Essential Linear Algebra

路致远

2024 年 12 月 30 日

1　线性空间与变换

“我打开线代书一看，满篇都是变换。”——阮东

1.1　线性空间

定义 1. 定义在集合 G，域 F 上的线性空间 V 是指集合 G 及其上的加法和与 F 中元素的数乘运算。
其中，G 上加法构成 *Able* 群，数乘运算满足结合和分配律。

1.2　线性空间的结构

1.2.1　空间内结构：直和

我们可以对线性空间进行直和分解。

定义 2. 线性子空间 W_i 构成线性空间 V 的直和分解

$$V = \bigoplus_{i=1}^{n} W_i$$

假若

$$\forall v \in V, \exists w_i \in W_i, s.t. v = \sum_{i=1}^{n} w_i$$

且这样的组合唯一。

命题 1. 若线性空间 $W=\sum_{i=1}^{m} W_i$，则以下命题等价

1.$W=\bigoplus_{i=1}^{m} W_i$

*2.*零元$0\in W$表示为W_i元素和的方法唯一

3.$W_i\cap\sum_{j\neq i} W_j=0$

4.$dim[W]=\sum_{i=1}^{m} dim[W_i]$

证明.

2⇒1:

若 $a\in W, a=\sum_{i=1}^{m} a_i=\sum_{i=1}^{m} b_i;\ a_i,b_i\in W_i$

则 $0=\sum_{i=1}^{m}(a_i-b_i)$

从而 $a_i-b_i=0$ 即 $a_i=b_i$

3⇒2:

若 $0=\sum_{i=1}^{m} a_i$

即 $-a_i=\sum_{j\neq i} a_j$

$\because -a_i\in W_i, \sum_{j\neq i} a_j\in\sum_{j\neq i} W_j$ 且 $W_i\cap\sum_{j\neq i} W_j=0$

$\therefore -a_i, (\sum_{j\neq i} a_j)=0$ 对所有 i 成立

$\therefore$ 零元$0\in W$表示为W_i元素和的方法唯一，即$0=0+0+\cdots+0$

4⇒3:

$dim[W_i\cap\sum_{j\neq i} W_j]=dim[W_i]+dim[\sum_{j\neq i} W_j]-dim[\sum_{i=1}^{m} W_j]=0$

从而 $W_i\cap\sum_{j\neq i} W_j$ 中只有零元。

1⇒4:

若 $dim[w]\neq\sum_{i=1}^{m} dim[W_i]$，即 $\exists i$ s.t. $W_i\cap\sum_{j\neq i} W_j\neq\{0\}$

对 $0\neq w\in W_i\cap\sum_{j\neq i} W_j, 0=w+(-w)=0+0$，与直和之定义矛盾。 □

例 1. $\{\hat{e}_i\}_{i=1}^{n}$ 是空间 V 的一组基 (最大无关组)，则空间可以直和分解为

$$V=\bigoplus_{i=1}^{n} span\{\hat{e}_i\}$$

例 2. 对于有限维线性空间 V 上的变换 A，V 可以依据 A 的根子空间进行直和分解。

这一例子的证明见后。

1.2.2 空间间结构：线性映射

定义 3. 映射 $\phi[\ ]$ 是线性映射 $\{V,F\}\to\{W,F\}$，假若

$$\phi[\lambda v+\mu u]=\lambda\phi[v]+\mu\phi[u];\ u,v\in V;\lambda,\mu\in F$$

即所谓保线性。

1.3 线性映射

为了更深入的了解线性映射 $\phi[\ ]$，我们可以对映射的映像空间 W 和原像空间 V 进行直和分解，即引入一组基 $\epsilon_i{}_{i=1}^n$和$e_j{}_{j=1}^m$
设 $\phi[e_j]=\sum\limits_{i=1}^n a_{ij}\epsilon_i$
我们把 a_{ij} 排成一纵列，

$$\begin{bmatrix} a_{1j}\\ a_{2j}\\ \cdots\\ a_{nj}\end{bmatrix}$$

并称之为其在映像空间 W 中的矩阵表示。应当注意的是，表示并不等于向量本身，同一向量允许不同表示。
把原像空间 V 的基横着排列起来，有

$$\phi[\begin{bmatrix} e_1 & e_2 & \cdots & e_m\end{bmatrix}]=\begin{bmatrix} a_{11} & & a_{m1}\\ a_{21} & & a_{m2}\\ \cdots & \cdots & \cdots\\ a_{n1} & & a_{mn}\end{bmatrix}$$

我们就得到了线性映射的矩阵表示。进一步地，对 $\forall x\in V,x=\sum\limits_{j=1}^m x_je_j$
由于变换保线性，映像之表示应是以上矩阵列向量之线性组合。由矩阵乘法定义，知道**矩阵右乘列向量，即将矩阵之列向量按右乘之列向量线性组合** (对左乘行向量亦是如此，这一点很重要，十分有助于思考)。
从而，映像之表示即

$$\phi[\begin{bmatrix} x_j\\ x_j\\ \cdots\\ x_j\end{bmatrix}]=\begin{bmatrix} a_{11} & & a_{m1}\\ a_{21} & & a_{m2}\\ \cdots & \cdots & \cdots\\ a_{n1} & & a_{mn}\end{bmatrix}\begin{bmatrix} x_j\\ x_j\\ \cdots\\ x_j\end{bmatrix}$$

特别的，我们喜欢这样的线性映射 $\psi : \{V, F\} \to \{V, F\}$，称之为线性变换
其映射到自身，在 V 的基下表示为一个方阵，此时，我们可以对空间进行更精细的分解。

2 线性变换

"国庆之后，好日子就到头了。"——李思

2.1 同构与同态

2.1.1 同构

类比中学所学过的群论知识，可以定义同构：

定义 4. $\phi[\ ] : \{V_1, F\} \to \{V_2, F\}$是同构映射，假若$\phi$是双射，此时称$V_1, V_2$同构
特别地，若 $V_1 = V_2$，称为自同构。

命题 2. 同构映射以可逆矩阵表示。

证明.

引理 1. 同构映射保持线性独立。

证明. 用反证法。设 $\{a_i\}_{i=1}^m \in V$ 是一组独立的元素，$\phi[\ ]$ 是同构映射。
假若 $\exists\{\lambda_i\}_{i=1}^m \in F$ 不全为零，s.t.

$$\sum_{i=1}^{m} \lambda_i \phi[a_i] = 0$$

由于保加法，$\phi[\sum\limits_{i=1}^{m} \lambda_i a_i] = 0$
而显然 $\phi[0] = 0$ 是唯一的。(保持加法唯一零元不变)

$$\therefore \sum_{i=1}^{m} \lambda_i a_i = 0$$

矛盾!

□

从而同构将一最大无关组映射到另一最大无关组。用线性独立的列向量表示两组独立元素，设空间的维数为 n，有

$$B = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{bmatrix} = T \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} = TA$$

其中，T 是矩阵表示。由秩不等式 (越乘越小) 和 AB 满秩

$$n \geq \text{rank}[A](\text{矩阵维数限制}), n = \text{rank}[B] = \text{rank}[TA] \leq \text{rank}[T]$$

$$\therefore \text{rank}[T] = n$$

从而可逆。 □

从而，我们得到有唯一解的线性方程组理论，寻找已知同构映射的原像的理论。
同构，顾名思义，具有完全相同的结构，是一个很高的要求。我们希望寻找弱一些的联系，并给出一般方程组的理论。同样类比中学群论知识，我们定义同态。

2.1.2 同态

定义 5. 线性映射$\phi[\ ] : \{V_1, F\} \rightarrow \{V_2, F\}$是同态映射，假若$\phi$是满射, 此时称$V_1, V_2$同态

一个直观的图 (请自动忽略图名)

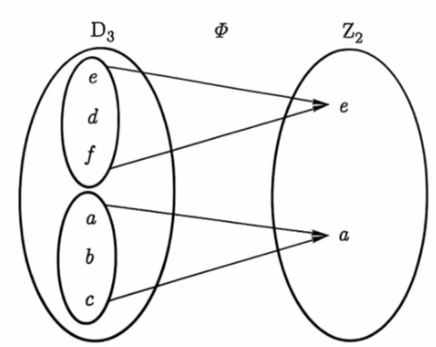

图 1.4 D_3 **群与二阶循环群的同态关系**

完全地类比群论，定义同态核

定义 6. V 之子集 $ker[\phi]$ 是同构映射 T 之同态核，假若 $T[x] = 0, \forall x \in ker[\phi]$

不难看出，同态核其实就是齐次方程的通解集合。类似陪集结构，同态核加上一个非核内元素，构成非齐次方程的一个解集。
以上，我们简要地介绍了思想，下面我们试图以之理解线性代数，使之更富逻辑，更加简明，以延续我们的好日子。

2.1.3 习题选讲

问题 1. *(秩不等式)* 试证：

$$rank[AB] \geq rank[A] + rank[B] - n$$

$$rank[ABC] \geq rank[AB] + rank[BC] - rank[B]$$

其中，A 是 $m \times n$ 矩阵，B 是 $n \times p$ 矩阵,C 是 $p \times q$ 矩阵

证明.

