\begin{pmatrix}1&0\\0&1\end{pmatrix},\ \sigma_1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\ \sigma_2=\begin{pmatrix}0&-\mathbf{i}\\\mathbf{i}&0\end{pmatrix},\ \sigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

解. 不妨略过 σ_0. 用第 j 行第 k 列表示 $\sigma_j\sigma_k$, 则有

σ_0	$\mathbf{i}\sigma_3$	$-\mathbf{i}\sigma_2$
$-\mathbf{i}\sigma_3$	σ_0	$\mathbf{i}\sigma_1$
$\mathbf{i}\sigma_2$	$-\mathbf{i}\sigma_1$	σ_0

. □

题 26. 验证如果 $A=(a_{ij})$ 和 $B=(b_{ij})$ 是, 则其乘积 $AB=\left(\sum_k a_{ik}b_{kj}\right)$ 仍是:

1. 对角阵, 即只有对角元非零的方阵;
2. 上三角阵, 即主对角线下方均为零的矩阵; 严格上三角阵, 即主对角元均为零的上三角阵;
3. 下三角阵, 即主对角线上方均为零的矩阵; 严格下三角阵, 即主对角元均为零的下三角阵;
4. 非负阵, 即各项均为非负实数的矩阵;
5. 右随机阵, 即每行行和均为 1 的非负阵;
6. 左随机阵, 即每列列和均为 1 的非负阵;
7. 置换阵, 即每行或每列元素恰有一个 1 且其余为 0 的方阵;
8. 循环阵, 即每行均为前一行各元素依次右移一个位置的方阵.

证明. 考虑2和5:

2. $$\sum_k a_{ik}b_{kj}=a_{ii}b_{ij}+\sum_{k<i}\underbrace{a_{ik}}_{=0}b_{kj}+\sum_{k>i}a_{ik}\underbrace{b_{kj}}_{=0},\quad \forall i\ge j.$$

5. $$AB\mathbf{1}=A\mathbf{1}=\mathbf{1}\quad\text{for}\quad \mathbf{1}=(1,1,\cdots,1)^\top.$$

其余从略 (留给读者). □

利用线性映射性质, 矩阵乘法满足结合律和分配律. 如果矩阵 A 和 B 满足 $AB = BA$, 则称它们可交换 (commuting); 即使 $AB = 0$, 也可能 $BA \neq 0$, 例如 $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ 和 $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. 通过直接计算可知, 矩阵转置满足 $(AB)^\top = B^\top A^\top$.

矩阵的初等变换可由矩阵乘法实现: 记 $E_{k\ell|n} = (\mathbb{1}_{[i=k;\, j=\ell]})_{1\le i,j\le n}$, 则 $A \in \mathbb{F}^{m\times n}$

1. 交换第 k 行和第 ℓ 行得到 $(I_m - E_{kk|m} - E_{\ell\ell|m} + E_{k\ell|m} + E_{\ell k|m})A$;
2. 第 k 行变为原先的 c 倍得到 $(I_m - E_{kk|m} + cE_{kk|m})A$;
3. 第 k 行加上 c 倍第 ℓ 行得到 $(I_m + cE_{k\ell|m})A$;

1. 交换第 k 列和第 ℓ 列得到 $A(I_n - E_{kk|n} - E_{\ell\ell|n} + E_{k\ell|n} + E_{\ell k|n})$;
2. 第 k 列变为原先的 c 倍得到 $A(I_n - E_{kk|n} + cE_{kk|n})$;
3. 第 k 列加上 c 倍第 ℓ 列得到 $A(I_n + cE_{\ell k|n})$.

作为2&2的推论, 用 $\operatorname{diag}(c_1,\ldots,c_n)$ 表示元素为 $c_1,\ldots,c_n$ 的 n 阶对角 (diagonal) 阵, 则 $A \in \mathbb{F}^{m\times n}$

- 第 $1,\ldots,m$ 行分别变为原先的 $c_1,\ldots,c_m$ 倍得到 $\operatorname{diag}(c_1,\ldots,c_m)A$;
- 第 $1,\ldots,n$ 列分别变为原先的 $c_1,\ldots,c_n$ 倍得到 $A\operatorname{diag}(c_1,\ldots,c_n)$.

任给 $A \in \mathbb{F}^{m\times n}$,

- $\begin{pmatrix} A & I_m \end{pmatrix} \xrightarrow{\text{行变换}} \begin{pmatrix} PA & P \end{pmatrix}$, 当 $PA = I_m$ 时即得 P 为 A 的左逆;
- $\begin{pmatrix} A \\ I_n \end{pmatrix} \xrightarrow{\text{列变换}} \begin{pmatrix} AQ \\ Q \end{pmatrix}$, 当 $AQ = I_n$ 时即得 Q 为 A 的右逆.

双射有逆映射, 对应于矩阵则是可逆 (invertible) 矩阵有逆 (inverse) 矩阵; 注意左逆和右逆都存在则必相等. 不难发现, 矩阵存在左 (右) 逆当且仅当它列 (行) 满秩, 所以有限阶矩阵可逆当且仅当它是满秩方阵. 对方阵 A 定义同阶的方阵 $\operatorname{adj} A$, 称为 A 的经典伴随 (classical adjoint), 其第 i 行第 j 列是 A 的 (j,i)-代数余子式, 则 $A \operatorname{adj} A = (\det A)I$. 当 A 是满秩方阵时,

$$A^{-1} = (\det A)^{-1} \operatorname{adj} A.$$

题 27. *设 A 和 B 是同阶方阵. 证明: 若 $A + B = AB$, 则 $AB = BA$.*

证明. 依题意, $(I - A)(I - B) = I - A - B + AB = I$, 故 $I = (I - B)(I - A) = I - B - A + BA$, 从而 $BA = A + B = AB$. (对于有限阶方阵, 左逆 = 右逆.) □

题 28. *(Woodbury 求逆公式) 设 A 和 B 是 (未必同阶的) 可逆矩阵, 且矩阵 U 和 V' 使 $B - V'A^{-1}U$ 可逆. 证明*

$$(A - UB^{-1}V')^{-1} = A^{-1} + A^{-1}U(B - V'A^{-1}U)^{-1}V'A^{-1}.$$

形式上 $\sum_{k=0}^\infty (UV')^k = I + U\sum_{k=0}^\infty (V'U)^k V'$. 华罗庚恒等式: $A^{-1} = (A + B)^{-1} + (A + AB^{-1}A)^{-1}$. (特例) Sherman–Morrison 公式: $(A + uv')^{-1} = A^{-1} - \dfrac{A^{-1}uv'A^{-1}}{1 + v'A^{-1}u}$, 其中 u 是列向量, v' 是行向量.

证明. 直接计算可得

$$\begin{aligned}
&(A - UB^{-1}V')(A^{-1} + A^{-1}U(B - V'A^{-1}U)^{-1}V'A^{-1}) - I \\
&= -UB^{-1}V'A^{-1} + U(B - V'A^{-1}U)^{-1}V'A^{-1} - UB^{-1}V'A^{-1}U(B - V'A^{-1}U)^{-1}V'A^{-1} \\
&= UB^{-1}\big(-I + \underbrace{B(B - V'A^{-1}U)^{-1} - V'A^{-1}U(B - V'A^{-1}U)^{-1}}_{=I}\big)V'A^{-1} = 0.
\end{aligned}$$

□

4.2 矩阵乘积的秩与行列式

题 29. *设 $A \in \mathbb{F}^{m\times n}$, $X \in \mathbb{F}^{n\times p}$, $B \in \mathbb{F}^{m\times p}$ 满足 $AX = B$. 证明:*

1. *A 的列向量能线性表出 B 的列向量, 从而 $\operatorname{rank} A \geq \operatorname{rank} B$;*
2. *若 X 行满秩, 则 B 的列向量也能线性表出 A 的列向量, 从而 $\operatorname{rank} A = \operatorname{rank} B$.*

证明. 记 $A = \begin{pmatrix}\alpha_1 & \dots & \alpha_n\end{pmatrix}$, $B = \begin{pmatrix}\beta_1 & \dots & \beta_p\end{pmatrix}$ 和 $X = (x_{ij})_{1\leq j\leq p}^{1\leq i\leq n}$.

1. 由 $AX = B$ 可知 $\sum_i x_{ij}\alpha_i = \beta_j$ $(\forall j)$, 所以列空间满足 $\operatorname{col} A \supset \operatorname{col} B$, 二者维数即为所求.
2. 利用1, 只需存在 $Y \in \mathbb{F}^{p\times n}$ 使得 $BY = A$. 因为 $\operatorname{rank} X = n$, 所以 X 的列向量张成 $\mathbb{F}^n$, 特别地有 $y_j \in \mathbb{F}^p$ 使
$$Xy_j = (\mathbb{1}_{[i=j]})_{1\leq i\leq n}, \quad j = 1, \dots, n.$$
取 $Y = \begin{pmatrix}y_1 & \dots & y_n\end{pmatrix}$, 则 $XY = I_n$, 进而 $BY = AXY = A$.

从映射的观点看, $B : \mathbb{F}^p \xrightarrow{X} \mathbb{F}^n \xrightarrow{A} \mathbb{F}^m$ 只能体现 A 的部分信息, 当 X 不造成信息压缩才能取等. □

题 30. *(Putnam 全美大学生数学竞赛 2009 年 A3)* 将 $\cos 1, \cos 2, \dots, \cos(n^2)$ 从左到右从上到下排成 $n \geq 3$ 阶方阵, 求其行列式.

解. 可以看出
$$\Big(\cos((i-1)n+j)\Big)_{1\leq i,j\leq n} = \Big(\cos((i-1)n)\Big)^{1\leq i\leq n}\Big(\cos j\Big)_{1\leq j\leq n} - \Big(\sin((i-1)n)\Big)^{1\leq i\leq n}\Big(\sin j\Big)_{1\leq j\leq n}$$
秩 $\leq 2 < n$, 因此所求行列式为零. □

题 31. [ix] *若 $\mathbb{R}^n$ 中 $n+2$ 个点两两之间的距离均为奇数, 则 $n+2 \equiv 0 \pmod{16}$.*

证明. 不妨设此 $n+2$ 个点为 $0, p_1, p_2, \dots, p_{n+1} \in \mathbb{R}^n$, 则 $|p_i|$ 和 $|p_i - p_j|$ 均为奇数. 注意到
$$(4k \pm 1)^2 = 16k^2 \pm 8k + 1,$$
对 $i \neq j$ 有 $2|p_i|^2 \equiv 2 \pmod{16}$ 和 $2p_i^\top p_j = |p_i|^2 + |p_j|^2 - |p_i - p_j|^2 \equiv 1 \pmod 8$. 令
$$P = \begin{pmatrix}p_1 & p_2 & \dots & p_{n+1}\end{pmatrix} \in \mathbb{R}^{n\times(n+1)}, \quad \mathbf{1} = (1, \cdots, 1)^\top \in \mathbb{R}^{n+1},$$
$$A = (a_{ij})_{1\leq i,j\leq n+1} = 2P^\top P - \mathbf{1}\mathbf{1}^\top = (2p_i^\top p_j - 1)_{1\leq i,j\leq n+1} \in \mathbb{R}^{(n+1)\times(n+1)}.$$
显然 $\operatorname{rank}(A + \mathbf{1}\mathbf{1}^\top) \leq n$, 所以 $\det(A + \mathbf{1}\mathbf{1}^\top) = 0$. 往证 $\det(A + \mathbf{1}\mathbf{1}^\top) \equiv n+2 \pmod{16}$. 根据**题7**,
$$\det(A + \mathbf{1}\mathbf{1}^\top) = \det A + \sum_{i=1}^{n+1} A_{ii} + \sum_{i\neq j} A_{ij},$$
其中 A_{ij} 是 A 的 (i,j)-代数余子式. 当 $j \neq k$ 且 $\ell \neq m$ 时, $a_{jk}a_{\ell m} \equiv 0 \pmod{64}$, 于是由行列式的 Leibniz 公式可得 $\det(A + \mathbf{1}\mathbf{1}^\top) \pmod{16} \equiv 1 + (n+1) - \sum_{i\neq j} a_{ji} = n + 2 - 2\sum_{i<j} a_{ij} \equiv n+2$. □

题 32. *(Frobenius 秩不等式) 设以下矩阵乘法有意义, 证明*
$$\operatorname{rank} AB + \operatorname{rank} BC \leq \operatorname{rank} ABC + \operatorname{rank} B.$$
(特例) Sylvester 秩不等式: $\operatorname{rank} A + \operatorname{rank} C \leq \operatorname{rank} AC + n$, $\forall A \in \mathbb{F}^{m\times n}$, $C \in \mathbb{F}^{n\times p}$.

