第 0 章　综述与杂叙

0.0　引言

在首章我们要总结许多有用的概念和结果，其中的一些内容给本书其余部分的材料提供了基础. 这部分材料中有一些已包含在规范的线性代数的初等课程之中，不过我们还另外增加了一些有用的内容，尽管这些内容在后面的阐释中并不出现. 读者可以将这一章当作本书第 1 章中主要部分开始讲述之前的一个温习热身；其后，它还可以被用作后续章节中遇到的记号和定义的一个方便的参考资料. 我们假设读者已经熟悉线性代数的基本概念以及类似矩阵乘法和矩阵加法这样的矩阵运算的技术手段.

0.1　向量空间

有限维的向量空间是矩阵分析的基本架构.

0.1.1　纯量域

构成向量空间的基础是它的**域**(field)，或者说是纯量的集合. 就我们的目的而言，典型的基础域是实数 $\mathbf{R}$ 或者复数 $\mathbf{C}$(见附录 A)，不过它也可以是有理数，以一个特殊的素数为模的整数，或者是某个另外的域. 当域被指定时，我们就用符号 $\mathbf{F}$ 来表示它. 为验证它是一个域，集合必须关于两个二元运算"加法"与"乘法"是封闭的. 这两个运算都必须满足结合律和交换律，且在该集合中每一运算都需有一个单位元；对于加法，每一元素在该集合中都必须有逆元存在，而对于乘法，除了加法的单位元之外，其他每一元素在该集合中也都必须有逆元存在；乘法关于加法必须是可分配的.

0.1.2　向量空间

域 $\mathbf{F}$ 上的一个**向量空间**(vector space)V 是一组对象(称为**向量**)的集合 V，它关于一个满足结合律和交换律的二元运算("加法")封闭，有一个单位元(**零向量**，记为 0)并且加法在该集合中有逆元. 该集合关于用纯量域 $\mathbf{F}$ 中的元素进行向量的"数乘"运算也是封闭的，且对所有 a，$b\in\mathbf{F}$ 以及所有 x，$y\in V$ 有下述性质：$a(x+y)=ax+ay$，$(a+b)x=ax+bx$，$a(bx)=(ab)x$，又对乘法单位元 $e\in\mathbf{F}$ 有 $ex=x$. 1

对给定的域 $\mathbf{F}$ 和给定的正整数 n，由 $\mathbf{F}$ 中的元素形成的 n 元数组的集合 $\mathbf{F}^n$ 在 $\mathbf{F}^n$ 中逐个元素相加的加法之下构成 $\mathbf{F}$ 上的一个向量空间. 我们约定 **$\mathbf{F}^n$ 中的元素总是表示成列向量**，通常称之为 n **元向量**(n-vector). 特殊的例子 $\mathbf{R}^n$ 和 $\mathbf{C}^n$ 是这本书中基本的向量空间，$\mathbf{R}^n$ 是一个实向量空间(即它是实数域上的向量空间)，而 $\mathbf{C}^n$ 既是实向量空间，又是复向量空间(即复数域上的向量空间). 实系数或者复系数(不高于一个指定次数或者任意次数的)多项式的集合以及在 $\mathbf{R}$ 或者 $\mathbf{C}$ 的子集合上的实值函数或者复值函数的集合(全都带有通常的函数加法以及数与函数的乘法的概念)也都是实的或者复的向量空间的例子.

0.1.3 子空间，生成子空间以及线性组合

域 $\mathbf{F}$ 上的向量空间 V 的一个**子空间**(subspace)是 V 的一个子集，其本身也是 $\mathbf{F}$ 上的一个向量空间，它有与 V 中同样的向量加法以及数乘运算．确切地说，V 的一个子集是一个子空间，当然，前提是它关于这两个运算是封闭的．例如，$\{[a,\ b,\ 0]^{\mathrm{T}}:\ a,\ b\in\mathbf{R}\}$ 是 $\mathbf{R}^3$ 的一个子空间，转置记号见(0.2.5)．子空间的交总是一个子空间，子空间的并不一定还是一个子空间．子集$\{0\}$和 V 永远是 V 的子空间，所以它们常被称为**平凡的子空间**(trivial subspace)；V 的一个子空间称为**非平凡的**(nontrivial)，如果它异于$\{0\}$和 V．V 的一个子空间称为**真子空间**(proper subspace)，如果它不等于 V．我们把$\{0\}$称为**零向量空间**(zero vector space)．由于一个向量空间总是包含零向量，故而子空间不可能是空的．

如果 S 是域 $\mathbf{F}$ 上向量空间 V 的一个子集，则 S 的生成子空间 $\mathrm{span}S$ 是 V 中所有包含 S 的子空间的交．如果 S 非空，那么 $\mathrm{span}S=\{a_1v_1+\cdots+a_kv_k:\ v_1,\ \cdots,\ v_k\in S,\ a_1,\ \cdots,\ a_k\in\mathbf{F},\ k=1,\ 2,\ \cdots\}$；如果 S 是空集，则它包含在 V 的每一个子空间中．V 的每个子空间的交就是子空间$\{0\}$，故而此定义确保 $\mathrm{span}S=\{0\}$．注意，即使 S 不是子空间，$\mathrm{span}S$ 也总是一个子空间；S 称为 $\mathrm{span}V$，如果 $\mathrm{span}S=V$．

域 $\mathbf{F}$ 上向量空间 V 中向量的一个**线性组合**(linear combination)是任意一个形如 $a_1v_1+\cdots+a_kv_k$ 的表达式，其中 k 是正整数，$a_1,\ \cdots,\ a_k\in\mathbf{F}$，而 $v_1,\ \cdots,\ v_k\in V$．于是，V 的一个非空子集 S 的生成子空间就由 S 中有限多个向量的所有线性组合组成．线性组合 $a_1v_1+\cdots+a_kv_k$ 是**平凡的**(trivial)，如果 $a_1=\cdots=a_k=0$；反之，它就是**非平凡的**(nontrivial)．根据定义，一个线性组合就是向量空间中**有限多个**元素的和．

设 S_1 和 S_2 是域 $\mathbf{F}$ 上一个向量空间的子空间．S_1 与 S_2 的**和**(sum)是子空间

$$S_1+S_2=\mathrm{span}\{S_1\cup S_2\}=\{x+y:x\in S_1,y\in S_2\}$$

如果 $S_1\cap S_2=\{0\}$，我们就说 S_1 与 S_2 的和是**直和**(direct sum)，并将它记为 $S_1\oplus S_2$；每一个 $z\in S_1\oplus S_2$ 都可以用唯一一种方式表示成 $z=x+y$，其中 $x\in S_1$，而 $y\in S_2$．

0.1.4 线性相关与线性无关

我们称域 $\mathbf{F}$ 上向量空间 V 中有限多个向量 $v_1,\ \cdots,\ v_k$ 是**线性相关的**(linearly dependent)，当且仅当存在不全为零的纯量 $a_1,\ \cdots,\ a_k\in\mathbf{F}$，使得 $a_1v_1+\cdots+a_kv_k=0$．于是，一组向量 $v_1,\ \cdots,\ v_k$ 是线性相关的，当且仅当 $v_1,\ \cdots,\ v_k$ 的某个非平凡的线性组合是零向量．通常更为方便的是说“$v_1,\ \cdots,\ v_k$ 是线性相关的”，而不用更为正式的说法“一组向量 $v_1,\ \cdots,\ v_k$ 是线性相关的”．称向量组 $v_1,\ \cdots,\ v_k$ 的**长度**(length)为 k．有两个或者更多个向量的向量组是线性相关的，如果这些向量中有一个是其余向量中某一些的线性组合．特别地，至少含有两个向量且有两个向量相等的向量组是线性相关的．两个向量是线性相关的，当且仅当其中一个是另一个的纯量倍数．仅由零向量组成的一组向量是线性相关的，因为对 $a_1=1$ 有 $a_10=0$．

域 $\mathbf{F}$ 上的向量空间 V 中的有限多个向量 $v_1,\ \cdots,\ v_k$ 的向量组是**线性无关的**(linearly independent)，如果它不是线性相关的．再次可以方便地说成“$v_1,\ \cdots,\ v_k$ 是线性无关的”，而不说“一组向量 $v_1,\ \cdots,\ v_k$ 是线性无关的”．

有时会遇到一组自然状态的向量，它们有无穷多个元素，例如，所有实系数多项式组成的向量空间中的单项式 1，t，t^2，t^3，…，以及以$[0, 2\pi]$为周期的复值连续函数组成的向量空间中的复指数 1，e^{it}，e^{2it}，e^{3it}，….

如果在一个向量组(有限或者无限的)中去掉某些向量，则得到的向量组是原来向量组的一个**子向量组**(sublist). 一个有无穷多个向量的向量组称为是线性相关的，如果它的某个有限的子向量组是线性相关的；它称为是线性无关的，如果它的任何一个有限的子向量组是线性无关的。线性无关的向量组的任意一个子向量组都是线性无关的，而具有一组线性相关的子向量组的任何一组向量都是线性相关的. 由于仅由零向量组成的向量组是线性相关的，因而包含零向量的任意一组向量都是线性相关的. 有可能一组向量是线性相关的，然而它的任意一个真子向量组都是线性无关的，见(1.4.P12). 一组空向量组不是线性相关的，故而它是线性无关的.

一个有限集合的**基数**(cardinality)是它的(必定不相同的)元素的个数. 对于向量空间 V 中一组给定的向量 v_1，…，v_k，集合$\{v_1, \cdots, v_k\}$的基数小于 k，当且仅当该组中至少有两个向量是相同的；如果 v_1，…，v_k 线性无关，那么集合$\{v_1, \cdots, v_k\}$的基数就是 k. 一组向量(有限或无限多个)的**生成子空间**(span)是该组中元素集合的生成子空间；一组向量张成 V，如果 V 是这组向量的生成子空间.

称向量的集合 S 是线性无关的，如果 S 中每一组有限多个不同的向量都是线性无关的；称 S 是线性相关的，如果 S 中某一组有限多个不同的向量是线性相关的.

0.1.5 基

向量空间 V 中一组以 V 作为其生成子空间的线性无关的向量称为 V 的一组**基**(basis). V 的每一个元素都可以用唯一一种方式表示成基中向量的线性组合，而只要向基中添加或者从基中删除任何一个向量，这一结论就不再成立. V 中一组线性无关的向量是 V 的一组基，当且仅当任何一组包含它作为真子集的向量都不会是线性无关的. 一组生成 V 的向量是 V 的一组基，当且仅当它的任何一个真子集都不可能生成 V. 空的向量组是零向量空间的基. 3

0.1.6 扩充成基

向量空间 V 中任何一组线性无关的向量看起来都可以用多于一种方式扩充成为 V 的一组基. 向量空间可以有非有限的基，例如，无限多个单项式 1，t，t^2，t^3，…就是所有实系数多项式组成的实向量空间的一组基，每一个多项式都是这组基中(有限多个)元素的唯一的线性组合.

0.1.7 维数

如果存在一个正整数 n，使得向量空间 V 的一组基恰好包含 n 个向量，那么 V 的每组基都恰好由 n 个向量组成，基的这个共同的基数就是向量空间 V 的**维数**(dimension)，并记为 $\dim V$. 在此情形，V 是**有限维的**(finite-dimensional)，反之则 V 是**无限维的**(infinite-dimensional). 对于无限维的情形，任何两组基的元素之间都有一个一一对应. 实向量空间 $\mathbf{R}^n$ 的维数是 n. 向量空间 $\mathbf{C}^n$ 关于域 $\mathbf{C}$ 的维数是 n，而关于域 $\mathbf{R}$ 的维数则是 $2n$. $\mathbf{F}^n$ 中的

基 e_1，…，e_n（每一个 n 元向量 e_i 的第 i 个元素是 1，而其他元素均为 0）称为**标准基**(standard basis).

将“V 是一个维数为 n 的有限维向量空间”简略说成“V 是一个 n 维向量空间”是很方便的．一个 n 维向量空间的任意一个子空间都是有限维的，其维数严格小于 n，如果它是真子空间．

设 V 是一个有限维向量空间，并设 S_1 与 S_2 是 V 的两个给定的子空间．则**子空间交引理**(subspace intersection lemma)说的是

$$\dim(S_1 \cap S_2) + \dim(S_1 + S_2) = \dim S_1 + \dim S_2 \tag{0.1.7.1}$$

将此恒等式改写成

$$\dim(S_1 \cap S_2) = \dim S_1 + \dim S_2 - \dim(S_1 + S_2) \geqslant \dim S_1 + \dim S_2 - \dim V \tag{0.1.7.2}$$

它揭示了一个有用的事实：如果 $\delta = \dim S_1 + \dim S_2 - \dim V \geqslant 1$，那么子空间 $S_1 \cap S_2$ 的维数至少是 δ，从而它包含 δ 个线性无关的向量，也就是 $S_1 \cap S_2$ 的基中的任意 δ 个元素．特别地，$S_1 \cap S_2$ 包含一个非零向量．归纳法指出，如果 S_1，…，S_k 是 V 的子空间，且如果 $\delta = \dim S_1 + \cdots + \dim S_k - (k-1)\dim V \geqslant 1$，那么

$$\dim(S_1 \cap \cdots \cap S_k) \geqslant \delta \tag{0.1.7.3}$$

从而 $S_1 \cap \cdots \cap S_k$ 包含 δ 个线性无关的向量，特别地，它包含一个非零向量．

0.1.8 同构

如果 U 和 V 是同一个纯量域 $\mathbf{F}$ 上的向量空间，又如果 $f: U \to V$ 是满足如下条件的一个可逆函数：对所有 x，$y \in U$ 以及所有 a，$b \in \mathbf{F}$ 都有 $f(ax+by) = af(x) + bf(y)$，那么 f 就称为是一个**同构**(isomorphism)，而 U 和 V 就称为是**同构的**(isomorphic)（即“构造相同的”）．同一个域上的两个有限维向量空间是同构的，当且仅当它们有相同的维数．于是，$\mathbf{F}$ 上任意一个 n 维向量空间都与 $\mathbf{F}^n$ 同构．这样一来，任意一个 n 维实向量空间都同构于

4

$\mathbf{R}^n$，而任意一个 n 维复向量空间都同构于 $\mathbf{C}^n$．特别地，如果 V 是域 $\mathbf{F}$ 上一个 n 维向量空间，它有指定的基 $\mathcal{B} = \{x_1, \cdots, x_n\}$，那么，由于任何元素 $x \in V$ 都可以用唯一的方式表示成 $x = a_1 x_1 + \cdots + a_n x_n$，其中每个 $a_i \in \mathbf{F}$，故而我们可以将 x 和一个 n 元向量 $[x]_{\mathcal{B}} = [a_1 \cdots a_n]^{\mathrm{T}}$ 等同起来．对任何基 $\mathcal{B}$，映射 $x \to [x]_{\mathcal{B}}$ 都是 V 与 $\mathbf{F}^n$ 之间的一个同构．

0.2 矩阵

这里研究的基本对象可以用两种重要的方式来考虑：视为纯量的矩形阵列，或者视为两个向量空间之间的线性变换（对每个空间给出指定的基）．

0.2.1 矩形阵列

矩阵(matrix)是从域 $\mathbf{F}$ 中取出的纯量组成的一个 $m \times n$ 阵列．如果 $m = n$，该矩阵就称为**方阵**(square)．$\mathbf{F}$ 上所有 $m \times n$ 矩阵的集合记为 $M_{m,n}(\mathbf{F})$，而 $M_{n,n}(\mathbf{F})$ 通常记为 $M_n(\mathbf{F})$．向量空间 $M_{n,1}(\mathbf{F})$ 与 $\mathbf{F}^n$ 等同．如果 $\mathbf{F} = \mathbf{C}$，那么 $M_n(\mathbf{C})$ 就进一步简写为 M_n，而 $M_{m,n}(\mathbf{C})$ 则简写为 $M_{m,n}$．矩阵通常用大写字母表示，而矩阵里的元素一般用双下标的小写字母表示．例如，如果

$$A=\begin{bmatrix}2 & -\frac{3}{2} & 0\\ -1 & \pi & 4\end{bmatrix}=[a_{ij}]$$

那么 $A\in M_{2,3}(\mathbf{R})$ 的元素是 $a_{11}=2$，$a_{12}=-3/2$，$a_{13}=0$，$a_{21}=-1$，$a_{22}=\pi$，$a_{23}=4$. 给定矩阵的**子矩阵**(submatrix)是位于给定矩阵中指定的行和列处的元素组成的阵列. 例如，$[\pi\quad 4]$是 A 的一个(位于第 2 行，第 2 列和第 3 列处的)子矩阵.

假设 $A=[a_{ij}]\in M_{n,m}(\mathbf{F})$. A 的**主对角线**(main diagonal)是元素列 a_{11}，a_{22}，…，a_{qq}，其中 $q=\min[n, m]$. 有时将 A 的主对角线表示为一个向量 $\mathrm{diag}A=[a_{ii}]_{i=1}^{q}\in\mathbf{F}^q$ 是很方便的. A 的第 p 条**超对角线**(superdiagonal)是元素列 $a_{1,p+1}$，$a_{2,p+2}$，…，$a_{k,p+k}$，其中 $k=\min\{n, m-p\}$，$p=0, 1, 2, \cdots, m-1$；而 A 的第 p 条**次对角线**(subdiagonal)是元素列 $a_{p+1,1}$，$a_{p+2,2}$，…，$a_{p+l,l}$，其中 $l=\min\{n-p, m\}$，$p=0, 1, 2, \cdots, n-1$.

0.2.2 线性变换

设 U 是一个 n 维向量空间，而 V 是一个 m 维向量空间，它们都以 $\mathbf{F}$ 为基域；设$\mathcal{B}_U$ 是 U 的一组基，而$\mathcal{B}_V$ 则是 V 的一组基. 我们可以利用同构 $x\to[x]_{\mathcal{B}_U}$ 和 $y\to[y]_{\mathcal{B}_V}$ 将 U 和 V 中的向量分别表示为 $\mathbf{F}$ 上的 n 元向量以及 m 元向量. 一个**线性变换**(linear transformation)是一个函数 T：$U\to V$，它使得任何纯量 a_1，a_2 以及向量 x_1，x_2 都有 $T(a_1x_1+a_2x_2)=a_1T(x_1)+a_2T(x_2)$. 矩阵 $A\in M_{m,n}(\mathbf{F})$以下述方式与一个线性变换 T：$U\to V$ 相对应，即 $y=T(x)$当且仅当$[y]_{\mathcal{B}_V}=A[x]_{\mathcal{B}_U}$. 矩阵 A 被说成是**表示线性变换** T(相对于基$\mathcal{B}_U$ 和$\mathcal{B}_V$)，表示矩阵 A 依赖于基的选取. 在研究矩阵 A 时，我们意识到是对于一组特别选定的基研究线性变换，不过通常并不需要明显提及基.

0.2.3 与矩阵或线性变换相关的向量空间

$\mathbf{F}$ 上的任意一个 n 维向量空间都可以等同于 $\mathbf{F}^n$，我们可以把 $A\in M_{m,n}(\mathbf{F})$视为从 $\mathbf{F}^n$ 到 $\mathbf{F}^m$ 的线性变换(也可视为一个阵列)$x\to Ax$. 这个线性变换的**定义域**(domain)是 $\mathbf{F}^n$，它的**值域**(range)是 $\mathrm{range}A=\{y\in\mathbf{F}^m: y=Ax\}$(对某个 $x\in\mathbf{F}^n$)，它的**零空间**(null space)是 null space$A=\{x\in\mathbf{F}^n: Ax=0\}$. A 的值域是 $\mathbf{F}^m$ 的一个子空间，A 的零空间是 $\mathbf{F}^n$ 的子空间 A 的零空间的维数记为 nullity$A$(即 A 的零化度)，range$A$ 的维数记为 rank$A$. 这些数由**秩-零化度定理**(rank-nullity theorem)联系在一起：对 $A\in M_{m,n}(\mathbf{F})$有

$$\dim(\mathrm{range}A)+\dim(\text{null space}A)=\mathrm{rank}A+\mathrm{nullity}A=n \qquad (0.2.3.1)$$

A 的零空间是 $\mathbf{F}^n$ 中的一组向量，这些向量的元素满足 m 个齐次线性方程.