引理 2. *(维数定理)* 若线性空间V上有同态映射$\phi[\,]:\{V,F\} \to \{Im[\phi],F\}$且其同态核为$ker[\phi]$,映像空间为$Im[\phi]$,则$dim[V] = dim[ker[\phi]] + dim[Im[\phi]]$;

证明. 记 $dim[ker[\phi]] = m, dim[Im[\phi]] = l, dim[V] = n$
取同态核基 $\{a_i\}_{i=1}^{m}$，将其补充为 V 的基 $\{a_i\}_{i=1}^{n}$
从而 $\forall x \in V, x = \sum\limits_{i=1}^{n} \lambda_i a_i$，$\phi[x] = \phi[\sum\limits_{i=m+1}^{n} \lambda_i a_i] = \sum\limits_{i=m+1}^{n} \lambda_i \phi[a_i]$
假若 $\exists \{\mu_i\}_{i=m+1}^{n}$ 不全为零 s.t. $\sum\limits_{i=m+1}^{n} \mu_i \phi[a_i] = 0$
即 $\phi[\sum\limits_{i=m+1}^{n} \mu_i a_i] = 0$ 从而 $\sum\limits_{i=m+1}^{n} \mu_i a_i \in ker[\phi]$
这与基的定义矛盾！
$\therefore \{\phi[a_i]\}_{i=m+1}^{n}$ 线性独立。由 $\forall x \in V, \phi[x] = \sum\limits_{i=m+1}^{n} \lambda_i \phi[a_i]$ 知道 $\{\phi[a_i]\}_{i=m+1}^{n}$ 是映像空间的一组基。
即 $l = n - m$，即 $dim[V] = dim[ker[\phi]] + dim[Im[\phi]]$

□

先证第一个不等式，考虑 AB 是 A 分别作用于 B 之列向量上所构成同态映射 $\phi_{\{B\} \to \{AB\}}$ 的映像空间

由维数定理：

$$dim[Im[A[B]]] = \text{rank}[AB] = dim[B\text{之列向量组}] - dim[ker[\phi_{\{B\}\to\{AB\}}]] = \text{rank}[B] - dim[ker[\phi_{\{B\}\to\{AB\}}]]$$

而由线性方程组理论 $dim[ker[\phi_{\{B\}\to\{AB\}}]] \leq min\{m,n\} - \text{rank}[A]$
代入得 $\text{rank}[AB] \geq \text{rank}[B] + \text{rank}[A] - min\{m,n\} \geq \text{rank}[B] + \text{rank}[A] - n$ 即证一式
再证二式：
同理，由维数定理有

$$\text{rank}[ABC] = \text{rank}[BC] - dim[ker[\phi_{\{BC\}\to\{ABC\}}]]$$

以及

$$\text{rank}[AB] = \text{rank}[B] - dim[ker[\phi_{B\to AB}]]$$

从另一角度看，对 B 左乘 C，实则是对 B 中的列向量线性组合，组合得到的 BC 中的列向量所张成的空间一定小于 B 中列向量所张成的。
从而

$$dim[ker[\phi_{B\to AB}]] \geq dim[ker[\phi_{\{BC\}\to\{ABC\}}]]$$

$$\therefore \text{rank}[ABC] \geq \text{rank}[BC] - dim[ker[\phi_{B\to AB}]] = \text{rank}[BC] + \text{rank}[AB] - \text{rank}[B]$$

□

2.2 根子空间

“我们所要做的其实是一种简单的量子场论。”——鲜于中之

2.2.1 准对角化

我们对于变换 (方阵) 的对角化是十分熟悉的，此时找到了一个特殊的对变换封闭的一维空间，其上的元素做变换，映像只是原像的简单数乘。然而，这样好的日子并不持续，我们知道存在不可对角化的矩阵。由此自然地想到，尽管一维不变空间不存在，可否有更高维数的不变空间？尽管简单数乘不存在，可否有简单矩阵的左乘替代之？
不难看出，如果存在这样的高维不变子空间，若在各空间之基所**直和**成的基下，方阵可以“准对角化”，或者说分块对角化。

(如果看不出，可以这样看)
对矩阵 A所表示的变换，记各子空间为$\{W_i\}_{i=1}^m$, 其基为$\{e_i^{(s)}\}_{s=1}^{k_i}$, 从而有
$Ax=\sum_{i=1}^m\sum_{s=1}^{k_i}x_i^{(s)}Ae_i^{(s)}=\sum_{i=1}^m\sum_{s=1}^{k_i}x_i^{(s)}[\sum_{l=1}^{k_i}a_i^{(sl)}e_i^{(l)}]=\sum_{i=1}^m\sum_{l=1}^{k_i}e_i^{(l)}[\sum_{s=1}^{k_i}x_i^{(s)}a_i^{(sl)}]$
写成这个基础下的矩阵形式：

$$A\begin{bmatrix}x_1^{(1)}\\ \cdots\\ x_1^{(k_1)}\\ x_2^{(1)}\\ \cdots\\ x_2^{(k_2)}\\ \cdots\end{bmatrix}=\begin{bmatrix}a_1^{(11)} & \cdots & a_1^{(k_11)} & & & & \cdots\\ \cdots & & \cdots & & & & \\ a_1^{(1k_1)} & \cdots & a_1^{(k_1k_1)} & & & & \\ & & & a_1^{(11)} & \cdots & a_1^{(k_11)} & \\ & & & \cdots & & \cdots & \\ & & & a_1^{(1k_1)} & \cdots & a_1^{(k_1k_1)} & \\ \cdots & & & & & & \end{bmatrix}\begin{bmatrix}x_1^{(1)}\\ \cdots\\ x_1^{(k_1)}\\ x_2^{(1)}\\ \cdots\\ x_2^{(k_2)}\\ \cdots\end{bmatrix}$$

现在你可以很明显地看出来了。
进一步地，我们想问，对角块有应当具有怎样的形式？

2.2.2 零化子与零化核

虽然我们已经知道答案是一个称为约当标准型的上三角形式 (有些书里写的是下三角形式，但这是无关紧要的)。但是为了更自然地引出，我们需要耐心一点。
首先来看我们熟知的知识，考虑方阵 A 的对角化:

$$(A-\lambda_i I)\vec{\lambda_i}=0$$

依据我们之前的理论，$\Lambda_i=span\vec{\lambda_i}$ 是矩阵 $(A-\lambda_i I)$ 所代表同态变换的同态核。然而，此时同态的结构不重要，我们更愿意称之为零化核。对应的，多项式 $\psi[A]=(A-\lambda_i I)$ 称为一个零化子
(这个是我乱定义的，就不正式地写为定义了)
例如，由凯莱-哈密顿定理，若多项式 $\psi[X]=\det[-A+XI]$，则 $\psi[A]$ 就是一个以全空间为零化核的零化子。
由凯莱-哈密顿定理，知道每一个方阵都存在全空间零化子。定义一个最小全空间零化子，其以全空间为零化核，且多项式形式的次数最低，记之为 $m[A]$。