[ix] Graham, R.L.; Rothschild, B.L.; Straus, E.G. (1974). Are there $n+2$ points in E^n with odd integral distances?.

证明. 作行列变换可得 (取等的充要条件为 $B = BCX - YAB$, 见题40.)

$$\begin{aligned}\operatorname{rank} AB + \operatorname{rank} BC &\le \operatorname{rank}\begin{pmatrix} BC & B \\ 0 & AB \end{pmatrix} \\ &= \operatorname{rank}\begin{pmatrix} BC & B \\ -ABC & 0 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} 0 & B \\ -ABC & 0 \end{pmatrix} = \operatorname{rank} ABC + \operatorname{rank} B.\end{aligned}$$

事实上, $\operatorname{im} B / \operatorname{im} BC \xrightarrow{A} \operatorname{im} AB / \operatorname{im} ABC$ 是良好定义的满射, 从而

$$\begin{aligned}\operatorname{rank} B - \operatorname{rank} BC &= \dim(\operatorname{im} B / \operatorname{im} BC) \\ &\ge \dim(\operatorname{im} AB / \operatorname{im} ABC) = \operatorname{rank} AB - \operatorname{rank} ABC.\end{aligned}$$

□

作为应用, $\begin{cases} \operatorname{rank} AB = \operatorname{rank} B \implies \operatorname{rank} ABC = \operatorname{rank} BC; \\ \operatorname{rank} BC = \operatorname{rank} B \implies \operatorname{rank} ABC = \operatorname{rank} AB. \end{cases}$

给定 $A \in \mathbb{F}^{n\times n}$, 考察 $X \in \mathbb{F}^{n\times n} \mapsto \left|AX\right| \in \mathbb{F}$, 则

$$\left|AX\right| \overset{\text{题4}}{=\!=} \left|AI_n\right|\left|X\right| = \left|A\right|\left|X\right|.$$

设 $f_1(x) = 1$ 和 $f_j(x) = \prod_{i=1}^{j-1}(x - x_i) = \sum_{i=1}^{j} a_{ij}x^{i-1}$, $j = 2, 3, \ldots, n$, 则

$$\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix}\begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \ddots & a_{2n} \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} f_1(x_1) & 0 & \cdots & 0 \\ f_1(x_2) & f_2(x_2) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ f_1(x_n) & f_2(x_n) & \ldots & f_n(x_n) \end{pmatrix},$$

进而两端求行列式可见 Vandermonde 行列式公式

$$\left|\left(x_i^{j-1}\right)_{1\le i,j\le n}\right| = \prod_{j=1}^{n} f_j(x_j) = \prod_{1\le i<j\le n}(x_j - x_i).$$

题 33. *(丘成桐大学生数学竞赛 2012 年笔试代数团体第 1 题) 证明 Cauchy 行列式*

$$\left|\left(\frac{1}{a_i + b_j}\right)_{1\le i,j\le n}\right| = \frac{\prod_{1\le i<j\le n}(a_j - a_i)(b_j - b_i)}{\prod_{i=1}^{n}\prod_{j=1}^{n}(a_i + b_j)}.$$

证明. 设 $p(x) = \prod_{i=1}^{n}(x + b_i)$ 和 $q_j(x) = \prod_{i\ne j}(x + b_i) = \sum_{i=1}^{n} c_{ij}x^{i-1}$, $j = 1, 2, \ldots, n$, 则

$$\operatorname{diag}\big(p(a_1), \ldots, p(a_n)\big)\left(\tfrac{1}{a_i+b_j}\right)_{1\le i,j\le n} = \left(q_j(a_i)\right)_{1\le i,j\le n} = \left(a_i^{j-1}\right)_{1\le i,j\le n}\left(c_{ij}\right)_{1\le i,j\le n},$$

且

$$\operatorname{diag}\big(q_1(-b_1), \ldots, q_n(-b_n)\big) = \left(q_j(-b_i)\right)_{1\le i,j\le n} = \left((-b_i)^{j-1}\right)_{1\le i,j\le n}\left(c_{ij}\right)_{1\le i,j\le n}.$$

两式两端求行列式, 利用乘法性质和 Vandermonde 行列式公式即得所欲.
也可以暴力递推 (见https://zhuanlan.zhihu.com/p/111537288), 或者视为多元多项式的比值后只需确定系数 (见https://www.zhihu.com/answer/128049228和https://www.cnblogs.com/xixifeng/p/3932238.html). □

当 $m < n$ 时, 对 $A \in \mathbb{F}^{m\times n}$ 和 $B \in \mathbb{F}^{n\times m}$ 成立如下 Cauchy–Binet 公式

$$\left|AB\right| = \sum_{\mathcal{I}\in\binom{[n]}{m}}\left|A\big[[m], \mathcal{I}\big]\right|\left|B\big[\mathcal{I}, [m]\big]\right|.$$

4.3 分块矩阵

分块矩阵 (block matrix) 指的是形如 $(A_{k\ell})_{1\le \ell\le q}^{1\le k\le p}$ 的矩阵, 其中 $A_{k\ell}$ 是 $m_k\times n_\ell$ 矩阵. 特别地, $\operatorname{diag}(A_1,\dots,A_p)=\begin{pmatrix}A_1 & & \\ & \ddots & \\ & & A_p\end{pmatrix}$ 称为分块对角阵 (block diagonal matrix). 如果有意义, 那么

$$(A_{k\ell})_{1\le \ell\le q}^{1\le k\le p}(B_{st})_{1\le t\le r}^{1\le s\le q}=\left(\sum_{s=1}^{q}A_{ks}B_{st}\right)_{1\le t\le r}^{1\le k\le p}.$$

题 34. *(三角分解) 任给矩阵 A, 证明存在置换阵 P、上三角阵 U 和对角元均为 1 的下三角阵 L 使得*

$$PA=LU. \qquad \textit{(这蕴涵 Gauss 消元的矩阵形式.)}$$

证明. 对 A 的行数 m 归纳. 当 $m=1$ 时, 取 $P=L=(1)$ 和 $U=A$ 即可. 假设结论对 $m-1$ 成立, 下面考虑 m 行的矩阵 A. 容易找出置换阵 $P_0\in\mathbb{F}^{m\times m}$ 和向量 $\ell\in\mathbb{F}^{m-1}$ 满足

$$\begin{pmatrix}1 & 0\\ -\ell & I_{m-1}\end{pmatrix}P_0A=\begin{pmatrix}a & \alpha^\top\\ 0 & A_1\end{pmatrix},$$

其中 $a\in\mathbb{F}$, $\alpha\in\mathbb{F}^{n-1}$, $A_1\in\mathbb{F}^{(m-1)\times(n-1)}$. 根据归纳假设, 可以得到置换阵 P_1、上三角阵 U_1 和对角元均为 1 的下三角阵 L_1 使得 $P_1A_1=L_1U_1$. 令 $P=\begin{pmatrix}1 & 0\\ 0 & P_1\end{pmatrix}P_0$, $L=\begin{pmatrix}1 & 0\\ P_1\ell & L_1\end{pmatrix}$, $U=\begin{pmatrix}a & \alpha^\top\\ 0 & U_1\end{pmatrix}$, 则

$$PA=\begin{pmatrix}1 & 0\\ 0 & P_1\end{pmatrix}\begin{pmatrix}1 & 0\\ \ell & I_{m-1}\end{pmatrix}\begin{pmatrix}a & \alpha^\top\\ 0 & P_1^{-1}L_1U_1\end{pmatrix}=LU.$$

依**题**26, P 是置换阵, 即证所欲. □

如果有意义, 那么 $(LU)^{-1}=U^{-1}L^{-1}$ 和 $\det(LU)=(\det L)(\det U)$ 计算起来很便捷.

矩阵打洞 (matrix holing)[x] 是处理分块矩阵的常用技巧, 旨在得到分块三角阵以便分析行列式、秩以及逆, 在**题**32中已初见端倪. 一般地,

- $\begin{pmatrix}I & 0\\ -A_{21}A_{11}^{-1} & I\end{pmatrix}\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}=\begin{pmatrix}A_{11} & A_{12}\\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12}\end{pmatrix}$;
- $\begin{pmatrix}I & -A_{12}A_{22}^{-1}\\ 0 & I\end{pmatrix}\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}=\begin{pmatrix}A_{11}-A_{12}A_{22}^{-1}A_{21} & 0\\ A_{21} & A_{22}\end{pmatrix}$;
- $\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix}I & -A_{11}^{-1}A_{12}\\ 0 & I\end{pmatrix}=\begin{pmatrix}A_{11} & 0\\ A_{12} & A_{22}-A_{21}A_{11}^{-1}A_{12}\end{pmatrix}$;
- $\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix}I & 0\\ -A_{22}^{-1}A_{21} & I\end{pmatrix}=\begin{pmatrix}A_{11}-A_{12}A_{22}^{-1}A_{21} & A_{12}\\ 0 & A_{22}\end{pmatrix}$.

设 $M=(M[i,j])_{i,j\in[n]}$, 其中 $[n]=\{1,\dots,n\}$. 对于 $\mathcal{I},\mathcal{J}\subset[n]$, 记 $M[\mathcal{I},\mathcal{J}]=(M[i,j])_{j\in\mathcal{J}}^{i\in\mathcal{I}}$ 和 $\mathcal{I}^{\complement}=[n]\setminus\mathcal{I}$. 当 $M[\mathcal{I},\mathcal{J}]$ 可逆, 定义 M 中 $M[\mathcal{I},\mathcal{J}]$ 的 Schur 补 (Schur complement) 为

$$M/M[\mathcal{I},\mathcal{J}]=M[\mathcal{I}^{\complement},\mathcal{J}^{\complement}]-M[\mathcal{I}^{\complement},\mathcal{J}]M[\mathcal{I},\mathcal{J}]^{-1}M[\mathcal{I},\mathcal{J}^{\complement}].$$

[x] 孟道骥; 王立云. (2006). 打洞技巧.