0.2.4 矩阵运算

矩阵加法对相同维数的阵列定义为逐个元素相加，并用+号表示("$A+B$")，它与线性变换(关于同一组基)的加法相对应，且它保持纯量域的交换性以及结合性. **零矩阵**(zero matrix)(所有元素均为零)是加法的单位元，而 $M_{m,n}(\mathbf{F})$则是 $\mathbf{F}$ 上的一个向量空间. 矩阵的乘法用矩阵并列("AB")表示，它与线性变换的复合相对应. 因此，矩阵乘法仅当 $A\in M_{m,n}(\mathbf{F})$且 $B\in M_{n,q}(\mathbf{F})$时才有定义. 它是结合的但通常不可交换. 例如

$$\begin{bmatrix}1 & 2\\6 & 8\end{bmatrix}=\begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\neq\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}=\begin{bmatrix}1 & 4\\3 & 8\end{bmatrix}$$

单位矩阵(identity matrix)

$$I=\begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}\in M_n(\mathbf{F})$$

是 $M_n(\mathbf{F})$ 中的乘法单位元，它的主对角线上的元素均为 1，而所有其他元素均为 0．单位矩阵以及它(作为一个**纯量矩阵**)的任意纯量倍数都与 $M_n(\mathbf{F})$ 中每个矩阵可交换，它们是有此性质的仅有的矩阵．矩阵乘法关于矩阵加法是可分配的．

在本书中始终用符号 0 表示下述诸量中的每一个：域的纯量零，向量空间的零向量，$\mathbf{F}^n$ 中的 n 元零向量(所有元素都等于 $\mathbf{F}$ 中的纯量零)以及 $M_{m,n}(\mathbf{F})$ 中的零矩阵(所有元素都等于纯量零)．符号 I 表示任意大小的单位矩阵．如果有出现混淆的可能，我们将用下标指明零矩阵以及单位矩阵的维数，例如 $0_{p,q}$，0_k 或者 I_k．

0.2.5 转置，共轭转置以及迹

如果 $A=[a_{ij}]\in M_{m,n}(\mathbf{F})$，则 A 的**转置**(transpose)(记为 A^{T})是 $M_{n,m}(\mathbf{F})$ 中的一个矩阵，后者的第 i 行第 j 列的元素是 a_{ji}，这就是说，行和列做了交换，反之亦然．例如

$$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}1 & 4\\2 & 5\\3 & 6\end{bmatrix}$$

当然有 $(A^{\mathrm{T}})^{\mathrm{T}}=A$．$A\in M_{m,n}(\mathbf{C})$ 的**共轭转置**(conjugate transpose)[有时称为**伴随**(adjoint)**矩阵**或者 Hermite **伴随**(Hermitian adjoint)**矩阵**]记为 A^*，其定义为 $A^*=\bar{A}^{\mathrm{T}}$，其中 $\bar{A}$ 是对 A 逐个元素取共轭复数．例如

$$\begin{bmatrix}1+i & 2-i\\-3 & -2i\end{bmatrix}^*=\begin{bmatrix}1-i & -3\\2+i & 2i\end{bmatrix}$$

转置与共轭转置都遵从**反序法则**(reverse-order law)：$(AB)^*=B^*A^*$ 以及 $(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}$．对乘积的复共轭而言无需反序：$\overline{AB}=\bar{A}\,\bar{B}$．如果 x，y 是同样大小的实向量或者复向量，那么 y^*x 是纯量，且其共轭转置与复共轭是相同的：$(y^*x)^*=\overline{y^*x}=x^*y=y^{\mathrm{T}}\bar{x}$．

许多重要的矩阵类都是用含有转置或者共轭转置的恒等式来定义的．例如，$A\in M_n(\mathbf{F})$ 称为是**对称的**(symmetric)，如果有 $A^{\mathrm{T}}=A$；它称为是**斜对称的**(skew symmetric)，如果有 $A^{\mathrm{T}}=-A$；称为是**正交的**(orthogonal)，如果有 $A^{\mathrm{T}}A=I$；$A\in M_n(\mathbf{C})$ 称为是 **Hermite 的**(Hermitian)，如果 $A^*=A$；它称为是**斜 Hermite 的**(skew Hermitian)，如果 $A^*=-A$；它称为是**本性 Hermite 的**(essentially Hermitian)，如果对某个 $\theta\in\mathbf{R}$，$e^{i\theta}A$ 是 Hermite 的；它称为是**酉的**(unitary)，如果 $A^*A=I$；它称为是**正规的**(normal)，如果 $A^*A=AA^*$．

每一个 $A\in M_n(\mathbf{F})$ 都可以用恰好一种方式写成 $A=S(A)+C(A)$，其中 $S(A)$ 是对称的，而 $C(A)$ 是斜对称的：$S(A)=\frac{1}{2}(A+A^{\mathrm{T}})$ 是 A 的**对称部分**，而 $C(A)=\frac{1}{2}(A-A^{\mathrm{T}})$ 则是 A 的**斜对称部分**．

每一个 $A\in M_{m,n}(\mathbf{C})$ 都可以用恰好一种方式写成 $A=B+\mathrm{i}C$，其中 B，$C\in M_{m,n}(\mathbf{R})$：$B=\frac{1}{2}(A+\overline{A})$ 是 A 的**实部**，而 $C=\frac{1}{2}(A-\overline{A})$ 则是 A 的**虚部**.

每一个 $A\in M_n(\mathbf{C})$ 都可以用恰好一种方式写成 $A=H(A)+\mathrm{i}K(A)$，其中 $H(A)$ 和 $K(A)$ 是 Hermite 的：$H(A)=\frac{1}{2}(A+A^*)$ 是 A 的 **Hermite 部分**，$\mathrm{i}K(A)=\frac{1}{2}(A-A^*)$ 是 A 的**斜 Hermite 部分**. 复矩阵或者实矩阵的表达式 $A=H(A)+\mathrm{i}K(A)$ 称为是它的 **Toeplitz 分解**(Toeplitz decomposition).

$A=[a_{ij}]\in M_{m,n}(\mathbf{F})$ 的**迹**(trace)是它的主对角线上元素之和 $\mathrm{tr}A=a_{11}+\cdots+a_{qq}$，其中 $q=\min\{m,n\}$. 对任何 $A=[a_{ij}]\in M_{m,n}(\mathbf{C})$，有 $\mathrm{tr}AA^*=\mathrm{tr}A^*A=\sum_{i,j}|a_{ij}|^2$，所以

$$\mathrm{tr}AA^*=0 \text{ 当且仅当 } A=0 \tag{0.2.5.1}$$

一个向量称为是**迷向的**(isotropic)，如果 $x^{\mathrm{T}}x=0$. 例如，$[1\quad \mathrm{i}]^{\mathrm{T}}\in\mathbf{C}^2$ 是一个非零的迷向向量. 在 $\mathbf{R}^n$ 中不存在非零的迷向向量.

0.2.6 矩阵乘法的基本知识

作为对于矩阵与向量乘法以及矩阵与矩阵乘法的寻常定义的补充，另外一些可供选择的视点可能是有用的.

1. 如果 $A\in M_{m,n}(\mathbf{F})$，$x\in\mathbf{F}^n$，$y\in\mathbf{F}^m$，那么(列)向量 Ax 是 A 的列的线性组合，这个线性组合的系数是 x 中的元素. 行向量 $y^{\mathrm{T}}A$ 则是 A 的行的线性组合，这一线性组合的系数是 y 中的元素.

2. 如果 b_j 是 B 的第 j 列，而 a_i^{T} 是 A 的第 i 行，那么 AB 的第 j 列是 Ab_j，而 AB 的第 i 行则是 $a_i^{\mathrm{T}}B$.

换言之，在矩阵乘积 AB 中，**用 A 左乘是乘以 B 的列，而用 B 右乘是乘以 A 的行.** 当其中有一个因子是对角矩阵时这一结果的一个重要的特例见(0.9.1).

假设 $A\in M_{m,p}(\mathbf{F})$，且 $B\in M_{n,q}(\mathbf{F})$. 设 a_k 是 A 的第 k 列，b_k 是 B 的第 k 列，则有下列结论.

3. 如果 $m=n$，那么 $A^{\mathrm{T}}B=[a_i^{\mathrm{T}}b_j]$：$A^{\mathrm{T}}B$ 位于 (i,j) 处的元素是纯量 $a_i^{\mathrm{T}}b_j$.

4. 如果 $p=q$，那么 $AB^{\mathrm{T}}=\sum_{k=1}^{n}a_kb_k^{\mathrm{T}}$：每一个求和项都是一个 $m\times n$ 矩阵，它是 a_k 与 b_k 的**外积**(outer product).

0.2.7 矩阵的列空间与行空间

$A\in M_{m,n}(\mathbf{F})$ 的值域也称为它的**列空间**(column space)，因为对于任何 $x\in\mathbf{F}^n$，Ax 都是 A 的列的一个线性组合(x 的元素是这个线性组合中的系数). $\mathrm{range}A$ 是 A 的列生成的子空间. 类似地，$\{y^{\mathrm{T}}A:\ y\in\mathbf{F}^m\}$ 称为 A 的**行空间**(row space). 如果 $A\in M_{m,n}(\mathbf{F})$ 的列空间包含在 $B\in M_{m,k}(\mathbf{F})$ 的列空间之中，那么就存在某个 $X\in M_{k,n}(\mathbf{F})$，使得 $A=BX$(反之亦然). X 的第 j 列中的元素告诉我们怎样将 A 的第 j 列表示成为 B 的列的一个线性组合.

如果 $A\in M_{m,n}(\mathbf{F})$，$B\in M_{m,q}(\mathbf{F})$，那么

$$\mathrm{range}A+\mathrm{range}B=\mathrm{range}[A\quad B] \tag{0.2.7.1}$$

如果 $A\in M_{m,n}(\mathbf{F})$，$B\in M_{p,n}(\mathbf{F})$，那么

$$\text{nullspace}A \cap \text{nullspace}B = \text{nullspace}\begin{bmatrix} A \\ B \end{bmatrix} \tag{0.2.7.2}$$

0.2.8　全 1 矩阵和全 1 向量

在 $\mathbf{F}^n$ 中，向量 $e=e_1+\cdots+e_n$ 的每个元素都是 1．矩阵 $J_n=ee^{\mathrm{T}}$ 的每个元素也都是 1．

0.3　行列式

在数学中，用单独一个数简要描述多元现象常常是有用的，而行列式函数就是其中之一例．其定义域是 $M_n(\mathbf{F})$（仅对方阵定义），且它可以用若干不同的方式表达．我们用 $\det A$ 来记 $A\in M_n(\mathbf{F})$ 的行列式．

0.3.1　按照行或者列用子式作 Laplace 展开

行列式可以用如下方法对 $A=[a_{ij}]\in M_n(\mathbf{F})$ 归纳定义．假设行列式已经对 $M_{n-1}(\mathbf{F})$ 有定义了，令 $A_{ij}\in M_{n-1}(\mathbf{F})$ 表示从 A 中去掉第 i 行和第 j 列得到的 $A\in M_n(\mathbf{F})$ 的子矩阵．那么，对于任何 i，$j\in\{1, \cdots, n\}$，我们有

$$\det A = \sum_{k=1}^{n}(-1)^{i+k}a_{ik}\det A_{ik} = \sum_{k=1}^{n}(-1)^{k+j}a_{kj}\det A_{kj} \tag{0.3.1.1}$$

第一个和是**用第 i 行子式作 Laplace 展开**(Laplace expansion by minors along row i)；而第
8
二个和是**用第 j 列子式作 Laplace 展开**(Laplace expansion by minors along column j)．这一归纳表述始于将 1×1 矩阵的行列式定义为单个元素的值．从而

$$\det[a_{11}] = a_{11}$$

$$\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

如此等等．注意，有 $\det A^{\mathrm{T}}=\det A$，$\det A^*=\overline{\det A}$（如果 $A\in M_n(\mathbf{C})$）以及 $\det I=1$．

0.3.2　交错和以及置换

$\{1, \cdots, n\}$ 的**置换**(permutation)是一个一对一函数 σ：$\{1, \cdots, n\}\to\{1, \cdots, n\}$．**恒等置换**(identity permutation)满足 $\sigma(i)=i$（对每个 $i=1, \cdots, n$）．$\{1, \cdots, n\}$ 有 $n!$ 个不同的置换，且所有这样的置换组成的集合关于函数的复合运算构成一个群．

与(0.3.1)中低维数的例子相一致，对 $A=[a_{ij}]\in M_n(\mathbf{F})$ 我们有交错表示

$$\det A = \sum_{\sigma}\left(\text{sgn}\sigma\prod_{i=1}^{n}a_{i\sigma(i)}\right) \tag{0.3.2.1}$$

其中的求和经过 $\{1, \cdots, n\}$ 的所有 $n!$ 个置换，而 $\text{sgn}\sigma$，即置换 σ 的“符号”(sign 或者 signum)是 $+1$ 或者 -1，要视从起始状态 $\{1, \cdots, n\}$ 到达 σ 所需要做的对换（两两交换）的

最少次数是偶数还是奇数而定．我们称置换 σ 是**偶的**(even)，如果 $\operatorname{sgn}\sigma=+1$；称 σ 是**奇的**(odd)，如果 $\operatorname{sgn}\sigma=-1$.

如果(0.3.2.1)中的 $\operatorname{sgn}\sigma$ 代之以 σ 的某个另外的函数，那么代替 $\det A$，我们就得到了**推广的矩阵函数**(generalized matrix function)．例如，A 的**积和式**(permanent)用 $\operatorname{per}A$ 表示，它是用恒等于 $+1$ 的函数取代 $\operatorname{sgn}\sigma$ 而得到的．

0.3.3 行和列的初等运算

有三种关于行和列的简单而基本的运算，称为**行和列的初等运算**(elementary row and column operation)，它们可以用来将矩阵(方阵或者矩形阵)变换成简单的形式，以方便求解线性方程、求秩、计算行列式与求方阵的逆之类的问题．我们集中讨论**行运算**(row operation)，行运算可以通过用矩阵从左边作用而实施之．**列运算**(column operation)依照类似的方式定义和使用，此时矩阵需从右边作用之． 9

类型 1：交换两行

对于 $i\neq j$，交换 A 的第 i 行和第 j 行可以用

$$\begin{bmatrix}
1 & & & & & & & & & \\
 & \ddots & & & & & & & & \\
 & & 0 & & \cdots & & 1 & & & \\
 & & & 1 & & & & & & \\
 & & \vdots & & \ddots & & \vdots & & & \\
 & & & & & 1 & & & & \\
 & & 1 & \cdots & \cdots & \cdots & 0 & & & \\
 & & & & & & & 1 & & \\
 & & & & & & & & \ddots & \\
 & & & & & & & & & 1
\end{bmatrix}$$

左乘 A 得到．其中两个位于对角线之外的 1，它们的位置是 (i,j) 以及 (j,i)，对角线上的两个 0 位于 (i,i) 以及 (j,j)，而所有未指出的元素均为 0.

类型 2：用一个非零的纯量乘以一行

用一个非零的纯量 c 乘以 A 的第 i 行可以用

$$\begin{bmatrix}
1 & & & & & & \\
 & \ddots & & & & & \\
 & & 1 & & & & \\
 & & & c & & & \\
 & & & & 1 & & \\
 & & & & & \ddots & \\
 & & & & & & 1
\end{bmatrix}$$

左乘 A 得到．其中 (i,i) 处的元素是 c，对角线上所有其他元素均为 1，而所有未指出的元素均为 0.

类型 3：将一行的一个纯量倍数加到另一行

对于 $i \neq j$，将 A 的第 i 行的 c 倍加到第 j 行可以用

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & c & & \ddots & \\ & & & & & 1 \end{bmatrix}$$

左乘 A 得到．其中 (j, i) 处的元素是 c，所有主对角线元素均为 1，而所有未指出的元素均为 0．所给出的矩阵描述了 $j > i$ 时的情形．

三个初等行(列)运算中的每一个运算的矩阵都正好是对单位矩阵 I 施行相应运算(对行运算用左乘；而对列运算用右乘)所产生的结果．类型 1 的运算对行列式的作用效果是用 -1 乘以它；类型 2 的运算的作用效果是用非零纯量 c 乘以它；而类型 3 的运算不改变行列式．有一行为零的方阵的行列式为零．方阵的行列式为零，当且仅当矩阵的行的
10 某个子集是线性相关的．

0.3.4 化简的行梯形矩阵

对每个 $A = [a_{ij}] \in M_{m,n}(\mathbf{F})$，在 $M_{m,n}(\mathbf{F})$ 中有一个(唯一的)标准型与之相对应，此即**化简的行梯形矩阵**(Reduced Row Echelon Form，RREF)，也称为 **Hermite 标准型**(Hermite normal form)．

如果 A 的一行不全为零，其**首项**(leading entry)是它的第一个非零元素．则 RREF 的确切的表述如下：

(a) 任何零行都出现在矩阵的底部．

(b) 任何非零行的首项都是 1．

(c) 首项所在的列中其余元素均为零．

(d) 首项从左向右以阶梯状的模式出现，这就是说，如果第 i 行是非零行，且 a_{ik} 是它的首项，那么要么 $i = m$，要么第 $i+1$ 行是零，要么第 $i+1$ 行的首项是 $a_{i+1,l}$，其中 $l > k$．

例如

$$\begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & \pi \\ 0 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

就有 RREF 之形式．

如果 $R \in M_{m,n}(\mathbf{F})$ 是 A 的 RREF，那么 $R = EA$，其中非奇异矩阵 $E \in M_m(\mathbf{F})$ 是与将 A 化简为 RREF 所执行的一系列行初等运算相对应的类型 1、类型 2 以及类型 3 的初等矩阵的乘积．

$A \in M_n(\mathbf{F})$ 的行列式不为零，当且仅当它的 RREF 是 I_n．$\det A$ 的值可以通过记录将其导向 RREF 所作的每一个初等运算的行列式的效果来加以计算．

对于线性方程组 $Ax=b$，其中 $A\in M_{m,n}(\mathbf{F})$以及 $b\in\mathbf{F}^m$ 是给定的，而 $x\in\mathbf{F}^n$ 是未知的，如果对 A 和 b 这两者执行同样的一串行初等运算，那么该方程组的解集不发生变化. 通过检查$[A\quad b]$的 RREF，就揭示出 $Ax=b$ 的解. 由于 RREF 是唯一的，对于给定的 A_1，$A_2\in M_{m,n}$ 以及给定的 b_1，$b_2\in\mathbf{F}^m$，线性方程组 $A_1x=b_1$ 和 $A_2x=b_2$ 有同样的解集，当且仅当$[A_1\quad b_1]$和$[A_2\quad b_2]$有同样的 RREF.

0.3.5　积性

行列式函数的一个关键性质是它是积性的：对于 A，$B\in M_n(\mathbf{F})$有

$$\det AB=\det A\det B$$

这可以利用对 A 和 B 两者作行化简的初等运算予以证明.

0.3.6　行列式的函数特征

如果把行列式视为一个矩阵的每一行(每一列)的函数，而让其他的行(列)固定不变，则 Laplace 展开(0.3.1.1)表明，行列式是任意给定的一行(列)的元素的线性函数. 我们如下总结这个性质：称函数 $A\to\det A$ 关于 A 的行(列)是**多重线性的**(multilinear). 11

行列式函数 $A\to\det A$ 是具有下列性质的唯一的函数 f：$M_n(\mathbf{F})\to\mathbf{F}$：

(a) 关于行变量是多重线性的；

(b) 交错性：对 A 作任何类型 1 的运算改变 $f(A)$的符号；

(c) 规范性：$f(I)=1$.

积和式函数(permanent function)也是多重线性的(作为另一类推广的矩阵函数)，而且它也是规范的，但它不是交错的.

0.4　秩

0.4.1　定义

如果 $A\in M_{m,n}(\mathbf{F})$，则 $\operatorname{rank}A=\dim\operatorname{range}A$ 是 A 的最长的线性无关列向量组的长度. 长度等于**秩**(rank)的线性无关的列向量组可能有多于一组. 一个值得注意的事实是 $\operatorname{rank}A^{\mathrm{T}}=\operatorname{rank}A$. 这样一来，可以用 A 的最长的线性无关的行向量组的长度来作为秩的等价定义：行秩=列秩.

