第一章 绪论

略

第二章 行列式

§2.1 二阶与三阶行列式

一、二元线性方程组与二阶行列式

用消元法解二元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{2.1.1}$$

得

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}} \tag{2.1.2}$$

我们用符号 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ 表示代数和 $a_{11}a_{22} - a_{12}a_{21}$，这样就有

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \tag{2.1.3}$$

符号 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ 表示代数和 $a_{11}a_{22} - a_{12}a_{21}$，我们称它为二阶行列式 。

行列式中的相关术语：

行列式的元素、行、列、主对角线、副对角线.

对角线法则 ：

二阶行列式是主对角线上两元素之积减去的副对角线上二元素之积所得的差.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{2.1.4}$$

例 1　求解二元线性方程组 $\begin{cases} 3x_1 - 2x_2 = 12 \\ 2x_1 + x_2 = 1 \end{cases}$

解：由于

$$D = \begin{vmatrix} 3 & -2 \\ 2 & 1 \end{vmatrix} = 3 - (-4) = 7 \neq 0$$

$$D = \begin{vmatrix} 12 & -2 \\ 1 & 1 \end{vmatrix} = 12 - (-2) = 14$$

$$D = \begin{vmatrix} 3 & 12 \\ 2 & 1 \end{vmatrix} = 3 - (24) = -21$$

因此

$$x_1 = \frac{D_1}{D} = \frac{14}{7} = 2 \quad , \quad x_2 = \frac{D_2}{D} = \frac{-21}{7} = -3 \quad .$$

二、三阶行列式

方程组 $\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{12}x_2 + a_{13}x_3 = b_2 \\ a_{31}x_1 + a_{12}x_2 + a_{13}x_3 = b_3 \end{cases}$ 的解为

$$x_1 = \frac{D_1}{D}\ ,\ x_2 = \frac{D_2}{D}\ ,\ x_3 = \frac{D_3}{D}$$

其中$D = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$　（2.1.5）

$$D_1 = b_1a_{22}a_{33} + a_{12}a_{23}b_3 + a_{13}b_2a_{32} - b_1a_{23}a_{32} - a_{12}b_2a_{33} - a_{13}a_{22}b_3$$

$$D_2 = a_{11}b_2a_{33} + b_1a_{23}a_{31} + a_{13}a_{21}b_3 - a_{11}a_{23}b_3 - b_1a_{21}a_{33} - a_{13}b_2a_{31}$$

$$D_3 = a_{11}a_{22}b_3 + a_{12}b_2a_{31} + b_1a_{21}a_{32} - a_{11}b_2a_{32} - a_{12}a_{21}b_3 - b_1a_{22}a_{31}$$

为了便于记忆和计算，我们用符号$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$表示代数和

$$a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}$$

并称它为三阶行列式

行列式中的相关术语

行列式的元素、行、列、主对角线、副对角线.

对角线法则

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

$$-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}$$

例 2　计算三阶行列式$\begin{vmatrix} 1 & 2 & -4 \\ -2 & 2 & 1 \\ -3 & 4 & -2 \end{vmatrix}$

解：按对角线法则，有

$$D=1\times2\times(-2)+2\times1\times(-3)+(-4)\times(-2)\times4$$

$$-1\times1\times4-2\times(-2)\times(-2)-(-4)\times2\times(-3)$$

$$=-4-6+32-4-8-24=-14$$

例 3　求解方程$\begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & x \\ 4 & 9 & x^2 \end{vmatrix}=0$

解：方程左端的三阶行列式

$$D=3x^2+4x+18-9x-2x^2-12=x^2-5x+6$$

由 $x^2-5x+6=0$，解得 $x=2$ 或 $x=3$

§2.2　全排列及其逆序数

引例　用 1、2、3 三个数字，可以组成多少个没有重复数字的三位数？

解：采用先选定百位数，再选定十位数，最后选定个位数的步骤。百位数有 3 种选法，十位数有 2 种选法，个位数有 1 种选法。因为 $3\times2\times1=6$，所以可以组成 6 个没有重复数字的三位数.这 6 个三位数是：

123，132，213，231，312，321

全排列

我们把 n 个不同的对象(称为元素)排成一列，叫做这 n 个元素的全排列(也简称排列). n 个不同元素的所有排列的总数，通常用P_{n}表示。

P_{n}的计算公式：$P_{\mathrm{n}} = n\cdot(n-1)\cdot(n-2)\cdots3\cdot2\cdot1 = n!$　(2.2.1)

举例：由 *a*, *b*, *c* 组成的所有排列为 *abc*　*acb*　*bac*　*bca*　*cab*　*cba*。*abb* 是排列吗？

标准排列（也叫自然排列）

在 n 个自然数的全排列中排列 $123\cdots n$ 称为标准排列.

逆序与逆序数

在一个排列中，如果某两个元素的先后次序与标准排列的次序不同，就说有 1 个逆序. 一个排列中所有逆序的总数叫做这个排列的逆序数.（以下我们只讨论 *n* 个自然数的全排列）

逆序数的计算

在排列$P_1P_2\cdots P_{\mathrm{n}}$中，如果$P_i$的前面有个大于$P_i$的数，就说元素$P_i$的逆序数是 t_i. 排列的逆序数为 $t=t_1+t_2+\cdots+t_n$. 例如:在排列 32514

中，$t_1=0$，$t_2=1$，$t_3=0$，$t_4=3$，$t_5=1$。排列 32514 的逆序数为 $t=0+1+0+3+1=5$.

奇排列与偶排列

逆序数为奇数的排列叫做奇排列，逆序数为偶数的排列叫做偶排列.例如：排列 32514 的逆序数是 5，它是奇排列.标准排列 12345 的逆序数是 0，它是偶排列.

练习：1. 选择 i 与 k 使得

1）1274i56k9 是偶排列；2）1i25k4897 是奇排列

2、如果排列$x_1x_2\dots x_{n-1}x_n$的逆序数为 k，求$x_{n-1}x_n\cdots x_2x_1$的逆序数为多少？

§2.3 *n* 阶行列式的定义

为了给出 n 阶行列式的定义，我们要先研究三阶行列式的结构。请思考：三阶行列式存在什么规律？

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \qquad (2.3.1)$$

三阶行列式的结构

(1) 行列式右边任一项除正负号外可以写成：$a_{1p_1}a_{2p_2}a_{3p_3}$，其中$p_1p_2p_3$是 1、2、3 的某个排列.

(2) 各项所带的正负号可以表示为$(-1)^t$，其中 t 为列标排列的逆序数.

三阶行列式可以写成 ：

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum(-1)^{\mathrm{t}} a_{1p_1}a_{2p_2}a_{3p_3} \ , \qquad (2.3.2)$$

其中 t 为排列 $p_1p_2p_3$ 的逆序数，$\sum$表示对 1、2、3 三个数的所有排列 $p_1p_2p_3$ 取和.

n 阶行列式的定义

由 n^2 个数$a_{ij}(i, j=1, 2, \cdots, n)$构成的代数和 $\sum(-1)^t a_{1p_1}a_{2p_2}a_{3p_3}$ 称为 n 阶行列式，记为

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

简记为 $\det(a_{ij})$，其中 $p_1p_2\cdots p_n$ 为自然数 $1, 2, \cdots, n$ 的一个排列，t 为这个排列的逆序数,$\sum$表示对所有排列 $p_1p_2\cdots p_n$ 取和。数 a_{ij} 为行列式 D 的(i, j)元，特别规定一阶行列式$|a|$的值就是 a.

例 1　证明行列式 $D = \begin{vmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}$

解：要使取自不同行不同列的 n 个元素的乘积不为零,第一行只能取 a_{11}，第二行只能取 a_{22}，第三行只能取 a_{33}，…,第 n 行只能取 $a_{nn.}$。这样的乘积项只有一个，即$a_{11}a_{22}\cdots a_{nn}$。因为它的列标排列为标准排列，其逆序数为 0，所以在它前面带有正号。因此

$$D = a_{11}a_{22}\cdots a_{nn}$$

注意：上三角形、下三角形及对角形行列式的值等于主对角线上 n 个元素的乘积。

例 2　证明 n 阶行列式$D=\begin{vmatrix} & & & \lambda_1 \\ & & \lambda_2 & \\ & \iddots & & \\ \lambda_n & & & \end{vmatrix}=(-1)^{\frac{n(n-1)}{2}}\lambda_1\lambda_2\cdots\lambda_n$

解：若记$\lambda_i=a_{i,\,n-i+1}$，则依行列式定义

$$D=\begin{vmatrix} & & & a_{1n} \\ & & a_{2,n-1} & \\ & \iddots & & \\ a_{n1} & & & \end{vmatrix}=(-1)^t a_{1n}a_{2,n-1}\cdots a_{n1}$$

$$=(-1)^t\lambda_1\lambda_2\cdots\lambda_n$$

其中 t 为排列 $n\cdot(n-1)\cdots 2\cdot 1$ 的逆序数，故

$$\mathrm{t}=0+1+2+\cdots+(n-1)=\frac{n(n-1)}{2}$$

因此　$D=(-1)^{\frac{n(n-1)}{2}}\lambda_1\lambda_2\cdots\lambda_n$

例 3：计算斜上三角行列式$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1\,\mathrm{n}-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2\,n-1} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & 0 & \cdots & 0 & 0 \end{vmatrix}$

解：$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1\,\mathrm{n}-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2\,n-1} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & 0 & \cdots & 0 & 0 \end{vmatrix}$

$=\sum sgn\begin{pmatrix} 1 & 2 & \cdots & n \\ & & & \end{pmatrix}a_{1n}a_{2,n-1}\cdots a_{n1}$

同理可计算斜下三角形行列式和斜对角形行列式。

例 4：计算$\begin{vmatrix} 1 & -1 & x & 0 \\ 0 & 3 & x & 2 \\ 0 & x^2 & x & 1 \\ x & 4 & 5 & 3 \end{vmatrix}$的最高项和常数项

练习：在 6 阶行列式中，$a_{23}a_{31}a_{42}a_{56}a_{14}a_{65}$和$a_{23}a_{43}a_{14}a_{51}a_{66}a_{25}$这两项分别为什么系数

§2.4　对换

在排列中，将任意两个元素对调，其余的元素不动，就得到另一个排列，这种对排列的变换方法称为对换. 将相邻两个元素对换，叫做相邻对换.

例如：在排列 21354 中，对换 1 与 4，得到的排列是 24351. 排列 21354 的逆序数是 2，排列 24351 的逆序数是 5。经过对换，排列的奇偶性发生了变化。

定理 1

一个排列中的任意两个元素对换，排列改变奇偶性

$$\cdots ij \cdots \to \ \cdots ji \cdots$$

$$\cdots i a_1 a_2 \cdots a_s j \cdots \ \to \ \cdots a_1 a_2 \cdots a_s j i \cdots \ \to \ \cdots j a_1 a_2 \cdots a_s i \cdots$$

推论: 奇排列变成标准排列的对换次数为奇数，偶排列变成标准排列的对换次数为偶数. 这是因为，由定理 1 知对换的次数就是排列奇偶性的变化次数，因此知推论成立.

推论: 全体 n 元排列的集合中，奇排列和偶排列各占一半.

定理 2

n 阶行列式也可定义为$\sum(-1)^{\mathrm{t}} a_{p_1 1} a_{p_2 2} \cdots a_{p_n n}$ ，其中 t 为行标排列 $p_1 p_2 \cdot\cdot\cdot p_n$ 的逆序数.

$$(-1)^{t(i_1 i_2 \cdots i_n)} a_{1 i_1} a_{2 i_2} \cdots a_{n i_n} \qquad (-1)^{\mathrm{t}(p_1 p_2 \cdots p_n)} a_{p_1 1} a_{p_2 2} \cdots a_{p_n n}$$

将列下标$i_1 i_2 \cdots i_n$经过 k 次对换变成标准排列的同时相应的行下标都变成了$p_1 p_2 \cdots p_n$，k 与$t(i_1 i_2 \cdots i_n)$和$\mathrm{t}(p_1 p_2 \cdots p_n)$有相同的奇偶性。

§2.5　行列式的性质

行列式的转置

将行列式 D 的行变为列后得到的行列式称为 D 的转置行列式，记为 D^T.即

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}，\text{则}D^T=\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix}.$$

显然，如果

$$D=\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}，\text{则}b_{ij}=a_{ji}(i,j=1,2,\cdots,n). \quad（2.5.1）$$

性质 1：行列式 D 与它的转置行列式 D^T 相等.

由此性质可知，行列式中的行与列具有同等的地位，行列式的性质凡是对行成立的对列也同样成立，反之亦然.

性质 2: 互换行列式的两行，行列式变号.

推论：如果行列式有两行(列)完全相同，则此行列式等于零.这是因为，把这两行互换，有 $D=-D$，故 $D=0$.

性质 3: 行列式的某一行(列)中所有的元素都乘以同一数 k，等于用数 k 乘此行列式.

推论：行列式中某一行(列)的所有元素的公因子可以提到行列式符号的外面

性质 4：行列式中如果有两行(列)元素成比例，则行列式等于零.

性质 5：若行列式的某一行(列)的元素都是两个数之和，则行列式等于两个行列式之和. 即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{in}+b_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$
=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{2.5.2}
$$

性质 6: 把行列式的某一行(列)的各元素乘以同一数然后加到另一列(行)对应的元素上去，行列式不变. 即

$$
\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1}+ka_{j1} & a_{i2}+ka_{j2} & \cdots & a_{in}+ka_{jn} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}
$$

符号规定

在计算行列式时，可以使用如下记号以便检查:

交换 i, j 两行记作 $r_i \leftrightarrow r_j$，交换 i, j 两列记作 $c_i \leftrightarrow c_j$.

第 i 行(或列)提出公因子 k，记作 $r_i \div k$(或 $c_i \div k$).

以数 k 乘第 j 行(列)加到第 i 行(列)上，记作 $r_i + kr_j$ ($c_i + kc_j$) .

例1　计算$D = \begin{vmatrix} 3 & 1 & -1 & 2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{vmatrix}$

$$
\text{解：} D = \begin{vmatrix} 3 & 1 & -1 & 2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{vmatrix} \xlongequal{c_1 \leftrightarrow c_2} - \begin{vmatrix} 1 & 3 & -1 & 2 \\ 1 & -5 & 3 & -4 \\ 0 & 2 & 1 & -1 \\ -5 & 1 & 3 & -3 \end{vmatrix}
$$

$$
\xlongequal[r_4+5r_1]{r_2-r_1} - \begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & -8 & 4 & -6 \\ 0 & 2 & 1 & -1 \\ 0 & 16 & -2 & 7 \end{vmatrix} \xlongequal{r_2 \leftrightarrow r_3} \begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & -8 & 4 & -6 \\ 0 & 16 & -2 & 7 \end{vmatrix}
$$

$$
r_4 + \frac{5}{4} r_3
$$

$$\xlongequal[r_4-8r_2]{r_3+r_2}\begin{vmatrix}1&3&-1&2\\0&2&1&-1\\0&0&8&-10\\0&0&-10&15\end{vmatrix}=\!=\!=\begin{vmatrix}1&3&-1&2\\0&2&1&-1\\0&0&8&-10\\0&0&0&5/2\end{vmatrix}=40$$

例2　计算$D=\begin{vmatrix}3&1&1&1\\1&3&1&1\\1&1&3&1\\1&1&1&3\end{vmatrix}$

解 $D=\begin{vmatrix}3&1&1&1\\1&3&1&1\\1&1&3&1\\1&1&1&3\end{vmatrix}\xlongequal{c_1+c_2+c_3+c_4}\begin{vmatrix}6&1&1&1\\6&3&1&1\\6&1&3&1\\6&1&1&3\end{vmatrix}$

$$\xlongequal{c_1\div 6}6\begin{vmatrix}1&1&1&1\\1&3&1&1\\1&1&3&1\\1&1&1&3\end{vmatrix}\xlongequal[r_4-r_1]{\substack{r_2-r_1\\r_3-r_1}}6\begin{vmatrix}1&1&1&1\\0&2&0&0\\0&0&2&0\\0&0&0&2\end{vmatrix}=6\times 8=48$$

例 3　计算$D=\begin{vmatrix}a&b&c&d\\a&a+b&a+b+c&a+b+c+d\\a&2a+b&3a+2b+c&4a+3b+2c+d\\a&3a+b&6a+3b+c&10a+6b+3c+d\end{vmatrix}$

解：$D\xlongequal[r_2-r_1]{\substack{r_4-r_3\\r_3-r_2}}\begin{vmatrix}a&b&c&d\\0&a&a+b&a+b+c\\0&a&2a+b&3a+2b+c\\0&a&3a+b&6a+3b+c\end{vmatrix}\xlongequal[r_3-r_2]{r_4-r_3}$

$$\begin{vmatrix}a&b&c&d\\0&a&a+b&a+b+c\\0&0&a&2a+b\\0&0&a&3a+b\end{vmatrix}\xlongequal{r_4-r_3}\begin{vmatrix}a&b&c&d\\0&a&a+b&a+b+c\\0&0&a&2a+b\\0&0&a&a\end{vmatrix}=a^4$$

例 4　证明 $D=D_1\cdot D_2$，其中$D=\begin{vmatrix} a_{11} & \cdots & a_{1k} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n1} & \cdots & c_{nk} & b_{n1} & \cdots & b_{nn} \end{vmatrix}$，

$D_1=\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}$，$D_2=\begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots \\ b_{n1} & \cdots & b_{nn} \end{vmatrix}$。

证：对 D_1 作运算 r_i+kr_j，把 D_1 化为下三角形行列式，设为

$D_1=\begin{vmatrix} p_{11} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ p_{k1} & \cdots & p_{kk} \end{vmatrix}=p_{11}\cdots p_{kk}$，对 D_2 作运算 c_i+kc_j，把 D_2 化为下三角形行列式，设为 $D_2=\begin{vmatrix} q_{11} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ q_{n1} & \cdots & q_{nn} \end{vmatrix}=q_{11}\cdots q_{nn}$，于是，对 D 的前 k 行作运算 r_i+kr_j，再对后 n 列作运算 c_i+kc_j，把 D 化为下三角形行列式$D=\begin{vmatrix} p_{11} & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{k1} & \cdots & p_{kk} & 0 & \cdots & 0 \\ c_{11} & \cdots & c_{1k} & q_{11} & \cdots & q_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n1} & \cdots & c_{nk} & q_{n1} & \cdots & q_{nn} \end{vmatrix}$，

故$D=p_{11}\cdots p_{kk}q_{11}\cdots q_{nn}=D_1\cdot D_2$.

例 5　计算 $2n$ 阶行列式 $D_{2n}=\begin{vmatrix} a & & & & & b \\ & \ddots & & & \iddots & \\ & & a & b & & \\ & & c & d & & \\ & \iddots & & & \ddots & \\ c & & & & & d \end{vmatrix}$，其中未写出的元素为零。

解：把 D_{2n} 中的第 $2n$ 行依次与 $2n-1$ 行、…、第 2 行对调(作 $2n-2$ 次相邻对换)，再把第 $2n$ 列依次与 $2n-1$ 列、…、第 2 列对调，得

$$D_{2n}=(-1)^{2(2n-2)}\begin{vmatrix} a & b & & & & & & \\ c & d & & & & & & \\ & & a & & & & & b \\ & & & \ddots & & & \iddots & \\ & & & & a & b & & \\ & & & & c & d & & \\ & & & \iddots & & & \ddots & \\ & & c & & & & & d \end{vmatrix}$$，根据例 4 的结果，

有$D_{2n}=D_2\cdot D_{2(n-1)}=(ad-bc)D_{2(n-1)}$，以此作递推公式，即得

$D_{2n}=(ad-bc)^2D_{2(n-1)}=\cdots=(ad-bc)^{n-1}D_2=(ad-bc)^n$

练习： 计算下面的行列式：

$$\begin{vmatrix} 1 & 2 & -3 & -4 \\ -1 & -2 & 5 & -8 \\ 0 & -1 & 2 & -1 \\ 1 & 3 & -5 & 10 \end{vmatrix} \qquad \begin{vmatrix} 1 & x & y & z \\ x & 1 & 0 & 0 \\ y & 0 & 1 & 0 \\ z & 0 & 0 & 1 \end{vmatrix}$$

§2.6　行列式按行(列)展开

余子式与代数余子式

在 n 阶行列式 $D=det(a_{ij})$中，把元素 a_{ij} 所在的第 i 行和第 j 列划去后，剩下来的 $n-1$ 阶行列式叫做元素 a_{ij} 的余子式，记作 M_{ij} ;记 $A_{ij}=(-1)^{i+j}M_{ij}$， A_{ij} 叫做元素 a_{ij} 的代数余子式.

例如，已知$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & \boldsymbol{a_{23}} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix}$，则 a_{23} 的余子式和代数余子式分别为$\begin{vmatrix} a_{11} & a_{12} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{54} & a_{55} \end{vmatrix}$，和$A_{23} = (-1)^{2+3} M_{23} = -M_{23}$.

引理：在 n 阶行列式 D 中，如果第 i 行元素除 a_{ij} 外都为零，那么这行列式等于 a_{ij} 与它的代数余子式 A_{ij} 的乘积，即 $D=a_{ij}\cdot A_{ij}$.

定理 3（行列式按行（列）展开法则）

行列式等于它的任一行（列）各元素与其对应的代数余子式乘积的和，即 $D = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}\ (i = 1, 2\cdots, n)$ 或

$$D = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}\ (j = 1, 2\cdots, n) \qquad (2.6.1)$$

推论：行列式某一行（列）的元素与另一行（列）的对应元素的代数余子式乘积之和等于零. 即

$$\sum_{k=1}^{n} a_{ik}A_{jk} = \begin{cases} D & \text{当 } \boldsymbol{i=j} \\ 0 & \text{当 } \boldsymbol{i\neq j} \end{cases} \qquad (2.6.2)$$

或者 $$\sum_{k=1}^{n} a_{ki}A_{kj} = \begin{cases} D & \text{当 } \boldsymbol{i=j} \\ 0 & \text{当 } \boldsymbol{i\neq j} \end{cases} \qquad (2.6.3)$$

例1　计算行列式 $D = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 2 \\ 3 & -1 & -1 & 0 \\ 1 & 2 & 0 & -5 \end{vmatrix}$

解：将 D 按第三列展开,应有

$$D = a_{13}A_{13} + a_{23}A_{23} + a_{33}A_{33} + a_{43}A_{43}$$ ，其中

a_{13}=3, a_{23}=1, a_{33}=-1, a_{43}=0， $A_{13}=(-1)^{1+3}\begin{vmatrix}1 & 0 & 2\\ 3 & -1 & 0\\ 1 & 2 & -5\end{vmatrix}=19$

$$A_{23}=(-1)^{2+3}\begin{vmatrix}1 & 2 & 4\\ 3 & -1 & 0\\ 1 & 2 & -5\end{vmatrix}=-63$$

$$A_{33}=(-1)^{3+3}\begin{vmatrix}1 & 2 & 4\\ 1 & 0 & 2\\ 1 & 2 & -5\end{vmatrix}=18$$

$$A_{43}=(-1)^{4+3}\begin{vmatrix}1 & 2 & 4\\ 1 & 0 & 2\\ 3 & -1 & 0\end{vmatrix}=-10$$

所以$D=3\times 19+(-63)\times 1+(-1)\times 18+0\times(-10)=-24$

例2 $D_n=\begin{vmatrix}1 & 1 & \cdots & 1\\ a_1 & a_2 & \cdots & a_n\\ a_1^2 & a_2^2 & \cdots & a_n^2\\ \cdots & \cdots & \cdots & \cdots\\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1}\end{vmatrix}$

$$=\begin{vmatrix}1 & 1 & \cdots & 1\\ 0 & a_2-a_1 & \cdots & a_n-a_1\\ 0 & a_2^2-a_1a_2 & \cdots & a_n^2-a_1a_n\\ \cdots & \cdots & \cdots & \cdots\\ 0 & a_2^{n-1}-a_1a_2^{n-2} & \cdots & a_n^{n-1}-a_1a_n^{n-2}\end{vmatrix}$$

$$=\begin{vmatrix}a_2-a_1 & a_3-a_1 & \cdots & a_2-a_1\\ a_2^2-a_1a_2 & a_3^2-a_1a_3 & \cdots & a_2^2-a_1n\\ \cdots & \cdots & \cdots & \cdots\\ a_2^{n-1}-a_1a_2^{n-2} & a_3^{n-1}-a_1a_3^{n-2} & \cdots & a_2^{n-1}-a_1a_n^{n-2}\end{vmatrix}$$

$$=(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)\begin{vmatrix}1 & 1 & \cdots & 1\\ a_2 & a_3 & \cdots & a_n\\ a_2^2 & a_3^2 & \cdots & a_n^2\\ \cdots & \cdots & \cdots & \cdots\\ a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2}\end{vmatrix}$$

$$=(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)D_{n-1}$$

于是$D_n=(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)D_{n-1}$，

$$= (a_2 - a_1)(a_3 - a_1)\cdots(a_n - a_1)(a_3 - a_2)\cdots(a_n - a_1)D_{n-2}$$

$$= \cdots$$

$$= \prod_{1\le i<j\le n}\left(a_j - a_i\right)$$

相关结果:

行列式按第 i 行展开，得

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \quad (2.6.4)$$

将元素 a_{i1} 换成 b_1, a_{i2} 换成 b_2, · · ·, a_{in} 换成 b_n，得

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ b_1 & b_2 & \cdots & b_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = b_1A_{i1} + b_2A_{i2} + \cdots + b_nA_{in} \quad (2.6.5)$$

如果第 i 行的元素为 $b_1, b_2, \cdots, b_n$，则有

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ b_1 & b_2 & \cdots & b_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = b_1A_{i1} + b_2A_{i2} + \cdots + b_nA_{in} \quad (2.6.6)$$

如果第 j 列的元素为 $b_1, b_2, \cdots, b_n$，则有

$$\begin{vmatrix} a_{11} & b_1 & \cdots & a_{1n} \\ a_{21} & b_2 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & b_n & \cdots & a_{nn} \end{vmatrix} = b_1A_{1j} + b_2A_{2j} + \cdots + b_nA_{nj} \quad (2.6.7)$$

例3　设$D = \begin{vmatrix} 3 & -5 & 2 & 1 \\ 1 & 1 & 0 & -5 \\ -1 & 3 & 1 & 3 \\ 2 & -4 & -1 & -3 \end{vmatrix}$，$D$ 的（i,j）的余子式和代数余子式依次记作 M_{ij} 和 A_{ij}，求 $A_{11}+A_{12}+A_{13}+A_{14}$ 及 $M_{11}+M_{21}+M_{31}+M_{41}$

解：$A_{12}+A_{12}+A_{12}+A_{12}=\begin{vmatrix}1&1&1&1\\1&1&0&-5\\-1&3&1&3\\2&-4&-1&-3\end{vmatrix}\overset{r_4+r_3}{\underset{r_3-r_1}{=\!=\!=}}$

$$\begin{vmatrix}1&1&1&1\\1&1&0&-5\\-2&2&0&2\\1&-1&0&0\end{vmatrix}\overset{\text{按第三列展开}}{=\!=\!=}\begin{vmatrix}1&1&-5\\-2&2&2\\1&-1&0\end{vmatrix}\overset{c_2+c_1}{=\!=\!=}\begin{vmatrix}1&2&-5\\-2&0&2\\1&0&0\end{vmatrix}$$

$$\overset{\text{按第三行展开}}{=\!=\!=}\begin{vmatrix}2&-5\\0&2\end{vmatrix}=4$$

$M_{11}+M_{21}+M_{31}+M_{41}=A_{11}-A_{21}+A_{31}-A_{41}$

$$=\begin{vmatrix}1&-5&2&1\\-1&1&0&-5\\1&3&1&3\\-1&-4&-1&-3\end{vmatrix}\overset{r_4+r_3}{=\!=\!=}\begin{vmatrix}1&-5&2&1\\-1&1&0&-5\\1&3&1&3\\0&-1&0&0\end{vmatrix}$$

$$\overset{\text{按第四行展开}}{=\!=\!=}-\begin{vmatrix}1&2&1\\-1&0&-5\\1&1&3\end{vmatrix}\overset{r_1-2r_3}{=\!=\!=}-\begin{vmatrix}-1&0&-5\\-1&0&5\\1&1&3\end{vmatrix}=0$$

例4　1、计算：$\begin{vmatrix}5&3&-1&2&0\\1&7&2&5&2\\0&-2&3&1&0\\0&-4&-1&4&0\\0&2&3&5&0\end{vmatrix}$

2、证明：$\begin{vmatrix}kc_1+a_1&ma_1+b_1&lb_1+c_1\\kc_2+a_2&ma_2+b_2&lb_2+c_2\\kc_3+a_3&ma_3+b_3&lb_3+c_3\end{vmatrix}=(klm+1)\begin{vmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\\a_3&b_3&c_3\end{vmatrix}$

拉普拉斯(Laplace)定理·行列式的乘法法则：

定义：在一个 n 阶行列式 D 中，任意选定 k 行或 k 列($1\le k\le n$)位于这些行和列的交叉点上的k^2个元素按照原来次序组成一个 k 级行列式 M，称为行列 D 的一个 k 级子式。当 $k<n$ 时，在 D 中划去这 k 行 k 列后余下的元素按照原来的次序组成的 n-k 级行列式M'，称为 k 级子式 M 的余子式。

从定义立刻看出，M 也是M'的余子式。所以 M 和M'可以称为 D 的一对互余的子式。

例 6：在四级行列式$\begin{vmatrix} 1 & 2 & 1 & 4 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{vmatrix}$中选定第一、三行，第二、四列得到：

$M=\begin{vmatrix} 2 & 4 \\ 0 & 1 \end{vmatrix}$，M 的余子式为：$M'=\begin{vmatrix} 0 & 2 \\ 0 & 1 \end{vmatrix}$

例 7：在五级行列式$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix}$中$\begin{vmatrix} a_{12} & a_{13} & a_{15} \\ a_{22} & a_{23} & a_{25} \\ a_{42} & a_{43} & a_{45} \end{vmatrix}$

$M'=\begin{vmatrix} a_{31} & a_{34} \\ a_{51} & a_{54} \end{vmatrix}$，两者是一对互余的子式。

定义 10：设 D 的 k 级子式 M 在 D 中所在的行、列指标分别是 $i_1, i_2, \cdots, \ i_k$； $j_1, j_2, \cdots, \ j_k$. 则 M 的余子式M'前面加上符号 $(-1)^{(i_1+i_2+i_k)+(j_1+j_2+j_k)}$后称作 M 的代数余子式。

例如，上述例 6 中 M 的代数余子式是：$(-1)^{(1+3)+(2+4)}M'=M'$；上述例 7 中 M 的代数余子式是：$(-1)^{(1+2+4)+(2+3+5)}M'=-M'$

因为 M 与M'位于行列式 D 中不同的行和不同的列，所以有下述：

引理：行列式 D 的任一子式 M 与它的代数余子式 A 的乘积中的每一项都是行列式 D 的展开式中的一项，而且符号也一致。

证明：首先考虑 M 位于行列式 D 的左上方（即第 1,2,…, k 行　和第 1,2,…,k 列）的情况。这时

$$D=\left|\begin{array}{cccc:ccc} a_{11} & a_{12} & \cdots & a_{1k} & a_{1k+1} & \cdots & a_{1n} \\ \vdots & M & & \cdots & \cdots & & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} & a_{kk+1} & \cdots & a_{kn} \\ \hdashline a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,k} & a_{k+1,k+1} & \cdots & a_{k+1,n} \\ \vdots & \vdots & & \vdots & \vdots & M' & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} & a_{n,k+1} & \cdots & a_{nn} \end{array}\right|$$

D 中 k 阶子式的余子式位于右下角，且其代数余子式为：

$$A=(-1)^{(1+2+\cdots k)+(1+2+\cdots k)}M'=M'$$

M 的每一项可写作$a_{1\alpha_1},a_{2\alpha_2},\cdots,a_{k\alpha_k}$,其中$\alpha_1$，$\alpha_2,\cdots,\alpha_k$是 1,2, …,k 的一个排列。所以这一项前面所带符号为：$(-\mathbf{1})^{t(\alpha_1\alpha_2\cdots\alpha_k)}$，对应的$M'$中每一项可写作：$a_{k+1,\beta_{k+1}},a_{k+2,\beta_{k+2}}\cdots,a_{n\beta_n}$，其中$\beta_{k+1}$，$\beta_{k+2}$，…，$\beta_n$，是$k+1$，$k+2,\cdots$，n的一个排列，这一项在$M'$中前面所带符号为：

$$(-\mathbf{1})^{t((\beta_{k+1}-k)(\beta_{k+2}-k)(\beta_n-k))}$$

这二项的乘积是：$\boldsymbol{a_{1\alpha}a_{2\alpha_2}\cdots a_{k\alpha_k}a_{k+1,\beta_{k+1}}\cdots a_{n\beta_n}}$，前面的符号是：$(-1)^{t(a_1a_2\cdots a_k)+t((\beta_{k+1}-k)(\beta_{k+2}-k)(\beta_n-k))}$。因为每个$\beta$比每个$\alpha$都大，所以上述符号等于：$(-1)^{t(\alpha_1\alpha_2\cdots\alpha_k\beta_{k+1}\cdots\beta_n)}$。因此这个乘积是行列式 D 中的一项而且符号相同。下面来证明一般情形：

设子式 M 位于 D 的第$i_1,i_2,\cdots$，i_k行；第$j_1,j_2,\cdots$，j_k列，这里

$$i_1<i_2<\cdots<i_k;\ j_1<j_2<\cdots<j_k$$

为了利用前面的结论，我们先把第i_1行依次与$i_1-1,i_1-2,\cdots,2,1$ 行对换。这样经过i_1-1次对换而将第i_1行换到第一行。再将i_2行依次与第$i_2-1,i_2-2,\cdots,2$ 行对换而换到第 2 行，共经i_2-2次对换，如此进行下去，一共经过$(i_1-1)+(i_2-2)+\cdots+(i_k-k)=(i_1+\cdots+i_k)-(1+2+\cdots+k)$次行对换把第$i_1,i_2,\cdots$，$i_k$行依次换到 1,2, …, k 行。

利用类似的列变换，可以将 M 的列换到 1,2, …, k 列。一共作了$(j_1-1)+(j_2-2)+\cdots+(j_k-k)=(j_1+\cdots+j_k)-(1+2+\cdots+k)$次列变换。

用D_1表示经上述行、列变换后得到的新行列式，那么

$$D_1=(-1)^{(i_1+\cdots+i_k)-(1+2+\cdots+k)+(j_1+\cdots+j_k)-(1+2+\cdots+k)}D$$
$$=(-1)^{(i_1+\cdots+i_k)+(j_1+\cdots+j_k)}D$$

由此可看出，D 和D_1的展开式中出现的项是一样的，只不过每一项都相差符号为$(-1)^{(i_1+\cdots+i_k)+(j_1+\cdots+j_k)}$

现在 M 位于D_1的左上角，所以 M$\cdot M'$中每一项都是D_1中的一项而且符号一致。但是$M\cdot A=(-1)^{i_1+\cdots+i_k+j_1+\cdots+j_k}M\cdot M'$,所以 M、A 中每一项都与 D 中的一项相等且符号一致。

定理 6(拉普拉斯定理): 设在行列式 D 中任意取定了$k(1\le k\le \mathrm{n}-1)$个行。由这 k 行元素所组成的一切 k 阶子式与它们的代数余子式的乘积的和等于行列式 D。

证明：设 D 中取定 k 行后所得的子式为$M_1,M_2,\cdots,\ M_t$,它 们的代数余子式分别为$A_1,A_2,\cdots,\ A_t$,定理要求证明$D=M_1A_1+M_2A_2+\cdots+M_tA_t$，由引理知，$M_iA_i$中的每一项都是 D 中一项而且符号相同，而且$M_iA_i$和$M_jA_j$(i≠j)无公共项。因此要证明（1）式成立，只要证明等式两边的项数相等就可以了。由定义知 D 中共有 n! 项，为了计算右边的项数，先算出 t 共有多少个。由组合公式知

$$t=C_n^k=\frac{n!}{k!\,(n-k)!}$$

因为M_i中共有 k! 项，A_i中共有(n-k)!项。所以右边共有$C_n^k\cdot k!\cdot(n-k)!=\mathrm{n}!$ 项。定理得证.

例 8：在行列式$D=\begin{vmatrix}1&2&1&4\\0&-1&2&1\\1&0&1&3\\0&1&3&1\end{vmatrix}$中取定第一、二行，求其子式和代数余子式.

解：$M_1=\begin{vmatrix}1&2\\0&-1\end{vmatrix}$，$M_2=\begin{vmatrix}1&1\\0&2\end{vmatrix}$，$M_3=\begin{vmatrix}1&4\\0&1\end{vmatrix}$，

$$M_4=\begin{vmatrix}2&1\\-1&2\end{vmatrix},M_5=\begin{vmatrix}2&4\\-1&1\end{vmatrix},M_6=\begin{vmatrix}1&4\\2&1\end{vmatrix}.$$

它们对应的代数余子式为：

$A_1=(-1)^{(1+2)+(1+2)}M_1{}' = M_1{}' \qquad A_2=(-1)^{(1+2)+(1+3)}M_2{}' = -M_2{}'$

$A_3=(-1)^{(1+2)+(1+4)}M_3{}' = M_3{}' \qquad A_4=(-1)^{(1+2)+(2+3)}M_4{}' = M_4{}'$

$A_5=(-1)^{(1+2)+(2+4)}M_5{}' = -M_5{}' \qquad A_6=(-1)^{(1+2)+(3+4)}M_6{}' = -M_6{}'$

根据拉普拉斯定理

$$D=\begin{vmatrix}1&2\\0&-1\end{vmatrix}\begin{vmatrix}1&3\\3&1\end{vmatrix}-\begin{vmatrix}1&1\\0&2\end{vmatrix}\begin{vmatrix}0&3\\1&1\end{vmatrix}+\begin{vmatrix}1&4\\0&1\end{vmatrix}\begin{vmatrix}0&1\\1&3\end{vmatrix}$$
$$+\begin{vmatrix}2&1\\-1&2\end{vmatrix}\begin{vmatrix}1&3\\0&1\end{vmatrix}-\begin{vmatrix}2&4\\-1&1\end{vmatrix}\begin{vmatrix}1&1\\0&3\end{vmatrix}+\begin{vmatrix}1&4\\2&1\end{vmatrix}\begin{vmatrix}1&0\\0&1\end{vmatrix}$$

= (-1)×(-8) − 2×(-3)+1×(-1)+ 5×1−6 × 3 +(-7) × 1

$=8+6-1+5-18-7$

$=-7$

定理 7 （行列式相乘规则）：两个 n 阶行列式

$$D_1=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}\text{和}D_2=\begin{vmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\b_{21}&b_{22}&\cdots&b_{2n}\\\cdots&\cdots&\cdots&\cdots\\b_{n1}&b_{n2}&\cdots&b_{nn}\end{vmatrix}$$

的乘积等于一个 n 阶行列式

$$C=\begin{vmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\cdots&\cdots&&\cdots\\c_{n1}&c_{n2}&\cdots&c_{nn}\end{vmatrix}$$

(其中c_{ij}是D_1的第 i 行元素分别与D_2的第 j 列的对应元素乘积之和：$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$)

证明：构造一个 2n 阶行列式

$$D_{2n} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & & \cdots & \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & & \cdots & \cdots & \cdots & & \cdots \\ 0 & 0 & \cdots & -1 & a_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}$$

根据拉普拉斯定理，将 D 按前 n 行展开。则因 D 中前 n 行除去左上角那个 n 级子式外，其余的 n 级子式都等于零。所以

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix} = D_1 D_2$$

现在来证 D=C。

对 D 作初等行变换。将第 n+1 行的a_{11}倍，第 n+2 行的a_{12}倍，⋯，第 2n 行的a_{1n}倍加到第一行，得

$$D = \begin{vmatrix} 0 & 0 & \cdots & 0 & c_{11} & c_{12} & \cdots & c_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}$$

再依次将第 n+1 行的a_{k1}(k=2,3,⋯,n)倍，第 n+2 行的a_{k2}倍， ⋯,第 2n 行的a_{kn}倍加到第 k 行，就得：

$$D=\begin{vmatrix}0&0&\cdots&0&c_{11}&c_{12}&\cdots&c_{1n}\\0&0&\cdots&0&c_{21}&c_{22}&\cdots&c_{2n}\\\vdots&\vdots&&\vdots&\vdots&\vdots&&\vdots\\0&0&\cdots&0&c_{\mathrm{n}1}&c_{\mathrm{n}2}&\cdots&c_{\mathrm{nn}}\\-1&0&\cdots&0&b_{11}&b_{12}&\cdots&b_{1n}\\0&-1&\cdots&0&b_{21}&b_{22}&\cdots&b_{2n}\\\vdots&\vdots&&\vdots&\vdots&\vdots&&\vdots\\0&0&\cdots&-1&b_{n1}&b_{n2}&\cdots&b_{nn}\end{vmatrix}$$

这个行列式的前 n 行也只可能有一个 n 级子式不为零，因此由拉普拉斯定理:

$$C=\begin{vmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\cdots&\cdots&&\cdots\\c_{n1}&c_{n2}&\cdots&c_{nn}\end{vmatrix}.$$

$$(-1)^{(1+2+\cdots+n)+(n+1+n+2+\cdots+2n)}\begin{vmatrix}-1&0&\cdots&0\\0&-1&\cdots&0\\\cdots&\cdots&&\cdots\\0&0&\cdots&-1\end{vmatrix}=C$$

定理得证。

上述定理也称为行列式的乘法定理。它的意义在第四章会进一步阐述。

习题 1：计算下列行列式：$D_4=$

$$\begin{vmatrix}1&1&1&1\\1+\sin\varphi_1&1+\sin\varphi_2&1+\sin\varphi_3&1+\sin\varphi_4\\\sin\varphi_1+\sin^2\varphi_1&\sin\varphi_2+\sin^2\varphi_2&\sin\varphi_3+\sin^2\varphi_3&\sin\varphi_4+\sin^2\varphi_4\\\sin^2\varphi_1+\sin^3\varphi_1&\sin^2\varphi_2+\sin^3\varphi_2&\sin^2\varphi_3+\sin^3\varphi_3&\sin^2\varphi_4+\sin^3\varphi_4\end{vmatrix}$$

解：在D_4的第 2 行中去掉与第 1 行成比例的分行，得到

$$\begin{vmatrix}1&1&1&1\\\sin\varphi_1&\sin\varphi_2&\sin\varphi_3&\sin\varphi_4\\\sin\varphi_1+\sin^2\varphi_1&\sin\varphi_2+\sin^2\varphi_2&\sin\varphi_3+\sin^2\varphi_3&\sin\varphi_4+\sin^2\varphi_4\\\sin^2\varphi_1+\sin^3\varphi_1&\sin^2\varphi_2+\sin^3\varphi_2&\sin^2\varphi_3+\sin^3\varphi_3&\sin^2\varphi_4+\sin^3\varphi_4\end{vmatrix}$$

在上行列式的第 3 行中去掉与第 2 行成比例的分行，得一新行列式，在此新行列式的第 4 行去掉与第 3 行成比例的分行，得到

$$D_4 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ \sin\varphi_1 & \sin\varphi_2 & \sin\varphi_3 & \sin\varphi_4 \\ \sin^2\varphi_1 & \sin^2\varphi_2 & \sin^2\varphi_3 & \sin^2\varphi_4 \\ \sin^3\varphi_1 & \sin^3\varphi_2 & \sin^3\varphi_3 & \sin^3\varphi_4 \end{vmatrix}$$

习题 2：证明：$\begin{vmatrix} 1+x_1y_1 & 1+x_1y_2 & 1+x_1y_3 \\ 1+x_2y_1 & 1+x_2y_2 & 1+x_3y_3 \\ 1+x_3y_1 & 1+x_2y_2 & 1+x_3y_3 \end{vmatrix} = 0$

提示：若没有常数 1，则至少两行成比例，故考虑消去常数 1

习题 3:证明 $\begin{vmatrix} 1+x_1^2 & x_1x_2 & x_1x_3 & x_1x_4 \\ x_2x_1 & 1+x_2^2 & x_2x_3 & x_2x_4 \\ x_3x_1 & x_3x_2 & 1+x_3^2 & x_3x_4 \\ x_4x_1 & x_4x_2 & x_4x_3 & 1+x_4^2 \end{vmatrix} = 1 + \sum_{i=1}^{4} x_i^2$

证：除主对角线之外，各行之间构成比例，加一行一列

$\begin{vmatrix} 1 & x_1 & x_2 & x_3 & x_4 \\ 0 & 1+x_1^2 & x_1x_2 & x_1x_3 & x_1x_4 \\ 0 & x_2x_1 & 1+x_2^2 & x_2x_3 & x_2x_4 \\ 0 & x_3x_1 & x_3x_2 & 1+x_3^2 & x_3x_4 \\ 0 & x_4x_1 & x_4x_2 & x_4x_3 & 1+x_4^2 \end{vmatrix}$再化成爪形

习题 4：计算：$D_n = \begin{vmatrix} 2 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 2 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 2 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 2 \end{vmatrix}$

$$D_n \xlongequal{r_2+(-1/2)r_1} \begin{vmatrix} 2 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 3/2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 2 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 2 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 2 \end{vmatrix}$$

$$\xlongequal{r_3+(-2/3)r_2}\begin{vmatrix}2&1&0&0&0&\cdots&0\\0&3/2&1&0&0&\cdots&0\\0&1&4/3&1&0&\cdots&0\\0&0&1&2&1&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&0&0&\cdots&2\end{vmatrix}$$

$$\xlongequal{r_4+(-3/4)r_3}\begin{vmatrix}2&1&0&0&\cdots&0\\0&3/2&1&0&\cdots&0\\0&0&4/3&1&\cdots&0\\0&0&0&5/4&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&0&\cdots&2\end{vmatrix}=\cdots$$

$$=\begin{vmatrix}2&1&0&0&\cdots&0\\0&3/2&1&0&\cdots&0\\0&0&4/3&1&\cdots&0\\0&0&0&5/4&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&0&\cdots&(n+1)/n\end{vmatrix}=n+1$$

习题 5， 计算$\begin{vmatrix}a_0&1&1&\cdots&1\\1&a_1&0&\cdots&0\\1&0&a_2&\cdots&0\\\vdots&\vdots&\vdots&\cdots&\vdots\\1&0&0&\cdots&a_n\end{vmatrix}(a_i\neq 0)$

解：可用列变换将第一列全变成零　$c_1-\frac{1}{a}c_2$，$-\frac{1}{a}c_2$，$c_1-\frac{1}{a}c_2\cdots$

$$a_1a_2\cdots a_n\left(a_0-\sum_{i=1}^{n}\frac{1}{a_i}\right)$$

习题 6， 计算$\begin{vmatrix}1+a_1&1&\cdots&1\\1&1+a_2&\cdots&1\\\vdots&\vdots&\cdots&\vdots\\1&1&\cdots&1+a_n\end{vmatrix}(a_i\neq 0)$

解：除主对角线外，其余元素全相同，可化为爪形！

$r_2-r_1,r_3-r_1,\cdots,r_n-r_1$即变成爪形

习题 7： 证明

$$\begin{vmatrix} a_0 & -1 & 0 & \cdots & 0 & 0 \\ a_1 & x & -1 & \cdots & 0 & 0 \\ a_2 & 0 & x & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n-2} & 0 & 0 & \cdots & x & -1 \\ a_{n-1} & 0 & 0 & \cdots & 0 & x \end{vmatrix} = \sum_{i=0}^{n-1} a_i x^{n-i-1}$$

解：

$$D_n=\begin{vmatrix} a_0 & -1 & 0 & \cdots & 0 & 0 \\ a_0x+a_1 & 0 & -1 & \cdots & 0 & 0 \\ a_0x^2+a_1x+a_2 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_0x^{n-2}+a_1x^{n-3}+\cdots+a_{n-2} & 0 & 0 & \cdots & 0 & -1 \\ a_0x^{n-1}+a_1x^{n-2}+\cdots+a_{n-2}x+a_{n-1} & 0 & 0 & \cdots & 0 & 0 \end{vmatrix}$$

$$= (-1)^{n-1}\begin{vmatrix} \sum_{i=0}^{n-1} a_i x^{n-i-1} & 0 & 0 & \cdots & 0 & 0 \\ a_0 & -1 & 0 & \cdots & 0 & 0 \\ a_0x+a_1 & 0 & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum_{i=0}^{n-2} a_i x^{n-2-i} & 0 & 0 & \cdots & 0 & -1 \end{vmatrix} = \sum_{i=0}^{n-1} a_i x^{n-i-1}$$

习题 8 计算: $\begin{vmatrix} x_1-m & x_2 & \cdots & x_n \\ x_1 & x_2-m & \cdots & x_n \\ \vdots & \vdots & \vdots & \vdots \\ x_1 & x_2 & \cdots & x_n-m \end{vmatrix}$

解：特点：每一行的和都相等，故可全加到第一列

$$D = \left(\sum x_i - m\right)\begin{vmatrix} 1 & x_2 & \cdots & x_n \\ 1 & x_2-m & \cdots & x_n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_2 & \cdots & x_n-m \end{vmatrix}$$

$$= \left(\sum x_i - m\right)\begin{vmatrix} 1 & x_2 & \cdots & x_n \\ 0 & -m & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -m \end{vmatrix}$$

习题 9：计算：$\begin{vmatrix} a_1 & a_2 & a_2 & \cdots & a_{n-1} & a_n \\ -x & x & 0 & \cdots & 0 & 0 \\ 0 & -x & x & \cdots & 0 & 0 \\ 0 & 0 & -x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -x & x \end{vmatrix}$

作业：1、$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ b+c & c+a & a+b \end{vmatrix}$ $\qquad$ $\begin{vmatrix} a & b & c & d \\ a & d & c & b \\ c & d & a & b \\ c & b & a & d \end{vmatrix}$

$$\begin{vmatrix} x+y & z+y & z+x \\ z+y & z+x & x+y \\ z+x & x+y & z+y \end{vmatrix}$$

2、若$\begin{vmatrix} k & 3 & 4 \\ -1 & k & 0 \\ 0 & k & 1 \end{vmatrix} = 0$，求 k 的值

3、 计算$\begin{vmatrix} a_1 & a_2 & a_2 & \cdots & a_{n-1} & a_n \\ -x & x & 0 & \cdots & 0 & 0 \\ 0 & -x & x & \cdots & 0 & 0 \\ 0 & 0 & -x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -x & x \end{vmatrix}$

4、计算 n 阶行列式$\begin{vmatrix} \alpha+\beta & \alpha & & 0 \\ \beta & \alpha+\beta & & \\ & & & \alpha \\ 0 & & \beta & \alpha+\beta \end{vmatrix}$

常用方法小结：

1、逐步用性质化为三角形行列式；

2、零元多的，可用行列式展开降阶；

3、分块的，可化为行列式之积；

4、递推法，或者由低阶猜测形式，用数学归纳法证明；

5、加边法（每行或者每列的和为 1）

6、利用特殊行列式，例如范德蒙行列式

7、利用定义，或者综合各种方法。

§2.7　克拉默法则

本节讨论 n 个未知数 n 个方程的线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots a_{2n}x_n=b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots a_{nn}x_n=b_n \end{cases} \qquad (*)$$

的求解问题

行列式$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$称为方程组$(*)$的系数行列式.

克拉默法则

如果线性方程组$(*)$的系数行列式 D 不等于零, 则方程组$(*)$有唯一解

$$x_1=\frac{D_1}{D},\ x_2=\frac{D_2}{D},\ \cdots x_n=\frac{D_n}{D} \qquad (2.7.1)$$

其中$D_j(j=1,2,\cdots,n)$是把系数行列式 D 中第 j 列的元素$a_{1j},a_{2j},\cdots,a_{nj}$对应地换为方程组的常数项$b_1,b_2,\cdots,b_n$后所得到的 n 阶行列式.

例 1 解线性方程组$\begin{cases} 2x_1+x_2-5x_3+x_4=8 \\ x_1-3x_2 \qquad -6x_4=9 \\ \qquad 2x_2-x_3+2x_4=-5 \\ x_1+4x_2-7x_3+6x_4=0 \end{cases}$

解　因为 D=27, $D_1=812=108, D_3=-27, D_4=27$，所以，所给方程组的唯一解为

$$x_1=\frac{D_1}{D}=3\,, x_2=\frac{D_2}{D}=-4\ \ , x_3=\frac{D_3}{D}=-1\ \ ,\ x_4=\frac{D_1}{D}=1$$

例 2　设曲线 $y=a_0+a_1x+a_2x^2+a_3x^3$ 通过四点(1, 3)、(2, 4)、(3, 3)、(4, –3)，求系数 a_0, a_1, a_2, a_3 .

解 把四个点的坐标代入曲线方程，得线性方程组

$$\begin{cases} a_0+a_1+a_2+a_3=3 \\ a_0+2a_1+4a_2+8a_3=4 \\ a_0+3a_1+9a_2+27a_3=3 \\ a_0+4a_1+16a_2+64a_3=-3 \end{cases}$$

因为 $D=12, D_1=36, D_2=-18, D_3=24, D_4=-6$，所以方程有唯一解

$$a_0=3, a_1=-\frac{3}{2}, a_2=2, a_3=-\frac{1}{2}$$

即曲线方程为:　　$\mathrm{y}=3-\frac{3}{2}x+2x^2-\frac{1}{2}x^3$

定理 4

如果线性方程组(＊)的系数行列式 D≠0，则方程组(＊)一定有解，且解是唯一的.

定理 4′

如果线性方程组(＊)无解或有两个不同的解，则它的系数行列式必为零.

讨论：常数项均为零的线性方程组称为齐次线性方程组，问齐次线性方程组有什么样的解?

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots a_{2n}x_n=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots a_{nn}x_n=0 \end{cases} (**),\ D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

定理 5

如果齐次线性方程组(＊＊)的系数行列式 $D\neq 0$，则齐次线性方程组(＊＊)没有非零解.

定理 5′

如果齐次线性方程组(＊＊)有非零解，则它的系数行列式必为零.