令 $\mathscr{S}(\mathcal{I}) = \operatorname{inv}(\mathcal{I}, \mathcal{I}^\complement) = (\sum_{i\in\mathcal{I}} i) - |\mathcal{I}|(|\mathcal{I}|+1)/2$, 则

$$\det M = (-1)^{\mathscr{S}(\mathcal{I})}(-1)^{\mathscr{S}(\mathcal{J})} \begin{vmatrix} M[\mathcal{I},\mathcal{J}] & M[\mathcal{I},\mathcal{J}^\complement] \\ M[\mathcal{I}^\complement,\mathcal{J}] & M[\mathcal{I}^\complement,\mathcal{J}^\complement] \end{vmatrix} = (-1)^{\mathscr{S}(\mathcal{I})+\mathscr{S}(\mathcal{J})}(\det M[\mathcal{I},\mathcal{J}])(\det M/M[\mathcal{I},\mathcal{J}]).$$

题 35. *(主主元变换[xi])* 设 $A = (A[i,j])_{i,j\in[n]} \in \mathbb{F}^{n\times n}$, 其中 $[n] = \{1,\dots,n\}$. 记 $A[\mathcal{I},\mathcal{J}] = (A[i,j])_{j\in\mathcal{J}}^{i\in\mathcal{I}}$, 其中 $\mathcal{I}$ 和 $\mathcal{J}$ 是 $[n]$ 的任意子集. 当 $A[\mathcal{I},\mathcal{I}]$ 可逆, 定义 $\operatorname{ppt}(A,\mathcal{I}) = B \in \mathbb{F}^{n\times n}$ 适合

$$\begin{pmatrix} B[\mathcal{I},\mathcal{I}] & B[\mathcal{I},\mathcal{I}^\complement] \\ B[\mathcal{I}^\complement,\mathcal{I}] & B[\mathcal{I}^\complement,\mathcal{I}^\complement] \end{pmatrix} = \begin{pmatrix} A[\mathcal{I},\mathcal{I}]^{-1} & -A[\mathcal{I},\mathcal{I}]^{-1}A[\mathcal{I},\mathcal{I}^\complement] \\ A[\mathcal{I}^\complement,\mathcal{I}]A[\mathcal{I},\mathcal{I}]^{-1} & A/A[\mathcal{I},\mathcal{I}] \end{pmatrix},$$

其中 $A/A[\mathcal{I},\mathcal{I}] = A[\mathcal{I}^\complement,\mathcal{I}^\complement] - A[\mathcal{I}^\complement,\mathcal{I}]A[\mathcal{I},\mathcal{I}]^{-1}A[\mathcal{I},\mathcal{I}^\complement]$ 是 $A[\mathcal{I},\mathcal{I}]$ 的 *Schur* 补, $\mathcal{I}^\complement = [n]\setminus\mathcal{I}$. 证明:

1. 若 $B = \operatorname{ppt}(A,\mathcal{I})$, 则 $\begin{pmatrix} B[\mathcal{I},\mathcal{I}] & B[\mathcal{I},\mathcal{I}^\complement] \\ B[\mathcal{I}^\complement,\mathcal{I}] & B[\mathcal{I}^\complement,\mathcal{I}^\complement] \end{pmatrix} = \begin{pmatrix} I_{|\mathcal{I}|} & 0 \\ A[\mathcal{I}^\complement,\mathcal{I}] & A[\mathcal{I}^\complement,\mathcal{I}^\complement] \end{pmatrix}\begin{pmatrix} A[\mathcal{I},\mathcal{I}] & A[\mathcal{I},\mathcal{I}^\complement] \\ 0 & I_{n-|\mathcal{I}|} \end{pmatrix}^{-1}$;
2. 若 $C \in \mathbb{F}^{n\times n}$ 满足 $\forall x,y\in\mathbb{F}^n$ 有
$$Ax = y \iff \begin{pmatrix} C[\mathcal{I},\mathcal{I}] & C[\mathcal{I},\mathcal{I}^\complement] \\ C[\mathcal{I}^\complement,\mathcal{I}] & C[\mathcal{I}^\complement,\mathcal{I}^\complement] \end{pmatrix}\begin{pmatrix} y[\mathcal{I}] \\ x[\mathcal{I}^\complement] \end{pmatrix} = \begin{pmatrix} x[\mathcal{I}] \\ y[\mathcal{I}^\complement] \end{pmatrix},$$
则 $C = \operatorname{ppt}(A,\mathcal{I})$; 逆命题同样成立;
3. 若 $\mathcal{I}_1, \mathcal{I}_2$ 是 $[n]$ 的不交子集, 则 $\operatorname{ppt}(\operatorname{ppt}(A,\mathcal{I}_1),\mathcal{I}_2) = \operatorname{ppt}(A,\mathcal{I}_1\cup\mathcal{I}_2)$;
4. 若 $A[k,k]\neq 0$, 则 $\operatorname{ppt}(A,\{k\}) = \operatorname{swp}(A,k)$, 其中 swp 见**算法2**. *(常用于线性回归.)*

Algorithm 2: Sweep

Input: $A = (A[i,j])_{i,j\in[n]} \in \mathbb{F}^{n\times n}$, $k\in[n]$

begin
- $A[\{k\}^\complement, k] \leftarrow A[\{k\}^\complement, k]/A[k,k]$;
- **for** $j \in [n]\setminus\{k\}$ **do**
  - $A[\{k\}^\complement, j] \leftarrow A[\{k\}^\complement, j] - A[k,j]A[\{k\}^\complement, k]$;
  - $A[k,j] \leftarrow -A[k,j]/A[k,k]$;
- $A[k,k] \leftarrow 1/A[k,k]$;

Output: $\operatorname{swp}(A,k) = A$

证明. 1. 不难算出 $\begin{pmatrix} A[\mathcal{I},\mathcal{I}] & A[\mathcal{I},\mathcal{I}^\complement] \\ 0 & I_{n-|\mathcal{I}|} \end{pmatrix}^{-1} = \begin{pmatrix} A[\mathcal{I},\mathcal{I}]^{-1} & -A[\mathcal{I},\mathcal{I}]^{-1}A[\mathcal{I},\mathcal{I}^\complement] \\ 0 & I_{n-|\mathcal{I}|} \end{pmatrix}$.

2. 我们有
$$Ax = y \iff \begin{pmatrix} A[\mathcal{I},\mathcal{I}] & A[\mathcal{I},\mathcal{I}^\complement] & -I_{|\mathcal{I}|} & 0 \\ A[\mathcal{I}^\complement,\mathcal{I}] & A[\mathcal{I}^\complement,\mathcal{I}^\complement] & 0 & -I_{n-|\mathcal{I}|} \end{pmatrix}\begin{pmatrix} x[\mathcal{I}] \\ x[\mathcal{I}^\complement] \\ y[\mathcal{I}] \\ y[\mathcal{I}^\complement] \end{pmatrix} = 0,$$

$$\begin{pmatrix} C[\mathcal{I},\mathcal{I}] & C[\mathcal{I},\mathcal{I}^\complement] \\ C[\mathcal{I}^\complement,\mathcal{I}] & C[\mathcal{I}^\complement,\mathcal{I}^\complement] \end{pmatrix}\begin{pmatrix} y[\mathcal{I}] \\ x[\mathcal{I}^\complement] \end{pmatrix} = \begin{pmatrix} x[\mathcal{I}] \\ y[\mathcal{I}^\complement] \end{pmatrix} \iff \begin{pmatrix} -I_{|\mathcal{I}|} & C[\mathcal{I},\mathcal{I}^\complement] & C[\mathcal{I},\mathcal{I}] & 0 \\ 0 & C[\mathcal{I}^\complement,\mathcal{I}^\complement] & C[\mathcal{I}^\complement,\mathcal{I}] & -I_{n-|\mathcal{I}|} \end{pmatrix}\begin{pmatrix} x[\mathcal{I}] \\ x[\mathcal{I}^\complement] \\ y[\mathcal{I}] \\ y[\mathcal{I}^\complement] \end{pmatrix} = 0,$$

[xi] Tsatsomeros, M.J. (2000). Principal pivot transforms: properties and applications.

二者解空间相同当且仅当系数矩阵只差了若干行变换 (有相同的简化行阶梯形矩阵), 即存在

$$P \in \mathbb{F}^{|\mathcal{I}|\times|\mathcal{I}|},\ Q \in \mathbb{F}^{|\mathcal{I}|\times(n-|\mathcal{I}|)},\ R \in \mathbb{F}^{(n-|\mathcal{I}|)\times|\mathcal{I}|},\ S \in \mathbb{F}^{(n-|\mathcal{I}|)\times(n-|\mathcal{I}|)}$$

使得 $\left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right) \in \mathbb{F}^{n\times n}$ 可逆并且

$$\begin{pmatrix} P & Q \\ R & S \end{pmatrix}\begin{pmatrix} A[\mathcal{I},\mathcal{I}] & A[\mathcal{I},\mathcal{I}^{\complement}] & -I_{|\mathcal{I}|} & 0 \\ A[\mathcal{I}^{\complement},\mathcal{I}] & A[\mathcal{I}^{\complement},\mathcal{I}^{\complement}] & 0 & -I_{n-|\mathcal{I}|} \end{pmatrix} = \begin{pmatrix} -I_{|\mathcal{I}|} & C[\mathcal{I},\mathcal{I}^{\complement}] & C[\mathcal{I},\mathcal{I}] & 0 \\ 0 & C[\mathcal{I}^{\complement},\mathcal{I}^{\complement}] & C[\mathcal{I}^{\complement},\mathcal{I}] & -I_{n-|\mathcal{I}|} \end{pmatrix}.$$

考察最后 $(n-|\mathcal{I}|)$ 列可知 $Q=0$ 和 $S=I_{n-|\mathcal{I}|}$, 再比较前 $|\mathcal{I}|$ 列则有 $P=-A[\mathcal{I},\mathcal{I}]^{-1}$ 和 $R=-A[\mathcal{I}^{\complement},\mathcal{I}]A[\mathcal{I},\mathcal{I}]^{-1}$, 于是由剩下 n 列可得 $C=\operatorname{ppt}(A,\mathcal{I})$. 反之, 已知 $C=\operatorname{ppt}(A,\mathcal{I})$, 如前定义 P,Q,R,S 即可.

3. 利用2, 从略.

4. 可以直接计算验证. □

题 36. *(Cochran 定理) 设 $A_1,\dots,A_m \in \mathbb{F}^{n\times n}$ 满足 $A_1+\cdots+A_m=I_n$. 证明: 若*

$$\operatorname{rank} A_1+\cdots+\operatorname{rank} A_m = n,$$

则 $A_i^2=A_i$ 且 $A_iA_j=0$, $\forall i\neq j$. 满足 $A^2=A$ 的矩阵 A 称为幂等 (idempotent) 的.

证明. 应用初等行列变换可得

$$\begin{pmatrix} 0 & & & \\ & A_1 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix} \to \begin{pmatrix} 0 & A_1 & \dots & A_m \\ & A_1 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix} \to \begin{pmatrix} I_n & A_1 & \cdots & A_m \\ A_1 & A_1 & & \\ \vdots & & \ddots & \\ A_m & & & A_m \end{pmatrix}$$
$$\to \begin{pmatrix} I_n & A_1 & \cdots & A_m \\ 0 & A_1-A_1^2 & \cdots & -A_1A_m \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -A_mA_1 & \cdots & A_m-A_m^2 \end{pmatrix} \to \begin{pmatrix} I_n & 0 & \cdots & 0 \\ 0 & A_1-A_1^2 & \cdots & -A_1A_m \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -A_mA_1 & \cdots & A_m-A_m^2 \end{pmatrix},$$

而分块对角阵的秩是对角块的秩之和, 故只能有 $\begin{pmatrix} A_1-A_1^2 & \cdots & -A_1A_m \\ \vdots & \ddots & \vdots \\ -A_mA_1 & \cdots & A_m-A_m^2 \end{pmatrix}=0.$ □

分块矩阵能优雅地给出许多结论[xii], 值得反复揣摩.

4.4 正交阵

由于现实世界常通过 Euclid 空间建模, 我们对实数域上的一类特殊的矩阵稍加关注.

任给 $A \in \mathbb{R}^{n\times n}$, 下述性质等价:

1. (保持长度) $|Ax|=|x|$, $\forall x \in \mathbb{R}^n$;

[xii] Garcia, S.R.; Horn, R.A. (2020). Block matrices in linear algebra.

2. (保持内积) $(Ax)^\top(Ay) = x^\top y,\ \forall x, y \in \mathbb{R}^n$;

3. $A^\top A = I_n$, 亦即 $A^{-1} = A^\top$;

4. $A = \begin{pmatrix}\alpha_1 & \dots & \alpha_n\end{pmatrix}$ 满足 $\alpha_i^\top \alpha_j = \mathbb{1}_{[i=j]},\ \forall i \neq j$.

这样的 A 称为正交 (orthogonal) 矩阵, 其列向量构成 $\mathbb{R}^n$ 的一组标准正交基 (orthonormal basis).

算法3表明任意向量组都可以标准正交化. 当 $w_1, \dots, w_r$ 标准正交, $P_W = \sum_{k=1}^r w_k w_k^\top$ 是向 $W = \operatorname{span}\{w_1, \dots, w_r\}$ 的正交投影 (projection), 满足 $P_W^2 = P_W = P_W^\top$, 并且 $P_W^\top(I - P_W) = 0$.