0.4.2　秩和线性方程组

设给定 $A\in M_{m,n}(\mathbf{F})$和 $b\in\mathbf{F}^n$. 线性方程组 $Ax=b$ 可能没有解，可能有一组解，或者有无穷多组解；这些就是仅有的可能性. 如果它至少有一组解，则该线性方程组是**相容的**(consistent)；如果它没有解，该线性方程组就是**不相容的**(inconsistent). 线性方程组 $Ax=b$ 是相容的，当且仅当 $\operatorname{rank}[A\quad b]=\operatorname{rank}A$. 矩阵$[A\quad b]\in M_{m,n+1}(\mathbf{F})$是**增广矩阵**(augmented matrix). 说线性方程组的增广矩阵和**系数矩阵** A 有相同的秩，等同于说 b 是 A 的列的线性组合. 在此情形，将 b 添加到 A 的列中并不增加它的秩. 线性方程组 $Ax=b$ 的一组**解**(solution)就是将 b 表示成 A 的列的线性组合时表示式中的系数作元素得到的向量 x.

0.4.3 RREF 与秩

初等运算不改变矩阵的秩，从而 A 的秩与 A 的 RREF 的秩相等，这个秩恰好等于其 RREF 中非零行的个数. 然而，事实上通过计算 RREF 来计算秩是不明智的. 在数值计算中间环节产生的舍入误差有可能使得 RREF 中的零行变成非零行出现，从而会影响到对于秩的判断.

0.4.4 秩的特征

关于一个给定矩阵 $A\in M_{m,n}(\mathbf{F})$ 的如下诸命题是等价的，每一个命题可能在某个不同情况下有用. 注意(b)和(c)中的关键点在于矩阵的列向量组或者行向量组的线性无关性.

(a) $\operatorname{rank}A=k$.

(b) A 的行向量中有 k 个(且不多于 k 个)是线性无关的.

(c) A 的列向量中有 k 个(且不多于 k 个)是线性无关的.

(d) 存在 A 的某个 $k\times k$ 子矩阵有非零的行列式，且 A 的每一个 $(k+1)\times(k+1)$ 子矩
12 阵的行列式都为零.

(e) $\dim(\operatorname{range}A)=k$.

(f) 存在 k 个(但不多于 k 个)线性无关的向量 $b_1,\cdots,b_k$，使得对每个 $j=1,\cdots,k$，线性方程组 $Ax=b_j$ 都是相容的.

(g) $k=n-\dim(\operatorname{nullspace}A)$(**秩-零化度定理**).

(h) $k=\min\{p: A=XY^{\mathrm{T}}$，对某个 $X\in M_{m,p}(\mathbf{F})$，$Y\in M_{n,p}(\mathbf{F})\}$.

(i) 对某些 $x_1,\cdots,x_p\in\mathbf{F}^m$，$y_1,\cdots,y_p\in\mathbf{F}^n$ 有 $k=\min\{p: A=x_1y_1^{\mathrm{T}}+\cdots+x_py_p^{\mathrm{T}}\}$.

0.4.5 有关秩的不等式

有如下关于秩的基本不等式.

(a) 如果 $A\in M_{m,n}(\mathbf{F})$，那么 $\operatorname{rank}A\leqslant\min\{m,n\}$.

(b) 若从一个矩阵中删去一行或多行(一列或多列)，则所得到的子矩阵的秩不大于原矩阵的秩.

(c) **Sylvester 不等式**：如果 $A\in M_{m,k}(\mathbf{F})$ 且 $B\in M_{k,n}(\mathbf{F})$，那么

$$(\operatorname{rank}A+\operatorname{rank}B)-k\leqslant\operatorname{rank}AB\leqslant\min\{\operatorname{rank}A,\operatorname{rank}B\}$$

(d) **秩-和不等式**(rank-sum inequality)：如果 $A,B\in M_{m,n}(\mathbf{F})$，那么

$$|\operatorname{rank}A-\operatorname{rank}B|\leqslant\operatorname{rank}(A+B)\leqslant\operatorname{rank}A+\operatorname{rank}B \tag{0.4.5.1}$$

其中第二个不等式中的等号成立，当且仅当 $(\operatorname{range}A)\cap(\operatorname{range}B)=\{0\}$ 以及 $(\operatorname{range}A^{\mathrm{T}})\cap(\operatorname{range}B^{\mathrm{T}})=\{0\}$. 如果 $\operatorname{rank}B=1$，那么

$$|\operatorname{rank}(A+B)-\operatorname{rank}A|\leqslant 1 \tag{0.4.5.2}$$

特别地，改变矩阵的一个元素至多将它的秩改变 1.

(e) **Frobenius 不等式**：如果 $A\in M_{m,k}(\mathbf{F})$，$B\in M_{k,p}(\mathbf{F})$，$C\in M_{p,n}(\mathbf{F})$，那么

$$\operatorname{rank}AB+\operatorname{rank}BC\leqslant\operatorname{rank}B+\operatorname{rank}ABC$$

其中的等号成立，当且仅当存在矩阵 X 和 Y，使得 $B=BCX+YAB$.

0.4.6 有关秩的等式

下面是与秩有关的一些等式.

(a) 如果 $A\in M_{m,n}(\mathbf{C})$，则 $\operatorname{rank}A^*=\operatorname{rank}A^{\mathrm{T}}=\operatorname{rank}\overline{A}=\operatorname{rank}A$.

(b) 如果 $A\in M_m(\mathbf{F})$ 与 $C\in M_n(\mathbf{F})$ 非奇异，又 $B\in M_{m,n}(\mathbf{F})$，那么 $\operatorname{rank}AB=\operatorname{rank}B=\operatorname{rank}BC=\operatorname{rank}ABC$，也就是说，用一个非奇异矩阵左乘或者右乘不改变原矩阵的秩.

(c) 如果 A，$B\in M_{m,n}(\mathbf{F})$，那么 $\operatorname{rank}A=\operatorname{rank}B$，当且仅当存在一个非奇异矩阵 $X\in M_m(\mathbf{F})$ 以及一个非奇异矩阵 $Y\in M_n(\mathbf{F})$，使得 $B=XAY$.

(d) 如果 $A\in M_{m,n}(\mathbf{C})$，那么 $\operatorname{rank}A^*A=\operatorname{rank}A$.

(e) **满秩分解**(full-rank factorization)：如果 $A\in M_{m,n}(\mathbf{F})$，那么 $\operatorname{rank}A=k$ 当且仅当对某个 $X\in M_{m,k}(\mathbf{F})$ 以及 $Y\in M_{n,k}(\mathbf{F})$ 有 $A=XY^{\mathrm{T}}$，X 与 Y 各自的列向量都是线性无关的. 对于某个非奇异的 $B\in M_k(\mathbf{F})$，等价的分解 $A=XBY^{\mathrm{T}}$ 也可能是有用的. 特别地，$\operatorname{rank}A=1$ 当且仅当对某个非零向量 $x\in\mathbf{F}^m$ 以及 $y\in\mathbf{F}^n$ 有 $A=xy^{\mathrm{T}}$.

(f) 如果 $A\in M_{m,n}(\mathbf{F})$，那么 $\operatorname{rank}A=k$ 当且仅当存在非奇异的矩阵 $S\in M_m(\mathbf{F})$ 以及 $T\in$ 13
$M_n(\mathbf{F})$，使得 $A=S\begin{bmatrix}I_k & 0\\ 0 & 0\end{bmatrix}T$.

(g) 设 $A\in M_{m,n}(\mathbf{F})$. 如果 $X\in M_{n,k}(\mathbf{F})$ 以及 $Y\in M_{m,k}(\mathbf{F})$，又如果 $W=Y^{\mathrm{T}}AX$ 是非奇异的，那么

$$\operatorname{rank}(A-AXW^{-1}Y^{\mathrm{T}}A)=\operatorname{rank}A-\operatorname{rank}AXW^{-1}Y^{\mathrm{T}}A \tag{0.4.6.1}$$

当 $k=1$ 时，这就是 **Wedderburn 秩 1 化简公式**(Wedderburn rank-one reduction formula)：如果 $x\in\mathbf{F}^n$ 且 $y\in\mathbf{F}^m$，又如果 $\omega=y^{\mathrm{T}}Ax\neq 0$，那么

$$\operatorname{rank}(A-\omega^{-1}Axy^{\mathrm{T}}A)=\operatorname{rank}A-1 \tag{0.4.6.2}$$

反之，如果 $\sigma\in\mathbf{F}$，$u\in\mathbf{F}^n$，$v\in\mathbf{F}^m$，且 $\operatorname{rank}(A-\sigma uv^{\mathrm{T}})<\operatorname{rank}A$，那么 $\operatorname{rank}(A-\sigma uv^{\mathrm{T}})=\operatorname{rank}A-1$，且存在 $x\in\mathbf{F}^n$ 以及 $y\in\mathbf{F}^m$，使得 $u=Ax$，$v=A^{\mathrm{T}}y$，$y^{\mathrm{T}}Ax\neq 0$ 以及 $\sigma=(y^{\mathrm{T}}Ax)^{-1}$.

0.5 非奇异性

一个线性变换或者矩阵称为是**非奇异的**(nonsingular)，如果它仅对输入 0 才得到输出 0. 反之则称它是**奇异的**(singular). 如果 $A\in M_{m,n}(\mathbf{F})$ 且 $m<n$，那么 A 必定是奇异的. $A\in M_n(\mathbf{F})$ 是**可逆的**(invertible)，如果存在一个矩阵 $A^{-1}\in M_n(\mathbf{F})$($A$ 的**逆**)，使得 $A^{-1}A=I$. 如果 $A\in M_n$ 且 $A^{-1}A=I$，那么 $AA^{-1}=I$；这就是说，A^{-1} 是**左逆**(left inverse)当且仅当它是**右逆**(right inverse)；A^{-1} 只要存在，就是唯一的.

回忆种种关于方阵何时才是非奇异的判别法是有用的. 对给定的 $A\in M_n(\mathbf{F})$，以下诸结论等价.

(a) A 是非奇异的.

(b) A^{-1} 存在.

(c) $\operatorname{rank}A=n$.

(d) A 的行线性无关.

(e) A 的列线性无关.

(f) $\det A \neq 0$.

(g) A 的值域的维数是 n.

(h) A 的零空间的维数是 0.

(i) 对每个 $b \in \mathbf{F}^n$，$Ax=b$ 都是相容的.

(j) 如果 $Ax=b$ 相容，那么解是唯一的.

(k) 对每个 $b \in \mathbf{F}^n$，$Ax=b$ 都有唯一解.

(l) $Ax=0$ 仅有解 $x=0$.

(m) 0 不是 A 的特征值(见第 1 章).

对于有限维向量空间 V 上的线性变换 $T: V \to V$，条件(g)和(h)是等价的；也就是说，对每一个 $y \in V$，$Tx=y$ 都有解 x，当且仅当满足 $Tx=0$ 的仅有的 x 是 $x=0$，当且仅当对每个 $y \in V$，$Tx=y$ 有唯一解 x.

$M_n(\mathbf{F})$ 中的非奇异矩阵构成一个群，称为**一般线性群**(general linear group)，常用 $GL(n, \mathbf{F})$来表示.

如果 $A \in M_n(\mathbf{F})$非奇异，那么$[(A^{-1})^{\mathrm{T}} A^{\mathrm{T}}]^{\mathrm{T}} = A(A^{-1}) = I$，所以$(A^{-1})^{\mathrm{T}} A^{\mathrm{T}} = I$，这就意味着$(A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1}$. 将$(A^{-1})^{\mathrm{T}}$ 或者$(A^{\mathrm{T}})^{-1}$写成 $A^{-\mathrm{T}}$是很方便的. 如果 $A \in M_n(\mathbf{C})$非
14 奇异，那么$(A^{-1})^* = (A^*)^{-1}$，从而可以确保将这两者写成 A^{-*}.

0.6 Euclid 内积与范数

0.6.1 定义

纯量$\langle x, y\rangle = y^* x$ 是 x，$y \in \mathbf{C}^n$ 的 **Euclid 内积**(Euclidean inner product)，也称**标准内积**(standard inner product)、**常用内积**(usual inner product)、**纯量积**(scalar product)、**点积**(dot product). $\mathbf{C}^n$ 上的 **Euclid 范数**(Euclidean norm)函数是实值函数$\|x\|_2 = \langle x, x\rangle^{1/2} = (x^* x)^{1/2}$，Euclid 范数也称为**常用范数**(usual norm)、**Euclid 长度**(Euclidean length). 这个函数的两个重要性质是：对所有非零的 $x \in \mathbf{C}^n$ 有$\|x\|_2 > 0$ 以及对所有 $x \in \mathbf{C}^n$ 和所有 $\alpha \in \mathbf{C}$ 有$\|\alpha x\|_2 = |\alpha| \|x\|_2$.

函数$\langle \cdot, \cdot \rangle$：$\mathbf{C}^n \times \mathbf{C}^n \to \mathbf{C}$ 关于第一个变量是线性的，关于第二个变量是**共轭线性的**(conjugate linear)；也就是说，对所有 α，$\beta \in \mathbf{C}$ 以及 y_1，$y_2 \in \mathbf{C}^n$ 有$\langle \alpha x_1 + \beta x_2, y\rangle = \alpha\langle x_1, y\rangle + \beta\langle x_2, y\rangle$以及$\langle x, \alpha y_1 + \beta y_2\rangle = \bar{\alpha}\langle x, y_1\rangle + \bar{\beta}\langle x, y_2\rangle$成立. 如果 V 是一个实的或者复的向量空间，且 $f: V \times V \to \mathbf{C}$ 关于第一个变量是线性的函数而关于第二个变量是共轭线性的函数，我们就称 f 在 V 上是**半双线性的**(sesquilinear)；f 是 V 上的**半内积**(semi-inner product)，如果它在 V 上是半双线性的，且对每个 $x \in V$ 都有 $f(x, x) \geqslant 0$；f 是 V 上的**内积**(inner product)，如果它在 V 上是半双线性的，且对每个非零的 $x \in V$ 都有 $f(x, x) > 0$. **内积空间**(inner product space)是元素对(V, f)，其中 V 是实的或者复的向量空间，而 f 则是 V 上的内积.

0.6.2 正交性和标准正交性

两个向量 x，$y \in \mathbf{C}^n$ 是**正交的**(orthogonal)，如果$\langle x, y\rangle = 0$. 在 $\mathbf{R}^2$ 和 $\mathbf{R}^3$ 中，“正交”

有常规的几何解释——“垂直”．一组向量 x_1，…，$x_m \in \mathbf{C}^n$ 说成是正交的，如果对所有不同的 i，$j \in \{1,\cdots,m\}$ 都有 $\langle x_i, x_j \rangle = 0$．一组非零的正交向量是线性无关的．Euclid 范数为 1 的向量称为是**标准化的**(normalized)(**单位向量**，unit vector)．对任何非零的 $x \in \mathbf{C}^n$，$x/\|x\|_2$ 是单位向量．一组正交向量是标准正交的向量组，如果它的每一个元素都是单位向量．一组标准正交的向量是线性无关的．这些概念中的每一个都在内积空间中有直接的推广．

0.6.3 Cauchy-Schwarz 不等式

Cauchy-Schwarz 不等式说的是，对所有 x，$y \in \mathbf{C}^n$ 有

$$|\langle x,y\rangle| \leqslant \|x\|_2 \|y\|_2$$

其中的等式成立，当且仅当其中有一个向量是另一个的纯量倍数．两个非零实向量 x，$y \in \mathbf{R}^n$ 之间的**角度**(angle)θ 定义为

$$\cos\theta = \frac{\langle x,y\rangle}{\|x\|_2 \|y\|_2}, \quad 0 \leqslant \theta \leqslant \pi \tag{0.6.3.1}$$

0.6.4 Gram-Schmidt 标准正交化

一个内积空间中的任意有限多个线性无关的向量可以用一组有同样生成空间的标准正交向量组替代．这种替代可以用许多方式来实现，其中有一种系统的方法可以实现，而且此法有一个有用的特殊性质．**Gram-Schmidt 方法**从一组向量 v_1，…，v_n 着手，产生出一组标准正交的向量 z_1，…，z_n(如果给定的向量组是线性无关的)，使得对每个 $k=1$，…，n 都有 $\mathrm{span}\{z_1,\cdots,z_k\}=\mathrm{span}\{x_1,\cdots,x_k\}$．向量 z_i 可以如下依次计算：设 $y_1=x_1$ 并将它标准化：$z_1=y_1/\|y_1\|_2$．设 $y_2=x_2-\langle x_2, z_1\rangle z_1$ (y_2 与 z_1 正交)并将它标准化：$z_2=y_2/\|y_2\|_2$．一旦 z_1，…，z_{k-1} 已经确定，则向量

$$y_k = x_k - \langle x_k, z_{k-1}\rangle z_{k-1} - \langle x_k, z_{k-2}\rangle z_{k-2} - \cdots - \langle x_k, z_1\rangle z_1$$ 15

与 z_1，…，z_{k-1} 正交；将其标准化：$z_k=y_k/\|y_k\|_2$．继续下去直到 $k=n$ 为止．如果记 $Z=[z_1 \ \cdots \ z_n]$ 以及 $X=[x_1 \ \cdots \ x_n]$，则 Gram-Schmidt 方法给出分解式 $X=ZR$，其中方阵 $R=[r_{ij}]$ 是非奇异的且是上三角的，即只要 $i>j$ 都有 $r_{ij}=0$．

如果向量 x_1，…，x_k 是标准正交的，且 x_1，…，x_k，x_{k+1}，…，x_n 是线性无关的，对后一组向量应用 Gram-Schmidt 方法就得出标准正交的向量组 x_1，…，x_k，z_{k+1}，…，z_n．

Gram-Schmidt 方法可以应用于任意一组有限多个向量，无论它们是线性无关的还是线性相关的．如果 x_1，…，x_n 线性相关，则 Gram-Schmidt 方法对于使得 x_k 是 x_1，…，x_{k-1} 的线性组合的最小的 k 值产生一个向量 $y_k=0$．

0.6.5 标准正交基

内积空间的一组**标准正交基**(orthonormal basis)是其元素构成一组标准正交向量的基．由于任何一组有限的有序基都可以用 Gram-Schmidt 方法转变成一组标准正交基，任意一个有限维的内积空间都有标准正交基，且任意一组标准正交向量都可以扩充成为一组标准正交基．与这样一组基打交道令人愉悦，因为内积的计算中交叉项相乘的结果全都等于零．

0.6.6 正交补

给定任意一个子集 $S\subset\mathbf{C}^n$，S 的**正交补**(orthogonal complement)是集合 $S^{\perp}=\{x\in\mathbf{C}^n: x^*y=0$，对所有 $y\in S\}$. 即使 S 不是子空间，$S^{\perp}$ 也总是一个子空间. 我们有$(S^{\perp})^{\perp}=\mathrm{span}S$，且当 S 是一个子空间时有$(S^{\perp})^{\perp}=S$. 恒有 $\dim S^{\perp}+\dim(S^{\perp})^{\perp}=n$. 如果 S_1 和 S_2 是子空间，那么就有$(S_1+S_2)^{\perp}=S_1^{\perp}\cap S_2^{\perp}$.

对给定的 $A\in M_{m,n}$，$\mathrm{range}A$ 是 $\mathrm{nullspace}A^*$ 的正交补. 于是，对给定的 $b\in\mathbf{C}^m$，线性方程组 $Ax=b$ 有一组解(不一定是唯一的)，当且仅当对每个满足 $A^*z=0$ 的 $z\in\mathbf{C}^m$ 都有 $b^*z=0$. 有时把这种等价性称为 **Fredholm 择一性**(alternative)(也称为**择一定理**)——下面两个命题中恰有一个为真：或者(1)$Ax=b$ 有解；或者(2)$A^*y=0$ 有满足 $y^*b\neq 0$ 的解.