因式分解有

$$m[A] = \prod_{i=1}^{n} (A - \mu_i I)^{t_i}$$

对于可对角化的矩阵，知道其具有独立的本征矢量，从而，若想要以全空间为零化核，就得把所有本征方程中的 $(A - \lambda_i I)$(所有相异的本征值) 含进去，这样一来，就知道了

命题 3. *$m[A] = \psi[A](\psi[X] = det(-A + XI))$，假若 A 可对角化，且具有两两不等的本征值。*

对于不可对角化的矢量，由于所有本征矢量远非完备的基，但 $\psi[A]$ 足以零化整个空间，这是我们所希望的。
我们意识到，之所以不完备，是因为存在这样的矢量 x，其虽不满足 $(A - XI)x = 0$，但是可以在更高次因子 $(A - XI)^n$ 的情况况下零化 $(A - XI)^n x = 0$
为了使基完备，我们引入如下矢量:

$$\vec{s}_{ij}^{(k)} \in V \ \ \text{s.t.} (A - \lambda_i I)^k \vec{s}_{ij}^{(k)} = 0 \text{且} (A - \lambda_i I)^{k-1} \vec{s}_{ij}^{(k)} \neq 0$$

当 $k = 1$ 时，即为本征矢量。
(对有限维的 A，这样的 $\vec{s}$ 总是存在的。这不难证明，只需证至少存在一个本征矢量即可，而方程 $f[\lambda] = \det[A - \lambda I] = 0$ 是一个有限次多项式方程，其在 $\mathbb{C}$ 上总是有解的。)
下面我们来构造这样的完备基。

2.2.3 约当标准型

一个自然的想法是以阶数为标准进行直和分解。我们将会看到，这是不行的。
首先，我们取“零粒子空间”，即由 $\vec{s}_{ij}^{(1)}$ 构成的本征矢量空间。

命题 4. *非简并 (相异) 本征值的本征矢独立。*

证明. 假设 $\vec{s}_{11}^{(1)}$ 和 $\vec{s}_{21}^{(1)}$ 分别是本征值 $\lambda_1 \neq \lambda_2$ 的本征矢量。
设 $\exists \mu_1, \mu_2$ 不全为零 s.t. $\mu_1 \vec{s}_{11}^{(1)} + \mu_2 \vec{s}_{21}^{(1)} = 0$

$$A[\mu_1 \vec{s}_{11}^{(1)} + \mu_2 \vec{s}_{21}^{(1)}] = \mu_1 \lambda_1 \vec{s}_{11}^{(1)} + \mu_2 \lambda_2 \vec{s}_{21}^{(1)} = \mu_1 (\lambda_1 - \lambda_2) \vec{s}_{11}^{(1)} = 0$$

考虑 $\vec{s}_{11}^{(1)} \neq 0, \lambda_1 - \lambda_2 \neq 0$

$$\therefore \mu_1 = 0 = \mu_2$$

矛盾！ □

对于简并的本征值，也可以有独立的本征矢，这样的一切矢量 $\{\vec{s}_{ij}^{(1)}\}$ 构成了“零粒子空间”$W^{(1)}$ 的基，其个数比全空间的基数小。
我们考虑“一粒子空间”$W^{(2)} = span[\vec{s}_{ij}^{(2)}]$
其中 $(A-\lambda_i)^2\vec{s}_{ij}^{(2)} = 0$，不难注意 $(A-\lambda_i)[(A-\lambda_i)\vec{s}_{ij}^{(2)}] = 0$, 即 $(A-\lambda_i)\vec{s}_{ij}^{(2)} \in W_1 \star$
我们将其中的独立矢量挑出来，组成一粒子空间 $W^{(2)}$ 的一组基 $\{\vec{s}_{ij}^{(2)}\}$。
$\cdots$ 以此类推，得到各个空间。
我们想要证明以下的命题

命题 5.

$$W^{(m)} + W^{(n)} = W^{(m)} \oplus W^{(n)}, for\ n \neq m$$

一个证明的尝试展示如下

证明. 不妨设 $m \geq n$，由式 $\star$, $\forall$ 基 $\vec{s}_{ik}^{(m)} \in W^{(m)}, (A-\lambda_i I)^{(m-n)}\vec{s}_{ik}^{(m)} \in W^{(n)}$
设 $\{\nu_{ij}\}$ 不全为零，$(A-\lambda_i I)^{(m-n)}\vec{s}_{ik}^{(m)} = \sum\limits_{j=1}^{l_i^{(n)}} \nu_{ij}\vec{s}_{ij}^{(n)}$
只需证任一 $W^{(m)}$ 基不能被 $W^{(n)}$ 基线性表示，用反证法:
设 $\exists\{\mu_{ij}\}$ 不全为零，$\vec{s}_{ik}^{(m)} = \sum\limits_j \mu_{ij}\vec{s}_{ij}^{(n)}$
代入得 $(A-\lambda_i I)^{(m-n)} \sum\limits_{j=1}^{l_i^{(n)}} \mu_{ij}\vec{s}_{ij}^{(n)} = \sum\limits_{j=1}^{l_i^{(n)}} \nu_{ij}\vec{s}_{ij}^{(n)}$
设所研究线性空间的维数为 p，由列向量 $\{\vec{s}_{ij}^{(n)}\}_{j=1}^{l_i^{(n)}}$ 所排列成的是 $p \times l_i^{(n)}$ 的矩阵 S。由于各列向量独立，可以对其进行归一正交化，对最后的结果并无影响，不过是 μ_i 和 ν_i 重新组合了而已。假设 S 已经归一正交化好。有矩阵形式：

$$S^\dagger(A-\lambda_i I)^{(m-n)}S \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \\ \cdots \\ \mu_{il_i^{(n)}} \end{bmatrix} = \begin{bmatrix} \nu_{i1} \\ \nu_{i2} \\ \cdots \\ \nu_{il_i^{(n)}} \end{bmatrix}$$

由矩阵不等式，记 $(A-\lambda_i I)^{(m-n)} = C$
需要证明 $S^\dagger CS\vec{\mu} = \vec{\nu}$ 无解或没有唯一解，而 $\mathrm{rank}[C] < p, \mathrm{rank}[S^\dagger] = \mathrm{rank}[S] = l_i^{(n)}$

我们可能想证 $\text{rank}[S^\dagger CS] < l_i^{(n)}$ 这是一个很符合直觉的想法
很遗憾，**但这是错的**。我们很容易就能举出一个反例:

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

□

这种分解方法**远非直和，命题 5 是大错特错**，而且其不变性质也极差，考虑：

$$A\vec{s}_{ij}^{(m)} = (A - \lambda_i I)\vec{s}_{ij}^{(m)} + \lambda_i \vec{s}_{ij}^{(m)}$$

由式 $\star$，知道 $(A - \lambda_i I)\vec{s}_{ij}^{(m)} \in W^{(m-1)}$
$\therefore A\vec{s}_{ij}^{(m)} = \lambda_i \vec{s}_{ij}^{(m)} + \sum_{j=1}^{l_i^{(m-1)}} p_{ij}^{m(m-1)} \vec{s}_{ij}^{(m-1)} \star\star$
这远非不变直和分解，更无法将矩阵准对角化。不难想象，在这组“基”下的表示矩阵，会出现 $[p_{ij}^{m(m-1)}]$ 这样的非对角分块。
但这并非毫无价值，式 $\star\star$ 启示我们，“m 粒子空间”与“m-1 粒子空间”“联合”不变。假若我们给每一个 $\vec{s}_{ij}^{(m)}$ 都指定一组“直系亲属” $\{(A - \lambda_i I)^n \vec{s}_{ij}^{(m)} = \vec{s}_{ij}^{(m-n)}\}_{n=0}^{m-1}$
这样的矢量构成一个不变子空间 J_{ij}，且其满足：

$$A\vec{s}_{ij}^{(t)} = (A - \lambda_i I)\vec{s}_{ij}^{(t)} + \lambda_i A\vec{s}_{ij}^{(t)} = \vec{s}_{ij}^{(t-1)} + \lambda_i \vec{s}_{ij}^{(t)}$$