Algorithm 3: Gram–Schmidt Process

Input: $v_1, \dots, v_p \in \mathbb{R}^n$

Output: $w_1, \dots, w_r \in \mathbb{R}^n$ 满足 $w_i^\top w_j = \mathbb{1}_{[i=j]}$ $(\forall i \neq j)$

```
begin
    r ← 0;
    for j = 1, ..., p do
        w ← v_j − Σ_{k=1}^{r} w_k w_k^⊤ v_j;
        if w ≠ 0 then
            r ← r + 1;
            w_r ← w/|w|;
```

题 37. *(QR 分解) 设 $A \in \mathbb{R}^{n\times p}$ 秩为 r. 证明: 存在各列标准正交的矩阵 $Q \in \mathbb{R}^{n\times r}$ 和主对角元非负的上三角阵 $R \in \mathbb{R}^{r\times p}$ 使得 $A = QR$.* *(可将 Q 扩充为正交阵并在 R 下方对应地添补零.)*

当 $r = p$ 时, Q 和 R 唯一, 且 R 的对角元严格大于零.

证明. 记 A 的列向量为 $v_1, \dots, v_p$, 经**算法3**得到 $w_1, \dots, w_r$ 作为 Q 的列. 断言 $R = Q^\top A$ 是满足题意的上三角阵. 事实上, 由构造过程可得 $v_j = \sum_{k=1}^{r_j} w_k w_k^\top v_j,\ \forall j$, 其中 $r_j \le j$, 从而当 $i \ge j$ 时

$$w_i^\top v_j = w_i^\top w_{r_j} w_{r_j}^\top v_j = \mathbb{1}_{[i=j=r_j]} w_{r_j}^\top v_j.$$

设 $r = p$, 且有 QR 分解 $A = Q_1R_1 = Q_2R_2$, 其中 $Q_1, Q_2 \in \mathbb{R}^{n\times p}$ 各列标准正交. 考察可逆矩阵 $A^\top A = R_1^\top R_1 = R_2^\top R_2$, 即知 R_1 和 R_2 可逆, 于是其对角元为正. 我们有 $(R_1^\top)^{-1}R_2^\top = R_1R_2^{-1}$, 它既是下三角阵又是上三角阵 (因为三角阵乘法和求逆满足封闭性), 从而是对角元为正的对角阵, 记为 D. 由 $R_1 = DR_2$ 可得 $D^{-1} = (R_2^\top)^{-1}R_1^\top = (R_2^\top)^{-1}R_2^\top D^\top = D$, 只能有 $D = I$, 亦即 $R_1 = R_2$. 此时, $Q_1 = AR_1^{-1} = AR_2^{-1} = Q_2$. □

上述证明中用到 $A \in \mathbb{R}^{n\times p}$ 与 $A^\top A$ 秩相等, 这通过比较二者 ker 易见.

题 38. *证明 $\mathbb{R}^n$ 中最多有 $n+1$ 个两两夹钝角的向量.* *(显然最多有 n 个两两正交的非零向量.)*

证明. 易见 $\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}, \begin{pmatrix}-n\\1\\\vdots\\1\end{pmatrix}, \begin{pmatrix}1\\-n\\\vdots\\1\end{pmatrix}, \dots, \begin{pmatrix}1\\1\\\vdots\\-n\end{pmatrix}$ 两两点积为负. 假设 $v_1, v_2, \dots, v_{n+2} \in \mathbb{R}^n$ 两两夹钝角, 下面将导出矛盾. 由于 $v_1, \dots, v_{n+1}$ 线性相关, 可得不全为零的 $c_1, \dots, c_{n+1} \in \mathbb{R}$ 使 $\sum_{i=1}^{n+1} c_i v_i = 0$. 考察 $\sum_{i=1}^{n+1} c_i \underbrace{v_i^\top v_{n+2}}_{<0} = 0$, 可知 c_i 有正有负. 令

$$u = \sum_{i:c_i>0} c_i v_i = -\sum_{j:c_j<0} c_j v_j,$$

则

$$0 \le u^\top u = -\sum_{i:c_i>0}\sum_{j:c_j<0} c_i c_j \underbrace{v_i^\top v_j}_{<0} < 0,$$

明所欲证. 另有一种做法是对 $v_1, \dots, v_n$ 作 Gram–Schmidt 正交化, 归纳地看出 $\forall k = 1, 2, \dots, n$,

$$w_k = v_k - \sum_{i<k} \frac{w_i^\top v_k}{w_i^\top w_i} w_i$$

与 $v_{k+1}, \dots, v_{n+1}, v_{n+2}$ 点积为负, 那么 $v_{n+j} = \sum_{i=1}^n \frac{w_i^\top v_m}{w_i^\top w_i} w_i$ $(j = 1, 2)$ 必然点积为正. □

数值计算中常用 Householder 反射变换进行 QR 分解. 任给单位 (unit) 向量 u, 反射 (reflection)

$$H_u = I - 2uu^\top$$

显然是正交阵. 为了将 A 的第一列 v_1 对齐 $e_1 = (\mathbb{1}_{[i=1]})$, 可使用 $u_1 = (v_1 - ce_1)/|v_1 - ce_1|$ 构造反射 H_{u_1}, 其中 $c = \pm|v_1|$, 于是 $H_{u_1}A = \begin{pmatrix} ce_1 & \begin{matrix} ** \\ B \end{matrix} \end{pmatrix}$. 对右下角子矩阵重复此操作, 则有

$$Q = (H_{u_p} \cdots H_{u_2} H_{u_1})^{-1} = H_{u_1} H_{u_2} \cdots H_{u_p}.$$

当 $A = Q$ 时, 有 Cartan–Dieudonné 定理[xiii]: 任意 n 阶正交阵均可表为不超过 n 个反射的乘积.

题 39. *(Putnam 全美大学生数学竞赛 2019 年 B3) 线性映射 A 的不动点指的是满足 $Av = v$ 的向量 v. 任给正交阵 $Q \in \mathbb{R}^{n\times n}$ 和单位向量 $u \in \mathbb{R}^n$, 证明 Q 或 QH_u 有不动点.*

证明. 假设 Q 没有不动点, 则 $I - Q$ 可逆. 记 $u^\perp = \{v \in \mathbb{R}^n : u^\top v = 0\}$. 下面尝试寻找 QH_u 的不动点 $u - v$, 其中 $v \in u^\perp$ 待定. 注意到 $H_u u = -u$ 和 $H_u v = v$, 应取 $v = Au$, 其中(与 Cayley 变换有关)

$$A = (I - Q)^{-1}(I + Q).$$

于是只需说明 $Au \in u^\perp$. 因为 $A^\top = -A$ (请读者验证), 由 $u^\top Au \xlongequal{\text{转置}} u^\top A^\top u = -u^\top Au$ 即证. □

5 矩阵的相抵与相似

本章对矩阵进行初步的分类, 提出一些刻画等价类 (equivalence class) 性质的不变量 (invariant).

回顾一下等价关系 (equivalence relation), 从而能清楚等价类的概念. 集合 S 上的二元关系 R 指的是 $S \times S$ 的子集, 一般记 $(x, y) \in R$ 为 xRy. 所谓等价关系, 即满足下述性质的二元关系:

- 反身 (reflexive): $\forall x \in S,\ xRx$;
- 对称 (symmetric): $\forall x, y \in S,\ xRy \implies yRx$;
- 传递 (transitive): $\forall x, y, z \in S,\ xRy \,\&\, yRz \implies xRz$.

等价关系诱导集合的划分 (partition), 使得每一类中的元素两两等价, 而等价类两两不交.

例: 三角形的全等 (congruence) 和整数的同余 (congruence) 都是等价关系.

[xiii] Gallier, J. (2011). *Geometric methods and applications: for computer science and engineering*, 2nd ed., §8.2.

5.1 相抵与相似

记 $\mathrm{GL}_n(\mathbb{F})=\{A\in\mathbb{F}^{n\times n}: A\text{ 可逆}\}$, 称为一般线性群 (general linear group).

矩阵空间 $\mathbb{F}^{m\times n}$ 上的行相抵 (row equivalence) 是等价关系:

$$A\text{ 与 }B\text{ 行相抵}\overset{\text{定义}}{\iff}\text{存在 }P\in\mathrm{GL}_m(\mathbb{F})\text{ 使得 }PA=B.$$

矩阵行相抵于它的简化行阶梯形, 也称为行相抵标准形 (row equivalence canonical form).

矩阵空间 $\mathbb{F}^{m\times n}$ 上的相抵 (equivalence) 是等价关系:

$$A\text{ 与 }B\text{ 相抵}\overset{\text{定义}}{\iff}\text{存在 }P\in\mathrm{GL}_m(\mathbb{F})\text{ 和 }Q\in\mathrm{GL}_n(\mathbb{F})\text{ 使得 }PAQ=B.$$

当 $\operatorname{rank}A=r$ 时, A 相抵于 $\begin{pmatrix}I_r&0\\0&0\end{pmatrix}$, 称为相抵标准形 (equivalence canonical form).

矩阵空间 $\mathbb{F}^{n\times n}$ 上的相似 (similarity) 是等价关系:

$$A\text{ 与 }B\text{ 相似}\overset{\text{定义}}{\iff}\text{存在 }P\in\mathrm{GL}_n(\mathbb{F})\text{ 使得 }P^{-1}AP=B.$$

相似类中矩阵乘法封闭: $(P^{-1}A_1P)(P^{-1}A_2P)=P^{-1}A_1A_2P$, 有时能加快计算速度.

题 40. [xiv] *任给 $A\in\mathbb{F}^{m\times p}$, $B\in\mathbb{F}^{q\times n}$, $C\in\mathbb{F}^{m\times n}$, 证明:*

1. *方程 $AX-YB=C$ 有解 $X\in\mathbb{F}^{p\times n}$, $Y\in\mathbb{F}^{m\times q}$ 当且仅当 $\left(\begin{smallmatrix}A&C\\0&B\end{smallmatrix}\right)$ 与 $\left(\begin{smallmatrix}A&0\\0&B\end{smallmatrix}\right)$ 相抵;*
2. *当 $m=p$ 且 $n=q$ 时, 方程 $AX-XB=C$ 有解 $X\in\mathbb{F}^{m\times n}$ 当且仅当 $\left(\begin{smallmatrix}A&C\\0&B\end{smallmatrix}\right)$ 与 $\left(\begin{smallmatrix}A&0\\0&B\end{smallmatrix}\right)$ 相似.*

证明. 必要性较直接, 只需算出

$$\begin{pmatrix}I_m&-Y\\0&I_q\end{pmatrix}\begin{pmatrix}A&0\\0&B\end{pmatrix}\begin{pmatrix}I_p&X\\0&I_n\end{pmatrix}=\begin{pmatrix}A&AX-YB\\0&B\end{pmatrix},$$

并注意到 $\left(\begin{smallmatrix}I_p&X\\0&I_n\end{smallmatrix}\right)^{-1}=\left(\begin{smallmatrix}I_p&-X\\0&I_n\end{smallmatrix}\right)$. 下面考虑充分性.

1. 设 $\left(\begin{smallmatrix}A&C\\0&B\end{smallmatrix}\right)=Q^{-1}\left(\begin{smallmatrix}A&0\\0&B\end{smallmatrix}\right)P$, 其中 $P\in\mathbb{F}^{(p+n)\times(p+n)}, Q\in\mathbb{F}^{(m+q)\times(m+q)}$ 均可逆, 则

$$\begin{aligned}V_0&=\left\{(S,T)\in\mathbb{F}^{(p+n)\times(p+n)}\times\mathbb{F}^{(m+q)\times(m+q)}:\left(\begin{smallmatrix}A&0\\0&B\end{smallmatrix}\right)S=T\left(\begin{smallmatrix}A&0\\0&B\end{smallmatrix}\right)\right\},\\V_1&=\left\{(S,T)\in\mathbb{F}^{(p+n)\times(p+n)}\times\mathbb{F}^{(m+q)\times(m+q)}:\left(\begin{smallmatrix}A&C\\0&B\end{smallmatrix}\right)S=T\left(\begin{smallmatrix}A&0\\0&B\end{smallmatrix}\right)\right\}\end{aligned}$$

通过线性双射 $(S,T)\mapsto(P^{-1}S,Q^{-1}T)$ 同构, 进而 $\dim V_0=\dim V_1$. 对于

$$\varphi_i:\left(\left(\begin{smallmatrix}S_{11}&S_{12}\\S_{21}&S_{22}\end{smallmatrix}\right),\left(\begin{smallmatrix}T_{11}&T_{12}\\T_{21}&T_{22}\end{smallmatrix}\right)\right)\in V_i\mapsto(S_{21},S_{22},T_{21},T_{22})\in\mathbb{F}^{n\times p}\times\mathbb{F}^{n\times n}\times\mathbb{F}^{q\times m}\times\mathbb{F}^{q\times q},\quad i=0,1,$$

有 $\ker\varphi_0=\ker\varphi_1$, 以及 $\operatorname{im}\varphi_1\subset\operatorname{im}\varphi_0$. 根据

$$\dim\operatorname{im}\varphi_i=\dim V_i-\dim\ker\varphi_i,\quad i=0,1,$$

可得

$$\operatorname{im}\varphi_1=\operatorname{im}\varphi_0\ni\varphi_0(-I_{p+n},-I_{m+q})=(0,-I_n,0,-I_q),$$

于是存在 $(S,T)\in V_1$ 适合 $S_{22}=-I_n$, 那么取 $(X,Y)=(S_{12},T_{12})$ 即可.