如果 $A\in M_{m,n}$且 $B\in M_{m,q}$，又有 $X\in M_{m,r}$以及 $Y\in M_{m,s}$，又如果 $\mathrm{range}X=\mathrm{nullspace}A^*$ 且 $\mathrm{range}Y=\mathrm{nullspace}B^*$，那么我们就有以下与(0.2.7.1)以及(0.2.7.2)相伴的结果：

$$\mathrm{range}A\cap\mathrm{range}B=\mathrm{nullspace}\begin{bmatrix}X^*\\Y^*\end{bmatrix}\tag{0.6.6.1}$$

0.7 集合与矩阵的分划

集合$\mathcal{S}$的一个**分划**(partition)是$\mathcal{S}$的子集组成的集合，它使得$\mathcal{S}$的每一个元素是其中一个且也仅为其中一个子集的元素. 例如，集合$\{1, 2, \cdots, n\}$的分划是其子集 α_1，…，α_t(称为**指标集**，index set)的一个集合，它使得介于 1 和 n 之间的每个整数都在这些指标集中且仅在其中一个之中. $\{1, 2, \cdots, n\}$的**顺序分划**(sequential partition)是这样一种分划，其指标
16 集有如下特殊的形式 $\alpha_1=\{1, \cdots, i_1\}$，$\alpha_2=\{i_1+1, \cdots, i_2\}$，…，$\alpha_t=\{i_{t-1}+1, \cdots, n\}$.

矩阵的**分划**(partition)是将矩阵划分成子矩阵的分解，它使得原始矩阵中的每个元素在且仅在一个子矩阵中. 分划矩阵对于理解有用的构造通常是一种便捷的工具. 例如，按照列所作的分划 $B=[b_1\cdots b_n]\in M_n(\mathbf{F})$揭示出矩阵乘积的表达式 $AB=[Ab_1\cdots Ab_n]$(根据 AB 的列来分划).

0.7.1 子矩阵

设 $A\in M_{m,n}(\mathbf{F})$. 对指标集 $\alpha\subseteq\{1, \cdots, m\}$和 $\beta\subseteq\{1, \cdots, n\}$，我们用 $A[\alpha, \beta]$来记 A 的位于由 α 所指定的那些行以及由 β 所指定的那些列的位置上的那些元素所构成的(子)矩阵. 例如

$$\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}[\{1,3\},\{1,2,3\}]=\begin{bmatrix}1&2&3\\7&8&9\end{bmatrix}$$

如果 $\alpha=\beta$，则子矩阵 $A[\alpha]=A[\alpha, \alpha]$就是 A 的一个**主子矩阵**(principal submatrix). 一个 $n\times n$ 矩阵有$\binom{n}{k}$个不同的 k 阶主子矩阵.

对 $A\in M_n(\mathbf{F})$以及 $k\in\{1, \cdots, n\}$，$A[\{1, \cdots, k\}]$是**前主子矩阵**(leading principal submatrix)，而 $A[\{k,\cdots,n\}]$则是**尾主子矩阵**(trailing principal submatrix).

通常用删除行或列的方法比用加入行或列的方法来指明子矩阵或者主子矩阵更为方便．这可以用设置指标集予以实现．设 $\alpha^c=\{1,\cdots,m\}\setminus\alpha$ 以及 $\beta^c=\{1,\cdots,n\}\setminus\beta$ 分别表示与 α 和 β 互补的指标集．那么 $A[\alpha^c,\beta^c]$就是从原矩阵中**删去**由 α 指定的行以及删去由 β 指定的列之后所得到的子矩阵．例如，子矩阵 $A[\alpha,\varnothing^c]$包含 A 中由 α 指定的那些行；$A[\varnothing^c,\beta]$则包含 A 中由 β 指定的那些列．

A 中一个 $r\times r$ 子矩阵的行列式称为一个**子式**(minor)；如果我们希望指出这个子矩阵的大小，就称它的行列式是一个 r **阶子式**(minor of size r)．如果一个 $r\times r$ 子矩阵是一个**主子矩阵**，那么它的行列式就是一个(r 阶)**主子式**(principal minor)；如果该子矩阵是一个**前主子矩阵**，那么它的行列式就是一个**前主子式**(leading principal minor)；如果该子矩阵是一个**尾主子矩阵**，那么它的行列式就是一个**尾主子式**(trailing principal minor)．我们约定，空的主子式是 1，即 $\det A[\varnothing]=1$.

像出现在 Laplace 展开式(0.3.1.1)中那样带有符号的子式$[(-1)^{i+j}\det A_{ij}]$称为**代数余子式**(cofactor)；如果希望指出其子矩阵的大小，就将这个带符号的行列式称为一个 r **阶代数余子式**(cofactor of size r).

0.7.2 分划，分块矩阵以及乘法

如果 $\alpha_1,\cdots,\alpha_t$ 构成$\{1,\cdots,m\}$的一个分划，而 $\beta_1,\cdots,\beta_s$ 构成$\{1,\cdots,n\}$的一个分划，那么矩阵 $A[\alpha_i,\beta_j]$($1\leqslant i\leqslant t$，$1\leqslant j\leqslant s$)就构成矩阵 $A\in M_{m,n}(\mathbf{F})$的一个分划．如果 $A\in M_{m,n}(\mathbf{F})$的分划以及 $B\in M_{n,p}(\mathbf{F})$的分划使得$\{1,\cdots,n\}$的两个分划完全相同，则这两个矩阵分划就称为是**共形的**(conformal)．在此情形有

$$(AB)[\alpha_i,\gamma_j]=\sum_{k=1}^{s}A[\alpha_i,\beta_k]B[\beta_k,\gamma_j] \tag{0.7.2.1}$$

其中各自的子矩阵 $A[\alpha_i,\beta_k]$以及 $B[\beta_k,\gamma_j]$的集合分别是 A 和 B 的共形分划．(0.7.2.1)的左边是乘积 AB 的一个子矩阵(用常规方法计算)，而右边的每一个求和项都是一个标准的矩阵乘积．于是，共形分划的矩阵之乘法与通常的矩阵之乘法极为相似．两个大小相同的分划矩阵 A，$B\in M_{m,n}(\mathbf{F})$之和有类似的令人愉悦的表达式，如果它们的行的分划(它们列的分划)是相同的： 17

$$(A+B)[\alpha_i,\beta_j]=A[\alpha_i,\beta_j]+B[\alpha_i,\beta_j]$$

如果对一个矩阵的行和列作顺序分划，所产生的分划矩阵就称为一个**分块矩阵**(block matrix)．例如，如果 $A\in M_n(\mathbf{F})$的行和列用同样的顺序分划 $\alpha_1=\{1,\cdots,k\}$，$\alpha_2=\{k+1,\cdots,n\}$加以划分，则所产生的分块矩阵是

$$A=\begin{bmatrix}A[\alpha_1,\alpha_1] & A[\alpha_1,\alpha_2]\\ A[\alpha_2,\alpha_1] & A[\alpha_2,\alpha_2]\end{bmatrix}=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}$$

其中的**分块**(block)是 $A_{ij}=A[\alpha_i,\alpha_j]$．在本书中始终使用分块矩阵进行计算，$2\times2$ 分块矩阵最为重要也最为有用.

0.7.3 分划矩阵的逆

在经过分划的非奇异矩阵 A 的逆矩阵中，知道对应的分块是有用的，也就是将分划矩

阵的逆表示成共形分划的形式．这可以用多种表面上看似不同但却等价的方式来实现——假设 $A\in M_n(\mathbf{F})$ 与 A^{-1} 的某些子矩阵也是非奇异的．为简略起见，设 A 被分划成 2×2 分块矩阵

$$A=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}$$

其中 $A_{ii}\in M_{n_i}(\mathbf{F})(i=1,\ 2)$，$n_1+n_2=n$．有关 A^{-1} 的对应分划表示的一个有用的表达式是

$$\begin{bmatrix}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & A_{11}^{-1}A_{12}(A_{21}A_{11}^{-1}A_{12}-A_{22})^{-1}\\ A_{22}^{-1}A_{21}(A_{12}A_{22}^{-1}A_{21}-A_{11})^{-1} & (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}\end{bmatrix}\tag{0.7.3.1}$$

假设其中涉及的所有逆矩阵都存在．有关 A^{-1} 的这个表达式可以用 A 作分划矩阵的乘法然后予以化简加以验证．按照一般指标集的记号，可以记

$$A^{-1}[\alpha]=(A[\alpha]-A[\alpha,\alpha^c]A[\alpha^c]^{-1}A[\alpha^c,\alpha])^{-1}$$

以及

$$\begin{aligned}A^{-1}[\alpha,\alpha^c]&=A[\alpha]^{-1}A[\alpha,\alpha^c](A[\alpha^c,\alpha]A[\alpha]^{-1}A[\alpha,\alpha^c]-A[\alpha^c])^{-1}\\&=(A[\alpha,\alpha^c]A[\alpha^c]^{-1}A[\alpha^c,\alpha]-A[\alpha])^{-1}A[\alpha,\alpha^c]A[\alpha^c]^{-1}\end{aligned}$$

这里再次假设有关的逆矩阵都存在．在这些表达式与 Schur 补之间有一个密切的关系，见(0.8.5)．注意 $A^{-1}[\alpha]$ 是 A^{-1} 的一个子矩阵，而 $A[\alpha]^{-1}$ 则是 A 的一个子矩阵的逆，一般说来，这两者是不相同的．

0.7.4 Sherman-Morrison-Woodbury 公式

假设非奇异的矩阵 $A\in M_n(\mathbf{F})$ 有一个已知的逆 A^{-1}，同时考虑 $B=A+XRY$，其中 X 是 $n\times r$ 矩阵，Y 是 $r\times n$ 矩阵，而 R 是 $r\times r$ 非奇异矩阵．如果 B 和 $R^{-1}+YA^{-1}X$ 是非奇异的，那么

$$B^{-1}=A^{-1}-A^{-1}X(R^{-1}+YA^{-1}X)^{-1}YA^{-1}\tag{0.7.4.1}$$

如果 r 比 n 小得多，那么对 R 和 $R^{-1}+YA^{-1}X$ 求逆比对 B 求逆要容易得多．例如，如果 x，$y\in\mathbf{F}^n$ 是非零向量，$X=x$，$Y=y^{\mathrm{T}}$，$y^{\mathrm{T}}A^{-1}x\neq -1$，且 $R=[1]$，那么(0.7.4.1)就变成一个对 A 作秩 1 修正的逆的公式：

$$(A+xy^{\mathrm{T}})^{-1}=A^{-1}-(1+y^{\mathrm{T}}A^{-1}x)^{-1}A^{-1}xy^{\mathrm{T}}A^{-1}\tag{0.7.4.2}$$

特别地，如果对 x，$y\in\mathbf{F}^n$ 以及 $y^{\mathrm{T}}x\neq -1$ 有 $B=I+xy^{\mathrm{T}}$，那么 $B^{-1}=I-(1+y^{\mathrm{T}}x)^{-1}xy^{\mathrm{T}}$．

0.7.5 零化度互补

假设 $A\in M_n(\mathbf{F})$ 非奇异，设 α 和 β 是 $\{1,\ \cdots,\ n\}$ 的非空子集，并用 $|\alpha|=r$ 和 $|\beta|=s$ 记 α 和 β 的基数．则**零化度互补**(complementary nullity)**法则**说的就是

$$\operatorname{nullity}(A[\alpha,\beta])=\operatorname{nullity}(A^{-1}[\beta^c,\alpha^c])\tag{0.7.5.1}$$

它等价于秩恒等式

$$\operatorname{rank}(A[\alpha,\beta])=\operatorname{rank}(A^{-1}[\beta^c,\alpha^c])+r+s-n\tag{0.7.5.2}$$

由于我们可以对行和列作排列，首先放置由 α 指定的 r 行以及由 β 指定的 s 列，所以只需考虑表达式

$$A=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix},\quad A^{-1}=\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}$$

即可，其中 A_{11} 和 B_{11}^{T} 是 $r\times s$ 矩阵，而 A_{22} 和 B_{22}^{T} 是 $(n-r)\times(n-s)$ 矩阵. 这样(0.7.5.1)说的就是 $\mathrm{nullity}A_{11}=\mathrm{nullity}B_{22}$.

这里的基本原则十分简单. 假设 A_{11} 的**零化度**(nullity)为 k. 如果 $k\geqslant 1$，设 $X\in M_{s,k}(\mathbf{F})$ 的列是 A_{11} 的零空间的一组基. 由于 A 非奇异，故而

$$A\begin{bmatrix}X\\ 0\end{bmatrix}=\begin{bmatrix}A_{11}X\\ A_{21}X\end{bmatrix}=\begin{bmatrix}0\\ A_{21}X\end{bmatrix}$$

是满秩的，所以 $A_{21}X$ 有 k 个线性无关的列. 但是

$$\begin{bmatrix}B_{12}(A_{21}X)\\ B_{22}(A_{21}X)\end{bmatrix}=A^{-1}\begin{bmatrix}0\\ A_{21}X\end{bmatrix}=A^{-1}A\begin{bmatrix}X\\ 0\end{bmatrix}=\begin{bmatrix}X\\ 0\end{bmatrix}$$

所以 $B_{22}(A_{21}X)=0$，从而 $\mathrm{nullity}B_{22}\geqslant k=\mathrm{nullity}A_{11}$，此命题当 $k=0$ 时平凡地成立. 从 B_{22} 出发用类似的推理可得 $\mathrm{nullity}A_{11}\geqslant\mathrm{nullity}B_{22}$. 不同的讨论方法参见(3.5.P13).

当然，(0.7.5.1)也告诉我们有 $\mathrm{nullity}A_{12}=\mathrm{nullity}B_{12}$，$\mathrm{nullity}A_{21}=\mathrm{nullity}B_{21}$，以及 $\mathrm{nullity}A_{22}=\mathrm{nullity}B_{11}$. 如果 $r+s=n$，那么 $\mathrm{rank}A_{11}=\mathrm{rank}B_{22}$ 且 $\mathrm{rank}A_{22}=\mathrm{rank}B_{11}$，而如果 $n=2r=2s$，则也有 $\mathrm{rank}A_{12}=\mathrm{rank}B_{12}$ 以及 $\mathrm{rank}A_{21}=\mathrm{rank}B_{21}$. 最后，(0.7.5.2)告诉我们，一个 $n\times n$ 非奇异矩阵的 $r\times s$ 子矩阵的秩至少为 $r+s-n$.

0.7.6 分划矩阵以及主秩矩阵中的秩

将 $A\in M_n(\mathbf{F})$ 分划为

$$A=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix},\quad A_{11}\in M_r(\mathbf{F}),\quad A_{22}\in M_{n-r}(\mathbf{F})$$

如果 A_{11} 非奇异，则当然有 $\mathrm{rank}[A_{11}\quad A_{12}]=r$ 以及 $\mathrm{rank}\begin{bmatrix}A_{11}\\ A_{21}\end{bmatrix}=r$. 值得注意的是其逆命题亦为真：

$$\text{如果 } \mathrm{rank}A=\mathrm{rank}[A_{11}\quad A_{12}]=\mathrm{rank}\begin{bmatrix}A_{11}\\ A_{21}\end{bmatrix},\quad \text{那么 } A_{11} \text{ 非奇异} \tag{0.7.6.1}$$

这可以由(0.4.6(c))推出：如果 A_{11} 是奇异的，那么 $\mathrm{rank}A_{11}=k<r$，且存在非奇异的 $S,T\in M_r(\mathbf{F})$，使得

$$SA_{11}T=\begin{bmatrix}I_k & 0\\ 0 & 0_{r-k}\end{bmatrix}$$

于是

$$\hat{A}=\begin{bmatrix}S & 0\\ 0 & I_{n-r}\end{bmatrix}A\begin{bmatrix}T & 0\\ 0 & I_{n-r}\end{bmatrix}=\begin{bmatrix}\begin{bmatrix}I_k & 0\\ 0 & 0_{r-k}\end{bmatrix} & SA_{12}\\ A_{21}T & A_{22}\end{bmatrix}$$

的秩是 r，正如它的第一个分块行和分块列所给出的那样. 因为 $\hat{A}$ 的第一个分块列的第 r 行是零，故 SA_{12} 中必定存在一列，它的第 r 个元素不为零，这就意味着 $\hat{A}$ 至少有 $r+1$ 列是线

性无关的．这与 rank $\hat{A}$＝rank$A$＝r 矛盾，所以 A_{11} 必定是非奇异的．

设 $A\in M_{m,n}(\mathbf{F})$，又假设 rank$A$＝r＞0．令 $A=XY^{\mathrm{T}}$ 是满秩分解，其中 X，$Y\in M_{m,r}(\mathbf{F})$，见(0.4.6c)．设 α，$\beta\subseteq\{1,\cdots,m\}$ 以及 γ，$\delta\subseteq\{1,\cdots,n\}$ 是基数为 r 的指标集．那么 $A[\alpha,\gamma]=X[\alpha,\varnothing^c]Y[\gamma,\varnothing^c]^{\mathrm{T}}\in M_r(\mathbf{F})$，它是非奇异的，只要 rank$X[\alpha,\varnothing^c]$＝rank$Y[\gamma,\varnothing^c]=r$．积性性质(0.3.5)确保有

$$\det A[\alpha,\gamma]\det A[\beta,\delta]=\det A[\alpha,\delta]\det A[\beta,\gamma] \tag{0.7.6.2}$$

假设 $A\in M_n(\mathbf{F})$ 且 rank$A=r$．我们说 A 是**主秩的**(rank principal)，如果它有一个非奇异的 $r\times r$ 主子矩阵．由(0.7.6.1)推出，如果存在某个指标集 $\alpha\subset\{1,\cdots,n\}$，使得

$$\operatorname{rank}A=\operatorname{rank}A[\alpha,\varnothing^c]=\operatorname{rank}A[\varnothing^c,\alpha] \tag{0.7.6.3}$$

(这就是说，如果 A 有 r 个线性无关的行，使得对应的 r 个列是线性无关的)，那么 A 就是主秩的．此外，$A[\alpha]$ 还是非奇异的．

如果 $A\in M_n(\mathbf{F})$ 是对称的或者斜对称的，或者如果 $A\in M_n(\mathbf{C})$ 是 Hermite 的或者斜 Hermite 的，那么对每个指标集 α 有 rank$A[\alpha,\varnothing^c]$＝rank$A[\varnothing^c,\alpha]$，所以 A 满足

20 (0.7.6.3)，从而它也是主秩的．

0.7.7 交换性，反交换性以及分块对角矩阵

两个矩阵 A，$B\in M_n(\mathbf{F})$ 称为是**交换的**(commute)，如果 $AB=BA$．交换性不是常规性质，但却是会经常遇到的重要情形．假设 $\Lambda=[\Lambda_{ij}]_{i,j=1}^s\in M_n(\mathbf{F})$ 是一个分块矩阵，其中 $\Lambda_{ij}=0$(如果 $i\neq j$)；$\Lambda_{ii}=\lambda_i I_{n_i}$(对某个 $\lambda_i\in\mathbf{F}$ 以及每个 $i=1,\cdots,s$)；又对 $i\neq j$ 有 $\lambda_i\neq\lambda_j$．与 Λ 共形地分划 $B=[B_{ij}]_{i,j=1}^s\in M_n(\mathbf{F})$．那么 $\Lambda B=B\Lambda$，当且仅当对每个 i，$j=1,\cdots,s$ 有 $\lambda_i B_{ij}=B_{ij}\lambda_j$，即对每个 i，$j=1,\cdots,s$ 有 $(\lambda_i-\lambda_j)B_{ij}=0$．这些恒等式能够满足，当且仅当只要 $i\neq j$ 就有 $B_{ij}=0$．于是，Λ 与 B 可交换，当且仅当 B 与 Λ 是共形的分块对角矩阵，见(0.9.2)．

两个矩阵 A，$B\in M_n(\mathbf{F})$ 称为是**反交换的**(anticommute)，如果 $AB=-BA$．例如，矩阵 $\begin{bmatrix}1&0\\0&-1\end{bmatrix}$ 和 $\begin{bmatrix}0&1\\1&0\end{bmatrix}$ 是反交换的．

0.7.8 映射 vec

将矩阵 $A\in M_{m,n}(\mathbf{F})$ 按照列进行分划：$A=[a_1\ \cdots\ a_n]$．映射 vec：$M_{m,n}(\mathbf{F})\to\mathbf{F}^{mn}$ 是

$$\operatorname{vec}A=[a_1^{\mathrm{T}}\ \cdots\ a_n^{\mathrm{T}}]^{\mathrm{T}}$$

也就是说，vec$A$ 是从左向右叠置 A 的列所得到的向量．在与矩阵方程有关的问题中，vec 算子可能是一个便利的工具．

0.8 再谈行列式

关于行列式的某些进一步的结果以及恒等式是有用的参考资料．

0.8.1 复合矩阵

令 $A\in M_{m,n}(\mathbf{F})$．设 $\alpha\subseteq\{1,\cdots,m\}$ 和 $\beta\subseteq\{1,\cdots,n\}$ 是基数为 $r\leqslant\min\{m,n\}$ 个元素