如果写成矩阵形式，即:

$$A\begin{bmatrix} \vec{s}_{ij}^{(1)} & \vec{s}_{ij}^{(2)} & \cdots & \vec{s}_{ij}^{(m)} \end{bmatrix} = \begin{bmatrix} \vec{s}_{ij}^{(1)} & \vec{s}_{ij}^{(2)} & \cdots & \vec{s}_{ij}^{(m)} \end{bmatrix} \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ & & & \cdots & \\ 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix}$$

从而自然地完成了循环分解、得出了约当标准型。
而考虑以全空间为零化核的零化子 $\psi[A] = \det[-A + XI]|_{X=A} = \prod_{i=1}^{n}(A - \lambda_i I)^{t_i}$
对其中的每一个因子 $(A - \mu_i I)^{t_i}$，可以首先找到 $\vec{s}_{ij}^{t_i}$，再依次左乘，得到整个 J_{ij}，并在这个基下，给出准对角化的 A 的一个对角块。

2.2.4 习题选讲：计算技巧

“证明结束后，应当撤走所有的梯子。”——高斯

首先做一个简单的计算，以熟悉我们刚刚所讲的东西。

问题 2. 试将:

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

相似变换为上三角阵。

解 . 本征方程

$$det[-A+\lambda I]=(\lambda-2)^2(\lambda-1)$$

零化子

$$\psi[A]=(A-2I)^2(A-1)$$

对 $(A-I)$，本征矢量 $\vec{s}_{11}^{(1)}=\begin{bmatrix}1\\0\\2\end{bmatrix}$

对 $(A-2I)^2=\begin{bmatrix}1&-1&0\\0&0&0\\2&-2&0\end{bmatrix},\vec{s}_{21}^{(2)}=\begin{bmatrix}1\\1\\0\end{bmatrix},\vec{s}_{21}^{(1)}=\begin{bmatrix}2\\2\\4\end{bmatrix}$

$$\therefore A\left[\vec{s}_{11}^{(1)}\vec{s}_{21}^{(1)}\vec{s}_{21}^{(2)}\right]=\left[\vec{s}_{11}^{(1)}\vec{s}_{21}^{(1)}\vec{s}_{21}^{(2)}\right]\begin{bmatrix}1&&\\&2&1\\&&2\end{bmatrix}$$

即

$$A=\frac{1}{4}\begin{bmatrix}1&2&1\\0&2&1\\2&4&0\end{bmatrix}\begin{bmatrix}1&&\\&2&1\\&&2\end{bmatrix}\begin{bmatrix}4&-4&0\\-2&2&1\\4&0&-2\end{bmatrix}$$

即为所求。

然后是一个作业题，虽然比较水，但是一个明确思路的好例子

问题 3. 试证: 约当标准型 J 与之转置相似。

这个例子可以用于证明任一方阵与之转置相似。同时，这也导致部分参考书的约当块是下三角形式。

分析 . 约当标准型的意义用下式 *(列向量的线性组合)* 说明:

$$[s_{ij}^{(n)}]J_{ij} = [\lambda_i s_{ij}^{(n)} + s_{ij}^{(n-1)}]$$

这是从左到右按照 n 从小到大排列的。如果将 n 从大到小排列，而保持以上线性组合关系，我们就会发现，J_{ij} 应当转置。
改变排列顺序的矩阵即是反对角矩阵。从而我们写下证明。

证明. 注意到

$$K = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ & & \cdots & & \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} s.t. K^2 = I, \text{即} K = K^{-1}$$

且其满足

$$KJK^{-1} = J^T$$

从而相似。 □

以上只是机械地计算。下面我们看两个巧算的例子;

问题 4. 求矩阵

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

之约当标准形。

分析 . 本题看似计算量较大，但注意到并不要求求出变换矩阵，于是我们通过观察，可以直接地写出原矩阵的约当标准型。
易知特征多项式 $\phi[\lambda] = [\lambda - (-1)][\lambda - 2]^5$
从而 A 有两个本征值。对应*-1* 本征值的本征矢量只有一个，给出约当标准型

$$J_{(-1)} = \begin{bmatrix} 1 \end{bmatrix}$$

而对于本征值 *2*，考虑 $B=A-2I=\begin{bmatrix}0&0&0&0&0&0\\1&0&0&0&0&0\\-1&0&0&0&0&0\\0&1&0&0&0&0\\1&1&1&1&0&0\\0&0&0&0&1&-3\end{bmatrix}$

观察行向量的线性相关性可以知道 $rank[B]=4$，从而该方程有 *2* 个独立的本征矢量，我们得到“零粒子空间” $dim[W^{(1)}]=2$

根子空间还剩下 *3* 个基向量待决定。

对于“一粒子空间” $W^{(2)}$，其至多有两个基向量，因为其必须要能“生成”一粒子空间。B^2 是不得不算的。

$$B^2=\begin{bmatrix}0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\1&0&0&0&0&0\\0&1&0&0&0&0\\1&1&1&1&-3&9\end{bmatrix}$$

同样，$rank[B^2]=3$，有三个独立的解，而其中两个是零粒子空间的，一粒子空间只有一个，即 $dim[W^{(2)}]=1$

同样的理由，剩下两个矢量分别包含于二粒子空间和三粒子空间，$dim[W^{(3)}]=dim[W^{(4)}]=1$

于是，我们清楚，有两个小约当块，一个对应一阶的循环空间，一个对应四阶的循环空间。约当块即*:*

$$J_{(2),1}=\begin{bmatrix}2\end{bmatrix}$$

$$J_{(2),2}=\begin{bmatrix}2&1&&\\&2&1&\\&&2&1\\&&&2\end{bmatrix}$$

而所求的约当标准型即 $diag\{J_{(-1)},J_{(2),1},J_{(2),2}\}$

而显然，以上的过程只是我们的分析，“n 粒子空间”不显于讲义的任意角落，明显不能书写。于是我们决定向高斯学习。

解．经计算，

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

即为所求约当标准型。

运用上面的经验，我们继续看这样一个直接计算更加困难的问题。

问题 5. 若 *14* 阶矩阵

$$N=\begin{bmatrix}0&1&&\\&0&1&\\&&\cdots&\\&&&0\end{bmatrix}$$

求 N^4 的约当标准型及根子空间。

分析．N 的本征值自然全是零，而 N 的乘方性质是我们所熟悉的，即乘方一次，对角线平移一格。N^4 则平移四格，以第一行为例，前四列都是 *0*，第五列为 *1*。
容易知道，$S_{11}^{(1)}=\begin{bmatrix}1&0&0&0&\cdots&0\end{bmatrix}^T$，$S_{21}^{(1)}=\begin{bmatrix}0&1&0&0&\cdots&0\end{bmatrix}^T$，$S_{31}^{(1)}=\begin{bmatrix}0&0&1&0&\cdots&0\end{bmatrix}^T$，
$S_{41}^{(1)}=\begin{bmatrix}0&0&0&1&\cdots&0\end{bmatrix}^T$ 是本征矢量，构成零粒子空间。
考察一粒子空间，有 $(N^4)^2=N^8$，前八列都是 *0*，第九列为 *1*。排除掉零粒子空间，有四个独立向量，分别在第 *5*、*6*、*7*、*8* 行为 *1*，分别记之 $S_{11}^{(2)},S_{21}^{(2)},S_{31}^{(2)},S_{41}^{(2)}$。
不难看出，$(N^4)S_{k1}^{(2)}=S_{k1}^{(1)}$，$k=1,2,3,4$
同理，我们可以写出 $S_{11}^{(3)},S_{21}^{(3)},S_{31}^{(3)},S_{41}^{(3)}$ 仍然满足上述生成关系，但对于三粒子空间，我们只剩两个独立的向量 *(*$14=3\times4\cdots2$*)*
其应该出现在 $S_{11}^{(4)},S_{21}^{(4)}$ 这两个最靠前的循环子空间，其分别在第 *13* 行和第 *14* 行为 *1*。
于是我们得到了答案。