[xiv] Flanders, H.; Wimmer, H.K. (1977). On the matrix equations $AX-XB=C$ and $AX-YB=C$.

2. 设 $\left(\begin{smallmatrix} A & C \\ 0 & B \end{smallmatrix}\right)$ 与 $\left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)$ 相似, 则

$$W_0 = \left\{T \in \mathbb{F}^{(m+n)\times(m+n)} : \left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)T = T\left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)\right\},$$
$$W_1 = \left\{T \in \mathbb{F}^{(m+n)\times(m+n)} : \left(\begin{smallmatrix} A & C \\ 0 & B \end{smallmatrix}\right)T = T\left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)\right\}$$

(请读者验证)满足 $\dim W_0 = \dim W_1$. 对于

$$\psi_i : \left(\begin{smallmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{smallmatrix}\right) \in W_i \mapsto (T_{21}, T_{22}) \in \mathbb{F}^{n\times m} \times \mathbb{F}^{n\times n}, \quad i = 0, 1,$$

有 $\ker\psi_0 = \ker\psi_1$, 以及 $\operatorname{im}\psi_1 \subset \operatorname{im}\psi_0$. 根据

$$\dim\operatorname{im}\psi_i = \dim W_i - \dim\ker\psi_i, \quad i = 0, 1,$$

可得 $\operatorname{im}\psi_1 = \operatorname{im}\psi_0 \ni \psi_0(-I_{m+n}) = (0, -I_n)$, 故存在 $T \in W_1$ 使 $T_{22} = -I_n$, 取 $X = T_{12}$ 即可.

通过分块矩阵暴算或许也能证明, 有兴趣的话不妨自行尝试. □

题 41. [xv]取 $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. 设 $\mathbb{K}^{n\times n}$ 的子空间 V 满足 $\dim V \geq pn + 1$, $p \in \mathbb{N}$. 证明存在 $A_0 \in V$ 使得

$$\operatorname{rank} A_0 \geq p + 1.$$

特别地, $\mathbb{K}^{n\times n}$ 的 $(n^2 - n + 1)$ 维子空间一定包含某个满秩矩阵 (北大信科高代 I 期中考试 (2016 秋)).

证明. 记 $r = \max\{\operatorname{rank} A : A \in V\}$. 取 $A_0 \in V$ 满足 $\operatorname{rank} A_0 = r$, 则有可逆矩阵 P 和 Q 使得 $A_0 = P\operatorname{diag}(I_r, 0)Q$. 记 $B^* = \overline{B}^\top$ 为矩阵 B 的共轭转置. 断言 V 与

$$W = \{P\left(\begin{smallmatrix} 0 & B \\ B^* & C \end{smallmatrix}\right)Q : B \in \mathbb{K}^{r\times(n-r)}, C \in \mathbb{K}^{(n-r)\times(n-r)}\}$$

交于 $\{0\}$; 事实上, 若 $A = P\left(\begin{smallmatrix} 0 & B \\ B^* & C \end{smallmatrix}\right)Q$ 适合 $A \pm A_0 \in V$, 考察列秩即知 $\operatorname{col}\left(\begin{smallmatrix} B \\ C \end{smallmatrix}\right) \subset \operatorname{col}\left(\begin{smallmatrix} I_r \\ B^* \end{smallmatrix}\right) \cap \operatorname{col}\left(\begin{smallmatrix} -I_r \\ B^* \end{smallmatrix}\right)$, 必有 $\left(\begin{smallmatrix} B \\ C \end{smallmatrix}\right) = \left(\begin{smallmatrix} I_r \\ B^* \end{smallmatrix}\right)B = \left(\begin{smallmatrix} I_r \\ -B^* \end{smallmatrix}\right)B$, 其中 B^*B 的对角元是 B 的列向量的模长平方, 从而 $A = 0$. 取 V 的一组基 $A_1, \ldots, A_{\dim V}$ 和 W 的一组基 $D_1, \ldots, D_{n(n-r)}$, 注意到 $\forall \lambda_i, \mu_j \in \mathbb{K}$,

$$\sum \lambda_i A_i = -\sum \mu_j D_j \in V \cap W \implies \lambda_i = \mu_j = 0,$$

可见二者合并后在 $\mathbb{K}^{n\times n}$ 中线性无关. 由此, $\dim V + n(n-r) \leq n^2$, 结合题设 $\dim V > np$ 立得 $r > p$. 对于 $p = n - 1$ 情形, 我们给出一个构造. 不妨 $\dim V = n^2 - n + 1$. 考察 V 中矩阵的 n^2 项, 可以选出 $n^2 - n + 1$ 个自由变量. 当 $n = 1$ 时命题显然. 对 n 归纳: 在 n 行中存在第 i 行均为自由变量(因为 $\max \geq \text{mean} = n - 1 + \frac{1}{n}$), 又设在 n 列中第 j 列自由变量个数最少($\leq n - 1$), 那么第 i 行和第 j 列的 $(2n-1)$ 项之外的自由变量个数 $\geq \dim V - (2n-2) = (n-1)^2 - (n-1) + 1$. 取第 i 行满足第 j 列是 1 而其他列是 0, 再用归纳假设得到满秩的 $(n-1)$ 阶子矩阵, 则相应的 n 阶矩阵 (行) 满秩. □

题 42. *(降秩公式[xvi])* 设 $A \in \mathbb{F}^{m\times n}$ 秩为 $r > k$, 且 $U \in \mathbb{F}^{m\times k}$, $R \in \mathbb{F}^{k\times k}$, $V \in \mathbb{F}^{n\times k}$ 均为列满秩矩阵. 证明

$$\operatorname{rank}(A - UR^{-1}V^\top) = \operatorname{rank} A - \operatorname{rank} UR^{-1}V^\top = r - k$$

当且仅当存在 $X \in \mathbb{F}^{n\times k}$ 和 $Y \in \mathbb{F}^{m\times k}$ 使得 $U = AX$, $V = A^\top Y$, $R = Y^\top AX$.

[xv] Meshulam, R. (1985). On the maximal rank in a subspace of matrices.

[xvi] Chu, M.T.; Funderlic, R.E.; Golub, G.H. (1995). A rank-one reduction formula and its applications to matrix factorizations.

证明. 不妨设 $A = \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$, 否则考虑 $A \leftarrow PAQ^\top$, $U \leftarrow PU$, $V \leftarrow QV$, $X \leftarrow (Q^\top)^{-1}X$, $Y \leftarrow (P^\top)^{-1}Y$, 其中 P 和 Q 是可逆矩阵, 适合 $PAQ^\top = \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$. 分出前 r 行, 记 $U = \left(\begin{smallmatrix} U_1 \\ U_2 \end{smallmatrix}\right)$ 和 $V = \left(\begin{smallmatrix} V_1 \\ V_2 \end{smallmatrix}\right)$, 则有

$$UR^{-1}V^\top = \begin{pmatrix} U_1R^{-1}V_1^\top & U_1R^{-1}V_2^\top \\ U_2R^{-1}V_1^\top & U_2R^{-1}V_2^\top \end{pmatrix} \implies A - UR^{-1}V^\top = \begin{pmatrix} I_r - U_1R^{-1}V_1^\top & -U_1R^{-1}V_2^\top \\ -U_2R^{-1}V_1^\top & -U_2R^{-1}V_2^\top \end{pmatrix}.$$

先证必要性. 因为

$$\operatorname{rank}(A - UR^{-1}V^\top) \ge \operatorname{rank}(I_r - U_1R^{-1}V_1^\top) \ge \operatorname{rank} I_r - \operatorname{rank} U_1R^{-1}V_1^\top \ge r - \operatorname{rank} UR^{-1}V^\top$$

应为等式, 所以 U_1 和 V_1 列满秩, 并且由 Cochran 定理 (**题36**) 知 $U_1R^{-1}V_1^\top$ 幂等, 从而 $R = V_1^\top U_1$. 考察列秩, 必有

$$\operatorname{col}(-U_1R^{-1}V_2^\top) \subset \operatorname{col}(I_r - U_1R^{-1}V_1^\top) \xrightarrow[\text{行变换}]{\text{左乘}(-V_1^\top)} \operatorname{col} V_2^\top \subset \operatorname{col} 0 = \{0\} \implies V_2 = 0.$$

类似地, 考察行秩, 必有 $U_2 = 0$. 充分性较简单, 注意到 $U_1(V_1^\top U_1)^{-1}V_1^\top$ 幂等即可, 留给读者. □

题 43. 设 $A, B \in \mathbb{R}^{n\times n}$ 在 $\mathbb{C}^{n\times n}$ 中相似. 证明 A, B 在 $\mathbb{R}^{n\times n}$ 中相似.

证明. 定义 $\mathbb{K}$-线性空间 $V_{\mathbb{K}} = \{X \in \mathbb{K}^{n\times n} : AX = XB\}$. 利用 Gauss 消元, $V_{\mathbb{R}}$ 和 $V_{\mathbb{C}}$ 有一组公共的基 $X_1, \ldots, X_p \in \mathbb{R}^{n\times n}$. 令 $f(t_1, \ldots, t_p) = \det(t_1X_1 + \cdots + t_pX_p)$. 依题意, f 在 $\mathbb{C}^p$ 上有非零取值, 从而是非零多项式, 那么在 $\mathbb{R}^p$ 上也有非零取值. 注意 $f(t_1, \ldots, t_p) \in \operatorname{span}\{t_1^{m_1} \cdots t_p^{m_p} : m_1 + \cdots + m_p = n\}$. □

5.2 特征值

题 44. 任给 $A \in \mathbb{C}^{n\times n}$, 证明存在 $\lambda \in \mathbb{C}$ 和 $v \in \mathbb{C}^n \setminus \{0\}$ 使得 $Av = \lambda v$.

证明. [xvii] 取定 $v_0 \in \mathbb{C}^n \setminus \{0\}$, 则 A^kv_0 $(k = 0, 1, \ldots, n)$ 在 $\mathbb{C}^n$ 中线性相关, 存在不全为零的复数 $a_0, a_1, \ldots, a_n$ 适合 $\sum_{k=0}^n a_kA^kv_0 = 0$. 根据代数基本定理, 存在 $z_1, \ldots, z_d \in \mathbb{C}$ 使得

$$\sum_{k=0}^n a_kz^k = \sum_{k=0}^d a_kz^k = a_d\prod_{j=1}^d (z - z_j), \quad d = \max\{k : a_k \ne 0\}.$$

令 $r = \max\{s : \prod_{j=1}^s (A - z_jI_n)\, v_0 \ne 0\} < d$ 和 $v = \prod_{j=1}^r (A - z_jI_n)\, v_0$, 那么 $Av = z_{r+1}v$. □

题 45. *(Schur 三角化)* 任给 $A \in \mathbb{C}^{n\times n}$, 证明存在酉 *(unitary)* 矩阵 $U \in \mathbb{C}^{n\times n}$ 使得 $U^{-1}AU$ 是上三角阵. 酉阵是正交阵在复数域上的推广, 满足 $U^{-1} = U^*$, 其中 $U^* = \overline{U}^\top$ 为 U 的共轭转置.