例 3　问λ取何值时，齐次线性方程组

$$\begin{cases} (5-\lambda)x+2y+2z=0 \\ 2x+(6-\lambda)y=0 \\ 2x+(4-\lambda)z=0 \end{cases}$$

有非零解?

解：若所给齐次线性方程组有非零解，则其系数行列式 D=0. 而

$$D=\begin{vmatrix} 5-\lambda & 2 & 2 \\ 2 & 6-\lambda & 0 \\ 2 & 0 & 4-\lambda \end{vmatrix}$$

$$=(5-\lambda)(6-\lambda)(4-\lambda)-4(4-\lambda)-4(6-\lambda)$$

$$=(5-\lambda)(2-\lambda)(8-\lambda)$$

由 D=0, 得λ=2、λ=5 或λ=8.当λ=2、λ=5 或λ=8 时，齐次线性方程组有非零解.

克拉默法则求解的局限性：

1、方程个数必须等于未知元的个数

2、必须$D \neq 0$

3、计算量大

第三章　　矩阵

矩阵概念

矩阵的运算

可逆矩阵

分块矩阵

矩阵的初等变换

矩阵的秩

§3.1 矩阵概念

引例

一个线性方程组与一个数表存在一一对应关系

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

这个数表就称为矩阵.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

3.1.1 矩阵的定义

数域 F 中，有$\mathrm{m} \times \mathrm{n}$个数$a_{ij}(i = 1,2,\cdots,m;\ \mathrm{j} = 1,2,\cdots,\mathrm{n})$排成 m 行 n 列的矩形数表称为$\mathrm{m} \times \mathrm{n}$矩阵，记作

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad (3.1)$$

其中 a_{ij}称为矩阵的第 i 行第 j 列的元素.

一般情况下，我们用大写字母 A，B，C 等表示矩阵. 矩阵 A 简记为$\mathrm{A} = (a_{ij})_{\mathrm{m}\times\mathrm{n}}$或 $\mathrm{A}_{\mathrm{m}\times\mathrm{n}}$元素属于实数域的矩阵叫实矩阵，复数域的叫复矩阵.

方阵　行矩阵　列矩阵

n 阶方阵：m=n

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1\mathrm{n}} \\ a_{21} & a_{22} & \cdots & a_{2\mathrm{n}} \\ \cdots & \cdots & \cdots & \cdots \\ a_{\mathrm{m}1} & a_{\mathrm{m}2} & \cdots & a_{\mathrm{mn}} \end{pmatrix}.$$

行矩阵：m=1

$$(a_1 \ a_2 \ \cdots \ a_{\mathrm{n}}).$$

列矩阵：n=1

$$\begin{pmatrix} \mathrm{b}_1 \\ \mathrm{b}_2 \\ \vdots \\ \mathrm{b}_{\mathrm{n}} \end{pmatrix}.$$

3.1.2 几种特殊的矩阵

1、零矩阵: $a_{ij}=0,\ i=1\cdots m, j=1\cdots n.$

$$\mathrm{O}=O_{m\times n}=\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

$$\mathrm{O}_{1\times3}=(0 \quad 0 \quad 0)，\ \mathrm{O}_{2\times4}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}，\ \mathrm{O}_{3\times2}=\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

2、对角矩阵: $\mathrm{m}=\mathrm{n},\ a_{ij}$不全为$0, a_{ij}=0,\ i$不等于$j$.

$$A_{n\times n}=\begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}.$$

3、单位矩阵

$\mathrm{m}=\mathrm{n},\ a_{ij}=1, a_{ij}=0,\ i$不等于$j$.

$$I=I_n=\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

4、三角形矩阵

$\mathrm{m}=\mathrm{n}, \forall \mathrm{i}<\mathrm{j}, a_{ij}=0, \forall \mathrm{i}\geq \mathrm{j}, a_{ij}$不全为 0.

$$A_{n\times n}=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ & a_{22} & \cdots & a_{2n}\\ & & \ddots & \\ & & & a_{nn}\end{pmatrix}.$$

$\mathrm{m}=\mathrm{n},\forall \mathrm{i}>\mathrm{j},a_{ij}=0,\forall \mathrm{i}\le \mathrm{j},a_{ij}$不全为 0.

$$B_{n\times n}=\begin{pmatrix}b_{11} & & & \\ b_{12} & b_{22} & & \\ \vdots & \vdots & \ddots & \\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{pmatrix}.$$

5、对称矩阵

n 阶方阵$A=(a_{ij})$满足$a_{ij}=a_{ji}(i,j=1,\cdots,n)$

$$\begin{pmatrix}1 & 0 & 6\\ 0 & 2 & -1\\ 6 & -1 & 3\end{pmatrix},\quad \begin{pmatrix}0 & -2\\ -2 & 0\end{pmatrix}.$$

6、反对称矩阵

n 阶方阵$A=(a_{ij})$满足$a_{ij}=-a_{ji}(i,j=1,\cdots,n)$

$$\begin{pmatrix}0 & 3 & -6\\ -3 & 0 & -1\\ 6 & 1 & 0\end{pmatrix},\quad \begin{pmatrix}0 & -2\\ 2 & 0\end{pmatrix}.$$

反对称矩阵主对角线元素全为 0.

7、行矩阵和列矩阵

$\mathrm{m}=1$，即A中只有一行的矩阵

$$A=[a_1 \quad a_2 \quad \cdots \quad a_n].$$

$\mathrm{n}=1$，即A中只有一列的矩阵

$$\mathrm{A}=\begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_m\end{bmatrix}.$$

3.1.3 矩阵与行列式区别

1、一个是值，一个是表.

2、行列式的 m=n，矩阵的行不一定于列，m，n.

3、当 m = n 时,可以定义方阵对应的行列式

$$A=\left(a_{ij}\right)_{n\times n}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \cdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \qquad (3.2)$$

$$|A|=|a_{ij}|=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \qquad (3.3)$$

4、当 m = n 时，可以定义方阵的迹，Tr(A).

$Tr(A)=\sum_i a_{ij}$(主对角元之和)

例 1 $\begin{bmatrix} 2 & 1 & 3 & 5 \\ -5 & 6 & 2 & 7 \end{bmatrix}$：2×4 实矩阵.

$\begin{bmatrix} 10 & 6i & 2 \\ -1 & 2 & 5 \\ 2 & 3 & 1 \end{bmatrix}$：3×3 实矩阵.

$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$：3×1 实矩阵.

$\begin{bmatrix} 2 & 3 & 5 & 9 \end{bmatrix}$：1×4 实矩阵

作业 1、试写出 4×5 矩阵 A，其元素$a_{ij}=2i-j$.

§3.2 矩阵的运算

3.2.1 同型矩阵

两个矩阵$A_{m\times n}$和 $B_{l\times k}$ 的行数和列数相等，即 m=l，n=k.

3.2.2 矩阵相等

$$A=(a_{ij})_{m\times n},B=(b_{ij})_{m\times n},\text{且}a_{ij}=b_{ij}\rightarrow A=B,$$

$$(i=1,\cdots,m;j=1,\cdots,n).$$

3.2.3 矩阵加法

$$A=(a_{ij})_{m\times n},B=(b_{ij})_{m\times n}\rightarrow A+B=(a_{ij}+b_{ij})_{m\times n}.$$

同型的两个矩阵相加等于其所有对应的元素相加.

$$\begin{aligned}A+B&=(a_{ij}+b_{ij})_{m\times n}\\&=\begin{pmatrix}a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n}\\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn}\end{pmatrix}.\end{aligned} \qquad (3.4)$$

矩阵加法的性质：

A,B,C,O均为$m\times n$矩阵

1. $A+B=B+A$
2. $(A+B)+C=A+(B+C)$
3. $A+O=O+A=A$
4. $A+(-A)=(-A)+A=O$
5. $A-B=A+(-B)=(a_{ij}-b_{ij})_{m\times n}$

> $-A$ 是 A 的负阵，
>
> $-A=(-a_{ij})_{m\times n}$).

3.2.4 矩阵数乘

$A=\left(a_{ij}\right)_{mn}, k\in F$（数域），$kA=\left(\mathrm{k}a_{ij}\right)_{mn}$.

$$kA=\begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}.\text{（3.5）}$$

数乘性质：$k,l\in F$

1. $(k+l)A=kA+lA$

2. $k(A+B)=kA+kB$

3. $k(lA)=(kl)A$

4. $1\cdot A=A$

数量矩阵： $kI_n=\begin{pmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{pmatrix}.$ （3.6）

3.2.5 矩阵乘法

$A=(a_{ik})_{m\times s}$，$B=\left(b_{kj}\right)_{s\times n}$,则A与B乘积的矩阵$C=\left(c_{ij}\right)_{m\times n}$,

其中$c_{ij}=\sum_{k=1}^{s}a_{ik}\,b_{kj}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{is}b_{sj}$,

$$(i=1,2,\cdots,m;j=1,2,\cdots,n).$$

记 $\mathrm{C}=\mathrm{AB}$

$$i\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{11} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{11} & \cdots & a_{is} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{11} & \cdots & a_{ms} \end{pmatrix}\begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & & \vdots & & \vdots \\ b_{s1} & \cdots & b_{sj} & \cdots & b_{sn} \end{pmatrix}$$

$$=\begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1n} \\ \vdots & \vdots & & \vdots & \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{in} \\ \vdots & \vdots & & \vdots & \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mn} \end{pmatrix}j$$

A的第i个行向量和B的第j个列向量做“点积”.

A的列数= B的行数.

例 2

$$A_{3\times4}=\begin{pmatrix}1&0&-1&2\\-1&1&3&0\\0&5&-1&4\end{pmatrix},B_{4\times3}=\begin{pmatrix}0&3&4\\1&2&1\\3&1&-1\\-1&2&1\end{pmatrix},\text{求}AB.$$

解：

$$C_{3\times3}=AB=\begin{pmatrix}1&0&-1&2\\-1&1&3&0\\0&5&-1&4\end{pmatrix}\begin{pmatrix}0&3&4\\1&2&1\\3&1&-1\\-1&2&1\end{pmatrix}=\begin{pmatrix}-5&6&7\\10&2&-6\\-2&17&10\end{pmatrix}$$

例 3 在线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m\end{cases}\qquad (3.7)\text{中令}$$

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix},X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},B=\begin{pmatrix}b_1\\b_2\\\vdots\\b_m\end{pmatrix}$$

则方程组可以表示为矩阵形式$AX=B$.

例 4 设$A=[a_1\quad a_2\quad\cdots\quad a_n]$，$B=\begin{bmatrix}b_1\\b_2\\\vdots\\b_n\end{bmatrix}$.

则$AB=\sum_{i=1}^{n}a_i\,b_i$ AB 是 1×1矩阵，即一个数.

BA =?

例 5 $\mathrm{A}=\begin{bmatrix}2&-1\\1&0\\-3&4\end{bmatrix}$，$\mathrm{B}=\begin{bmatrix}1&-2&-5\\3&4&0\end{bmatrix}$.

则$\mathrm{AB}=?$　　$\mathrm{BA}=?$　　若$\mathrm{B}=\begin{bmatrix}1 & -2\\ 3 & 4\end{bmatrix}$呢？

问题：单位矩阵左乘或右乘 A？对角矩阵 diag$[\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_n]$.

左乘或右乘 A？（记住结论！）

注意：

1、只有A的列数与B的行数相等时，AB才有意义.

（1）　$A_{34}\cdot B_{45}=C_{35}$　　$B_{45}\cdot A_{34}$无意义

（2）　$A_{34}\cdot B_{43}=C_{35}$　　$B_{43}\cdot A_{34}=C_{44}$　　阶数不同

2、矩阵乘法不满足交换律，即$\mathrm{AB}\neq\mathrm{BA}$.

$$A=\begin{pmatrix}1 & 1\\ -1 & -1\end{pmatrix}, B=\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}$$

$$AB=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix} \neq BA=\begin{pmatrix}2 & 2\\ -2 & -2\end{pmatrix}$$

若$AB=BA$，称A与B可交换.

3、两个不为零的矩阵的乘积可以是零矩阵.

从而　$AB=0 \nRightarrow A=0$ 或$B=0$

4、矩阵乘法不满足消去律，即

$$AB=AC \nRightarrow B=C$$

如上例

$$A=\begin{pmatrix}1 & 1\\ -1 & -1\end{pmatrix},\ B=\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix},\ AB=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

取

$$C=\begin{pmatrix}-2 & -2\\ 2 & 2\end{pmatrix},\quad AC=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},\quad 显然B\neq C$$

一般来讲，矩阵乘法没有定义与之对应的逆运算，即“除法”.

（以后章节学习矩阵在什么条件下可逆）

乘法性质：设下列矩阵都可以进行有关运算.

1、$A(BC)=(AB)C$

2、$A(B+C)=AB+AC$

$(B+C)A=BA+CA$

3、$I_nA_{np}=A_{np}$，$A_{np}I_p=A_{np}$

4、$k(AB)=(kA)B=A(kB)$ $\qquad kI_n=\begin{pmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{pmatrix}.$

性质 4 的特殊情形$A_{n\times n}$,有$kA=(kI_n)A=A(kI_n)$.

n 阶数量矩阵与 n 阶方阵作乘法是可交换的.

3.2.6 矩阵转置

$$A=\left(a_{ij}\right)_{m\times n}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad (3.8)$$

$$A^T=\left(a_{ji}\right)_{n\times m}=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix} \qquad (3.9)$$

设$X=(x_1 \quad x_2 \quad \cdots \quad x_n)$

则有

$$XX^T=(x_1 \quad x_2 \quad \cdots \quad x_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}={x_1}^2+{x_2}^2+\cdots+{x_n}^2$$

$$X^TX=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}(x_1 \quad x_2 \quad \cdots \quad x_n)=\begin{pmatrix} {x_1}^2 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & {x_2}^2 & \cdots & x_2x_n \\ \cdots & \cdots & \cdots & \cdots \\ x_nx_1 & x_nx_2 & \cdots & {x_n}^2 \end{pmatrix}$$

对称矩阵：$A^T=A.$

反对称矩阵：$A^T=-A.$

转置矩阵的性质：

1、$(A^T)^T = A$

2、$(A+B)^T = A^T + B^T$

3、$(kA)^T = kA^T$

4、$(AB)^T = B^T A^T$

4 式可以推广到一般情况：$(A_1A_2A_3\cdots A_n)^T = {A_n}^T{A_{n-1}}^T\cdots {A_1}^T.$

作业

1、证B^TB和BB^T均是对称矩阵.

2、设$A=\left(a_{ij}\right)_{n\times n}, X=[x_1 \quad x_2 \quad \cdots \quad x_n]^T$，求$X^TAX.$

3、$A^3-B^3=(A-B)(A^2+AB+B^2)$是否成立？

4、设 A 是 n 阶方阵，求证：

（1）$\mathrm{A}+\mathrm{A}^{\mathrm{T}}$是对称矩阵；$\mathrm{A}-\mathrm{A}^{\mathrm{T}}$是反对称矩阵.

（2）A可表示为对称矩阵元和反对称矩阵之和.

5、设 X 是 n 阶行向量，A 是 n 阶方阵，若对任意的 X 均有$\mathrm{X}^{\mathrm{T}}\mathrm{AX}=0$，则 A 满足什么性质？

例 6 求矩阵 X，使$3\begin{pmatrix}1 & -2 & 1\\ 4 & 1 & 0\end{pmatrix}+2X^T=\begin{pmatrix}2\\ -1\end{pmatrix}(1 \quad -2 \quad 3).$

解：$2X^T=\begin{pmatrix}2\\ -1\end{pmatrix}(1 \quad -2 \quad 3)-3\begin{pmatrix}1 & -2 & 1\\ 4 & 1 & 0\end{pmatrix}$

$$=\begin{pmatrix}2 & -4 & 6\\ -1 & 2 & -3\end{pmatrix}-\begin{pmatrix}3 & -6 & 3\\ 12 & 3 & 0\end{pmatrix}$$

$$=\begin{pmatrix}-1 & 2 & 3\\ -13 & -1 & -3\end{pmatrix}.$$

所以 $$x=\frac{1}{2}\begin{pmatrix}-1 & 2 & 3\\ -13 & -1 & -3\end{pmatrix}^T=\begin{pmatrix}-\frac{1}{2} & -\frac{13}{2}\\ 1 & -\frac{1}{2}\\ \frac{3}{2} & -\frac{3}{2}\end{pmatrix}.$$

例 7 求与$A=\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}$可交换的矩阵.

解：设$B=\begin{pmatrix}x & y\\ z & u\end{pmatrix}$与 A 可交换，

由$AB=BA$, 即$\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}\begin{pmatrix}x & y\\ z & u\end{pmatrix}=\begin{pmatrix}x & y\\ z & u\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}$

得$\begin{cases} x+z=x \\ y+u=x+2y \\ 2z=z \\ 2u=z+2u \end{cases} \Rightarrow \begin{cases} z=0 \\ u=x+y \end{cases}$

所以$B=\begin{pmatrix}x & y\\ 0 & x+y\end{pmatrix}$ x, y 是任意数

例 8 设 A，B 为同阶对称矩阵，则 AB 为对称矩阵的充分必要条件是 AB = BA.

证明： “⇐”（充分）

因为 $A^T=A$，$B^T=B$

又 $AB=BA$

则有 $(AB)^T=B^TA^T=BA=AB$

“⇒”（必要）

因为 $A^T=A$，$B^T=B$

若 $(AB)^T=AB$

则有 $AB=(AB)^T=B^TA^T=BA$.

$\forall A_{m\times n}$，A^TA和AA^T都是对称矩阵.

3.2.7 方阵的幂

A 为 n 阶方阵：$\underbrace{A^k=AA\cdots A}_{k}$

1、 $A^0=I_n$

2、 $A^kA^l=A^{k+l}$

3、 $(A^k)^l=A^{kl}$

4、 设$f(x)=a_0+a_1x+\cdots+a_mA^m$

定义矩阵多项式 $f(A)=a_0I+a_1A+\cdots+a_mA^m$

$$a_0, a_1, \cdots, a_m 为常系数$$

3.2.8 方阵行列式运算

A，B 为 n 阶方阵

1、 $|A^T| = |A|^T = |A|$

2、 $|A^k| = |A|^k$

3、 $|AB| = |A| \cdot |B| = |AB|$

4、 $|kA| = k^n|A|$

5、 $||A| \cdot B| = |A|^n \cdot |B|$

补充：设 A，B 均为同阶方阵，方阵的迹满足：

（1）$: Tr(A^T) = Tr(A)$

（2）$: Tr(AB) = Tr(BA).$

推论：（1）$Tr((A_1A_2\cdots A_n)^T) = Tr\left({A_n}^T \cdots {A_2}^T {A_1}^T\right)$

（2） $Tr(A_1A_2\cdots A_n) = Tr(\ A_nA_1A_2\cdots A_{n-1})$

$$= Tr(\ A_2\cdots A_nA_1)$$

注意：$Tr(A_1A_2\cdots A_n) \neq Tr(\ A_nA_{n-1}A_{n-2}\cdots A_1).$

例 9 设$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$，$n \in N$, 求$A^n$.

解：方法 1：数学归纳法.

方法 2：设$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = I + B$

易算$B^2 = 0$, 所以$B^k = 0 (k \geq 2)$, 且 B 与 I 可交换.

由二项式定理

$$A^n = (I + B)^n = I^n + C_n^1 I^{n-1} B + C_n^2 I^{n-2} B^2 + \cdots + B^n$$

$$= I + nB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 2n \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2n \\ 0 & 1 \end{pmatrix}.$$

§ 3.3 可逆矩阵

1、在各种数域（例如 Q、R、C 等）中，可进行加、减、乘、除（除数非 0 时）运算.实际上，减法是通过加法和负元来定义的.

2、对矩阵来说，我们可以定义加法、负元（从而定义减法）和乘法.那么是否也可以定义矩阵的除法呢？这取决于是否可以定义类似数的乘法的逆元（倒数）？

3、由于单位矩阵乘以任意矩阵 A 还等于 A，因此单位矩阵 I 相当于数的乘法中的数字，因此矩阵 A 的逆元应满足条件：

$$A^{-1}A = AA^{-1} = I$$

我们将看到，与数不同，并不是每一个非零的矩阵都可逆！

定义（可逆矩阵） 设 A 是 n 阶方阵，若存在一个方阵 B，使 $AB = BA = I_n$，则称 A 为可逆矩阵，或非奇异矩阵，并称 B 为 A 的逆矩阵；否则称 A 为不可逆矩阵，或奇异矩阵.

例 $A = \begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix}$，$B = \begin{bmatrix} -1 & 3/2 \\ 1 & -1 \end{bmatrix}$，则$\begin{cases} AB = \\ BA = \end{cases}$

几个问题:

矩阵可逆的充要条件是什么？

1、矩阵可逆的充要条件是什么？

2、逆矩阵是否唯一？

3、怎样求矩阵的逆？

定理 3.3.1（逆矩阵的唯一性） 若 n 阶方阵 A 可逆，则逆矩阵唯一.

证明：设 A 有两个逆矩阵，B 和 C 则：

$$AB = BA = I_n, AC = CA = I_n \quad \Rightarrow \quad B = BI_n = B(AC) = I_nC = C.$$

基于以上定理，可将 A 的唯一性的逆矩阵记作A^{-1}，

$$AA^{-1} = A^{-1}A = I_n.$$

注意：并不是所有的非零矩阵均有逆矩阵，例如：

$A=\begin{pmatrix}1 & 2\\ 0 & 0\end{pmatrix}$对任意的矩阵 B= $\begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{pmatrix}$,

有$AB=\begin{pmatrix}b_{11}+2b_{21} & b_{12}+2b_{22}\\ 0 & 0\end{pmatrix}$均不可能等于单位矩阵.

逆矩阵的求法之一（待定系数法）

例 10 求$A=\begin{pmatrix}2 & 1\\ -1 & 0\end{pmatrix}$的逆矩阵.

解：设$B=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$是 A 的逆矩阵，则由$AB=I_n$得：

$$AB=\begin{pmatrix}2 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\begin{pmatrix}2a+c & 2b+d\\ -c & -b\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

$$\Rightarrow\begin{cases}2a+c=1\\ 2b+d=0\\ -a=0\\ -b=1\end{cases}\Rightarrow\begin{cases}a=0\\ b=-1\\ c=1\\ d=2\end{cases}\Rightarrow B=\begin{pmatrix}0 & -1\\ 1 & 2\end{pmatrix}$$

又$BA=\begin{pmatrix}0 & -1\\ 1 & 2\end{pmatrix}\begin{pmatrix}2 & 1\\ -1 & 0\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$

因此 B 是 A 的逆矩阵.

可逆矩阵的一些性质

定理 3.3.2 设 A、B 均为 n 阶可逆矩阵，则：

（1）A 的逆矩阵A^{-1}亦为可逆矩阵，且$(A^{-1})^{-1}=A$.

（2）A 的转置矩阵A^T亦为可逆矩阵，且$(A^T)^{-1}=(A^{-1})^T$.

（3）$|A^{-1}|=|A|^{-1}$.

（4）AB的乘积亦可逆，且$(AB)^{-1}=B^{-1}A^{-1}$.

（5）$(kA)^{-1}={A^{-1}}/{k}$.

推论：若$A_1,A_2,\cdots,A_m$均是 n 阶可逆矩阵，则$A_1,A_2,\cdots,A_m$亦可逆且$(A_1,A_2,\cdots,A_m)^{-1}={A_m}^{-1},{A_{m-1}}^{-1},\cdots,{A_2}^{-1}{A_1}^{-1}$.

为了研究矩阵可逆的条件，我们先定义“伴随矩阵”.

定义（伴随矩阵）：将 n 阶方阵$\mathrm{A}=\left(a_{ij}\right)$的代数余子式$A_{ij}$按如下方式排成方阵：

$$\begin{bmatrix} A_{11} & A_{21} & & A_{n1} \\ A_{12} & A_{22} & & A_{n2} \\ & & & \\ A_{1n} & A_{2n} & & A_{nn} \end{bmatrix}$$ 称为$\mathrm{A}=\begin{bmatrix} a_{11} & a_{12} & & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix}$的伴随矩阵，并记作$A^*$.（注意$A^*$的行、列指标！）

由行列式的展开定理，易得A^*的如下性质：

$$\mathrm{A}A^*=A^*A=|A|I_n\,.$$

例 11 求$\mathrm{A}=\begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 1 & 2 & 4 \end{vmatrix}$的伴随矩阵$A^*$.

矩阵可逆的充要条件：

定理 3.3.3 n 阶可逆的充分条件是$|A|\neq 0$，且当$|A|\neq 0$时，$A^{-1}=\frac{1}{|A|}A^*$.

证明（1）必要性：设 A 可逆，则$\mathrm{A}A^{-1}=I_n\Rightarrow |A|\,|A^{-1}|=1$, 故$|A|\neq 0$.

（2）充分性：设$|A|\neq 0$，则由A^*的性质有：$\mathrm{A}\frac{A^*}{|A|}=\frac{A^*}{|A|}A=1$, 故 A 可逆，且$A^{-1}=\frac{1}{|A|}A^*$.

定理 3.3.4 n 阶可逆 A 可逆的充要条件是：存在 n 阶方阵 B，使$\mathrm{AB}=\mathrm{I}$（或 $\mathrm{BA}=\mathrm{I}$），且$\mathrm{B}=A^{-1}$.

证明 （1）必要性：设 A 可逆，则可取$\mathrm{B}=A^{-1}$, 则$\mathrm{AB}=\mathrm{I}$.

（2）充分性：设存在 B，使$AB=I$，则$|A||B|=1$，故$|A|\neq 0$.

由定理 3.3.3 可知 A 可逆，即A^{-1}存在，于是：

$$B=IB=(A^{-1}A)B=A^{-1}(AB)=A^{-1}I=A^{-1}.$$

说明：可逆的定义需要$\mathrm{AB}=\mathrm{I}$且$\mathrm{BA}=\mathrm{I}$两个条件，此定理说明：两个条件中只要满足一个即可！

逆矩阵的求法之二（伴随矩阵法）

（1）$A^{-1}=\frac{1}{|A|}A^*$，其中$A^*=\begin{bmatrix}A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \vdots & \vdots & & \vdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{bmatrix}$，$A_{ij}$是$a_{ij}$的代数余子式.

（2）特别地，$\mathrm{A}=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$时，$A^{-1}=\frac{1}{ad-bc}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$.

例 12　求$\mathrm{A}=\begin{bmatrix}1 & 2 & 3\\ 2 & 1 & 2\\ 1 & 3 & 3\end{bmatrix}$的逆矩阵.

例 13　设$\mathrm{A}=\mathrm{diag}[\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_n]$, 且$\neq 0$，求$A^{-1}$.

例 14　$\mathrm{A}=\begin{bmatrix}1 & 2 & 3\\ 2 & 2 & 1\\ 3 & 4 & 3\end{bmatrix}, B=\begin{bmatrix}2 & 1\\ 5 & 3\end{bmatrix}, C=\begin{bmatrix}1 & 3\\ 2 & 0\\ 3 & 1\end{bmatrix}$, 求矩阵 X 使之满足：$\mathrm{AXB}=\mathrm{C}$.

例 15　设 3 阶矩阵 A、B 满足:$A^{-1}BA=aA+BA$, 且$\mathrm{A}=\mathrm{diag}\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{7}\end{bmatrix}$, 求 B. 注意：$(A+B)^{-1}=A^{-1}+B^{-1}$.

例 16　设方阵 A 满足：$A^2-A+2I=0$，求证：A和$\mathrm{A}+2\mathrm{I}$均可逆，并求它们的逆！

回头再看 Cramer 法则

线性代数$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\ \quad\cdots\cdots\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n\end{cases}$

可简记作$\mathrm{AX}=\mathrm{B}$，其中：

$\mathrm{A}=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{\mathrm{n}1} & \cdots & a_{nn}\end{bmatrix}, X=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}, B=\begin{bmatrix}b_1\\ \vdots\\ b_n\end{bmatrix}$

则：若$|A|\neq 0$，A 即可逆，于是A^{-1}存在，$A^{-1}(AX)=A^{-1}B\Rightarrow X=A^{-1}B$，将$A^{-1}=\frac{1}{|A|}A^*$代入，得：$X=\frac{1}{|A|}A^*B$.

§3.4　分块矩阵

对于行数和列数比较高的矩阵，为了简化运算，常用的一种技巧是分块法，使大矩阵的运算化成各个小矩阵的运算，具体做法如下：

定义（分块矩阵） 将矩阵 A 用若干条横线和纵线分成许多小矩阵，每一个小矩阵称为 A 的分块或子块，以子块为元素的形式上的矩阵称为分块矩阵.

例 17　$\mathrm{A}=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$，若记：$A_{11}=\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$　$A_{21}=\begin{bmatrix} a_{31} \\ a_{41} \end{bmatrix}$

$A_{11}=\begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix}$　$A_{22}=\begin{bmatrix} a_{32} & a_{33} \\ a_{42} & a_{43} \end{bmatrix}$　则$\mathrm{A}=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$.

一般地，对于一个$\mathrm{m}\times\mathrm{n}$的矩阵 A，若行分成了 s 块，列分成了 t 块，则得到 A 的一个$\mathrm{s}\times\mathrm{t}$分块矩阵，记作：$\mathrm{A}=(A_{kl})_{\mathrm{s}\times\mathrm{t}}$.

常见的分块形式：

（1）按行分块：$\mathrm{A}=\left(a_{ij}\right)_{\mathrm{m}\times\mathrm{n}}, \mathrm{A}=\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$,

其中$\alpha_1=[a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}]$是行向量.

（2）按列分块，$\mathrm{A}=\left(a_{ij}\right)_{\mathrm{m}\times\mathrm{n}}$，$\mathrm{A}=[\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n]$,

其中$\beta_1=\begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}$是列向量.

（3）对角块矩阵（准对角矩阵）：若 n 阶方阵 A 的非零元素均集中在主对角线附近，则可分成如下形式，其中$\mathrm{A_i}$是$\mathrm{r_i}$阶小方阵.

$$\mathrm{A}=\begin{bmatrix}A_1 & & & O\\ & A_2 & & \\ & & \cdots & \\ O & & & A_s\end{bmatrix}$$（由上一章可得：$|A|=|A_1||A_2|\cdots|A_s|$）.

$$A=\begin{bmatrix}1&2&0&0&0&0\\0&4&0&0&0&0\\0&0&0&-1&1&0\\0&0&2&5&8&0\\0&0&0&3&-2&0\\0&0&0&0&0&9\end{bmatrix}=\begin{bmatrix}A_1 & & \\ & A_2 & \\ & & A_s\end{bmatrix}$$

分块矩阵的加法和数乘和转置：

设 A、B 是两个同型矩阵，且用同样的方法将 A、B 分块：

$$\mathrm{A}=\begin{bmatrix}A_{11} & \cdots & A_{1s}\\ \vdots & & \vdots\\ A_{k1} & \cdots & A_{ks}\end{bmatrix},\mathrm{B}=\begin{bmatrix}B_{11} & \cdots & B_{1s}\\ \vdots & & \vdots\\ B_{k1} & \cdots & B_{ks}\end{bmatrix}$$

则：由矩阵的加法和数乘的定义直接有：

$$\mathrm{A}+\mathrm{B}=\begin{bmatrix}A_{11}+B_{11} & \cdots & A_{1s}+B_{1s}\\ \vdots & & \vdots\\ A_{k1}+B_{k1} & \cdots & A_{ks}+B_{ks}\end{bmatrix}$$（对应块相加）

$$\lambda\mathrm{A}=\begin{bmatrix}\lambda A_{11} & \cdots & \lambda A_{1s}\\ \vdots & & \vdots\\ \lambda A_{k1} & \cdots & \lambda A_{ks}\end{bmatrix}$$（数乘所有块）

$$A^T=\begin{bmatrix}{A_{11}}^T & \cdots & {A_{1s}}^T\\ \vdots & & \vdots\\ {A_{k1}}^T & \cdots & {A_{ks}}^T\end{bmatrix}$$

分块矩阵的乘法：

若 A 是 m*k 的矩阵，B 是 k*n 的矩阵，从而 AB 有意义，如果对 A 和 B 分块使得 A 的列分法与 B 的行分法完全一致，即：

$$A=\begin{bmatrix} \overbrace{A_{11}}^{k} & \overbrace{A_{12}}^{k} & \cdots & \overbrace{A_{1s}}^{k} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{r1} & A_{r1} & \cdots & A_{rs} \end{bmatrix}\begin{matrix} \}m_1 \\ \}m_1 \\ \\ \}m_1 \end{matrix},\quad B=\begin{bmatrix} \overbrace{B_{11}}^{n_1} & \overbrace{B_{12}}^{n_2} & \cdots & \overbrace{B_{1s}}^{n_p} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{sp} \end{bmatrix}\begin{matrix} \}k_1 \\ \}k_2 \\ \\ \}k_s \end{matrix}$$

定义$C_{ij}=\sum_{t=1}^{p}A_{it}B_{tj}$(A 的行分法与 B 的列分法一致！)

$$C=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots \\ c_{r1} & c_{r2} & \cdots & c_{rp} \end{bmatrix}$$

易证$C=AB$!即分块矩阵相乘可将小块看成元素再按通常的矩阵乘法相乘!

例 18 $A=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}, B=\begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 1 & 0 \\ -1 & -1 \end{bmatrix}$, 求 AB.（两种方法做!）

例 19 设 A、B 均为 n 阶准对角矩阵且按相同的方式分块：

$$A=\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \cdots & \\ & & & A_r \end{bmatrix},\quad B=\begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \cdots & \\ & & & B_r \end{bmatrix}$$, 试求 AB.

例 20 A 为 n 阶准对角矩阵：$A=\begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \cdots & \\ & & & A_r \end{bmatrix}$, 试求 A 可逆的充要条件？并在可逆时求出其真逆矩阵！

解：$|A|\neq 0 \Rightarrow |A_1||A_2|\cdots|A_r|\neq 0.$

即：A 可逆的充要条件是每个分块的均可逆.

注意到$A\begin{bmatrix} {A_1}^{-1} & & & \\ & {A_2}^{-1} & & \\ & & \ddots & \\ & & & {A_r}^{-1} \end{bmatrix}=1$,

故$A^{-1}=\begin{bmatrix} {A_1}^{-1} & & & \\ & {A_2}^{-1} & & \\ & & \ddots & \\ & & & {A_r}^{-1} \end{bmatrix}$

补充：若 A 为斜对角矩阵，即：$\mathrm{A}=\begin{bmatrix} & & & A_1 \\ & & A_2 & \\ & ⋰ & & \\ A_n & & & \end{bmatrix}$呢？

例 21 $A=\begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 2 & 1 \end{bmatrix}$，求$A^{-1}$.

例 22 设$\mathrm{P}=\begin{bmatrix} A & 0 \\ C & B \end{bmatrix}$是一个 n 阶方阵,A 和 B 分别是 r 阶和 s 阶（r+s=n）可逆方阵，求证 P 可逆，并求出P^{-1}.

例 23 设 A、B、C、D 均是 n 阶方阵，且 A 可逆，又 AC=CA，试证明 $\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|AD-CB|$.

作业

1、求矩阵 X，使得：$\begin{pmatrix} 2 & 3 & -1 \\ 1 & 2 & 0 \\ -1 & 2 & -1 \end{pmatrix}X=\begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 3 & 0 \end{pmatrix}$.

2、设 A, B 为 n 阶方阵，若$\mathrm{AB}=\mathrm{A}+\mathrm{B}$，证明 $\mathrm{A}-\mathrm{I}$可逆，且$\mathrm{AB}=\mathrm{BA}$.

3、设$\mathrm{A}=\begin{pmatrix} 3 & 4 & 0 & 0 \\ 4 & -3 & 0 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 2 \end{pmatrix}$求$\mathrm{A}^{2\mathrm{k}}$, $|\mathrm{A}^{2\mathrm{k}}|$, k 为正整数.

4、求矩阵$\mathrm{A}=\begin{pmatrix} 3 & -3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{pmatrix}$的逆矩阵.

§3.5　矩阵的初等变换

3.5.1 初等变换、初等矩阵

前面给出过两种矩阵求逆的方法：待定系数法、伴随矩阵法，计算量都十分巨大，这节课介绍一种用初等变换求逆的方法.

定义（初等变换） 矩阵的行（列）初等变换是指对一个矩阵施行的如下三类变换：

1、交换矩阵的两行（列）；

2、用非零的数乘矩阵的某一行（列）；

3、用某个数乘矩阵的某一行（列）后加到另一行（列）.

矩阵的行初等变换和列初等变换统称为矩阵的初等变换.

类似的行列式的等值变形，引入初等变换的如下记号：

$$\xrightarrow{r_i \leftrightarrow r_j} \qquad \xrightarrow{kr_i} \qquad \xrightarrow{r_i + kr_j}$$

$$\xrightarrow{c_i \leftrightarrow c_j} \qquad \xrightarrow{kr_i} \qquad \xrightarrow{c_i + kc_i}$$

互换　　　　　　数乘　　　　　　数乘后再加到另一行（列）

定义（初等矩阵） 对 n 阶单位矩阵I_n施行一次初等变换得到的矩阵，称作 n 阶初等矩阵.

三种初等变换对应三种初等矩阵：

$$P_{ij}=\left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
\cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots & 1 & \cdots & \cdots & \cdots \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
\cdots & \cdots & \cdots & 1 & \cdots & \cdots & \cdots & 0 & \cdots & \cdots & \cdots \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right]\begin{array}{l}
\\ \\ \\ \text{第 i 行} \\ \\ \\ \\ \text{第 j 行} \\ \\ \\ \\
\end{array}$$

I_n的第 i, j 行（列）互换.

$$D_i(k)=\left[\begin{array}{ccccccc}
1 & & & & & & 0 \\
 & \ddots & & & & & \\
 & & 1 & & & & \\
 & & & k & & & \\
 & & & & 1 & & \\
 & & & & & \ddots & \\
0 & & & & & & 1
\end{array}\right]\begin{array}{l}
\\ \\ \\ \text{第 i 行} \\ \\ \\ \\
\end{array}$$

I_n的第 i 行（列）倍乘 k

$$T_{ij}(c)=\left[\begin{array}{ccccccc}
1 & & & & & & \\
 & \ddots & & & & & \\
 & & 1 & \cdots & c & & \\
 & & & 1 & \vdots & & \\
 & & & & 1 & & \\
 & O & & & & \ddots & \\
 & & & & & & 1
\end{array}\right]\begin{array}{l}
\\ \\ \text{第 i 行} \\ \\ \text{第 j 行} \\ \\ \\
\end{array}$$

或

$$\left[\begin{array}{ccccccc}
1 & & & & & & \\
 & \ddots & & & & O & \\
 & & 1 & & & & \\
 & & \vdots & 1 & & & \\
 & & c & \cdots & 1 & & \\
 & & & & & \ddots & \\
 & & & & & & 1
\end{array}\right]\begin{array}{l}
\\ \\ \text{第 i 行} \\ \\ \text{第 j 行} \\ \\ \\
\end{array}$$

I_n的第 j 行（或 j 列）倍乘 c 后加到第 i 行（或 i 列）上

定理 3.5.1　n 阶初等矩阵均为可逆矩阵，且其逆与其本身属同类型的初等矩阵.

即：(1) ${P_{ij}}^{-1} = P_{ij}$

(2) $\left(D_i(k)\right)^{-1} = D_i\left(\frac{1}{k}\right), k \neq 0$

(3) $\left(T_{ij}(c)\right)^{-1} = T_{ij}(-c)$

定理 3.5.2　对矩阵$\mathrm{A}_{\mathrm{m}\times\mathrm{n}}$施行一次行（列）初等变换的结果等于用一个 m(n)阶初等矩阵左（右）乘 A.

证：　仅对行初等变换证明之，（列类似）. 将 A 如下分块

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_1 \\ \vdots \\ A_1 \\ \vdots \\ A_1 \end{bmatrix}$$

$$P_{ij}A = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & & \ddots & & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \\ A_n \end{bmatrix}$$

$$T_{ij}(c)A = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & c & & \\ & & & \ddots & & & \\ & & 0 & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix}$$

$$=\begin{bmatrix} A_1 \\ \vdots \\ A_i + cA_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix}.$$

3.5.2 行等价、简化行梯形矩阵、主 1

定义（行等价） 若 A 经有限次行初等变换化到矩阵 B，则称 A 与 B 等价.

由于初等变换的逆也是初等变换，故行等价是一个等价关系，即：

（1）自反性：A 与自身等价

（2）对称性：若 A 与 B 行等价，则 B 与 A 行等价

（3）传递性：若 A 与 B 行等价，B 与 C 行等价，A 与 C 行等价

定义（简化行阶梯形，或行标准形） 若矩阵 U 满足如下条件：

（1）U 的所有零行在非零行的下面；

（2）每个非零行的 第一个非零元素（自左至右）都等 于 1，称作主 1；

（3）每一个非零行的主 1，是它所在列唯一的非零元素；

（4）每 一非零行的主 1 所在的列在上一行主 1 所在列的右边.

定理 3.5.3 任意一个矩阵$A_{m\times n}$一定与一个简化行梯形矩阵 U 行等价. 即：存在有限个初等矩阵E_1，E_2，…，E_l，使得：$E_l \cdots E_2 E_1 A = U$.

定理的证明：对行数进行数归！

（1）找出第一个非零列，将非零列元素所在的行换到第一行，并归一；

（2）将该列其余元素变为零；

（3）对右下部分用数归；

（4）将第一行对应其它主 1 的位置化零.

（证明的思路亦是实际计算过程!）

引理 3.5.4 可逆的 n 阶简化行阶梯形矩阵 U 必为I_n.

证：U 可逆$\Rightarrow |U| \neq 0 \Rightarrow$U不存在零行，每行都有主 1，

故 U 中有 n 个主 1，主 1 的个数等于列数，故$\mathrm{U} = I_n$.

引理 3.5.5 两个行等价 n 阶矩阵或者均可逆，或者均不可逆.

证：$E_l \cdots E_2 E_1 A = B \Rightarrow |E_l| \cdots |E_2||E_1||\mathrm{A}| = |\mathrm{B}|$

又$|E_i| \neq 0$，故$\begin{cases} |A| = 0 \Rightarrow |B| = 0 \\ |A| \neq 0 \Rightarrow |B| \neq 0 \end{cases}$

定理 3.5.6 设 A 为 n 阶方阵，则下列条件等价:

(1) A 可逆

(2) A 与I_n行等价

(3) A 等于有限个初等矩阵的积。

证:$(1) \Rightarrow (2)$:设$E_k \cdots E_2 E_1 A = U$, 则 U 可逆,由引理 3.5.4 有$\mathrm{U} = I_n$，

$(2) \Rightarrow (3)$ ： 由条件$E_k \cdots E_2 E_1 A = I_n$，

$A = {E_1}^{-1} {E_2}^{-1} \cdots {E_k}^{-1} I_n = {E_1}^{-1} {E_2}^{-1} \cdots {E_k}^{-1}$.

$(3) \Rightarrow (1)$:设$A = E_1 E_2 \cdots E_k$，则$|A| \neq 0$，故可逆.

（此定理给出逆矩阵的第 3 种求法：初等变换法）

初等变换求逆矩阵：

假设 A 可逆，则$E_k \cdots E_2 E_1 A = I_n$，构造分块矩阵：$[A \;\vdots\; I_n]$，再注意到$E_k \cdots E_2 E_1 A = A^{-1}$.

则：

$E_k \cdots E_2 E_1 [A \;\vdots\; I_n] = [E_k \cdots E_2 E_1 A \;\vdots\; E_k \cdots E_2 E_1] = [I_n \;\vdots\; A^{-1}]$

即：对$[A \;\vdots\; I_n]$实行行初等变换，将 A 变成I_n，此时，

右半部分的I_n即变成了A^{-1}!

若事先不知 A 是否可逆，则在上述方法中变换若干步之后，A 中出现了零行或零列.

由于$\begin{bmatrix} A \\ \cdots \\ I_n \end{bmatrix} E_1E_2\cdots E_l = \begin{bmatrix} AE_1E_2\cdots E_k \\ \cdots\cdots \\ E_1E_2\cdots E_k \end{bmatrix} = \begin{bmatrix} I_n \\ \cdots \\ A^{-1} \end{bmatrix}$, 故亦为用列变换的方法求逆矩阵！

例 24 $\mathrm{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 4 & 2 \\ 1 & 1 & -1 \end{bmatrix}$，A 是否可逆？若可逆，求$A^{-1}$.

例 25 $\mathrm{A} = \begin{bmatrix} 1 & 2 & -3 \\ 1 & -2 & 1 \\ 5 & -2 & -3 \end{bmatrix}$，A 是否可逆？

若允许进行列初等变换，简化行梯形矩阵还可以继续简化. 为此，我们回顾一下行标准型的结构：

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

（1）将主 1 所在的列依次对换到第 1 列，第 2 列，⋯.

（2）用主将所在行的其它非零元素化为零.

定理 3.5.7 任意一个$\mathrm{m} \times n$的矩阵，经过初等变换（包括行初等变换和列初等变换）一定可以化简成如下形式：

$$\begin{pmatrix} \mathrm{I_r} & \mathrm{O_{r\times(n-r)}} \\ \mathrm{O_{(m-r)\times r}} & \mathrm{O_{(m-r)\times(n-r)}} \end{pmatrix}$$

其中$0 \le \mathrm{r} \le \min(\mathrm{m}, \mathrm{n})$. 上述形式称为 A 的标准型. 换言之：对于任意矩阵 A，一定存在可逆方阵 P，Q 使得：

$$\mathrm{PAQ} = \begin{pmatrix} \mathrm{I_r} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} \end{pmatrix}$$

注： 前面已证明，对于可逆方阵，简化行梯形就是单位矩阵，也即标准型. 因此可逆矩阵只需要行初等变换就可化为标准型.

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

3.5.3 等价、标准型

定义（等价） 若 A 经有限次初等变换（包括行和列）变化到矩阵 B，则称 A 与 B 等价.

A 和 B 等价说明存在可逆矩阵 P、Q，使$PAQ=B$.

显然，矩阵的等价亦是一个等价关系（自反、对称、传递）

定理 3.5.8 亦可表述为：任意矩阵 A 必定与矩阵 $\begin{bmatrix}I_r & 0\\ 0 & 0\end{bmatrix}$ 等价！

问题 一个矩阵 A 的标准形是否唯一？即：用不同的步骤得到的I_r的阶数是否相等？（矩阵的秩！）

分块矩阵的初等变换

这里仅就常用的2×2分块矩阵为例来讨论. 对分块矩阵的初等变换，相应定义三种分块初等矩阵为：

（1）分块对换初等矩阵

$\begin{pmatrix}0 & I_s\\ I_r & 0\end{pmatrix}$， I_r, I_s分别为 r 与 s 阶单位矩阵.

（2）分块倍乘初等矩阵

$\begin{pmatrix}kI_r & 0\\ 0 & I_s\end{pmatrix}$ 或 $\begin{pmatrix}I_r & 0\\ 0 & kI_s\end{pmatrix}$，k 为非零数.

（3）分块倍加初等矩阵

$$\begin{pmatrix} I_r & 0 \\ C_1 & I_s \end{pmatrix} \quad 或 \quad \begin{pmatrix} I_r & C_2 \\ 0 & I_s \end{pmatrix}$$

将这些分块初等矩阵左(右)乘矩阵 A，则相当于对矩阵 A 做相应的行(列)块的初等变换. 下面举例：

例如：设 A 为 r 阶可逆矩阵，D 为 s 阶矩阵，对分块矩阵$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$，

为使其中的矩阵 C 和 B 化为零矩阵，令倍加分块初等矩阵为

$$P = \begin{pmatrix} I_r & 0 \\ -CA^{-1} & I_s \end{pmatrix},$$

于是 $$\begin{pmatrix} I_r & 0 \\ -CA^{-1} & I_s \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix},$$

$$Q = \begin{pmatrix} I_r & -A^{-1}B \\ 0 & I_s \end{pmatrix}$$

$$\begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}\begin{pmatrix} I_r & -A^{-1}B \\ 0 & I_s \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{pmatrix}$$

例 26 设A, B, C, D为 n 阶方阵，A 为可逆矩阵，且$AC = CA$，证明：$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD - CB|$.

证：利用上面例题的结论，并对两边的矩阵取行列式，并利用倍加分块初等矩阵的行列式等于 1 及$AC = CA$，得

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} A & B \\ 0 & D - CA^{-1}B \end{vmatrix} = |A||D - CA^{-1}B| = |AD - ACA^{-1}B|$$

例 27 求$\begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$的逆.

解：

$$\left(\begin{array}{cc:cc} A & 0 & I & 0 \\ C & B & 0 & I \end{array}\right) \to \left(\begin{array}{cc:cc} I & 0 & A^{-1} & 0 \\ B^{-1}C & I & 0 & B^{-1} \end{array}\right) \to \left(\begin{array}{cc:cc} I & 0 & A^{-1} & 0 \\ 0 & I & B^{-1}CA^{-1} & B^{-1} \end{array}\right)$$

例 28 A,B 皆为 n 阶方阵，证明$|I - AB|$=$|I - BA|$.

证：（1）若 A 或 B 可逆（不妨设 A 可逆），则

$$|I - AB| = |A(I - BA)A^{-1}| = |A||A^{-1}||I - BA| = |I - BA|$$

（2）若 A，B 皆不可逆，构造分块矩阵 $\begin{pmatrix} I_n & A \\ B & I_n \end{pmatrix}$,

用倍加分块初等矩阵把 B 消为零矩阵，

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

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

对上面两式取行列式，得

$$\begin{vmatrix} I_n & A \\ B & I_n \end{vmatrix} = |I_n||I_n - BA| = |I_n||I_n - AB|,$$

故有 $$|I - AB| = |I - BA|$$

§ 3.6 矩阵的秩

3.6.1 子式、秩

秩是矩阵的一个重要数量特性，矩阵的许多性质可通过秩来刻画.

定义（子式） 矩阵$A_{m\times n}$ 中任取 k 行 k 列$\left(k\leq m，n\right)$，位于这些行列式交叉点处的 k^2个元素按原相对位置构成的 k 阶行列式称矩阵 A 的一个 k 阶子式，k 阶子式的个数:$C_m^k C_n^k$.

定义（秩） 矩阵$A_{m\times n}$中不等于 0 的子式的最大阶数，称为 A 矩的秩，记作 r(A). 规定零矩阵 r(0)=0. 显然，只有零矩阵 r=0,非零矩阵至少$r\geq 1$.

由定义有：

（1）对$A_{m\times n}$，$r(A)\leq \min\{m,n\}$

（2）$r(kA)=r(A)$，$(k\neq 0)$

（3）对 n 阶方阵，r(A)=n（满秩)⇔A 可逆！

（4）若 A 的所有 k+1 阶子式等于零，则 $r(A)\leq k$；若 A 有一个 k 阶子式非零，则$r(A)\geq k$.

例 29 求下列矩阵的秩：

$$A=\begin{bmatrix}1&2&1&3&4\\0&2&5&1&2\\0&0&0&1&1\\0&0&0&0&0\end{bmatrix};\quad B=\begin{bmatrix}1&3&-9&3\\0&1&-3&4\\-2&-3&9&6\end{bmatrix}$$

由上面例子可见，通过定义求秩，需要要求若干个行列式.

3.6.2 初等变换不改变秩，主 1 的个数

下面寻找秩的其它求法——初等变换法.

定理 3.6.1 初等变换不改变矩阵的秩.

证明：一，二类易证.以第三类行初等变换为例证之， 设 r(A)=r

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \underset{r_i+(-k)r_j}{\overset{r_i+kr_j}{\rightleftarrows}}$$

$$A=\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+ka_{j1} & \cdots & a_{in}+ka_{jn} \\ \vdots & & \vdots \\ a_{j1} & & a_{jn} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{jm} \end{bmatrix}$$

①B 不存在>r 阶子式，r (B) ≤ r(A)

②B 存在 r+1 阶子式，取一个 r+1 阶子式 $\left\{\begin{matrix} \text{不含第 i 行} \\ \text{含第 i 行，不含第 j 行} \\ \text{含 i，j 行} \end{matrix}\right\}$

r (B) ≤ r(A)

可逆方阵乘以 A 会改变 A 的秩吗？

推论 3.6.2 行等价(或列等价，或等价)的矩阵，秩相等.

即：$r(PA) = r(AQ) = r(PAQ) = r(A)$

推论 3.6.3 矩阵的秩等于其简化行梯形中主 1 的个数.

矩阵的标准形的秩显然等于其中I_r的阶数，故对任意一个矩阵，用不同的途径得到的标准形中I_r的阶数都等于原矩阵的秩.

推论 3.6.4 矩阵的标准形唯一,且矩阵的秩等于其标准形中I_r的阶数.

例 30 $A=\begin{bmatrix}1&0&2&1&0\\7&1&14&7&1\\0&5&1&4&6\\2&1&1&-10&-2\end{bmatrix}$的秩.

定理 3.6.5（推论 3.6.2 的加强）：同型矩阵 A、B 等价⇔r(A)=r(B).

证明： ⇒: 显然，

⇐：r(A)=r(B)，则均等价于$\begin{bmatrix}I_r&0\\0&0\end{bmatrix}$由等价的传递性即得！

定理 3.6.6 $r(A^T)=r(A)$.

证：设$r(A)=r$，$A_{m\times n}$，则 $PAQ=\begin{bmatrix}I_r&O_{r\times(n-r)}\\O_{(m-r)\times r}&O_{(m-r)\times(n-r)}\end{bmatrix}$

取转置即可证之.

3.6.3 矩阵秩的一些不等式

由定义可知：对矩阵$A_{m\times n}$ 有： $r(A)\le\min\{m,n\}$. 下面研究矩阵运算后秩满足的一些不等式：

定理 3.6.7 矩阵 A 数乘 k 后，秩满足：

$$r(kA)=\begin{cases}0&,k=0\\r(A)&,k\neq 0\end{cases}$$

证：k=0 时 kA=0；$k\neq 0$ 时 kA 与 A 等价.