的指标集．其位于 α，β 处的元素是 $\det A[\alpha,\beta]$的那个$\binom{m}{r}\times\binom{n}{r}$矩阵称为 A 的第 r 个**复合矩阵**(compound matrix)，并记之为 $C_r(A)$．在构造 $C_r(A)$的行和列时，我们按照字典顺序对指标集排序，也就是说，{1，2，4}先于{1，2，5}先于{1，3，4}，如此等等．例如，如果

$$A=\begin{bmatrix}1&2&3\\4&5&6\\7&8&10\end{bmatrix}\qquad(0.8.1.0)$$

那么 $C_2(A)=$

$$\begin{bmatrix}\det\begin{bmatrix}1&2\\4&5\end{bmatrix}&\det\begin{bmatrix}1&3\\4&6\end{bmatrix}&\det\begin{bmatrix}2&3\\5&6\end{bmatrix}\\\det\begin{bmatrix}1&2\\7&8\end{bmatrix}&\det\begin{bmatrix}1&3\\7&10\end{bmatrix}&\det\begin{bmatrix}2&3\\8&10\end{bmatrix}\\\det\begin{bmatrix}4&5\\7&8\end{bmatrix}&\det\begin{bmatrix}4&6\\7&10\end{bmatrix}&\det\begin{bmatrix}5&6\\8&10\end{bmatrix}\end{bmatrix}=\begin{bmatrix}-3&-6&-3\\-6&-11&-4\\-3&-2&2\end{bmatrix}$$

如果 $A\in M_{m,k}(\mathbf{F})$，$B\in M_{k,n}(\mathbf{F})$，且 $r\leqslant\{m,k,n\}$，由 Cauchy-Binet 公式(0.8.7)得出

$$C_r(AB)=C_r(A)C_r(B)\qquad(0.8.1.1)$$

这就是第 r 个复合矩阵的**积性性质**(multiplicativity property)． 21

定义 $C_0(A)=1$．我们有 $C_1(A)=A$；如果 $A\in M_n(\mathbf{F})$，则有 $C_n(A)=\det A$．

如果 $A\in M_{m,k}(\mathbf{F})$且 $t\in\mathbf{F}$，那么 $C_r(tA)=t^rC_r(A)$

如果 $1\leqslant r\leqslant n$，那么 $C_r(I_n)=I_{\binom{n}{r}}\in M_{\binom{n}{r}}$

如果 $A\in M_n$ 非奇异且 $1\leqslant r\leqslant n$，那么 $C_r(A)^{-1}=C_r(A^{-1})$

如果 $A\in M_n$ 且 $1\leqslant r\leqslant n$，那么 $\det C_r(A)=(\det A)^{\binom{n-1}{r-1}}$

如果 $A\in M_{m,n}(\mathbf{F})$且 $r=\operatorname{rank}A$，那么 $\operatorname{rank}C_r(A)=1$

如果 $A\in M_{m,n}(\mathbf{F})$且 $1\leqslant r\leqslant\min\{m,n\}$，那么 $C_r(A^{\mathrm{T}})=C_r(A)^{\mathrm{T}}$

如果 $A\in M_{m,n}(\mathbf{C})$且 $1\leqslant r\leqslant\min\{m,n\}$，那么 $C_r(A^*)=C_r(A)^*$．

如果 $\Delta=[d_{ij}]\in M_n(\mathbf{F})$是上(下)三角矩阵(见(0.9.3))，那么 $C_r(\Delta)$是上(下)三角矩阵．它的主对角线上的元素是从 d_{11}，…，d_{nn} 中选取 r 个元素组成的$\binom{n}{r}$个可能的乘积，也就是说，它们是$\binom{n}{r}$个按照字典顺序排列的纯量 $d_{i_1i_1}$，…，$d_{i_ri_r}$，其中 $1\leqslant i_1<\cdots<i_r\leqslant n$．如果 $D=\operatorname{diag}(d_1,\cdots,d_n)\in M_n(\mathbf{F})$是对角矩阵，则 $C_r(D)$亦然，它的主对角线元素是从 d_1，…，d_n 中选取 r 个元素组成的$\binom{n}{r}$个可能的乘积，也就是说，它们是$\binom{n}{r}$个按照字典顺序排列的纯量 d_{i_1}，…，d_{i_r}，其中 $1\leqslant i_1<\cdots<i_r\leqslant n$．关于复合矩阵的详尽的讨论请参见 Fiedler(1986)专著的第 6 章．

0.8.2 转置伴随矩阵与逆矩阵

如果 $A\in M_n(\mathbf{F})$ 且 $n\geqslant 2$，那么 A 的诸代数余子式构成的转置矩阵

$$\operatorname{adj}A=[(-1)^{i+j}\det A[\{j\}^c,\{i\}^c]] \tag{0.8.2.0}$$

称为 A 的**转置伴随**(adjugate)**矩阵**，也称为 A 的**经典伴随**(classical adjoint)**矩阵**．例如，$\operatorname{adj}\begin{bmatrix}a & b\\ c & d\end{bmatrix}=\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$.

利用关于行列式的 Laplace 展开式进行计算揭示出转置伴随矩阵的基本性质：

$$(\operatorname{adj}A)A=A(\operatorname{adj}A)=(\det A)I \tag{0.8.2.1}$$

由此可见，如果 A 是非奇异的，则 $\operatorname{adj}A$ 是非奇异的，而且有 $\det(\operatorname{adj}A)=(\det A)^{n-1}$.

如果 A 非奇异，那么

$$\operatorname{adj}A=(\det A)A^{-1},\quad 即\quad A^{-1}=(\det A)^{-1}\operatorname{adj}A \tag{0.8.2.2}$$

例如，$\begin{bmatrix}a & b\\ c & d\end{bmatrix}^{-1}=(ad-bc)^{-1}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$(如果 $ad\neq bc$)．特别地，$\operatorname{adj}(A^{-1})=A/\det A=(\operatorname{adj}A)^{-1}$.

如果 A 是奇异的，且 $\operatorname{rank}A\leqslant n-2$，那么 A 的每个 $n-1$ 阶子式都是零，所以 $\operatorname{adj}A=0$.

如果 A 是奇异的，且 $\operatorname{rank}A=n-1$，那么 A 的某个 $n-1$ 阶子式不等于零，所以 $\operatorname{adj}A\neq 0$ 且 $\operatorname{rank}\operatorname{adj}A\geqslant 1$．此外，$A$ 的某 $n-1$ 列是线性无关的，所以，恒等式 $(\operatorname{adj}A)A=(\det A)I=0$ 确保 $\operatorname{adj}A$ 的零空间的维数至少为 $n-1$，从而 $\operatorname{rank}\operatorname{adj}A\leqslant 1$．我们断言 $\operatorname{rank}\operatorname{adj}A=1$．满秩分解(0.4.6(e))确保对某个非零的 $\alpha\in\mathbf{F}$ 以及非零的 x，$y\in\mathbf{F}^n$ 有 $\operatorname{adj}A=\alpha xy^{\mathrm{T}}$，它们可以如下来确定：计算

22

$$(Ax)y^{\mathrm{T}}=A(\operatorname{adj}A)=0=(\operatorname{adj}A)A=x(y^{\mathrm{T}}A)$$

并得出结论 $Ax=0$ 以及 $y^{\mathrm{T}}A=0$，这就是说，x(或 y)作为 A(或 A^{T})的一维零空间的非零元素被确定到相差一个非零的纯量倍数.

函数 $A\to\operatorname{adj}A$ 在 M_n 上连续($\operatorname{adj}A$ 的每个元素都是关于 A 的元素的多项式)，且 M_n 中的每个矩阵都是非奇异矩阵的极限，故而转置伴随矩阵的性质可以由连续性以及反函数的性质推导出来．例如，如果 A，$B\in M_n$ 非奇异，那么 $\operatorname{adj}(AB)=(\det AB)(AB)^{-1}=(\det A)(\det B)B^{-1}A^{-1}=(\det B)B^{-1}(\det A)A^{-1}=(\operatorname{adj}B)(\operatorname{adj}A)$．这样连续性就确保有

$$\operatorname{adj}(AB)=(\operatorname{adj}B)(\operatorname{adj}A),\quad 对所有 A,B\in M_n \tag{0.8.2.3}$$

对任何 $c\in\mathbf{F}$ 以及任何 $A\in M_n(\mathbf{F})$，有 $\operatorname{adj}(cA)=c^{n-1}\operatorname{adj}A$．特别地，有 $\operatorname{adj}(cI)=c^{n-1}I$ 以及 $\operatorname{adj}0=0$.

如果 A 非奇异，那么

$$\operatorname{adj}(\operatorname{adj}A)=\operatorname{adj}((\det A)A^{-1})=(\det A)^{n-1}\operatorname{adj}A^{-1}=(\det A)^{n-1}(A/\det A)=(\det A)^{n-2}A$$

故而连续性就确保有

$$\operatorname{adj}(\operatorname{adj}A)=(\det A)^{n-2}A,\quad 对所有 A\in M_n \tag{0.8.2.4}$$

如果 $A+B$ 非奇异，那么 $A(A+B)^{-1}B=B(A+B)^{-1}A$，故而连续性就确保有

$$A\operatorname{adj}(A+B)B=B\operatorname{adj}(A+B)A,\quad 对所有 A,B\in M_n \tag{0.8.2.5}$$

令 A，$B\in M_n$ 并假设 A 与 B 可交换．如果 A 非奇异，那么 $BA^{-1}=A^{-1}ABA^{-1}=$

$A^{-1}BAA^{-1}=A^{-1}B$，所以 A^{-1} 与 B 可交换. 但是 $BA^{-1}=(\det A)^{-1}B\mathrm{adj}A$ 且 $A^{-1}B=(\det A)^{-1}(\mathrm{adj}A)B$，所以 $\mathrm{adj}A$ 与 B 可交换. 连续性确保，只要 A 与 B 可交换，那么 $\mathrm{adj}A$ 也就与 B 可交换，即便 A 是奇异的，此结论依然成立.

如果 $A=[a_{ij}]$是上三角的，那么 $\mathrm{adj}A=[b_{ij}]$也是上三角的，且每一个 $b_{ii}=\prod_{j\neq i}a_{jj}$；如果 A 是对角的，则 $\mathrm{adj}A$ 亦然.

转置伴随矩阵是 $\det A$ 的梯度的转置：

$$(\mathrm{adj}A)=\left[\frac{\partial}{\partial a_{ij}}\det A\right]^{\mathrm{T}} \tag{0.8.2.6}$$

如果 A 非奇异，由(0.8.2.6)即推出有

$$\left[\frac{\partial}{\partial a_{ij}}\det A\right]^{\mathrm{T}}=(\det A)A^{-1} \tag{0.8.2.7}$$

如果 $A\in M_n$ 非奇异，那么 $\mathrm{adj}A^{\mathrm{T}}=(\det A^{\mathrm{T}})A^{-\mathrm{T}}=(\det A)A^{-\mathrm{T}}=[(\det A)A^{-1}]^{\mathrm{T}}=(\mathrm{adj}A)^{\mathrm{T}}$. 连续性保证有

$$\mathrm{adj}A^{\mathrm{T}}=(\mathrm{adj}A)^{\mathrm{T}},\quad \text{对所有 } A\in M_n(\mathbf{F}) \tag{0.8.2.8}$$

类似的推理给出

$$\mathrm{adj}A^{*}=(\mathrm{adj}A)^{*},\quad \text{对所有 } A\in M_n \tag{0.8.2.9}$$ 23

设 $A=[a_1\cdots a_n]\in M_n(\mathbf{F})$按照它的列加以分划，并设 $b\in\mathbf{F}^n$. 定义

$$(A\underset{i}{\leftarrow}b)=[a_1\ \cdots\ a_{i-1}\ \ b\ \ a_{i+1}\ \cdots\ a_n]$$

这就是说，$(A\underset{i}{\leftarrow}b)$表示这样一个矩阵，它的第 i 列为 b 而其余的列与 A 的列相同. 对 $\det(A\underset{i}{\leftarrow}b)$的第 i 列的子式用 Laplace 展开式(0.3.1.1)检验，表明它是向量$(\mathrm{adj}A)b$ 的第 i 个元素，也就是说

$$[\det(A\underset{i}{\leftarrow}b)]_{i=1}^{n}=(\mathrm{adj}A)b \tag{0.8.2.10}$$

将这个向量恒等式应用到 $C=[c_1\ \cdots\ c_n]\in M_n(\mathbf{F})$的每一列上就给出矩阵恒等式

$$[\det(A\underset{i}{\leftarrow}c_j)]_{i,j=1}^{n}=(\mathrm{adj}A)C \tag{0.8.2.11}$$

0.8.3 Cramer 法则

当 $A\in M_n(\mathbf{F})$非奇异时，Cramer 法则是对 $Ax=b$ 的解的特定元素给出解析表达式的有用的方法. 恒等式

$$A[\det(A\underset{i}{\leftarrow}b)]_{i=1}^{n}=A(\mathrm{adj}A)b=(\det A)b$$

由(0.8.2.9)推出. 如果 $\det A\neq 0$，我们就对解向量 x 的第 i 个元素 x_i 得到 Cramer 法则

$$x_i=\frac{\det(A\underset{i}{\leftarrow}b)}{\det A}$$

Cramer 法则也可以由行列式的积性直接得出. 方程组 $Ax=b$ 可以改写成

$$A(I\underset{i}{\leftarrow}x)=A\underset{i}{\leftarrow}b$$

在两边取行列式(利用积性)就给出

$$(\det A)\det(I \underset{i}{\leftarrow} x) = \det(A \underset{i}{\leftarrow} b)$$

但是 $\det(I \underset{i}{\leftarrow} x)=x_i$，这就得到了公式.

0.8.4 逆矩阵的子式

Jacobi 恒等式推广了非奇异矩阵的逆的转置伴随公式，且把 A^{-1} 的子式与 $A\in M_n(\mathbf{F})$ 的子式联系在一起：

$$\det A^{-1}[\alpha^c,\beta^c] = (-1)^{p(\alpha,\beta)}\frac{\det A[\beta,\alpha]}{\det A} \tag{0.8.4.1}$$

其中 $p(\alpha,\beta)=\sum_{i\in\alpha}i+\sum_{j\in\beta}j$. 我们一般约定 $\det A[\varnothing]=1$. 对主子矩阵，Jacobi 恒等式有简单的形式

$$\det A^{-1}[\alpha^c] = \frac{\det A[\alpha]}{\det A} \tag{0.8.4.2}$$

0.8.5 Schur 补和行列式公式

设给定 $A=[a_{ij}]\in M_n(\mathbf{F})$，并假设 $\alpha\subseteq\{1, \cdots, n\}$是使得 $A[\alpha]$非奇异的一个指标集. 基于利用 α 和 α^c 对 A 作的 2 分划，就有一个关于 $\det A$ 的重要公式

24

$$\det A = \det A[\alpha]\det(A[\alpha^c]-A[\alpha^c,\alpha]A[\alpha]^{-1}A[\alpha,\alpha^c]) \tag{0.8.5.1}$$

它推广了关于 2×2 矩阵的行列式的一个熟悉的公式. 特殊的矩阵

$$A/A[\alpha] = A[\alpha^c]-A[\alpha^c,\alpha]A[\alpha]^{-1}A[\alpha,\alpha^c] \tag{0.8.5.2}$$

称为 $A[\alpha]$在 A 中的 **Schur 补**，它也出现在(0.7.3.1)中的逆矩阵的分划形式中. 在方便的时候，取 $\alpha=\{1, \cdots, k\}$，并将 A 写成 2×2 分块矩阵 $A=[A_{ij}]$，其中 $A_{11}=A[\alpha]$，$A_{22}=A[\alpha^c]$，$A_{12}=A[\alpha, \alpha^c]$，以及 $A_{21}=A[\alpha^c, \alpha]$. 公式(0.8.5.1)可以通过计算恒等式

$$\begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \tag{0.8.5.3}$$

两边的行列式加以验证，此恒等式含有大量有关 Schur 补 $S=[s_{ij}]=A/A_{11}=A_{22}-A_{21}A_{11}^{-1}A_{12}$ 的信息.

(a) 如果将 A 的前 k 行(列)的线性组合用这样一种方式添加到后 $n-k$ 行(列)，使得在其左下角(右上角)产生出一个零分块，那么 Schur 补 S(仅仅)出现在右下角；这就是 **Gauss 分块消元法**(block Gaussian elimination)，这是(仅有的)可能，因为 A_{11}是非奇异的. A 的任何包含 A_{11}作为其主子矩阵的子矩阵在消元法之前及之后都有相同的行列式，这一方法产生出(0.8.5.3)中的分块对角形. 于是，对任何指标集 $\beta=\{i_1, \cdots, i_m\}\subseteq\{1, \cdots, n-k\}$，如果我们构作平移指标集$\tilde{\beta}=\{i_1+k, \cdots, i_m+k\}$，那么 $\det A[\alpha\cup\tilde{\beta}, \alpha\cup\tilde{\gamma}]$(之前的行列式)$=\det(A_{11}\oplus S[\beta, \gamma])$(之后的行列式)，所以

$$\det S[\beta,\gamma] = \det A[\alpha\cup\tilde{\beta},\alpha\cup\tilde{\gamma}]/\det A[\alpha] \tag{0.8.5.4}$$

例如，如果 $\beta=\{i\}$且 $\gamma=\{j\}$，那么对 $\alpha=\{1, \cdots, k\}$我们有

$$\det S[\beta,\gamma] = s_{ij} = \det A[\{1,\cdots,k,k+i\},\{1,\cdots,k,k+j\}]/\det A_{11} \tag{0.8.5.5}$$

所以 S 的所有元素都是 A 的子式的比值.

(b) $\operatorname{rank}A=\operatorname{rank}A_{11}+\operatorname{rank}S\geqslant\operatorname{rank}A_{11}$，且 $\operatorname{rank}A=\operatorname{rank}A_{11}$，当且仅当 $A_{22}=A_{21}A_{11}^{-1}A_{12}$.

(c) A 是非奇异的，当且仅当 S 是非奇异的，因为 $\det A=\det A_{11}\det S$. 如果 A 是非奇异的，那么 $\det S=\det A/\det A_{11}$.