解．注意到：

$$N^4=\begin{bmatrix}1&0&0&0&0&0&0&0&0&0&0&0&0&0\\0&0&0&0&1&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&1&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0&0&1&0&0\\0&1&0&0&0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&1&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&1&0&0&0&0\\0&0&0&0&0&0&0&0&0&0&0&0&1&0\\0&0&1&0&0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&1&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0&1&0&0&0\\0&0&0&0&0&0&0&0&0&0&0&0&0&1\\0&0&0&1&0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&1&0&0&0&0&0&0\end{bmatrix}$$

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

即为所求。

2.3 习题选讲：DN 分解

(这个名词是我瞎造的，不要引用)

由之前的作业，我们知道，对任意 n 阶方阵 A，$A=D+N$，且 D，N 可交换，即一个变换

是由其可对角化部分 D 和幂零部分 N 组成。

问题 6. 若 A 是 n 维复线性空间 V 上的一个线性变换，D 是 A 的可对角化部分，试证，对多项式 $g[A]$，其可对角化部分为 $g[D]$

证明. 设 $g[A] = \sum\limits_{k=0}^{h} g_k A^k$，采用我们之前使用的基 $\{s_{ij}^{(n)}\}$，在此基下，A 表示为一个约当标准型有 $As_{ij}^{(n)} = \lambda_i s_{ij}^{(n)} + s_{ij}^{(n-1)}$

引理 3. $A^k s_{ij}^{(n)} = \sum\limits_{m=0}^{k} C_k^m \lambda_i^m s_{ij}^{(n-k+m)}$

证明. 用归纳法证明之：$A^1 s_{ij}^{(n)} = \lambda_i s_{ij}^{(n)} + s_{ij}^{(n-1)}$
若 $A^{(k-1)} s_{ij}^{(n)} = \sum\limits_{m=0}^{k-1} C_{k-1}^m \lambda_i^m s_{ij}^{(n-k+1+m)}$
则 $AA^{(k-1)} s_{ij}^{(n)} = \sum\limits_{m=0}^{k-1} C_{k-1}^m \lambda_i^{m+1} s_{ij}^{(n-k+1+m)} + \sum\limits_{m=0}^{k-1} C_{k-1}^m \lambda_i^m s_{ij}^{(n-k+m)} = \sum\limits_{m=1}^{k-1} (C_{k-1}^m + C_{k-1}^{m-1}) \lambda_i^m s_{ij}^{(n-k+m)} + s_{ij}^{n-k} + s_{ij}^{n} = \sum\limits_{m=0}^{k} C_k^m \lambda_i^m s_{ij}^{(n-k+m)}$
从而证得断言。

□

将 A 分解为可对角化/幂零:

$$As_{ij}^{(n)} = (D+N)s_{ij}^{(n)} = \lambda_i s_{ij}^{(n)} + s_{ij}^{(n-1)}$$

即

$$Ds_{ij}^{(n)} = \lambda_i s_{ij}^{(n)}$$

同理，由引理

$$A^k s_{ij}^{(n)} = \sum_{m=0}^{k} C_k^m \lambda_i^m s_{ij}^{(n-k+m)} = \lambda_i^k s_{ij}^{(n)} + \sum_{m=0}^{k-1} C_k^m \lambda_i^m s_{ij}^{(n-k+m)}$$

$$\therefore (\text{g[A] 的可对角部分}) D[g[A]] s_{ij}^{(n)} = \sum_{k=0}^{h} g_k \lambda_i^k s_{ij}^{(n)}$$

而 $Ds_{ij}^{(n)}=\lambda_i s_{ij}^{(n)}, g[D]s_{ij}^{(n)}=\sum\limits_{k=0}^{h} g_k\lambda_i^k s_{ij}^{(n)}$

$$\therefore D[g[A]]=g[D]$$

□

现在我们已经基本清楚如何使用这样的分解。下面做一个简单的练习：

问题 7. 若 n 阶实方阵 A 满足 $A^2+I=0$，试证明:$A\sim\begin{bmatrix}0_{\frac{n}{2}\times\frac{n}{2}} & -I_{\frac{n}{2}\times\frac{n}{2}}\\ I_{\frac{n}{2}\times\frac{n}{2}} & 0_{\frac{n}{2}\times\frac{n}{2}}\end{bmatrix}$;

证明. 首先证 n 是偶数。$\det[A^2]=[\det[A]]^2=\det[-I]=(-1)^n$，从而 n 为偶数。
设 A 的约当标准型为：
$\exists P$ 可逆，s.t. $PAP^{-1}=D+N$，其中，D 是对角的，N 是 m+1 阶幂零的，且 $m+1\leq n$
考虑 D,N 可交换，$A^2=(D+N)^2=D^2+2DN+N^2=-I$，其中只有 D^2 对对角元有贡献，从而 $D=diag[\pm i]$
因为 A 是实矩阵，$\mathrm{tr}[A]=\mathrm{tr}[D]\in\mathbb{R}$，故 D 之对角元正负参半。不妨正值全部排列于左上，负值全部排列于右下。容易知道 $D^{-1}=-D$
$\because -AA=I\ \therefore A^{-1}=-A=-D-N$
而注意到 $(D+N)[(-D)(I+DN+D^2N^2+\cdots+D^mN^m)]=I$
(事实上，这一步是形式化地读写 $\frac{1}{D+N}$ 再做“级数展开”得到)
$\therefore [(-D)(I+DN+D^2N^2+\cdots+D^mN^m)]=-D-N$
即

$$N+DN^2-N^3-DN^4+N^5+\cdots+(-D^{m+1})N^m=-N$$

两侧同乘以 N^{m-1}，有：

$$N^m=-N^m$$

$\therefore N^m=0$
带入上上式，有

$$N+DN^2-N^3-DN^4+N^5+\cdots+(-D^m)N^{m-1}=-N$$

两侧同乘以 N^{m-2}，有：

$$N^{m-1}=-N^{m-1}$$

$\therefore N^{m-1}=0$
依此类推，得到 $N=0$

$\therefore$A 可对角化，且 $A \sim diag[+i,+i,\cdots,+i,-i,-i,\cdots,-i] = \Lambda$
而对于 $B = \begin{bmatrix} 0_{\frac{n}{2}\times\frac{n}{2}} & -I_{\frac{n}{2}\times\frac{n}{2}} \\ I_{\frac{n}{2}\times\frac{n}{2}} & 0_{\frac{n}{2}\times\frac{n}{2}} \end{bmatrix}$ 带入本征方程 $\det[B-\lambda_i I]=0$(这个的快速算法就是直接对中心的二阶行列式展开计算)
解得 $\lambda_i = \pm i$，带入本征方程，可以求得 n 个独立的本征矢量，从而 B 可对角化。即

$$B \sim \Lambda \sim A$$

□

然后以此复习一下线性方程组理论，用一个较简单的水题看看它在坏日子里面有什么变化

问题 8. *A 是有限维线性空间 V 上的一个变换，且 $rank[A]=1$，试证：A 要么可以对角化，要么是幂零的。*

证明. 设 $dim[V]=n$，由线性方程组理论，$dim[ker[\phi_A]] = n - \mathrm{rank}[A] = n-1$
设核基为 $\{s^i\}_{i=1}^{n-1}$，s^n 为之余下的一个基

$$\therefore As^i = 0, for 1 \le i \le n-1$$

可设

$$As^n = \lambda s^n + \sum_{i=1}^{n-1} \mu_i s^i$$

1. 若 $\lambda \neq 0$
则

$$A[s^n + \sum_{i=1}^{n-1} \frac{\mu_i}{\lambda} s^i] = \lambda[s^n + \sum_{i=1}^{n-1} \frac{\mu_i}{\lambda} s^i]$$