为研究矩阵加法和乘法运算后秩的变化，先看如下引理：

引理 3.6.8 分块矩阵 $\begin{bmatrix}A&0\\0&B\end{bmatrix}$的秩 r=r(A)+r(B).

证： 设$P_1AQ_1=\begin{bmatrix}I_{r(A)}&0\\0&0\end{bmatrix}$ $P_2BQ_2=\begin{bmatrix}I_{r(B)}&0\\0&0\end{bmatrix}$

构造$P=\begin{bmatrix}P_1&0\\0&P_2\end{bmatrix}$ $,Q=\begin{bmatrix}Q_1&0\\0&Q_2\end{bmatrix}$.

定理 3.6.9 设 $A_{m\times n}$ ，$B_{n\times k}$ ，则$r(AB)\le \min\{r(A), r(B)\}$.

证：设 r(A)=r,则$A=P^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}Q^{-1}$，

$AB=P^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}(Q^{-1}B)=P^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} C1 \\ C2 \end{bmatrix}=\ P^{-1}\begin{bmatrix} C1 \\ 0 \end{bmatrix}$

故$r(AB)\le r=r(A)$

同理$r(AB)\le r(B)$. 证毕.

定理 3.6.10 设$A_{m\times n}$，$B_{n\times k}$，则$r(AB)\ge r(A)+r(B)-n$.

特别地，若 AB=0，则$r(A)+r(B)\le n$.

证 I：构造：$\begin{bmatrix} I_n & 0 \\ -A & I_m \end{bmatrix}\begin{bmatrix} I_n & B \\ A & 0 \end{bmatrix}\begin{bmatrix} I_n & -B \\ 0 & I_k \end{bmatrix}=\cdots$

而$r\left(\begin{bmatrix} I_n & B \\ A & 0 \end{bmatrix}\right)\ge r(A)+r(B)$,⋯证毕.

证 II：假设 r(A)=r，r(B)=s ， 即存在 P1，Q1 和 P2，Q2，

使 $P_1AQ_1=\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$， $P_2BQ_2=\begin{bmatrix} I_s & 0 \\ 0 & 0 \end{bmatrix}$

$$P_1ABQ_2=\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}Q_1^{-1}\,P_2^{-1}\begin{bmatrix} I_s & 0 \\ 0 & 0 \end{bmatrix}$$

$$=\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}\begin{bmatrix} I_s & 0 \\ 0 & 0 \end{bmatrix}$$

$$=\begin{bmatrix} c_{11} & \cdots & c_{1s} & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & 0 & 0 & 0 \\ c_{r1} & \cdots & c_{rs} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}=\begin{bmatrix} D_{r\times s} & 0 \\ 0 & 0 \end{bmatrix}$$

$r(D_{r\times s})\ge r(C_{n\times n})-(n-s)-(n-r)=r+s-n$

定理 3.6.11 设 $A_{m\times n}$，$B_{m\times n}$，则$r(A+B)\le r(A)+r(B)$.

证：构造： $[A \ \vdots \ B]\begin{bmatrix} I_n \\ I_n \end{bmatrix} = A + B \Rightarrow r(A+B) \le r([A \ \vdots \ B])$

又$r([A \ \vdots \ B]) \le r\left(\begin{bmatrix} A & B \\ 0 & B \end{bmatrix}\right)$

而$\begin{bmatrix} A & B \\ 0 & B \end{bmatrix} = \begin{bmatrix} I_m & I_m \\ 0 & I_m \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$

故$r\left(\begin{bmatrix} A & B \\ 0 & B \end{bmatrix}\right) = r\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right) = r(A) + r(B)$

综上有 $r(A+B) \le r(A) + r(B)$.

例 31 A、B 为 n 阶方阵，则 r(AB)=r(BA)成立否？

解：例如： $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$，B=$\begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix}$.

例 32 设 A 是 n 阶幂等矩阵，即$A^n = A$，求证：

$r(A) + r(I - A) = n$.

证：想办法利用和、积关系：$A + (I - A) = I$， $A(I - A) = 0$.

例 33 设 A、B 为 n 阶矩阵，且$r(A - I) = p$, $r(B - I) = q$，求证：

$r(AB - I) \le p + q$.

证：想办法将$AB - I$中凑出$A - I$和$B - I$，再用不等式

$$AB - I = A(B - I) + A - I = \ldots$$

补充例题

例 34 设 $\alpha = [1 \quad 2 \quad 3], \beta = \left[1, \frac{1}{2}, \frac{1}{3}\right]$，求$A = \alpha^T\beta$ ，以及A^n.

例 35 设 α是三维行向量，则 $\alpha^T\alpha = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$ 求$\alpha\alpha^T$.

例 36 设 A 可逆，且每行元素之和为a，求证A^{-1}每行元素之和为$\frac{1}{a}$.

解：构造 $[1 \quad 1 \quad \cdots]^T$，将 A 左乘之，又$|A| \ne 0$，各列全加到第一列上，可证明$a \ne 0$.

例 37 设 n 阶矩阵 A 可逆，且$n \geq 2$，求$(A^*)^*$与 A 之关系.

利用$A^{-1} = \frac{1}{|A|}A^*$.

例 38 设方阵 A 满足$A^2 + A - 4I = 0$，求$(A - I)^{-1}$.

因子分解时凑出$A - I$来.

例 39 设 $A = \begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}$，且 $r(A^*)=1$，求 a、b 满足的关系.

先求出A^*来比较繁，可考虑利用秩的不等式！$AA^*=|A|I$.

例 40 设 $ABA^{-1} = BA^{-1}+3I$，且 $A^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -3 & 0 & 8 \end{bmatrix}$，求 B.

解：$\because A^{-1}= \frac{A^*}{|A|}$，要求$A^{-1}$必须先求$|A|$

由$AA^*=|A|I$可求出$|A|$

求解时先化简，再求值！

$B=A^{-1}B+ 3I \Longrightarrow B$.

例 41 设 A 为 n 阶矩阵，求证：

（1）$r(A)=n$ 时， $r(A^*)=n$

（2）$r(A)=n-1$ 时，$r(A^*)=1$

（3）$r(A)<n-1$ 时，$r(A^*)=0$ 换言之，A^*的秩只可能取值 0，1，n.

第四章：线性空间

线性空间的定义和例子 子空间

向量组的线性无关性

n 元向量组与矩阵的关系

基、维数、坐标

§4.1 线性空间的定义和例子

线性空间是最基本的数学概念之一. 也是一个抽象的概念，是某一类事物从量的方面的一个抽象，通过研究向量空间的普遍性质来解决具体问题.

第三章研究了数域 F 上的$m \times n$ 矩阵的集合$F^{m\times n}$并定义了矩阵的线性运算——加法和数乘. 这些线性运算满足一些性质，数学中经常会遇到许多具有同样性质的类似的集合，抽象出来，即：

定义 4. 1. 1（线性空间） 设 F 是一个数域，V 是一非空集合，若下列条件被满足，则称 V 是 F 上的一个线性空间，或称作向量空间：

i) V 中定义了加法，即对$\forall\alpha,\beta\in V$ 存在唯一元素$\gamma\in V$与之对应，称作$\alpha+\beta$之和，并记作：$\gamma=\alpha+\beta$.

ii) V 中定义了数乘，即对$\forall k\in F$，$\forall\alpha\in V$，存在唯一元$\beta\in V$，与它们对应，称 k 与α的积，并记作$\beta=k\alpha$.

iii) 上述加法和数乘满足如下运算律：

1° $\alpha+\beta=\beta+\alpha$ （加法交换律）

2° $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$ （加法结合律）

3° V 中存在零元，记作 0，满足：

$\forall\alpha\in V$，$\alpha+0=\alpha$ （存在零元）

4° $\forall\alpha\in V$，$\exists\alpha'\epsilon V$，使$\alpha+\alpha'=0$

称α'是α的负元素 （存在负元）

5° $k(\alpha+\beta)=k\alpha+k\beta$ （数乘分配律Ⅰ）

6° $(k+l)\alpha=k\alpha+l\alpha$ （数乘分配律Ⅱ）

7° $(kl)\alpha=k(l\alpha)$ （数乘结合律）

8° $1\alpha=\alpha$ （单位律）

这时，我们将 V 中的元素称为“向量”，F 中的元素称为“数量”或“纯量”.

F 为实数域时，称 V 为“实线性空间”；F 为复数域时，称 V 为“复线性空间”.

从线性空间的定义可推出如下性质：

（1）V 中的零向量（用 0 表示）是唯一的.

证：设$\exists 0,0'$使$\alpha+0=\alpha$，$\alpha+0'=\alpha$，则$0'+\alpha=\alpha$

于是：$0'=0'+0=0$.

（2）$\forall\alpha\in V$，其负向量唯一.

证：设α有两个负向量α'，α''，则$\alpha'+\alpha=\alpha''+\alpha=0$

故：$\alpha'=\alpha'+0=\alpha'+(\alpha+\alpha'')=(\alpha'+\alpha)+\alpha''=0+\alpha''=\alpha''$.

由于负向量唯一，可将α的负向量记作$-\alpha$，即：

$\alpha+(-\alpha)=(-\alpha)+\alpha=0$.

并定义减法：$\alpha-\beta\triangleq\alpha+(-\beta)$.

(3)$\forall\alpha\in V$有$0\alpha=0$；$\forall k\in F$有，$k0=0$.

证：$0\alpha=0\alpha+0=0\alpha+(0\alpha-0\alpha)=(0\alpha+0\alpha)-0\alpha$

$$=(0+0)\alpha-0\alpha=0\alpha-0\alpha=0$$

$k0=k0+0=k0+(k0-k0)=(k0+k0)-k0$

$$=k(0+0)-k0=k0-k0=0.$$

（4）$\forall k\in F$，$\forall\alpha\in V$，有$k(-\alpha)=(-k)\alpha=-k\alpha$.

证：$k(-\alpha)+k\alpha=k(-\alpha+\alpha)=k0=0$

故$k(-\alpha)=-k\alpha$

同理，$(-k)\alpha+k\alpha=(-k+k)\alpha=0\alpha=0$

故$(-k)\alpha=-k\alpha$.

特别地：$(-1)\alpha=-\alpha$.

$$k\alpha = 0 \Rightarrow k = 0 \text{ 或 } \alpha = 0.$$

证：若 k=0,显然 $k\alpha = 0$; 则$\alpha = 1\alpha = \left(\frac{1}{k}k\right)\alpha = 0.$

判断线性空间的方法：

1、有加法和数乘运算，且封闭；

2、满足八条运算律.

例 1 实数域上的全体$m \times n$矩阵，对矩阵的加法和数乘运算构成是了实数域上的线性空间，记作$R^{m\times n}$.

特别地，数域 F 上的全体$n \times 1$矩阵的集合，以及全体 $1\times n$矩阵的集合，分别构成 F 上的线性空间. 前者称 F 上的 n 元列向量空间，后者称 F 上 n 元行向量空间，并用同一个记号F^n来表示之，即：（n 维有序数组）

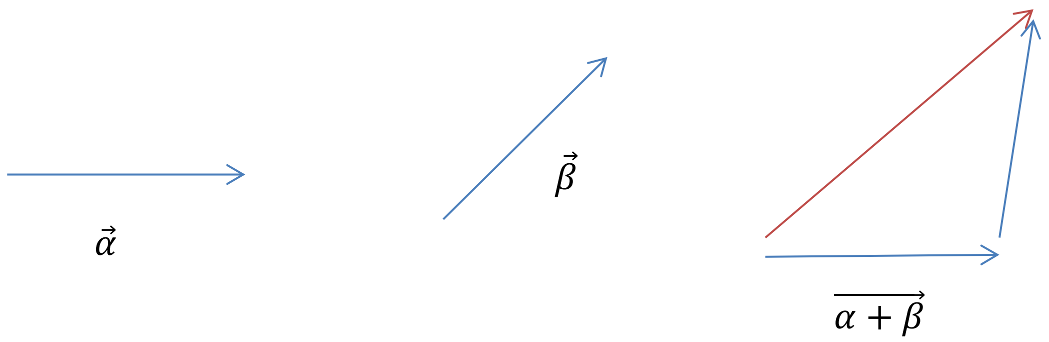
$$F^n = \{[a_1, a_2, \cdots, a_n]^T : a_i \in F, i = 1,2,\cdots,n\} \qquad (4.1)$$

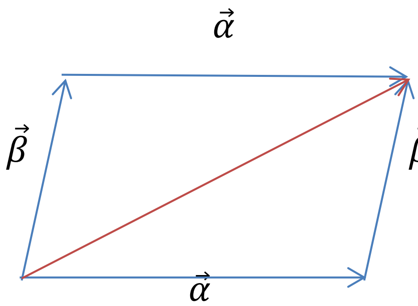
或 $F^n = \{[a_1, a_2, \cdots, a_n] : a_i \in F, i = 1,2,\cdots,n\}$

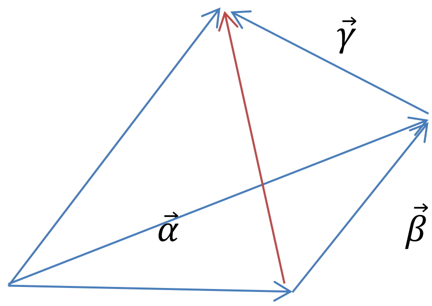
（n 维基本向量 ……）.

例 2 任意数域 F 按照通常数的加法和乘法，构成自身上一个线性空间.

例 3 三维空间中的矢量(三维空间中的有向线段)，对于通常的矢量加法（三角形规则），和实数与矢量的数量乘法（有向线段的长度倍乘），构成实数域上的线性空间.







例 4　R 上次不超过 n 的多项式的全体，记作 $\mathrm{P[x]_n}$，即：

$\mathrm{P[x]_n} = \{\mathrm{p} = a_n x^{\mathrm{n}} + \cdots + a_1 x + a_0 : a_{\mathrm{n}}, \cdots, a_1, a_0 \in \mathrm{R}\}$

对于通常多项多的加法，以及数与多项式的乘法，是否构成线性空间？

次数等于 n 的全体 $\mathrm{Q}[x]_{\mathrm{n}}$呢？

例 5　R 上如下的正弦函数集合：

$$\mathrm{S}[x] = \{\mathrm{s} = \mathrm{A}\sin(x + \mathrm{B}) : \mathrm{A}, \mathrm{B} \in \mathrm{R}\}$$

对于通常的函数加法，以及数与函数的普通乘法，是否构成一个线性空间？

$$\begin{aligned} \because S_1 + S_2 &= A_1 \sin(x + B_1) + A_2 \sin(x + B_2) \\ &= (a_1 \cos x + b_1 \sin x) + (a_2 \cos x + b_2 \sin x) \\ &= (a_1 + a_2) \cos x + (b_1 + b_2) \sin x \\ &= \mathrm{A}\sin(x + B) \ \in \mathrm{S}[x]. \end{aligned}$$

更一般地，区间[a，b]上的一切实连续函数的集合，对通常的加法和乘法运算，构成 R 上的一个实线性空间.

例 6　n 元有序数组的全体

$$\mathrm{S^n} = \{x = (x_1, x_2, \cdots, x_{\mathrm{n}}) : x_{\mathrm{i}} \in \mathrm{R}, \mathrm{i} = 1,2, \cdots, \mathrm{n}\}$$

对通常有序数组的加法和如下数乘：

$$\lambda_0 (x_1, x_2, \cdots, x_{\mathrm{n}}) = (0,0, \cdots, 0)$$

是否构成线性空间？

例 7　全体正实数$\mathrm{R^+}$，定义如下的加法和数乘运算：

加法：$\forall a, \mathrm{b} \in \mathrm{R^+}$, $a \oplus \mathrm{b} = a\mathrm{b}$,

数乘：$\forall \mathrm{k} \in \mathrm{R},\ a \in \mathrm{R}^+,\ \mathrm{k} \circ a = a^{\mathrm{k}}$.

验证R^+对上述加法和数乘构成 R 上的线性空间.

加法的 0 元素是？加法的负向量是？

下面一一验证八条线性运算规律：

(1)$a \oplus \mathrm{b} = a\mathrm{b} = \mathrm{b}a = \mathrm{b} \oplus a$;

(2)$(a \oplus \mathrm{b}) \oplus \mathrm{c} = (a\mathrm{b}) \oplus \mathrm{c} = (a\mathrm{b})\mathrm{c} = a \oplus (\mathrm{b} \oplus \mathrm{c})$;

(3)R^+中存在零元素 1，对任何$a \in \mathrm{R}^+$，有$a \oplus 1 = a \cdot 1 = a$;

(4)$\forall a \in \mathrm{R}^+$，有负元素$a^{-1} \in \mathrm{R}^+$，使$a \oplus a^{-1} = a \cdot a^{-1} = 1$;

(5)$1 \circ a = a^1 = a$;

(6)$\lambda \circ (\mu \circ a) = \lambda \circ a^{\mu} = (a^{\mu})^{\lambda} = a^{\lambda\mu} = (\lambda\mu) \circ a$;

(7)$(\lambda + \mu) \circ a = a^{\lambda+\mu} = a^{\lambda}a^{\mu} = a^{\lambda} \oplus a^{\mu} = \lambda \circ a \oplus \mu \circ a$;

(8)$\lambda \circ (a \oplus \mathrm{b}) = \lambda \circ (a\mathrm{b}) = (a\mathrm{b})^{\lambda} = a^{\lambda}\mathrm{b}^{\lambda} = a^{\lambda} \oplus \mathrm{b}^{\lambda} = \lambda \circ a \oplus \lambda \circ \mathrm{b}$.

所以R^+对所定义的运算构成线性空间.

§ 4.2 子空间

定义（子空间） 设 V 是数域 F 上的线性空间，W 是 V 的一个非空子集合. 若 W 对于 V 中的加法和数乘运算也构成 F 上的线性空间，则称 W 是 V 的子空间.

例 8 任意的线性空间 V 都有两个子空间，一个是 V 自身，另一个是 $\{0\}$，称作零元素空间. 这两个也叫 V 的平凡子空间.

例 9 次数$\leq n$的多项式集合$\mathrm{P}[x]_{\mathrm{n}}$是多项式空间$\mathrm{P}[\mathrm{x}]$的一个子空间.

定理 4.2.1 线性空间 V 的非空子集合 W 构成子空间的充要条件是：

$(1)\forall\alpha,\beta\in W,\alpha+\beta\in W$

$(2)\forall\alpha\in W,\ \forall k\in F,k\alpha\in W$

证：$(\Rightarrow)$ ：显然，

$(\Leftarrow)$ ：$W\subseteq V$，故加法和数乘的运算法则是遵守的，只需证（3）存在零元和（4）存在负元.

（1）（2）可合写成：$\forall\alpha$，$\beta\in W$，$\forall k$，$l\in F$，$k\alpha+l\beta\in W$.

$R^{2\times3}$的下列子集是否构成子空间？为什么？

$(1)W_1=\left\{\begin{pmatrix}1&b&0\\0&c&d\end{pmatrix}\middle|b,c,d\in R\right\}$;

$(2)W_2=\left\{\begin{pmatrix}a&b&0\\0&0&c\end{pmatrix}\middle|a+b+c=0,b,c,d\in R\right\}$.

解：$A=B=\begin{pmatrix}1&0&0\\0&0&0\end{pmatrix}\in W_1$

$$A+B=\begin{pmatrix}2&0&0\\0&0&0\end{pmatrix}\notin W_1.$$

即W_1对矩阵加法不封闭，不构成子空间.

因$\begin{pmatrix}0&0&0\\0&0&0\end{pmatrix}\in W_2$，即$W_2$非空.

$A=\begin{pmatrix}a_1&b_1&0\\0&0&c_1\end{pmatrix},B=\begin{pmatrix}a_2&b_2&0\\0&0&c_2\end{pmatrix}\in W_2$

$a_1+b_1+c_1=0,\qquad a_2+b_2+c_2=0,$

$$A+B=\begin{pmatrix} a_1+a_2 & b_1+b_2 & 0 \\ 0 & 0 & c_1+c_2 \end{pmatrix}.$$

满足$(a_1+a_2)+(b_1+b_2)+(c_1+c_2)=0$，

即$A+B\in W_2$，对任意$k\in R$有

$$kA=\begin{pmatrix} ka_1 & kb_1 & 0 \\ 0 & 0 & kc_1 \end{pmatrix}，k\,a_1+kb_1+kc_1=0.$$

即$kA\in W_2$，故W_2是$R^{2\times 3}$的子空间.

例 10　$R^{n\times n}$的如下子集呢?

（1）$W_1=\{A: A^T=A，A\in R^{n\times n}\}$

（2）$W_2=\{A: |A|\neq 0，A\in R^{n\times n}\}$

（3）$W_3=\{A: A^2=A，A\in R^{n\times n}\}$

下面考虑子空间的交与和：

定理 4.2.2 设W_1,W_2是域 F 上的线性空间 V 的两个子空间，则它们的交$W_1\cap W_2$亦是 V 的一个子空间

证：设$\forall k,l\in F,\forall\alpha,\beta\in W_1\cap W_2$

　　则：$k\alpha+l\beta\in W_1, k\alpha+l\beta\in W_2$（ W_1和W_2是子空间）

　　　故　$k\alpha+l\beta\in W_1\cap W_2$.

定理 4.2.3　设W_1,W_2是域 F 上的线性空间 V 的两个子空间，则它们的和：$W_1+W_2\triangleq\{\alpha_1+\alpha_2，\alpha_1\in W_1,\alpha_2\in W_2\}$亦是 V 的一个子空间

证：W_1+W_2 是 V 非空子集合，设$\forall k,l\in F,\forall\alpha,\beta\in W_1+W_2$,

　　有$\alpha_1\in W_1$，$\alpha_2\in W_2$使得$\alpha=\alpha_1+\alpha_2$和$\beta_1\in W_1$, $\beta_2\in W_2$

　　使得$\beta=\beta_1+\beta_2$

故 $k\alpha+l\beta=(k\alpha_1+l\beta_1)+(k\alpha_2+l\beta_2)\in W_1+W_2$.

以上两个定理的结论，可以推广到任意有限个子空间的交与和.

例 11 通常的三维空间R^3中，xoy平面L_1与xoz平面L_2均是R^3的子空间$L_1\cap L_2$是 x 轴，L_1+L_2是全空间 R^3，均是子空间.

例 12　令 $F^{n\times n}$ 表示 F 上所有 n 阶矩阵构成的线性空间，求证：F 上全体 n 阶对称矩阵的集合 S 与反对称矩阵的集合 T 均是$F^{n\times n}$子空间，且有：$S\cap T=\{0\}$，$S+T=F^{n\times n}$.

§4.3 向量组的线性无关性

一、线性组合

向量组之间的线性无关在研究线性空间的构造时，起着极为重要的作用. 以下提到线性空间 V，都是指给定数域 F 上的.

定义 4.3.1（线性组合） 设α_1，α_2，…，α_m是线性空间 V 中的 m 个向量k_1，k_2，…，k_m是 F中任意 m 个数，则称和式：

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m \qquad (4.2)$$

为向量α_1，α_2，…，α_m的一个线性组合.

若$\alpha\in V$是α_1，α_2，…，α_m的线性组合，我们也说：

α 可由α_2，…，α_m线性表示.

例 13 设$\alpha_1=[1，2，-1]^T$，$\alpha_2=[0，1，2]^T$，$\alpha_3=[1，1，-1]^T$是R^3的向量，由于$\alpha_1-2\alpha_2+3\alpha_3=[4，3，-8]^T$，故：

$[4，3，-8]^T$可由α_1，α_2，α_3表示！

例 14　R^3中的任意向量都是三维基本向量的线性组合

证：$\alpha=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\alpha_1\begin{bmatrix}1\\0\\0\end{bmatrix}+\alpha_2\begin{bmatrix}0\\1\\0\end{bmatrix}+\alpha_3\begin{bmatrix}0\\0\\1\end{bmatrix}$.

例 15　零向量 0 可由 V 中任意一组向量线性表示

证：$0=0\alpha_1+0\alpha_2+\cdots+0\alpha_m$.

例 16　向量组$\{\alpha_1，\alpha_2，\cdots，\alpha_m\}$中每一个向量均可由这一组向量线性表示

证：　$\alpha_i = 0\alpha_1 + \cdots + 0\alpha_{i-1} + 1\alpha_i + 0\alpha_{i+1} + \cdots 0\alpha_m.$

习惯上，常常将线性组合写成矩阵的形式：

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = [\alpha_1，\alpha_2，\cdots，\alpha_m]\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_m \end{bmatrix}.$$

注意：这只是记号，一般来说$[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_m]$不是数域上的矩阵！

例 17　$A \in R^{m\times n}, X = [x_1，x_2，\cdots，x_n]^T \in R^n$，则AX是 A 的列向量的线性组合

证：将 A 分块：$A = [\alpha_1，\alpha_2，\cdots，\alpha_n]$

$$AX = [\alpha_1，\alpha_2，\cdots，\alpha_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \qquad (4.3)$$

定理 4.3.2　若向量 r 可$\mathrm{\alpha_1}，\mathrm{\alpha_2}，\cdots，\mathrm{\alpha_r}$线性表示，而每一个$\mathrm{\alpha_i}$又都可由$\mathrm{\beta_1}，\mathrm{\beta_2}，\cdots，\mathrm{\beta_s}$线性表示，则r 亦可由$\mathrm{\beta_1}，\mathrm{\beta_2}，\cdots，\mathrm{\beta_s}$线性表示.

证：$r = \sum_i k_i\alpha_i, \alpha_i = \sum_j l_{ij}\beta_j \Rightarrow r = \sum_j\left(\sum_i k_i l_{ij}\right)\beta_i$

或：$\mathrm{r} = [\mathrm{\alpha_1}，\mathrm{\alpha_2}，\cdots，\mathrm{\alpha_r}]\begin{bmatrix} \mathrm{k_1} \\ \mathrm{k_2} \\ \vdots \\ \mathrm{k_r} \end{bmatrix}$

$$= [\mathrm{\beta_1}，\mathrm{\beta_2}，\cdots，\mathrm{\beta_s}]\begin{bmatrix} \mathrm{l_{11}} & \mathrm{l_{21}} & \cdots & \mathrm{l_{r1}} \\ \mathrm{l_{12}} & \mathrm{l_{22}} & \cdots & \mathrm{l_{r2}} \\ \vdots & \vdots & & \vdots \\ \mathrm{l_{1s}} & \mathrm{l_{2s}} & \cdots & \mathrm{l_{rs}} \end{bmatrix}\begin{bmatrix} \mathrm{k_1} \\ \mathrm{k_2} \\ \vdots \\ \mathrm{k_r} \end{bmatrix}. \qquad (4.4)$$

定理 4.3.3　设$\mathrm{\alpha_1}，\mathrm{\alpha_2}，\cdots，\mathrm{\alpha_m} \in \mathrm{V}$，则由$\mathrm{\alpha_1}，\mathrm{\alpha_2}，\cdots，\mathrm{\alpha_m}$的所有线性组合构成的集合：

$$L(\alpha_1, \alpha_2, \cdots, \alpha_m) = \{k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m : k_i \in F\} \quad (4.5)$$

是 V 的一个子空间！

证：任取β、$\gamma \in L(\alpha_1, \alpha_2, \cdots, \alpha_m)$，k、$l \in F$.

则$\beta = p_1\alpha_1 + p_2\alpha_2 + \cdots + p_m\alpha_m$，$\gamma = q_1\alpha_1 + q_2\alpha_2 + \cdots + q_m\alpha_m$

故$k\beta + l\gamma = (kp_1 + lq_1)\alpha_1 + \cdots + (kp_m + lq_m)\alpha_m \in L(\alpha_1, \cdots, \alpha_m)$.

常称$L(\alpha_1, \alpha_2, \cdots, \alpha_m)$为由$\alpha_1, \alpha_2, \cdots, \alpha_m$生成的子空间，并称$\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$为$L(\alpha_1, \alpha_2, \cdots, \alpha_m)$的生成元组

举例：$\left\{[1, 0, 0]^T, [0, 1, 0]^T\right\}$.

定义（向量组等价） $L(\alpha_1, \alpha_2, \cdots, \alpha_r) = L(\beta_1, \beta_2, \cdots, \beta_s)$则称两个向量组$\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$与$\{\beta_1, \beta_2, \cdots, \beta_s\}$等价.

注： 1、显然是个等价关系，满足自反、对称、传递三种规律

2、两向量组等价，所含向量个数未必相同

定理 4.3.4 （等价的另一种表述）：V 中两向量组 $S_1 = \{\alpha_1, \alpha_2, \cdots, \alpha_r\}$与 $S_2 = \{\beta_1, \beta_2, \cdots, \beta_s\}$等价的充要条件是：$S_1$中每个$\alpha_i$均可由 $\beta_1, \beta_2, \cdots, \beta_s$线性表示，且$S_2$中每个$\beta_i$可由 $\alpha_1, \alpha_2, \cdots, \alpha_r$线性表示

证：必要性：设S_1与S_2等价，则$\alpha_i \in S_1 = S_2$，$\beta_i \in S_2 = S_1$.

充分性：任取$\alpha \in L(S_1)$,则 α 可由$\alpha_1, \alpha_2, \cdots, \alpha_r$线性表示，由**定理 4.3.5** 得 α 亦可由 $\beta_1, \beta_2, \cdots, \beta_s$线性表示，故$\alpha \in L(S_1)$亦即：

$$L(S_1) \subseteq L(S_2)$$

同理可证$L(S_2) \subseteq L(S_1)$,故$L(S_1) = L(S_2)$,

由于$\mathrm{k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m}$可写成：$[\alpha_1，\alpha_2，\cdots，\alpha_m]\begin{bmatrix}k_1\\k_2\\\vdots\\k_n\end{bmatrix}$的形式，

故上述定理亦可表述为：两向量组等价的充要条件是：存在两个矩阵$\mathrm{K_{s\times r}}$和$\mathrm{M_{r\times s}}$,使得：

$$[\alpha_1，\alpha_2，\cdots，\alpha_r]=[\beta_1，\beta_2，\cdots，\beta_s]K$$
$$[\beta_1，\beta，\cdots，\beta_s]=[\alpha_1，\alpha_2，\cdots，\alpha_r]M$$

同时成立 .

例 18 $\mathrm{R^3}$中向量组$\mathrm{T_1=\{e_1,e_2,e_3\}}$与$\mathrm{T_2=\{e_1,e_2,e_3，\beta\}}$，其中$\mathrm{e_i}$是三维基本向量，$\beta=[\mathrm{b_1,b_2,b_3}]\neq \mathrm{e_i}$，则$\mathrm{T_1}$与$\mathrm{T_2}$是否等价?

定义（线性相关、线性无关） 设$\alpha_1，\alpha_2，\cdots，\alpha_m\epsilon$线性空间 V.若 F 中存在不全为零的 m 个数$\mathrm{k_1，k_2，\cdots，k_m}$,使得：

$$\mathrm{k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m=0} \qquad （4.6）$$

时称$\alpha_1，\alpha_2，\cdots，\alpha_m$线性相关，否则，称$\alpha_1，\alpha_2，\cdots，\alpha_m$线性无关.

由定义：线性无关是说：$\mathrm{k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m=0}$对$\mathrm{k_i}$来说只有零解，即：

由之可推出：$\mathrm{k_1=k_2=\cdots=k_m=0}$，只要找到一组不全为 0 的$\mathrm{k_i}$,则$\alpha_1，\alpha_2，\cdots，\alpha_m$线性相关.

例 19 R^3中，$\alpha_1=[1，2，-1]$，$\alpha_2=[-1，0，1]$，$\alpha_3=[-1，4，1]$线性相关，因为：$2\alpha_1+3\alpha_2-\alpha_3=0$.

例 20 对于两个非零向量α_1，α_2来说，线性相关$\Leftrightarrow$ α_1与α_2成比例.

证：① $\alpha_1=\mathrm{k}\alpha_2\Rightarrow\alpha_1-\mathrm{k}\alpha_2=0$.

②$\exists$不全为0的k_1, k_2,使$k_1\alpha_1 - k_2\alpha_2 = 0$,设$k_1 \neq 0$,则 $\alpha_1 = -\frac{k_2}{k_1}\alpha_2$.

例 21 平面内不共线的两矢量线性无关，R^3中不共面的三个矢量线性无关.

例 22 n 维基本向量组 $\{e_1, e_2, \cdots, e_n\}$线性无关.

例 23 R^3中$\alpha_1 = [1, 2, 1]$，$\alpha_2 = [1, -2, 2]$，$\alpha_3 = [2, 0, 1]$是否线性无关?

解：设$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_3\alpha_3 = 0$

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

克莱姆法则！n 个 n 维向量类似!

n 个 m 维向量呢？需要线性

定理 4.3.6 向量组$\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$线性相关的充要条件是：其中至少有一个向量可由其它向量线性表示.

证明：

必要性:存在不全为0的k_i,使$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = 0$,设$k_i \neq 0$,则$\alpha_i = -\frac{k_1}{k_i}\alpha_1 + \left(-\frac{k_2}{k_i}\alpha_2\right) + \cdots + \left(-\frac{k_1}{k_i}\alpha_m\right)$.（和式不含$\alpha_i$）

充分性：设$\alpha_i = l_1\alpha_1 + l_2\alpha_2 + \cdots + l_m\alpha_m$,（和式不含$\alpha_i$），

则：$l_1\alpha_1 + l_2\alpha_2 + \cdots - 1\alpha_i + l_m\alpha_m = 0$．证毕.

定理 4.3.6 的逆否表述:向量组$\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$线性无关的充要条件是：其中任何一个向量都不能由其它向量线性表示

此定理说明：线性相关的向量组是“不独立”的，而线性无关的向量组则彼此“独立”.

定理 4.3.7 若$\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$线性无关，而$\{\beta, \alpha_1, \cdots, \alpha_m\}$线性相关.则：$\beta$可由$\alpha_1, \alpha_2, \cdots, \alpha_m$唯一地线性表示出来.

证：存在不全为 0 的 k,k_1，k_2，$\cdots$，k_m,使$k\beta+k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m$=0，但$k\neq 0$,因为$k=0\Rightarrow\alpha_1$，$\alpha_2$，$\cdots$，$\alpha_m$线性相关.矛盾.

故$\beta=-\frac{k_1}{k}\alpha_1+(-\frac{k_2}{k}\alpha_2)+\cdots+(-\frac{k_1}{k}\alpha_m)$ 可由它们线性表示.

假设有两种不同的表示：$\begin{cases}\beta= k_1\alpha_1+\cdots+ k_m\alpha_m\\ \beta= k_1'\alpha_1+\cdots+ k_m'\alpha_m\end{cases}$.

则$(k_1-k_1')\alpha_1+(k_2-k_2')\alpha_2+\cdots+(k_m-k_m')\alpha_m=0$.

因此，k_1=k_1'，k_2=k_2',⋯. k_m=k_m'.

定理4.3.7′ 若{β，α_1，α_2，$\cdots$，α_m}线性相关，且β可由 α_1，α_2，$\cdots$，α_m唯一线性表示，则{α_1，α_2，$\cdots$，α_m}必线性无关.

证：反证法.由于$\beta=\sum_i l_i\alpha_i$,假设{α}线性相关，

则存在k_i使$\sum_i k_i\alpha_i=0$，故

$\beta=\sum_i l_i\alpha_i$

$\beta=\sum_i l_i\alpha_i$+0

$\beta=\sum_i(l_i+k_i)\alpha_i$ 不唯一！矛盾！

一些性质：

1、含零向量的向量组一定线性相关

2、若{α_1，α_2，$\cdots$，α_m}中有两向量成比例，则 α_1，α_2，$\cdots$，α_m 线性相关.

3、若$S_1\subseteq S_2$，则S_1相关$\Rightarrow S_2$相关；S_2无关$\Rightarrow S_1$无关. (反证法)

例 24 设$\{\alpha_1，\alpha_2，\ldots，\alpha_n\}$线性无关. $\{\beta_1，\beta_2，\ldots，\beta_n\}$中每个$\beta_i$均可由$\alpha_1$，$\alpha_2$，…，$\alpha_n$线性表示，即：

$$[\beta_1\ \beta_2\ \ldots\ \beta_n]=[\alpha_1\ \alpha_2\ \ldots\ \alpha_n]K \qquad (4.7)$$

其中 K 为 n 阶方阵，求β_1，β_2， …，β_n线性无关的充要条件.

解：①令$[\beta_1\ \ \beta_2\ \ \dots\ \ \beta_n][x_1\ \ x_2\ \ \dots\ \ x_n]^T=0$

则$[\alpha_1\ \ \alpha_2\ \ \dots\ \ \alpha_n]K[x_1\ \ x_2\ \ \dots\ \ x_n]^T=0$

又因为，$\alpha_1\ \ \alpha_2\ \ \dots\ \ \alpha_n$线性无关 故$K[x_1\ \ x_2\ \ \dots\ \ x_n]^T=0$，

若$|K|\neq 0 \Rightarrow [x_1\ \ x_2\ \ \dots\ \ x_n]^T=0$,即$\beta_1\ \ \beta_2\ \ \dots\ \ \beta_n$线性无关（充分条件）.

②若$|K|=0$，则存在 $PAQ=\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$

故$[\beta_1\ \ \beta_2\ \ \dots\ \ \beta_n]Q=[\alpha_1\ \ \alpha_2\ \ \dots\ \ \alpha_n]P^{-1}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$.

因为，Q 中没有零列，$\Rightarrow \beta_1\ \ \beta_2\ \ \dots\ \ \beta_n$线性相关.

综上，$\beta_1\ \ \beta_2\ \ \dots\ \ \beta_n$线性无关$\Leftrightarrow |K|\neq 0$.

在学习线性方程组的一般理论后，$|K|=0$ 的情况也可以由方程组来讨论，且$\{\alpha_i\}$，$\{\beta_i\}$数目不等的情况，亦可用方程组来讨论！

例 25 设α_1，α_2， …，α_n线性无关，求证：$\beta_1=\alpha_1+\alpha_2$，$\beta_2=\alpha_2+\alpha_3,\dots,\ \beta_n=\alpha_n+\alpha_1$,线性无关的充要条件是：n 是奇数.

证：$$[\beta_1\ \ \beta_2\ \ \dots\ \ \beta_n]=[\alpha_1\ \ \alpha_2\ \ \dots\ \ \alpha_n]\begin{bmatrix} 1 & 0 & 0 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$K_n\neq 0 \Leftrightarrow \beta_1$，$\beta_2$， …，$\beta_n$线性无关

按第一行展开， $K_n=1+(-1)^{n+1}$.

例 26 R^3中向量组$T_1=\{e_1,e_2,e_3\}$与$T_2=\{e_1,e_2,e_3,\ \beta\}$，其中$e_i$是三维基本向量，$\beta=[b_1,b_2,b_3]\neq e_i$，则$T_1$与$T_2$是否等价?

定义（极大无关组） 若向量组 S= $\{\alpha_1，\alpha_2，\dots，\alpha_n\}$的一个部分组 T= $\{\alpha_{i1}，\alpha_{i2}，\dots，\alpha_{ir}\}(r\leq m)$满足如下条件：

（1）α_{i1}，α_{i2}， …，α_{ir}线性无关

（2）S 中每一个α_i均可由α_{i1}，α_{i2}， …，α_{ir}线性表示

则称 T 是 S 的一个极大线性无关部分组，简称极大线性无关组或极大无关组.

说明：①由于$T \subseteq S$，故 T 中每个向量均可由 S 表示，又根据定义，S 亦可由 T 表示，故可知 S 和 T 等价.

②若任取 S 中不属于 T 的向量，加到 T 中，则 T 就会变成线性相关.

③以上分析说明：T 是按照与 S 等价的前提下，能够选取的最大的线性无关的向量组，后面会用“秩”来说明.

例 27　求$\alpha_1 = [1，2，4]$，$\alpha_2 = [-1，2，0]$，$\alpha_3 = [0，4，4]$的极大线性无关组

解：$\alpha_1 \neq 0$，故$\{\alpha_1\}$线性无关，又$\alpha_2 \neq k\alpha_1$,故$\{\alpha_1，\alpha_2\}$线性无关

又$\alpha_3 = \alpha_1 + \alpha_2$，故$\{\alpha_1，\alpha_2\}$即极大无关组.

同理可验证$\{\alpha_1，\alpha_3\}$，$\{\alpha_2，\alpha_3\}$也都是.

推广此方法可有如下定理：

定理 4.3.8（极大无关组的存在性）：V 中不全为 0 的有限向量组 S=$\{\alpha_1，\alpha_2， …，\alpha_n\}$一定存在极大无关组.

证：不失一般性，设$\alpha_{i1} = \alpha_1 \neq 0$，则令$S_1 = \{\alpha_1\}$

考察α_2若α_2能被s_1表示，则令$s_2 = s_1$

若α_2不能被s_1表示，则令$s_2 = s_1 \cup \{\alpha_2\} = \{\alpha_{i1}, \alpha_{i2}\}$

如此下去，由于个数有限，总可完成.

注：筛选法十分繁琐，对 n 元向量组，我们以后发展其它方法！（初等变换法）

例 28　求向量组R^n的一个极大无关组.

解：令$e_1, e_2, \cdots, e_n$代表R^n中的基本向量，

即： $e_1 = [1,0,0,\cdots,0]^T, e_2 = [0,1,0,\cdots,0]^T, \cdots, e_n = [0,0,0,\cdots,1]^T$

则易证明：

①$e_1, e_2, \cdots, e_n$线性无关

②$\forall \alpha = [x_1, x_2, \cdots, x_n]^T \in R^n$，$\alpha = x_1e_1 + x_2e_2 + \cdots + x_ne_n$

故$\{e_1, e_2, \cdots, e_n\}$是$R^n$的一个极大无关组.

$\{e_1, e_1 + e_2, \cdots, e_1 + e_2 + \cdots + e_n\}$是$R^n$的一个极大无关组吗？

向量组的极大无关组并不唯一！

定理 4.3.9 V 中两向量组:

$T_1 = \{\alpha_1，\alpha_2，\ldots，\alpha_{r1}\}, T_2 = \{\beta_1，\beta_2，\ldots，\beta_{r2}\}$.

若T_1能被T_2线性表示，且$r_1 > r_2$,则向量组T_1线性相关!

证：按条件，有：$[\alpha_1，\alpha_2，\ldots，\alpha_{r1}] = [\beta_1，\beta_2，\ldots，\beta_{r2}]K_{r_2\times r_1}$

其中 K 为$r_2 \times r_1$矩阵，即：$A = BK$.

故$r(K) \le r_2$,存在可逆$P_{r_2\times r_2}, Q_{r_1\times r_1}$,使：

$$PKQ = \begin{bmatrix} I_{r(k)} & 0 \\ 0 & 0 \end{bmatrix}$$

即：$KQ = P^{-1}\begin{bmatrix} I_{r(k)} & 0 \\ 0 & 0 \end{bmatrix} = \lfloor K'_{r_2\times r_2} \quad \vdots \quad O_{r_2\times(r_1-r_2)} \rfloor$

故 AQ=$[\alpha_1，\alpha_2，\ldots，\alpha_{r1}]Q = [\beta_1，\beta_2，\ldots，\beta_{r2}, O_{r2\times(r1-r2)}]$

考虑第r_2+1列，$q_{1r_2+1}\alpha_1 + q_{2r_2+1}\alpha_{2+}q_{r_1r_2+1}\alpha_{r_1} = 0$

而$|Q| \ne 0$无零列，故$\alpha_1，\alpha_2，\ldots，\alpha_{r1}$线性相关.证毕！

即：若一组向量可由另一组表示 ，且比另一组更多，则它们必线性相关！

定理 4.3.9 的等价表述（逆否命题）：设 V 中两组向量组：

$$T_1=\{\alpha_1, \alpha_2, \ldots, \alpha_{r1}\}\text{与}T_2=\{\beta_1, \beta_2, \ldots, \beta_{r2}\}$$

若T_1能被T_2线性表示，且T_1线性无关.则$r_1 \le r_2$.

推论 4.3.10 若两个线性无关的向量组等价，则它们所含向量个数相等.

推论 4.3.11 两个等价的向量组，它们的极大无关组所含向量的个数相等. 特别地，同一个向量组的任意两个极大无关组所含的向量的个数相等.

由于一个向量组的极大无关组中向量个数的唯一性，可引入向量组的“秩”的概念，即：

定义（向量组的秩） 一个向量组的极大线性无关组所含的向量的个数，称为向量组的秩. 规定由零向量构成的向量组$\{0\}$的秩为零.

推论 4.3.12 等价的向量组秩相等.（反之未必）

设向量组 T 的秩为 r，则由定理 4.3.7 可知：T 中任意一个线性无关向量组所含向量个数不超过 r（极大无关组中的个数）这个也是“极大”一词的由来.

T 中与 T 等价的任意一个向量组所含向量的个数≥ r.

例 29 由于$\{e_1, e_2, \cdots, e_n\}$是$R^n$的一个极大无关组.故$R^n$的秩为 n，故$R^n$中任意 n+1 个向量一定线性相关！

作业：

1.若$\alpha_1, \alpha_2, \alpha_3$线性无关，问$\alpha_1-\alpha_2$，$\alpha_2-\alpha_3$，$\alpha_3-\alpha_1$是否线性无关？

2.若$\alpha_1, \alpha_2, \alpha_3, \alpha_4$线性无关，问$\alpha_1+\alpha_2$，$\alpha_2+\alpha_3$，$\alpha_3+\alpha_4$，$\alpha_4+\alpha_1$是否线性无关？

3.证明$\alpha_1+\alpha_2, \alpha_2+\alpha_3, \alpha_3+\alpha_1$线性无关的充分必要条件是$\alpha_1, \alpha_2, \alpha_3$线性无关.

4.设向量β可以由$\alpha_1,\alpha_2,\cdots$，α_r线性表示，但不能由$\alpha_1,\alpha_2,\cdots$，α_{r-1}线性表示.证明向量组$\{\alpha_1,\alpha_2,\cdots,\alpha_{r-1},\alpha_r\}$与向量组$\{\alpha_1,\alpha_2,\cdots,\alpha_{r-1},\beta\}$等价.

5.设向量组$\alpha_1,\alpha_2,\cdots$，α_r线性无关，而$\alpha_1,\alpha_2,\cdots$，α_r，β,γ线性相关.证明，或者β 与 γ中至少有一个可以由$\alpha_1,\alpha_2,\cdots$，α_r线性表示，或者向量组$\{\alpha_1,\alpha_2,\cdots,\alpha_r,\beta\}$与$\{\alpha_1,\alpha_2,\cdots,\alpha_r,\gamma\}$等价.

§4.4 n 元向量组与矩阵的关系

N 元向量空间是F_n一种极其重要的线性空间，这是由于：

1、N 元向量空间是最“直观”、最“具体”的向量空间之一；

2、N 元向量与矩阵之间有密切的联系 n 元向量空间的性质可以通过研究矩阵的性质得到，而且计算方便.

3、(后面会学到）任意一个一般的向量空间都与某个 n 元向量的空间“ 同构”(元素一一对应)，因此，研究清楚了 n 元向量空间的性质,即可得到任意一个一般的向量空间的性质.

矩阵$A_{m\times n}$可按行（或列）分组成行向量组（或列向量组）：

A=$[\alpha_1,\alpha_2,\cdots,\alpha_n]$或 A=$\begin{bmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_m\end{bmatrix}$（其中$\alpha_i$是列向量，$\beta_i$向是行向量）.

定理 4.4.1 矩阵$A_{m\times n}$的行（列）初等变换不改变其列（行）向量的线性相关性和线性组合关系.

具体说来，即：若$A_{m\times n}$经有限次初等变换成$B_{m\times n}$，则：

1.A 的列向量组$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$中任意一个部分组

$\{\alpha_{i1},\alpha_{i2},\cdots,\alpha_{is}\}$与 B 的列向量组$\{\beta_1,\beta_2,\cdots,\beta_n\}$中相对应的部分组

$\{\beta_{i1},\beta_{i2},\cdots,\beta_{is}\}$同为线性无关或线性相关的向量组.

2.若 A 的一个列向量α_j可由$\alpha_{i1},\alpha_{i2},\cdots,\alpha_{is}$线性表示，则 B 中与$\alpha_j$相对应的列向量可由$\beta_{i1},\beta_{i2},\cdots,\beta_{is}$线性表示，且表示系数完全一致.

证明：依条件，存在可逆矩阵 $P_{m\times m}$，使$PA=B$.

即： $P[\alpha_1,a_2,\cdots,\alpha_n]=[\beta_1,\beta_2,\cdots,\beta_n]$，

即：$P\alpha_i=\beta_i$，$\alpha_j=\left[\alpha_{i1},\ \alpha_{i2},\cdots,\alpha_{is}\right]X$,

$$P\alpha_j=\beta_j=P\left[\alpha_{i1},\ \alpha_{i2},\cdots,\alpha_{is}\right]X=\left[\beta_{i1},\ \beta_{i2},\cdots,\beta_{is}\right]X. \qquad (4.8)$$

（对行向量组，转置即可得证！）

设$A_{m\times n}\in F^{m\times n}$，其秩$r(A)=\ r$，则对 A 有限次行初等变换可化为简化行梯形B，且B中有 r 个主 1：

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

可见，B 中主 1 所在的r列是B的列向量组的极大线性无关组.由上述定理 4.4.1，A 的列向量组中与B的极大无关组对应的向量组即 A 的极大无关组. 故A的列向量组的秩等于r.

由于$r(A^T)=r(A)$，故对A^T同样讨论，即有：A 的行向量组的秩也等于r，由此可得如下定理：

定理 4.4.2 矩阵的行向量组与列向量组有相等的秩，且都等于这个矩阵的秩.

推论 4.4.3 设$A_{m\times n}$，则：

（1）$m>n$时，A 的行向量组必定线性相关；

（2）$m<n$时，A 的列向量组必定线性相关；

（3） $m=n$时，A 的行（列）向量组线性无关的充分必要条件是：$|A|\neq 0$;

（4）$r(A)=m$时，行向量组线性无关；$r(A)<m$时，行相关；

（5）$r(A)=n$时，列向量组线性无关；$r(A)<n$时，列相关.

由于矩阵的秩与其行（列）向量组的秩相等，因此亦可反过来用其行（列）向量组的秩来研究矩阵的秩. 下面给出矩阵秩不等式的重新推导. 借助于线性空间的概念，重新理解秩的不等式.

定理 4.4.4　设$A_{m\times n}$，$B_{m\times n}$，则$r(A+B)\leq r(A)+r(B)$.

证：设 A 和 B 的列向量组的极大无关组分别为：

$$\alpha_1,\alpha_2,\cdots,\alpha_{r(A)}\text{和}\beta_1,\beta_2,\cdots,\beta_{r(B)}$$

则$A+B$的每一个列向量均可由向量组

$C=\{\alpha_1,\alpha_2,\cdots,\alpha_{r(A)},\beta_1,\beta_2,\cdots,\beta_{r(B)}\}$ 表示，

即$r(A+B)\leq r(C)\leq r(A)+r(B)$.

定理 4.4.5　设$A_{m\times n}$，$B_{n\times k}$，则 $r(AB)\leq\min\{r(A), r(B)\}$.

证 ：设$C=AB$，对 C 和 A 按行分块 ，有 ：

$$[r_1,r_2,\cdots,r_n]=[\alpha_1,\alpha_2,\cdots,\alpha_n]\begin{bmatrix}b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots\\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{bmatrix}.$$

即：C 的每一个列向量均可由 A 的列向量线性表示. 故：

$r(C)\leq r(A)$，同理可证$r(C)\leq r(B)$，证毕.

定理 4.4.6 设$A_{m\times n}$，$B_{n\times k}$，则$r(C)\geq r(A)+r(B)-n$.

此定理亦可用向量组的理论证明.但不比用矩阵理论证明更简单，不再赘述，可参见陈仲《大学数学》 P_{152}.

由于初等变换不改变线性关系，我们可以得到如下三类问题的一种解法：

1. 判断一给定的 n 元向量组是否线性无关；

2. 求一给定的 n 元向量组的极大无关组；

3. 求一给定的 n 元向量如何用已知的线性无关的 n 元向量组线性表示.

后面将给出一些典型例题.

三类最基本的题型：

1. 判断给定 n 元向量组的线性相关性：

例 30 设$\alpha_1=[1,4,1,0,2]^T$，$\alpha_2=[2,5,-1,3,2]^T$, $\alpha_3=[-1,2,5,6,2]^T$，$\alpha_4=[0,2,2,-1,0]^T$. 问向量组$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$线性相关还是线性无关?

解：解法一：按线性无关的定义，设k_i满足：

$k_1\alpha_1+k_2\alpha_2+k_3\alpha_3+k_4\alpha_4=0$求解$k_i$看有无非零解.

即：$[\alpha_1,\alpha_2,\alpha_3,\alpha_4]\begin{bmatrix}k_1\\k_2\\k_3\\k_4\end{bmatrix}=0$，

即：$AK=0$.（线性方程组的一般理论求解.）

解法二：向量组的秩⇒ 矩阵的秩，行变换不改变列向量间的线性关系！

将矩阵$[\alpha_1,\alpha_2,\alpha_3,\alpha_4]\xrightarrow{\text{行变换}}$简化行梯形

若排成的矩阵 A 正好是方阵，则可直接由|A|是否为 0 判断！

2. 求给定的 n 元向量的极大无关组：

例 31 同上一题，$\alpha_1=[1,4,1,0,2]^T$， $\alpha_2=[2,5,-1,3,2]^T$，$\alpha_3=[-1,2,5,6,2]^T$，$\alpha_4=[0,2,2,-1,0]^T$, 求$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$的极大无关组.

解：由定理 4.4.1 行变换不改变列向量的线性关系，方法仍是：

将列向量排成矩阵$[\alpha_1,\alpha_2,\alpha_3,\alpha_4]$，通过线性变换化为简化行梯形，则主 1 对应的列向量即极大无关组.

若题目给的是行向量，则可转置成列向量再求解.