证明. 对 n 归纳. 当 $n = 1$ 时, 结论平凡. 一般地, (利用**题44**) 取 $\lambda \in \mathbb{C}$ 和单位向量 $u \in \mathbb{C}^n$ 使得 $Au = \lambda u$. 将 $(u) \in \mathbb{C}^{n\times 1}$ 扩充为酉矩阵 $U_0 \in \mathbb{C}^{n\times n}$, 则有 $AU_0 = U_0\begin{pmatrix} \lambda & ** \\ 0 & B \end{pmatrix}$. 根据归纳假设, 存在 $n-1$ 阶酉阵 U_1 使得 $U_1^{-1}BU_1$ 是上三角阵. 令 $U = U_0 \operatorname{diag}(1, U_1)$, 则 $U^{-1}AU = \begin{pmatrix} \lambda & ** \\ 0 & U_1^{-1}BU_1 \end{pmatrix}$. □

设 $A \in \mathbb{F}^{n\times n}$, $v \in \mathbb{F}^n \setminus \{0\}$, $\lambda \in \mathbb{F}$, 若 $Av = \lambda v$, 则称 (λ, v) 为 A 的一个特征对 (eigenpair), 其中 λ 称为特征值 (eigenvalue), v 称为特征向量 (eigenvector). 也可将 $Av = \lambda v$ 改写成 $(\lambda I_n - A)v = 0$, 作为齐次线性方程组存在非零解的充要条件是

$$\det(\lambda I_n - A) = 0.$$

[xvii] Axler, S. (1995). Down with determinants!.

注意**题44**表明特征对的存在性不依赖行列式! 称

$$p_A(x)=\det(xI_n-A)=x^n-(\operatorname{tr}A)x^{n-1}+\cdots+(-1)^n\det A$$

为 A 的特征多项式 (characteristic polynomial), 其中 $\operatorname{tr}A$ 是 A 的对角元之和, 称为 A 的迹 (trace).

- 对任意 $k\in[n]=\{1,\ldots,n\}$, 特征多项式 $p_A(x)$ 中 x^{n-k} 的系数为 $(-1)^k\sum_{\mathcal{I}\in\binom{[n]}{k}}\det A[\mathcal{I},\mathcal{I}]$.
- 若 $\dim\ker(\lambda_0 I-A)=m$, 则 A 相似于 $\begin{pmatrix}\lambda_0 I_m & *\\ 0 & *\end{pmatrix}$, 进而 $(x-\lambda_0)^m$ 是 $\det(xI-A)$ 的因式.
- 任给多项式 $p(x)=\sum_{k=0}^n a_kx^k$, 其中 $a_n=1$, 则相应的 Frobenius 友阵 (companion matrix)

$$\begin{pmatrix}0&0&\cdots&0&-a_0\\1&0&&0&-a_1\\0&1&\ddots&&\vdots\\\vdots&\ddots&\ddots&0&-a_{n-2}\\0&\cdots&0&1&-a_{n-1}\end{pmatrix}$$

  特征多项式即为 $p(x)$, 并且其特征对形如 $(\lambda,(c\sum_{k=i}^n a_k\lambda^{k-i})_{1\le i\le n})$.
- 任给多项式 f, 若 (λ,v) 是 A 的特征对, 则 $(f(\lambda),v)$ 是 $f(A)$ 的特征对.
- 通过 Schur 三角化 (**题45**), 任给复方阵 A 和多项式 f, 成立谱映射 (spectral mapping) 定理: 若 A 的特征值为 $\lambda_1,\ldots,\lambda_n$, 则 $f(A)$ 的特征值为 $f(\lambda_1),\ldots,f(\lambda_n)$.

题 46. *任给正整数 n, 证明 $\sqrt{n}$ 要么是整数要么是无理数.*

证明. [xviii] 考察矩阵 $A=\begin{pmatrix}-k&n\\1&-k\end{pmatrix}$, 其特征多项式为 $p_A(x)=\det\begin{pmatrix}x+k&-n\\-1&x+k\end{pmatrix}=(x+k)^2-n$. 设 $\sqrt{n}$ 不是整数, 取 $k=\lfloor\sqrt{n}\rfloor$ 为不超过 $\sqrt{n}$ 的最大整数, 则 A 的最大特征值 $\lambda_1=\sqrt{n}-k$ 绝对值严格小于 1. 容易算出 $V_{\lambda_1}(A)=\{v\in\mathbb{C}^2: Av=\lambda_1 v\}$ 可写为 $V_{\lambda_1}(A)=\{\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right): x=\sqrt{n}y\}$. 若存在 $v_0\in V_{\lambda_1}(A)\cap\mathbb{Z}^2\setminus\{0\}$, 则 $A^tv_0\in\mathbb{Z}^2\setminus\{0\}$ $(\forall t\in\mathbb{N})$, 而 $\lim_{t\to\infty}A^tv_0=\lim_{t\to\infty}\lambda_1^tv_0=0$, 矛盾! □

利用行列式的可乘性质, 相似矩阵的特征多项式相同, 并且 $\begin{cases}\text{迹 (= 特征值之和)}\\ \text{行列式 (= 特征值之积)}\end{cases}$ 相同.

题 47. *(Sylvester 行列式定理) 任给 $A\in\mathbb{F}^{m\times n}$ 和 $B\in\mathbb{F}^{n\times m}$, 证明*

$$x^n\det(xI_m-AB)=x^m\det(xI_n-BA).$$

特别地, AB 与 BA 有相同的非零特征值 (以及迹), 并且 I_m-AB 和 I_n-BA 行列式相等.

证明. 注意到

$$\begin{pmatrix}I_m&0\\-B&I_n\end{pmatrix}\begin{pmatrix}0&A\\0&BA\end{pmatrix}=\begin{pmatrix}0&A\\0&0\end{pmatrix}=\begin{pmatrix}AB&A\\0&0\end{pmatrix}\begin{pmatrix}I_m&0\\-B&I_n\end{pmatrix},$$

于是 $\begin{pmatrix}0&A\\0&BA\end{pmatrix}$ 与 $\begin{pmatrix}AB&A\\0&0\end{pmatrix}$ 相似, 二者特征多项式即为所求. □

[xviii] Kalman, D.; Mena, R.; Shahriari, S. (1997). Variations on an irrational theme — geometry, dynamics, algebra.

任给 $A, B \in \mathbb{F}^{n\times n}$, 有

$$\operatorname{tr}(AB - BA) = \operatorname{tr} AB - \operatorname{tr} BA = 0.$$

题 48. *我们进一步考察加法换位子 (commutator) 与迹零 (traceless) 矩阵. 证明:*

1. *在 $\mathbb{F}[x] = \{\mathbb{F}$-系数多项式$\}$ 上取线性映射 $A = \frac{\mathrm{d}}{\mathrm{d}x}$ 和 $B = (f(x) \mapsto xf(x))$, 则有 $AB - BA = I$;*
2. *当 $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$ 时, 在 $\mathbb{F}^{2\times 2}$ 中 $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ 和 $B = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ 满足 $AB - BA = I$;*
3. *若 $C \in \mathbb{F}^{n\times n}$ 对角元均为零, 且 $A \in \mathbb{F}^{n\times n}$ 是对角元互异的对角阵, 则存在 $B \in \mathbb{F}^{n\times n}$ 使得 $AB - BA = C$;*
4. *集合 $\{AB - BA : A, B \in \mathbb{F}^{n\times n}\}$ 在相似变换下不变;*
5. *(Shoda 定理) 只要 $C \in \mathbb{F}^{n\times n}$ 满足 $\operatorname{tr} C = 0$, 就有 $A, B \in \mathbb{F}^{n\times n}$ 使得 $AB - BA = C$.*

证明. 前略. 对于5, 只需说明 $C = (c_{ij})$ 相似于某个对角元均为零的矩阵. [xix)] 现逐个消除 C 的非零对角元. 若 c_{kk} 非零, 则存在 $c_{\ell\ell} \notin \{0, c_{kk}\}$, 从而 $C_{k\ell} = \left(\begin{smallmatrix} c_{kk} & c_{k\ell} \\ c_{\ell k} & c_{\ell\ell} \end{smallmatrix}\right)$ 不是 I_2 的常数倍. 存在 $v \in \mathbb{F}^2$ 使得 $C_{k\ell}v$ 不是 v 的常数倍, 那么 $v, C_{k\ell}v$ 构成 $\mathbb{F}^2$ 的一组基, 于是 $S = \begin{pmatrix} v & C_{k\ell}v \end{pmatrix}$ 可逆. 定义 $P \in \mathbb{F}^{n\times n}$ 如下: (k,ℓ)-主子矩阵为 S, 第 $i \notin \{k,\ell\}$ 个对角元为 1, 其余元素均为零. 不难验证 $P^{-1} \in \mathbb{F}^{n\times n}$ 如下: (k,ℓ)-主子矩阵为 S^{-1}, 第 $i \notin \{k,\ell\}$ 个对角元为 1, 其余元素均为零. 根据分块矩阵乘法性质, $P^{-1}CP$ 相比 C 只有第 k,ℓ 行与第 k,ℓ 列发生了变化, 且 (k,ℓ)-主子矩阵变为 $S^{-1}C_{k\ell}S = \left(\begin{smallmatrix} 0 & * \\ 1 & * \end{smallmatrix}\right)$.

当 $\mathbb{F} = \mathbb{R}$ 时, 也可操作如下. 若 c_{kk} 非零, 则存在 $c_{\ell\ell}$ 与之符号相反. 构造 Givens 旋转阵 $G_\theta \in \mathbb{R}^{n\times n}$ 如下: (k,ℓ)-主子矩阵为 $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, 第 $i \notin \{k,\ell\}$ 个对角元为 1, 其余元素均为零. 易见 $G_\theta^{-1} = G_\theta^\top$, 于是 $G_\theta^{-1}CG_\theta$ 相比 C 只有第 k,ℓ 行与第 k,ℓ 列发生了变化, 且第 k 个对角元变为

$$c_{kk}\cos^2\theta + c_{\ell\ell}\sin^2\theta - (c_{k\ell} + c_{\ell k})\sin\theta\cos\theta.$$

根据介值定理, 存在 $\theta \in (0, \frac{\pi}{2})$ 使 $G_\theta^{-1}CG_\theta$ 第 k 个对角元为零. □

注: 如果 $D \in \mathbb{F}^{n\times n}$ 满足对任意 $v \in \mathbb{F}^n$ 都存在 $c_v \in \mathbb{F}$ 使得 $Dv = c_v v$, 则存在 $c_1, \ldots, c_n \in \mathbb{F}$ 使得 $D = DI_n = \operatorname{diag}(c_1, \ldots, c_n)$, 并且 $(c_1, \ldots, c_n)^\top = D\mathbf{1} = (c_1, \cdots, c_1)^\top$, 其中 $\mathbf{1} = (1, \cdots, 1)^\top$.

5.3 对角化

若矩阵相似于对角阵, 则称之可对角化 (diagonalizable). 记 $P = \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix}$ 和 $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, 则

$$P^{-1}AP = \Lambda \iff AP = P\Lambda \iff Av_i = \lambda_i v_i \ (\forall i).$$

因此, $A \in \mathbb{F}^{n\times n}$ 可对角化当且仅当 $\mathbb{F}^n$ 存在一组基由 A 的特征向量构成. 断言此条件等价于

$$\sum_\lambda \dim\ker(\lambda I_n - A) = n.$$

事实上, 从属于不同特征值的特征向量必然线性无关.

证明. 若 $Av_i = \lambda_i v_i$, 则 $\sum_{i=1}^k v_i = 0 \overset{\text{左乘}A^j}{\Longrightarrow} \sum_{i=1}^k \lambda_i^j v_i = 0 \ (\forall j) \underset{\text{Vandermonde}}{\overset{\lambda_i\text{互异}}{\Longrightarrow}} v_i = 0 \ (\forall i)$. □

xix) Lax, P.D. (2007). *Linear algebra and its applications*, 2nd ed., Appendix 13.
Garcia, S.R.; Horn, R.A. (2023). *Matrix mathematics: a second course in linear algebra*, 2nd ed., §4.5.