令 $s^n + \sum\limits_{i=1}^{n-1} \frac{\mu_i}{\lambda} s^i$ 为一个新基，与核基一起构成完备的基，在该基下，A 被对角化。
2. 若 $\lambda = 0$
即 $As^n = \sum\limits_{i=1}^{n-1} \mu_i s^i$，则 $A^2 s^i = 0, for 1 \le i \le n$
即为幂零的。 □

最来看一个无穷维的简单例子 (以防止被期末偷袭)

问题 9. *若 V 是无限可微实函数所组成之线性空间，导子 $\mathcal{D}$ 是 V 上一个线性变换。求之本征值、本征矢及根子空间。*

解． 本征方程 $(\mathcal{D}-\lambda)f(x)=0$
解得 $f(x)=\mathbb{R}e^{\lambda x}$，即为所求对应本征值 λ 的本征矢量所在空间。
根子方程 $(\mathcal{D}-\lambda)^n f(x)=0$
由微分方程理论，根子空间为 $e^{\lambda x}\mathbb{R}[x]$，$\mathbb{R}[x]$ 是 x 的实系数多项式。
即为所求。

3 内积空间

“长度是什么，我们还没有定义好。”——艾颖华

3.1 幺正变换与正定性

笔者并没有太多超出李思老师讲义的内容，故不作详述，只将重点罗列如下，并讨论一些解题技巧。
(1) 幺正变换的定义与性质，厄密矩阵的定义与性质，正定的定义与性质 (正定的定义基于实对称，这一点值得注意)；
(2)SVD，QR 分解，极分解，广义逆与 MP 逆。
(3) 最小二乘法、格拉姆矩阵。

3.2 习题选讲：对称构造

首先看一个作业题；

问题 10. 试证明两正定矩阵 A,B 之积 AB 仍正定。

分析． 本题看似简单，实则并不显然。主要的原因就是 AB 缺乏对称性，使得许多工作难以开展。同时知道，相似变换不改变正定性，我们希望构造一个对称性较好的、同时与 AB 相似的矩阵。首先想到的是 $AB+BA$，但对于多项式，其很难与一个单项式相似。从而，我们转向 ABA，但是由于矩阵次数的冲突 *(一个是“2 次”，一个是“3 次”)*，并不相似。于是我们对称地调整次数，终于得到 $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ 这一符合要求的矩阵。

证明. 由于 A 是实对称正定矩阵
$\therefore \exists P$ 正交 $s.t. A=Pdiag(\lambda_i)P^T, \lambda_i\in\mathbb{R}^+$
令 $A^{-\frac{1}{2}}=Pdiag(\frac{1}{\sqrt{\lambda_i}})P^T, A^{\frac{1}{2}}=Pdiag(\sqrt{\lambda_i})P^T$;
$A^{-\frac{1}{2}}$ 也是实对称正定矩阵, 因为其本征值 $\frac{1}{\sqrt{\lambda_i}}>0$ 且 $(A^{-\frac{1}{2}})^T=(Pdiag(\frac{1}{\sqrt{\lambda_i}})P^T)^T=Pdiag(\frac{1}{\sqrt{\lambda_i}})P^T=A^{-\frac{1}{2}}$

同理可知,$A^{\frac{1}{2}}$ 也是实对称正定矩阵;
又 $A^{-\frac{1}{2}}A^{\frac{1}{2}}=1$ 即 $A^{-\frac{1}{2}}=(A^{\frac{1}{2}})^{-1}$
且有 $AA^{-\frac{1}{2}}=A^{\frac{1}{2}}$
注意到:

$$M=A^{-\frac{1}{2}}ABA^{\frac{1}{2}}=Pdiag(\frac{1}{\sqrt{\lambda_i}})P^TPdiag(\lambda_i)P^TBPdiag(\sqrt{\lambda_i})P^T=A^{\frac{1}{2}}BA^{\frac{1}{2}}$$

$$\therefore M^T=(A^{\frac{1}{2}}BA^{\frac{1}{2}})^T=(A^{\frac{1}{2}}BA^{\frac{1}{2}})=M$$

考虑二次型, 由于 B 是正定的:

$$\langle x|\,M\,|x\rangle=\langle x|\,(A^{\frac{1}{2}}BA^{\frac{1}{2}})\,|x\rangle=\langle y|\,B\,|y\rangle>0$$

其中 $|y\rangle=A^{\frac{1}{2}}\,|x\rangle\,,\langle y|=\langle y|\,{A^{\frac{1}{2}}}^T=\langle y|\,A^{\frac{1}{2}}$
因此,M 也是实对称正定矩阵, 其具有全是正实数的本征值
而 $M=A^{-\frac{1}{2}}ABA^{\frac{1}{2}}\sim AB$
相似的矩阵具有相同的本征值, 从而 AB 具有有全是正实数的本征值.

□

一个扩展练习如下：

问题 11. 若 S 和 T 是 n 阶正定实对称方阵，P 为 n 阶实方阵，试证明：

$$S-P^TT^{-1}P\text{与}T-PS^{-1}P^T$$

有相同的正定性，即同为正定或非正定。

分析 . 我们看到，这是对 S、T 对称的结论，但具体的表达式却是不对称的。我们用以下的构造，把 S，T 放到一个扩大矩阵的对角线上，而把这两个表达式视作分块对角化的结果。事实上，这样的构造虽然说不出道理，但却是常见的，若有余力应当记住这样的构造。

证明. 注意到：

$$\begin{bmatrix} I & -T^{-1}P \\ 0 & I \end{bmatrix}^T \begin{bmatrix} T & P \\ P^T & S \end{bmatrix}^T \begin{bmatrix} I & -T^{-1}P \\ 0 & I \end{bmatrix}=\begin{bmatrix} T & 0 \\ 0 & S-P^TT^{-1}P \end{bmatrix}$$

$$\begin{bmatrix} I & 0 \\ -S^{-1}P^T & I \end{bmatrix}^T \begin{bmatrix} T & P \\ P^T & S \end{bmatrix}^T \begin{bmatrix} I & 0 \\ -S^{-1}P^T & I \end{bmatrix} = \begin{bmatrix} T - PS^{-1}P^T & 0 \\ 0 & S \end{bmatrix}$$

证讫。 □

为了熟悉这样的构造，有一个这样的练习：

问题 12. 若 S 是 n 阶半正定实对称方阵，且

$$S = \begin{bmatrix} S^1_{r\times r} & S^2_{r\times(n-r)} \\ {S^2_{r\times(n-r)}}^T & S^3_{(n-r)\times(n-r)} \end{bmatrix}$$

试证明：

$$det[S] \leq det[S^1] \cdot det[S^3]$$

提示：先证明若 S 和 T 都是半正定是对称方阵，则 $\det[S+T] \geq \det[S]$

3.3 习题选讲：几何观点

问题 13. 对任一 n 阶实对称矩阵 A 和 B，记 A 的最大本征值为 $\Lambda[A]$，最小本征值为 $\lambda[A]$。试证：$\forall\alpha \in [0,1]$，

$$\alpha\Lambda[A] + (1-\alpha)\Lambda[B] \geq \Lambda[\alpha A + (1-\alpha)B],$$
$$\alpha\lambda[A] + (1-\alpha)\lambda[B] \leq \lambda[\alpha A + (1-\alpha)B]$$

分析 . 本题直接从代数上求解十分困难，因为其涉及组合矩阵的最大特征值，最大特征值的出现位置难以确定。乍看无从下手，但若我们想起可对角化矩阵的几何意义：伸缩变换；以及特征值的几何意义：主轴伸缩比例；从而得出下面的引理，这个问题就迎刃而解了。类似地，SVD 的几何意义、行列式的几何意义都应清楚，这有时很有助于解题。

引理 4.

$$\Lambda[A] = \max_{\forall x\in V}[\frac{x^TAx}{\|x\|^2}], \lambda[A] = \min_{\forall x\in V}\forall x \in V[\frac{x^TAx}{\|x\|^2}]$$