3. 求一给定的 n 元向量如何用其它 n 元向量组线性表示：

例 32　同上一题，$\alpha_1=[1,4,1,0,2]^{\mathrm{T}}$，　$\alpha_2=[2,5,-1,3,2]^{\mathrm{T}}$，

$$\alpha_3=[-1,2,5,6,2]^{\mathrm{T}}，\alpha_4=[0,2,2,-1,0]^{\mathrm{T}},$$

求α_3能否由α_1，α_2，α_4线性表示？若能，则求之！

解：方法一：设$\alpha_3=x_1\alpha_1+x_2\alpha_2+x_4\alpha_4$，则写成矩阵形式，即

$[\alpha_1,\alpha_2,\alpha_4][x_1,x_2,x_3]^T=\alpha_3$,即:$AX=B$的形式.

若 A 是可逆方阵，则$X=A^{-1}B$；否则，需要线性方程组的一般理论.

方法二：由定理 4.4.1，行变换不改变列向量的线性关系，方法仍是：

将列向量排成矩阵，再用行初等变换：$\left[\alpha_1,\alpha_2,\alpha_4，\ \vdots\alpha_3\right]\to$ 简化行梯形.

注：上一题的解法实际提供了已知方阵 A，B，且 A 可逆，求$A^{-1}B$的一种方法：

将 A，B 写成分块矩阵，然后对之进行行初等变换：

$[A\vdots B]\to\ P_l\cdots P_2P_1[A\vdots B]=[I\vdots X]$

则 $X=A^{-1}B$，这是因为：

$$\begin{cases}P_l\cdots P_2P_1A=I\\P_l\cdots P_2P_1B=X\end{cases}\Rightarrow\begin{cases}P_l\cdots P_2P_1=A^{-1}\\X=A^{-1}B\end{cases}.$$

更多典型例题：

例 33 讨论下列向量组的线性相关性：

（1）$[1,2]^{\mathrm{T}}$，$[2,7]^{\mathrm{T}}$，$[-4,5]^{\mathrm{T}}$　　（R^2）

（2）$[1,3,5]^{\mathrm{T}}$，$[1,1,0]^{\mathrm{T}}$，$[-1,1,5]^{\mathrm{T}}$　　（R^3）

（3）$[1,1,3,1]^{\mathrm{T}}$，$[4,1,-3,2]^{\mathrm{T}}$，$[1,0,-1,2]^{\mathrm{T}}$　　（R^4）

例 34 设α_1=$[2,2,-1]^{\mathrm{T}}$，$\alpha_2=[2,-1,2]^{\mathrm{T}}$，$\alpha_3$=$[-1,2,2]^{\mathrm{T}}$

$\beta_1 = [1,0,-4]^T$，$\beta_2 = [4,3,2]^T$，

求证：$\{\alpha_1,\alpha_2,\alpha_3\}$线性无关，并将$\beta_1$，$\beta_2$用其表示出.

解：可判断行列式，但还须求线性表示，故直接对

$[{\alpha_1}^T,{\alpha_2}^T,{\alpha_3}^T, \vdots\ {\beta_1}^T,{\beta_2}^T]$进行行初等变换.

例 35 设$\alpha_1 = [1,-1,2,4]^T,\alpha_2 = [0,3,1,2]^T,\alpha_3 = [3,0,7,14]^T$,

$\alpha_4 = [1,-2,2,0]^T$，$\alpha_5 = [2,1,5,10]^T$，求$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5\}$的一个极大无关组，并将其它向量用极大无关组线性表示.

例 36 设$A_{m\times n}$，求证：若存在$B_{n\times m}$，使$BA = I_n$，则A的列向量组线性无关.

证：方法一：按定义，设存在$[x_1,x_2,\cdots,x_n]^T = X$，使$AX = 0$，

$$则BAX = IX = 0 \Rightarrow X = 0$$

方法二：$r(A) \geq r(AB)$

例 37 设$T_1 = \{\alpha_1,\alpha_2,\alpha_3\}$，$T_2 = \{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$，

$T_3 = \{\alpha_1,\alpha_2,\alpha_3,\alpha_5\}$，且$r(T_1) = r(T_2) = 3$，$r(T_3) = 4$，

求证：$T_4 = \{\alpha_1,\alpha_2,\alpha_3,\alpha_5 - \alpha_4\}$的秩$r(T_4) = 4$.

例 38 设$\alpha_1 = [1,0,2,3]^T,\alpha_2 = [1,-1,a+2,1]^T$, $\alpha_3 = [1,2,4,a+8]^T$,

$\alpha_4 = [1,1,3,5]^T$，$\beta = [1,1,b+3,5]^T$.

（1）a，b 为何值时，β可由$\alpha_1,\alpha_2,\alpha_3,\alpha_4$唯一线性表示？

（2）a，b 为何值时，β可由$\alpha_1,\alpha_2,\alpha_3,\alpha_4$线性表示，但不唯一？

（3）a，b 为何值时，β不可由$\alpha_1,\alpha_2,\alpha_3,\alpha_4$线性表示？

解： $[\alpha_1,\alpha_2,\alpha_3,\alpha_4，\beta] \rightarrow$ 简化行梯形.

例 39（矩阵的满秩分解）：设 $A\in F^{m\times n}$，r(A)=r，求证：

存在 P，Q 使得 $A=P_{m\times r}Q_{r\times n}$，其中 P, Q 的秩均为 r.

证明：

$$A=P'\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}Q'=P'\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{m\times n}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times n}Q'$$

$$=[P_1,P_2,\cdots,P_r,P_{r+1},\cdots,P_m]\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{m\times n}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times n}\begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_r \\ Q_{r+1} \\ \vdots \\ Q_n \end{bmatrix}.$$

$$=[P_1,\cdots,P_r,0,\cdots,0]\begin{bmatrix} Q_1 \\ \vdots \\ Q_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}=[P_1,\cdots,P_r]\begin{bmatrix} Q_1 \\ \vdots \\ Q_r \end{bmatrix}$$，证毕.

证法 2：

$$PA=\begin{bmatrix} 0 & 1 & * & 0 & * & 0 & * \\ & & 0 & 1 & * & 0 & * \\ & & & 0 & 0 & \cdots & * \\ & & & & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}=\begin{bmatrix} C_{r\times n} \\ 0_{(m-r)\times n} \end{bmatrix}.$$

$r(C)=r(A)=r.$

$$A=P^{-1}\begin{bmatrix} C \\ 0 \end{bmatrix}=\begin{bmatrix} B_{m\times r} & B_{m\times(m-r)}' \end{bmatrix}\begin{bmatrix} C_{r\times n} \\ 0_{(m-r)\times n} \end{bmatrix}=BC.$$

B是可逆矩阵P^{-1}的前r列，而P^{-1}的列向量是线性无关的，故B的r列向量线性无关，即$r(B)=r$, 同时$r(C)=r$.

作业

1.求下列向量组的极大无关组：

$[1,-2,-1,0,\ 2]^T$, $[1,-2,-1,-3,3]^T$, $[2,-1,0,2,3]^T$, $[3,3,3,3,4]^T$.

2. 已知向量组： (1)$\beta_1=\begin{bmatrix}0\\1\\-1\end{bmatrix}, \beta_2=\begin{bmatrix}a\\2\\1\end{bmatrix}, \beta_3=\begin{bmatrix}b\\1\\0\end{bmatrix}$, 和向量组

(2)$\alpha_1=\begin{bmatrix}1\\2\\-3\end{bmatrix}, \alpha_2=\begin{bmatrix}3\\0\\1\end{bmatrix}, \alpha_3=\begin{bmatrix}9\\6\\-7\end{bmatrix}$的秩相等，且$\beta_3$可以由

$\alpha_1, \alpha_2, \alpha_3$线性表示，求$a$和 b 的值.

§4.5　线性空间的基、维数、坐标

设 V 是数域 F 上的线性空间，若存在有限个向量$\alpha_1, \dots, \alpha_m \in$ V 使：L=$(\alpha_1, \alpha_2, \dots, \alpha_m)$=V，则称 V 是有限维线性空间，否则，则称 V 是无限维线性空间.

下面仅限于讨论有限维空间

定义(基)　设 V 是 F 上有限维线性空间，若 V 中一向量组$\{\alpha_1, \dots, \alpha_n\}$满足：（1）$\alpha_1, \alpha_2, \dots, \alpha_n$线性无关；

（2）$\forall\alpha \in V$，α 均能被$\alpha_1, \alpha_2, \dots, \alpha_n$线性表示.

则称$\{\alpha_1, \alpha_2, \dots, \alpha_n\}$是 V 的一个基.

注：条件(2)称作向量组(在空间 V 中)的完备性. 故“基”即“完备的线性无关组”.

理解：将 V 看成无穷个向量构成的向量组，则基就是其极大无关组.

另外，把握了 V 的基，即可由基的线性组合生成整个空间 V！

例 40　F^n中，自然向量组

$\left\{e_1=[1,0,\dots,0]^T, e_2,=[0,1,\dots,0]^T, e_n=[0,0,\dots,1]^T\right\}$，显然构成$F^n$的一个基，并称为自然基.

例 41　$F^{m\times n}$中，如 F 的 mn 个矩阵

$$E_{ij}=\begin{bmatrix}0 & \cdots & 0 & \cdots & 0\\ \vdots & & \vdots & & \vdots\\ 0 & \cdots & 1 & \cdots & 0\\ \vdots & & \vdots & & \vdots\\ 0 & \cdots & 0 & \cdots & 0\end{bmatrix}$$（只有(i,j)元等于 1，其他均为 0.）

构成的集合$\{E_{ij}\}$，显然是$F^{m\times n}$的一个基. 这个基中共含有$m\times n$个向量.一般来说，线性空间的基并不唯一，例如:

$\{g_1=[1,0,\ldots,0,0]^T, g_2=[1,1,\ldots,0,0]^T, g_3=g_3=[1,1,\ldots,1,1]^T\}$

也构成F^n 的一组基. 但是，类似向量组的极大无关组所含向量个数相同，同样可以证明如下定理：

定理 4.5.1 非零线性空间 V 的所有基都含有相同数目的向量.

由上述定理，我们可非零线性空间 V 以有如下定义：

定义(维数) 非零线性空间 V 的基中所含向量的个数称为 V 的维数，记作 dimV. 零空间没有基，规定其维数为 0.

理解：若将 V 看成无穷个向量构成的向量组，则维数即为该向量组的秩！

例：$\dim F^n$=n，$\dim F^{m\times n}$=$m\times n$.

定理 4.5.2 设 dimV=n，则：

（1）V 中线性无关向量组所含向量个数≤n.(线性无关组)

（2）V 中生元组所含向量个数≥ n.(完备组)

（3）V 的子空间的维数≤ n.

线性空间的基并不唯一，一般地，有如下定理：

定理 4.5.3 设 dimV=n，则 V 中任意 n 个线性无关的向量构成的向量组，都是 V 的一个基.

证: 设$\{e_1,e_2,\ldots,e_n\}$是 V 的一个基. $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$线性无关. 即需要证明$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$亦是 V 的一个基.

注意到$\{\alpha_1,\alpha_2,\dots,\alpha_n\}=\{e_1,e_2,\dots,e_n\}K_{n\times n}$. 由上一节的例题，$\{e_1,e_2,\dots,e_n\}$与$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$均线性无关，则 K 可逆，故：$\{e_1,e_2,\dots,e_n\}=\{\alpha_1,\alpha_2,\dots,\alpha_n\}K^{-1}$故$\forall\alpha\in V$均可用$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$线性表示，证毕.

可得到如下定理:

定理 4.5.4 若$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$是 V 的一个基，则对$\forall\alpha\in V$，$\alpha$ 均可被$\alpha_1,\alpha_2,\dots,\alpha_n$以唯一的方式线性表示.

由于此定理，我们可以引入“坐标”的概念.（以下谈到线性空间的基，均可指有序基，即：基中的向量具有确定的次序.）

定义（坐标） 设 V 是 F 上的线性空间，$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$是 V 的一个基，则对$\forall\alpha\in$"V"，$\alpha$可唯一写成：

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n \qquad (4.9)$$

或：$\alpha=[\alpha_1,\alpha_2,\dots,\alpha_n]\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$，其中$x_i\in F$

称$[x_1,x_2,\dots,x_n]^T$是向量关于基$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$的坐标.

注意:（1）向量的坐标是与基的选取有关的.

（2）当给定一组基时，向量与其坐标之间是一一对应的，即任意给定一个向量，其坐标是唯一确定的；反之，任意给定一个坐标(例如$[x_1,x_2,\dots,x_n]^T$)，则与之对应的向量也是唯一确定的（$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$，且若 $\alpha\neq\beta$，则坐标亦不相同.

向量和坐标之间，更进一步地，还有如下定理：

定理 4.5.5 设$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$是 F 上的一个基，V 中两个向量$\alpha$，$\beta$关于这个基的坐标分别是$[x_1,x_2,\dots,x_n]^T$和$[y_1,y_2,\dots,y_n]^T$,

则：$\alpha+\beta$与 $k\alpha$关于这个基的坐标分别是：

$[x_1,x_2,\dots,x_n]^T+[y_1,y_2,\dots,y_n]^T=[x_1+y_1,\dots,x_n+y_n]^T$

$k[x_1,\ \dots,x_n]^T=[kx_1,\ \dots,kx_n]^T$.

注： 由于向量与坐标之间一一对应，故此定理说明：

向量之间的线性相关性和线性组合关系与向量的坐标之间的线性相关性和线性组合关系完全一致！

由于给定基时，向量与坐标一一对应，且向量间的线性关系与坐标间的线性关系也完全一致. 故对一般的向量，三类基本问题：

1. 判断向量组是否线性无关；
2. 判断向量组的极大无关组；
3. 求某向量用给定的向量组线性表示；

可转化为 n 元向量组的问题来求解.（对于 1、3 两类问题亦可按定义用线性方程组来求解，但不如初等变换法方便.）

例 42 R^4中，$\alpha_1=[1,-1,-1,-1]^T$，$\alpha_2=[-1,1,-1,-1]^T$，$\alpha_3=[-1,-1,1,-1]^T$，$\alpha_4=[-1,-1,-1,1]^T$，$\beta=[1,2,1,1]^T$,则：

（1）求证：$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$是$R^4$的一个基.

（2）求向量β关于基$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$的坐标.

$$\beta=x_1\alpha_1+x_2\alpha_2+x_3\alpha_3+x_4\alpha_4=[\alpha_1,\alpha_2,\alpha_3,\alpha_4]\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}.$$

（答案：β 的坐标为$\left[-\frac{3}{4},-\frac{1}{4},-\frac{3}{4},-\frac{3}{4}\right]^T$）.

注： 求一个向量关于一组基的坐标，实际上就是第三类基本问题：求一个向量如何用给定的向量线性表示.

坐标变换：

向量的坐标与基的选择有关. 对不同的基，同一个向量具有不同的坐标. 为寻求坐标间的关系，先研究两个基之间的关系.

设$\{\mathrm{\alpha}_1,\mathrm{\alpha}_2,\dots,\mathrm{\alpha}_\mathrm{n}\}$与$\{\mathrm{\beta}_1,\mathrm{\beta}_2,\dots,\mathrm{\beta}_\mathrm{n}\}$是线性空间的两个基，则按基的定义，每个$\mathrm{\beta}_\mathrm{j}(\mathrm{j}=1,\dots,,\mathrm{n})$均可被$\mathrm{\alpha}_1,\mathrm{\alpha}_2,\dots,\mathrm{\alpha}_\mathrm{n}$线性表示，即：

$$\begin{cases}\beta_1=C_{11}\alpha_1+C_{21}\alpha_2+\cdots+C_{n1}\alpha_n\\ \beta_2=C_{12}\alpha_1+C_{22}\alpha_2+\cdots+C_{n2}\alpha_n\\ \qquad\cdots\cdots\cdots\\ \beta_n=C_{1n}\alpha_1+C_{2n}\alpha_2+\cdots+C_{nn}\alpha_n\end{cases}\qquad(4.10)$$

写成矩阵形式，即$C=\begin{bmatrix}C_{11}&C_{12}&\cdots&C_{1n}\\C_{21}&C_{22}&\cdots&C_{2n}\\\vdots&\vdots&&\vdots\\C_{n1}&C_{n2}&\cdots&C_{nn}\end{bmatrix}$

$$[\mathrm{\beta}_1,\mathrm{\beta}_2,\dots,\mathrm{\beta}_\mathrm{n}]=[\mathrm{\alpha}_1,\mathrm{\alpha}_2,\dots,\mathrm{\alpha}_\mathrm{n}]\mathrm{C}$$

注：习惯上，基排成行矩阵，放在左边；坐标排成列矩阵，放在右边.

定义（过渡矩阵） 称如上的 C 矩阵为由基$\{\mathrm{\alpha}_1,\mathrm{\alpha}_2,\dots,\mathrm{\alpha}_\mathrm{n}\}$到基$\{\mathrm{\beta}_1,\mathrm{\beta}_2,\dots,\mathrm{\beta}_\mathrm{n}\}$的过渡矩阵.

定理 4.5.6 设方阵 C 是由 n 维线性空间 V 的基$\{\mathrm{\alpha}_1,\mathrm{\alpha}_2,\dots,\mathrm{\alpha}_\mathrm{n}\}$到$\{\mathrm{\beta}_1,\mathrm{\beta}_2,\dots,\mathrm{\beta}_\mathrm{n}\}$的过渡矩阵，则 C 可逆且$\mathrm{C}^{-1}$是由基$\{\mathrm{\beta}_1,\mathrm{\beta}_2,\dots,\mathrm{\beta}_\mathrm{n}\}$到基$\{\mathrm{\alpha}_1,\mathrm{\alpha}_2,\dots,\mathrm{\alpha}_\mathrm{n}\}$的过渡矩阵.

证：由定理条件：$[\beta_1,\beta_2,\dots,\beta_n]=[\alpha_1,\alpha_2,\dots,\alpha_n]C$，

而$\{\mathrm{\beta}_1,\mathrm{\beta}_2,\dots,\mathrm{\beta}_\mathrm{n}\}$是一个基，故$\mathrm{\alpha}_\mathrm{i}(\mathrm{i}=1,\dots,\mathrm{n})$均可由$\mathrm{\beta}$表示，

即：$[\mathrm{\alpha}_1,\mathrm{\alpha}_2,\dots,\mathrm{\alpha}_\mathrm{n}]=[\mathrm{\beta}_1,\mathrm{\beta}_2,\dots,\mathrm{\beta}_\mathrm{n}]\mathrm{D}$，代入可得:

$$[\mathrm{\beta}_1,\mathrm{\beta}_2,\dots,\mathrm{\beta}_\mathrm{n}]=[\mathrm{\alpha}_1,\mathrm{\alpha}_2,\dots,\mathrm{\alpha}_\mathrm{n}]\mathrm{C},$$

又 $\{\mathrm{\beta}\}$是基，故$\mathrm{DC}=\mathrm{I}_\mathrm{n}$，故 C 可逆，且$\mathrm{D}=\mathrm{C}^{-1}$，证毕.

$$\left.\begin{array}{r}[\beta_1,\beta_2,\dots,\beta_n]=[\alpha_1,\alpha_2,\dots,\alpha_n]C\\ [\alpha_1,\alpha_2,\dots,\alpha_n]=[\alpha_1,\alpha_2,\dots,\alpha_n]C^{-1}\end{array}\right\}\text{基的变换公式}$$

下面来看由基的变换诱导出的坐标变换公式：

设向量$\mathrm{\alpha}$关于基$\{\mathrm{\alpha}_1,\mathrm{\alpha}_2,\dots,\mathrm{\alpha}_\mathrm{n}\}$的坐标是$[x_1,x_2,\dots,x_n]^T$，关于基$\{\mathrm{\beta}_1,\mathrm{\beta}_2,\dots,\mathrm{\beta}_\mathrm{n}\}$的坐标是 $[y_1,y_2,\dots,y_n]^T$，则有

$$\alpha=[\alpha_1,\alpha_2,\dots,\alpha_n][x_1,x_2,\dots,x_n]^T=[\beta_1,\beta_2,\dots,\beta_n][y_1,y_2,\dots,y_n]^T$$

设基$\{\alpha_1, \alpha_2, \dots, \alpha_n\}$到$\{\beta_1, \beta_2, \dots, \beta_n\}$的过渡矩阵为 C，则：
$[\beta_1, \beta_2, \dots, \beta_n] = [\alpha_1, \alpha_2, \dots, \alpha_n]C.$
代入上式：并用到坐标的唯一性，以及 C 可逆，即有：

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = C \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad 或： \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = C^{-1} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

上面两式称为坐标变换公式！

注:（1）坐标的变换正好和基的变换相反！

（2）抓“不变量”！向量是不依赖于基的选择的不变量！

例 43　已知R^3中两个基：

$$\left\{\alpha_1 = [1,0,1]^T，\alpha_2 = [1,1,-1]^T，\alpha_3 = [1,-1,1]^T\right\}$$

$$与\left\{\beta_1 = [3,0,1]^T，\beta_2 = [2,0,0]^T，\beta_3 = [0,2,-2]^T\right\}$$

（1）求由基$\{\alpha_1, \alpha_2, \alpha_3\}$到基$\{\beta_1, \beta_2, \beta_3\}$的过渡矩阵.

（2）求向量$\xi=[1,0,-1]^T$在上述两个基下的坐标.

$$\left(答案：C = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}, X = [-1,1,1]^T，Y = [-1,2,0]^T\right)$$

注：求过渡矩阵实际上也是第 3 类基本问题：求一个基中的

向量如何用另一个基线性表示！

线性空间的同构

设$\alpha_1, \alpha_2, \cdots, \alpha_n$是 n 维线性空间$V_n$的一组基，在这组基下，$V_n$中的每个向量都有唯一确定的坐标. 而向量的坐标可以看作R^n中的元素，因此，向量与它的坐标之间的对应就是V_n到R^n的一个映射.

由于R^n中的每个元素都有V_n中的向量与之对应，同时V_n中不同的向量的坐标不同，因而对应R^n中的不同元素. 我们称这样的映射V_n与R^n的一个 1-1 对应的映射. 这个对应的重要性表现在它与运算的关系上

设 $$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$$
$$\beta = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n$$

即向量α，$\beta \in V$在基$\alpha_1, \alpha_2, \cdots, \alpha_n$下的坐标分别为$(a_1, a_2, \cdots, a_n)^T$和$(b_1, b_2, \cdots, b_n)^T$，则

$$\alpha + \beta = (a_1 + b_1)\alpha_1 + (a_2 + b_2)\alpha_2 + \cdots + (a_n + b_n)\alpha_n \qquad (4.12)$$

$$k\alpha = ka_1\alpha_1 + ka_2\alpha_2 + \cdots + ka_n\alpha_n \qquad (4.13)$$

于是$\alpha + \beta$与$k\alpha$的坐标分别为

$$(a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n)^T = (a_1, a_2, \cdots, a_n)^T + (b_1, b_2, \cdots, b_n)^T$$

$$(ka_1, ka_2, \cdots, ka_n)^T = k(a_1, a_2, \cdots, a_n)^T$$

上式表明：在向量用坐标表示后，它们的运算就归结为坐标的运算，因而线性空间V_n的讨论就归结为R^n的讨论.

下面更确切地说明这一点.

定义 设U，V是两个线性空间，如果它们的元素之间有一一对应关系，且这个对应关系保持线性组合的对应关系，那么就称线性空间 U 与 V 同构.

（1） 存在 1-1 对应的映射f;

（2）$\forall\alpha$，β，有 $f(\alpha + \beta) = f(\alpha) + f(\beta)$;

（3）$\forall\alpha \in V$，$\forall k \in F$ ，有$f(k\alpha) = kf(\alpha)$.

n 维线性空间$V = \{\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \mid x_1, x_2, \cdots, x_n \in R\}$与 n 维数组向量空间$R^n$同构.

（1）V_n中的元素α与R^n中的元素$(x_1, x_2, \cdots, x_n)^T$形成一一对应关系.

$$\begin{array}{lccccc} V_n & \alpha = & x_1\alpha_1 + & x_2\alpha_2 + & \cdots + & x_n\alpha_n \\ & & \updownarrow & \updownarrow & & \updownarrow \\ R^n & x = (& x_1, & x_2, & \cdots, & x_n)^T \end{array}$$

（2）设 $\alpha \leftrightarrow (x_1, x_2, \cdots, x_n)^T$

$\beta \leftrightarrow (y_1, y_2, \cdots, y_n)^T$

$$\alpha+\beta \leftrightarrow (x_1, x_2, \cdots, x_n)^T + (y_1, y_2, \cdots, y_n)^T$$
$$\lambda\alpha \leftrightarrow \lambda(x_1, x_2, \cdots, x_n)^T$$

结论： 1．数域 F 上任意两个 n 维线性空间都同构.

2．同维数的线性空间必同构.

3．同构的线性空间之间具有反身性、对称性与传递性.

同构的意义

在线性空间的抽象讨论中，无论构成线性空间的元素是什么，其中的运算是如何定义的，我们所关心的只是这些运算的代数性质．从这个意义上可以说，同构的线性空间是可以不加区别的，而有限维线性空间唯一本质的特征就是它的维数.

作业

1.设三维向量γ在基$\{\alpha_1,\alpha_2,\alpha_3\}$下的坐标为$[1\quad 2\quad 3]^T$，求$\gamma$关于基$\{\alpha_1+\alpha_2,\alpha_1+\alpha_2+\alpha_3,\alpha_1-\alpha_2\}$的坐标.

2.设R^3中有两组基：$\{\alpha_1=[1,1,0]^T,\alpha_2=[0,1,1]^T,\alpha_3=[0,0,1]^T\}$和$\{\beta_1,\beta_2,\beta_3\}$，已知$\{\alpha_1,\alpha_2,\alpha_3\}$到$\{\beta_1,\beta_2,\beta_3\}$的过渡矩阵A为

$A=\begin{bmatrix}1 & 1 & -2\\ -2 & 0 & 3\\ 4 & -1 & -6\end{bmatrix}$，求$\{\beta_1,\beta_2,\beta_3\}$.

3.设R^4中两组基$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$和$\{\beta_1,\beta_2,\beta_3,\beta_4\}$，$\forall\gamma\in R^4$，$\gamma$关于两组基的坐标分别为$X=[x_1,x_2,x_3,x_4]^T$和$Y=[y_1,y_2,y_3,y_4]^T$，已知有坐标变换公式：$y_1=x_1,y_2=x_2-x_1,y_3=x_3-x_2,y_4=x_4-x_3$．求从基$\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$到$\{\beta_1,\beta_2,\beta_3,\beta_4\}$的过渡矩阵C.

第五章：线性方程组

线性方程组的基本概念

线性方程组的基本概念

线性方程组解的结构

§5.1　线性方程组解的结构

克莱姆法则的应用限制：

1、方程个数必须等于未知量的个数；

2、必须系数矩阵行列式非零；

3、需要计算 n+1 个 n 阶行列式、计算量大.

线性方程组的一般形式：

含 m 个方程，n 个未知量的线性方程组一般形式是：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \qquad (5.1)$$

称 m×n 型线性方程组.

若所有 $\mathrm{b_i} = 0$，称齐次线性方程组，若 $\mathrm{b_i}$不全为 0，称非齐次线性方程组.

线性方程组的一个解指的是一个n元向量$[c_1, c_2, \dots, c_n]^T$，当$x_1, x_2, \dots, x_n$分别用$c_1, c_2, \dots, c_n$代入后线性方程组中每个方程都变为恒等式.线性方程组的解的全体称它的解集合.

解方程组就是求其解集合.

若两个线性方程组的解集合相同，则称它们是同解方程组。

线性方程组理论、主要解决三个问题：

1、有没有解？有解的条件是什么？

2、如果有解、究竟有多少解？如何求解？

3、当解集合不止含有一个解时、解与解之间的关系如何？

在具体讨论这三个问题之前，先看线性方程组与矩阵、向量的联系：

原线性方程组可写成$AX = b$的形式，其中$X = [x_1, x_2, \cdots, x_n]^T$.

定义系数矩阵，增广矩阵如下：

系数矩阵：

$$A=\begin{bmatrix} a_{11} & a_{12} & L & a_{1n} \\ a_{21} & a_{22} & L & a_{2n} \\ M & M & O & M \\ a_{m1} & a_{m2} & L & a_{mn} \end{bmatrix} \qquad (5.2)$$

增广矩阵：

$$\mathrm{B}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\mathrm{n}} & b_1 \\ a_{21} & a_{22} & \cdots & a_{1\mathrm{n}} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{\mathrm{m}1} & a_{\mathrm{m}2} & \cdots & a_{\mathrm{mn}} & b_m \end{bmatrix} = [A \vdots b] \qquad (5.3)$$

另一方面，可将系数矩阵按列分块：

$A=[A_1,A_2,\cdots,A_n]$，其中，$\mathrm{A_i}$是列向量.

则原方程组可写成:$\mathrm{A}_1 x_1+A_2 x_2+\cdots+A_n x_n=b$

故原方程组有解的充要条件列向量 b 可写成 A 的列向量组的线性组合，也即向量组$\{\mathrm{A}_1,\mathrm{A}_2,\cdots,\mathrm{A}_\mathrm{n},\mathrm{b}\}$ 与向量组$\{\mathrm{A}_1,\mathrm{A}_2,\cdots,\mathrm{A}_\mathrm{n}\}$等价，即两个向量组的秩相等.

从上述对向量表示的分析，可以讨论线性方程组有解的充要条件，见 hust 线性代数第 4.1.3 节，但给不出求解方法.

后面主要用矩阵理论来讨论.

消元法与初等变换：

中学已学过消元法求解二元、三元等简单线性方程组，例如：

例一：$\begin{cases} 2x_1+2x_2+3x_3=3 \\ x_1+2x_2+4x_3=1 \\ 2x_1+x_2-x_3=5 \end{cases}$ $\left(\begin{cases} 2x_1+2x_2+3x_3=3 \\ x_2+2.5x_3=-0.5 \\ x_3=-1 \end{cases}\right)$

例二：$\begin{cases} x+y+z=3 \\ 2x+3y+z=8 \\ 3x-y+7z=1 \end{cases}$ $\left(\begin{cases} x+y+z=3 \\ y-z=2 \\ 0=0 \end{cases}\right)$

例三：$\begin{cases} x+y+z=3 \\ 2x+3y+z=8 \\ x+3y-z=5 \end{cases}$ $\left(\begin{cases} x+y+z=3 \\ y-z=2 \\ 0=-1 \end{cases}\right)$

以上三例中，对方程组实行了下列三种变换：

1、交换两个方程的位置；

2、用非零的数乘某一个方程；

3、用一个数乘某一个方程后加到另一个方程.

这三种变换统称线性方程组的初等变换.

定理 5.1.1 初等变换将一个线性方程组变成一个与他同解的线性方程组.

证明：设$Ax=b(\ I)$，初等变换成$A'x=b'(\mathrm{II})$，则必有：

$\mathrm{A}'=\mathrm{PA}$，$\mathrm{b}'=\mathrm{Pb}$，且 P 可逆.

若列向量 C 是原方程的一个解，则$\mathrm{AC}=\mathrm{b}$.

$\mathrm{PAC}=\mathrm{Pb}$即$\mathrm{A}'\mathrm{C}=b'$，故C也是(Ⅱ)的一个解.

反之，若 C 是(Ⅱ)的一个解,则$\mathrm{A}'\mathrm{C}=b'$，左乘P^{-1}有

$\mathrm{AC}=\mathrm{b}$. 故 C 也是(Ⅰ)的解.故(Ⅰ)与(Ⅱ)同解，证毕！

§5.2　Gauss 消元法

通过对上面的三个例子、以及定量 5.1.1、求解线性方程组可以通过对其增广矩阵进行初等变换，使增广矩阵变为简化行梯形矩阵.

同三类初等变换将一个线性方程组化成增广矩阵是标准形的线性方程组，这个过程称为 Gauss 消元.

即：对 Ax=b：

$$[A \ \vdots \ b] \ \longrightarrow \ [C \ \vdots \ d]$$（行标准型）

则：$\mathrm{Ax} = \mathrm{b} \iff \mathrm{Cx} = \mathrm{d}.$

例：用 Gauss 消元法求解：

(1) $\begin{cases} x_1 + x_2 = 1 \\ x_1 - x_2 = 3 \\ -x_1 + 2x_2 = -2 \end{cases}$　　(2) $\begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 = 4 \\ x_1 + 3x_2 \qquad - 3x_4 = 1 \\ x_2 - x_3 + x_4 = -3 \\ 7x_2 - 3x_3 - x_4 = 3 \end{cases}$

(3) $\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 2 \\ x_1 + x_2 + x_3 + 2x_4 + 2x_5 = 3 \\ x_1 + x_2 + x_3 + 2x_4 + 3x_5 = 2 \end{cases}$

解：(1) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 3 \\ -1 & 2 & -2 \end{bmatrix} \xrightarrow[c_2 - c_1]{c_3 + c_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 2 \\ 0 & 1 & 1 \end{bmatrix}$

$$\xrightarrow[c_2 \leftrightarrow c_3]{c_2 + 2c_3} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$　　无解.

(2) $\begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 = 4 \\ x_1 + 3x_2 \qquad - 3x_4 = 1 \\ x_2 - x_3 + x_4 = -3 \\ 7x_2 - 3x_3 - x_4 = 3 \end{cases}$

$$\begin{bmatrix}1&-2&3&-4&4\\1&3&0&-3&1\\0&1&-1&1&-3\\0&7&-3&-1&3\end{bmatrix}\xrightarrow[c_3\leftrightarrow c_2]{\substack{c_2-c_1\\c_4-7c_3}}\begin{bmatrix}1&-2&3&-4&4\\0&1&-1&1&-3\\0&5&-3&1&-3\\0&0&1&-2&6\end{bmatrix}$$

$$\xrightarrow[c_4-c_3]{c_3-5c_2}\begin{bmatrix}1&-2&3&-4&4\\0&1&-1&1&-3\\0&0&1&-2&6\\0&0&0&0&0\end{bmatrix}\xrightarrow{\substack{c_1+2c_2-c_3\\c_2+c_3}}$$

$$\begin{bmatrix}1&0&0&0&-8\\0&1&0&-1&3\\0&0&1&-2&6\\0&0&0&0&0\end{bmatrix}$$

(3) $\begin{cases}x_1+x_2+x_3+x_4+x_5=2\\x_1+x_2+x_3+2x_4+2x_5=3\\x_1+x_2+x_3+2x_4+3x_5=2\end{cases}$

$$\begin{bmatrix}1&1&1&1&1&2\\1&1&1&2&2&3\\1&1&1&2&3&2\end{bmatrix}\xrightarrow{\substack{c_3-c_2\\c_2-c_1}}\begin{bmatrix}1&1&1&1&1&2\\0&0&0&1&1&1\\0&0&0&0&1&-1\end{bmatrix}$$

$$\xrightarrow[c_2-c_3]{c_1-c_2}\begin{bmatrix}1&1&1&0&0&1\\0&0&0&1&0&2\\0&0&0&0&1&-1\end{bmatrix}.$$

一般地，设$Ax=b$.（增广：$\mathrm{B}=\left[\mathrm{A}\ \vdots\ \mathrm{b}\right]$）高斯消元法变成：

$Cx=d$.（增广：$\mathrm{D}=\left[\mathrm{C}\ \vdots\ \mathrm{d}\right]$）D的一般形式为：

$$D=\begin{bmatrix}1 & 0 & \cdots & 0 & c_{1r+1} & \cdots & c_{1n} & d_1\\ 0 & 1 & \cdots & 0 & c_{2r+1} & \cdots & c_{2n} & d_2\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & c_{rr+1} & \cdots & c_{rn} & d_r\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d_{r+1}\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0\end{bmatrix}=[C\vdots d\,]$$

注：为了书写方便，得到上面的 D 时有可能会交换列的位置. 交换列相当于交换方程组中未知量的位置，显然不会影响其解. 实际解题时，没有必要交换增广矩阵的列！

对应的方程组 $Cx=d$ 即为：

$$\begin{cases}x_{i1} \qquad\qquad + C_{1r+1}x_{i_{r+1}} + \cdots + C_{1n}x_{i_n} = d_1\\ \quad x_{i2} \qquad + C_{2r+1}x_{i_{r+1}} + \cdots + C_{2n}x_{i_n} = d_2\\ \qquad\qquad \vdots\\ \qquad\quad x_{ir} + C_{rr+1}x_{i_{r+1}} + \cdots + C_{rn}x_{i_n} = d_r\\ \qquad\qquad\qquad 0 = d_{r+1}\\ \qquad\qquad\qquad 0 = 0\\ \qquad\qquad\qquad \vdots\\ \qquad\qquad\qquad 0 = 0\end{cases}$$

其中i_1，$i_2,\cdots,i_n$是$1,2,\cdots,$n的一个排列，这是由于可能交换过列的位置，相应地的位置也要调整.

讨论：（注意：$\mathrm{r}\le\min\{\mathrm{m},\mathrm{n}\}\ \le \mathrm{n}$ ，故只有下面三种情形！）

（1）情形 1：$\mathrm{d}_{r+1}\neq 0$

此时 $\mathrm{Cx}=\mathrm{d}$ 无解. 相应地，$\mathrm{Ax}=\mathrm{b}$也无解.

$\mathrm{d}_{r+1}\neq 0$说明：$\mathrm{r(D)}\ \neq\ \mathrm{r(c)}$，即：$\mathrm{r(B)}\ \neq\ \mathrm{r(A)}$.

（2）情形 2：$\mathrm{d}_{r+1}=0$。此时再分两种情况：

（2.1） $\mathrm{d_{r+1}}=0$，且 $\mathrm{r}=\mathrm{n}$：

此时$Cx=d$有唯一解 $x_{id}=d_s$

相应地，$Ax=b$也有唯一解，解与$Cx=d$相同.

另， $\mathrm{d_{r+1}}=0$且$\mathrm{r}=\mathrm{n}$，说明： $\mathrm{r(D)}\ =\ \mathrm{r(C)}=\mathrm{n}$

亦即：$\mathrm{r}(\mathrm{B}) = \mathrm{r}(\mathrm{A}) = \mathrm{n}$.

(2.2) $\mathrm{d}_{\mathrm{r+1}} = 0$，且$\mathrm{r} < \mathrm{n}$：

此时$Cx = d$ 的解（亦即$\mathrm{Ax} = \mathrm{b}$的解）可写成：

$$\begin{cases} x_{i1} = d_1 - C_{1r+1}t_{r+1} - \cdots - C_{1n}t_n \\ x_{i2} = d_2 - C_{2r+1}t_{r+1} - \cdots - C_{2n}t_n \\ \quad\vdots \\ x_{ir} = d_r - C_{rr+1}t_{r+1} - \cdots - C_{rn}t_n \\ x_{i_{r+1}} = t_{r+1} \\ \quad\vdots \\ x_{i_n} = t_n \end{cases}$$

由于$\mathrm{t}_{\mathrm{i}_{\mathrm{r+1}}}, \dots, \mathrm{t}_{\mathrm{i}_\mathrm{n}}$可以为任意的数值，故$\mathrm{Ax} = \mathrm{b}$有无穷多个解。上述解称$\mathrm{A}x = \mathrm{b}$的通解(或一般解). 与 D 中主一对应的未知量$x_{i_{r+1}}, \dots, x_{i_n}$为自由未知量.

再来看一个例题:

求解：$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 2 \\ 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 = -2 \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 = 2 \\ x_2 + 2x_3 + 2x_4 + x_5 = 3 \end{cases}$

（写出通解，并指出主未知量，自由未知量！）

综合上述讨论，即得到线性方程组的有解的判定定理：

定理 5.2.1 线性方程组$AX = b$有解的充要条件是：其系数矩阵于增广矩阵的秩相等，即：$\mathrm{r}(\mathrm{A}) = \mathrm{r}([\mathrm{A} \vdots \mathrm{B}])$！

定理 5.2.2 设线性方程组 $A_{m\times n}X_{n\times 1}=B_{m\times 1}$的系数矩阵与增广矩阵秩均为r，则：

（1）$\mathrm{r} = \mathrm{n}$时，方程组有唯一解.

（2）$\mathrm{r} < \mathrm{n}$时，方程组有无穷多解.

推论 5.2.3 齐次方程组$A_{m\times n}X_{n\times 1} = 0$,若$\mathrm{r}(\mathrm{A}) < \mathrm{n}$，则有非零解.

推论 5.2.4 齐次方程组$A_{m\times n}X_{n\times 1} = 0$,若$\mathrm{m} < \mathrm{n}$,则有非零解.

定理 5.2.3 n 个 n 元方程组成的齐次线性方程组:$A_{n\times n}X_{n\times 1}=0$ 有非零解的充要条件是:$|\mathrm{A}|=0$.

证明：若$|\mathrm{A}|\neq 0$,

则由克莱姆法则，$AX=0$只有零解没有非零解；

若$|\mathrm{A}|=0$,

则$\mathrm{r(A)}<\mathrm{n}$. 由推论 5.2.3，AX=0 有非零解.

综上，$|A|=0 \iff AX=0$有非零解, 证毕.

§5.3　线性方程组解的结构

本节研究：当线性方程组有不止一个解时，解与解之间的关系如何？

齐次线性方程组解的结构：

$m\times n$型齐次线性方程组：

$$\begin{cases}\text{矩阵形式：} A_{m\times n}x_{n\times 1}=0\\ \text{向量形式：} x_1A_1+x_2A_2+\cdots+x_nA_n=0\end{cases}$$

齐次：至少有一个解：$x=0$

解集合记作：$N(A)=\{x: Ax=0\}$.

定理 5.3.1　设x_1和x_2是 $Ax=0$ 的两解，则$\forall k_1$，$k_2\in F$，

$k_1x_1+k_2x_2$亦是$Ax=0$的一个解.

证：$A(k_1x_1+k_2x_2)=k_1Ax_1+k_2Ax_2=0$，故$k_1x_1+k_2x_2\in N(A)$

上述定理说明：齐次线性方程组的解的全体，构成数域 F 的一个子空间. 即：解集合构成一个线性空间，称之为解空间.

定义（基础解系）　齐次线性方程组解空间的一个基称为该方程组的一个基础解系.

注：由定义，一个基础解系满足如下三个条件：

(1). 其中每个向量均是$AX=0$之解.

(2). 基础解系中的向量线性无关.

(3). 解空间中任意一个向量均可以用基础解系中的向量表示.

设$\{x_1,x_2,\dots,x_s\}$是$Ax=0$的一个基础解系，则$Ax=0$的任意一个解均可写成：$k_1x_1+k_2x_2+\cdots+k_sx_s$，此式称$Ax=0$的通解.

解空间的维数

若$r(A)=n$. 由上节课，$Ax=0$只有零解，故解空间是零空间.

若$r(A)<n$. 则由初等变换的：（设$r(A)=r$）

$\mathrm{B}=\begin{bmatrix}\mathrm{A} & \vdots & 0\end{bmatrix} \longrightarrow \mathrm{D}=\begin{bmatrix}\mathrm{C} & \vdots & 0\end{bmatrix}$，其中：

$$D=\begin{bmatrix}1 & 0 & \cdots & 0 & c_{1r+1} & \cdots & c_{1n} & 0\\ 0 & 1 & \cdots & 0 & c_{2r+1} & \cdots & c_{2n} & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & c_{rr+1} & \cdots & c_{rn} & 0\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0\end{bmatrix}=[C \ \vdots \ 0]$$

同解方程组 $CX = 0$ 如下：

$$\left\{\begin{array}{llllllllll}
x_{i_1} & & & + & C_{1r+1}x_{i_{r+1}} & + & \cdots & + & C_{1n}x_{i_n} & =0\\
 & x_{i_2} & & + & C_{2r+1}x_{i_{r+1}} & + & \cdots & + & C_{2n}x_{i_n} & =0\\
 & & & & \vdots & & & & & \\
 & & x_{i_r} & + & C_{rr+1}x_{i_{r+1}} & + & \cdots & + & C_{rn}x_{i_n} & =0\\
 & & & & & & & & 0 & =0\\
 & \downarrow & & & & & \downarrow & & 0 & =0\\
 & & & & & & & & \vdots & \\
 & \text{主未知量} & & & & & \text{自由未知量} & & 0 & =0
\end{array}\right.$$

其中i_1，$i_2,\ldots$，i_n是 1,2， …, n 的一个排列，这是由于可能交换过列的位置，相应地 x_i 的位置也要调整.

其解为(亦即 Ax=0 之解)：

$$\left\{\begin{array}{llllll}
x_{i_1} & = & -C_{1r+1}t_{i_{r+1}} & - & \cdots & - & C_{1n}t_{i_n}\\
x_{i_2} & = & -C_{2r+1}t_{i_{r+1}} & - & \cdots & - & C_{2n}t_{i_n}\\
 & & \vdots & & & & \\
x_{i_r} & = & -C_{rr+1}t_{i_{r+1}} & - & \cdots & - & C_{rn}t_{i_n}\\
x_{i_{r+1}} & = & t_{i_{r+1}} & & & & \\
 & \vdots & & & & & \\
x_{i_n} & = & t_{i_n} & & & &
\end{array}\right.$$

让自由未知变量依次取$(1,0,\dots,0)$，$(0,1,\dots,0)$， …， $(0,0,\dots,1)$可得到$\mathrm{n}-\mathrm{r}$个解向量$x_{r+1}, x_{r+2}, \cdots$， x_n 从而：

解可写成列向量的形式，即：

定理 5.3.2 设$A_{m\times n}=0$中矩阵的秩为$r(A)$，则它的解空间的维数是：$\dim N(A)=n-r(A)$.

$$\begin{bmatrix} x_{i_1} \\ x_{i_2} \\ \vdots \\ x_{i_r} \\ x_{i_{r+1}} \\ x_{i_{r+2}} \\ \vdots \\ x_{i_n} \end{bmatrix} = t_{i_{r+1}} \underset{\alpha_{r+1}}{\begin{bmatrix} -c_{1r+1} \\ -c_{2r+1} \\ \vdots \\ -c_{rr+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}} + t_{i_{r+2}} \underset{\alpha_{r+2}}{\begin{bmatrix} -c_{1r+2} \\ -c_{2r+2} \\ \vdots \\ -c_{rr+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}} + \cdots + t_{i_n} \underset{\alpha_n}{\begin{bmatrix} -c_{1n} \\ -c_{2n} \\ \vdots \\ -c_{rn} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}} \quad (5.4)$$

显然：(1)这$n-r$个解向量α_{r+1}，$\alpha_{r+2},\cdots$， α_n线性无关；

(2)任意一个解都可用α_{r+1}， $\alpha_{r+2},\cdots$， α_n线性表示.

故$\{\alpha_{r+1}, \alpha_{r+2},\cdots, \alpha_n\}$构成了$N(A)$的一个基，

故$\dim(A)=n-r=n-r(A)$，证毕.

例：求$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ 2x_1 + 2x_2 + x_3 + 3x_4 = 0 \\ x_1 + x_2 + 2x_3 = 0 \end{cases}$ 的基础解系和通解.

解： $$A=\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 3 \\ 1 & 1 & 2 & 0 \end{bmatrix} \to \begin{array}{c} \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{matrix} x_1 & x_2 & x_3 & x_4 \end{matrix} \end{array}$$

主未知量　　自由未知量

解法一：$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}-x_2-2x_4\\x_2\\x_4\\x_4\end{bmatrix}=x_2\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}+x_4\begin{bmatrix}-2\\0\\1\\1\end{bmatrix}$

由 (x_2,x_4) 依次取（1, 0）和（0, 1）来得到

故基础解系是：$\left\{[-1,1,0,0]^{\mathrm{T}}，[-2,0,1,1]^{\mathrm{T}}\right\}$

解法二：由 A 的标准形可知：

列向量组$\{A_1,A_2,A_3,A_4\}$的极大无关组是$\{A_1,A_3\}$，故A_2和A_4均可用A_1,A_3线性表示，具体来说，即：

$\begin{cases}A_2=A_1\\A_4=2A_1-A_3\end{cases}$ 重新改写，即

$\begin{cases}-A_1+A_2=0\\-2A_1+A_3+A_4=0\end{cases}$ 与$x_1A_1+x_2A_2+\cdots+x_4A_4=0$的形式比较.

故$\begin{cases}x_1=[-1,1,0,0]^T\\x_2=[-2,0,1,1]^T\end{cases}$ 是 $AX=0$ 的两个解.

基础解系是：$\left\{[-1,1,0,0]^{\mathrm{T}}，[-2,0,1,1]^{\mathrm{T}}\right\}$.

总结：求齐次线性方程组$AX=0$基础解系的两种方法：

1、求出$AX=0$的通解，并写成向量形式即可得基础解系（换句话说、即依次只取某一自由未知量为 1，其他自由未知量为 0，这样来得到$n-r$个解向量.）

2、将 A 化为标准型，找出 A 的列向量的极大无关组. 将极大无关组之外的其它列向量用极大无关组线性表示出来，然后移项，写成线性相关的形式，由线性组合系数即可得到基础解系.

例 1 设$A_{m\times n}$，$B_{n\times k}$,有$AB=0$，求证：$r(A)+r(B)\le n$.

证：方法一：矩阵秩的不等式.

方法二：$B=[B_1,B_2,\dots,B_k]$，则$B_i\in N(A)$.

例 2 设$A_{m\times n}$是实矩阵. 求证：$r(A^TA)=r(A)$.

证：构造方程组：$Ax=0$与$(A^TA)x=0$，证明同解.

例 3 设$A \in C^{m\times n}$, 求证：$r(A^* A) = r(A)$.

其中：$A^* := (A^T)^c$ （转置复共轭）

例 4 已知$A_{m\times n}$的秩$r(A) = m < n. Ax = 0$的一个基础解系为：$[b_1, b_2, \dots, b_{n-m}]$， 求方程组$BY = 0$的一个基础解系，其中$B = [b_1, b_2, \dots, b_{n-m}]^T$.

证：$Ab_1 = 0 \Longrightarrow \cdots \Longrightarrow BA^T = 0$.

例 5 求证：方程组$A_{m\times n}Y_{n\times 1} = b_{m\times 1}$有解的充要条件是:

$A^T X_{m\times 1} = 0$的任意一个解都满足$b^T X$=0.

证：（必要性）：设$AY = b$有解，则∃Y_0，使$b = AY_0$.

对$A^T X_{m\times 1}$=0的任意一个解X_0，都有 $b^T X_0 = Y_0^T A^T X_0 = 0$.

（充分性）：∀X，有$A^T X = 0$, $b^T X = 0$，故

$A^T X = 0 \Leftrightarrow \begin{bmatrix} A^T \\ b^T \end{bmatrix} X = 0$. 故$r(A) = r(\begin{bmatrix} A & \vdots & b \end{bmatrix})$，故$AY = b$有解.

非齐次线性方程组解的结构

m× n型非齐次线性方程组：

$$\begin{cases} \text{矩阵形式：} & A_{m\times n}X_{n\times 1} = b_{m\times 1} \\ \text{向量形式：} & x_1A_1 + x_2A_2 + \cdots + x_nA_n = b \end{cases} \qquad (A_i，b \in F^m)$$

例 6 设满足x_1, x_2满足：$Ax_1 = b$，$Ax_2 = b$

则对 $x_1 + x_2$ 有：$A(x_1 + x_2) = b + b = 2b \neq b$.

非齐次线性方程组的解，不构成线性空间.

但是，任意两个解的差，构成一个线性空间.

定义(导出方程组) 称$AX = 0$为非齐次线性方程组$AX = b$的导出方程组，简称导出组.

定理 5.3.3 (1) $AX = b$的两个解之差是其导出组$AX = 0$的解；

(2) $AX = b$的一个解与其导出组$Ax = 0$的一个解之和，

仍是$AX=b$的一个解.

证：(1)设x_1，x_2是$AX=b$的两个解. 则：

$$A(x_1-x_2)=Ax_1-Ax_2=b-b=0.\text{故}x_1-x_2\in N(A)$$

(2)设Y_1是$AX=b$之一解，x_1是$AX=0$之一解.

则$A(Y_1+x_1)=AY_1+Ax_1=b+0=b$. 证毕.

定理 5.3.4 若r_0是$AX=b$的一个解(称作特解)，则$AX=b$的所有解都可写成$r_0+\alpha$的形式(称作通解)，其中α是导出组(即$AX=0$)的解.

证：按定理 5.3.3(2). $r_0+\alpha$是$AX=b$的一个解.

对$AX=b$的任意一个解r，$r-r_0$是导出组之解$\alpha=r-r_0$

则 $r=r_0+\alpha$. 证毕.

一般地，设$Ax=b$.（增广：$\mathrm{B}=\begin{bmatrix}\mathrm{A} & \vdots & \mathrm{b}\end{bmatrix}$）高斯消元法变成：

$Cx=d$.（增广：$\mathrm{D}=\begin{bmatrix}\mathrm{C} & \vdots & \mathrm{d}\end{bmatrix}$）D 的一般形式为：

$$D=\begin{bmatrix}1 & 0 & \cdots & 0 & c_{1r+1} & \cdots & c_{1n} & d_1\\ 0 & 1 & \cdots & 0 & c_{2r+1} & \cdots & c_{2n} & d_2\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & c_{rr+1} & \cdots & c_{rn} & d_r\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d_{r+1}\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0\end{bmatrix}=[C\ \vdots\ d\,]$$

注：为了书写方便，得到上面的 D 时有可能会交换列的位置. 交换列相当于交换方程组中未知量的位置，显然不会影响其解. 实际解题时

没有必要交换增广矩阵的列！

$$\begin{cases} x_{i1} = d_1 - C_{1r+1}t_{r+1} - \cdots - C_{1n}t_n \\ x_{i2} = d_2 - C_{2r+1}t_{r1} - \cdots - C_{2n}t_n \\ \qquad\vdots \\ x_{ir} = d_r - C_{rr+1}t_{r+1} - \cdots - C_{rn}t_n \\ x_{i_{r+1}} = t_{r+1} \\ \qquad\vdots \\ x_{i_n} = t_n \end{cases}$$

$$\begin{bmatrix} x_{i_1} \\ x_{i_2} \\ \vdots \\ x_{i_r} \\ x_{i_{r+1}} \\ x_{i_{r+2}} \\ \vdots \\ x_{i_n} \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + t_{i_{r+1}} \begin{bmatrix} -c_{1r+1} \\ -c_{2r+1} \\ \vdots \\ -c_{rr+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + t_{i_{r+2}} \begin{bmatrix} -c_{1r+2} \\ -c_{2r+2} \\ \vdots \\ -c_{rr+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + t_{i_n} \begin{bmatrix} -c_{1n} \\ -c_{2n} \\ \vdots \\ -c_{rn} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

由上述定理可知：若已求出非齐次线性方程组$AX = b$的一个特解r_0，以及其导出组$AX = 0$的一组基础解系$\alpha_1, \alpha_2, \ldots,$ $\alpha_{\mathrm{n-r}}$其中$r = r(A)$，则$AX = b$的通解为：$r_0 + k_1\alpha_1 + k_2\alpha_2 + \cdots + k_{n-r}\alpha_{n-r}$.