题 49. 设 $A \in \mathbb{F}^{n\times n}$ 幂等. 证明 A 相似于 $\operatorname{diag}(I_r, 0)$, 其中 $r = \operatorname{rank} A$.

证明. 由于 $\dim\ker A = n - \operatorname{rank} A$, 而 $\operatorname{rank} A = \dim \operatorname{im} A$, 只需说明 $\operatorname{im} A = \ker(I - A)$. 一方面, $(I - A)A = 0$, 于是 $\operatorname{im} A \subset \ker(I - A)$. 另一方面, 若 $x \in \mathbb{F}^n$ 满足 $(I - A)x = 0$, 则有 $x = Ax$, 即得 $\ker(I - A) \subset \operatorname{im} A$. □

作为推论, 幂等阵 A 满足 $\operatorname{tr} A = r = \operatorname{rank} A$.

题 50. 设 $A \in \mathbb{F}^{n\times n}$ 是循环 *(circulant)* 阵, 即存在 $a_0, a_1, \dots, a_{n-1} \in \mathbb{F}$ 使得 $A = \sum_{k=0}^{n-1} a_k C^k$, 其中 $C = (\mathbb{1}_{[i+1\equiv j \pmod n]})_{1\le i,j\le n}$. 证明: 若 $\mathbb{F} = \mathbb{C}$, 则 A 可对角化.

证明. 不难算出 $\det(\lambda I - C) = \lambda^n - 1$, 于是 C 有 n 个互异的特征值 $\lambda_k = \mathrm{e}^{2\pi(k/n)\mathrm{i}}$, 进而 C 相似于 $\operatorname{diag}(\lambda_0, \lambda_1, \dots, \lambda_{n-1})$. 令 $f(x) = \sum_{k=0}^{n-1} a_k x^k$, 则 $A = f(C)$ 相似于 $\operatorname{diag}(f(\lambda_0), f(\lambda_1), \dots, f(\lambda_{n-1}))$. 特别地, $\det A = f(\lambda_0)f(\lambda_1)\cdots f(\lambda_{n-1})$. □

题 51. 任给 $A \in \mathbb{F}^{m\times m}$, $B \in \mathbb{F}^{n\times n}$, $C \in \mathbb{F}^{m\times n}$, 证明: 若 $\left(\begin{smallmatrix} A & C \\ 0 & B \end{smallmatrix}\right)$ 可对角化, 则 A 和 B 均可对角化.

证明. 设 $P \in \mathbb{F}^{m\times m}$, $Q \in \mathbb{F}^{m\times n}$, $R \in \mathbb{F}^{n\times m}$, $S \in \mathbb{F}^{n\times n}$ 使得 $\left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right)$ 可逆且 $\left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right)^{-1}\left(\begin{smallmatrix} A & C \\ 0 & B \end{smallmatrix}\right)\left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right) = D$ 是对角阵. 考察后 n 行可知 $\begin{pmatrix} R & S \end{pmatrix}$ 的列向量均为 B 的特征向量, 其极大线性无关组即可对角化 B. 设 $X \in \mathbb{F}^{m\times m}$, $Y \in \mathbb{F}^{m\times n}$, $Z \in \mathbb{F}^{n\times m}$, $W \in \mathbb{F}^{n\times n}$ 满足 $\left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} X & Y \\ Z & W \end{smallmatrix}\right)$, 则

$$D = D^\top = \begin{pmatrix} X^\top & Z^\top \\ Y^\top & W^\top \end{pmatrix}^{-1} \begin{pmatrix} A^\top & 0 \\ C^\top & B^\top \end{pmatrix} \begin{pmatrix} X^\top & Z^\top \\ Y^\top & W^\top \end{pmatrix}.$$

考察前 m 行, 如前得到 $A^\top$ 可对角化, 从而 A 可对角化. □

题 52. 证明: 若方阵 A, B 均可对角化, 且 $AB = BA$, 则 AB 可对角化.

证明. 取可逆阵 P 使得 $P^{-1}AP = \operatorname{diag}(\lambda_1 I_{n_1}, \dots, \lambda_s I_{n_s})$, 其中 $\lambda_1, \dots, \lambda_s$ 是互异的标量. 记 $P^{-1}BP = (B_{ij})_{1\le i,j\le s}$, 其中 B_{ij} 是 $n_i \times n_j$ 矩阵. 由 $AB = BA$ 可得 $\lambda_i B_{ij} = \lambda_j B_{ij}$, 从而 $B_{ij} = 0$, $\forall i \ne j$. 易见 $P^{-1}BP$ 可对角化, 于是**题51**表明 B_{ii} 均可对角化. 取 n_i 阶可逆阵 Q_i 使得 $Q_i^{-1}B_{ii}Q_i = D_i$ 是对角阵, 记 $Q = \operatorname{diag}(Q_1, \dots, Q_s)$, 则有

$$\begin{cases} Q^{-1}P^{-1}APQ = \operatorname{diag}(\lambda_1 I_{n_1}, \dots, \lambda_s I_{n_s}) \\ Q^{-1}P^{-1}BPQ = \operatorname{diag}(D_1, \dots, D_s) \end{cases} \implies (PQ)^{-1}AB(PQ) = \operatorname{diag}(\lambda_1 D_1, \dots, \lambda_s D_s).$$

□

若 $A \in \mathbb{C}^{n\times n}$ 共轭对称, 即

$$A^* = A,$$

则称 A 自伴 (self-adjoint), 也称 A 为 Hermite 矩阵. 酉相似不改变自伴性, 而**题45**表明复矩阵总是酉相似于上三角阵, 所以 Hermite 矩阵总是酉相似于实对角阵. 实 Hermite 矩阵即实对称矩阵, 能够正交对角化 (在**题45**的证明中将 $\mathbb{C}$ 改为 $\mathbb{R}$, 酉阵改为正交阵).

题 53. 设 A 是复方阵. 证明 A 可酉对角化当且仅当 A 正规 *(normal)*, 即满足 $A^*A = AA^*$.

证明. 必要性显然. 下证充分性. 注意到酉相似不改变正规性, 利用 Schur 三角化 (**题45**), 只需考虑 A 是上三角阵的情形. 记 $A = (a_{ij})_{1\le i,j\le n}$, 其中 $a_{ij} = 0$, $\forall i > j$. 比较 $A^*A = AA^*$ 对角元可得 $\sum_{j=i}^n |a_{ij}|^2 = \sum_{j=1}^i |a_{ji}|^2$, $\forall i$. 依次考察 $i = 1, \dots, n$, 必有 $a_{ij} = 0$, $\forall j > i$. □

设标准正交基 $u_1,\dots,u_n$ 将方阵 A 对角化, 即 $\begin{pmatrix}u_1 & \dots & u_n\end{pmatrix}^* A\begin{pmatrix}u_1 & \dots & u_n\end{pmatrix} = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$, 则成立下述谱分解 (spectral decomposition):

$$A = \begin{pmatrix}u_1 & \dots & u_n\end{pmatrix}\operatorname{diag}(\lambda_1,\dots,\lambda_n)\begin{pmatrix}u_1 & \dots & u_n\end{pmatrix}^* = \lambda_1 u_1 u_1^* + \dots + \lambda_n u_n u_n^*.$$

将 (λ_i, u_i) $(1 \le i \le n)$ 排列为 $(\lambda_{(j)}, u_{jk})$ $(1 \le k \le m_j;\ 1 \le j \le s)$, 其中 $\lambda_{(j)}$ 互不相同, 则 $\ker(\lambda_{(j)} I - A) = \operatorname{span}\{u_{jk} : 1 \le k \le m_j\}$ 两两正交. 如果解析函数 f 在 A 的特征值处有定义, 则有

$$f(A) = \begin{pmatrix}u_1 & \dots & u_n\end{pmatrix}\operatorname{diag}(f(\lambda_1),\dots,f(\lambda_n))\begin{pmatrix}u_1 & \dots & u_n\end{pmatrix}^* = f(\lambda_1) u_1 u_1^* + \dots + f(\lambda_n) u_n u_n^*.$$

6 二次型与矩阵的合同

6.1 二次型

关于 $x_1,\dots,x_n$ 的二次型 (quadratic form) 指的是 $\operatorname{span}\{x_i x_j : 1 \le i \le j \le n\}$ 中的元素, 它总能唯一地表示为 $x^\top A x$, 其中 $x = (x_1,\dots,x_n)^\top$, 矩阵 A 是对称的 (symmetric, 即 $A^\top = A$). 作变换 $x = Py$, 则有

$$x^\top A x = y^\top P^\top A P y.$$

如果存在可逆阵 P 使得 $P^\top A P = B$, 则称 A 与 B 合同 (congruent), 这是等价关系. 任意对称矩阵都合同于某个对角阵, 于是任意二次型都可以表为平方项的加权和 $\sum_{i=1}^n c_i y_i^2$. 在实数域上, 成立惯性 (inertia) 定理: 设对称矩阵 $A \in \mathbb{R}^{n\times n}$ 的特征值为 $\lambda_1,\dots,\lambda_n$, 若 A 合同于 $\operatorname{diag}(c_1,\dots,c_n)$, 则

$$\begin{cases} |\{i : c_i > 0\}| = |\{i : \lambda_i > 0\}| & \text{称为正惯性指数,} \\ |\{i : c_i < 0\}| = |\{i : \lambda_i < 0\}| & \text{称为负惯性指数.} \end{cases}$$

正负惯性指数合称为惯性指数 (signature), 二者之和是秩. 惯性指数相同的实对称矩阵彼此合同.

题 54. *设 $W = (w_{ij})_{i,j\in[n]}$ 是反对称实矩阵, 其中 $[n] = \{1,\dots,n\}$. 令 $E = \{(i,j) \in [n]\times[n] : w_{ij} > 0\}$, 则有关联矩阵 $B = (b_{ie})_{e\in E}^{i\in[n]}$, 其中 $b_{ie} = \begin{cases} \sqrt{w_{ij}} & \textit{if } e = (i,j); \\ -\sqrt{w_{ji}} & \textit{if } e = (j,i); \\ 0 & \textit{otherwise.} \end{cases}$ 此外, 度矩阵 D 是 $d_i = \sum_{j\in[n]} |w_{ij}|$ 组成的对角阵.*

1. *证明 $BB^\top = D - (|w_{ij}|)_{i,j\in[n]}$, 称之为 Laplace 矩阵, 记作 L.*
2. *(图的 Dirichlet 能量[xx]) 证明二次型 $(x_1,\dots,x_n) \mapsto \sum_{(i,j)\in E} w_{ij}(x_i - x_j)^2$ 的矩阵是 L.*

证明. 2. 原式 $= \frac{1}{2}\sum_{i,j\in[n]} |w_{ij}|(x_i^2 + x_j^2 - 2x_i x_j) = \frac{1}{2}\big(\sum_i d_i x_i^2 + \sum_j d_j x_j^2\big) - \sum_{i,j\in[n]} |w_{ij}| x_i x_j = x^\top L x$, $\forall x = (x_1,\dots,x_n)^\top$. 易得 0 是 L 的最小特征值. □

题 55. *求二次型 $X = (x_{ij})_{1\le i,j\le n} \mapsto \operatorname{tr} X^2$ 的惯性指数.*

解. 注意到 $\operatorname{tr} X^2 = \sum_{i,j} x_{ij} x_{ji} = \sum_i x_{ii}^2 + \sum_{j>i} \big\{(\frac{x_{ij}+x_{ji}}{\sqrt{2}})^2 - (\frac{x_{ij}-x_{ji}}{\sqrt{2}})^2\big\}$, 所求正负惯性指数分别为 $\frac{n^2+n}{2}$ 和 $\frac{n^2-n}{2}$. □

[xx] Boyd, S.; Vandenberghe, L. (2018). *Introduction to applied linear algebra: vectors, matrices, and least squares*, §7.3.