假设 A 是非奇异的. 那么在(0.8.5.3)的两边取逆就对逆矩阵给出一种不同于(0.7.3.1)的表示法：

$$A^{-1}=\begin{bmatrix}A_{11}+A_{11}^{-1}A_{12}S^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}S^{-1}\\ -S^{-1}A_{21}A_{11}^{-1} & S^{-1}\end{bmatrix}\tag{0.8.5.6}$$

其中它还告诉我们 $A^{-1}[\{k+1,\cdots,n\}]=S^{-1}$，所以

$$\det A^{-1}[\{k+1,\cdots,n\}]=\det A_{11}/\det A\tag{0.8.5.7}$$ 25

这是 Jacobi 恒等式(0.8.4.1)的一种形式. 另一种形式可以通过用转置伴随矩阵表示逆矩阵得到，这就给出

$$\det((\operatorname{adj}A))[\{k+1,\cdots,n\}]=(\det A)^{n-k-1}\det A_{11}\tag{0.8.5.8}$$

当 α^c 由单独一个元素组成时，$A[\alpha]$在 A 中的 Schur 补是一个纯量，且(0.8.5.1)转化为恒等式

$$\det A=A[\alpha^c]\det A[\alpha]-A[\alpha^c,\alpha](\operatorname{adj}A[\alpha])A[\alpha,\alpha^c]\tag{0.8.5.9}$$

此式甚至在 $A[\alpha]$为奇异时也成立. 例如，如果 $\alpha=\{1,\cdots,n-1\}$，那么 $\alpha^c=\{n\}$且 A 被表示成**加边矩阵**(bordered matrix)的形式

$$A=\begin{bmatrix}\widetilde{A} & x\\ y^{\mathrm{T}} & a\end{bmatrix}$$

其中 $a\in\mathbf{F}$，x，$y\in\mathbf{F}^{n-1}$，而$\widetilde{A}\in M_{n-1}(\mathbf{F})$；(0.8.5.9)就是加边矩阵的行列式的 **Cauchy 展开式**

$$\det\begin{bmatrix}\widetilde{A} & x\\ y^{\mathrm{T}} & a\end{bmatrix}=a\det\widetilde{A}-y^{\mathrm{T}}(\operatorname{adj}\widetilde{A})x\tag{0.8.5.10}$$

Cauchy 展开式(0.8.5.10)与 A 的带符号的 $n-2$ 阶子式(adj $\widetilde{A}$的元素)以及一行与一列元素的一个双线性型有关，Laplace 展开式(0.3.1.1)与 A 的带符号的 $n-1$ 阶子式以及一行或者一列元素的一个线性型有关. 如果 $a\neq0$，我们可以利用$[a]$在 A 中的 Schur 补表示

$$\det\begin{bmatrix}\widetilde{A} & x\\ y^{\mathrm{T}} & a\end{bmatrix}=a\det(\widetilde{A}-a^{-1}xy^{\mathrm{T}})$$

将这个恒等式的右边与(0.8.5.10)的右边等同起来，并置 $a=-1$ 就给出**秩 1 扰动的行列式的 Cauchy 公式**

$$\det(\widetilde{A}+xy^{\mathrm{T}})=\det\widetilde{A}+y^{\mathrm{T}}(\operatorname{adj}\widetilde{A})x\tag{0.8.5.11}$$

(a)中所讨论的 Schur 补的唯一性可以用来推导出一个与 Schur 补(在一个 Schur 补的范围内)有关的恒等式. 假设非奇异的 $k\times k$ 分块 A_{11}被分划成为 2×2 分块矩阵 $A_{11}=[\mathcal{A}_{ij}]$，它的左上部的 $l\times l$ 块$\mathcal{A}_{11}$是非奇异的. 记 $A_{21}=[\mathcal{A}_1\quad\mathcal{A}_2]$，其中$\mathcal{A}_1$ 是$(n-k)\times l$ 的，又记 $A_{12}^{\mathrm{T}}=[\mathcal{B}_1^{\mathrm{T}}\quad\mathcal{B}_2^{\mathrm{T}}]$，其中$\mathcal{B}_1$ 是 $l\times(n-k)$的. 这给出加细的分划

$$A=\begin{bmatrix}\mathcal{A}_{11} & \mathcal{A}_{12} & \mathcal{B}_1\\ \mathcal{A}_{21} & \mathcal{A}_{22} & \mathcal{B}_2\\ \mathcal{A}_1 & \mathcal{A}_2 & \mathcal{A}_{22}\end{bmatrix}$$

现在将 A 的前 l 行的线性组合添加到接下来的 $k-l$ 行，从而将$\mathcal{A}_{21}$化简为一个零分块．其结果是

$$A'=\begin{bmatrix}\mathcal{A}_{11} & \mathcal{A}_{12} & \mathcal{B}_1\\ 0 & A_{11}/\mathcal{A}_{11} & \mathcal{B}'_2\\ \mathcal{A}_1 & \mathcal{A}_2 & A_{22}\end{bmatrix}$$

在这里我们已经确认所得到的A'的 2×2 分块就是$\mathcal{A}_{11}$在 A_{11}中的(必定非奇异的)Schur 补．现在将A'的前 k 行的线性组合添加到最后 $n-k$ 行，从而将$[\mathcal{A}_1\quad \mathcal{A}_2]$化简为一个零分块．结果得到

$$A''=\begin{bmatrix}\mathcal{A}_{11} & \mathcal{A}_{12} & \mathcal{B}_1\\ 0 & A_{11}/\mathcal{A}_{11} & \mathcal{B}'_2\\ 0 & 0 & A/A_{11}\end{bmatrix}$$

在这里我们确认所得到的A''的 3×3 分块就是 A_{11}在 A 中的 Schur 补．A''的右下方的 2×2 分块必定是 $A/\mathcal{A}_{11}$，即$\mathcal{A}_{11}$在 A 中的 Schur 补．此外，$A/\mathcal{A}_{11}$的右下角分块必定是 $A_{11}/\mathcal{A}_{11}$在 $A/\mathcal{A}_{11}$中的 Schur 补．这个结果就是 **Schur 补的商性质**(quotient property)：

$$A/A_{11}=(A/\mathcal{A}_{11})/(A_{11}/\mathcal{A}_{11}) \tag{0.8.5.12}$$

如果(0.8.5.3)中的四个分块 A_{ij} 都是方阵且有同样的大小，又如果 A_{11}与 A_{21}可交换，那么

$$\det A=\det A_{11}\det S=\det(A_{11}S)=\det(A_{11}A_{22}-A_{11}A_{21}A_{11}^{-1}A_{12})=\det(A_{11}A_{22}-A_{21}A_{12})$$

如果 A_{11}与 A_{12}可交换，则同样的结论可以通过计算 $\det A=(\det S)(\det A)=\det(SA_{11})$得出．根据连续性，只要 A_{11}与 A_{21}可交换或者 A_{11}与 A_{12}可交换，恒等式

$$\det A=\det(A_{11}A_{22}-A_{21}A_{12}) \tag{0.8.5.13}$$

就成立，即使 A_{11}为奇异亦然．如果 A_{22}与 A_{12}可交换或者 A_{22}与 A_{21}可交换，则利用 A_{22}的 Schur 补所作的类似的讨论就证明了

$$\det A=\det(A_{11}A_{22}-A_{12}A_{21}) \tag{0.8.5.14}$$

如果 A_{22}与 A_{21}可交换或者 A_{22}与 A_{12}可交换．

0.8.6 Sylvester 行列式恒等式与 Kronecker 行列式恒等式

考虑(0.8.5.4)的两个推论．如果我们置

$$B=[b_{ij}]=[\det A[\{1,\cdots k,k+i\},\{1,\cdots,k,k+j\}]]_{i,j=1}^{n-k}$$

那么 B 的每个元素都是形如(0.8.5.10)的加边矩阵的行列式：$\tilde{A}$是 A_{11}，x 是 A_{12}的第 j 列，y^{T} 是 A_{21}的第 i 行，而 a 是 A_{22}的位于(i,j)处的元素．恒等式(0.8.5.5)告诉我们 $B=(\det A_{11})S$，所以

$$\det B=(\det A_{11})^{n-k}\det S=(\det A_{11})^{n-k}(\det A/\det A_{11})=(\det A_{11})^{n-k-1}\det A$$

关于 B 的这一结论就是**加边行列式的 Sylvester 恒等式**：

$$\det B=(\det A[\alpha])^{n-k-1}\det A \tag{0.8.6.1}$$

其中 $B=[\det A[\alpha\cup\{i\},\alpha\cup\{j\}]]$，而 i，j 则是不包含在 α 中的指标．

如果 $A_{22}=0$，那么 B 的每个元素都是形如(0.8.5.10)的一个加边矩阵的行列式(对于 $a=0$)．在此情形，Schur 补 $A/A_{11}=-A_{21}A_{11}^{-1}A_{12}$ 的秩至多为 k，故而 B 的每个 $(k+1)\times(k+1)$ 子矩阵的行列式都是零．关于 B 的这一结论就是**加边行列式的 Kronecker 定理**(Kronecker's theorem for bordered determinant)．

0.8.7 Cauchy-Binet 公式

由于与矩阵乘法公式外表上的相似性，这个有用的公式能够记得住．这并非偶然，因为它与复合矩阵的积性(0.8.1.1)等价．设 $A\in M_{m,k}(\mathbf{F})$，$B\in M_{k,n}(\mathbf{F})$，以及 $C=AB$．进一步设 $1\leqslant r\leqslant\min\{m, k, n\}$，又设 $\alpha\subseteq\{1,\cdots,m\}$ 与 $\beta\subseteq\{1,\cdots,n\}$ 是指标集，它们每一个的基数都是 r．C 的 α，β 子式的一个表达式是

$$\det C[\alpha,\beta]=\sum_{\gamma}\det A[\alpha,\gamma]\det B[\gamma,\beta]$$

其中的和式取遍基数为 r 的所有指标集 $\gamma\subseteq\{1,\cdots,k\}$．

0.8.8 子式之间的关系

设给定 $A\in M_{m,n}(\mathbf{F})$ 以及一个基数为 k 的固定的指标集 $\alpha\subseteq\{1,\cdots,m\}$．当 $\omega\subseteq\{1,\cdots,n\}$ 取遍基数为 k 的**有序**指标集时，诸子式 $\det A[\alpha,\omega]$ 并非是代数无关的，因为在子矩阵中间子式的个数要比不同的元素更多．在这些子式之间已知有二次关系．设 $i_1, i_2, \cdots, i_k\in\{1,\cdots,n\}$ 是 k 个不同的指标，不一定按自然顺序排列，又设 $A[\alpha; i_1,\cdots,i_k]$ 表示这样一个矩阵，它的行由 α 指出，而它的第 j 列是 $A[\alpha,\{1,\cdots,n\}]$ 的第 i_j 列．这个记号与我们先前的记号之间的区别在于，它的列可能如同在 $A(\{1,3\}; 4,2)$ 中那样不是按自然顺序出现，$A(\{1,3\}; 4,2)$ 的第 1 列是 A 的第 1 行第 4 列以及第 3 行第 4 列元素．这样一来，对每个 $s=1,\cdots,k$ 以及所有不同的指标序列 $i_1,\cdots,i_k\in\{1,\cdots,n\}$ 以及 $j_1,\cdots,j_k\in\{1,\cdots,n\}$，我们就有关系式

$$\begin{aligned}&\det A[\alpha;i_1,\cdots,i_k]\det A[\alpha;j_1,\cdots,j_k]\\&=\sum_{t=1}^{k}\det A[\alpha;i_1,\cdots,i_{s-1},j_t,i_{s+1},\cdots,i_k]\det A[\alpha;j_1,\cdots,j_{t-1},i_s,j_{t+1},\cdots,j_k]\end{aligned}$$

0.8.9 Laplace 展开定理

用给定的一行或者一列的子式计算的 Laplace 展开式(0.3.1.1)包含在有关行列式的一族自然的表达式中．设 $A\in M_n(\mathbf{F})$，又设给定 $k\in\{1,\cdots,n\}$，并令 $\beta\subseteq\{1,\cdots,n\}$ 是任意一个给定的基数为 k 的指标集．那么

$$\det A=\sum_{\alpha}(-1)^{p(\alpha,\beta)}\det A[\alpha,\beta]\det A[\alpha^c,\beta^c]=\sum_{\alpha}(-1)^{p(\alpha,\beta)}\det A[\beta,\alpha]\det A[\beta^c,\alpha^c]$$

其中的和式取遍基数为 k 的所有指标集 $\alpha\subseteq\{1,\cdots,n\}$，而 $p(\alpha,\beta)=\sum_{i\in\alpha}i+\sum_{j\in\beta}j$．选取 $k=1$ 以及 $\beta=\{i\}$ 或者 $\{j\}$ 就给出(0.3.1.1)中的展开式．

0.8.10 行列式的导数

设 $A(t)=[a_1(t)\cdots a_n(t)]=[a_{ij}(t)]$ 是一个 $n\times n$ 复矩阵，其元素是 t 的可微函数，并定

28 义 $A'(t)=[a'_{ij}(t)]$. 由行列式的多重线性(0.3.6(a))以及导数的定义可得

$$\frac{\mathrm{d}}{\mathrm{d}t}\det A(t)=\sum_{j=1}^{n}\det(A(t)\underset{j}{\leftarrow}a'_j(t))=\sum_{j=1}^{n}\sum_{i=1}^{n}((\mathrm{adj}A(t))^{\mathrm{T}})_{ij}a'_{ij}(t)$$
$$=\mathrm{tr}((\mathrm{adj}A(t))A'(t)) \qquad (0.8.10.1)$$

例如，如果 $A\in M_n$ 且 $A(t)=tI-A$，那么 $A'(t)=I$，且有

$$\frac{\mathrm{d}}{\mathrm{d}t}\det(tI-A)=\mathrm{tr}((\mathrm{adj}A(t))I)=\mathrm{tr}\ \mathrm{adj}(tI-A). \qquad (0.8.10.2)$$

0.8.11 Dodgson 恒等式

设 $A\in M_n(\mathbf{F})$. 定义 $a=\det A[\{n\}^c]$，$b=A[\{n\}^c,\{1\}^c]$，$c=A[\{1\}^c,\{n\}^c]$，$d=\det A[\{1\}^c]$以及 $e=\det A[\{1,n\}^c]$. 如果 $e\neq 0$，那么 $\det A=(ad-bc)/e$.

0.8.12 转置伴随与复合

设 $A, B\in M_n(\mathbf{F})$. 设 $\alpha\subseteq\{1,\cdots,m\}$以及 $\beta\subseteq\{1,\cdots,n\}$是基数为 $r\leqslant n$ 的指标集. 第 r 个转置伴随矩阵 $\mathrm{adj}_r(A)\in M_{\binom{n}{r}}(\mathbf{F})$的$(\alpha,\beta)$元素是

$$(-1)^{p(\alpha,\beta)}\det A[\beta^c,\alpha^c] \qquad (0.8.12.1)$$

其中 $p(\alpha,\beta)=\sum_{i\in\alpha}i+\sum_{j\in\beta}j$. $\mathrm{adj}_r(A)$的行和列是按照字典顺序排列指标集所得到的，正如对第 r 个复合矩阵那样. 例如，利用(0.8.1.0)中的矩阵 A，我们就有

$$\mathrm{adj}_2(A)=\begin{bmatrix}10&-6&3\\-8&5&-2\\7&-4&1\end{bmatrix}$$

第 r 个转置伴随矩阵的**积性性质**是

$$\mathrm{adj}_r(AB)=\mathrm{adj}_r(B)\mathrm{adj}_r(A) \qquad (0.8.12.2)$$

定义 $\mathrm{adj}_n(A)=1$. 我们有 $\mathrm{adj}_0(A)=\det A$ 以及 $\mathrm{adj}_1(A)=A$. 第 r 个转置伴随矩阵和第 r 个复合矩阵由恒等式

$$\mathrm{adj}_r(A)C_r(A)=C_r(A)\mathrm{adj}_r(A)=(\det A)I_{\binom{n}{r}}$$

联系在一起，(0.8.9)中的恒等式是它的特例. 特别地，有 $C_r(A)^{-1}=(\det A)^{-1}\mathrm{adj}_r(A)$，如果 A 非奇异.

矩阵之和的行列式可以用第 r 个转置伴随矩阵以及第 r 个复合矩阵来表示：

$$\det(sA+tB)=\sum_{k=0}^{n}s^k t^{n-k}\mathrm{tr}(\mathrm{adj}_k(A)C_r(B)) \qquad (0.8.12.3)$$

29 特别地，$\det(A+I)=\sum_{k=0}^{n}\mathrm{tr}\ \mathrm{adj}_k(A)=\sum_{k=0}^{n}\mathrm{tr}C_r(B)$.

0.9 特殊类型的矩阵

某种特殊类型的矩阵频繁地出现并有重要的性质. 我们将其中的一些罗列在这里以备参考和术语查询.

0.9.1 对角矩阵

矩阵 $D=[d_{ij}]\in M_{n,m}(\mathbf{F})$是**对角的**(diagonal)，如果只要 $j\neq i$，就有 $d_{ij}=0$. 如果一个对角矩阵的所有对角元素都是正(非负)实数，我们就称之为**正(非负)对角矩阵**(positive (non-negative) diagonal matrix). 术语**正对角矩阵**表示该矩阵是对角矩阵，且有正的对角元素，它不涉及有正的对角元素的一般性的矩阵. 单位矩阵 $I\in M_n$ 是一个正对角矩阵. 对角方阵 D 是**纯量矩阵**(scalar matrix)，如果它的对角元素全都相等，也就是说，对某个 $\alpha\in\mathbf{F}$ 有 $D=\alpha I$. 用纯量矩阵左乘或者右乘一个矩阵与对应的纯量乘该矩阵有同样的效果.

如果 $A=[a_{ij}]\in M_{n,m}(\mathbf{F})$，且 $q=\min\{m, n\}$，那么 $\mathrm{diag}A=[a_{11}, \cdots, a_{qq}]^{\mathrm{T}}\in\mathbf{F}^q$ 就表示 A 的对角元素组成的向量(0.2.1). 反过来，如果 $x\in\mathbf{F}^q$ 且 m 和 n 是满足 $\min\{m, n\}=q$ 的正整数，那么 $\mathrm{diag}x\in M_{n,m}(\mathbf{F})$就表示满足 $\mathrm{diag}A=x$ 的 $n\times m$ 对角矩阵 A，为使得 $\mathrm{diag}x$ 有良好的定义，必须对 m 和 n 两者加以指定. 对任何 $a_1, \cdots, a_n\in\mathbf{F}$，$\mathrm{diag}(a_1, \cdots, a_n)$ 总是表示这样的矩阵 $A=[a_{ij}]\in M_n(\mathbf{F})$：对每个 $i=1, \cdots, n$ 都有 $a_{ii}=a_i$，而对 $i\neq j$ 则有 $a_{ij}=0$.

假设 $D=[d_{ij}]$，$E=[e_{ij}]\in M_n(\mathbf{F})$是对角矩阵，又设给定 $A=[a_{ij}]\in M_n(\mathbf{F})$. 那么(a) $\det D=\prod_{i=1}^n d_{ii}$；(b)$D$ 是非奇异的，当且仅当所有 $d_{ii}\neq 0$；(c)用 D **左乘** A 就是用 D 的对角元素乘以 A 的行(DA 的第 i 行即是 d_{ii} 乘以 A 的第 i 行)；(d)用 D **右乘** A 就是用 D 的对角元素乘以 A 的列，也就是说，AD 的第 j 列即是 d_{jj} 乘以 A 的第 j 列；(e)$DA=AD$ 当且仅当只要 $d_{ii}\neq d_{jj}$ 就有 $a_{ij}=0$；(f)如果 D 的所有对角元素都不相同，且 $DA=AD$，那么 A 是对角的；(g)对任何正整数 k，$D^k=\mathrm{diag}(d_{11}^k, \cdots, d_{nn}^k)$；(h)任何两个同样大小的对角矩阵 D 和 E 都是可交换的：$DE=\mathrm{diag}(d_{11}e_{11}, \cdots, d_{nn}e_{nn})=ED$.

0.9.2 分块对角矩阵与直和

形如

$$A=\begin{bmatrix} A_{11} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_{kk} \end{bmatrix}$$

的一个矩阵 $A\in M_n(\mathbf{F})$称为**分块对角的**(block diagonal)，其中 $A_{ii}\in M_{n_i}(\mathbf{F})$，$i=1, \cdots, k$，$\sum_{i=1}^{k} n_i=n$，而分块对角线上方与下方所有的块都是零块. 将这样一个矩阵写成

$$A=A_{11}\oplus A_{22}\oplus\cdots\oplus A_{kk}=\bigoplus_{i=1}^{k} A_{ii}$$

是很方便的. 这就是矩阵 $A_{11}, \cdots, A_{kk}$ 的**直和**(direct sum). 分块对角矩阵的许多性质都是对角矩阵性质的推广. 例如，$\det(\oplus_{i=1}^k A_{ii})=\prod_{i=1}^k \det A_{ii}$，所以，$A=\oplus A_{ii}$是非奇异的，当且仅当每个 $A_{ii}(i=1, \cdots, k)$都是非奇异的. 此外，两个直和 $A=\oplus_{i=1}^k A_{ii}$ 与 $B=\oplus_{i=1}^k B_{ii}$ 可交换(每个 A_{ii} 与 B_{ii} 都有同样的大小)，当且仅当每一对 A_{ii} 与 B_{ii} 可交换，$i=1, \cdots, k$. 又有 $\mathrm{rank}(\oplus_{i=1}^k A_{ii})=\sum_{i=1}^{k}\mathrm{rank}A_{ii}$.