（与前面用 Gauss 消元法求得的通解组互印证！）

例 7　求下面方程组的通解：

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 7 \\ 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 = -2 \\ x_2 + 2x_3 + 2x_4 + 6x_5 = 23 \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 = 12 \end{cases}$$

（两种方法求解基础解系）

例 8　设R^4中列向量组$\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$，$\alpha_1, \alpha_2, \alpha_4$线性无关, 且

$\alpha_3 = 3\alpha_1 - \alpha_2 - 2\alpha_4$，$\beta = \alpha_1 + 2\alpha_2 + 3\alpha_3 + 4\alpha_4$，设

$\mathrm{A} = [\alpha_1, \alpha_2, \alpha_3, \alpha_4]$，求 $\mathrm{AX} = \beta$的通解.

用向量和方程组两种方法

例 9 设γ是非齐次方程组$AX=b$的一个特解，$\alpha_1,\alpha_2,\dots,\alpha_{n-r}$是其导出组$AX=0$的一个基础解系，其中$r=r(A)$，求证：

（1）γ，$\alpha_1,\alpha_2,\dots,\alpha_{n-r}$线性无关.

（2）γ，$\gamma+\alpha_1$，$\dots,\gamma+\alpha_{n-r}$也线性无关.

（3）γ，$\gamma+\alpha_1$，$\dots,\gamma+\alpha_{n-r}$也是$AX=b$的解集合的一个极大无关组.

证：(1) 设$k\gamma+k_1\alpha_1+\cdots+k_{n-r}\alpha_{n-r}=0$. 左乘 A 有: $k=0$.

例 10 设$AX=b$中，$r(A)=r$，若$\alpha_0,\alpha_1,\alpha_2,\dots,\alpha_{n-r}$是$AX=b$的线性无关的解，求其通解.

解：$\alpha_1-\alpha_0,\alpha_2-\alpha_0,\dots,\ \alpha_{n-r}-\alpha_0$

回过头来再看第四章的一个习题：

例：设α_1=$[1,0,2,3]^T$， α_2=$[1,-1,a+2,1]^T$，α_3=$[1,2,4,a+8]^T$，

α_4=$[1,1,3,5]^T$，β=$[1,1,b+3,5]^T$，

（1）a，b为何值时，β可由$\alpha_1,\alpha_2,\alpha_3,\alpha_4$唯一线性表示？

（2）a，b为何值时，β可由$\alpha_1,\alpha_2,\alpha_3,\alpha_4$线性表示，但不唯一？

（3）a，b为何值时，β不可由$\alpha_1,\alpha_2,\alpha_3,\alpha_4$线性表示？

$$\begin{bmatrix}1 & 1 & 1 & 1 & 1\\0 & -1 & 2 & 1 & 1\\2 & a+2 & 4 & 3 & b+3\\3 & 1 & a+8 & 5 & 5\end{bmatrix}\rightarrow$$

$$\begin{bmatrix}1 & 0 & 3 & 2 & 2\\0 & -1 & 2 & 1 & 1\\0 & 0 & a+1 & 0 & 0\\0 & 0 & 0 & a+1 & a+b+1\end{bmatrix}$$

（分别用向量的语言和线性方程组的语言来讨论！）

作业：

1.设非齐次线性方程组(I)和(II)分别为

(I) $\begin{cases} x_1+x_2-2x_4=-6 \\ 4x_1-x_2-x_3-x_4=1 \\ 3x_1-x_2-x_3=3 \end{cases}$；　(II) $\begin{cases} x_1+mx_2-x_3-x_4=-5 \\ nx_2-x_3-2x_4=-11 \\ x_3-2x_4=-t+11 \end{cases}$

(1)求方程组(I)的通解；(2)m，n，t 为何值时(I)和(II)是同解方程组.

2. 设非齐次线性方程组$AX=b$，系数矩阵 A 为5×3阶，$r(A)=2$, 且 η_1, η_2 是该方程组的两个解，有$\eta_1+\eta_2=[1,3,0]^T$，$2\eta_1+3\eta_2=[2,5,1]^T$,求该方程组的通解.

3. A 是$m\times n$阶的矩阵，它的 m 个行是某个 n 元齐次线性方程组的基础解系，又 B 是一个可逆的 m 阶矩阵，证明 BA 的行向量也构成该齐次线性方程组的一个基础解系.

第六章:线性变换

§6.1　线性映射和线性变换的定义和运算

主要知识点：线性映射、像、核、秩、零度、加法与数乘、乘法；线性泛函、对偶空间、对偶向量；线性变换 、可逆线性变换、线性变换群

线性变换是线性空间到自身的一类特殊映射。对有限维线性空间，线性变换与矩阵有极为密切的联系。

在第四章学过线性空间中的同构映射，这里我们先看一类更广泛的映射:线性映射。线性泛函和线性变换各是一类特殊的线性映射。

线性映射

定义：设 V 和 U 是数域 F 上的线性空间，若映射σ：$V \to U$满足下列两个条件：

（1）. $\forall \alpha, \beta \in V, \sigma(\alpha+\beta) = \sigma(\alpha) + \sigma(\beta)$

（2）. $\forall \alpha \in V, k \in F, \sigma(k\alpha) = k\sigma(\alpha)$

则称σ为 V 到 U 的线性映射

定义中条件（1）（2）可合写成$\sigma(k\alpha + l\beta) = k\sigma(\alpha) + l\,\sigma(\beta)$

可见：线性映射实际上是保持线性运算的一类特殊映射，具体来说，有：（1）$\sigma(0) = 0'$

（2）$\forall \alpha \in V, \sigma(-\alpha) = -\sigma(\alpha)$

（3）$\sigma(\sum k_i\alpha_i) = \sum k_i\,\sigma(\alpha_i)$因此若有：$\beta = \sum k_i\alpha_i$，则可推出$\sigma(\beta) = \sum k_i\,\sigma(\alpha_i)$

（4）若α_1，α_2，…，α_n线性相关，则$\sigma(\alpha_1)$，$\sigma(\alpha_i)$，…，$\sigma(\alpha_n)$也线性相关。注意：线性无关时不一定成立。例如$\forall \alpha \in V, \sigma(\alpha) = 0'$

同构映射是一一对应的映射。

例：设$A_{m\times n}\in F_{m\times n}$，则如下定义的映射σ：$F^n\to F^m$. $\forall\alpha\in F^n,\sigma(\alpha)=\ A\alpha$是一个线性映射。（随简单，但是有限维的同时！以后证明：有限维时，线性变换和矩阵之间可建立一一对应的关系！）

例：$\mathrm{V}=\varphi[a,b]$是定义在$[a,b]$上的实连续函数空间，定义σ：$\mathrm{V}\to\mathrm{R}$: $\forall f(x)\in\varphi[a,b],\sigma(f(x))=\int_a^b f(x)dx$.是$\varphi[a,b]$到 R 的一个线性映射。

例：设$P_n[x]$是次数不大于 n 的多项式空间，定义σ: $P_n[x]\to P_n[x]$，$\forall f(x)\in P_n[x]$，$\sigma(f(x))=f'(x)$是$P_n[x]$到$P_n[x]$的一个线性映射。

例：下列映射是线性映射？

（1）$\sigma([x,y,z]^T)=[y,z]^T$

（2）$\sigma([x_1,x_2,x_3])=[0,0]$

（3）$\sigma([x_1,x_2,x_3]^T=[1+x_1,x_3])$

（4）$\sigma([x_1,x_2,x_3])=[x_3,x_1+x_2]$

（5）$\sigma([x_1,x_2,x_3])=[1+x_1,x_2]$

（6）$\sigma([x,y,z])=[x+y,2z,x]$

定理 6.11 设σ：$\mathrm{V}\to\mathrm{U}$是 F 上线性空间 V 到 U 的线性映射，S 与H'分别是 V 和 U 的子空间，则：

（1）S 的像：$\sigma(S)=\{\sigma(\alpha),\alpha\in \mathrm{S}\}$是 U 的子空间。

（2）H'的像源：$H=\{\alpha\in V,\sigma(\alpha)\in H'\}$是 V 的子空间。

证明：（1） $k\sigma(\alpha)+l\,\sigma(\beta)=\sigma(k\alpha+l\beta)\ \in\sigma(S)$

（2）$\forall\alpha,\beta\in H,\sigma(\alpha),\sigma(\beta)\in H'.\sigma(k\alpha+l\beta)=\ k\sigma(\alpha)+l\,\sigma(\beta)\ \in H'$ 故$k\alpha+l\beta\in H$

定义（像，核）：V 的全体在σ下的像向量集合$\sigma(V)$是 U 的一个子空间，称映射σ的像，U 中的零向量$0'$在σ之下的像源集合是 V 的一个子空间，称映射σ的核

像 $I_m(\sigma)=\{\sigma(\alpha),\alpha\in V\}$　　核$\mathrm{Ker}(\alpha)=\{\alpha\in V,\sigma(\alpha)=0'\}$

定理 6.12 设σ：$V\rightarrow U$是 F 上线性空间 V 到 U 的线性映射,则：

（1）σ是单映射$\Leftrightarrow\mathrm{Ker}(\sigma)=\{0\}$

（2）σ是满映射$\Leftrightarrow\ I_m(\sigma)=U$

证：（1）σ是单映射，则$0'$的像源只有一个，只能是 0，故$\mathrm{Ker}(\sigma)=\{0\}$。反之，若$\mathrm{Ker}(\sigma)=\{0\}$，则设$\alpha,\beta\in V$，若$\sigma(\alpha)=\sigma(\beta)$，则$\sigma(\alpha-\beta)=0'$，故$\alpha-\beta=0$，即$\alpha=\beta$

（2）满映射的定义

定义（秩，零度）：设σ是 F 上有限线性空间 V 到有限线性空间 U 的线性映射$\sigma:V\rightarrow U$，则σ的像与核的维数分别称为σ的秩和零度，记作$r(\sigma)$和$N(\sigma)$.即：$r(\sigma)=\dim(\mathrm{Ker}(\sigma))$ ；$N(\sigma)=\ \dim(\ I_m(\sigma))$

定理 6.13　设σ是 F 上有限线性空间 V 到有限线性空间 U 的线性映射$\sigma:V\rightarrow U$，则其秩和零度满足：$\boldsymbol{r(\sigma)+N(\sigma)=dim(V)}$

证明：（1）$N(\sigma)=dim(V)$时，$I_m(\sigma)=\{0\}$，$r(\sigma)=0$.故成立。

（3）$N(\sigma)<dim(V)$时，设$N(\sigma)=k,dim(V)=n$。则：选取$\mathrm{Ker}(\sigma)$的一个基$\{\alpha_1，\alpha_2，\cdots，\alpha_k\}$,则中 V 必有不能用这组基表示的向量，任取一个记作β_{k+1}，如此下去，可构造

$\{\alpha_1，\alpha_2，\cdots，\alpha_k,\beta_{k+1},\cdots,\beta_n\}$成为 V 的一组基。再证明：$\{\sigma(\beta_{k+1}),\cdots,\sigma(\beta_n)\}$构成 $I_m(\sigma)$的基： $\forall\beta\in I_m(\sigma)$，有：$\exists\alpha\in V$，使$\beta=\sigma(\alpha)$。则 $\alpha=k_1\alpha_1+\cdots+k_k\alpha_k+k_{k+1}\beta_{k+1}+\cdots+k_n\beta_n$ 。故$\sigma(\alpha)=k_{k+1}\sigma(\beta_{k+1})+\cdots+k_n\sigma(\beta_n)$.

另一方面设$k_{k+1}\sigma(\beta_{k+1})+\cdots+k_n\sigma(\beta_n)=0$，则$\sigma(\sum k\beta)=0$

$k_{k+1}\beta_{k+1}+\cdots+k_n\beta_n=k_1\alpha_1+\cdots+k_k\alpha_k$，故所有$k_i=0$

例：再来看$F^n\to F^m$的如下定义的线性映射$\sigma: X\to Y=AX$，则$A\in F^{m\times n}$

则：$\mathrm{Ker}(\sigma)=\{X,X\in F^n 且 AX=0\}$ $I_m(\sigma)=\{AX,\forall X\in F^n\}$

即$\mathrm{Ker}(\sigma)$是$AX=0$的解空间，$I_m(\sigma)$是 A 的列向量的所有线性组合：$A_1X_1+\cdots+A_nX_n$

故：$dim(Ker(\sigma))=n-r(A)$ ；$N(\sigma)=n-r(A)$. $dim(I_m(\sigma))=r(A)$,即$r(\sigma)=r(A)$

故$r(\sigma)+N(\sigma)=r(A)+(n-r(A))=n=dim(V)$

可见：线性映射与矩阵和线性方程之间有深刻的练习（以后会讲！）故线性映射的秩，零度与矩阵的秩，齐次方程的解空间也有深刻的练习。这也是秩，零度名称的由来。τ

线性映射的线性运算（对应矩阵的线性运算）

定义（加法）：设σ和τ是 F 上 V 到 U 的两线性映射，则定义$\sigma+\tau$为：

$$\forall\alpha\in V,(\sigma+\tau)(\alpha)=\sigma(\alpha)+\tau(\alpha)$$

注：一个映射由它对象源集合里任何一个元素的操作来定义！

易证明：$\sigma+\tau$仍是 V 到 U 的一个线性映射，即

$$(\sigma+\tau)(k\alpha+l\beta)=k(\sigma+\tau)(\alpha)+l(\sigma+\tau)(\beta)$$

且，线性映射的加法满足如下运算律：

(1) $\sigma+\tau=\tau+\sigma$ （交换律）

(2) $\sigma+(\tau+p)=(\sigma+\tau)+p$ （结合律）

(3)定义零映射：$(\forall\alpha\in V,\sigma(\alpha)=0)$存在，且 $0+\sigma=\sigma$（存在零元）

(4)定义$-\sigma$: $(-\sigma)(\alpha)=-\sigma(\alpha)$,则$(-\sigma)+\sigma=0$ （存在负元）

定义（数乘）：设σ是 F 上 V 到 U 的线性映射，$k\in$F,则定义$k\sigma$为:

$$\forall\alpha\in V,(k\sigma)(\alpha)=k\sigma(\alpha)$$

易证明：$k\sigma$仍是 V 到 U 的线性映射，即：

$$k\sigma(m\alpha+l\beta)=m(k\sigma)(\alpha)+l(k\sigma)(\beta)$$

且数乘满足如下运算律：

（1）$k\sigma(\alpha+\tau)=\ k\sigma+k\tau$

（2）$(k+l)\sigma=k\sigma+l\sigma$

(3) $(kl)\sigma=k(l\sigma)$

(4) $1\sigma=\sigma$

定理 6.14　F 上 V 到 U 的线性映射全体，线性映射的加法与数乘，构成 F 上的一个线性空间。（注：以后可知：n 维 V 到 m 维 U 的线性映射σ 与 $m\times n$ 矩阵一一对应，故这个线性空间将与$F^{m\times n}$同构。

定义（乘法）：设σ是 F 上 V 到 U 的线性映射，τ是 F 上 U 到 W 的线性映射, 则定义σ和τ的乘积$\tau\sigma$如下：

$$\forall\alpha\in V:\quad(\tau\sigma)(\alpha)=\tau\sigma(\alpha)$$

易证明：（1）$\tau\sigma$仍是 V 到 U 的线性映射

（2）一般地，并不能定义$\sigma\tau$，除非 W=V

定义（线性泛函）：数域 F 上线性空间 V 到 F 的一个线性映射，称为 V 上的一个 F 值线性泛函。

注：（1）F 是自身上的一个线性空间！

（2）非数集到数集的映射一般均称为泛函。

由定理 6.14 可引入

定义（对偶空间）：V 上 F 值线性泛函的全体，构成一个线性空间称为 V 的对偶空间，记作V^*。设$\{u_1$，u_2，$\cdots$，$u_n\}$是 V 中的一个基。由于线性映射是通过对任意向量的映射来定义的，又具有线性性质，故只要给出了对 V 中基的映射，就能确定下一个线性映射。因此。可如下定义：n 个线性泛函W^i：$W^i(u_j) = \delta_j^i \ (i, j = 1,2 \cdots, n)$

注：（1）首先：每个W^i是恰当定义的，因为$\forall \alpha \in \mathrm{V}$，有

$$W^i(\alpha) = W^i\left(\sum_j k_j v_j\right) = \sum_j k_j W^i(v_j) = k_i$$

（2）$W^1, W^2, \cdots, W^n$线性无关！这是因为：假设$\sum_i k_i W^i = 0$,作用到u_j上有：$k_j = 0$

（3）$\forall \alpha \in V^*$，$\sigma\alpha$可用$W^1, W^2, \cdots, W^n$来展开！这是因为：σ和$\sum_i \sigma(v_i) W^i$作用在任意$\alpha \in V$上，结果都相等，故$\sigma = \sum_i \sigma(v_i) W^i$

可见:$\{W^1, W^2, \cdots, W^n\}$构成$V^*$的一个基，称 V 中$\{u\}$的对偶基。且有：

$$\dim\mathrm{V} = \dim V^*$$

例：设 V 中基$\{v\}$到基$\{v'\}$的过渡矩阵为 C，求证V^*中$\{v\}$的对偶基$\{\omega\}$到$\{v'\}$的对偶基ω'的过渡矩阵是：$(C^{-1})^T$.

例：三维矢量空间 V 不共面的三个矢量$\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}$可构成一个基，任意一个矢量 $\boldsymbol{v}$，可定义$\mathrm{V} \to \mathrm{R}$的一个线性映射：$\boldsymbol{v}\boldsymbol{v}' = \boldsymbol{v} \cdot \boldsymbol{v}'$

故$V^* = \mathrm{V}$,对偶基可选作：$\left\{\boldsymbol{b_1} = \frac{\boldsymbol{1}}{\boldsymbol{\Omega}}(\boldsymbol{a_2} \times \boldsymbol{a_3}), \boldsymbol{b_2} = \frac{\boldsymbol{1}}{\boldsymbol{\Omega}}(\boldsymbol{a_3} \times \boldsymbol{a_1}), \boldsymbol{b_3} = \frac{\boldsymbol{1}}{\boldsymbol{\Omega}}(\boldsymbol{a_1} \times \boldsymbol{a_2})\right\}$

例：量子力学中，右矢$|\varphi\rangle$的集合构成一个向量空间，其对偶空间是坐矢$\langle\varphi|$的集合。

例：广义相对论中，协变向量构成的向量空间是逆变向量构成的向量空间的对偶空间。

定义（线性变换）：线性空间 V 到其自身的线性映射称为 V 上的线性变换。（注：一般地，一个集合到其自身的一个映射，都称其为变换）

例：任给定一个$\mathrm{A} \in F^{m \times n}$,则如下定义的映射σ：$F^n \to F^n, \forall x \in F^n, \sigma(x) = Ax$是$F^n$上的一个线性变换。

例：设$\boldsymbol{n} = l\boldsymbol{i} + m\boldsymbol{j} + n\boldsymbol{k}$是$R^3$中的一个单位向量（），则：

定义的σ是R^3的一个线性变换。

由于线性变换是一类特殊的线性映射，因此定理 6.1.1-6.1.3 对线性变换仍然成立。即

对 V 上的线性变换σ：

（1）像$I_m(\sigma)$和核 Ker(σ)分别构成线性空间；

（2）秩 r(σ)和零度 N(σ)满足：r(σ) +N(σ)=dimV

同样，线性变换的线性运算(加法和数乘)与线性映射的运算完全一样，因此，对应于 6.1.4，有：

定理 6.1.5:　线性空间 V 上的线性变换全体，对线性变换的加法和数乘构成数域 F 上的线性空间，通常记作 L(V)

定义（线性变换的乘法）：设$\sigma,\tau\in$ L(V),则σ与τ的复合映射$\sigma^{\circ}\tau$：$\forall\alpha\in$ V,$(\sigma^{\circ}\tau)(\alpha)=\sigma(\tau(\alpha))$ 是 V 的一个变换，可证明，这是一个线性变换，称为σ与τ的乘积，记作$\sigma\tau$。（注：（1）以后将证明：该乘法对应于方阵的乘法，（2）一般的，$\sigma\tau\neq\tau\sigma$

定理 6.1.6：若 V 上的线性变换σ存在逆映射σ^{-1} ，则此逆映射σ^{-1}也是一个线性变换。此时，称 σ 是可逆线性变换。

证：令ε表示恒等变换，则：$\forall\alpha,\beta\in$ V ，由于σ^{-1}存在，故：

$$k\sigma^{-1}(\alpha)+l\sigma^{-1}(\beta)\in V$$

$$k\sigma^{-1}(\alpha)+l\sigma^{-1}(\beta)=\varepsilon(k\sigma^{-1}(\alpha)+l\sigma^{-1}(\beta))$$

$$=\ \sigma^{-1}\,\sigma(\cdots)=\ \sigma^{-1}(k\alpha+l\beta)$$

注：以后将证明，可逆线性变换与可逆方阵一一对应。

易证明如下定理:

定理 6.1.7:　V 上所有可逆线性变换的全体，对线性变换的乘法构成群称 V 上的线性变换群。

对有限维度线性空间上的线性变换有：

定理 6.1.8：设σ是 n 为线性空间 V 上的线性变换，则如下论述等价：

（1）σ是单映射

（2）σ是满映射

（3）σ是双映射

（4）σ是可逆映射

证：循环法：$(1)\Rightarrow(2)$: $\mathrm{Ker}(\sigma)=\{0\}$, 故$N(0)=0$, 故$I_m(\sigma)=V$

$(2)\Rightarrow(3)$：$I_m(\sigma)=V$ ，故 $\mathrm{r}(\sigma)=\mathrm{n}$,故$\mathrm{Ker}(\sigma)=\{0\}$，故是单射。

$(3)\Rightarrow(4)$：双映射有逆映射，由定理 6.1.6 即得。

$(4)\Rightarrow(1)$：σ^{-1} 存在，故$\forall\alpha\in\mathrm{Ker}(\sigma)$有：$\sigma(\alpha)=0$

故$\alpha=\sigma^{-1}(\sigma(\alpha))=\sigma^{-1}(0)=0$.故$\mathrm{Ker}(\sigma)=\{0\}$。故是单射

例：以F^n上的线性变换 σ：$F^n\to F^n,\forall x\in F^n,\sigma(x)=Ax$为例，其中 $\mathrm{A}\in F^{n\times n}$,讨论定理 6.1.8.

解：σ是单映射。说明$\mathrm{Ker}(\sigma)=0$，即$Ax=0$只有零解，故$\mathrm{r}(A)=\mathrm{n}$

σ是满映射.说明$I_m(\sigma)=F^n$，即$Ax=A_1x_1+A_2x_2+\cdots+A_nx_n$构成$F^n$，故$\{A_1,A_2\cdots,A_\mathrm{n}\}$线性无关。$\mathrm{r}(A)=\mathrm{n}$。

σ可逆，说明 A 有逆矩阵，故$\mathrm{r}(A)=\mathrm{n}$

例：F^n中：定义：$\sigma([x_1,x_2,\cdots,x_n])=[0,x_1,\cdots,x_{n-1}]$

求证　σ是F^n的一个线性变换，且$\sigma^n=0$

求 $\mathrm{Ker}(\sigma)$和$I_m(\sigma)$的维数，即求$\mathrm{N}(\sigma)$和$\mathrm{r}(\sigma)$!

§6.2 线性映射和线性变换的矩阵表示

线性映射的矩阵表示　线性映射空间与矩阵空间的同构　线性变换的矩阵表示　矩阵表示的坐标变换、矩阵的相似

线性映射的矩阵表示

设σ是F上线性空间V到U的线性映射，且$dimV=n,dimU=m$, $\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是V的一组基，$\{\beta_1,\beta_2,\cdots,\beta_m\}$是U的一组基，则$\sigma(\alpha_i)\in$ U,可用$\{\beta\}$来展开，即：

$$\sigma(\alpha_i)=\sum_{j=1}^{m}\beta_j a_{ji}=\beta_1 a_{1i}+\beta_2 a_{2i}+\cdots+\beta_m a_{mi}$$

写成矩阵的形式，即：$\sigma(\alpha_i)=[\beta_1,\beta_2,\cdots,\beta_m][a_{1i},a_{2i},\cdots,a_{mi}]^T$

将$\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)$ 关于基 $\{\beta_1,\beta_2,\cdots,\beta_m\}$的坐标为列排成如下$m\times n$的矩阵

$$A=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix}$$

即：$[\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)]=[\beta_1,\beta_2,\cdots,\beta_m]A$

称A为σ关于基$\{\alpha\}$和$\{\beta\}$的矩阵.（通过对基的映射来定义矩阵）

引理 6.2.1 设$\sigma: V\to U$是一线性映射。$\alpha_1，\alpha_2，\cdots，\alpha_n\in V$, $X=[x_1,x_2,\ldots,x_n]^T\in F^n, A\in F^{n\times k}$. 则

（1）$\sigma([\alpha_1，\alpha_2，\cdots，\alpha_n]X)=[\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)]\,X$

（2）$\sigma([\alpha_1，\alpha_2，\cdots，\alpha_n]A)=[\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)]\,A$

证：（1）$\sigma([\alpha_1，\alpha_2，\cdots，\alpha_n]X)=\sigma(\sum_i\alpha_i x_i)=\sum_i\sigma(\alpha_i x_i)=[\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)]\,X$

（2）对 A 按列分组，由（1）即得。

定理 6.2.2　设$dimV=n,dimU=m$，$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是 V 的一组基，$\{\beta_1,\beta_2,\cdots,\beta_m\}$是 U 的一组基，线性映射 $\sigma:V\to U$在上述基下的矩阵是$A_{m\times n}$,则若$\alpha\in V$ 在上述基下的坐标为$X=[x_1,x_2,\cdots,x_n]^T$

则$\sigma(\alpha)$在上述基下的坐标为: $Y=AX$

证：$\sigma=[\alpha_1，\alpha_2，\cdots，\alpha_n]X$，故有引理 6.2.1 有：$\sigma(\alpha)=\sigma([\alpha_1，\alpha_2，\cdots，\alpha_n]X=[\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)]X=[\beta_1,\beta_2,\cdots,\beta_m]AX$

对 $\alpha\to\sigma(\alpha)$，$X\to Y=AX$。称线性映射的矩阵表示。

关于线性映射与矩阵运算之间的对应关系，有如下定理：

定理 6.2.3　设$\sigma:V\to U$, $\tau:V\to U,\rho:U\to W$是三个线性映射，且在各定的V,U,W的某个基下的矩阵分别是A,B,C 。则 $\sigma+\tau,k\sigma,\rho\sigma$在同样的基下的矩阵分别是$A+B,kA,CA$

证：设V的基是$\{\alpha_1，\alpha_2，\cdots，\alpha_n\},n=dimV$

U的基是$\{\beta_1,\beta_2,\cdots,\beta_m\},m=dimU$

W的基是$\{\gamma_1，\gamma_2，\cdots，\gamma_k\},k=dimW$

则$A\in F^{m\times n},B\in F^{m\times n},C\in F^{k\times n}$

$$[\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)]=[\beta_1,\beta_2,\cdots,\beta_m]A$$

$$[\tau(\alpha_1),\tau(\alpha_2),\cdots,\tau(\alpha_n)]=[\beta_1,\beta_2,\cdots,\beta_m]B$$

$$[\rho(\beta_1),\rho(\beta_2),\cdots,\rho(\beta_m)]=[\gamma_1，\gamma_2，\cdots，\gamma_k]C$$

乘法：

$[(\rho\sigma)(\alpha_1),\cdots,(\rho\sigma)(\alpha_n)]=[\rho(\sigma(\alpha_1)),\cdots,\rho(\sigma(\alpha_n))]$

$=\rho([\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)])=\rho([\beta_1,\beta_2,\cdots,\beta_m]A)$

$=[\gamma_1，\gamma_2，\cdots，\gamma_k]CA$

加法和数乘类似证之！

线性映射空间与矩阵空间的同构

上面指出了:给定一个线性映射, 存在一个矩阵与之对应.下面将证明反之亦成立。即: 给出一个矩阵，也存在一个线性映射与之对应不仅如此，这个对应关系还可建立两个线性空间的同构。先来看如下引理，它说明的是:基的映射唯一地确定了一个线性空间的线性映射。

引理 6.2.4： $\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是 F 上 n 维线性空间 V 的一个基，则对于线性空间 U 中的任意 n 维向量$\beta_1,\beta_2,\cdots,\beta_n$，存在唯一一个线性映射$\sigma:V\to U$，使得$\sigma(\alpha_i)=\beta_i$, $i=1,2,3,\cdots,n$

证：$\forall\alpha\epsilon V$，若α在上述基下的坐标为$[x_1,x_2,\cdots,x_n]^T$，则定义$\sigma:V\to U$，$\sigma(\alpha_i)=\sum_{i=1}^{n}x_i\beta_i$

可证其为线性映射，且$\sigma(\alpha_i)=\beta_i$。设另有τ,$\tau(\alpha_i)=\beta_i$. 则$\forall\alpha\epsilon V,\tau(\alpha)=\sum x_i\tau(\alpha_i)=\sum x_i\beta_i=\sigma(\alpha)$.故$\tau=\sigma$

引理 6.2.5：F 上 n 维线性空间 V 到 m 维线性空间 U 的全体映射构成的线性空间（记作$L(V\to U)$ 与 F 上的$\mathrm{m}\times\mathrm{n}$矩阵$F^{m\times n}$用构）

证：取 V 的基$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$，U 的基$\{\beta_1,\beta_2,\cdots,\beta_m\}$，则对于$L(V\to U)$中任意一个线性映射$\sigma$，让$\sigma$在上述基下的矩阵 A 与$\sigma$对应，于是，建立起来一个$L(V\to U)$中到$F^{m\times n}$的映射$f:f(\sigma)=A$ 。先证明

该映射是双映射。即要证明对任意$A \in F^{m\times n}$ ，有且仅有一个σ满足$f(\sigma)=A$ 任取$A \in F^{m\times n}$，令$\gamma_i=\sum_{j=1}^{m}\beta_j A_{ji},(i=1,2,3,\cdots,n)$

由引理 6.2.4，存在唯一的$\sigma \in L(V\to U)$，使得$\sigma(\alpha_i)=\gamma_i,(i=1,2,3,\cdots,n) \Leftrightarrow$ σ在上述基下的矩阵为 A，即$f(\sigma)=A$. 故$f: L(V\to U)\to F^{m\times n}$是一个双映射，再有定理 6.2.3 可知$f$是同构映射。即：

$$L(V\to U)\simeq F^{m\times n}$$

线性变换的矩阵表示

设σ是 n 维线性空间 V 上的线性变换，即$\sigma \in L(V)$，σ: V → V $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$是 V 的一组基，则

$$\sigma(\alpha_i)=\sum_{j=1}^{n}\alpha_j a_{ji}=\alpha_1 a_{1i}+\alpha_2 a_{2i}+\cdots+\alpha_n a_{ni}$$

排成列矩阵的形式即为：

$$[\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)]=[\alpha_1, \alpha_2, \cdots, \alpha_n]A$$

其中：

$$A=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

称为线性变换σ在基$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$下的矩阵

$\forall \alpha \epsilon V$, 若α在上述基下的坐标为$[x_1,x_2,\cdots,x_n]^T$, 则$\sigma(\alpha)$在上述基下的坐标为: $Y=AX$

对于线性变换的加法、数乘、乘法，定理 6.2.3 仍然适用

设σ, τ ∈ $L(V)$，且在 V 的$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$基下，σ和τ的矩阵分别是$A,B$，则

（1）$\sigma+\tau$在该基下的矩阵为$A+B$

（2）$k\sigma$在该基下的矩阵为kA

（3）$\sigma\tau$在该基下的矩阵为$AB \Rightarrow$ 若 σ 可逆，则σ^{-1}的矩阵为A^{-1}

定理 6.2.5 亦适用：设 V 为 n 维线性空间，则：

$$L(V) \backsimeq F^{n\times n}$$

总结：

（1）n 维向量空间：$V \backsimeq F^{n}$

（2）n 维 V 到 m 维 U 的线性映射：$L(V\to U) \backsimeq F^{m\times n}$

（3）n 维 V 上的线性变换： $L(V) \backsimeq F^{n\times n}$

即：（1）向量α与$X\in F^{n}$一一对应

（2）n 维 V 到 m 维 U 的线性映射α与$A\in F^{m\times n}$一一对应

（3）V 上的线性变换α与$A\in F^{n\times n}$一一对应

例：求R^3中的线性变换：σ：$\sigma(\begin{bmatrix} x \\ y \\ -z \end{bmatrix}) = \begin{bmatrix} x-y \\ 2x \\ x+z \end{bmatrix}$在$\{e_1,e_2,e_3\}$下的矩阵

解：$\sigma(e_1)$=...

例：由$\hat{n}=l\hat{\imath}+m\hat{\jmath}+n\hat{k}$定义的$R^3$中的线性变换：$\sigma:\alpha\to(\hat{\alpha}\cdot\hat{n})\hat{n}$在$\{\hat{\imath},\hat{\jmath},\hat{k}\}$下的矩阵是？

例：设$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是 V 一个基，求$\sigma:\alpha\to\sigma(\alpha)=k\alpha$关于这个基的矩阵。

解：$\sigma(\alpha_j)=k\alpha_j=[\alpha_1，\alpha_2，\cdots，\alpha_n][0,\cdots,k,\cdots,0]^T$故矩阵为$kI$

贴别地，恒等变换关于任何一个基的矩阵都是单位矩阵 I

零变换 0 关于任何一个一个基的矩阵都是零矩阵 0

矩阵表示的坐标变换、矩阵的相似

向量的坐标与基的选择有关，类似地，线性变换的矩阵也与基的选择有关。不同的基下，线性变换的矩阵不同，下面推到坐标变换

定理 6.2.6： 设 C 是 F 上 V 的基$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$到基$\{\beta_1,\beta_2,\cdots,\beta_m\}$的过渡矩阵，V 上的线性变换σ关于基$\{\alpha_i\}$与基$\{\beta_i\}$的矩阵分别是 A 和 B，则有：$B=C^{-1}AC$

证：依条件有：$[\beta_1,\beta_2,\cdots,\beta_n]=[\alpha_1，\alpha_2，\cdots，\alpha_n]C,$

$$[\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)]=[\alpha_1，\alpha_2，\cdots，\alpha_n]A$$

$$[\sigma(\beta_1),\sigma(\beta_2),\cdots,\sigma(\beta_n)]=[\beta_1,\beta_2,\cdots,\beta_n]B$$

$$\sigma([\beta_1,\beta_2,\cdots,\beta_n])=[\beta_1,\beta_2,\cdots,\beta_n]B=[\alpha_1，\alpha_2，\cdots，\alpha_n]CB$$

又

$$\sigma([\beta_1,\beta_2,\cdots,\beta_n])=\sigma([\alpha_1，\alpha_2，\cdots，\alpha_n]C)$$

$$=[\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)]C=[\alpha_1，\alpha_2，\cdots，\alpha_n]AC$$

故$CB=AC\Rightarrow B=C^{-1}AC$

定义（相似矩阵）：设$\mathrm{A,B}\in F^{n\times n}$，若存在可逆的$\mathrm{C}\in F^{n\times n}$，使得：$B=C^{-1}AC$, 则称矩阵 A 与 B 相似，记作$\mathrm{A\sim B}$，C 称相似变换矩阵。显然方阵的相似是一种等价关系，即：

(1) $\mathrm{A\sim A}$

(2) $\mathrm{A\sim B\Rightarrow B\sim A}$

(3) $\mathrm{A\sim B, B\sim C\Rightarrow A\sim C}$

（注意区分矩阵的相似，等价，合同）

利用上述定义，定理 6.2.6 可表述成：V 上的线性变换关于两个基的矩阵式相似的。反之，两个相似的矩阵可看作同一个线性变换在不同基之下的矩阵。

矩阵相似的一些性质：

(1) $\mathrm{A}\sim\mathrm{B}\Rightarrow\mathrm{tr(A)}=\mathrm{tr(B)}$　(2) $\mathrm{A}\sim\mathrm{B}\Rightarrow|A|=|B|$

例：σ：$[x,y,z]^T\to[x-y,2x,x+z]^T$，求σ关于基$\{e_1,e_2,e_3\}$和基$\{g_1=[1,0,0]^T,g_2=[1,1,0]^T,g_3=[1,1,1]^T\}$的矩阵。（分别按定义和按$B=C^{-1}AC$，求$\{g\}$下的矩阵）

例：设$\{\alpha_1，\alpha_2，\alpha_3\}$是 V 的基, $\sigma,\tau\in\mathrm{L(V)}$ 分别定义如下：

$$\begin{cases}\sigma(\alpha_1)=4\alpha_1+3\alpha_2-3\alpha_3\\ \sigma(\alpha_2)\ =3\alpha_1+4\alpha_2-3\alpha_3\\ \sigma(\alpha_3)\ =-\alpha_1-2\alpha_2+\alpha_3\end{cases}\qquad\qquad\begin{cases}\tau(\alpha_1)=-\alpha_1+\alpha_2\\ \tau(\alpha_2)\ =-\alpha_1+3\alpha_3\\ \tau(\alpha_3)\ =\alpha_1-\alpha_2\end{cases}$$

（1）分别求$\sigma,\tau,3\sigma-2\tau,\sigma\tau,\tau\sigma,\sigma^{-1}$的矩阵

（2）可逆否？

例：$\sigma\in\mathrm{L}(F^3),\sigma$：$[x,y,z]^T\to[x+y,x-y+z,2z]^T$

（1）求σ在$\{e_1,e_2,e_3\}$下的矩阵

（2）求$\{e_1,e_2,e_3\}$到$\{g_1=[1,0,0]^T,g_2=[1,1,0]^T,g_3=[1,1,1]^T\}$的过渡矩阵和逆矩阵

（3）求σ在$\{g_1,g_2,g_3\}$下的矩阵

（4）求向量$\alpha=[1,2,3]^T$在基$\{g_1,g_2,g_3\}$下的坐标

（5）求σ分别在上述两组基下的坐标

例：设 $\mathrm{A,B}\in F^{n\times n}$，且 A 可逆，求证：AB = BA

例：设$A=\begin{bmatrix}2&0&0\\0&0&1\\0&1&x\end{bmatrix}\sim B=\begin{bmatrix}2&0&0\\0&y&0\\0&0&-1\end{bmatrix}$，求 x 和 y

§6.3、本征值、本征向量

线性变换的本征值、本征向量　矩阵的本征值、本征向量　本征多项式、本征方程

问题的提出：$\sigma\in L(v)$,由于 v 中不同基下σ的矩阵不同，一个问题是：能否找到 V 中的一个基，使得σ在该基下的矩阵是简单的对角形式？或等价的，对$\mathrm{A}\in F^{m\times n}$ 能否找到可逆的矩阵 P，使得$P^{-1}AP$成为简单的对角形式？这些都与本征值和本征向量有关。

定义（本征值、本征向量）：设$\sigma\in L(v),\lambda\in\mathrm{F}$，若 V 中存在非零向量α，使得：$\sigma(\alpha)=\lambda\alpha$, 则称λ是$\sigma$ 的一个本征值，而α成σ的属于本征值λ的本征向量

定理 6.3.1　σ的属于本征值λ的本征向量与零向量组成的集合 $E_\lambda=\{\alpha\in\mathrm{V}:\sigma(\alpha)=\lambda\alpha\}$ 是 V 的子空间，σ的属于本征值λ的本征子空间。

证：$\forall\alpha$，$\beta\in E_\lambda$，$\sigma(\alpha)=\lambda\alpha$，$\sigma(\beta)=\lambda\beta$。故$\sigma(k\alpha+l\beta)=k\sigma(\alpha)+l\sigma(\beta)=\lambda(k\alpha+l\beta)$，故$k\alpha+l\beta\in E_\lambda$

由于线性变换与数域上的方阵存在一一对应关系，故，平行地有：

定义（矩阵的本征值，本征向量）：设$\lambda \in \mathrm{F}$，若存在$x \in F^n$,$\mathrm{x} \neq 0$，使得$Ax = \lambda x$，则称λ是A的一个本征值，称x是A的属于本征值λ的一个本征向量。（注：同样有：若x是，则 kx也是！）

定理 6.3.1' n 阶矩阵 A 的属于本征值λ的全部本征向量与F^n中的零向量组成的集合是F^n的一个子空间，称A的属于本征值λ的本征子空间

线性变换的本征值、本征向量与矩阵的本征值、本征向量一一对应:

$$\tau(\alpha) = \lambda\alpha \Leftrightarrow \tau([\alpha_1, \alpha_2, \cdots, \alpha_n][x_1, x_2, \cdots, x_n]^T) = [\alpha_1, \alpha_2, \cdots, \alpha_n]AX$$
$$= \lambda[\alpha_1, \alpha_2, \cdots, \alpha_n]X \Leftrightarrow AX = \lambda X$$

本征多项式、本征方程

由于选定 V 的一个基之后，线性变换的本征值、本征向量与矩阵的本征值、本征向量的分别一一对应，故求出了矩阵的本征值、本征向量，即可得线性变换的本征值、本征向量

以下只讨论矩阵的本征值、本征向量!

设 $A \in F^{n \times n}$，$x \in F^n$，则$AX = \lambda X$可写成：$(\lambda \mathrm{I} - \mathrm{A})\mathrm{X} = 0$

而按$\mathrm{X} \neq 0$定义，故必须有：

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0$$

由行列式理论，上式展开之后可得到 F 上的一个 λ 的 n 次多项式，即：$f(\lambda) = |\lambda I - A| = \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1 \lambda + b_0$

注：由行列式可看出，λ^n前的系数$b_n = 1$！

定义（本征多项式，本征方程）：称$f(\lambda)=|\lambda I-A|$为矩阵 A 的本征多项式，而方程$f(\lambda)=0$成为矩阵 A 的本征方程。

显然，$f(\lambda)=$在 F 中的根就是 A 的全部本征值。而由代数学基本定理可知：在复数域 C 中， $f(\lambda)=0$有 n 个根（重根按重数计算）。从而 A 由 n 个本征值（可能有重值）。设 A 在 C 中的 n 个本征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$，则有：$f(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$

定理 6.3.2: 复数域上矩阵 A 的特征值满足如下关系：

(1) $\lambda_1+\lambda_2+\cdots+\lambda_n=\mathrm{tr}(\mathrm{A})=-b_{n-1}$

(2) $\lambda_1\lambda_2\cdots\lambda_n=|A|=(-1)^n b_0$

证：从行列式的展开来看λ^{n-1}的系数，必须在$\mathrm{n}-1$个主对角元中取λ，在剩下的一个对角元中取$-a_{ij}$，故有

又，从韦达定理看 的系数，由：

故

（2）在$f(\lambda)=|\lambda I-A|$中取$\lambda=0$，即：$\mathrm{f}(0)=b_0=|-A|=(-1)^n\mathrm{A}$

又，从韦达定理看常数项的系数，有: $b_0=(-1)^n\lambda_1\lambda_2\cdots\lambda_n$

故

$$\lambda_1\lambda_2\cdots\lambda_n=\prod_{i=1}^{n}\lambda_i=|A|=(-1)^n b_0$$

由于定理 6.3.2，A 的本征多项式可写为：$f(\lambda)=|\lambda I-A|=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^n|A|$

推论 6.3.3 矩阵 A 可逆的充要条件是：A 的全部本征值都不为零。

证：A 可逆$\Leftrightarrow|A|\neq 0\Leftrightarrow\prod_{i=1}^{n}\lambda_i\neq 0\Leftrightarrow\forall i,\lambda_i\neq 0$

定理 6.3.4：设 A 与 B 相似，则 A，B 的本征多项式相等！

证： $\mathrm{A} \sim \mathrm{B} \Leftrightarrow \mathrm{B}=C^{-1} A C$，故 $|\lambda I-B|=\left|\lambda I-C^{-1} A C\right|=\left|C^{-1}(\lambda I-A) C\right|=|\lambda I-A|$

推论 6.3.5：相似的矩阵，本征值相同！（若有重根，则重数也相同）(由于相似的矩阵可看成同一个线性变换在不同基下的表现形式，故此理论理所当然)。不仅如此，相似的矩阵其本征向量也是一一对应的：设 α 满足 $\mathrm{A} \alpha=\lambda_i \alpha$。则：$\mathrm{B}\left(C^{-1} \alpha\right)=C^{-1} A C C^{-1} \alpha=C^{-1} A \alpha=C^{-1} \lambda_i \alpha=\lambda_i\left(C^{-1} \alpha\right)$，即若 A 由以本征向量，则 B 也有属于相同本征值的本征向量$C^{-1} \alpha$

注意：推论的逆命题不一定成立！即本征值相同的矩阵未必相似。

反例：$\mathrm{A}=\begin{bmatrix}0 & 1 \\ 0 & 0\end{bmatrix} \quad B=\begin{bmatrix}0 & 0 \\ 0 & 0\end{bmatrix}$

总结：相似矩阵 $\Rightarrow$有相同的本征多项式，相同的本征值，相同的迹，相同的行列式，相同的秩。并有一一对应的本征向量。

例 1：设 A 与 B 相似，求证 A 的多项式 *f(A)*与 B 的多项式 *f(B)*相似，即 $A \sim B \Rightarrow f(A) \sim f(B)$

证：(1) $A \sim B \Rightarrow \mathrm{B}=C^{-1} A C \Rightarrow \mathrm{k}\, \mathrm{B}=C^{-1} k A C \Rightarrow k A \sim k B$

(2) $A \sim B \Rightarrow B+k I=C^{-1}(A+k I) C \Rightarrow B+k I \sim A+k I$

(3) $A \sim B \Rightarrow \mathrm{B}^k=\left(C^{-1} A C\right)^k=C^{-1} A^k C \Rightarrow \mathrm{B}^k \sim A^k$

(4) $A \sim B \Rightarrow f(B)=f\left(C^{-1} A C\right)=C^{-1} f(A) C \Rightarrow f(B) \sim f(A)$

故：$A \sim B \Rightarrow|\lambda I-A|=|\lambda I-B|$

例 2：设$A_{m \times n}, B_{n \times m}, m \geq n$，则 AB 和 BA 的本征多项式满足：

$$|\lambda I-A B|=\lambda^{m-n}|\lambda I-B A|$$

证：构造$\mathrm{P}=\begin{bmatrix}I & 0 \\ -A & I\end{bmatrix}$，$\mathrm{Q}=\begin{bmatrix}\lambda I & 0 \\ \lambda A & \lambda I\end{bmatrix}$

则$\mathrm{PQ}=\begin{bmatrix}\lambda I_n & B \\ 0 & \lambda I-AB\end{bmatrix}$, $\mathrm{QP}=\begin{bmatrix}\lambda I-AB & B \\ 0 & \lambda I_m\end{bmatrix}$

由$|\mathrm{PQ}|=|\mathrm{QP}|$既得证。

讨论：（1）若 A,B 均为方阵，结论可写为：

（2）若 A,B 均为方阵，且 A 可逆，则可知$AB\sim BA$ ，故必有相同的本征多项式

上述例的结论的一个应用是：求由行向量$\alpha=[a_1,\cdots,a_n]$和列向量$\beta=\begin{bmatrix}b_1\\ \vdots \\ b_n\end{bmatrix}$构成的方阵 A=βα 的本征值,由于$|\lambda I-A|=|\lambda I-\beta\alpha|=\lambda^{n-1}|\lambda I-\alpha\beta|=\lambda^{n-1}(\lambda-\alpha\beta)$，可知 A 的本征值是αβ和 0（n-1 重根）.

例 3：设 A 有一个本征值λ_o，且属于λ_o的一个本征向量是α_0，则：

（1）kA有一个本征值$k\lambda_o$，且α_0也是kA的特征向量。

（2）A^k一个本征值${\lambda_o}^k$，且α_0也是A^k的特征向量

（3）$\mathrm{g(A)}=b_kA^k+b_{k-1}A^{k-1}+\cdots+b_1A\ +b_0I$有一个本征值$g(\lambda_o)$

（4）若 A 可逆，则A^{-1}有一个本征值${\lambda_o}^{-1}$

（5）若 A 满足矩阵方程$\mathrm{g(A)}=b_kA^k+b_{k-1}A^{k-1}+\cdots+b_1A\ +b_0I=0$，则$\lambda_o$满足同样的多项式方程：$g(\lambda_o)=0$

证：（1）$\mathrm{A}\alpha_0=\lambda_o\alpha_0$，$(k\mathrm{A})\alpha_0=k\lambda_o\alpha_0$

（2）$A^k\alpha_0=A^{k-1}(\lambda_o\alpha_0)=A^{k-2}\left({\lambda_o}^2\alpha_0\right)=\cdots={\lambda_o}^k\alpha_0$

（3）$\mathrm{g(A)}\alpha_0=\cdots$

（4）$A^{-1}\alpha_0=\frac{1}{\lambda_o}A^{-1}A\alpha_0={\lambda_o}^{-1}\alpha_0$

（5）由（3）的结论$\mathrm{g(A)}\alpha_0=g(\lambda_o)$ ，故$g(\lambda_o)\alpha_0=0$，而$\alpha_0\neq 0$故$g(\lambda_o)=0$

例 4：设$A \in C^{3\times3}$，A 的三个本征值是 2,3,5，求$|2A^2+I|$

解： $2A^2+I$有特征值：9,19,51. 故$|2A^2+I| = 9\times19\times51$

定理 6.3.6（Hamilton---Caylay 定理）,设$A \in F^{n\times n}$，$f(\lambda)=|\lambda I-A|$，则：

$$f(A) = A^n-(a_{11}+a_{22}+\cdots+a_{nn})A^{n-1}+\cdots+(-1)^n|A|I=0$$

证：设$B(\lambda)$是$\lambda I-A$的伴随矩阵，由于$B(\lambda)$的元素是$|\lambda I-A|$的代数余子式，故$B(\lambda)=\lambda^{n-1}B_0+\lambda^{n-2}B_1+\cdots+\lambda B_{n-2}+B_{n-1}$，$B_i\in F^{n\times n}$，设$f(\lambda)=\lambda^n+C_1\lambda^{n-1}+\cdots+C_{n-1}\lambda+C_n$，$C_i\in F$

由$B(\lambda)(\lambda I-A)=|\lambda I-A|I=f(\lambda)I$，比较的$\lambda$系数有：$B_0=I,B_1-B_0A=C_1I,\cdots,B_{n-1}-B_{n-2}A=C_{n-1}I,-B_{n-1}A=C_nI$

以 $A^n,A^{n-1},\cdots,A,I$，分别乘上面各式有：$B_0A^n=A^n,B_1A^{n-1}-B_0A^n=C_1A^{n-1},\cdots,-B_{n-1}A=C_nI$,各式相加，即有：$f(A)=0$

本征值、本征向量的求法

由于：

$$AX = \lambda X(X\neq0) \Leftrightarrow (\lambda I-A)X=0(X\neq0) \Leftrightarrow |\lambda I-A|=0$$

故：（1）求解 A 的本征值 $\Leftrightarrow$求解 A 的本征方程

（2）求 A 关于λ_i的本征向量$\Leftrightarrow$求$(\lambda_i I-A)X=0$的非零解

即：$(\lambda_i I-A)X=0$的解空间就是 A 的属于本征值λ_i的本征子空间E_{λ_i}，而$(\lambda_i I-A)X=0$的一个基础解析就是E_{λ_i}的一个基！

求解步骤：（1）由$f(\lambda)=|\lambda I-A|=0$求出 A 的本征值$\lambda_1$，$\lambda_2$，$\cdots$，$\lambda_n$(注意可能有重根)

（2）对每个本征值，求$(\lambda_i I-A)X=0$的非零解（肯定有!），得到 A 的属于本征值λ_i的全部本征向量。

例 3：求$A=\begin{bmatrix}-2 & -1 & 9\\ -9 & 4 & 0\\ -3 & 1 & 1\end{bmatrix}$的本征值和本征向量

答案：$\lambda_1=1$，$\lambda_2=1+3i,\lambda_3=1-3i$

$E_{\lambda_1}=\{k[3,9,2]^T\}\ E_{\lambda_2}=\{k[1-i,3,1]^T\}\ E_{\lambda_3}=\{k[1+i,3,1]^T\}$

例 4：求$A=\begin{bmatrix}-3 & 1 & -1\\ -7 & 5 & -1\\ -6 & 6 & -2\end{bmatrix}$的本征值和本征向量

答案：$\lambda_1=\lambda_2=-2,\lambda_3=4$

$E_{\lambda_1}=\{k[1,1,0]^T\}\qquad E_{\lambda_3}=\{k[0,1,1]^T\}$

例 5：求$A=\begin{bmatrix}1 & -3 & 3\\ 3 & -5 & 3\\ 6 & -6 & 4\end{bmatrix}$的本征值和本征向量

答案：$\lambda_1=\lambda_2=-2,\lambda_3=4$

$E_{\lambda_1}=\{k_1[1,1,0]^T+k_2[0,1,1]^T\}\qquad E_{\lambda_3}=\{k[1,1,2]^T\}$

以上三个例子:

第一例只有单根, 每个单根对应一个线性无关的本征向量

第二例有二重根, 该二重根只对应一个线性无关的本征向量

第三例也有二重根, 但该二重根对应了两个线性无关的本征向量

可见: 一个 r 重本征值对应着多少个线性无关的本征向量？而一个本征向量又能否对应不同的本征值？都是待深究的问题。

例 6 证明：

(1) 若$A^2=I$，则 A 的本征值只能是±1

(2) 若$A^2=A$，则 A 的本征值只能是0 或 1

(3) 若$A^m=0\ (m\in Z^+)$，则 A 的本征值只能是0

证：（1）$\mathrm{A}\alpha = \lambda\alpha \Rightarrow \mathrm{A}^2\alpha = \lambda^2\alpha = \mathrm{I}\alpha \quad \lambda^2 = 1$

（2）$\mathrm{A}\alpha = \lambda\alpha \Rightarrow \mathrm{A}^2\alpha = \lambda^2\alpha = \mathrm{A}\alpha = \lambda\alpha \quad \lambda^2 = \lambda$

（3）$\mathrm{A}\alpha = \lambda\alpha \Rightarrow \mathrm{A}^m\alpha = \lambda^m\alpha = 0 \quad \lambda^m = 0$

例 7:　3 阶方阵 A 的本征值为 1,2,3，求$(2A)^{-1}, A^*$的本征值。

例 8：3 阶方阵 A 的本征值为 1,-1,2，C~A,而$B = C^2 - 5C + 2I$,求$|B|$, $|C^{-1} - 5I|$

例 9：　$\mathrm{A} = \begin{bmatrix} a & -1 & a \\ 5 & b & 3 \\ 1-a & 0 & -a \end{bmatrix}$,$|A| = 1, \alpha = [-1,-1,1]^T$是 A 的伴随矩阵$A^*$的本征值为$\lambda_0$的本征向量。求$a, b, \lambda$

解：$A^*\alpha = \lambda_0\alpha \Rightarrow AA^*\alpha = \lambda_0 A\alpha \Rightarrow \lambda_0 A\alpha = -\alpha$

例 10：设$A_n = (a_{ij})_{n\times n}$的本征值是$\lambda_i$，则$\sum_{i=1}^n \lambda_i^2$等于（　）

（A）$\sum_{i=1}^n a_{ii}^2$　　　　(B)$(\sum_{i=1}^n a_{ii})^2$

(C) $\sum_{i,j=1}^n a_{ij}^2$　　　　(D) $\sum_{i,j=1}^n a_{ij}a_{ji}$

§6.4 矩阵的对角化

可对角化的充要条件、对角化的求法　不同本征值的本征向量线性无关　本征值的重数与本征子空间的维数　可对角化更深刻的条件

这节课将研究矩阵与一个对角矩阵相似的充要条件，从而回答上节课最开始提出的问题：

能否找到 V 的一个基，使σ在该基下的矩阵是简单的对角形式？或等价地，对$\mathrm{A} \in F^{n\times n}$，能否找到可逆的 P，使得$P^{-1}AP$是简单的对角形式？

定义：设是 F 上线性空间 V 的一个线性变换，若存在 V 的一个基使σ在该基下的矩阵是简单的对角矩阵，则称σ可以对角化。平行地有如下定义：

定义：设 A 是 F 上一个 n 阶矩阵，若存在一个可逆的 n 阶矩阵 P，使$P^{-1}AP$是对角矩阵，则称矩阵 A 可以对角化。

可对角化的条件、对角化的求法

定理 6.4.1：F 上一个 n 阶矩阵 A 可对角化的充要条件是：A 在F^n中具有 n 个线性无关的本征向量。

证：A 可对角化⇔存在一个可逆的 n 阶矩阵 P，使$P^{-1}AP = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_n]$ ⇔ 存在一个可逆的 n 阶矩阵 P, 使$AP = P\mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_n]$(将 P 按列分块，$\mathrm{P} = [P_1, P_2, \cdots, P_n]$)⇔ 存在可逆的 P,使得$\mathrm{A}[P_1, P_2, \cdots, P_n] = [\lambda_1 P_1, \lambda_2 P_2, \cdots, \lambda_n P_n]$ ⇔ 存在 n 个线性无关的向量$P_1, P_2, \cdots, P_n$，使得$\mathrm{A}P_1 = \lambda_1 P_1, \mathrm{A}P_2 = \lambda_2 P_2, \cdots, \mathrm{A}P_n = \lambda_n P_n$.