题 56. 设 $\mathbb{K}$ 是 $\mathbb{R}$ 或 $\mathbb{C}$, 矩阵 $A \in \mathbb{K}^{n\times n}$ 自伴, 特征值排列为 $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$. 对 $k = 1, \ldots, n$, 证明:

1. (Courant–Fischer–Weyl 最小最大原理)
$$\lambda_k(A) = \max_{\dim V=k} \min_{v\in V\setminus\{0\}} \frac{v^*Av}{v^*v}, \quad \lambda_{n-k+1}(A) = \min_{\dim V=k} \max_{v\in V\setminus\{0\}} \frac{v^*Av}{v^*v},$$
其中 $\frac{v^*Av}{v^*v}$ 称为 Rayleigh–Ritz 商;

2. [xxi]若 $P \in \mathbb{K}^{n\times n}$ 可逆, 则存在 $\theta_k \in [\lambda_n(P^*P), \lambda_1(P^*P)]$ 使得
$$\lambda_k(P^*AP) = \theta_k\lambda_k(A).$$
由 $\lambda_n(P^*P) > 0$ 可知 $\lambda_k(P^*AP)$ 与 $\lambda_k(A)$ 同号.

证明. 1. 通过谱分解可得标准正交基 $v_1, \ldots, v_n \in \mathbb{K}^n$ 适合 $Av_i = \lambda_i(A)v_i$. 第一个等式由下述两点保证:

- 取 $V_k = \operatorname{span}\{v_1, \ldots, v_k\}$, 则 $v \in V_k$ 总可以写成 $v = c_1v_1 + \cdots + c_kv_k$, 其中 $c_1, \ldots, c_k \in \mathbb{K}$, 进而
$$\begin{aligned} v^*Av &= v^*(c_1\lambda_1(A)v_1 + \cdots + c_k\lambda_k(A)v_k) \\ &= |c_1|^2\lambda_1(A) + \cdots + |c_k|^2\lambda_k(A) \geq (|c_1|^2 + \cdots + |c_k|^2)\lambda_k(A) = \lambda_k(A)v^*v. \end{aligned}$$
这表明 $\min_{v\in V_k\setminus\{0\}} \frac{v^*Av}{v^*v} \geq \lambda_k(A)$.
- 任取 k 维子空间 V 和它的一组基 $u_j = \sum_{i=1}^n v_ib_{ij}$, $j = 1, \ldots, k$, 其中 $b_{ij} \in \mathbb{K}$ 使 $B = (b_{ij})_{1\leq j\leq k}^{1\leq i\leq n}$ 列满秩. 记 B 的第 j 列为 β_j. 对 B 作列变换 Gauss 消元可得 $z_1, \ldots, z_k \in \mathbb{K}$ 使 $\sum_{j=1}^k z_j\beta_j$ 第一个非零分量出现在第 $\ell \geq k$ 行, 从而 $V \ni \sum_{j=1}^k z_ju_j = \sum_{i=\ell}^n v_i \sum_{j=1}^k b_{ij}z_j$, 且不难验证此向量的 Rayleigh 商不超过 $\lambda_\ell(A)$, 故 $\min_{v\in V\setminus\{0\}} \frac{v^*Av}{v^*v} \leq \lambda_\ell(A) \leq \lambda_k(A)$.

注意到 $\lambda_{n-k+1}(A) = -\lambda_k(-A)$, 第一个等式应用于 $-A$ 即得第二个等式.

2. 在1中取 $k = n$ 立得 $\lambda_n(P^*P) \leq \frac{v^*P^*Pv}{v^*v} \leq \lambda_1(P^*P)$, 从而
$$\lambda_n(P^*P)\frac{(Pv)^*A(Pv)}{(Pv)^*(Pv)} \leq \frac{v^*P^*APv}{v^*v} \leq \lambda_1(P^*P)\frac{(Pv)^*A(Pv)}{(Pv)^*(Pv)}.$$
对于 k 维子空间 V, 线性双射 $v \in V \mapsto Pv \in P(V)$ 像集仍是 k 维子空间, 所以应用1即证所欲. □

6.2 正定矩阵

设 $\mathbb{K}$ 是 $\mathbb{R}$ 或 $\mathbb{C}$, 矩阵 $A \in \mathbb{K}^{n\times n}$ 自伴. 若
$$x^*Ax \geq 0, \ \forall x \in \mathbb{K}^n,$$
则称 A 半正定 (positive semi-definite); 若 $x^*Ax > 0$, $\forall x \neq 0$, 则称 A 正定 (positive definite). 利用谱分解, A 半正定当且仅当 A 特征值均非负, A 正定当且仅当 A 特征值均为正数. 主子式判据并不常用.

任给矩阵 P, 易见 P^*P 半正定.

[xxi] Wimmer, H.K. (1983). On Ostrowski's generalization of Sylvester's law of inertia.

题 57. *(Cholesky 分解) 设方阵 B 半正定. 证明: 存在对角元非负的下三角矩阵 L 使得 $B = LL^*$.*

证明. 存在方阵 A 使得 $B = A^*A$, 比如 $A = \sqrt{B}$. 通过 QR 分解 (**题37**) 得到酉阵 Q 和上三角阵 R 适合 $A = QR$, 则 $B = R^*Q^*QR = R^*R$. 取 $L = R^*$ 即证. □

易得 Hadamard 不等式 $\det B \le \prod_{j=1}^{n} B[j,j]$, 其中 $B[j,j]$ 是 B 的第 j 个对角元.

题 58. *设 $A \in \mathbb{K}^{n\times n}$ 半正定. 证明 $(\det A)^{1/n} = \inf\{(\det B)^{-1/n} \operatorname{tr} AB/n : B \in \mathbb{K}^{n\times n}$正定$\}$.*

证明. 设酉矩阵 U 满足 $U^*AU = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. 任给正定矩阵 B, 记 U^*BU 对角元为 $b_1, \ldots, b_n$, 则 Hadamard 不等式给出 $\det B \le \prod_j b_j$, 进而 $\operatorname{tr} AB/n = \sum_j \lambda_j b_j/n \ge (\prod_j \lambda_j b_j)^{1/n} \ge (\det A \det B)^{1/n}$. 取 $B_\varepsilon = (A+\varepsilon I)^{-1}$, 则 $(\det B_\varepsilon)^{-1/n} \operatorname{tr} AB_\varepsilon/n = \{\prod_j(\lambda_j+\varepsilon)\}^{1/n} \sum_j \frac{\lambda_j}{\lambda_j+\varepsilon}/n \to (\det A)^{1/n}$ $(\varepsilon \searrow 0)$. □

题 59. *(Brunn–Minkowski 不等式: 行列式版本) 设 $A, B \in \mathbb{K}^{n\times n}$ 半正定. 证明*

$$(\det(A+B))^{1/n} \ge (\det A)^{1/n} + (\det B)^{1/n}.$$

证明. 取正定矩阵 P_k 使得 $(\det P_k)^{-1/n} \operatorname{tr}(A+B)P_k/n \to (\det(A+B))^{1/n}$ $(k \to \infty)$. 注意到
左端 $= (\det P_k)^{-1/n} \operatorname{tr} AP_k/n + (\det P_k)^{-1/n} \operatorname{tr} BP_k/n \ge (\det A)^{1/n} + (\det B)^{1/n}$ 即证. □

题 60. *(奇异值分解, singular value decomposition) 设 $A \in \mathbb{K}^{m\times n}$ 秩为 r, 其中 $\mathbb{K}$ 是 $\mathbb{R}$ 或 $\mathbb{C}$. 证明: 存在 $\sigma_1 \ge \cdots \ge \sigma_r > 0$ 以及 (各自空间中) 标准正交向量组 $u_1, \ldots, u_r \in \mathbb{K}^m$ 和 $v_1, \ldots, v_r \in \mathbb{K}^n$ 使得*

$$A = \sigma_1 u_1 v_1^* + \cdots + \sigma_r u_r v_r^* = U\Sigma V^*,$$

其中 $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, $U = \begin{pmatrix} u_1 & \ldots & u_r \end{pmatrix}$, $V = \begin{pmatrix} v_1 & \ldots & v_r \end{pmatrix}$.

进一步有 $A = (UV^*)(V\Sigma V^*)$, 称作极分解 (polar decomposition, 类似复数的极坐标分解).

当 $m \le n$ 时, $A = (UU^*)\big((U, U_\perp)(\operatorname{diag}(\Sigma, I_{m-r}), 0)(V, V_\perp)^*\big)$ 是正交投影阵和行满秩矩阵的乘积, 其中 $U_\perp$ 和 $V_\perp$ 由相应的正交补空间的基构成列.

证明. 半正定矩阵 A^*A 谱分解可得标准正交基 $v_1, \ldots, v_n \in \mathbb{K}^n$ 和实数 $\lambda_1 \ge \cdots \ge \lambda_n$ 适合

$$A^*A = \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix} \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix}^* = \lambda_1 v_1 v_1^* + \cdots + \lambda_n v_n v_n^*,$$

根据惯性定理和 $\operatorname{rank} A^*A = \operatorname{rank} A$ 可知 $\lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n$. 令 $\sigma_j = \sqrt{\lambda_j}$ 和 $u_j = Av_j/\sigma_j$. 不难验证 $u_1, \ldots, u_r$ 标准正交, 且 $Av_j = 0$, $\forall j > r$, 从而

$$A = A(v_1 v_1^* + \cdots + v_n v_n^*) = \sigma_1 u_1 v_1^* + \cdots + \sigma_r u_r v_r^*.$$

注: 可扩充 U 和 V 为酉 (正交) 阵, Σ 添补零. □

题 61. *(Fisher 不等式) 设 $A_1, \ldots, A_m$ 是 $[n] = \{1, \ldots, n\}$ 的两两不同的非空子集, 且 $|A_j \cap A_k| = t$ 不依赖 j, k. 证明 $m \le n$.*

证明. 先考虑存在某个 A_j 是 t 元集合的情况, 不妨设 $|A_1| = t \ge 1$. 此时必有 $A_1 \subset A_k$ $(\forall k)$, 进而 $A_k \setminus A_1$ $(k = 2, \ldots, m)$ 是 $[n] \setminus A_1$ 的两两不同的非空子集, 于是归纳可得

$$m - 1 = |\{A_k \setminus A_1 : k = 2, \ldots, m\}| \le |[n] \setminus A_1| = n - t \implies m \le n - (t-1) \le n.$$

以下设 $\min_{j\in[n]} |A_j| > t \ge 0$. 令 $x_{ij} = \mathbb{1}_{[i\in A_j]}$, 往证 $X = (x_{ij})_{i\in[n],\, j\in[m]}$ 列满秩.

方法 1. 令 $\mathbf{1}_m = (1, \cdots, 1)^\top \in \mathbb{R}^m$. 注意到

$$X^\top X = \Big(\sum_{i\in[n]} x_{ij}x_{ik}\Big)_{j,k\in[m]} = \operatorname{diag}(|A_j| - t)_{j\in[m]} + t\mathbf{1}_m\mathbf{1}_m^\top$$

正定, 必有 $\operatorname{rank} X \geq \operatorname{rank} X^\top X = m$.

方法 2.[xxii]设 $\beta = (\beta_1, \ldots, \beta_m)^\top \in \mathbb{R}^m$ 满足 $X\beta = 0$. 一方面,

$$\sum_{j\in[m]} |A_j|\beta_j = \mathbf{1}_n^\top X\beta = 0.$$

另一方面,

$$|A_k|\beta_k + t\sum_{j\in[m]\setminus\{k\}} \beta_j = (\mathbb{1}_{[i\in A_k]})_{i\in[n]}^\top X\beta = 0, \quad \forall k \in [m];$$

此 m 式相加, 可得

$$\sum_{j\in[m]} |A_j|\beta_j + t(m-1)\sum_{j\in[m]} \beta_j = 0.$$

当 $t = 0$ 时, $|A_k|\beta_k = 0 \implies \beta_k = 0$. 当 $m = 1$ 时, $|A_1|\beta_1 = 0 \implies \beta_1 = 0$. 当 $t \geq 1$ 且 $m > 1$ 时, 容易得到 $\sum_{j\in[m]} \beta_j = 0$, 进而 $(|A_k| - t)\beta_k = 0 \implies \beta_k = 0$. 综上, $\ker X = 0$. □

[xxii]Mathew, R.; Mishra, T.K. (2020). A combinatorial proof of Fisher's inequality.