证明. 由于 A 是实对称矩阵，其正交相似于对角阵，故

$$\frac{x^TAx}{\|x\|^2} = \frac{1}{[x^TP^T][Px]}[x^TP^T]\begin{bmatrix} \Lambda & & & & \\ & \lambda & & & \\ & & \lambda_3 & & \\ & & & \dots & \\ & & & & \lambda_n \end{bmatrix}[Px]$$

令 $Px = y = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}$

有

$$\frac{x^T Ax}{\|x\|^2} = \frac{\Lambda y_1^2 + \lambda y_2^2 + \sum\limits_{i=3}^{n} \lambda_i^2 y_i^2}{\sum\limits_{i=1}^{n} y_i^2} \in [\Lambda, \lambda]$$

□

证明. 由引理

$$\Lambda[\alpha A + (1-\alpha)B] = \underset{\forall x \in V}{max}[\frac{x^T(\alpha A + (1-\alpha)B)x}{\|x\|^2}] = \underset{\forall x \in V}{max}[\alpha[\frac{x^T Ax}{\|x\|^2}] + (1-\alpha)[\frac{x^T Bx}{\|x\|^2}]]$$

$$\leq \alpha \underset{\forall x \in V}{max}[[\frac{x^T Ax}{\|x\|^2}]] + (1-\alpha)\underset{\forall x \in V}{max}[[\frac{x^T Bx}{\|x\|^2}]] = \alpha\Lambda[A] + (1-\alpha)\Lambda[B]$$

□

4 其它习题选讲

“科学就是物理学和集邮。”——费曼

问题 14. *(实虚部)*

设 λ_0 是 n 阶实正交方阵 A 的一个特征值 *(可能是复数)*。

(1) 试证 $|\lambda_0| = 1$

(2) 假设 $\lambda_0 \notin \mathbb{R}$，$z \in \mathbb{C}^n$ 是 λ_0 的一个复特征向量。记 $z = x + iy$，这里 $x, y \in \mathbb{R}^n$。证明 $x \perp y$ 并且 $\|x\| = \|y\|$。

本题中，$\langle x|$ 表示 $|x\rangle$ 的转置共轭；

证明. (1)

由矩阵之正交性有，设 $|0\rangle$ 是其本征值为 λ_0 的本征矢量，

$\langle 0| A^\dagger A |0\rangle = \langle 0|0\rangle = \langle 0|0\rangle \cdot \lambda_0^* \lambda_0$

$\therefore \lambda_0^* \lambda_0 = 1$ 即 $|\lambda_0| = 1$

(2)

设 $\lambda_0 = \alpha + i\beta$, 注意到

$A|z\rangle = A(|x\rangle + i|y\rangle) = (\alpha + i\beta)(|x\rangle + i|y\rangle) = (\alpha|x\rangle - \beta|y\rangle) + i(\beta|x\rangle + \alpha|y\rangle)$
由实虚部分别相等有（A 设为实矩阵）
$A|x\rangle = \alpha|x\rangle - \beta|y\rangle$
$A|y\rangle = \alpha|y\rangle + \beta|x\rangle$
同理，由矩阵之正交性有
$\langle x|A^\dagger = \alpha\langle x| - \beta\langle y|$
$\langle y|A^\dagger = -\alpha\langle y| - \beta\langle x|$
$\because \langle x|y\rangle = \langle y|A^\dagger A|x\rangle = -\langle x|A^\dagger A|y\rangle = \langle x|A^\dagger A|y\rangle$
$\therefore \langle x|y\rangle = 0$, 即 $|x\rangle, |y\rangle$ 正交.
$\therefore \langle x|x\rangle = \langle x|A^\dagger A|x\rangle = (\alpha\langle x| - \beta\langle y|)(\alpha|x\rangle - \beta|y\rangle) = \alpha^2\langle x|x\rangle + \beta^2\langle y|y\rangle - 2\alpha\beta\langle x|y\rangle = \alpha^2\langle x|x\rangle + \beta^2\langle y|y\rangle$
即 $\beta^2\langle y|y\rangle = (1-\alpha^2)\langle x|x\rangle$
由 (a) 中已证 $|\lambda_0| = 1$
$\therefore 1-\alpha^2 = \beta^2$
即 $\langle x|x\rangle = \langle y|y\rangle$, 即 $\|x\| = \|y\|$ □

问题 15. *(矩阵的迹)*
设 A 是 n 阶反对称方阵，即 $A^T = -A$。证明 $e^A \in SO(n)$ 是 n 阶特殊正交方阵。

证明. 先证正交性: 由泰勒展开
$e^A = \sum\limits_{k=0}^{\infty} \frac{A^k}{k!}$
由转置的运算规律及反对称性
$(e^A)^T = \sum\limits_{k=0}^{\infty} \frac{{A^T}^k}{k!} = \sum\limits_{k=0}^{\infty} \frac{(-A)^k}{k!} = e^{A^T} = e^{-A}$
$\therefore e^A \cdot (e^A)^T = 1$, 即是正交的。
再证行列式为 1:

引理 5. 反对称阵具有纯虚共轭或为零的本征值;

证明. 设 A 具有本征值 $\lambda = \alpha + i\beta; \alpha, \beta \in \mathbb{R}$, 对应的本征矢量为 $|z\rangle = |x\rangle + i|y\rangle; |x\rangle, |y\rangle \in \mathbb{R}^n$
$\because A|z\rangle = A(|x\rangle + i|y\rangle) = (\alpha + i\beta)(|x\rangle + i|y\rangle) = (\alpha|x\rangle - \beta|y\rangle) + i(\beta|x\rangle + \alpha|y\rangle)$
由实虚部分别相等有（A 设实矩阵）
$A|x\rangle = \alpha|x\rangle - \beta|y\rangle$
$A|y\rangle = \alpha|y\rangle + \beta|x\rangle$
$\therefore \langle x|A|x\rangle = \alpha\langle x|x\rangle - \beta\langle x|y\rangle, \langle y|A|y\rangle = \alpha\langle y|y\rangle + \beta\langle x|y\rangle$
两式相加, 有

$\langle x| A |x\rangle + \langle y| A |y\rangle = \alpha(\langle x|x\rangle + \langle y|y\rangle)$
同取转置, 有:
$(\langle x| A |x\rangle + \langle y| A |y\rangle)^T = \alpha(\langle x|x\rangle + \langle y|y\rangle)$
而考虑反对称性:
$(\langle x| A |x\rangle + \langle y| A |y\rangle)^T = \langle x| A^T |x\rangle + \langle y| A^T |y\rangle = -(\langle x| A |x\rangle + \langle y| A |y\rangle)$
$\therefore \langle x| A |x\rangle + \langle y| A |y\rangle = \alpha(\langle x|x\rangle + \langle y|y\rangle) = 0$
由内积之正定性,$\alpha = 0$, 即是纯虚的.
而考虑 A 是实矩阵, 其特征方程是实系数多项式方程, 解必为共轭复数或 0;
$\therefore$ 反对称阵具有纯虚共轭或为零的本征值;

□

由引理, 及约当标准型,

$\exists P$ 可逆,$A = P \begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \cdots & * \\ & & \lambda_3 & \cdots \\ & & & \cdots \\ & & & \lambda_n \end{bmatrix} P^{-1} = P\Lambda P^{-1}$

其中,λ_i 是之本征值, 由泰勒展开:

$e^A = \sum\limits_{k=0}^{\infty} \frac{A^k}{k!} = \sum\limits_{k=0}^{\infty} P\frac{\Lambda^k}{k!}P^{-1}$

而 $\Lambda^k = \begin{bmatrix} \lambda_1^k & * & \cdots & * \\ & \lambda_2^k & \cdots & * \\ & & \lambda_3^k & \cdots \\ & & & \cdots \\ & & & \lambda_n^k \end{bmatrix}$

$\therefore e^A = P \begin{bmatrix} e_1^\lambda & * & \cdots & * \\ & e_2^\lambda & \cdots & * \\ & & e_3^\lambda & \cdots \\ & & & \cdots \\ & & & e_n^\lambda \end{bmatrix} P^{-1}$