为了更深刻地揭示矩阵可对角化的条件，我们继续研究：属于不同本征值的本征向量之间的关系

定理 6.4.1：设α_1，α_2，$\cdots$，α_m是 n 阶矩阵 A 的属于不同本征值的本征向量，则α_1，α_2，$\cdots$，α_m线性无关

证明：归纳法。当 m=1 时，按定义$\alpha_1 \neq 0$，故α_1线性无关.

假设定理对$\mathrm{m}-1$成立，则考虑 m 个不同本征值的情况。设 m 个不同本征值为$\lambda_1, \lambda_2, \cdots, \lambda_m$，设$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = 0$

用 $(\mathrm{A}-\lambda_m)I$ 左乘得，$k_1(\lambda_1-\lambda_m)\alpha_1 + k_2(\lambda_2-\lambda_m)\alpha_2 + \cdots + k_{m-1}(\lambda_{m-1}-\lambda_m)\alpha_{m-1} = 0$,而按归纳法假设，$\alpha_1$，$\alpha_2$，$\cdots$，$\alpha_{m-1}$线

性无关　，又λ_i互不相等，故$k_1=k_2=\cdots=k_{m-1}=0\Rightarrow k_m=0$，故$\alpha_1$，$\alpha_2$，$\cdots$，$\alpha_m$也线性无关。

由**定理 6.4.1 和定理 6.4.2**，立即得到如下推论

推论 6.4.3：n 阶矩阵 A 有 n 个互不相同的本征值，则 A 可对角化（注：此为充分条件，非必要！）

定理 6.4.4： 设$\lambda_1,\lambda_2,\cdots,\lambda_k$是 A 的 k 个互不相同的本征值，而$\alpha_{i1}$，$\alpha_{i2}$，$\cdots$，$\alpha_{is_i}$是 A 的$s_i$个线性无关的本征向量，$i=1,2,\cdots,k$,则$\alpha_{11}$，$\alpha_{12}$，$\cdots$，$\alpha_{1s_i}$，$\alpha_{21}$，$\alpha_{22}$，$\cdots$，$\alpha_{2s_i}$，$\alpha_{k1}$，$\alpha_{k2}$，$\cdots$，$\alpha_{ks_i}$线性无关。

证：设$C_{11}\alpha_{11}+\cdots+C_{1s_1}\alpha_{1s_1}+\cdots+C_{k1}\alpha_{k1}+\cdots+C_{ks_k}\alpha_{ks_k}=0$，令$\alpha_i=C_{i1}\alpha_{i1}+C_{i2}\alpha_{i2}\cdots+C_{is_i}\alpha_{is_i}$，$i=1,2,\cdots,k$.

则$\alpha_1+\alpha_2+\cdots\alpha_k=0$, α_i是 n 阶矩阵 A 的属于本征值λ_i的本征向量.故所有的$\alpha_i=0$（若存在非零的α_i，与定理 6.2.4 矛盾），又α_{i1}，α_{i2}，$\cdots$，α_{is_i}线性无关，从而$C_{i1}=C_{i2}=\cdots=C_{is_i}=0$.

即：将属于 k 个互不相同的本征值的 k 组各自线性无关的本征向量合并在一起，仍然线性无关。

本征值的重数与本征子空间的维数

定理 6.4.5 ： 设λ是矩阵A_n的 S 重本征值，E_λ是 A 的属于λ的本征子空间，则$\dim E_\lambda\le S$

证：设$\dim E_\lambda=l$, $\{\alpha_1$，α_2，$\cdots$，$\alpha_l\}$是$\dim E_\lambda$的一个基，则可通过逐步扩充（见定理 6.1.3 的证明），得到F^n的一个基如下：

$$\{\alpha_1，\alpha_2，\cdots，\alpha_l,\beta_{l+1},\beta_{l+2},\cdots,\beta_n$$

通过考虑 $A\alpha_1$，$A\alpha_2$，$\cdots$，$A\alpha_l, A\beta_{l+1}, A\beta_{l+2}, \cdots, A\beta_n$ ，可知：

$$\mathrm{A}\left[\alpha_1，\alpha_2，\cdots，\alpha_l, \beta_{l+1}, \beta_{l+2}, \cdots, \beta_n\right]$$

$$=\left[\alpha_1，\alpha_2，\cdots，\alpha_l, \beta_{l+1}, \beta_{l+2}, \cdots, \beta_n\right]\begin{bmatrix}\lambda I_l & C \\ 0 & B\end{bmatrix}$$

即$AP = P\begin{bmatrix}\lambda I_l & C \\ 0 & B\end{bmatrix}$, $P^{-1}AP = \begin{bmatrix}\lambda I_l & C \\ 0 & B\end{bmatrix}$，其中$\begin{bmatrix}\lambda I_l & C \\ 0 & B\end{bmatrix}$记作$\breve{A}$

故$|\lambda I - A| = \left|xI - \breve{A}\right| = (x-\lambda)^l q(x)$。可见$\lambda$的重数不小于 l.

即 $S \geq l$ ，即$\dim E_\lambda \leq S$.

此定理说明：A_n的线性无关的本征向量的个数≤ n.

可对角化更深刻的充要条件

定理 6.4.6：设复数域 C 上的 n 阶矩阵 A 上有 m 个互不相同的本征值，重数分别是$S_1, S_2, \cdots, S_m$（$\sum_{i=1}^{m} S_i = n$），则矩阵 A 可对角化的充要条件是对每一个本征值λ_i, A 的属于λ_i的本征子空间$E_{\lambda i}$的维数正好等于λ_i的重数S_i，即$\dim E_{\lambda i} = S_i (i = 1,2,\cdots,m)$

证：这是定理 6.4.1 和 6.4.5 的直接结果。⟸ 设 $\dim E_{\lambda i} = S_i$，则有 n 个线性无关的本征向量，则矩阵 A 可对角化 ⇒设矩阵 A 可对角化，则 A 有 n 个线性无关的本征向量.而 A 得所有线性无关的本征向量的个数是$\sum_{i=1}^{m} \dim E_{\lambda i}$ ，故$\sum_{i=1}^{m} \dim E_{\lambda i} \geq n$, 又由于定理 6.4.5 $\dim E_{\lambda i} \leq S_i$，而 $\sum_{i=1}^{m} S_i = n$，故只能是$\dim E_{\lambda i} = S_i$

矩阵对角化的方法

综合上述定理，可按如下步骤求与矩阵 A 相似的对角矩阵以及对角化的过渡矩阵 P：

（1）现有$|\lambda I - A| = 0$，求出数域 F 中A_n的全部本征值λ：及其重数S_i

（2）对每一个本征值λ_i，求出$(\lambda I - A)X = 0$的基础解系

（3）若每一个本征值λ_i，都有$\dim E_{\lambda i} = S_i$，则矩阵 A 可对角化，否则，不能对角化。

（4）可对角化时，以这些基础解系中的解向量作为列向量，排成矩阵 P，则 P 必可逆，且满足$AP = P\mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_n]$

$$即P^{-1}AP = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_n]$$

例 1 判断上节的例 3，例 4，例 5 的矩阵是否可对角化？若可，求出对应的对角矩阵和相应的过渡矩阵。

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

矩阵对角化的一些用途：

（1）求A^n: 设$\mathrm{A} = C^{-1}\Lambda\, C$，则$A^n = C^{-1}\Lambda^n\, C$，对角矩阵$A^n$易求！

（2）求$A^n\beta$：可先按（1）求出A^n，或，将β用 A 的本征向量展开。

例 2: 设A_3有本征值$\lambda_1 = 1, \lambda_2 = 2,, \lambda_3 = 3$, 对应的本征向量是：

$$x_1 = [1 \quad 1 \quad 1]^T, x_2 = [1 \quad 2 \quad 4]^T, x_3 = [1 \quad 3 \quad 9]^T$$

（1）求A^{-1}（2）求A^n

（3）设$\beta = [1 \quad 1 \quad 3]^T$，求$A^n\beta$

例 3：求解：
$$\begin{cases} \dfrac{dx_1}{dt} = 4x_1 + x_3 \\ \dfrac{dx_2}{dt} = 2x_1 + 3x_2 + 3x_3 \\ \dfrac{dx_3}{dt} = x_1 + 4x_3 \end{cases}$$

提示：令$X = [x_1 \quad x_2 \quad x_3]^T$ ，则可写成 $\dfrac{d\mathrm{X}}{dt} = AX$

若 A 可对角化，即$\mathrm{A} = C^{-1}\Lambda\, C$ ，则

$$\frac{dX}{dt} = C^{-1}\Lambda\, CX \Longrightarrow C\frac{dX}{dt} = \Lambda\, CX$$

令$CX = Y$，即$\frac{dY}{dt} = \Lambda Y$，对角形式易求解。

例 4：（1）设 A 的本征值均相等，且可对角化，求证 A=kI

（2）求证$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$不可对角化

例 5：设$A = \begin{bmatrix} 3 & 3 & -2 \\ -k & -1 & k \\ 4 & 2 & -3 \end{bmatrix}$，求 k 值，使 A 可对角化，并求出与 A 相似的对角矩阵，以及相应的过渡矩阵。

例 6：设$\alpha = [a_1, a_2, \cdots, a_n]^T$，$\beta = [b_1, b_2, \cdots, b_n]^T$ $(a_i \neq 0, b_i \neq 0)$，且

$$\alpha^T\beta = 2, A = \alpha\beta^T$$

（1）求 A 的本征值和本征向量

（2）求证 A 可对角化，并求出与 A 相似的对角矩阵

解：（1）两种方法：$|\lambda I - AB|$与$|\lambda I - BA|$之关系，或$A^2 = 2A$。对$\lambda = 2$。直接可验证$A\alpha = 2\alpha$

例 7：设$r(A_n) = r$ 且$A^2 = A$

（1）求 A 的本征值

（2）证明 A 可对角化，并写出对角矩阵

提示：由$(I - A)A = 0 \Rightarrow r(I - A) = n - r(A)$

例 8：求相似变换的过渡矩阵 P，使得$P^{-1}\begin{bmatrix} 0 & -\sqrt{2}/2 & 1/2 \\ \sqrt{2}/2 & 0 & -1/2 \\ -1/2 & 1/2 & 0 \end{bmatrix}P = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

解：法一：将 A,B 分别对角化

法二：解关于P_{ij}的方程。（用 AP=PB 求解）

例 9 ：求相似变换的过渡矩阵 P，使以下三个变换同时成立：

$$P^{-1}\begin{bmatrix}0 & -i & 0\\ i & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\mathrm{P}=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{bmatrix}\quad P^{-1}\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & -i\\ 0 & i & 0\end{bmatrix}\mathrm{P}=\frac{1}{\sqrt{2}}\begin{bmatrix}0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}$$

$$P^{-1}\begin{bmatrix}0 & 0 & i\\ 0 & 0 & -0\\ -i & 0 & 0\end{bmatrix}\mathrm{P}=\frac{1}{\sqrt{2}}\begin{bmatrix}0 & -1 & 0\\ 1 & 0 & -1\\ 0 & 1 & 0\end{bmatrix}$$

解：有第一个方程，求出 P 带有三个参数k_1,k_2,k_3 ，再将 P 带入后两个方程确定 。（用 *AP=PB* 求解）

例 10 ：设$\mathrm{A}\sim\begin{bmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{bmatrix},\mathrm{B}=(A-\lambda_1 I)(A-\lambda_2 I)(A-\lambda_3 I)$，求 B

法 一 ： $\mathrm{A}=P^{-1}\mathrm{diag}[\lambda_1,\lambda_2,\lambda_3]P\Rightarrow\ \mathrm{B}=P^{-1}(\mathrm{diag}[\lambda_1,\lambda_2,\lambda_3]-\lambda_1 I)(\mathrm{diag}[\lambda_1,\lambda_2,\lambda_3]-\lambda_2 I)(\mathrm{diag}[\lambda_1,\lambda_2,\lambda_3]-\lambda_3 I)P=P^{-1}OP=O$

法二：令$f(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)$，则又 Halmiton-caylay 定理，*f(A)=0,*故 *B= f(A)=0*

法三：令$f(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)$，由于 *A* 可对角化，而 *f(A)*与 *A* 有相同的本征向量，故 *f(A)* 亦可对角化，将本征向量排列成矩阵，有：$A=P\begin{bmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3\end{bmatrix}P^{-1}$

故 $f(A)=P\begin{bmatrix}f(\lambda_1) & & \\ & f(\lambda_2) & \\ & & f(\lambda_3)\end{bmatrix}P^{-1}$ ， 而 $f(\lambda_i)=0$ ， 故 $f(A)=0$，即 *B=0*

法一和法三：对 A 相似于对角形的特殊情况，证明 Halmiton-caylay 定理！

Jordan 标准型简介

不是每个方阵都可对角化. 当不可对角化时，相似变换可将方阵尽可能简单地化为何种形式?各种标准型理论。这里简单介绍 **Jordan 标准型**

定义（Jordan 块）：形如$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}_{r\times r}$的方阵称为一个 r 阶的 Jordan 块，记作 J(λ)

定义（Jordan 矩阵）：主对角线子块为 Jordan 块$J_i(\lambda_i)$的准对角矩阵称为 **Jordan 矩阵**

$$J=\begin{bmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_m(\lambda_m) \end{bmatrix}$$

注：（1）Jordan 矩阵中的 Jordan 子块不必互异，各个λ_i不必互异

（2）Jordan 矩阵是上三角矩阵，故其本征值就是主对角线上的元素

一些 Jordan 块的例子 $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

一些 Jordan 矩阵的例子

$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} J_1(1) & & \\ & J_2(5) & \\ & & J_3(2) \end{bmatrix}$,其中$J_1(1)=\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

$J_2(5)=[5]$ $J_3(2)=\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$

注：对角矩阵也是一类特殊的 Jordan 矩阵

这里我们不加证明，给出如下定理:

定理 6.4.7：复数域上 C 内，每个 n 阶矩阵 A 都相似于一个 Jordan 矩阵，且若不计较主对角线上各 Jordan 子块的顺序的话，该 Jordan 矩阵是唯一的。

即：$\forall \mathrm{A} \in C^{m\times n}$, $\exists P$可逆，

$$使得P^{-1}AP = J = \begin{bmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_m(\lambda_m) \end{bmatrix}$$

若矩阵 A 与 Jordan 矩阵 J 相似，我们称 J 是 A 的 **Jordan 标准型**。从而上述定理可以表述成：任意一个 C 上的 n 阶矩阵必存在 Jordan 标准型，且唯一。（显然可对角化的矩阵，其相似对角形就是其 Jordan 标准型）

下面不加证明地给出 Jordan 标准型的求法之一：

由于 Jordan 矩阵的本征值就是主对对角线上的元素，因此要求 A 的 Jordan 标准型，第一步先求 A 的本征方程的根。

设$|(\lambda_1 I - \mathrm{A})| = \mathrm{f(A)} = (\lambda - \lambda_1)^{S_1}(\lambda - \lambda_2)^{S_2} \cdots (\lambda - \lambda_m)^{S_m}$

显然$S_1 + S_2 + \cdots + S_m = n$

在 A 的 Jordan 标准型中，将属于同一λ_i的 Jordan 块放在一起，有

$$J_A = \begin{bmatrix} A_1(\lambda_1) & & & \\ & A_2(\lambda_2) & & \\ & & \ddots & \\ & & & A_m(\lambda_m) \end{bmatrix}$$

其中$A_i(\lambda_i)$包含了一系列本征值为λ_i的 Jordan 块。

$$P^{-1}AP = J_A \Longrightarrow A[P_1, P_2, \cdots, P_n] = [P_1, P_2, \cdots, P_n]\, J_A$$

假设$A_l(\lambda_l)$中含有k_l个 Jordan 块（$k_l = dimE_\lambda$）

即：$A_l(\lambda_l)=\begin{bmatrix} J_{l1}(\lambda_l) & & & \\ & J_{l2}(\lambda_l) & & \\ & & \ddots & \\ & & & J_{lk_l}(\lambda_l) \end{bmatrix}$

其中每个$J_{li}(\lambda_l)$均是一个 Jordan 块

具体看来每一个 Jordan 块$J_{li}(\lambda_l)$ ，与之相对应的列向量P_j:

设某个λ_l的 Jordan 块是$t\times t$的矩阵，对应的P_j用α, β标记

$$\mathrm{A}[\alpha_1,\beta_2,\cdots,\beta_t]=[\alpha_1,\beta_2,\cdots,\beta_t]\begin{bmatrix} \lambda_l & 1 & & & \\ & \lambda_l & 1 & & \\ & & \lambda_l & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_l \end{bmatrix}_{t\times t}$$

写成方程组即：$\begin{cases} A\alpha_1=\lambda_l\alpha_1 \\ A\beta_2=\alpha_1+\lambda_l\beta_2 \\ \vdots \\ A\beta_t=\beta_{t-1}+\lambda_l\beta_t \end{cases} \Rightarrow \begin{cases} (A-\lambda_l I)\alpha_1=0 \\ (A-\lambda_l I)\beta_2=\alpha_1 \\ \vdots \\ (A-\lambda_l I)\beta_t=\beta_{t-1} \end{cases}$

（$[\alpha_1,\beta_2,\cdots,\beta_t]$称为一个 Jordan 链）

可证明$(A-\lambda_l I)X=\beta_t$是无解的方程！可见每一个 Jordan 块对应了λ_l的一个本征向量α，可证明λ_l的本征子空间E_λ的一个基中，每一个本征向量也对应一个 Jordan 块。还可证明，通过线性无关的α，结合上述方程来求解 P，得到的所有的α，β放在一起构成的向量组线性无关。故 Jordan 标准型的求法如下：

（1） 求解$|(\lambda I-\mathrm{A})|=0$，得到 A 的所有的互异本征值$\lambda_i$,及其重数$S_i$；

（2） 对每一个λ_i，由$|(\lambda_i I-\mathrm{A})|X=0$，求出$E_{\lambda_i}$的一个基：

$$\{\alpha_{i1},\alpha_{i2},\cdots,\alpha_{ir_i}\}(r_i=dimE_{\lambda_i})$$

（3） 若$r_i = S_i$（对于单根，定是如此），则与本征值λ_i对应的Jordan 块全是1×1的矩阵，共有S_i个；若$r_i < S_i$，在E_{λ_i}的所有基中，先选取α_{i1}，由：$(A - \lambda_i I)\beta_2 = \alpha_{i1}$，$(A - \lambda_i I)\beta_3 = \beta_2, \cdots,$直到$(A - \lambda_i I)X = \beta_t$无解，求出与$\alpha_{i1}$相对应的 Jordan 链 。继续求$\alpha_{i2}$相对应的 Jordan 链直到$\alpha_{ir_i}$。（实际上求到$\alpha_i$与$\beta$的个数加起来等于$S_i$即可停止，因为可证明总数正好是$S_i$）。对所有的 i（即所有相异的本征值）重复上述步骤！

（4） 将对所有不同的 i，所有不同的α_{ij}，按上述步骤求出的所有α，β当作列向量排成矩阵，即得矩阵 P, 使：

$$P^{-1}AP = J_A$$

例 11 ：求$A = \begin{bmatrix} -3 & 2 & -2 \\ -7 & 6 & -3 \\ 1 & -1 & 2 \end{bmatrix}$的$J_A$和 P 答案：$J_A = \begin{bmatrix} 1 & & \\ & 2 & 1 \\ & & 2 \end{bmatrix}$

例 12 ：求 $A = \begin{bmatrix} 2 & 1 & 0 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$的 J_A 和 P 答案：$J_A = \begin{bmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}$

例 13 ：求$A = \begin{bmatrix} 2 & -1 & -1 \\ 2 & -1 & -2 \\ -1 & 1 & 2 \end{bmatrix}$的$J_A$和 P 答案：$J_A = \begin{bmatrix} 1 & & \\ & 1 & 1 \\ & & 1 \end{bmatrix}$

例 14 ：求 Jordan 块的幂: $\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}^n$ $\begin{bmatrix}\lambda & 1 & \\ & \lambda & 1\\ & & \lambda\end{bmatrix}^n$

解：

$$\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}^2=\begin{bmatrix}\lambda^2 & 2\lambda\\ 0 & \lambda^2\end{bmatrix},\ \begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}^3=\begin{bmatrix}\lambda^3 & 3\lambda^2\\ 0 & \lambda^3\end{bmatrix},\cdots,\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}^n=\begin{bmatrix}\lambda^n & n\lambda\\ 0 & \lambda^n\end{bmatrix}$$

$$\begin{bmatrix}\lambda & 1 & \\ & \lambda & 1\\ & & \lambda\end{bmatrix}^2=\begin{bmatrix}\lambda^2 & 2\lambda & 1\\ & \lambda^2 & 2\lambda\\ & & \lambda^2\end{bmatrix},\ \cdots,$$

$$\begin{bmatrix}\lambda & 1 & \\ & \lambda & 1\\ & & \lambda\end{bmatrix}^n=\begin{bmatrix}\lambda^n & n\lambda^{n-1} & \frac{n}{2}(n-1)\lambda^{n-2}\\ & \lambda^n & n\lambda^{n-1}\\ & & \lambda^n\end{bmatrix}$$

例 15 ：设$[\alpha_1,\beta_2,\cdots,\beta_t]$是方阵 A 的属于本征值为$\lambda$的一个 Jordan 链，求证：$[\alpha_1,\beta_2,\cdots,\beta_t]$线性无关！

提示：利用性质：$(A-\lambda I)^k P_{k-1}=\alpha_1$; $(A-\lambda I)^k\beta_k=0,$

注：将此方法稍加推广即可证明前面给出的 Jordan 标准型的求法中，得到的所有α，β组成的向量组是线性无关的！

第七章　内积空间

内积的定义、欧空间

Schmidt 标准正交化方法、

正交矩阵（群）、正交变换（群）

对称变换、实对称矩阵的对角化

酉空间

§7.1、内积的定义、欧空间

内积、双线性泛函、欧几里得空间　*数学中常见的“空间”简介　度规

模（长度）、夹角　正交、标准正交基

内积、欧几里得空间

如前所述，线性代数起源于物理学中的矢量. 但，三维空间中的矢量还具有长度、夹角等性质. 为描述这些，必须在线性空间中附加更丰富的结构. 为此，我们先引入内积的概念，通过内积，可将长度、夹角等度量系统推广到更一般的、所有的线性空间中.

定义（内积） 设 V 是实数域 R 上的线性空间，若$\forall\alpha$，$\beta\epsilon V$，有唯一确定的一个实数（记作$(\alpha，\beta)$）与之对应，且满足：

（1）　$(\alpha，\beta)=(\beta，\alpha)$

（2）　$(\alpha+\beta，\gamma)=(\alpha，\gamma)+(\beta，\gamma)$

（3）　$(k\alpha，\beta)=k(\alpha，\beta)$

（4）　$\forall\alpha\neq 0$，$(\alpha，\alpha)>0$

则称$(\alpha，\beta)$为向量α与β的内积，V 称对这个内积来说的实内积空间，或称欧几里得空间.

注：内积是$V\times V\rightarrow R$的一个双线性映射，即双线性泛函！

内积的性质：

定理 7.1.1　设 V 是一欧空间，则：

（1）　$(\alpha，\beta+\gamma)=(\alpha，\beta)+(\alpha，\gamma)$

（2）　$(\alpha，k\beta)=k(\alpha，\beta)$　（由之可得$(\alpha，0)=(0，\alpha)=0$）

（3）　$(\alpha，\alpha)=0$当且仅当$\alpha=0$

（证明，直接由定义可得！）

例 1　三维空间的矢量（有方向的线段）构成的线性空间中，规定

$(\vec{\alpha}，\vec{\beta}):=|\vec{\alpha}||\vec{\alpha}|\cos\theta$　　（7.1.1）

可证明这是一个内积，此即通常的“点乘”.

例 2　R^n中，$X=[x_1，x_2，\cdots，x_n]^T$，$Y=[y_1，y_2，\cdots，y_n]^T$

若规定

$$(X，Y):=\sum_{i=1}^{n}x_i y_i=X^TY \qquad (7.1.2)$$

求证这是R^n的一个内积，故R^n对这样定义的内积来说构成了一个偶空间.

例 3　在R^n中，规定 $(X，Y):=\sum_{l=1}^{n}lx_l y_l$　　（7.1.3）

求证这也是R^n的一个内积，故R^n对这样定义的内积来说也构成一个欧空间.

可见：同一线性空间可定义不同的内积，使之成为不同的欧空间.

以后凡提到欧空间R^n时，恒指对上一例的内积而言的欧空间！

例 4　实$f(x)$，$g(x)\epsilon\varphi[a，b]$，规定：$(f,g)\coloneqq\int_a^b f(x)g(x)\,dx$，则满足内积的定义。故$\varphi[a，b]$可看成一个欧空间.

***数学中常见的“空间”简介**

数学分析研究的是函数的性质（连续性、微分积分等），泛函分析研究泛函的性质，即抽象集合到数的映射的性质。极限、连续性用到的$\varepsilon-\delta$语言，依赖于$|x-y|$，$|f(x)-f(y)|$，因此需在抽象集合中附加度量结构：

1.度量空间（距离空间）：定义了任意两元素间“距离”的集合.

$d(a, b)=d(b, a)$，$d(a, b)\geq 0$，$d(a, c)\leq d(a, b)+d(b, c)$

有了度量（距离）的概念，就可研究序列的极限、映射的连续性等。

2.赋范线性空间：定义了任意一个向量的“范数”（长度）的线性空间. $\|\alpha\|\geq 0$，$\|\beta\|=k|\alpha|$，$\|\alpha+\beta\|\leq\|\alpha\|+\|\beta\|$

赋范线性空间属于度量空间，因为由范数可以诱导定义距离：$d(\alpha, \beta)\coloneqq\|\alpha-\beta\|$.

3.Banach 空间：完备的（聚点在自身的）赋范线性空间.

空间的“完备”性（不同于向量组的完备性）简单说是指，空间中所有收敛序列的极限仍在该空间内。例如有理数就不完备，考虑序列：$a_1=1$，$a_{n+1}=1/(1+a_n)$，即：1，$\frac{1}{1+1}$，$\frac{1}{1+\frac{1}{1+1}}$，$\cdots$

4.内积空间：定义了内积的线性空间.（欧空间和酉空间）

内积空间属于赋范线性空间（从而也属于度量空间），因为由内积可诱导定义范数（从而定义距离）：

$$\|\alpha\|\coloneqq\sqrt{(\alpha, \alpha)} \qquad \mathrm{d}(\alpha, \beta)\coloneqq\|\alpha-\beta\|=\sqrt{(\alpha-\beta,\alpha-\beta)}$$

5.Hilbert 空间：完备的内积空间.

量子力学中物理态矢的集合即构成一个 Hilbert 空间.

闭区间$[a, b]$上所有平方可积的函数集合，也构成一个 Hilbert 空间.

平方可积：$|f(x)|^2$可积！记作$L^2[a，b]$，内积：

$$(f，g) \coloneqq \int_a^b f(x) g^*(x) dx$$

泛函分析中，线性映射和线性变换习惯被称作线性算符或线性算子.

度规矩阵

设 V 是 n 维欧空间,$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是一个基,则令$a_{ij} = (\alpha_i，\alpha_j)$，得到一个实对称矩阵.

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} (\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_n) \\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_n) \\ \vdots & \vdots & & \vdots \\ (\alpha_n,\alpha_1) & (\alpha_n,\alpha_2) & \cdots & (\alpha_n,\alpha_n) \end{bmatrix} \quad (7.1.4)$$

称欧空间在基$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$下的度规矩阵（或度量矩阵）.

可见　当给定一个基后，由内积就可以完全确定下度规矩阵. 反之，给定一个基下的度规矩阵，也可以完全确定内积. 这是由于：

$$(\alpha，\beta) = \left(\sum_{i=1}^n x_i \alpha_i，\sum_{i=1}^n y_i \alpha_i\right) = \sum_{ij} x_i (\alpha_i，\alpha_j) y_i = X^T A Y \quad (7.1.5)$$

与线性变换的矩阵类似，度规矩阵依赖于基的选择，可证：设基的过渡矩阵为 C，则 A 的变换为：$\mathrm{A} \rightarrow C^T A C$.

例 5　写出欧空间R^n在基$\{e_1，e_2，\cdots，e_n\}$和基.

$\{e_1，e_1 + e_2，e_1 + e_2 + e_3，\cdots，e_1 + e_2 + \cdots + e_n\}$下的度规矩阵.

解：

$e_1 = [1，0，\cdots，0]^T$，$e_2 = [0，1，\cdots，0]^T$，$\cdots$，$e_n = [0，0，\cdots，1]^T$

设分别为 A 和 B，则：

$a_{ij} = (e_i，e_j)$

$$b_{ij}=\left(\sum_{k=1}^{i}e_k,\ \sum_{l=1}^{i}e_j\right)=Min(i,\ j)$$

故$A=\begin{bmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&1\end{bmatrix}$，$B=\begin{bmatrix}1&1&1&\cdots&1\\1&2&2&\cdots&2\\1&2&3&\cdots&3\\\vdots&\vdots&\vdots&&\vdots\\1&2&3&\cdots&n\end{bmatrix}$

模、夹角

向量的模（或长度）

定义（模） 欧空间中，$\sqrt{(\alpha,\ \alpha)}$称向量α的模或长度，记作$|\alpha|$. 由定义，零向量长度为零；非零向量长度是一正数.

定义（单位向量）：模等于 1 的向量称单位向量.

例 6 欧空间R^n中，$|x|=\sqrt{x^Tx}=\sqrt{{x_1}^2+{x_2}^2+\cdots+{x_n}^2}$.

例 7 实φ$[a,\ b]$中，$|f(x)|=\sqrt{\int_a^b f^2(x)}\,dx$.

模的性质

定理 7.1.2:欧空间 V 中：

（1） $|k\alpha|=|k||\alpha|$.

（2） 柯西不等式：$\left|(\alpha,\ \beta)\le|\alpha||\beta|\right|$，等号当且仅当$\alpha$，$\beta$线性相关成立.

（3） 三角不等式：$|\alpha+\beta|\le|\alpha|+|\beta|$，等号当且仅当$\alpha$，$\beta$线性相关成立.

证：（1）$|k\alpha|=\sqrt{(k\alpha,\ k\alpha)}=\sqrt{k^2(\alpha,\ \alpha)}=|k||\alpha|$，证毕.

（2）线性相关时，易证等号成立. 下面考虑α，β线性无关，则$\forall \mathrm{k} \in \mathrm{R}$，取$\alpha - k\beta = \gamma \neq 0$，则$|\gamma| > 0$，即$|\gamma|^2 > 0$，而$|\gamma|^2 = (\alpha - k\beta，\alpha - k\beta) = |\alpha|^2 - 2k(\alpha，\beta) + k^2|\beta|^2$，由于此时$|\beta|^2 > 0$，故对$\forall \mathrm{k} \in \mathrm{R}$，上式都为正，必有判别式：

$\Delta = 4(\alpha，\beta)^2 - 4|\alpha|^2|\beta|^2 < 0$，证毕.

（3）$|\alpha + \beta|^2 = (\alpha + \beta，\alpha + \beta) = |\alpha|^2 + 2(\alpha，\beta) + |\beta|^2$

$\leq |\alpha|^2 + 2|\alpha||\beta| + |\beta|^2 = (|\alpha| + |\beta|)^2$，证毕.

例 8　对欧空间R^n应用柯西不等式，有：

$$(X，Y)^2 = (\sum_{i=1}^n x_i\, y_i)^2 \leq (\sum_{i=1}^n x_i{}^2)(\sum_{i=1}^n \mathrm{y}_i{}^2) \quad (7.1.6)$$

例 9 对欧空间$\varphi[a，b]$应用柯西不等式，有：

$$\left|\int_a^b f(x)g(x)\right| \leq \sqrt{\int_a^b f^2(x)dx \int_a^b g^2(x)dx}\ .$$

两向量间的夹角：

由于柯西不等式：对$\alpha \neq 0$，$\beta \neq 0$有　$\dfrac{|(\alpha，\beta)|}{|\alpha||\beta|} \leq 1$　，亦即：

$$-1 \leq \frac{(\alpha，\beta)}{|\alpha||\beta|} \leq 1 \qquad (7.1.7)$$

故可引入夹角的概念.

定义（夹角）　设α，β是欧空间 V 中的两非零向量，则α，β间的夹角θ为：

$$\cos\theta = \frac{(\alpha，\beta)}{|\alpha||\beta|} \quad ，其中0 \leq \theta \leq \pi.$$

注：$|\alpha + \beta|^2 = |\alpha|^2 + |\beta|^2 + 2|\alpha||\beta|\cos\theta$，余弦定理.

正交和标准正交基

定义（正交） 若$(\alpha, \beta) = 0$，则称α和β正交.

可见：零向量与所有向量都正交.

定义 欧空间 V 中的一个向量组，若所有向量均非零，且两两正交，则称该向量组为一个正交向量组；进一步，若正交向量组中每一个向量都是单位向量，则称该向量组为一个标准正交向量组.

正交向量组的性质：

定理 7.1.3 设$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$是一个正交向量组，则

$\alpha_1, \alpha_2, \cdots, \alpha_n$线性无关。

即：正交向量组中的向量必线性无关。

证：设$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n = 0$

则：$(\alpha_i, k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n) = (\alpha_i, 0) = 0$

即：$k_i(\alpha_i, \alpha_i) = 0$，故$k_i = 0(i = 1, 2, \cdots, n)$，证毕。

推论 7.1.4 n 维欧空间 V 中正交向量组所含向量个数不大于 n.

由定理 7.1.3，在 n 维欧空间 V 中若一个正交向量组正好含有 n 个向量，则这个正交组是 V 的一个基。引入如下定义：

定义 由 V 中相互正交的向量构成的基，承做 V 的正交基；若正交基中每个向量均是单位向量，则称该基是 V 的标准正交基.

$\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\}$：$(\varepsilon_i, \varepsilon_j) = \delta_{ij}$

例 10 欧空间R^n中，

$$e_1 = [1, 0, \cdots, 0]^T, e_2 = [0, 1, \cdots, 0]^T, \cdots, e_n = [0, 0, \cdots, 1]^T$$

是一个标准正交基；

$\{e_1, e_1+e_2, e_1+e_2+e_3, \cdots, e_1+e_2+\cdots+e_n\}$不是标准基.

例 11 设

$$\alpha=[1, 2, -1, 1]^T, \beta=[2, 3, 1, -1]^T, \cdots, \gamma=[-1, -1, -2, 2]^T$$

求α，β，γ的长度，每两向量的内积，每两向量的夹角，以及与

α，β，γ均正交的所有向量和单位向量.

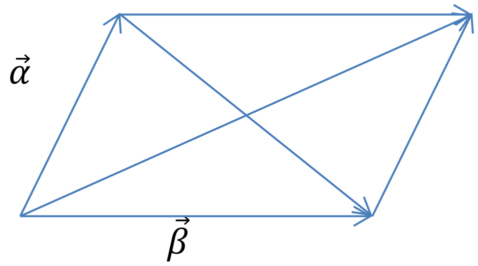
解：X与α，β，γ正交，有：$\alpha^T X=0$，$\beta^T X=0$，$\gamma^T X=0$，

故$\begin{bmatrix}\alpha^T\\ \beta^T\\ \gamma^T\end{bmatrix}X=0$。

例 12 欧空间R^n中，求证：$|\alpha+\beta|^2+|\alpha-\beta|^2=2|\alpha|^2+2|\beta|^2$ 。

在平面上上式得几何含义是什么？

解：



§7.2、 Schmidt 标准正交化方法、正交矩阵（群）、正交变换（群）

Schmidt 标准正交化方法　正交矩阵（群）　正交变换（群）

若已知 n 维欧空间 V 的一个标准正交基$\{\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_n\}$，则 V 在该基下的度规矩阵，V 中向量的坐标，V 中任意两个向量的内积，V 中向量长度的坐标表达式均变得特别简单：

度规 A：$a_{ij}=(\varepsilon_i，\varepsilon_j)=\delta_{ij}$, 故$A=I_n$.

向量α的坐标：$x_i=(\varepsilon_i，\alpha)$，　$\alpha=\sum x_i\,\varepsilon_i=\sum(\varepsilon_i，\alpha)\,\varepsilon_i$

向量$\alpha=[\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_n]X$，$\beta=[\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_n]Y$

则$(\alpha，\beta)=X^TAY=X^TI_nY=X^TY=x_1y_1+x_2y_2+\cdots+x_ny_n$,

$|\alpha|^2=(\alpha，\alpha)=X^TX={x_1}^2+{x_2}^2+\cdots+{x_n}^2$.

可见： n 维欧空间中使用标准正交基将给计算带来许多方便。那么，在任意的 n 维欧空间中是否一定存在标准正交基？如果存在，又如何找出一个标准正交基？答案由 Schmidt 标准正交化方法给出：

Schmidt 标准正交化方法：

定理 7.2.1　n 维欧空间 V 中，从任意一族基$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$出发，可以构造一组标准正交基$\{\gamma_1，\gamma_2，\cdots，\gamma_n\}$使$\gamma_i$为$\alpha_1，\alpha_2，\cdots，\alpha_n$的线性组合。

证：构造

$$\begin{cases} \beta_1=\alpha_1 \\ \beta_2=\alpha_2-\frac{(\alpha_2，\beta_1)}{(\beta_1，\beta_1)}\beta_1 \\ \beta_3=\alpha_3-\frac{(\alpha_3，\beta_1)}{(\beta_1，\beta_1)}\beta_1-\frac{(\alpha_3，\beta_2)}{(\beta_2，\beta_2)}\beta_2 \\ \vdots \\ \beta_n=\alpha_n-\sum_{i=1}^{n-1}\frac{(\alpha_n，\beta_i)}{(\beta_i，\beta_i)}\beta_i \end{cases} \qquad (7.2.1)$$

由归纳法可证所有的β_i均为零（α_1，α_2，$\cdots$，α_n线性无关得到），并且每个β_i均为β_1到β_{i-1}正交，故$\{\beta_1，\beta_2，\cdots，\beta_n\}$是一个正交基，进一步，构造：$\gamma_1=\beta_1/|\beta_1|$，$\gamma_n=\beta_n/|\beta_n|$, 则$\{\gamma_1，\gamma_2，\cdots，\gamma_n\}$即一个标准正交基！

例 13　试从R^3的基$\alpha_1=(1，1，0)^T$，$\alpha_2=(2，0，1)^T$，$\alpha_3=(2，2，1)^T$出发构造一组标准正交基.

例 14　从R^3的基$\alpha_1=(1，1，1)^T$，$\alpha_2=(1，2，1)^T$，$\alpha_3=(0，-1，1)^T$构造标准正交基.

例 15　行列式

$$g(\alpha_1，\alpha_2，\cdots，\alpha_n)=\begin{vmatrix}(\alpha_1，\alpha_1) & (\alpha_1，\alpha_2) & \cdots & (\alpha_1，\alpha_n)\\(\alpha_2，\alpha_1) & (\alpha_2，\alpha_2) & \cdots & (\alpha_2，\alpha_n)\\ \vdots & \vdots & \cdots & \vdots\\(\alpha_n，\alpha_1) & (\alpha_n，\alpha_2) & \cdots & (\alpha_n，\alpha_n)\end{vmatrix}$$

(7.2.1)

称欧空间中向量α_1，α_2，$\cdots$，α_n的 Gram 行列式.求证：

α_1，α_2，$\cdots$，α_n施行正交变化过程得到正交向量组β_1，β_2，$\cdots$，β_n，则有：

$$g(\alpha_1，\alpha_2，\cdots，\alpha_n)=g(\beta_1，\beta_2，\cdots，\beta_n).$$

提示：α到β的过渡矩阵是上三角矩阵，故度规$|C^TGC|=|G|$.

例 16 求证：在欧空间R^n中，n 个列向量x_1，x_2，⋯，x_n线性无关的充要条件是：

$$g(x_1,\ x_2,\ \cdots,\ x_n)=\begin{vmatrix}(x_1,\ x_1) & (x_1,\ x_2) & \cdots & (x_1,\ x_n)\\ (x_2,\ x_1) & (x_2,\ x_2) & \cdots & (x_2,\ x_n)\\ \vdots & \vdots & \cdots & \vdots\\ (x_n,\ x_1) & (x_n,\ x_2) & \cdots & (x_n,\ x_n)\end{vmatrix}\neq 0$$

证：令$A_{n\times n}=[x_1,\ x_2,\ \cdots,\ x_n]$，注意到$(x_i,\ x_j)={x_i}^T x_j$

故上述行列式$=|A^TA|=|A|^2$。

例 17 求证：欧空间中，n 个向量α_1，α_2，⋯，α_n线性无关的充要条件是 Gram 行列式$g(\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n)\neq 0$.

提示：选取标准正交基，将$g(\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n)$用$\alpha_1$，$\alpha_2$，⋯，$\alpha_n$的坐标表示，并用上例的结论！

例 18 求证 Gram 行列式满足不等式：$g(\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n)\geq 0$.

提示：证明同上例。

正交矩阵、正交变换

为了后面的需要我们引入正交矩阵的概念，并讨论正交矩阵与标准正交基的关系。正交矩阵是数学和物理学中一类极其重要的矩阵。

定义（正交矩阵） $A^TA=I_n$的 n 阶实矩阵 A 称为正交矩阵。

正交矩阵的性质：

（1）$|A|=\pm 1$，可逆！　　　　（2）$A^{-1}=A^T$

（3）$AA^T=A^TA=I$　　　　（4）A^T（即A^{-1}）也是正交矩阵

定理 7.2.2 若A，B均为 n 阶正交矩阵，则AB亦是。

证：$(AB)^T(AB)=B^TA^TAB=B^TB=I$（物理上有用！）

由于上述正交矩阵的性质，以及定理 7.2.2，可得到如下重要结论：

定理 7.2.3 所有 n 阶实正交矩阵的集合，对于矩阵的乘法构成一个群，称实正交矩阵群，记作$\mathrm{O}(n)$.

证：封闭性，结合律，单位元（I_n），逆元.

回顾上一章：

（1）数域上的所有 n 阶可逆矩阵的集合，对于矩阵的乘法构成一个群，称为一般线性矩阵群，记作$\mathrm{GL}(n，F)$.

（2）对应地，数域中的 n 维线性空间上的所有非奇异（可逆）线性变换，对于线性变换的乘法（相继变换）也构成一个群. 该群与（1）的矩阵群同构，称为一般线性变换群，记作$\mathrm{GL}(n，F)$.

n 阶实正交矩阵集是 n 阶实矩阵集的子集，故$\mathrm{O}(n)$群是$\mathrm{GL}(n)$群的一个子集。

实正交矩阵 A，若$|A| = 1$，则称 A 为幺模实正交矩阵. 显然有：

定理 7.2.3' 所有 n 阶幺模实正交矩阵的集合，对于矩阵的乘法构成群，称幺模实正交矩阵群，记作$\mathrm{SO}(n)$.

$\mathrm{O}(n)$，$\mathrm{SO}(n)$对应的线性变换群，以及几何意义，后面再讨论.

正交矩阵与标准正交基的关系：

定理 7.2.4 n 阶实矩阵 A 是正交矩阵的充要条件是：A 的 n 个行（列）向量构成R^n的标准正交矩阵基.

证明：将 A 按行分组：$A = [A_1，A_2，\cdots，A_n]$，则：

$A^TA=\begin{bmatrix}(A_1, A_1) & \cdots & (A_1, A_n)\\ \vdots & \vdots & \vdots\\ (A_n, A_1) & \cdots & (A_n, A_n)\end{bmatrix}$列向量得证！同理，按列分组，由

AA^T可证行向量.

定理 7.2.5 设$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$是 n 维欧空间的一个标准正交基，P 是一个 n 阶实矩阵，$[\beta_1, \beta_2, \cdots, \beta_n]=[\alpha_1, \alpha_2, \cdots, \alpha_n]P$，则：$\{\beta_1, \beta_2, \cdots, \beta_n\}$是 V 的标准正交基的充要条件是 P 为一正交矩阵.

证：考虑

$$(\beta_i, \beta_j)=\left(\sum_s \alpha_s P_{si}, \sum_t \alpha_{\mathrm{t}} P_{tj}\right)=\sum_{s,t} P_{si}P_{tj}(\alpha_s, \alpha_{\mathrm{t}})$$
$$=\sum_{s,t} P_{si}P_{tj}\delta_{st}=\sum_{s,t} P_{si}P_{tj}=(P^TP)_{ij}$$

小结：

（1）标准正交基之间的过渡矩阵是正交矩阵。

（2）正交矩阵把一个标准正交基变为另一个标准正交基。

（3）正交矩阵的行向量（或列向量）是标准正交基。

下面考虑正交变换：

几何和物理学中经常会要求线性变换不改变向量的长度，因此引入：

定义（正交变换） 设σ是欧空间 V 中的线性变换，若∀α∈V，有：$|\sigma(\alpha)|=|\alpha|$，则称σ为一个正交变换.

关于正交变换与正交矩阵、标准正交基的关系，有如下定理

定理 7.2.5 设σ是欧空间 V 中的线性变换，则如下论述等价：

（1）σ是正交矩阵.

（2）∀α，β∈V，有$\big(\sigma(\alpha), \sigma(\beta)\big)=(\alpha, \beta)$.

（3）σ将 V 的一个标准正交基变成 V 的另一个标准正交基.

（4）σ关于 V 的任意一个标准正交基的矩阵是正交矩阵.

证：循环证明：

（1）$\Longrightarrow$（2）考虑：$\big(\sigma(\alpha+\beta)，\sigma(\alpha+\beta)\big)=(\alpha+\beta，\alpha+\beta)$

（2）$\Longrightarrow$（3）考虑：$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$，$\big(\sigma(\alpha_i)，\sigma(\alpha_j)\big)=(\alpha_i，\alpha_j)=\delta_{ij}$

（3）$\Longrightarrow$（4）$[\sigma(\alpha_i)，\cdots，\sigma(\alpha_n)]=[\alpha_1，\alpha_2，\cdots，\alpha_n]P$，由定理 7.2.4 即得.

（4）$\Longrightarrow$（1）设$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是标准正交基，$\sigma$的矩阵为 P，则$\forall\alpha\epsilon V$，

设α坐标为x，则$\sigma(\alpha)$坐标为 Px，注意此时度规为I_n，

故$\big(\sigma(\alpha)，\sigma(\alpha)\big)=(\mathrm{P}x)^T\mathrm{P}x=x^Tx=(\alpha，\alpha)$，证毕.

注意：正交变换保持长度、夹角、距离等度量不变！

（1）由于对正交变换，$\ker(\sigma)=\{\alpha:\sigma(\alpha))=0\}$，而$|\alpha|=|\sigma(\alpha)|$，

故$\ker(\sigma)=\{\alpha:\sigma(\alpha))=0\}$，即$\ker(\sigma)=\{0\}$，

故正交变换是单映射，也即：正交变换是可逆线性变换.

（2）$\big|\sigma\big(\sigma^{-1}(\alpha)\big)\big|=|\sigma^{-1}(\alpha)|=|\alpha|$，故逆变换也是正交变换.

（3）$|\sigma\circ\tau(\alpha)|=\big|\sigma\big(\tau(\alpha)\big)\big|=|\tau(\alpha)|=\tau(\alpha)$，故乘积，亦是正交变换.

（4）恒等变换$\varepsilon(\alpha)=\alpha$是正交变换.

综上有如下定理：

定理 7.2.6　n 维欧空间上的正交变换全体，对线性变换的乘法来说构成一个群，称作实正交变换群.

显然，在选定一组标准正交基之后，n 维实正交矩阵群与 n 维实正交矩阵群是同构的. 因此 n 维实正交矩阵群也用同样的标记$\mathrm{O}(n)$.

标准正交基、正交矩阵、正交变换三者关系：

标准正交基$\{\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_n\}$：$\left(\varepsilon_i，\varepsilon_j\right)=\delta_{ij}$

正交矩阵 A：$A^TA=AA^T=I$

正交变换α：保持长度（及内积）不变

标准正交基与正交矩阵：

A_n是正交矩阵$\Leftrightarrow A_n$的行（列）向量组是R^n的标准正交基.

正交矩阵与正交变换：

σ是正交变换$\xLeftrightarrow{\text{通过标准正交基联系}}$σ在标准正交基下的矩阵是正交矩阵.

正交变换与标准正交基

σ是正交变换⇔σ将标准正交基变为另一个标准正交基.

为进一步看清正交变换的几何意义，我们分别考虑二维和三维实线性空间中的正交变换实例：（正交变换：R^n中的“转动”和“反射”变换！）

注意到所有的 n 维实线性空间都同构（且都与R^n同构），故我们以二维矢量空间（平面上有方向的线段集合）和三维矢量空间（空间中有方向的线段集合）来考虑：

例 19 在R^2中，将一个矢量旋转一个角θ的变换，是R^2的一个线性变换，且是一个正交变换，这个变换关于标准正交基$\{\hat{\imath}，\hat{\jmath}\}$的矩阵是

$$\mathrm{P}=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

故R^2中的每一个转动，对应SO(2)中的一个矩阵，反之可证明，每一个SO(2)中的矩阵，必定对应R^2中的一个转动（只需注意到SO(2)中

的矩阵将标准正交基变为标准正交基，故只可能是转动或反射，再注意到行列式即得）. 故SO(2)是二维转动群！

例 20 考虑R^3中，关于 xy 平面的反射变换 H，关于原点的反射变换 P，关于 z 轴旋转180°的变换π，易证明这些都是线性变换，且都是正交变换（保持矢量长度不变）. 求关于标准正交基$\{\hat{\imath}，\hat{\jmath}，\hat{k}\}$的矩阵.

解：$H=\begin{bmatrix}1 & & \\ & 1 & \\ & & -1\end{bmatrix}$ $P=\begin{bmatrix}-1 & & \\ & -1 & \\ & & -1\end{bmatrix}$ $\pi=\begin{bmatrix}-1 & & \\ & -1 & \\ & & 1\end{bmatrix}$

故$P=H\pi$（宇称变换等于镜面反射后再旋转180°！）

R^3中宇称变换和旋转变换都是O(3)中的元素，反之，可证明任意一个O(3)中的元素必对应一个宇称或旋转或转动后再做宇称变换（由于正交变换将$\{\hat{\imath}，\hat{\jmath}，\hat{k}\}$仍变成标准正交基，或者说，保持长度和距离均不变，故只能是转动或反射）.进一步，注意到$\sigma(\hat{\imath})\cdot\left(\sigma(\hat{\jmath})\times\sigma\left(\hat{k}\right)\right)=|A|$，故SO(3)中的元素还保持手性，故只能是转动.