如果 $A\in M_n$ 与 $B\in M_n$ 是非奇异的，那么 $(A\oplus B)^{-1}=A^{-1}\oplus B^{-1}$ 且 $(\det(A\oplus B))\times(A\oplus B)^{-1}=(\det A)(\det B)(A^{-1}\oplus B^{-1})=((\det B)(\det A)A^{-1}\oplus(\det A)(\det B)B^{-1})$，所以连续性推理就确保有

$$\operatorname{adj}(A\oplus B)=(\det B)\operatorname{adj}A\oplus(\det A)\operatorname{adj}B \tag{0.9.2.1}$$

0.9.3 三角矩阵

矩阵 $T=[t_{ij}]\in M_{n,m}(\mathbf{F})$ 是**上三角的**(upper triangular)，如果只要 $i>j$ 就有 $t_{ij}=0$. 如果只要 $i\geqslant j$ 就有 $t_{ij}=0$，那么就称 T 是**严格上三角的**(strictly upper triangular). 类似地，T 是**下三角的**(lower triangular)(**严格下三角的**，strictly lower triangular)，如果它的转置是上三角的(严格上三角的). 三角矩阵或者是下三角的，或者是上三角的；**严格三角矩阵**(strictly triangular matrix)既是严格上三角的，又是严格下三角的. **单位三角矩阵**(unit triangular matrix)是一个主对角线上元素均为 1 的三角矩阵(上三角的或者下三角的). 有时用名词**右**(代替上)和**左**(代替下)来描述三角矩阵.

设给定 $T\in M_{n,m}(\mathbf{F})$. 如果 T 是上三角的，那么当 $n\leqslant m$ 时有 $T=[R\quad T_2]$，而当 $n\geqslant m$ 时则有 $T=\begin{bmatrix}R\\0\end{bmatrix}$；$R\in M_{\min\{n,m\}}(\mathbf{F})$ 是上三角的，而 T_2 是任意的($n=m$ 时它是空的). 如果 T 是下三角的，那么当 $n\leqslant m$ 时有 $T=[L\quad 0]$，而当 $n\geqslant m$ 时则有 $T=\begin{bmatrix}L\\T_2\end{bmatrix}$；$L\in M_{\min\{n,m\}}(\mathbf{F})$ 是下三角的，而 T_2 是任意的(当 $n=m$ 时它是空的).

正方的三角矩阵与正方的对角矩阵共享这样的性质：其行列式等于其对角元素之乘积. 正方的三角矩阵与相同大小的其他的正方的三角矩阵不一定是可交换的. 然而，如果 $T\in M_n$ 是三角的，其对角元素均不相同，又与 $B\in M_n$ 可交换，那么 B 必定是与 T 类型相同的三角矩阵(2.4.5.1).

对每个 $i=1,\cdots,n$，用一个下三角矩阵 L 左乘 $A\in M_n(\mathbf{F})$ $(A\rightarrow LA)$ 就是用 A 的前 i 行的一个线性组合来替代 A 的第 i 行. 对 A 执行有限多次第 3 类行运算(0.3.3)的结果是一个矩阵 LA，其中 L 是一个单位下三角矩阵. 关于列运算以及用上三角矩阵作的右乘运算可以得出相应的命题.

三角矩阵的秩至少是其主对角线上非零元素的个数(也可以大于这个数). 如果一个正方的三角矩阵是非奇异的，那么它的逆也是同一类型的三角矩阵；同样大小且同种类型的正方的三角矩阵的乘积是同样类型的三角矩阵；这样一个矩阵乘积中的每个位于 (i,i) 处的对角元素都是其因子的位于 (i,i) 处的元素之乘积.

0.9.4 分块三角矩阵

形如

$$A=\begin{bmatrix}A_{11} & \bigstar & \bigstar\\ & \ddots & \bigstar\\ \mathbf{0} & & A_{kk}\end{bmatrix} \tag{0.9.4.1}$$

的矩阵 $A\in M_n(\mathbf{F})$ 是**分块上三角的**(block upper triangular)，其中 $A_{ii}\in M_{n_i}(\mathbf{F})$，$i=1,\cdots,$

k，$\sum_{i=1}^{k} n_i = n$，且位于分块对角线下方所有的分块都为零．进而，它是严格分块上三角的，如果它所有的对角分块都是零块．一个矩阵是**分块下三角的**(block lower triangular)，如果它的转置是分块上三角的；它是**严格分块下三角的**(strictly block lower triangular)，如果它的转置是严格分块上三角的．我们说一个矩阵是**分块三角的**(block triangular)，要么它是分块下三角的，要么它是分块上三角的；一个矩阵既是分块下三角的又是分块上三角的，当且仅当它是分块对角的．

对角块均为 1×1 或者 2×2 的分块上三角矩阵称为**拟上三角的**(upper quasitriangular)．一个矩阵是**拟下三角的**(lower quasitriangular)，如果它的转置是拟上三角的；它是**拟三角的**(quasitriangular)，如果它是拟上三角的，或者是拟下三角的．既是拟上三角又是拟下三角的矩阵就是**拟对角的**(quasidiagonal)．

考虑(0.9.4.1)中的分块三角方阵 A．我们有 $\det A = \det A_{11}\cdots\det A_{kk}$，且 $\operatorname{rank} A \geqslant \operatorname{rank} A_{11}+\cdots+\operatorname{rank} A_{kk}$．如果 A 非奇异(即对所有 $i=1$，…，k，A_{ii} 均是非奇异的)，那么 A^{-1} 就是一个与 A 共形分划的分块三角矩阵，其对角块是 A_{11}^{-1}，…，A_{kk}^{-1}．

如果 $A\in M_n(\mathbf{F})$ 是上三角的，那么，对$\{1，\cdots，n\}$的**任意一个**顺序分划 α_1，…，α_t，$[A(\alpha_i，\alpha_j)]_{i,j=1}^{t}$ 都是分块上三角的(0.7.2)．

0.9.5 置换矩阵

方阵 P 是一个**置换矩阵**(permutation matrix)，如果每一行以及每一列中恰好都仅有一个元素等于 1，而所有其他元素均为 0．用这样一个矩阵来作乘法的效果等同于对被乘的矩阵的行或者列作置换．例如

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}=\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

就描述了一个置换矩阵作用在一个向量时对该向量的行(元素)所产生的置换：它把第 1 个元素安置到第 2 个位置上，把第 2 个元素安置到第 1 个位置上，而把第 3 个元素保留在第 3 个位置上．用一个 $m\times m$ 置换矩阵 P 左乘矩阵 $A\in M_{m,n}$ 是对 A 的行进行置换，而用一个 $n\times n$ 置换矩阵 P 右乘 A 是对 A 的列进行置换．执行第 1 类初等运算(0.3.3)的矩阵是一种特殊类型的置换矩阵的例子，称为**对换**(transposition)．任何置换矩阵都是对换的乘积．

置换矩阵的行列式等于± 1，所以置换矩阵是非奇异的．尽管置换矩阵不一定可交换，但两个置换矩阵的乘积仍然是一个置换矩阵．由于单位矩阵是置换矩阵，且对每个置换矩阵 P 有 $P^{\mathrm{T}}=P^{-1}$，$n\times n$ 置换矩阵的集合是 $GL(n，\mathbf{C})$的一个基数为 $n!$ 的子群．

由于用 $P^{\mathrm{T}}=P^{-1}$ 右乘就是按照与用 P 左乘对行作置换的同样的方式对列作置换，因而变换 $A\rightarrow PAP^{\mathrm{T}}$ 以同样的方式对 $A\in M_n$ 的行和列(从而也对主对角线上的元素)进行置换．在以 A 为系数矩阵的线性方程组中，这一变换等同于用同样的方式对变量和方程重新编号．对某个置换矩阵 P 使得 PAP^{T} 成为三角矩阵的矩阵 $A\in M_n$ 称为**本性三角的**(essentially triangular)，这些矩阵与三角矩阵有许多共同点．

如果 $\Lambda\in M_n$ 是对角的，而 $P\in M_n$ 是置换矩阵，那么 $P\Lambda P^{\mathrm{T}}$ 是对角矩阵．

$n\times n$ **反序矩阵**(reversal matrix)是置换矩阵

$$K_n=\begin{bmatrix} & & 1\\ & \iddots & \\ 1 & & \end{bmatrix}=[\kappa_{ij}]\in M_n \tag{0.9.5.1}$$

其中 $\kappa_{i,n-i+1}=1$(对 $i=1,\cdots,n$)，而所有其他的元素均为零. K_nA 的行是按照相反次序排列的 A 的行，而 AK_n 的列则是按照相反次序排列的 A 的列. 反序矩阵有时也称为 **sip 矩阵**(**标准对合置换**(standard involutory permutation))、**后向恒等式**(backward identity)或者**交换矩阵**(exchange matrix).

对任何 $n\times n$ 矩阵 $A=[a_{ij}]$，元素 $a_{i,n-i+1}$(对 $i=1,\cdots,n$)构成它的**反对角线**(counterdiagonal)，有时也称为**第二对角线**(second diagonal)、**后向对角线**(backward diagonal)、**交叉对角线**(cross diagonal)、**右对角线**(dexter-diagonal)或者**反对角线**(antidiagonal).

广义置换矩阵(generalized permutation matrix)是一个形如 $G=PD$ 的矩阵，其中 $P,D\in M_n$，P 是一个置换矩阵，而 D 则是非奇异的对角矩阵. $n\times n$ 广义置换矩阵的集合是 $GL(n,\mathbf{C})$ 的一个子群.

0.9.6 循环矩阵

形如

$$A=\begin{bmatrix} a_1 & a_2 & & \cdots & a_n\\ a_n & a_1 & a_2 & \cdots & a_{n-1}\\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2}\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ a_2 & a_3 & \cdots & a_n & a_1 \end{bmatrix} \tag{0.9.6.1}$$

的矩阵 $A\in M_n(\mathbf{F})$是一个**循环矩阵**(circulant matrix). 每一行都是前一行向前循环一步得到的，每一行中的元素都是第 1 行中元素的一个循环排列. $n\times n$ 置换矩阵

$$C_n=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0\\ \vdots & 0 & 1 & & \vdots\\ & & \ddots & \ddots & 0\\ 0 & & & & 1\\ 1 & 0 & & \cdots & 0 \end{bmatrix}=\begin{bmatrix} 0 & I_{n-1}\\ 1 & 0_{1,n-1} \end{bmatrix} \tag{0.9.6.2}$$

是**基本循环置换矩阵**(basic circulant permutation matrix). 矩阵 $A\in M_n(\mathbf{F})$可以写成形式

$$A=\sum_{k=0}^{n-1}a_{k+1}C_n^k \tag{0.9.6.3}$$

(它是关于矩阵 C_n 的多项式)，当且仅当它是循环的. 我们有 $C_n^0=I=C_n^n$，而系数 $a_1,\cdots,a_n$ 则是 A 的第 1 行的元素. 这个表达式表明阶为 n 的循环矩阵是一个交换代数：循环矩阵的线性组合和乘积是循环矩阵，非奇异的循环矩阵的逆是循环矩阵，任何两个同样大小的循环矩阵是可交换的.

0.9.7 Toeplitz 矩阵

形如

$$A=\begin{bmatrix} a_0 & a_1 & a_2 & \cdots & \cdots & a_n \\ a_{-1} & a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{-2} & a_{-1} & a_0 & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & a_1 \\ a_{-n} & a_{-n+1} & \cdots & \cdots & a_{-1} & a_0 \end{bmatrix}$$

的矩阵是一个 Toeplitz 矩阵. 对某个给定的序列 a_{-n}，a_{-n+1}，$\cdots$，a_{-1}，a_0，a_1，a_2，$\cdots$，a_{n-1}，$a_n\in\mathbf{C}$，元素 a_{ij} 与 a_{j-i} 相等. A 中向下沿着与主对角线平行的对角线上的元素是常数. 根据它们对标准基$\{e_1,\cdots,e_{n+1}\}$的作用效果，Toeplitz 矩阵

$$B=\begin{bmatrix} 0 & 1 & & \mathbf{0} \\ & 0 & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & 0 \end{bmatrix} \quad 和 \quad F=\begin{bmatrix} 0 & & & \mathbf{0} \\ 1 & 0 & & \\ & \ddots & \ddots & \\ \mathbf{0} & & 1 & 0 \end{bmatrix}$$

称为**后向位移**(backward shift)和**前向位移**(forward shift). 此外，$F=B^{\mathrm{T}}$ 且 $B=F^{\mathrm{T}}$. 矩阵 $A\in M_{n+1}$ 可以写成

$$A=\sum_{k=1}^{n}a_{-k}F^k+\sum_{k=0}^{n}a_kB^k \tag{0.9.7.1}$$

当且仅当它是 Toeplitz 矩阵. Toeplitz 矩阵自然出现在与三角矩有关的问题中.

利用适当大小的反序矩阵 K(0.9.5.1)，注意前向平移矩阵与后向平移矩阵是相关联的：$F=KBK=B^{\mathrm{T}}$ 以及 $B=KFK=F^{\mathrm{T}}$. 表达式(0.9.7.1)确保对任何 Toeplitz 矩阵 A 有 $KA=A^{\mathrm{T}}K$，即有 $A^{\mathrm{T}}=KAK=KAK^{-1}$.

上三角 Toeplitz 矩阵 $A\in M_{n+1}(\mathbf{F})$可以表示成 B 的多项式：

$$A=a_0I+a_1B+\cdots+a_nB^n$$

这个表达式(以及 $B^{n+1}=0$ 这一事实)使得我们清楚地看出，为什么阶为 n 的上三角 Toeplitz 矩阵是一个交换代数：上三角 Toeplitz 矩阵的线性组合以及乘积是上三角 Toeplitz 矩阵；A 是非奇异的，当且仅当 $a_0\neq 0$，且在此情形 $A^{-1}=b_0I+b_1B+\cdots+b_nB^n$ 也是上三角 Toeplitz 矩阵，其中 $b_0=a_0^{-1}$，而对 $k=1$，$\cdots$，n 有 $b_k=a_0^{-1}\left(\sum_{m=0}^{k-1}a_{k-m}b_m\right)$. 任何两个同样大小的上三角 Toeplitz 矩阵可交换.

0.9.8 Hankel 矩阵

形如

$$A=\begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_n \\ a_1 & a_2 & \cdots & \iddots & a_{n+1} \\ a_2 & & \iddots & & \vdots \\ \vdots & a_n & & & a_{2n-1} \\ a_n & a_{n+1} & \cdots & a_{2n-1} & a_{2n} \end{bmatrix}$$

的矩阵 $A\in M_{n+1}(\mathbf{F})$ 是 **Hankel 矩阵**．对某个给定的序列 a_0，a_1，a_2，…，a_{2n-1}，a_{2n}，每个元素 a_{ij} 都与 a_{i+j-2} 相等．A 的沿着与主对角线垂直的对角线上的元素是常数．Hankel 矩阵自然出现在与幂矩有关的问题中．利用适当大小的反序矩阵 K(0.9.5.1)，注意，对任何 Toeplitz 矩阵 A，KA 与 AK 都是 Hankel 矩阵；对任何 Hankel 矩阵 H，KH 和 HK 都是 Toeplitz 矩阵．由于 $K=K^{\mathrm{T}}=K^{-1}$，且 Hankel 矩阵都是对称的，这就意味着任何 Toeplitz 矩阵是带有特殊构造的两个对称矩阵的乘积：一个是反序矩阵，另一个是 Hankel 矩阵．

0.9.9 Hessenberg 矩阵

矩阵 $A=[a_{ij}]\in M_n(\mathbf{F})$ 称为是一个**上 Hessenberg 型**(upper Hessenberg form)或者**上 Hessenberg 矩阵**(upper Hessenberg matrix)，如果对所有 $i>j+1$ 都有 $a_{ij}=0$：

$$A=\begin{bmatrix} a_{11} & & & & \bigstar \\ a_{21} & a_{22} & & & \\ & a_{32} & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & a_{n,n-1} & a_{nn} \end{bmatrix}$$

一个上 Hessenberg 矩阵 A 说成是**未约化的**(unreduced)，如果它所有次对角线元素都不等于零，也就是说，如果对所有 $i=1$，…，$n-1$ 都有 $a_{i+1,i}\neq 0$．这样一个矩阵的秩至少为 $n-1$，由于它的前 $n-1$ 列是线性无关的．

设 $A\in M_n(\mathbf{F})$ 是未约化的上 Hessenberg 矩阵．那么，对所有 $\lambda\in\mathbf{F}$，$A-\lambda I$ 都是未约化的上 Hessenberg 矩阵，故而对所有 $\lambda\in\mathbf{F}$ 有 $\operatorname{rank}(A-\lambda I)\geqslant n-1$．

矩阵 $A\in M_n(\mathbf{F})$ 是下 Hessenberg 矩阵，如果 A^{T} 是上 Hessenberg 矩阵．

0.9.10 三对角矩阵，双对角矩阵以及其他构造的矩阵

既是上 Hessenberg 型又是下 Hessenberg 型的矩阵 $A=[a_{ij}]\in M_n(\mathbf{F})$ 称为**三对角的**(tridiagonal)，也就是说，A 是三对角的，如果只要 $|i-j|>1$，就有 $a_{ij}=0$：

$$A=\begin{bmatrix} a_1 & b_1 & & 0 \\ c_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ 0 & & c_{n-1} & a_n \end{bmatrix} \tag{0.9.10.1}$$

A 的行列式可以归纳地加以计算：从 $\det A_1=a_1$ 出发，$\det A_2=a_1a_2-b_1c_1$，接下来计算一系列 2×2 矩阵乘积

35

$$\begin{bmatrix} \det A_{k+1} & 0 \\ \det A_k & 0 \end{bmatrix}=\begin{bmatrix} a_{k+1} & -b_kc_k \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \det A_k & 0 \\ \det A_{k-1} & 0 \end{bmatrix},\quad k=2,\cdots,n-1$$

Jacobi 矩阵是次对角线元素均为正数的实对称的三对角矩阵．

上双对角矩阵(upper bidiagonal matrix)$A\in M_n(\mathbf{F})$ 是满足 $c_1=\cdots=c_{n-1}=0$ 的三对角矩阵(0.9.10.1)．矩阵 $A\in M_n(\mathbf{F})$ 是**下双对角的**(lower bidiagonal)，如果 A^{T} 是上双对角的．

分块三对角矩阵或者分块双对角矩阵有与(0.9.10.1)中的式样相似的块状结构，其对角线上的块都是正方的，而超对角分块与次对角分块的大小则由离它们最近的对角块的大小来决定．

矩阵 $A=[a_{ij}]\in M_n(\mathbf{F})$是**全对称的**(persymmetric)，如果对所有 i，$j=1$，…，n 都有 $a_{ij}=a_{n+1-j,n+1-i}$，也就是说，全对称矩阵是关于反对角线对称的矩阵．它的另一种也是很有用的刻画是：A 是全对称的，如果 $K_nA=A^TK_n$，其中 K_n 是反序矩阵(0.9.5.1)．如果 A 是全对称且可逆的，那么 A^{-1} 也是全对称的，这是因为 $K_nA^{-1}=(AK_n)^{-1}=(K_nA^T)^{-1}=A^{-T}K_n$．Toeplitz 矩阵是全对称的．我们说 $A\in M_n(\mathbf{F})$是**斜全对称的**(skew persymmetric)，如果 $K_nA=-A^TK_n$．非奇异的斜全对称矩阵的逆是斜全对称的．

满足 $K_nA=A^*K_n$ 的复矩阵 $A\in M_n(\mathbf{F})$是**全 Hermite 的**(perhermitian)．A 是**斜全 Hermite**(skew perhermitian)，如果 $K_nA=-A^*K_n$．非奇异的全 Hermite(斜全 Hermite)矩阵的逆是全 Hermite(斜全 Hermite)的．

矩阵 $A=[a_{ij}]\in M_n(\mathbf{F})$是**中心对称的**(centrosymmetric)，如果对所有 i，$j=1$，…，n 都有 $a_{ij}=a_{n+1-i,n+1-j}$．等价地，A 是中心对称的，如果 $K_nA=AK_n$；A 是**斜中心对称的**(skew centrosymmetric)，如果 $K_nA=-AK_n$．中心对称的矩阵是关于其几何中心为对称的，正如下面的例子

$$A=\begin{bmatrix}1&2&3&4&5\\0&6&7&8&9\\-1&-2&-3&-2&-1\\9&8&7&6&0\\5&4&3&2&1\end{bmatrix}$$

所描述的那样．如果 A 是非奇异的，且是中心对称的(斜中心对称的)，那么 A^{-1} 也是中心对称的(斜中心对称的)，这是由于 $KA^{-1}=(AK_n)^{-1}=(K_nA)^{-1}=A^{-1}K_n$．如果 A 与 B 是中心对称的，那么 AB 是中心对称的，这是由于 $K_nAB=AK_nB=ABK_n$．如果 A 与 B 是斜中心对称的，那么 AB 是中心对称的．

中心对称的矩阵 $A\in M_n(\mathbf{F})$有特殊的分块结构．如果 $n=2m$，那么

$$A=\begin{bmatrix}B&K_mCK_m\\C&K_mBK_m\end{bmatrix},\quad B,C\in M_m(\mathbf{F})\qquad(0.9.10.2)$$

如果 $n=2m+1$，那么

$$A=\begin{bmatrix}B&K_my&K_mCK_m\\x^T&\alpha&x^TK_m\\C&y&K_mBK_m\end{bmatrix},\quad B,C\in M_m(\mathbf{F}),x,y\in\mathbf{F}^m,\alpha\in\mathbf{F}\qquad(0.9.10.3)$$

满足 $K_nA=\overline{A}K_n$ 的复矩阵 $A\in M_n$ 是**中心 Hermite 的**(centrohermitian)；它是**斜中心 Hermite 的**(skew centrohermitian)，如果 $K_nA=-\overline{A}K_n$．一个非奇异的中心 Hermite(斜中心 Hermite)矩阵的逆是中心 Hermite(斜中心 Hermite)的．中心 Hermite 矩阵的乘积是中心 Hermite 的． 36

0.9.11 Vandermonde 矩阵与 Lagrange 插值

Vandermonde 矩阵 $A\in M_n(\mathbf{F})$有形式

$$A=\begin{bmatrix}1&x_1&x_1^2&\cdots&x_1^{n-1}\\1&x_2&x_2^2&\cdots&x_2^{n-1}\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&x_n&x_n^2&\cdots&x_n^{n-1}\end{bmatrix}\qquad(0.9.11.1)$$

其中 x_1，…，$x_n\in\mathbf{F}$，也就是说，$A=[a_{ij}]$，其中 $a_{ij}=x_i^{j-1}$. 事实上有

$$\det A=\prod_{\substack{i,j=1\\i>j}}^{n}(x_i-x_j)\tag{0.9.11.2}$$

所以 Vandermonde 矩阵是非奇异的，当且仅当诸参数 x_1，…，x_n 各不相同.