$\therefore \det(e^A) = \det(P) exp[\sum\limits_{k=1}^{n} \lambda_k] \det(P^{-1})$

由引理,$\sum\limits_{k=1}^{n} \lambda_k = 0$, 即 $\det(e^A) = 1$

$\therefore e^A \in SO(3)$ □

以上证明过于笨拙，我们可以用迹的性质来简化行列式为 1 的证明 (大家应该都是这么做

的，但这是用于复习的好例子)。

引理 6. 对 $A \in F^{n\times n}$,

$$det[e^A] = e^{tr[A]}$$

为了证明这个引理，我们需要引入迹的运算性质。

引理 7.
(1)$tr[A+B] = tr[A] + tr[B], tr[A]^* = tr[A^\dagger]$
(2)(轮换性)$tr[AB] = tr[BA], tr[ABC] = tr[BCA] = tr[CAB]\cdots$
(3)(相似不变迹) 若 P *可逆，*$tr[A] = tr[P^{-1}AP]$

证明.
(1)
显然；
(2)
$\mathrm{tr}[\prod_{s=1}^{l} A^{(s)}] = \sum_{i_1,i_2,\cdots,i_l} a^{(1)}_{i_1 i_2} a^{(2)}_{i_2 i_3} \cdots a^{(l)}_{i_l i_1} = \sum_{i_1,i_2,\cdots,i_l} a^{(2)}_{i_2 i_3} a^{(3)}_{i_3 i_4} \cdots a^{(l)}_{i_l i_1} a^{(1)}_{i_1 i_2} = \mathrm{tr}[[\prod_{s=2}^{l} A^{(s)}]A^{(1)}]$
证讫。
(3)
由 (2),$\mathrm{tr}[P^{-1}AP] = \mathrm{tr}[APP^{-1}] = \mathrm{tr}[AI] = \mathrm{tr}[A]$ 证讫。 □

对引理 6 的证明：

证明. 由约当标准型，A 可以相似地化为上三角形式

$$A = P \begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \cdots & * \\ & & \lambda_3 & \cdots \\ & & & \cdots \\ & & & \lambda_n \end{bmatrix} P^{-1} = P\Lambda P^{-1}$$

$$\therefore e^A = P \begin{bmatrix} e^\lambda_1 & * & \cdots & * \\ & e^\lambda_2 & \cdots & * \\ & & e^\lambda_3 & \cdots \\ & & & \cdots \\ & & & e^\lambda_n \end{bmatrix} P^{-1}$$

$\therefore \det[e^A] = e^{\sum_{i=1}^{n}\lambda_i} = e^{\mathrm{tr}[\Lambda]} = e^{tr[A]}$

$\square$

从而，对原问题的证明就要简单的多。

证明. 由于 A 是反对称矩阵，$\mathrm{tr}[A] = 0$
$\therefore \det[e^A] = e^0 = 1$

$\square$

与迹有关的其它习题：(homework14 3. 7.)

问题 16. 设 A, B 是 n 阶厄密方阵。证明 $tr(AB)^2$ 和 $tr(A^2B^2)$ 都是实数并且

$$tr(AB)^2 \leq tr(A^2B^2)$$

等号成立当且仅当 $AB = BA$。

证明. 首先证实数，由于 AB 是厄密矩阵

$$\mathrm{tr}[ABAB]^* = \mathrm{tr}[(ABAB)^\dagger] = \mathrm{tr}[BABA] = \mathrm{tr}[ABAB] \in \mathbb{R}$$

$$\mathrm{tr}[AABB]^* = \mathrm{tr}[(AABB)^\dagger] = \mathrm{tr}[BBAA] = \mathrm{tr}[AABB] \in \mathbb{R}$$

再证不等式: 结合我们之前提到的对称化构造思想，这里要研究差值，故采取反称构造 $C = iAB - iBA$
$S = C^\dagger C = -ABAB - BABA + ABBA + BAAB, \mathrm{tr}[C'^2] = -2\mathrm{tr}[ABAB] + 2\mathrm{tr}[AABB]$，只需证明 C' 正定即可。而如我们所希望的，$S$ 具有这样的形式：$S = C^\dagger C$，正定性是易于证明的。 $\square$

问题 17. 设 n 阶复方阵 A 的所有特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$。证明

$$\sum_{i=1}^{n} |\lambda_i|^2 \leq tr(A^\dagger A)$$

等号成立当且仅当 A 是正规矩阵。

本题十分简单，只需用到相似不变迹的性质。权当一个小练习：

证明. 由讲义命题，A 可以酉相似地化为上三角形式

$$S^{-1}AS = [a_{ij}] = \begin{bmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ & \lambda_2 & \cdots & a_{2n} \\ & & \cdots & \\ & & \cdots & \lambda_n \end{bmatrix}$$

$\therefore \mathrm{tr}[A^\dagger A] = \mathrm{tr}[S^{-1}A^\dagger AS] = \sum_{i,j} a_{ij}a_{ij}^* \geq \sum_i a_{ii}a_{ii}^* = \sum\limits_{i=1}^{n} |\lambda_i|^2$
当且仅当 A 为正规矩阵时，可以对角化，非对角元全为零，取得等号。

□

最后一个练习：

问题 18. 试证：若 A，B 为 n 阶实方阵，则

$$tr[AB(AB)^T] \leq tr[AA^T] \cdot \Lambda_{BB^T}$$

其中 Λ_{BB^T} 是 BB^T 的最大特征根。

证明. 轮换到需要的形式即可。

$$\mathrm{tr}[AB(AB)^T] = \mathrm{tr}[ABB^TA^T] = \mathrm{tr}[(A^TA)(BB^T)]$$

考虑 BB^T 是正定实方阵，正交对角化，有：

$$\mathrm{tr}[(A^TA)(BB^T)] = \mathrm{tr}[P(A^TA)P^TP(BB^T)P^T] = \mathrm{tr}[P(A^TA)P^T\Lambda] = \mathrm{tr}[A'\Lambda] = \sum_{i=1}^{n} a'_{ii}\lambda_i \leq \Lambda_{BB^T}\sum_{i=1}^{n} a'_{ii}$$

而

$$\sum_{i=1}^{n} a'_{ii} = \mathrm{tr}[P(A^TA)P^T] = \mathrm{tr}[A^TA] = \mathrm{tr}[AA^T]$$

$$\therefore \mathrm{tr}[AB(AB)^T] \leq \mathrm{tr}[AA^T] \cdot \Lambda_{BB^T}$$

□

5 早期结论复习

现将一些早期而常用的结论罗列如下，便于复习。
1.$\det[I - AB] = \det[I - BA]$
2.$\nexists$方阵$A, B, \mathrm{s.t.} AB - BA = I$

3. 存在可逆 m 阶方阵 P 和可逆 n 阶方阵 Q 使得 $A = P\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}Q$(相抵标准型)

4. 上三角行列式的逆仍是上三角行列式

5.(范德蒙德行列式？)

6 写在最后

笔者是清华大学 24 级物理系的一名同学，本文依据李思老师：线性代数 (理科类) 课程上课内容和笔者所思所想结合而来。结合李老师课堂讲义阅读体验更佳。笔者的本意是希望通过一些做题经验感触，更有逻辑地叙述线性代数，并结合笔者所熟悉的知识，使得线性代数更加具体而非抽象，这也是标题叫 Essential Linear Algebra 的原因。笔者无意编写一套严谨的数学笔记 (毕竟是物理系的)，希望看上去更像“欧拉”而非“柯西”，这是有助于笔者这种平凡人类理解的。

第二版修改了部分错误，增补了 DN 分解、结论汇总和部分习题。特别感谢刘宏润、刘乐融、周启轩、王博樊、曹乐宸 (复旦) 等同学和助教老师的对本文的建议、纠错和帮助。

限于笔者水平，此笔记必有诸多谬误疏漏有待指正，以及不足之处冀望各位高水平人士为之补全。笔者万分荣幸，感激不尽。如果你觉得这篇笔记对你有帮助，且愿意请我和咖啡的话，我个人是十分欢迎的。