故：三维幺模正交矩阵群SO(3)就是三维转动群！

§7.3、对称变换、实对称矩阵的对角化

对称变换与对称矩阵　　　　　　　　　　实对称矩阵的本征值全为实数

实对称矩阵属于不同本征值的本征向量正交

实对称矩阵必可对角化（本征向量组的完备性）

除正交变换外，欧空间中另一类重要的变换是对称变换：

定义（对称变换）. σ是欧空间 V 上的一个线性变换，若σ满足：∀α，β∈V，$\big(\sigma(\alpha), \beta\big)=\big(\alpha, \sigma(\beta)\big)$，则称σ是 V 上的一个对称变换.

对称变换和对称矩阵有如下联系：

定理 7.3.1 欧空间 V 上的线性变换σ是对称变换的充要条件是：σ关于 V 的标准正交基的矩阵是对称矩阵.

证明：(⟹)：$\big(\sigma(\alpha_i), \alpha_j\big)=\big(\alpha_i, \sigma(\alpha_j)\big)$，故 $\left(\sum_k \alpha_k A_{ki}, \alpha_j\right)=\left(\alpha_i, \sum_l \alpha_l A_{lj}\right)$

即：$A_{ji}=A_{ij}$，故$A=A^T$

(⟸)：取 $\alpha=\sum_i \alpha_i x_i$ ，$\beta=\sum_j \alpha_j y_j$

故 $$\big(\sigma(\alpha), \beta\big)=\sum_{i,j} x_i\big(\sigma(\alpha_i), \alpha_j\big)y_j=\sum_{i,j} x_i A_{ji} y_j$$

$$\big(\alpha, \sigma(\beta)\big)=\sum_{i,j} x_i\big(\alpha_i, \sigma(\alpha_j)\big)y_j=\sum_{i,j} x_i A_{ij} y_j$$

而$A=A^T$，故$\big(\sigma(\alpha), \beta\big)=\big(\alpha, \sigma(\beta)\big)$，证毕.

注意：对称变换的集合不构成群，实对称矩阵的集合也不构成群！

（1） $(AB)^T=B^TA^T=BA\neq AB$

（2） $\big(\sigma\circ\tau(\alpha), \beta\big)=\big(\sigma(\tau(\alpha)), \beta\big)=\big(\tau(\alpha), \sigma(\beta)\big)$

$$=\left(\alpha,\ \tau\left(\sigma(\beta)\right)\right)=\left(\alpha,\ \tau\circ\sigma(\beta)\right)\neq\left(\alpha,\ \sigma\circ\tau(\beta)\right)$$

欧空间中对称变换（对应于实对称矩阵）的重要性在于：

1. 对称矩阵的本征值一定为实数；

2. 对称矩阵的所有本征向量可选为两两正交的单位向量；

3. 对称变换（实对称矩阵）一定可以对角化。（本征向量组完备）

由于选定标准正交基之后，对称变换和实对称矩阵一一对应，为此我们下面只考虑实对称矩阵：

定理 7.3.2 实对称矩阵的本征值全是实数.

证：设$Ax=\lambda x$ 由于实对称：$A=A^{+}:=\overline{A^{T}}$

故$(Ax)^{+}x=(x^{+}A^{+})x=x^{+}(Ax)$，即：$(\lambda x)^{+}x=x^{+}(\lambda x)$

即：$\bar{\bar{\lambda}}(x^{+}x)=\lambda(x^{+}x)$，而$x^{+}x>0$，故$\bar{\bar{\lambda}}=\lambda$。

由于实对称矩阵的本征值都是实数，故实对称矩阵的本征向量都可选为R^{n}中的向量（实的列向量）：$(\lambda I-A)=0$.

定理 7.3.3 实对称矩阵的属于不同本征值的本征向量相互正交.

证：设$\mu\neq\lambda$是两本征值，$Ax=\mu x$，$Ay=\lambda y$,

考虑$(Ax)^{T}y=x^{T}A^{T}y=x^{T}Ay=x^{T}(Ay)$

故$(\mu x)^{T}y=x^{T}(\lambda y)$，即：$(\mu-\lambda)(x^{T}y)=0$

而$\mu\neq\lambda$，故$x^{T}y=\left(x,\ y\right)$，正交。

定理 7.3.4 实对称矩阵必可通过正交矩阵对角化。即：设 A 为实对称矩阵，则必存在实正交矩阵 C，使$C^{-1}AC=C^{T}AC$为对角矩阵.

证：用数归. 当A_{n}的阶$\mathrm{n}=1$时显然成立.

假设$n-1$阶时成立，考虑n阶：A至少有一个本征值和单位本征向量，设为λ_1和P_1：$AP_1=\lambda_1 P_1$，$|P_1|=1$.

则通过逐步添加，可从P_1构造 R^n的一个基：$\{P_1, q_2, \cdots, q_n\}$，用Schmidt 标准正交化方法可得$R^n$的一个标准正交基$\{P_1, P_2, \cdots, P_n\}$，令$P=[P_1, P_2, \cdots, P_n]$，$P$为正交矩阵，故

$AP=[AP_1, AP_2, \cdots, AP_n]=[P_1, P_2, \cdots, P_n]B=PB$

其中$B=\begin{bmatrix}\lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & *\end{bmatrix}$，而$B=P^{-1}AP=P^TAP$，故B亦为实对称，

故$B=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & *\end{bmatrix}=\begin{bmatrix}\lambda_1 & 0 \\ 0 & B_{n-1}\end{bmatrix}$，$B_{n-1}$为$n-1$阶实对称，

故$B_{n-1}={Q_{n-1}}^{-1}F_{n-1}Q_{n-1}={Q_{n-1}}^TF_{n-1}Q_{n-1}$，令$Q=\begin{bmatrix}1 & 0 \\ 0 & Q_{n-1}\end{bmatrix}$，则$Q^{-1}=Q^T$，且$Q^{-1}BQ=\begin{bmatrix}\lambda_1 & 0 \\ 0 & {Q_{n-1}}^{-1}B_{n-1}Q_{n-1}\end{bmatrix}=F_n$，故令$C=PQ$即得证！

推论 7.3.5 实对称矩阵的任意一个本征值λ的重数S_λ，与该本征值λ对应的本征子空间E_λ的维数$\dim E_\lambda$相等.

即：$S_\lambda=\dim E_\lambda$

注：本征值是 n 重根，必然有 n 个线性无关的本征向量，全部本征向量构成的向量组是完备的，可作为线性空间R^n的一个基！

用正交矩阵对角化实对称矩阵（即对实对称矩阵 A，求正交矩阵 P 使$P^{-1}AP=P^TAP=diag[\lambda_1, \lambda_2, \cdots, \lambda_n]$）的步骤如下：

（1）由$|\lambda I-A|=0$，求出A_n的所有不同本征值λ_1，λ_2，$\cdots$，λ_n，以及它们的重数S_1，S_2，$\cdots$，S_n，显然$\lambda_i\in R$，且 $\sum_{i=1}^{m}S_i=n$

（2)对每一个λ_i，由$|\lambda_i I - A|x = 0$，求其基础解系α_{i1}，α_{i2}，$\cdots$，α_{iS_i}，再用 Schmidt 方法将其标准正交化为P_{i1}，P_{i2}，$\cdots$，P_{iS_i}.

（3）将求出的 n 个相互正交的单位本征向量排成矩阵：

$$P = \left[P_{11}, P_{12}, \cdots, P_{1S_1}, P_{21}, P_{22}, \cdots, P_{2S_2}, \cdots, P_{m1}, P_{m2}, \cdots, P_{mS_m}\right]$$

则$P^T = P^{-1}$，且

$$P^{-1}AP = P^TAP = diag\left[\lambda_1, \lambda_1, \cdots, \lambda_1, \lambda_2, \lambda_2, \cdots, \lambda_2, \cdots, \lambda_m, \lambda_m, \cdots, \lambda_m\right]$$

例 21 $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$，$B = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 2 & -2 \\ 0 & -2 & 3 \end{bmatrix}$，用正交矩阵分别将 A 和 B 相似对角化.

例 22 两个n阶实对称矩阵A，B可对易（即：AB = BA）的充要条件是 A 和 B 可具有相同的完备本征向量组（n 个，线性无关），或者说，A 和 B 可同时对角化（用同一个相似变换矩阵对角化）.

证：（充分性⟸）：两种证法.

方法一：设$P^{-1}AP = F_1$，$P^{-1}BP = F_2$，

则$\mathrm{AB} - \mathrm{BA} = P(F_1F_2 - F_2F_1)P^{-1}$.

方法二：设A，B有共同的线性无关本征向量组$\{P_1, P_2, \cdots, P_n\}$，

对应的本征值分别是λ_i和μ_i，$(i = 1, 2, \cdots, n)$,

则$(AB - BA)P_i = (\lambda_i\mu_i - \mu_i\lambda_i)P_i = 0$.

$\forall \mathrm{x} \in R^n$，有 $x = \sum_{i=1}^{n} k_iP_i$ ，故$(AB - BA)x = 0$.

由于x任意，故$\mathrm{AB} - BA = 0$。（方法二可推广到对称变换）

（必要性$\Longrightarrow$）：假设$AB=BA$，由于 A，B 均是实对称矩阵，可对角化，故设 A 的一组线性无关本征向量组为$\{P_1, P_2, \cdots, P_n\}$对应本征值$\lambda_1, \lambda_2, \cdots, \lambda_n$。

（1）若λ_i非 A 的重本征值，则

$\mathrm{A}(BP_i)=(AB)P_i=(BA)P_i=B(AP_i)=B(\lambda_i P_i)=\lambda_i(BP_i)$

故$BP_i \epsilon E_{\lambda_i}$，即：$BP_i=\mu_i P_i$，$P_i$亦是 B 的本征向量.

（2）若λ_i是 A 的重本征值，设重数为S_i，则$\{P_1, P_2, \cdots, P_n\}$中对应$\lambda_i$的本征向量必有$S_i$个，重新命名为$\{q_{i1}, q_{i2}, \cdots, q_{is}\}$，化成$E_{\lambda_i}$，则同样由$ABq_{ij}=BAq_{ij}=\lambda_i Bq_{ij}$可知：$Bq_{i1}, Bq_{i2}, \cdots, Bq_{is} \in E_{\lambda_i}$，但一般地，$Bq_{il} \neq \mu_i q_{il}$，选取 $\widetilde{q_l}=\sum_{l=1}^{S_i} x_l\, q_{il}$ ，使$B\widetilde{q_l}=\mu\widetilde{q_l}$，则与$q_{im}$作内积有：

$(q_{im}, B\widetilde{q_l})=\mu(q_{im}, \widetilde{q_l})$，即：

$$\sum_{l=1}^{S_i}({q_{im}}^T Bq_{il})\, x_l=\sum_{l=1}^{S_i}\mu\delta_{ml}\, x_l$$

其中${q_{im}}^T Bq_{il}$记作C_{ml}

$B=B^T \Longrightarrow C_{ml}=C_{lm}$，排成$S_i \times S_i$矩阵 C，$C^T=C$，$x=[x_1, \cdots, x_{S_i}]^T$，方程化为$Cx=\mu x \Longrightarrow (\mu I-C)x=0$，必有$S_i$个线性无关的$x$，故必有$S_i$个线性无关的$\widetilde{q_l}$，使其同时是 A，B 的本征向量，证毕！

例 23 设 V 是一欧空间，$\gamma \epsilon \mathrm{V}$是一固定非零向量，则定义$\sigma$如下：

$\forall \alpha \epsilon \mathrm{V}, \sigma(\alpha) \coloneqq \alpha-\frac{2(\alpha, \gamma)}{(\gamma, \gamma)}\gamma$，求证：$\sigma$是 V 的一个正交变换，且$\sigma^2=\varepsilon$，$\varepsilon$是恒等变换.

提示：考虑 n 行意义，注意到$\alpha-\frac{(\alpha，\gamma)}{(\gamma，\gamma)}\gamma$与$\gamma$正交.

例 24 证明：正交矩阵本征值的模为 1。

证：设$A^TA=I$，$Ax=\lambda x$,则考虑上式得厄米共轭（注意，考虑转置不够，因为λ，x可能含复数）$x^+A^+=\bar{\lambda}x^+$，故$x^+A^+Ax=\bar{\lambda}x^+\lambda x$，即：$x^+x=(\bar{\lambda}\lambda)x^+x$，由于$x$不是零向量，故$x^+x=|x_1|^2+|x_2|^2+\cdots+|x_n|^2>0$，故$\bar{\lambda}\lambda=1$，即：$|\lambda|=1$.

§ 7.4、酉空间

酉空间：复内积空间、厄米共轭 度规矩阵　　长度、正交、标准正交基　　施密特标准正交化方法　　幺正矩阵（群）、幺正变换（群）　　厄米矩阵、厄米变换、厄米矩阵必可对角化

前面讨论了欧空间，即定义了内积的实线性空间. 实际上通过恰当定义内积可将欧空间的很多结论都推广到复线性空间，即么正空间. 么正空间的理论平行于欧空间，故不再细讲证明，但借此机会补充几个无限维线性空间的例子.

定义（内积） 设 V 是复数域 C 上的线性空间，若$\forall\alpha$，$\beta\in V$，有唯一确定的一个复数（记作(α,β)）与之对应，且满足：

（5）　$(\alpha,\beta)=\overline{(\beta,\alpha)}$，“——”表示共轭复数，即复共轭。

（6）　$(\gamma,\alpha+\beta)=(\gamma,\alpha)+(\gamma,\beta)$

（7）　$(\alpha,k\beta)=k(\alpha,\beta)$

（8）　$\forall\alpha\neq 0$，(α,α)为正实数.

（9）　则称(α,β)为向量α与β的内积，V 称对这个内积来说的复内积空间，或称幺正（酉）空间.

注：（1）尤其注意和（1）条与欧空间的区别！

（2）可见，当限定在 R 上时，上述定义退化到欧空间的内积。

内积的性质：

定理 7.4.1 设 V 是一酉空间，则：

（1）$(\alpha+\beta，\gamma)=(\alpha，\gamma)+(\beta，\gamma)$

（2）$(k\alpha，\beta)=\bar{k}(\alpha，\beta)$ （由之可得$(\alpha，0)=(0，\alpha)=0$）

（3）$(\alpha，\alpha)=0$当且仅当$\alpha=0$

（证明，直接由定义可得！注意（2）与欧空间的区别）

厄米共轭：

定义（厄米共轭矩阵） 设$A\in C^{n\times n}$，则$\overline{A^T}$称为 A 的厄米共轭矩阵，记作A^+.即：$A^+:=\overline{A^T}$.

由转置运算的性质易推出厄米共轭运算的性质：

$(AB)^+=B^+A^+$，$(A_1A_2\cdots A_n)^+={A_n}^+{A_{n-1}}^+\cdots {A_1}^+$ （7.4.1）

定义（厄米共轭转换） 设σ是酉空间 V 上的线性变换，若存在另一个变换τ满足∀α，β∈V，$(\tau(\alpha)，\beta)=(\alpha，\sigma(\beta))$，则称$\tau$为$\sigma$的厄米共轭变换，记作$\tau=\sigma^+$。（可证明$\sigma^+$肯定存在，$\sigma\to A$，则$\sigma^+\to A^+$）

例 25 C^n中，$X=[x_1，x_2，\cdots，x_n]^T$，$Y=[y_1，y_2，\cdots，y_n]^T$

若规定

$$(X，Y):=\sum_{i=1}^{n}\overline{x_i}\,y_i=X^+Y$$

求证这是C^n的一个内积，故C^n对这样定义的内积来说构成了一个幺正空间.

例 26 在C^n中，规定 $$(X，Y):=\sum_{l=1}^{n}l\overline{x_l}\,y_l$$

求证这也是C^n的一个内积，故C^n对这样定义的内积来说也构成一个欧空间.

同样：同一个复线性空间可定义不同的内积使之成为不同的幺正空间.

以后凡提到幺正空间C^n时，恒指对前一例的内积而言的幺正空间！

例 27 复值$f(x)$，$g(x)\epsilon\varphi[a，b]$，规定：$(f,g)\coloneqq\int_a^b\overline{f(x)}g(x)\,dx$，则满足内积的定义. 故$\varphi[a，b]$可看成一个幺正空间.

度规矩阵

设 V 是 n 维欧空间,$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是一个基,则令$a_{ij}=(\alpha_i，\alpha_j)$，得到一个厄米矩阵.

$$\mathrm{A}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}=\begin{bmatrix}(\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_n)\\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_n)\\ \vdots & \vdots & & \vdots\\ (\alpha_n,\alpha_1) & (\alpha_n,\alpha_2) & \cdots & (\alpha_n,\alpha_n)\end{bmatrix}$$

(7.4.2)

称幺正空间在基$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$下的度规矩阵（或度量矩阵）.

可见：当给定一个基后，由内积就可以完全确定下度规矩阵. 反之，给定一个基下的度规矩阵，也可以完全确定内积. 这是由于：

$$(\alpha，\beta)=\left(\sum_{i=1}^{n}x_i\,\alpha_i，\sum_{i=1}^{n}y_i\,\alpha_i\right)=\sum_{ij}\overline{x_i}(\alpha_i，\alpha_j)y_i=X^+AY$$

(7.4.3)

与线性变换的矩阵类似，度规矩阵依赖于基的选择，可证：设基的过渡矩阵为 C，则 A 的变换为：$\mathrm{A}\rightarrow C^+AC$。

例 28 证明

$$\left\{\alpha_1=[1，1+i，-1]^T，\alpha_2=[1-i，-1，2-i]^T，\alpha_3=[i，2+i，3]^T\right\}$$

是C^3的一个基，并写出幺正空间C^3在该基下的度规矩阵.

向量的模（或长度）

定义（模） 酉空间中，$\sqrt{(\alpha，\alpha)}$称向量α的模或长度，记作$|\alpha|$. 由定义，零向量长度为零；非零向量长度是一正数.

定义（单位向量） 模等于 1 的向量称单位向量.

例 29 酉空间C^n中，$|x| = \sqrt{x^+ x} = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$（7.4.4）

例 30 复φ$[a，b]$中，$|f(x)| = \sqrt{\int_a^b |f\,(x)|^2\,dx}$.

模的性质

定理 7.4.2 酉空间 V 中：

（4） $|k\alpha| = |k||\alpha|$

（5） 柯西不等式：$\left|(\alpha，\beta) \leq |\alpha||\beta|\right|$，等号当且仅当$\alpha$，$\beta$线性相关成立.

（6） 三角不等式：$|\alpha + \beta| \leq |\alpha| + |\beta|$，等号当且仅当$\alpha$，$\beta$线性相关成立.

证：（1）$|k\alpha| = \sqrt{(k\alpha，k\alpha)} = \sqrt{|k|^2(\alpha，\alpha)} = |k||\alpha|$，证毕.

（2)线性相关时，易证等号成立. 下面考虑α，β线性无关，则$\forall \mathrm{k} \in \mathrm{C}$，取$\alpha - k\beta = \gamma \neq 0$，则$|\gamma| > 0$，即$|\gamma|^2 > 0$，而

$|\gamma|^2 = (\alpha - k\beta，\alpha - k\beta) = |\alpha|^2 - k(\alpha，\beta) - \overline{k(\alpha，\beta)} + |k|^2|\beta|^2$，

由于$\forall \mathrm{k} \in \mathrm{C}$，上式都为正，故可取$k = (\beta，\alpha)/(\beta，\beta)$，代入有：

$|\alpha|^2 - \frac{2|(\alpha，\beta)|^2}{|\beta|^2} + \frac{|(\alpha，\beta)|^2}{|\beta|^2}$，证毕.

（3）$|\alpha+\beta|^2=(\alpha+\beta，\alpha+\beta)=|\alpha|^2+(\alpha，\beta)+\overline{(\alpha，\beta)}+|\beta|^2$
$\leq|\alpha|^2+2|(\alpha，\beta)|+|\beta|^2\leq(|\alpha|+|\beta|)^2$，证毕.

例 31 对酉空间C^n应用柯西不等式，有：

$$(X，Y)^2=\left(\sum_{i=1}^{n}\overline{x_i}\,y_i\right)^2\leq\left(\sum_{i=1}^{n}|x_i|^2\right)\left(\sum_{i=1}^{n}|y_i|^2\right)$$

例 32 对酉空间$\varphi[a，b]$应用柯西不等式，有：

$$\left|\int_a^b\overline{f(x)}g(x)\right|\leq\sqrt{\int_a^b|f(x)|^2dx\int_a^b|g(x)|^2dx}$$

酉空间中没有夹角的概念：

虽然柯西不等式：对$\alpha\neq0，\beta\neq0$有 $\dfrac{|(\alpha，\beta)|}{|\alpha||\beta|}$ ，但$(\alpha，\beta)$为复数，故无法引入夹角 $-1\leq\dfrac{(\alpha，\beta)}{|\alpha||\beta|}\leq1$ ，即:一般地，在酉空间中无夹角的概念，但是，仍可引入“正交”的概念.

正交和标准正交基

定义（正交） 若$(\alpha，\beta)=0$，则称α和β正交.

可见：零向量与所有向量都正交.

定义 酉空间 V 中的一个向量组，若所有向量均非零，且两两正交，则称该向量组为一个正交向量组；进一步，若正交向量组中每一个向量都是单位向量，则称该向量组为一个标准正交向量组.

正交向量组的性质：

定理 7.4.3 设$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是一个正交向量组，则

$\alpha_1，\alpha_2，\cdots，\alpha_n$线性无关.

即：正交向量组中的向量必线性无关.

证：设$k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n=0$

则：$(\alpha_i，k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n)=(\alpha_i，0)=0$

即：$k_i(\alpha_i，\alpha_i)=0$，故$k_i=0(i=1，2，\cdots，n)$，证毕.

推论 7.4.4 n 维酉空间 V 中正交向量组所含向量个数不大于 n.

由定理 7.4.3，在 n 维酉空间 V 中若一个正交向量组正好含有 n 个向量，则这个正交组是 V 的一个基。引入如下定义：

定义 由 V 中相互正交的向量构成的基，承做 V 的正交基；若正交基中每个向量均是单位向量，则称该基是 V 的标准正交基.

$\{\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_n\}$：$(\varepsilon_i，\varepsilon_j)=\delta_{ij}$

若已知 n 维酉空间 V 的一个标准正交基$\{\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_n\}$，则 V 在该基下的度规矩阵，V 中向量的坐标，V 中任意两个向量的内积，V 中向量长度的坐标表达式均变得特别简单：

度规 A：$a_{ij}=(\varepsilon_i，\varepsilon_j)=\delta_{ij}$, 故$A=I_n$.

向量α的坐标：$x_i=(\varepsilon_i，\alpha)$，　$\alpha=\sum_i x_i\,\varepsilon_i=\sum_i(\varepsilon_i，\alpha)\,\varepsilon_i$

向量$\alpha=[\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_n]X$，$\beta=[\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_n]Y$

则$(\alpha，\beta)=X^+AY=X^+I_nY=X^+Y=\overline{x_1}y_1+\overline{x_2}y_2+\cdots+\overline{x_n}y_n$,

$|\alpha|^2=(\alpha，\alpha)=X^+X=|x_1|^2+|x_2|^2+\cdots+|x_n|^2$.

（可数）无限维线性空间的两个例子：

例 33 $[-\pi，\pi]$上定义的函数集合：$\left\{\frac{1}{\sqrt{2\pi}}e^{ikx}: k=0，\pm 1，\pm 2，\cdots\right\}$在内积$(f，g)=\int_{-\pi}^{\pi}\overline{f(x)}g(x)\,dx$之下是标准正交的. 另一方面，所有满足$f(-\pi)=f(\pi)$的$[-\pi，\pi]$上的连续函数均可用上述函数组一致逼近（希尔伯特《数理方法》卷一P_{53}），故上述函数组是$[-\pi，\pi]$上连续的2π为周期的函数构成的酉空间上的一个标准正交基.（无穷维）

例 34 $[-1，1]$上定义的正数集合$\{x^n：n=0，1，2，\cdots，+\infty\}$

维尔斯特拉斯逼近定理（希尔伯特《数理方法》卷一P_{51}）知：其构成了$[-1，1]$上实连续函数构成的欧空间上的一个基，但非标准正交！按施密特正交化方案可构成正交基，即勒让德多项式！

Schmidt 标准正交化方法与欧空间完全相同：

定理 7.4.4 n 维酉空间 V 中，从任意一组基$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$出发，可以构造一组标准正交基$\{\gamma_1，\gamma_2，\cdots，\gamma_n\}$使$\gamma_i$为$\alpha_1，\alpha_2，\cdots，\alpha_n$的线性组合.

证：构造

$$\left\{\begin{aligned} \beta_1 &= \alpha_1 \\ \beta_2 &= \alpha_2-\frac{(\alpha_2，\beta_1)}{(\beta_1，\beta_1)}\beta_1 \\ \beta_3 &= \alpha_3-\frac{(\alpha_3，\beta_1)}{(\beta_1，\beta_1)}\beta_1-\frac{(\alpha_3，\beta_2)}{(\beta_2，\beta_2)}\beta_2 \\ &\vdots \\ \beta_n &= \alpha_n-\sum_{i=1}^{\mathrm{n}-1}\frac{(\alpha_n，\beta_i)}{(\beta_i，\beta_i)}\beta_i \end{aligned}\right. \tag{7.4.5}$$

由归纳法可证所有的β_i均为零（$\alpha_1，\alpha_2，\cdots，\alpha_n$线性无关得到），并且每个$\beta_i$均为$\beta_1$到$\beta_{i-1}$正交，故$\{\beta_1，\beta_2，\cdots，\beta_n\}$是一个正交基，进

一步，构造：$\gamma_1=\beta_1/|\beta_1|$，⋯，$\gamma_n=\beta_n/|\beta_n|$，则$\{\gamma_1，\gamma_2，\ \cdots，\gamma_n\}$即一个标准正交基！

例 35　求证：在酉空间C^n中，n 个列向量$x_1，x_2，\cdots，x_n$线性无关的充要条件是：

$$g(x_1，x_2，\ \cdots，x_n)=\begin{vmatrix}(x_1，x_1) & (x_1，x_2) & \cdots & (x_1，x_n)\\(x_2，x_1) & (x_2，x_2) & \cdots & (x_2，x_n)\\ \vdots & \vdots & \cdots & \vdots\\(x_n，x_1) & (x_n，x_2) & \cdots & (x_n，x_n)\end{vmatrix}\neq 0$$

证：令$A_{n\times n}=[x_1，x_2，\ \cdots，x_n]$，注意到$(x_i，x_j)=x_i{}^+x_j$

故上述行列式$=|A^+A|=\overline{|A|}|A|\geq 0.$

例 36 求证：酉空间中，n 个向量$\alpha_1，\alpha_2，\cdots，\alpha_n$线性无关的充要条件是 Gram 行列式$g(\alpha_1，\alpha_2，\cdots，\alpha_n)\neq 0.$

提示：选取标准正交基，将$g(\alpha_1，\alpha_2，\ \cdots，\alpha_n)$用$\alpha_1，\alpha_2，\ \cdots，\alpha_n$的坐标表示，并用上例的结论！

幺正矩阵（群）

幺正矩阵和幺正变换是物理学中极其重要的矩阵和算符.

定义（幺正矩阵）　满足$A^+A=I_n$的 n 阶矩阵 A 称为幺正矩阵.

注：可见实的幺正矩阵就是正交矩阵.

幺正矩阵的性质：

（1）$|A|=e^{i\theta}$，可逆！　　　　（2）$A^{-1}=A^+$

（3）$AA^+=A^+A=I$　　　　（4）A^+（即A^{-1}）也是幺正矩阵

定理 7.4.5　若A，B均为 n 阶幺正矩阵，则AB亦是.

证：$(AB)^+(AB)=B^+A^+AB=B^+B=I$（物理上有用！）

由于上述幺正矩阵的性质，以及定理 7.4.5，可得到如下重要结论：

定理 7.4.6 所有 n 阶幺正矩阵的集合，对于矩阵的乘法构成一个群，称幺正矩阵群，记作$\mathrm{U}(n)$.

证：封闭性，结合律，单位元（I_n），逆元。

幺正矩阵 A，若$|A| = 1$，则称 A 为幺模幺正矩阵. 显然有：

定理 7.4.6’ 所有 n 阶幺模幺正矩阵的集合，对于矩阵的乘法构成群，称幺模幺正矩阵群，记作$\mathrm{SU}(n)$.

$\mathrm{U}(n)$，$\mathrm{SU}(n)$群都是$GL(n，G)$群的子群，它们对应的线性变换群后面再讨论. 正交矩阵与标准正交基的关系：

定理 7.4.7 n 阶复矩阵 A 是幺正矩阵的充要条件是：A 的 n 个行（列）向量构成C^n的标准正交矩阵基.

证明：将 A 按行分组：$\mathrm{A} = [A_1，A_2，\cdots，A_n]$，则：

$A^+A = \begin{bmatrix} (A_1，A_1) & \cdots & (A_1，A_n) \\ \vdots & \vdots & \vdots \\ (A_n，A_1) & \cdots & (A_n，A_n) \end{bmatrix}$列向量得证！同理，按列分组，由$AA^T$可证行向量。

定理 7.4.8 设$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是 n 维酉空间的一个标准正交基，P 是一个 n 阶复矩阵，$[\beta_1，\beta_2，\cdots，\beta_n] = [\alpha_1，\alpha_2，\cdots，\alpha_n]P$，则：$\{\beta_1，\beta_2，\cdots，\beta_n\}$是 V 的标准正交基的充要条件是 P 为一幺正矩阵.

证：考虑

$$(\beta_i，\beta_j) = \left(\sum_s \alpha_s P_{si}，\sum_t \alpha_\mathrm{t} P_{tj}\right) = \sum_{s,t} \overline{P_{s\iota}} P_{tj}(\alpha_s，\alpha_\mathrm{t})$$

$$= \sum_{s,t} \overline{P_{s\iota}} P_{tj}\delta_{st} = \sum_{s,t} \overline{P_{s\iota}} P_{tj} = (P^+P)_{ij}$$

幺正变换（群）：

类似正交变换，通过要求线性变换不改变向量的长度引入：

定义（幺正变换） 设σ是酉空间 V 中的线性变换，若$\forall\alpha\epsilon V$，有：

$|\sigma(\alpha)|=|\alpha|$，则称σ为一个幺正变换.

关于幺正变换与幺正矩阵、标准正交基的关系，有如下定理

定理 7.4.9 设σ是酉空间 V 中的线性变换，则如下论述等价：

（1）σ是幺正矩阵.

（2）$\forall\alpha，\beta\epsilon V$，有$\left(\sigma(\alpha)，\sigma(\beta)\right)=(\alpha，\beta)$.

（3）σ将 V 的一个标准正交基变成 V 的另一个标准正交基.

（4）σ关于 V 的任意一个标准正交基的矩阵是幺正矩阵.

注意：幺正变换保持长度、正交性、距离等度量不变！

证：循环证明：

（1）$\Longrightarrow$（2）考虑：$\left(\sigma(\alpha+\beta)，\sigma(\alpha+\beta)\right)=(\alpha+\beta，\alpha+\beta)$

（2）$\Longrightarrow$（3）考虑：$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$，$\left(\sigma(\alpha_i)，\sigma(\alpha_j)\right)=(\alpha_i，\alpha_j)=\delta_{ij}$

(3) $\Longrightarrow$(4) $[\sigma(\alpha_i)，\cdots，\sigma(\alpha_n)]=[\alpha_1，\alpha_2，\cdots，\alpha_n]P$，由定理 7.4.8 即得.

（4)$\Longrightarrow$(1)设$\{\alpha_1，\alpha_2，\cdots，\alpha_n\}$是标准正交基，σ的矩阵为 P，则$\forall\alpha\epsilon V$，设$\alpha$坐标为$x$，则$\sigma(\alpha)$坐标为 $\mathrm{P}x$，注意此时度规为I_n，

故$\left(\sigma(\alpha)，\sigma(\alpha)\right)=(\mathrm{P}x)^{+}\mathrm{P}x=x^{+}x=(\alpha，\alpha)$，证毕.

由于选定了标准正交基之后，幺正变换和幺正矩阵一一对应，且该对应保持乘法关系不变，故平行地有：

定理 7.4.10 n 维酉空间上的正交变换全体，对线性变换的乘法来说构成一个群，称作幺正变换群.

由于 n 维幺正变换群与 n 维幺正变换群是同构。因此 n 维幺正变换群也用同样的标记$\mathrm{U}(n)$.

标准正交基、幺正矩阵、幺正变换三者关系：

标准正交基$\{\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_n\}$：$(\varepsilon_i，\varepsilon_j)=\delta_{ij}$

正交矩阵 A：$A^{+}A=AA^{+}=I$

正交变换α：保持长度（及内积）不变

标准正交基与幺正矩阵：

A_n是幺正矩阵$\Leftrightarrow A_n$的行（列）向量组是C^n的标准正交基.

幺正矩阵与幺正变换：

σ是幺正变换$\xLeftrightarrow{\text{通过标准正交基联系}}\sigma$在标准正交基下的矩阵是幺正矩阵.

幺正变换与标准正交基

σ是幺正变换$\Leftrightarrow\sigma$将标准正交基变为另一个标准正交基.

$\mathrm{SO}(n)$变换是 n 维实空间R^n的转动变换，类似地，$\mathrm{SU}(n)$可“看成”n 维复空间C^n的“转动”变换。

厄米矩阵、厄米变换

除幺正变换外，欧空间中另一类重要的变换是厄米变换（连续的幺正变换的生成元即为厄米变换）. 为此我们引入：

定义（厄米矩阵） 设 A 是一个 n 阶方阵，若$A=A^{+}$，则称 A 为一个厄米矩阵.

注：厄米矩阵即厄米共轭矩阵，实的厄米矩阵即对称矩阵.

注意：厄米矩阵不一定可逆，但其转置于复共轭矩阵仍然厄米：

即：若$A=A^{+}$，则$A^{T}=\bar{A}$亦是厄米矩阵：$(A^{T})^{+}=\overline{(A^{T})^{T}}=\bar{A}=A^{T}$。

定义（厄米变换） σ是酉空间 V 中的线性变换，若σ满足：$\forall\alpha，\beta\epsilon V$ $\left(\sigma(\alpha)，\beta\right)=\left(\alpha，\sigma(\beta)\right)$，则称σ是 V 上的一个厄米变换.

例 37 厄米矩阵的例子：$\begin{bmatrix}2 & 1+i\\1-i & 1\end{bmatrix}$，$\begin{bmatrix}1 & 1+i & -i\\1-i & 2 & i\\i & -i & 0\end{bmatrix}$, 一般形式：$\begin{bmatrix}a_{11} & a_{12} & a_{13} & \cdots\\\overline{a_{12}} & a_{22} & a_{23} & \cdots\\\overline{a_{13}} & \overline{a_{23}} & a_{33} & \cdots\\\vdots & \vdots & \vdots & \ddots\end{bmatrix}$，其中$a_{ij}\in R$.

厄米算符的两个例子：（注意区分“复值函数”和“复变函数”）

例 38 $[a，b]$上的复值连续函数的集合在$(f，g):=\int_{\mathrm{a}}^{b}\overline{f(x)}g(x)\,dx$之下构成一个酉空间，试判断$i\frac{d}{dx}$是否是一个厄米算符？

解：$i\frac{d}{dx}$易证是线性算符，间证$i\frac{d}{dx}=p$, 则

$$\begin{aligned}(pf，g)&=\int_{a}^{b}\overline{\left(i\frac{df}{dx}\right)}g(x)\,dx=-i\int_{a}^{b}\overline{\left(\frac{df}{dx}\right)}g(x)\,dx\\&=\int_{a}^{b}\overline{f(x)}i\frac{d}{dx}g(x)-i\int_{a}^{b}\frac{d}{dx}\left(\overline{f(x)}g(x)\right)dx\\&=(f，pg)-i\left[\overline{f(x)}g(x)\right]\Big|_{a}^{b}.\end{aligned}$$

故在$[a，b]$上满足$f(a)=f(b)$的复值函数构成的酉空间上是厄米的！

（量子力学中周期性边界条件与动量算符的厄米性）

对算符$\left(i\frac{d}{dx}\right)^{2}=-\left(\frac{d}{dx}\right)^{2}$亦有同样的结论.

例 39 $-\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right]+q(x)y=\lambda p(x)y$的方程称施图姆–刘维尔型（S–L 型）方程，对$p(x)\geq 0$且不恒为零，可用$\sqrt{p}y\to U$变成：

$-\frac{d}{dx}\left[\phi(x)\frac{dU}{dx}\right]+\psi(x)U=\lambda U$ 的形式. 记：$L:=-\frac{d}{dx}\left[\phi(x)\frac{dU}{dx}\right]+\psi(x)$

则可写成$LU(x)=\lambda U(x)$, 试问$[a，b]$上的复值连续函数构成的酉空间 V 上，L 仍是厄米的？

解：$(U_1，LU_2)-(LU_1，U_2)=-\int\frac{d}{dx}\left[\phi(x)\left(\overline{U_1}\frac{dU_2}{dx}-\frac{d\overline{U_1}}{dx}U_2\right)\right]dx$

故在边界条件$\left[\phi(x)\left(\overline{U_1}\frac{dU_2}{dx}-\frac{d\overline{U_1}}{dx}U_2\right)\right]\Big|_a^b=0$之下，L 是厄米变换（算符）.（数理方法：课程的主要内容）

定理 7.4.11 酉空间 V 上的线性变换σ是厄米变换的充要条件是:σ关于 V 的标准正交基的矩阵是厄米矩阵.

证明：(⟹)：$(\sigma(\alpha_i)，\alpha_j)=(\alpha_i，\sigma(\alpha_j))$，故 $\left(\sum_k\alpha_kA_{ki}，\alpha_j\right)=\left(\alpha_i，\sum_l\alpha_lA_{lj}\right)$

即：$\overline{A_{ji}}=A_{ij}$，故$A=A^+$

(⟸)：取 $\alpha=\sum_i\alpha_ix_i$ ， $\beta=\sum_j\alpha_jy_j$

故 $(\sigma(\alpha)，\beta)=\sum_{i，j}\overline{x_i}(\sigma(\alpha_i)，\alpha_j)y_j=\sum_{i，j}\overline{x_i}\,\overline{A_{ji}}y_j$

$$(\alpha，\sigma(\beta))=\sum_{i，j}\overline{x_i}(\alpha_i，\sigma(\alpha_j))y_j=\sum_{i，j}\overline{x_i}A_{ij}y_j$$

而$A=A^+$，故$(\sigma(\alpha)，\beta)=(\alpha，\sigma(\beta))$，证毕.

类似对称矩阵和对称变换，厄米矩阵和厄米变换的集合也不构成群！

（1）$(AB)^+=B^+A^+=BA\neq AB$

（2）$(\sigma\circ\tau(\alpha)，\beta)=(\sigma(\tau(\alpha))，\beta)=(\tau(\alpha)，\sigma(\beta))$

$$=\left(\alpha，\tau\left(\sigma(\beta)\right)\right)=\left(\alpha，\tau\circ\sigma(\beta)\right)\neq\left(\alpha，\sigma\circ\tau(\beta)\right)$$

平行于对称矩阵和对称变换，对厄米矩阵和厄米变换也有如下重要性质：1.厄米变换的本征值一定为实数；

2.厄米变换的所有本征向量可选为两两正交的单位向量；

厄米变换（厄米矩阵）一定可以对角化.（本征向量组完备）

由于选定标准正交基之后，厄米变换和厄米矩阵一一对应，为此下面只考虑厄米矩阵：

定理 7.4.12　厄米矩阵的本征值全是实数.

证：设$Ax=\lambda x$　由于厄米性：$A=A^{+}:=\overline{A^{T}}$

故$(Ax)^{+}x=(x^{+}A^{+})x=x^{+}(Ax)$，即：$(\lambda x)^{+}x=x^{+}(\lambda x)$

即：$\bar{\bar{\lambda}}(x^{+}x)=\lambda(x^{+}x)$，而$x^{+}x>0$，故$\bar{\bar{\lambda}}=\lambda$.

注意:虽然厄米矩阵的本征值都是实数，故本征向量不一定为实 n 元向量，一般仍然是C^{n}中的复 n 元向量：$(\lambda I-A)=0$.

推论:厄米矩阵 A 的行列式为实数，即$|A|$为实数.

定理 7.4.13　厄米矩阵的属于不同本征值的本征向量相互正交.

证：设$\mu\neq\lambda$是两本征值，$Ax=\mu x$，$Ay=\lambda y$,

　　考虑$(Ax)^{+}y=x^{+}A^{+}y=x^{+}Ay=x^{+}(Ay)$

　　故$(\mu x)^{+}y=x^{+}(\lambda y)$，即：$(\mu-\lambda)(x^{+}y)=0$

　　而$\mu\neq\lambda$，故$x^{+}y=\left(x，y\right)=0$，正交.

定理 7.4.14 厄米矩阵必可通过幺正矩阵对角化.即：设 A 为厄米矩阵，则必存在一幺正矩阵 C，使$C^{-1}AC=C^{+}AC$为对角矩阵.

证：用数归.当A_{n}的阶$\mathrm{n}=1$时显然成立.

假设$n-1$阶时成立，考虑n阶：A至少有一个本征值和单位本征向量，设为λ_1和P_1：$AP_1=\lambda_1 P_1$，$|P_1|=1$.

则通过逐步添加，可从P_1构造C^n的一个基：$\{P_1, q_2, \cdots, q_n\}$，用 Schmidt 标准正交化方法可得$C^n$的一个标准正交基$\{P_1, P_2, \cdots, P_n\}$，令$P=[P_1, P_2, \cdots, P_n]$，$P$为幺正矩阵，故

$AP=[AP_1, AP_2, \cdots, AP_n]=[P_1, P_2, \cdots, P_n]B=PB$

其中$B=\begin{bmatrix}\lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & *\end{bmatrix}$，而$B=P^{-1}AP=P^{+}AP$，故 B 亦为厄米矩阵，

故$B=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & *\end{bmatrix}=\begin{bmatrix}\lambda_1 & 0 \\ 0 & B_{n-1}\end{bmatrix}$，$B_{n-1}$为$n-1$阶厄米矩阵，

故$B_{n-1}={Q_{n-1}}^{-1}F_{n-1}Q_{n-1}={Q_{n-1}}^{+}F_{n-1}Q_{n-1}$，令$Q=\begin{bmatrix}1 & 0 \\ 0 & Q_{n-1}\end{bmatrix}$，则$Q^{-1}=Q^{+}$，且$Q^{-1}BQ=\begin{bmatrix}\lambda_1 & 0 \\ 0 & {Q_{n-1}}^{-1}B_{n-1}Q_{n-1}\end{bmatrix}=F_n$，故令$\mathrm{C}=\mathrm{PQ}$即得证！

推论 7.4.15 厄米矩阵的任意一个本征值λ的重数S_λ，与该本征值λ对应的本征子空间E_λ的维数$\dim E_\lambda$相等.

即：$S_\lambda=\dim E_\lambda$

注：本征值是 n 重根，必然有 n 个线性无关的本征向量，故全部本征向量构成的向量组是完备的，可作为C^n的一个基！

用幺正矩阵对角化厄米矩阵（即对厄米矩阵 A，求一幺正矩阵 P，使$P^{-1}AP=P^{+}AP=diag[\lambda_1, \lambda_2, \cdots, \lambda_n]$）的步骤如下：

（1）由$|\lambda I-A|=0$，求出A_n的所有不同本征值λ_1，λ_2，$\cdots$，λ_n，以及它们的重数S_1，S_2，$\cdots$，S_n，显然$\lambda_i\in R$，且 $\sum_{i=1}^{\mathrm{m}} S_i=n$

（2）对每一个λ_i，由$|\lambda_i I-A|x=0$，求其基础解系α_{i1}，α_{i2}，$\cdots$，α_{iS_i}，再用 Schmidt 方法将其标准正交化为P_{i1}，P_{i2}，$\cdots$，P_{iS_i}.

（3）将求出的 n 个相互正交的单位本征向量排成矩阵： $P=$

$[P_{11}，P_{12}，\cdots，P_{1S_1}，P_{21}，P_{22}，\cdots，P_{2S_2}，\cdots，P_{m1}，P_{m2}，\cdots，P_{mS_m}]$

则$P^{+}=P^{-1}$，且

$P^{-1}AP=P^{+}AP$

$=diag[\lambda_1，\lambda_1，\cdots，\lambda_1，\lambda_2，\lambda_2，\cdots，\lambda_2，\cdots，\lambda_m，\lambda_m，\cdots，\lambda_m]$

例 40 对前面定义的 S–L 型方程中的线性变换（算符）L：

$L\coloneqq-\frac{d}{dx}\left[\phi(x)\frac{dU}{dx}\right]+\psi(x)$，在前述边界条件下是厄米的，不仅如此，若$\phi(x)$，$\psi(x)$均为恒正，则 L 的本征值$\lambda_n\geq 0$.

证：设$LU(x)=\lambda_n U(x)$，则：

$(U_n，LU_n)=\lambda_n(U_n，U_n)=\int_a^b\phi(x)\overline{U_n{}'}\,U_n{}'dx+\int_a^b\psi(x)\overline{U_n}\,U_n dx$

故$\lambda_n\geq 0$.

一个有用的例子是$\phi(x)=1$，$\psi(x)=V(x)\geq 0$，则$L=-\left(\frac{d}{dx}\right)^2+V(x)$为能量算符.

无穷维酉空间中，厄米变换（厄米算符）的本征向量组的完备性在某些情况下亦成立，例如：

例 41 $\{e^{ikx}: k=0,\pm1,\pm2,\cdots\}$显然是$i\frac{d}{dx}$（以及$-\left(\frac{d}{dx}\right)^2$）的本征值向量组. 由前面的讨论知该向量组是$[-\pi，\pi]$上周期为$2\pi$的函数构成的酉空间的一个基，显然是完备的.（可数无穷维）

例 42 （不可数无穷维的一个例子，本征值为连续谱）$\{e^{ikx}: k\in R\}$是$(-\infty，+\infty)$上平方可积函数构成的酉空间$L^2(-\infty，+\infty)$上的一个基（傅里叶变换），可证$i\frac{d}{dx}$亦是$L^2(-\infty，+\infty)$上的厄米算符。上述基正的就是$i\frac{d}{dx}$的本征向量组，即：其本征向量组完备，

$f(x)=\int_{-\infty}^{+\infty}\tilde{f}(k)e^{ikx}dk.$

下一章最后一节将给出：无穷维酉空间中，某特定情况下厄米变换的本征向量组的完备性之证明.

例 43 $A=\begin{bmatrix}0 & -i & 1\\ -i & 0 & 0\\ 1 & 0 & 0\end{bmatrix}, B=\begin{bmatrix}2 & i & 0\\ -i & 1 & i\\ 0 & -i & 2\end{bmatrix}$, 用幺正矩阵分别将 A 和 B 相似对角化.

例 44 两个 n 阶厄米矩阵 A，B可对易（即：$AB=BA$）的充要条件是 A 和 B 可具有相同的完备本征向量组（n 个，线性无关），或者说，A 和 B 可同时对角化（用同一个相似变换矩阵对角化）

证：(充分性 $\Leftarrow$)：两种证法，

方法一：设$P^{-1}AP=F_1$，$P^{-1}BP=F_2$，

则$\mathrm{AB}-\mathrm{BA}=P(F_1F_2-F_2F_1)P^{-1}$.

方法二：设A，B有共同的线性无关本征向量组$\{P_1，P_2，\cdots，P_n\}$，对应的本征值分别是λ_i和μ_i，$(i=1，2，\cdots，n)$,

则$(AB-BA)P_i=(\lambda_i\mu_i-\mu_i\lambda_i)P_i=0$.

$\forall \mathrm{x}\in R^n$，有 $x=\sum_{i=1}^{n}k_iP_i$ ，故$(AB-BA)x=0$.

由于x任意，故$AB-BA=0$.（方法二可推广到对称变换）

（必要性$\Longrightarrow$）：假设$AB=BA$，由于 A，B 均是实对称矩阵，可对角化，故设 A 的一组线性无关本征向量组为$\{P_1，P_2，\cdots，P_n\}$对应本征值$\lambda_1，\lambda_2，\cdots，\lambda_n$.

（1）若λ_i非 A 的重本征值，则

$A(BP_i)=(AB)P_i=(BA)P_i=B(AP_i)=B(\lambda_i P_i)=\lambda_i(BP_i)$

故$BP_i \epsilon E_{\lambda_i}$，即：$BP_i=\mu_i P_i$，$P_i$亦是 B 的本征向量.

(2)若λ_i是 A 的重本征值，设重数为S_i，则$\{P_1，P_2，\cdots，P_n\}$中对应λ_i的本征向量必有S_i个，重新命名为$\{q_{i1}，q_{i2}，\cdots，q_{is}\}$，化成$E_{\lambda_i}$，则同样由$ABq_{ij}=BAq_{ij}=\lambda_i Bq_{ij}$可知：$Bq_{i1}，Bq_{i2}，\cdots，Bq_{is}\in E_{\lambda_i}$，

但一般地，$Bq_{il}\neq \mu_i q_{il}$，选取 $\widetilde{q_\iota}=\sum_{l=1}^{S_i} x_l\, q_{il}$ ，使$B\widetilde{q_\iota}=\mu\widetilde{q_\iota}$，则与$q_{im}$作内积有：

$(q_{im}，B\widetilde{q_\iota})=\mu(q_{im}，\widetilde{q_\iota})$，即： $\sum_{l=1}^{S_i}(q_{im}{}^T Bq_{il})\,x_l=\sum_{l=1}^{S_i}\mu\delta_{ml}\,x_l$

其中$q_{im}{}^T Bq_{il}$记作C_{ml}，

$B=B^T\Longrightarrow C_{ml}=C_{lm}$，排成$S_i\times S_i$矩阵 C，$C^T=C$，$x=[x_1，\cdots，x_{S_i}]^T$，方程化为$Cx=\mu x\Longrightarrow(\mu I-C)x=0$，必有$S_i$个线性无关的$x$，故必有$S_i$个线性无关的$\widetilde{q_\iota}$，使其同时是 A，B 的本征向量，证毕！

例 45 证明：幺正矩阵本征值的模为 1.

证：设$A^+A=I$，$Ax=\lambda x$，则考虑上式得厄米共轭 $x^+A^+=\bar{\lambda}x^+$，故$x^+A^+Ax=\bar{\lambda}x^+\lambda x$，即：$x^+x=(\bar{\lambda}\lambda)x^+x$，由于$x$不是零向量，故

$x^+x=|x_1|^2+|x_2|^2+\cdots+|x_n|^2>0$，故$\bar{\lambda}\lambda=1$，即：$|\lambda|=1$. 由于幺正变换不改变向量的长度，故上述结论很容易理解.

例 46 求证：n 阶幺正矩阵 U 一定可以通过某幺正变换 P 相似对角化，即存在$P^+P=1$，使$P^{-1}\mathrm{U}P$是对角矩阵.（注意，正交矩阵无此性质）

证明：完全类似于厄米矩阵的证明. 用数归：

当U_n的阶$\mathrm{n}=1$时显然成立.

假设$\mathrm{n}-1$阶时成立，考虑 n 阶：U 至少有一个本征值和单位本征向量，设为λ_1和P_1：$UP_1=\lambda_1P_1$，$|P_1|=1$.

则通过逐步添加，可从P_1构造 C^n的一个基：$\{P_1, q_2, \cdots, q_n\}$，用 Schmidt 标准正交化方法可得$C^n$的一个标准正交基$\{P_1, P_2, \cdots, P_n\}$，令$P=[P_1, P_2, \cdots, P_n]$，$P$为幺正矩阵，故

$UP=[UP_1, UP_2, \cdots, UP_n]=[P_1, P_2, \cdots, P_n]B=PB$

其中$B=\begin{bmatrix}\lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & *\end{bmatrix}$，而$B=P^{-1}AP=P^+AP$，故 B 亦为幺正矩阵，

故$B=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & *\end{bmatrix}=\begin{bmatrix}\lambda_1 & 0 \\ 0 & B_{n-1}\end{bmatrix}$，$B_{n-1}$为$\mathrm{n}-1$阶幺正矩阵，

故$B_{n-1}={Q_{n-1}}^{-1}F_{n-1}Q_{n-1}={Q_{n-1}}^{+}F_{n-1}Q_{n-1}$，令$Q=\begin{bmatrix}1 & 0 \\ 0 & Q_{n-1}\end{bmatrix}$，则$Q^{-1}=Q^+$，且$Q^{-1}BQ=\begin{bmatrix}\lambda_1 & 0 \\ 0 & {Q_{n-1}}^{-1}B_{n-1}Q_{n-1}\end{bmatrix}=F_n$，故令$\mathrm{C}=\mathrm{PQ}$即得证！

例 47 （厄米矩阵与幺正矩阵之联系）设 A 是厄米矩阵，求证e^{iA}为幺正矩阵；设 U 是幺正矩阵，求证存在厄米矩阵 A，使得$U=e^{iA}$.

证：（1）设$A^{+}=A$，则：$A=P^{-1}FP$ （$P^{+}=P$，F实对角）

故$e^{iA}=exp(iP^{-1}FP)=P^{-1}e^{iF}P$，

$\left(e^{iA}\right)^{+}\left(e^{iA}\right)=P^{-1}e^{-iF}PP^{-1}e^{iF}P=P^{-1}e^{-iF}e^{iF}P=1$，故$e^{iA}$幺正。

（2）设 U 幺正，则由前两个例题：$U=P^{-1}FP$，（$P^{+}=P$，$|\lambda_i|=1$）

故$F=\begin{bmatrix} e^{i\lambda_1} & & \\ & e^{i\lambda_2} & \\ & & \ddots \end{bmatrix}=e^{iF'}$，$F'=\begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{bmatrix}\in R^{n\times n}$，故

$U=P^{-1}FP=P^{-1}e^{iF'}P=e^{iP^{-1}F'P}=e^{iA}(A=P^{-1}F'P)$，而

$A^{+}=P^{+}{F'}^{+}P=P^{-1}F'P=A$厄米.