如果 x_1，…，x_n 是不相同的，Vandermonde 矩阵(0.9.11.1)的逆 $A^{-1}=[\alpha_{ij}]$是

$$\alpha_{ij}=(-1)^{i-1}\frac{S_{n-i}(x_1,\cdots,\hat{x}_j,\cdots,x_n)}{\prod\limits_{k\neq j}(x_k-x_j)},\quad i,j=1,\cdots,n$$

其中 $S_0=1$，而如果 $m>0$，则 $S_m(x_1$，…，$\hat{x}_j$，…，$x_n)$是 $n-1$ 个变量 $x_k(k=1$，…，n，$k\neq j)$的第 m 个初等对称函数，见(1.2.14).

Vandermonde 矩阵出现在**插值问题**(interpolation problem)中，该问题是寻求次数至多为 $n-1$ 且系数在 $\mathbf{F}$ 中的多项式 $p(x)=a_{n-1}x^{n-1}+a_{n-2}x^{n-2}+\cdots+a_1x+a_0$，它满足

$$\begin{aligned}p(x_1)&=a_0+a_1x_1+a_2x_1^2+\cdots+a_{n-1}x_1^{n-1}=y_1\\p(x_2)&=a_0+a_1x_2+a_2x_2^2+\cdots+a_{n-1}x_2^{n-1}=y_2\\&\ \ \vdots\qquad\vdots\qquad\qquad\vdots\qquad\qquad\qquad\vdots\quad\vdots\\p(x_n)&=a_0+a_1x_n+a_2x_n^2+\cdots+a_{n-1}x_n^{n-1}=y_n\end{aligned}\tag{0.9.11.3}$$

其中 x_1，…，x_n 和 y_1，…，y_n 是 $\mathbf{F}$ 中给定的元素. 插值条件(0.9.11.3)是有 n 个未知系数 a_0，…，a_{n-1} 以及 n 个方程的方程组，它们有形式 $Aa=y$，其中 $a=[a_0\ \cdots\ a_{n-1}]^{\mathrm{T}}\in\mathbf{F}^n$，$y=[y_1\cdots y_n]^{\mathrm{T}}\in\mathbf{F}^n$，而 $A\in M_n(\mathbf{F})$是 Vandermonde 矩阵(0.9.11.1). 这个插值问题总会有解，如果诸点 x_1，x_2，…，x_n 各不相同，这是因为在此情形 A 是非奇异的.

如果诸点 x_1，x_2，…，x_n 各不相同，则插值多项式的系数原则上可以通过求解方程组
37 (0.9.11.3)得到，但通常更为有用的是将插值多项式 $p(x)$表示成为 **Lagrange 插值多项式**(interpolation polynomial)

$$L_i(x)=\frac{\prod\limits_{j\neq i}(x-x_j)}{\prod\limits_{j\neq i}(x_i-x_j)},\quad i=1,\cdots,n$$

的线性组合. 每一个多项式 $L_i(x)$的次数都是 $n-1$ 且有如下性质：如果 $k\neq i$，则有 $L_i(x_k)=0$，但 $L_i(x_i)=1$. **Lagrange 插值公式**

$$p(x)=y_1L_1(x)+\cdots+y_nL_n(x)\tag{0.9.11.4}$$

提供了满足诸方程(0.9.11.3)的一个阶至多为 $n-1$ 的多项式.

0.9.12 Cauchy 矩阵

Cauchy 矩阵 $A\in M_n(\mathbf{F})$是一个形如 $A=[(a_i+b_j)^{-1}]_{i,j=1}^n$ 的矩阵，其中 a_1，…，a_n 和 b_1，…，b_n 是满足 $a_i+b_j\neq0$(对所有 i，$j=1$，…，n)的纯量. 事实上有

$$\det A=\frac{\prod\limits_{1\leqslant i<j\leqslant n}(a_j-a_i)(b_j-b_i)}{\prod\limits_{1\leqslant i\leqslant j\leqslant n}(a_i+b_j)}\tag{0.9.12.1}$$

所以，A 是非奇异的，当且仅当对所有 $i\neq j$ 都有 $a_i\neq a_j$ 以及 $b_i\neq b_j$. **Hilbert 矩阵** $H_n=[(i+$

$j-1)^{-1}]_{i,j=1}^{n}$是这样一种矩阵：它既是 Cauchy 矩阵，又是 Hankel 矩阵. 事实上有

$$\det H_n = \frac{(1!2!\cdots(n-1)!)^4}{1!2!\cdots(2n-1)!} \tag{0.9.12.2}$$

所以，Hilbert 矩阵总是非奇异的. 它的逆矩阵 $H_n^{-1}=[h_{ij}]_{i,j=1}^{n}$的元素是

$$h_{ij} = \frac{(-1)^{i+j}(n+i-1)!(n+j-1)!}{((i-1)!(j-1)!)^2(n-i)!(n-j)!(i+j-1)} \tag{0.9.12.3}$$

0.9.13 对合矩阵，幂零矩阵，射影矩阵，共轭对合矩阵

矩阵 $A\in M_n(\mathbf{F})$是

- **对合矩阵**(involution)，如果 $A^2=I$，即 $A=A^{-1}$(也用术语**对合的**(involutory))
- **幂零矩阵**(nilpotent)，如果对某个正整数 k 有 $A^k=0$. 最小的这样的 k 称为 A 的**幂零指数**(index of nilpotence)
- **射影矩阵**(projection)，如果 $A^2=A$(也用术语**幂等的**(idempotent))

现在假设 $\mathbf{F}=\mathbf{C}$. 矩阵 $A\in M_n$ 是

- **Hermite 射影矩阵**(Hermitian projection)，如果 $A^*=A$ 且 $A^2=A$(也用术语**正交射影**(orthogonal projection)**矩阵**，见(4.1.P19))
- **共轭对合矩阵**(coninvolution)，如果 $A\bar{A}=I$，即 $\bar{A}=A^{-1}$(也用术语**共轭对合的**(coninvolutory)) 38

0.10 基的变换

设 V 是域 $\mathbf{F}$ 上的一个 n 维向量空间，而向量组$\mathcal{B}_1=\{v_1, v_2, \cdots, v_n\}$是 V 的一组基. 任何向量 $x\in V$ 都可以表示成 $x=\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n$，这是因为 V 由$\mathcal{B}_1$ 生成. 如果还有另外一种用同样的基给出的表示 $x=\beta_1v_1+\beta_2v_2+\cdots+\beta_nv_n$，那么

$$0 = x - x = (\alpha_1-\beta_1)v_1+(\alpha_2-\beta_2)v_2+\cdots+(\alpha_n-\beta_n)v_n$$

由此推出所有的 $\alpha_i-\beta_i=0$，这是因为向量组$\mathcal{B}_1$ 线性无关. 给定基$\mathcal{B}_1$，由 V 到 $\mathbf{F}^n$ 的线性映射

$$x\rightarrow [x]_{\mathcal{B}_1} = \begin{bmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{bmatrix},\quad 其中\ x=\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n$$

有良好的定义，且是一对一的满映射. 纯量 α_i 是 x 关于基$\mathcal{B}_1$ 的**坐标**(coordinate)，而列向量$[x]_{\mathcal{B}_1}$则是 x 唯一的$\mathcal{B}_1$ **坐标表示**($\mathcal{B}_1$-coordinate representation).

设 T：$V\rightarrow V$ 是一个给定的线性变换. 一旦我们知道了 n 个向量 $Tv_1, Tv_2, \cdots, Tv_n$，T 在 $x\in V$ 上的作用就被确定了，因为任何 $x\in V$ 都有一个唯一的表示 $x=\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n$，而根据线性则有 $Tx=T(\alpha_1v_1+\cdots+\alpha_nv_n)=T(\alpha_1v_1)+\cdots+T(\alpha_nv_n)=\alpha_1Tv_1+\cdots+\alpha_nTv_n$. 于是，只要知道了$[x]_{\mathcal{B}_1}$，$Tx$ 的值就确定了.

设$\mathcal{B}_2=\{w_1, w_2, \cdots, w_n\}$也是 V 的一组基(与$\mathcal{B}_1$ 可能不同也可能相同)，并假设 Tv_j 的$\mathcal{B}_2$ 坐标表示是

$$[Tv_j]_{\mathcal{B}_2}=\begin{bmatrix}t_{1j}\\ \vdots \\ t_{nj}\end{bmatrix},\quad j=1,2,\cdots,n$$

这样一来，对任何 $x\in V$ 我们有

$$[Tx]_{\mathcal{B}_2}=\left[\sum_{j=1}^{n}\alpha_j Tv_j\right]_{\mathcal{B}_2}=\sum_{j=1}^{n}\alpha_j[Tv_j]_{\mathcal{B}_2}=\sum_{j=1}^{n}\alpha_j\begin{bmatrix}t_{1j}\\ \vdots \\ t_{nj}\end{bmatrix}=\begin{bmatrix}t_{11} & \cdots & t_{1n}\\ \vdots & \ddots & \vdots \\ t_{n1} & \cdots & t_{nn}\end{bmatrix}=\begin{bmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{bmatrix}$$

$n\times n$ 阵列$[t_{ij}]$与 T 有关，且与基$\mathcal{B}_1$ 以及$\mathcal{B}_2$ 的选取有关，但它与 x 无关．定义 T 的$\mathcal{B}_1$-$\mathcal{B}_2$ **基表示**($\mathcal{B}_1$-$\mathcal{B}_2$ basis representation)为

$$_{\mathcal{B}_2}[T]_{\mathcal{B}_1}=\begin{bmatrix}t_{11} & \cdots & t_{1n}\\ \vdots & \ddots & \vdots \\ t_{n1} & \cdots & t_{nn}\end{bmatrix}=\left[[Tv_1]_{\mathcal{B}_2}\cdots[Tv_n]_{\mathcal{B}_2}\right]$$

我们刚刚指出了，对任何 $x\in V$ 有$[Tx]_{\mathcal{B}_2}={}_{\mathcal{B}_2}[T]_{\mathcal{B}_1}[x]_{\mathcal{B}_1}$．在$\mathcal{B}_2=\mathcal{B}_1$ 这一重要的特殊情形，我们有$_{\mathcal{B}_1}[T]_{\mathcal{B}_1}$，它称为是 T 的$\mathcal{B}_1$ **基表示**(basis representation)．

考虑恒等线性变换 I：$V\rightarrow V$，它对所有 x 定义为 $Ix=x$．这样就对所有 $x\in V$ 有

$$[x]_{\mathcal{B}_2}=[Ix]_{\mathcal{B}_2}={}_{\mathcal{B}_2}[I]_{\mathcal{B}_1}[x]_{\mathcal{B}_1}={}_{\mathcal{B}_2}[I]_{\mathcal{B}_1}[Ix]_{\mathcal{B}_1}={}_{\mathcal{B}_2}[I]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[I]_{\mathcal{B}_2}[x]_{\mathcal{B}_2}$$

依次选取 $x=w_1$，w_2，…，w_n，这个恒等式使得我们可以辨识出$_{\mathcal{B}_2}[I]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[I]_{\mathcal{B}_2}$的每一列，而且还证明了

$$_{\mathcal{B}_2}[I]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[I]_{\mathcal{B}_2}=I_n$$

如果从$[x]_{\mathcal{B}_1}=[Ix]_{\mathcal{B}_1}=\cdots$出发来做同样的计算，我们就求得

$$_{\mathcal{B}_1}[I]_{\mathcal{B}_2}\,{}_{\mathcal{B}_2}[I]_{\mathcal{B}_1}=I_n$$

于是，每一个形如$_{\mathcal{B}_2}[I]_{\mathcal{B}_1}$的矩阵都是可逆的，且$_{\mathcal{B}_1}[I]_{\mathcal{B}_2}$就是它的逆．反过来，每一个可逆矩阵 $S=[s_1\,s_2\cdots s_n]\in M_n(\mathbf{F})$都有$_{\mathcal{B}_1}[I]_{\mathcal{B}}$的形式(对某个基$\mathcal{B}$)．我们可以取$\mathcal{B}$为由$[\tilde{s}_i]_{\mathcal{B}_1}=s_i(i=1，2，…，n)$所定义的向量$\{\tilde{s}_1，\tilde{s}_2，…，\tilde{s}_n\}$．向量组$\mathcal{B}$是线性无关的，因为 S 是可逆的．

注意

$$_{\mathcal{B}_2}[I]_{\mathcal{B}_1}=\left[[Iv_1]_{\mathcal{B}_2}\cdots[Iv_n]_{\mathcal{B}_2}\right]=\left[[v_1]_{\mathcal{B}_2}\cdots[v_n]_{\mathcal{B}_2}\right]$$

所以$_{\mathcal{B}_2}[I]_{\mathcal{B}_1}$描述了基$\mathcal{B}_1$ 的元素是怎样由基$\mathcal{B}_2$ 的元素构成的．现在设 $x\in V$ 并计算

$$\begin{aligned}_{\mathcal{B}_1}[T]_{\mathcal{B}_2}[x]_{\mathcal{B}_2}&=[Tx]_{\mathcal{B}_2}=[I(Tx)]_{\mathcal{B}_2}={}_{\mathcal{B}_2}[I]_{\mathcal{B}_1}[Tx]_{\mathcal{B}_1}={}_{\mathcal{B}_2}[I]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[T]_{\mathcal{B}_1}[x]_{\mathcal{B}_1}\\&={}_{\mathcal{B}_2}[I]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[T]_{\mathcal{B}_1}[Ix]_{\mathcal{B}_1}={}_{\mathcal{B}_2}[I]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[T]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[I]_{\mathcal{B}_2}[x]_{\mathcal{B}_2}\end{aligned}$$

相继选取 $x=w_1$，w_2，…，w_n，我们就得出结论

$$_{\mathcal{B}_2}[T]_{\mathcal{B}_2}={}_{\mathcal{B}_2}[I]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[T]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[I]_{\mathcal{B}_2}\qquad(0.10.1.1)$$

这个恒等式指出了，如果基改变成了$\mathcal{B}_2$，T 的$\mathcal{B}_1$ 基表示会怎样改变．为此，我们将矩阵$_{\mathcal{B}_2}[I]_{\mathcal{B}_1}$称为$\mathcal{B}_1$-$\mathcal{B}_2$ **基变换矩阵**($\mathcal{B}_1$-$\mathcal{B}_2$ change of basis matrix)．

任何矩阵 $A\in M_n(\mathbf{F})$都是某个线性变换 T：$V\rightarrow V$ 的一个基表示，因为如果$\mathcal{B}$是 V 的任意一组基，我们就能利用$[Tx]_{\mathcal{B}}=A[x]_{\mathcal{B}}$确定 Tx．对于这个 T，计算表明有$_{\mathcal{B}}[T]_{\mathcal{B}}=A$．

0.11 等价关系

设 S 是一个给定的集合，而 Γ 是 $S\times S=\{(a,b): a\in S, b\in S\}$ 的一个给定的子集．这样 Γ 就按照如下的方式定义了 S 上的一个**关系**(relation)：我们说 a 与 b **有关系**(记为 $a\sim b$)，如果 $(a,b)\in\Gamma$．S 上的一个关系称为**等价关系**(equivalence relation)，如果它是：(a)**自反的**(reflexive)(对每个 $a\in S$ 有 $a\sim a$)，(b)**对称的**(symmetric)(只要 $b\sim a$ 就有 $a\sim b$)，(c)**传递的**(transitive)(只要 $a\sim b$ 且 $b\sim c$，就有 $a\sim c$)．S 上的一个等价关系以一种自然的方式给出 S 的一个不相交的分划：对任何 $a\in S$，如果我们用 $S_a=\{b\in S: b\sim a\}$ 来定义它的**等价类**(equivalence class)，那么就有 $S=\bigcup_{a\in S}S_a$，且对每个 $a, b\in S$，或者有 $S_a=S_b$(如果 $a\sim b$)，或者有 $S_a\cap S_b=\varnothing$(如果 $a\nsim b$)．反之，S 的任何不相交的分划都可以用来定义 S 上的一个等价关系． 40

等价关系～	$A\sim B$
相合	$A=SBS^{T}$
酉相合	$A=UBU^{T}$
* 相合	$A=SBS^{*}$
共轭相似	$A=SB\bar{S}^{-1}$
等价	$A=SBT$
酉等价	$A=UBV$
对角等价	$A=D_1BD_2$
相似	$A=SBS^{-1}$
酉相似	$A=UBU^{*}$
三角等价	$A=LBR$

右边的表列出了在矩阵分析中出现的一些等价关系．因子 D_1，D_2，S，T，L 以及 R 都是方阵，且是非奇异的；U 和 V 是酉矩阵；L 是下三角的；R 是上三角的；D_1 和 D_2 是对角的；而对于等价、酉等价、三角等价或者对角等价来说，A 和 B 不一定是方阵．

只要在矩阵分析中出现一种有意义的等价关系，辨识出其等价类的一组特殊的代表元是很有用的(这种等价关系的一个**标准型**(canonical form)或者**正规型**(normal form))．换言之，我们常会希望有有效的判别法(不变量)能用来确定两个给定的矩阵是否属于同一个等价类．

抽象地说，集合 S 上**等价关系～的一个标准型**(canonical form for an equivalence relation～)是 S 的一个子集 C，它满足 $S=\bigcup_{a\in C}S_a$ 以及 $S_a\cap S_b=\varnothing$(只要 $a, b\in C$ 且 $a\neq b$)；**元素** $a\in S$ **的标准型**(canonical form of an element $a\in S$)是满足 $a\in S_c$ 的唯一的元素 $c\in C$．

对矩阵分析中一个给定的等价关系而言，重要的是用巧妙而简洁的方式选择标准型，有时我们会尝试多种方式以使得标准型能量身定做般地适用于特定的要求．例如，Jordan 标准型和 Weyr 标准型对相似来说是不同的标准型，Jordan 标准型在涉及矩阵的幂的问题中有良好的效用，而 Weyr 标准型则在与交换性有关的问题中有良好的效果．

S 上等价关系～的**不变量**(invariant)是 S 上一个函数 f，它使得只要 $a\sim b$ 就有 $f(a)=f(b)$．S 上一个等价关系～的一族不变量 $\mathcal{F}$ 称为是**完备的**(complete)，如果对所有 $f\in\mathcal{F}$ 有 $f(a)=f(b)$ 成立，当且仅当 $a\sim b$．一族完备的不变量常被称作**完备不变量系**(complete system of invariants)．例如，对酉等价来说，矩阵的奇异值就是一个完备不变量系． 41