第八章 二次型和厄米型

二次型的定义和标准型

二次型的规范型和惯性定理

二次型的正定性

厄米型

本征值问题与极值问题的等价性

§8.1 二次型的定义和标准形

实、复二次型　　二次型的矩阵——对称矩阵　　　线性替换

矩阵的合同与二次型的等价　　　二次型的标准型

二次型的定义

定义（二次型） 数域F上n个变量$x_1, x_2, \cdots, x_n$的不含一次项与常数项的二次多项式称二次齐次多项式，或简称为二次型.

即

$$f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{1n}x_1x_n +$$

$$a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n + \cdots +$$

$$2a_{n-1n-1}x_{n-1}^2 + 2a_{n-1n}x_{n-1}x_n + a_{nn}x_n^2$$

$$= \sum_{i=1}^{n} \sum_{j=i+1}^{n} 2\, x_i a_{ij} x_j + \sum_{i=1}^{n} a_{ii} {x_i}^2$$

（其中$a_{ij} \epsilon F$）

若$F = R$，则称实二次型，若$F = C$，则称复二次型.

由于负数域上遇到的一般是厄米型，复二次型的用处不大，故本章只讨论实二次型.复二次型的结论只在适当的地方顺带指出，不作专门讨论.从此处之后，除非特别指明，否则提到的二次型都是实二次型.

例 1　$f(x_1, x_2, x_3) = x_1^2 + x_1x_2 + x_1x_3 + x_2^2 + x_2x_3 + x_3^2$

二次型的矩阵和矩阵表示：

$$f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots + a_{1n}x_1x_n$$

$$+a_{21}x_2x_1 + a_{22}x_2^2 + a_{23}x_2x_3 + \cdots + a_{2n}x_2x_n$$

$$+a_{31}x_3x_1 + a_{32}x_3x_2 + a_{33}x_3^2 + \cdots + a_{3n}x_3x_n$$

$$+\cdots$$

$$+a_{n1}x_nx_1+a_{n2}x_nx_2+a_{n3}x_nx_3+\cdots+a_{nn}x_n^2$$

$$=\sum_{\mathrm{i}=1}^{\mathrm{n}}\sum_{j=1}^{n}a_{ij}x_ix_j \qquad （其中 a_{ij}=a_{ji}）$$

将上述a_{ij}排成矩阵$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$，则$A=A^T$，称为二次型$f(x_1,x_2,\cdots,x_n)$的矩阵。若引入$X=[x_1,x_2,\cdots,x_n]^T$，则二次型可写成：$f(x_1,x_2,\cdots,x_n)=f(\mathrm{X})=X^TAX.$

二次型与对称矩阵的一一对应：

任意矩阵$\frac{1}{2}(B+B^T)+\frac{1}{2}(B-B^T)=C+D \quad C=C^T$，$D=-D^T$

故$X^TBX=X^TCX+X^TDX=X^TCX$，只有对称部分有贡献.

即：将二次齐次多项式写成矩阵形式时，若不限定矩阵的性质，则得到的矩阵并不唯一，可任意补上一个反对称的矩阵.但是，若按上面的方法限定引入的矩阵是对称矩阵时，这个对称矩阵是唯一的：

定理 8.1.1 设A，B是对称矩阵，若对任意$\mathrm{X}\epsilon R^n$都有$X^TAX=X^TBX$，则$\mathrm{A}=\mathrm{B}$.

证明：将左右两边展开成多项式，比较对应项的系数即得！

上页的分析表明，每个二次型可确定一个对称矩阵.上述定理又保证这个对称矩阵是唯一的，即每个二次型唯一确定了一个对称矩阵。反之，给定一个对称矩阵，由X^TAX也唯一确定了一个二次型.。故：

结论：二次型与对称矩阵一一对应！

例 2 欧空间选定一组基，有内积$(\alpha，\alpha)$即可确定一个二次型X^TAX.

例 3 欧空间中标准正交基下对称变换的期望值$(\alpha,\sigma(\alpha))$确定了一个二次型.

例 4 求下列二次型的矩阵形式：

$$(1)\ f(x_1,x_2,x_3)=3x_1^2+2x_2^2-x_3^2+2x_1x_2-4x_2x_3$$

$$(2)\ f(x_1,x_2,x_3)=2x_1x_2-2x_2x_3+2x_1x_3$$

$$(3)\ f(x_1,x_2,x_3)=x_1^2+3x_2^2-5x_3^2$$

线性替换：

定义（线性替换） 数域R上的两组变量$x_1,x_2,\cdots,x_n$和$y_1,y_2,\cdots,y_n$之间的关系：

$$\begin{cases} x_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n \\ \cdots \\ x_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n \end{cases} (c_{ij}\epsilon R)$$

称数域R上由变量$x_1,x_2,\cdots,x_n$到变量$y_1,y_2,\cdots,y_n$的一个线性变换.

进一步，若系数组成的矩阵$\mathrm{C}=(c_{ij})_{m\times n}$是可逆的，或称非退化的.

若引入向量$\mathrm{X}=[x_1,x_2,\cdots,x_n]^T.\mathrm{Y}=[y_1,y_2,\cdots,y_n]^T$，则可写成矩阵形式：

$$X=CY$$

引入$\mathrm{X}=\mathrm{CY}$的线性替换后，二次型X^TAX可用Y表示为：

$$X^TAX=(CY)^TA(CY)=Y^T(C^TAC)Y$$

合同矩阵和等价二次型：

定义（合同） 设A，B是R上的n阶方阵，若存在R上可逆的C，使：$C^TAC=B$，则称A与B合同，A和B是合同矩阵.

注意：（1）区分矩阵的等价、相似、合同！

（2）由于C可逆，故合同是一个等价关系.

（3）合同矩阵必有相同的秩.

定义（等价二次型） 设$f(x_1,x_2,\cdots,x_n)$与$g(y_1,y_2,\cdots,y_n)$是R上两个二次型，若可通过非退化的线性替换将其中一个化为另一个，则称这两个二次型等价.

显然有如下定理：

定理 8.1.2 两个二次型等价当且仅当它们的矩阵合同.

证：（等价⇒合同）存在$X=CY$（C可逆）使$X^TAX=Y^TBY$，则$Y^TC^TACY=Y^TBY$.而C^TAC亦是对称矩阵，故$C^TAC=B$.

（合同⇒等价） 存在可逆的C使$\mathrm{A}=C^TBC$，则令$\mathrm{Y}=\mathrm{CX}$有：

$$Y^TBY=X^TC^TBCX=X^TAX$$

注：（1）等价的二次型可看成变量选取不同，本质上无差别.

（2）合同矩阵有相同的秩，故可定义二次型的秩为它的矩阵的秩.

二次型的标准型：

若一个二次型$f(x_1，x_2，\cdots，x_n)$仅含平方项，即$f(x_1,x_2,\cdots,x_n)=a_1x_1^2+a_2x_2^2+\cdots+a_nx_n^2$，（其中$a_i$可能为零），则称该二次型为标准型.

若一个二次型$f(x_1,x_2,\cdots,x_n)$仅含平方项，即$f(x_1,x_2,\cdots,x_n)=a_1x_1^2+a_2x_2^2+\cdots+a_nx_n^2$，（其中$a_i$可能为零），则称该二次型为标准型.

一个有用的问题是：对任意一个二次型，能否通过一个非退化的线性替换将其化为标准型？或者等价地用矩阵语言说，对任意一个对称矩阵，能否通过合同变换将其变为对角矩阵？ 是！

定理 8.1.3　若A是一个R(或C)上的对称矩阵，则必存在R(或C)上的一个可逆矩阵C，使$C^TAC=B$是对角矩阵.

定理 8.1.3’ 设　$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\quad(a_{ij}=a_{ji})$是R(或C)上的一个n元二次型，则必存在非退化的线性替换$X=CY$将其化作标准型：$b_1y_1^2+b_2y_2^2+\cdots+b_ny_n^2$.（$b_i$可能为零！）

注：如前所言，C上的二次型用处不大，故C上的上述定理亦无大用处.

证明方法一（行列对称初等变换法）：　对A的阶数n用数归.

$n=1$时，A已是对角矩阵，故C为任意的一阶可逆矩阵都成立.

设$n-1$阶对称矩阵定理成立，考察n阶矩阵A：

（1）设$a_{11}\neq 0$，则进行初等变换：$C_j-\frac{a_{1j}}{a_{11}}C_1$，$r_j-\frac{a_{1j}}{a_{11}}r_1$

用初等矩阵表示即：

$$\begin{bmatrix}1 & & & \\ & \ddots & & \\ -\frac{a_{1j}}{a_{11}} & & 1 & \\ & & & \ddots\end{bmatrix}\begin{bmatrix}a_{11} & \cdots & a_{1j} & \cdots \\ \vdots & & & \\ a_{1j} & & & \\ \vdots & & & \end{bmatrix}\begin{bmatrix}1 & & -\frac{a_{1j}}{a_{11}} & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots\end{bmatrix}=$$

$$\begin{bmatrix}a_{11} & \cdots & 0 & \cdots \\ \vdots & & & \\ 0 & & & \\ \vdots & & & \end{bmatrix}$$

故$\cdots E_2^TE_1^TAE_1E_2\cdots=\begin{pmatrix}a_{11} & 0\\ 0 & A_{n-1}\end{pmatrix}$，即存在可逆$C_1$：

$C_1^TAC_1=\begin{pmatrix}a_{11} & 0\\ 0 & A_{n-1}\end{pmatrix}$

再用归纳法，可证存在可逆的C，使$C^TAC=diag[\cdots]$.

（2）若$a_{11}=0$，但存在某$a_{jj}\neq 0$.

进行初等变换：　$C_j\leftrightarrow C_1$，$r_j\leftrightarrow r_1$

用初等矩阵表示即：

$$\begin{bmatrix} 0 & & 1 & \\ & \ddots & & \\ 1 & & 0 & \\ & & & \ddots \end{bmatrix}\begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & a_{jj} & \\ & & & \ddots \end{bmatrix}\begin{bmatrix} 0 & & 1 & \\ & \ddots & & \\ 1 & & 0 & \\ & & & \ddots \end{bmatrix}=\begin{bmatrix} a_{jj} & & & \\ & \ddots & & \\ & & 0 & \\ & & & \ddots \end{bmatrix}$$

即：通过$E^TA\mathrm{E}$化为第一种情况.故得证.

（3）若所有$a_{ii}=0$.则必有某$a_{1j}\neq 0$，否则为$\mathrm{n}-1$元二次型.

作初等变换：$C_1+1\cdot C_j$，$r_1+1\cdot r_j$，用初等矩阵表示即：

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

$$=\begin{bmatrix} 2a_{1j} & \cdots & * & \cdots \\ \vdots & & & \\ * & & & \\ \vdots & & & \end{bmatrix}$$

即：通过行列对称的初等变换逐步化为低阶的对称矩阵.

证明方法二（配方法）：对二次型的变量个数 n 用数归.

$\mathrm{n}=1$时，$f(x_1)=a_{11}x_1^2$，显然已是标准型.设$\mathrm{n}-1$元二次型可通过非退化线性替换化为标准型，看 n 元情形：

（1）设某个$a_{jj}\neq 0$，则对x_j配方：

$f(x_1,x_2,\cdots,x_n)=$

$a_{jj}\left(x_j+\frac{a_{j1}}{a_{jj}}x_1+\cdots+\frac{a_{jn}}{a_{jj}}x_n\right)^2+f'\left(x_1,\cdots,x_{j-1},x_{j+1},\cdots,x_n\right)$

而$x_j\to x_j+\frac{a_{j1}}{a_{jj}}x_1+\cdots+\frac{a_{jn}}{a_{jj}}x_n$，其它$x_i\to x_i$，是可逆线性替换.

f'是$\mathrm{n}-1$元二次型，故由归纳假设可证n元二次型亦可化为标准型.

（2）若所有$a_{ii}=0$，则必有$a_{1j}\neq 0$，否则f中不含x_1，则令

$x_1=y_1+y_j$，$x_j=y_1-y_j$，其它$x_i=y_i$，是可逆线性替换，故

$2a_{1j}x_1x_j=2a_{1j}{y_1}^2-2a_{1j}{y_j}^2$化为第一种情况，亦证得.

证明方法三（仅对实二次型成立，用正交变换法）：上一章证明了，对实二次型必可通过正交变换相似对角化，即：存在$P^TP=1$，使$P^{-1}AP=diag$，而$P^T=P^{-1}$，故存在可逆的 P，使$P^TAP=diag.$

注：以上定理证明了二次型必定可化为标准型，但标准型不唯一！

既然二次型的标准型并不唯一，那么不同的标准型哪些性质一样呢？显然标准型中所含的平方项的个数一样，都等于二次型的秩. 也就是合同对角化后，主对角上 0 的个数一样. 下一节将证明，不仅 0 的个数一样，而且正数和负数的个数也一样，即惯性定理.

小结：将二次型对角化并求线性替换矩阵的三种方法：

1. 行列对称初等变换法（做一次行变换立即做一次相同的列变换）：对$\begin{bmatrix}A\\I_n\end{bmatrix}$的列及相应的行进行初等变换得到$\begin{bmatrix}C\\D\end{bmatrix}$，则

$\begin{cases}\cdots E_2^TE_1^TAE_1E_2\cdots=D\\ I_nE_1E_2\cdots=C\end{cases}$，即$C^TAC=D$，故引入$X=CY$，有：

$X^TAX=YC^TACY=YDY.$

2．配方法：

当没有平方项时，一个技巧是引入$\begin{cases}x_1=y_1+y_2\\x_2=y_1-y_2\end{cases}$

3.正交变换法：

即用正交变换将 A 相似对角化.注意：前两种方法得到的合同变换矩阵不一定是正交矩阵，故在几何上要改变长度（除空间的转动之外还包含伸缩等变换），因此改变形状，而正交变换法只包含了转动变换，故不改变形状！（应用：二次曲面的化简）. 具体含义见下面例题！

例 5　$\mathrm{A}=\begin{bmatrix}i&1&i\\1&1&3\\i&3&1+i\end{bmatrix}$，求可逆的 C，使$C^TAC=diag.$

例 6 分别用配方法和行列对称初等变换法将如下二次型化为标准型：

$f(x_1, x_2, x_3) = {x_1}^2 + 7{x_2}^2 + 8{x_3}^2 - 6x_1x_2 + 4x_1x_3 - 10x_2x_3$

$f(x_1, x_2, x_3) = x_1x_2 + 4x_1x_3 + x_2x_3$

例 7 在直角坐标系下，曲面方程为

$2{x_1}^2 + 5{x_2}^2 + 5{x_3}^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3 = 1$，试用主轴变换化简之并确定由面类型和主轴方向.

提示：$X \to Y$本质是$(\hat{\imath}, \hat{\jmath}, \hat{k}) \to (\widehat{\imath'}, \widehat{\jmath'}, \hat{k})$是一个坐标系的转动变换，主轴方向由新的坐标$\mathrm{Y} = [1, 0, 0]^T$，$[0, 1, 0]^T$，$[0, 0, 1]^T$来确定.

更一般的含有一次项的二次曲面：先将二次部分化为标准形，再平移坐标，化简一次项！除二次曲面的化简外，主轴变换的另一用途是转动惯量的化简！

§8.2 二次型的规范型和惯性定理

规范型 惯性定理 惯性指数、号差

二次型的规范型

上一节给出了二次型必可合同变换为标准型.但指出：标准型不唯一.

设$f(x_1, x_2, \cdots, x_n)=\sum_{i,j} a_{ij}x_ix_j=x^TAX$,

则可通过$X=CY(|C|\neq 0)$变成标准型：（r为二次型的秩）

$$f(x_1, x_2, \cdots, x_n)=C_1{y_1}^2+C_2{y_2}^2+\cdots+C_r{y_r}^2.$$

（1)考虑复数域 C 上：总可引入$y_1=\frac{1}{\sqrt{C_1}}z_1, y_2=\frac{1}{\sqrt{C_2}}z_2, \cdots, y_r=\frac{1}{\sqrt{C_r}}z_r$,

$y_{r+1}=z_{r+1}, \cdots, y_n=z_n$，$f(x_1, x_2, \cdots, x_n)={z_1}^2+{z_2}^2+\cdots+{z_r}^2$.

定理 8.2.1 复二次型一定可通过非退化线性替换化为规范型，且规范型唯一！

推论 8.2.2 复对称矩阵必合同于$diag[1, \cdots, 1, 0, \cdots, 0]$.

$$\begin{cases}\text{复对称矩阵合同} \Leftrightarrow \text{秩相等}\\ \text{复二次型等价} \Leftrightarrow \text{秩相等}\end{cases}.$$

（2）考虑实数域 R 上：

$C_i(i=1, 2, \cdots, n)$可为正或负.作变量的调换之后，可令

$\begin{cases} C_1, C_2, \cdots, C_p\text{为正}\\ C_{p+1}, C_{p+2}, \cdots, C_r\text{为负}\end{cases}$ 即：

$$f(x_1, x_2, \cdots, x_n)=d_1{x_1}^2+\cdots+d_p{x_p}^2-d_{p+1}{x_{p+1}}^2-\cdots-d_r{x_r}^2$$

其中$d_i>0(i=1, 2, \cdots, r)$，总可引入：

$y_1=\frac{1}{\sqrt{d_1}}z_1$, $y_2=\frac{1}{\sqrt{d_2}}z_2$, $\cdots$, $y_r=\frac{1}{\sqrt{d_r}}z_r$, $y_{r+1}=z_{r+1}$, $\cdots$, $y_n=z_n$,

则 $f(x_1, x_2, \cdots, x_n)={z_1}^2+{z_2}^2+\cdots+{z_p}^2-{z_{p+1}}^2-\cdots-{z_r}^2$.

上述系数为± 1的平方和形式，称为实二次型的规范型. 以下只研究实二次型.

定理8.2.3　R上任意一个n元二次型都可通过非退化的线性替换化为如下的规范型（其中V为二次型的秩）：

$${t_1}^2+{t_2}^2+\cdots+{t_p}^2-{t_{p+1}}^2-\cdots-{t_r}^2$$

即：任意一个实二次型必定与一个规范型等价.

惯性定理，规范型的唯一性

规范型中，平方项的个数等于二次型的秩，因此由二次型唯一确定. 一个问题是，通过不同的方法化为规范型后，正负1的个数是不是也唯一由二次型确定？若是，则规范型唯一.

定理8.2.4（惯性定理） 一个秩r的n元实二次型$f(x_1, x_2, \cdots, x_n)$，无论采用何种非退化的线性替代将其化为规范型，其中正项的个数p，负项的个数$\mathrm{r}-\mathrm{p}$都是唯一确定的！

注：由于上述定理，故p等于二次型的矩阵的正本征值个数，$\mathrm{r}-\mathrm{p}$等于负本征值个数，r等于非零本征值的个数。

证明：（反证法）假设$X=C_1Y$，$X=C_2Z$分别化为：（C_1，C_1可逆）

$$\begin{cases} f(x_1, x_2, \cdots, x_n)={y_1}^2+{y_2}^2+\cdots+{y_p}^2-{y_{p+1}}^2-\cdots-{y_r}^2. \\ f(x_1, x_2, \cdots, x_n)={z_1}^2+{z_2}^2+\cdots+{z_q}^2-{z_{q+1}}^2-\cdots-{z_r}^2 \end{cases}$$

且$p\neq q$，不妨设$p>q$，由于$Z={C_2}^{-1}C_1Y$，

考虑关于$\mathrm{Y}=[y_1, y_2, \cdots, y_n]$的线性方程：

$$\begin{cases} z_1 = 0 \\ z_2 = 0 \\ \quad\vdots \\ z_q = 0 \\ y_{p+1} = 0 \\ \quad\vdots \\ y_n = 0 \end{cases}$$

方程个数：$q + (n - p) = n + (q - p) < n$，未知元个数：n，故必有非零解，取其中之一为$\tilde{Y}$，可由$Z = {C_2}^{-1}C_1Y$得到$\tilde{Z} = {C_2}^{-1}C_1\tilde{Y} \neq 0$，将$\tilde{Y}$代入，可知二次型方程的值$f > 0$，而将$\tilde{Z}$代入又有$f < 0$，故矛盾！必有$p = q$，证毕.

定义：实二次型的规范型中正项个数 p 称该二次型的正惯性指数，负项个数$\mathrm{r} - \mathrm{p}$称该二次型的负惯性指数，$\mathrm{s} = \mathrm{p} - (r - p) = 2p - r$ 称二次型的符号差，简称号差.

定理 8.2.4’（惯性定理） 任意一个 n 阶实对称矩阵 A 必定合同于一个如下的被称为规范型的矩阵：$\begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix}$，且规范型唯一.

注：其中$\mathrm{r} = \mathrm{r}(A)$，亦等于非零本征值的个数，而 p 等于正本征值的个数.

注：惯性定理亦可用标准型来表述：即无论用何种非退化的线性替换将二次型化为标准型，其中正平方项的个数一定相同（负平方项个数亦相同）.因为如若不然的话，将这两种标准型再分别规范化，就与定理 8.2.4 矛盾.

推论 8.2.5 两个实二次型等价当且仅当他们有相同的秩和正惯性指数；两个实对称矩阵合同当且仅当他们有相同的秩和正惯性指数.

由上述定理，可将二次型分类：同一类的二次型等价：

推论 8.2.6 一切 n 元实二次型可分为$\frac{1}{2}(n+1)(n+2)$个类，属同一类的二次型相互等价，不同类的二次型不等价.

证：规范型$C_{r,\ p}=\begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix}$，r 可取0，1，…，n，固定 r 后，p 可取0，1，…，r，故共有$1+2+\cdots+(n+1)=\frac{1}{2}(n+1)(n+2)$中不同的规范型.

例 8 将上节课的二次型化为规范型.

例 9 3 元实二次型的类：

$$\begin{bmatrix} 0 & & \\ & 0 & \\ & & 0 \end{bmatrix},\ \begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix},\ \begin{bmatrix} -1 & & \\ & 0 & \\ & & 0 \end{bmatrix},\ \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix},\ \begin{bmatrix} 1 & & \\ & -1 & \\ & & 0 \end{bmatrix},$$
$$\begin{bmatrix} -1 & & \\ & -1 & \\ & & 0 \end{bmatrix},\ \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix},\ \begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \end{bmatrix},\ \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix},$$
$$\begin{bmatrix} -1 & & \\ & -1 & \\ & & -1 \end{bmatrix}.$$

§ 8.3 二次型正定性

正定二次型的定义　　正定的一些充要条件　　负定、准正定、准负定

正定二次型的定义

R 上的 n 元二次型 $f(x_1, x_2, \cdots, x_n) = \sum_{i,j} a_{ij}x_ix_j = x^TAx$ 可看成定义在 R 上的 n 个变量的实函数。

定义（正定二次型，正定矩阵） R 上的实二次型$f(x_1, x_2, \cdots, x_n)$，若对于任意一组不全为 0 的实数$C_1, C_2, \cdots, C_n$，恒有 $f(C_1, C_2, \cdots, C_n) > 0$，则称该二次型为正定二次型。设 A 是一实对称矩阵，若二次型x^TAx是正定的，则称 A 为正定矩阵。

注：即：$\forall x \neq 0$，总有$x^TAx > 0$，类比内积$(\alpha, \alpha) > 0$.

例 10 欧空间中的度规矩阵 A 是一个正定矩阵.

$(\forall\alpha \neq 0, (\alpha, \alpha) > 0)$

正定二次型的性质，及判定方法

定理 8.3.1 n 元二次型$f(x_1, x_2, \cdots, x_n)$正定的充要条件是正惯性指数$p = n$.

证明：$(\Leftarrow)$ 假设$p = n$，则由$X = CY$得：

$f(x) = g(x) = {y_1}^2 + {y_2}^2 + \cdots + {y_n}^2$, $\forall\tilde{X} \neq 0$, 有$\tilde{Y} = C^{-1}\tilde{X}$, 代入有：

$f(\tilde{x}) = g(\tilde{y}) > 0$，故正定.

$(\Rightarrow)$ 设$f(x)$正定，则假设$p < n$，有：$X = CY$，

$f(x) = g(y) = {y_1}^2 + \cdots + {y_p}^2 - {y_{p+1}}^2 - \cdots - {y_r}^2$ ，

取$\tilde{Y} = [0, 0, \cdots, 1]^T$, 有：$\tilde{X} = C\tilde{Y} \neq 0$, 代入得：$f(\tilde{x}) = g(\tilde{y}) \leq 0$，矛盾，故$p = n$.

定理 8.3.1’ n 阶实对称矩阵正定的充要条件是其正惯性指数$p=n.$

推论 8.3.2 n 阶实对称矩阵正定的充要条件是其没有零本征值（满秩），且本征值全为正.

推论 8.3.3 n 阶实对称矩阵 A 正定的充要条件是其与I_n合同，或者说：存在可逆矩阵 D，使$A=D^TD.$

推论 8.3.4 若 A 为正定矩阵，则$|A|>0.$

证明：$|A|=\prod_{i=1}^{n}\lambda_i>0$ ，或$|A|=|D^TD|=|D|^2>0.$

推论 8.3.4 不是充分条件，一个实用的判定方法是：

定理 8.3.5 实二次型$f(x)=x^TAx$正定的充要条件是 A 的一切顺序主子式都大于零，即：

$A_1=a_{11}>0$，$A_2=\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}>0$，$A_3=\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}>0$，$\cdots$，

$A_n=\begin{vmatrix}a_{11} & \cdots & a_{13}\\ \vdots & \vdots & \vdots\\ a_{31} & \cdots & a_{33}\end{vmatrix}=|A|>0.$

注： 主子式：取相同的行和列构成的子式.（子式：相同数目的行和列）顺序主子式：按顺序依次取相同的前个k行和列构成的子式.

证明：（必要性）设$f(x)=x^TAx$正定，则可令

$X=\left[x_1,\ x_2,\ \cdots,\ x_k,\ 0,\ 0,\ \cdots,\ 0\right]^T$代入 $f(x)$得到一个k元的二次型：$\mathrm{g}\left(x_1,\ x_2,\ \cdots,\ x_k\right)=f(x)=x^TAx$，由于$f$正定，故无论

$X_k=\left[x_1,\ x_2,\ \cdots,\ x_k\right]^T$怎样选取$\mathrm{g}(X_k)>0$，即$\mathrm{g}(X_k)$正定.

由推论 8.3.4 可知 k 阶顺序主子式$A_k>0.$

(注：同样的方法可证明若 A 正定，则 A 的一切主子式均大于 0)

（充分性）：设 A 的一切顺序主子式均大于零. 对阶数 n 用数归.

$n=1$时，$a_{11}>0$，显然$a_{11}x_1x_1$正定，充分性成立.

假设 n-1 元次型充分性成立，则对 n 元次型$f(x)=x^TAx$. 令

$A=\begin{bmatrix} A_{n-1} & b \\ b & a_{nn} \end{bmatrix}$，则由归纳假设$A_{n-1}$正定，故存在

$Q_{n-1}^TA_{n-1}Q_{n-1}=I_{n-1}$, 取 Q= $\begin{bmatrix} Q_{n-1} & Y \\ 0 & 1 \end{bmatrix}$，有

$Q^TAQ=\begin{bmatrix} Q_{n-1}^TA_{n-1}Q_{n-1} & Q_{n-1}^TA_{n-1}Y+Q_{n-1}^Tb \\ YA_{n-1}Q_{n-1}+b^TQ_{n-1} & Y^TA_{n-1}Y+Y^Tb+b^TY+a_{nn} \end{bmatrix}$，

取$Y=-A_{n-1}b$可知: $Q^TAQ=\begin{bmatrix} I_{n-1} & 0 \\ 0 & a_{nn}-b^TA_{n-1}^{-1}b \end{bmatrix}$, 由于$|A|>0$，

故$\left|Q^TAQ\right|>0$，即：$a_{nn}-b^TA_{n-1}^{-1}b>0$，故$Q^TAQ$正定，从而 A 正定，证毕.

例 11　判别二次型

$f\left(x_1, x_2, x_3\right)=5{x_1}^2+{x_2}^2+5{x_3}^2+4x_1x_2-8x_1x_3-4x_2x_3$的正定性.

方法：1. 顺序主子式，2. 配方，3. 行列对称的初等变换，4. 本征值

例 12　求 t 的取值范围，使下面二次型为正定型：

$f\left(x_1, x_2, x_3\right)={x_1}^2+{x_2}^2+5{x_3}^2+2tx_1x_2-2x_1x_3+4x_2x_3$

负定、准正定、准负定、不定二次型

定义：R 上的实二次型$f\left(x_1, x_2, \cdots, x_n\right)=x^TAx$，若对于任意一组不全为 0 的实数$C_1$，$C_2$，$\cdots$，$C_n$，恒有：

（1）$f\left(C_1, C_2, \cdots, C_n\right)<0$，则称该二次型为负定二次型，称 A 为负定矩阵.

（2）$f\left(C_1, C_2, \cdots, C_n\right)\geq 0$，则称该二次型为准正定二次型，称 A 为准正定矩阵.

（3）$f\left(C_1, C_2, \cdots, C_n\right) \leq 0$，则称该二次型为准负定二次型，称 A 为准负定矩阵. 若$f(x)$既不是准正定的，也不是准负定的，则称为不定二次型.

显然一个二次型$f(x) = x^T A x$是负定的，当且仅当二次型

$-f(x) = x^T(-A)x$是正定的. 由此根据正定二次型的定理直接可得：

定理 8.3.6　设$f\left(x_1, x_2, \cdots, x_n\right) = x^T A x$是一个 n 元实二次型，则下列论述等价：

（1）$f(x)$是负定的；

（2）$f(x)$的负惯性指数是 n；

（3）$f(x) = x^T A x$中矩阵 A 的 n 个本征值全部小于零；

（4）$f(x) = x^T A x$中矩阵 A 的顺序主子式A_k满足：

$A_1 = a_{11} < 0$，$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$，$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0$，⋯，

即：$(-1)^k A_k > 0$.

对准正定二次型有如下类似的定理：

定理 8.3.7　设$f\left(x_1, x_2, \cdots, x_n\right) = x^T A x$是一个 n 元实二次型，则下列论述等价：

（1）$f(x)$是准正定的；

（2）$f(x)$的正惯性指数等于秩；

（3）$f(x) = x^T A x$中矩阵 A 的本征值全部大于或等于零；

（4）$f(x) = x^T A x$中矩阵 A 的所有主子式都大于或等于零：

注：尤其要注意第（4）条与正定的区别：要所有主子式，不仅是顺序主子式.

例 13 $f\left(x_1,\ x_2,\ x_3\right)={x_1}^2+{x_2}^2+{x_3}^2+2x_1x_2+4x_1x_3+4x_2x_3$

$\mathrm{A}=\begin{bmatrix}1 & 1 & 2\\1 & 1 & 2\\2 & 2 & 1\end{bmatrix}$，$D_1=1>0$，$D_2=0$，$D_3=0$

但$f\left(x_1,\ x_2,\ x_3\right)=(x_1+x_2+2x_3)^2-3{x_3}^2$ 不定！

证明：$(1)\Leftrightarrow(2)$ 与定理 8.3.1 的证明类似，只须将证明中的 n 改为 r，并将>改为≥即可.

$(2)\Leftrightarrow(3)$ 显然等价，因为非零本征值个数等于秩，而正本征值个数等于正惯性指数.

$(1)\Leftrightarrow(4)$的证明：

$(1)\Longrightarrow(4)$ 设$f\left(x_1,\ x_2,\ \cdots,\ x_n\right)=x^TAx$准正定，则任取 k 个$x_i$非零，其余的$x_i$均为 0，代入得$f\left(x_1,\ x_2,\ \cdots,\ x_n\right)=g\left(x_{i1},\ x_{i2},\ \cdots,\ x_{ik}\right)$可知：k 元二次型$g$必准正定（恒大于或等于 0）. 故由第（3）条的推论可知，g的矩阵行列式必大于或等于 0，而 g 的矩阵行列式正好是 A 的 k 阶主子式. 故 A 的任意 k 阶主子式均大于零$(k=1,2,\cdots,n)$.

$(4)\Longrightarrow(1)$ 设 A 的所有主子式均大于零，则构造$B=\lambda I+A(\lambda>0)$

令A_k和B_k分别是 A 和 B 的 k 阶顺序主子式构成的矩阵，于是：

$$|B_k|=|\lambda I_k+A_k|=\begin{vmatrix}\lambda+a_{11} & a_{12} & \cdots & a_{1k}\\ a_{21} & \lambda+a_{22} & \cdots & a_{2k}\\ \vdots & \vdots & \vdots & \vdots\\ a_{k1} & a_{k2} & \cdots & \lambda+a_{kk}\end{vmatrix}$$

$$=\lambda^k+p_1\lambda^{k-1}+\cdots+p_{k-1}\lambda+p_k$$

其中p_i为A_k中一切i级主子式之和，由题设可知 A 的一切主子式均大于或等于 0，故$p_i\geq 0$，故$\lambda>0$时$|B_k|>0$，$B=\lambda I+A$为正定矩阵.

假设A不是准正定的，则存在$\tilde{x} \in R^n$使$\tilde{x}^T A\tilde{x} = -C < 0$，那么，取$\lambda = C/\tilde{x}^T\tilde{x} > 0$，则有：$\tilde{x}^T B\tilde{x} = \tilde{x}^T(\lambda I + A)\tilde{x} = \lambda\tilde{x}^T x - C = C - C = 0$，矛盾，故A必为准正定矩阵，证毕.

注意到二次型$f(x)$为准负定的充要条件是$-f(x)$为准正定，故可得到与上述定理类似的结论.

例14 多元函数的极值问题与二次型的正定性之关系：

解：二元Taylor公式：

$$f(x,\ y) = f(0,\ 0) + \sum_{k=1}^{n} \frac{1}{k!}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^k f(0,\ 0) + R_n$$

极值点$f_x'(x_0,\ y_0) = f_y'(x_0,\ y_0) = 0$，故

$df(x_0,\ y_0) = f_{xx}''(x_0,\ y_0)dx^2 + 2f_{xy}''(x_0,\ y_0)dxdy +$

$f_{yy}'(x_0,\ y_0)dy^2$　二次型正定⟹极小值；二次型负定⟹极大值；不定⟹非极值，多元函数类似讨论！

例15 若对称矩阵A可逆，求证A^{-1}与A合同.

证：$A = C^T F C$　　$A^{-1} = C^T F^{-1} C = C^T F^{-1} F F^{-1} C = (F^{-1}C)^T F (F^{-1}C)$

例16 求证：（1）数域F上的n阶反对称必与如下形式矩阵合同：

$$\begin{bmatrix} 0 & 1 & & & & & & \\ -1 & 0 & & & & O & & \\ & & \ddots & & & & & \\ & & & 0 & 1 & & & \\ & & & -1 & 0 & & & \\ & & & & & 0 & & \\ & O & & & & & \ddots & \\ & & & & & & & 0 \end{bmatrix}$$

（2）反对称矩阵的秩为偶数；

（3）下上两个n阶反对称矩阵合同当且仅当它们秩相等.

证：（1）1.从上到下找第一个非零行，行列式对称初等变换将零行零列移到最后面的行和列，2.非零元分别归一到 1 和-1 并移到前面，3.用 1 和-1 将它们所在的行和列的其它元素化为零，4.对后面的行列重复以上三步即得证.

例 17 若A，B均为同阶正定矩阵，$k,l\in R^+$, 求证$kA+kB$正定.

证明：用正定的定义：$\forall x\in R^n\neq 0$, 有：

$x^T(kA+lB)x=k(x^TAx)+l(x^TBx)>0$, 证毕.

例 18 设 A 为正定矩阵，求证A^{-1}，kA，A^m，A^*均正定（$k\in R^+$，$m\in Z^+$）.

证：由本征值来判定！注意$A^*=|A|A^{-1}$.

例 19 设 A 为实对称矩阵，且$A^3-5A^2+A-5I=0$，求证 A 为正定.

证：若 A 满足多项式方程$f(A)=0$，则 A 的所有本征值亦满足，因为若$A\alpha=\lambda\alpha$，则$f(A)\alpha=f(\lambda)\alpha=0$，$\alpha\neq 0$，故$f(\lambda)=0$，而

$\lambda^3-5\lambda^2+\lambda-5=0\Longrightarrow(\lambda-5)(\lambda^2+1)=0$，$\lambda$只可能是 5 或$\pm i$.

又 A 为实对称，故 A 的所有本征值均为 5（n 重根），大于零，故 A 正定.

例 20 A 为实对称矩阵，求证：（1）当ε充分小时，$I+\varepsilon\mathrm{A}$正定；（2）当 t 充分大时，$\mathrm{t}I+\mathrm{A}$正定. 提示：$I+\varepsilon\mathrm{A}$和$\mathrm{t}I+\mathrm{A}$的本征值分别为

$1+\varepsilon\lambda$和$t+\lambda$.

例 21 设 A 为 n 阶正定矩阵，求证$|A+2I|>2^n$. 提示：$A+2I$的本征值为$\lambda+2$.

例 22 设 A 为一正定矩阵，求证存在另一个正定矩阵 S 使$A=S^2$.

证：注意到正定矩阵是实对称矩阵，故必有：

$A=P^{-1}FP$，$F=diag\left[\lambda_1, \cdots, \lambda_n\right]$，A 正定，故$\lambda_i>0$，故$F=\tilde{F}^2$，

其中$\tilde{F}=diag\left[\sqrt{F_1}, \cdots, \sqrt{F_n}\right]$，即：

$A=P^{-1}\tilde{F}^2P=P^{-1}\tilde{F}PP^{-1}\tilde{F}P=\left(P^{-1}\tilde{F}P\right)^2$

例 23 已知$A=\begin{bmatrix}13&14&4\\14&24&18\\4&18&29\end{bmatrix}$正定，求另一正定矩阵 B 使$B^2=A$.

解：方法同上一题.

例 24 设$f(x)=x^TAx$，$g(x)=x^TBx$是两实二次型且 g 正定，求证存在一个非退化的线性替换$X=CY$分别将f和g化为$\lambda_1{y_1}^2+\cdots+\lambda_n{y_n}^2$和${y_1}^2+\cdots+{y_n}^2$（同时标准化）.

证：首先，存在C_1使${C_1}^TBC_1=I_n$，此时${C_1}^TAC_1=A'$，A'实对称，故存在正交矩阵C_2，使${C_2}^TA'C_2=F$，此时${C_2}^TI_nC_2={C_2}^{-1}I_nC_2=I_n$，综合上述，令$C=C_1C_2$，则$\begin{cases}C^TAC={C_2}^T{C_1}^TAC_1C_2={C_2}^TA'C_2=F\\C^TBC={C_2}^T{C_1}^TBC_1C_2={C_2}^{-1}I_nC_2=I_n\end{cases}$.

例 25 设 A, B 为同阶的正定和非正定矩阵，且$B\neq 0$，求证

$|A+B|>|A|+|B|$.

证：类似上题，可证存在 C 可逆，使：

$C^TAC=I$，$C^TBC=F=diag\left[\lambda_1, \cdots, \lambda_n\right]$，$\lambda_i>0$不全为零，故

$$\begin{aligned}|A+B|&=\left|{C^{-1}}^T C^T(A+B)CC^{-1}\right|=|C^{-1}|^2|I+F|>|C^{-1}|^2(|I|+|F|)\\&=\left|{C^{-1}}^TIC^{-1}\right|+\left|{C^{-1}}^TFC^{-1}\right|=|A|+|B|.\end{aligned}$$

例 26 设 A 是一个 n 阶正实矩阵，求证$f\left(x_1, x_2, \cdots, x_n\right)=\begin{vmatrix}A&X\\X^T&0\end{vmatrix}$是负定矩阵.

证:$f(\mathrm{x})=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & x_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n \\ x_1 & x_2 & \cdots & x_n & 0 \end{vmatrix}$,令$M_{ij}$为 A 的$a_{ij}$的余子式,$A_{ij}$为 A 的$a_{ij}$的代数余子式，故${x_i}^2$的系数为

$(-1)^{n+1+i}(-1)^{n+i}M_{ii}=-(-1)^{i+i}M_{ii}=A_{ii}$，$x_ix_j$的系数为

$(-1)^{n+1+i}(-1)^{n+j}\left(M_{ij}+M_{ji}\right)=-\left(A_{ij}+A_{ji}\right)$，故$\mathrm{f(x)}=-X^TA^*X$

（注：这一点亦可用另一方法来证：

$\begin{bmatrix} A & X \\ X^T & 0 \end{bmatrix}\begin{bmatrix} I & -A^{-1}X \\ 0 & 1 \end{bmatrix}=\begin{bmatrix} A & 0 \\ X^T & -X^TA^{-1}X \end{bmatrix}$)，取行列式即：

$f(\mathrm{x})=-|A|X^TA^{-1}X=-X^TA^*X$，由于 A 正定，故$A^*$正定，故$\mathrm{f(x)}$负定，证毕.

例 27 若 A 正定，求证$|A|\le a_{nn}|A_{n-1}|$，其中$|A_{n-1}|$是 A 的 n-1 阶顺序主子式.

证：A 正定，故A_{n-1}亦正定，故$\begin{bmatrix} A_{n-1} & X \\ X^T & 0 \end{bmatrix}$负定.

$$而A=\begin{vmatrix} a_{11} & \cdots & a_{1n-1} & a_{1n} \\ \vdots & & \vdots & \vdots \\ a_{n-11} & \cdots & a_{n-1n-1} & a_{n-1n} \\ a_{n1} & \cdots & a_{nn-1} & 0 \end{vmatrix}+\begin{vmatrix} a_{11} & \cdots & a_{1n-1} & 0 \\ \vdots & & \vdots & \vdots \\ a_{n-11} & \cdots & a_{n-1n-1} & 0 \\ a_{n1} & \cdots & a_{nn-1} & a_{nn} \end{vmatrix}$$

$\le a_{nn}|A_{n-1}|$.

例 28 若 A 正定，求证$|A|\le a_{11}a_{22}\cdots a_{nn}$.

证：反复利用上一例题的结论.

例 29 设 A 为一 n 阶实矩阵，求证： $|A|^2\le\prod_{i=1}^{\mathrm{n}}({a_{1i}}^2+{a_{2i}}^2+\cdots+{a_{ni}}^2)$

证：$|A|=0$时显然成立，$|A|\ne 0$时，A^TA为正定矩阵，利用上一题的结论有$|A|^2=|A^TA|\le(A^TA)_{11}(A^TA)_{22}\cdots(A^TA)_{nn}$，证毕.

§8.4 厄米性

厄米型的定义、等价　N 元厄米型可化为 2n 元二次型　　厄米型的标准形、规范形

惯性定理　　　　厄米型的正定性

厄米型的定义：

定义：设 A 是 C 上的一个 n 阶厄米矩阵，$X \in C^n$，则

$f(x_1, x_2, \cdots, x_n) = f(x) = X^+AX$，称为复域 C 上的一个厄米型，A 称为厄米型的矩阵.

$$\begin{aligned}
f(x_1, x_2, \cdots, x_n) = {} & a_{11}|x_1|^2 + a_{12}\overline{x_1}x_2 + a_{13}\overline{x_1}x_3 + \cdots + a_{1n}\overline{x_1}x_n \\
& + a_{21}\overline{x_2}x_1 + a_{22}|x_2|^2 + a_{23}\overline{x_2}x_3 + \cdots + a_{2n}\overline{x_2}x_n \\
& + a_{31}\overline{x_3}x_1 + a_{32}\overline{x_3}x_2 + a_{33}|x_3|^2 + \cdots + a_{3n}\overline{x_3}x_n \\
& + \cdots \\
& + a_{n1}\overline{x_n}x_1 + a_{n2}\overline{x_n}x_2 + a_{n3}\overline{x_n}x_3 + \cdots + a_{nn}|x_n|^2 \\
= {} & \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{x_i}x_j \qquad （其中 a_{ij} = \overline{a_{ji}}）
\end{aligned}$$

显然$a_{ij}\overline{x_i}x_j = \overline{(a_{ji}\overline{x_j}x_i)}$；$\overline{f} = (X^+AX)^+ = X^+AX = f$,

即：厄米型是$C^n \to R$的一个映射，取值为实数！

定理 8.4.1　设A，B是对称矩阵，若对任意$X \epsilon C^n$都有$X^+AX = X^+BX$，则$A = B$.

证明：1.依次取$X = [1，0，\cdots，0]^T$，$[0，1，\cdots，0]^T$，$\cdots$，$[0，0，\cdots，1]^T$代入$X^+AX = X^+BX$,即可证明$a_{ii} = b_{ii}(i = 1,2,\cdots,n)$.

2.再取$x_i=x_j=1$，其余$x_k=0(k=1,2,\cdots n,k\neq i,j)$，代入

$X^+AX=X^+BX$，有$a_{ij}+a_{ji}=b_{ij}+b_{ji}$.

3.再取$x_i=i$，$x_j=1$，其余$x_k=0(k=1,2,\cdots n,k\neq i,j)$，代入

$X^+AX=X^+BX$，有$-a_{ij}+a_{ji}=-b_{ij}+b_{ji}$.综上有：$A=B$.

注：与实数域上不同！R 上$X^TAX=X^TBX$只能得到 A，B 对称部分相等！原因：任意复矩阵$B=\frac{1}{2}(B+B^+)+\frac{1}{2}(B-B^+)=F+G$

$\overline{X^TGX}=(X^+GX)^+=X^+(-G)X=-(X^+GX)$. 纯虚数.

结论：厄米型与厄米矩阵一一对应！

例 30　酉空间选定一组基，有内积$(\alpha，\alpha)$即可确定一个厄米型X^+AX.

例 31 酉空间中标准正交基下厄米变换的期望值$(\alpha,\sigma(\alpha))$确定了一个厄米型.

n 元厄米型可化为 2n 元二次型：

注意到厄米型的值为实数，故令$x_i=u_i+iv_i$，$a_{ij}=b_{ij}+ic_{ij}$, 即：

$X=U+iV$，$A=B+iC$，有：

$$X^+AX=\sum_{i，j}\overline{x_i}\,a_{ij}x_j=Re\left(\sum_{i，j}\overline{x_i}\,a_{ij}x_j\right)$$

$$=\sum_{i，j}\ Re\left(b_{ij}+ic_{ij}\right)(u_i-iv_i)\left(u_j+iv_j\right)$$

$$=\sum_{i，j}\ \left(b_{ij}u_iu_j+b_{ij}v_iv_j+c_{ij}v_iu_j-c_{ij}u_iv_j\right)$$

故将u_i，v_i排成 2n 元实列向量$\begin{bmatrix}U\\V\end{bmatrix}$后上式为一个二次型.

亦可写成更紧促的形式：

引入$\mathrm{Y}=\begin{bmatrix}Re(x)\\Im(x)\end{bmatrix}=\begin{bmatrix}\frac{1}{2}(x+\overline{x})\\\frac{1}{2i}(x-\overline{x})\end{bmatrix}\in R^{2n}$

$A'=\begin{bmatrix}Re(A) & -Im(A)\\Im(A) & Re(A)\end{bmatrix}=\begin{bmatrix}\frac{1}{2}\left(A+\overline{A}\right) & \frac{i}{2}\left(A-\overline{A}\right)\\\frac{1}{2i}\left(A-\overline{A}\right) & \frac{1}{2}\left(A+\overline{A}\right)\end{bmatrix}$,则$B=B^T$,且有:

$X^+AX=Y^TA'Y$(直接展开右边可验证)

线性替换：

定义（线性替换） 数域C上两组变量$x_1,x_2,\cdots,x_n$和$y_1,y_2,\cdots,y_n$之间的关系：

$$\begin{cases}x_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n\\x_2=c_{21}y_1+c_{22}y_2+\cdots+c_{2n}y_n\\\qquad\qquad\cdots\\x_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n\end{cases}\left(c_{ij}\epsilon C\right)$$

称数域C上由变量$x_1,x_2,\cdots,x_n$到变量$y_1,y_2,\cdots,y_n$的一个线性变换.

进一步，若系数组成的矩阵$\mathrm{C}=(c_{ij})_{m\times n}$是可逆的，则称该线性替换是可逆的，或称非退化的.

若引入列向量$X=[x_1,x_2,\cdots,x_n]^T.Y=[y_1,y_2,\cdots,y_n]^T$，则可写成矩阵形式：

$$X=CY$$

引入$X=CY$的线性替换后，厄米型X^TAX可用Y表示为：

$$X^+AX=(CY)^+A(CY)=Y^+(C^+AC)Y$$

合同矩阵和等价厄米型：

定义（合同） 设A，B是 C 上的两n阶方阵，若存在 C 上可逆的C，使：$C^+AC=B$，则称A与BH合同，A和B是H合同矩阵.

注意：（1）区分矩阵的等价、相似、合同、H 合同！

（2）由于C可逆，故 H 合同是一个等价关系.

（3）H 合同矩阵必有相同的秩.

定义（等价厄米型） 设$f(x_1,x_2,\cdots,x_n)$与$g(y_1,y_2,\cdots,y_n)$是 C 上两个厄米型，若可通过非退化的线性替换将其中一个化为另一个，则称这两个厄米型等价.

显然有如下定理：

定理 8.4.2 两个厄米型等价当且仅当它们的矩阵 H 合同.

证：（等价⇒合同）存在$X=CY$（C可逆）使$X^{+}AX=Y^{+}BY$，则$Y^{+}C^{+}ACY=Y^{+}BY$.故$C^{+}AC=B$.

（合同⇒等价） 存在可逆的C使$A=C^{+}BC$，则令$Y=CX$有：

$$Y^{+}BY=X^{+}C^{+}BCX=X^{+}AX$$

注：（1）等价的厄米型可看成变量选取不同，本质上无差别.

（2）合同矩阵有相同的秩，故可定义厄米型的秩为它的矩阵的秩.

厄米型的标准型：

若一个厄米型$f\left(x_1，x_2，\cdots，x_n\right)$仅含平方项，即$f(x_1,x_2,\cdots,x_n)=a_1|x_1|^2+a_2|x_2|^2+\cdots+a_n|x_n|^2$，（其中$a_i$可能为零），则称该厄米型为标准形.

一个有用的问题是：对任意一个厄米型，能否通过一个非退化的线性替换将其化为标准型？或者等价地用矩阵语言说，对任意一个厄米矩阵，能否通过 H 合同变换将其变为对角矩阵？ 是！

定理 8. 4. 3 若A是一个复域C上的一个厄米矩阵，则必存在复域C上的一个可逆矩阵C，使$C^{+}AC=B$是对角矩阵.

定理 8.4.3’ 设$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\overline{x_i}x_j\left(a_{ij}=\overline{a_{ji}}\right)$是复域C上的一个n元厄米型，则必存在非退化的线性替换$X=CY$将其化作标准型：$b_1|y_1|^2+b_2|y_2|^2+\cdots+b_n|y_n|^2$.（$b_i$可能为零！）

注：合同对角化后必为实矩阵，因为C^+AC亦厄米，故主对角元全为实！

证明方法一（行列复共轭初等变换法）： 对A的阶数n用数归.

$n=1$时，A已是对角矩阵，故C为任意的一阶可逆矩阵都成立.

设$n-1$阶对称矩阵定理成立，考察n阶矩阵A：

（1）设$a_{11}\neq 0$，则进行初等变换：$C_j-\frac{a_{1j}}{a_{11}}C_1$，$r_j-\frac{\overline{a_{1j}}}{a_{11}}r_1$

用初等矩阵表示即：

$$\begin{bmatrix} 1 & & & \\ & \ddots & & \\ -\frac{\overline{a_{1j}}}{a_{11}} & & 1 & \\ & & & \ddots \end{bmatrix}\begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots \\ \vdots & & & \\ \overline{a_{1j}} & & & \\ \vdots & & & \end{bmatrix}\begin{bmatrix} 1 & & -\frac{a_{1j}}{a_{11}} & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \end{bmatrix}=$$

$$\begin{bmatrix} a_{11} & \cdots & 0 & \cdots \\ \vdots & & & \\ 0 & & & \\ \vdots & & & \end{bmatrix}$$

故$\cdots E_2^+E_1^+AE_1E_2\cdots=\begin{pmatrix} a_{11} & 0 \\ 0 & A_{n-1} \end{pmatrix}$，即存在可逆$C_1$：

$C_1^+AC_1=\begin{pmatrix} a_{11} & 0 \\ 0 & A_{n-1} \end{pmatrix}$

再用归纳法，可证存在可逆的C，使$C^+AC=diag[\cdots]$.

（2）若$a_{11}=0$，但存在某$a_{jj}\neq 0$.

进行初等变换： $C_j\leftrightarrow C_1$，$r_j\leftrightarrow r_1$

用初等矩阵表示即：

$$\begin{bmatrix}0 & & 1 & \\ & \ddots & & \\ 1 & & 0 & \\ & & & \ddots\end{bmatrix}\begin{bmatrix}0 & & & \\ & \ddots & & \\ & & a_{jj} & \\ & & & \ddots\end{bmatrix}\begin{bmatrix}0 & & 1 & \\ & \ddots & & \\ 1 & & 0 & \\ & & & \ddots\end{bmatrix}=\begin{bmatrix}a_{jj} & & & \\ & \ddots & & \\ & & 0 & \\ & & & \ddots\end{bmatrix}$$

即：通过$E^{+}A$E化为第一种情况.故得证.

（3）若所有$a_{ii}=0$.则必有某$a_{1j}\neq 0$，否则为$\mathrm{n}-1$元厄米型.

作初等变换：$C_1+1\cdot C_j$，$r_1+1\cdot r_j$ （$Re\left(a_{1j}\right)\neq 0$时）

或 $C_1-i\cdot C_j$，$r_1+i\cdot r_j$ （$Im\left(a_{1j}\right)\neq 0$时）

用初等矩阵表示即：

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

$$=\begin{bmatrix}2Re\left(a_{1j}\right) & \cdots & * & \cdots \\ \vdots & & & \\ * & & & \\ \vdots & & & \end{bmatrix}$$

即：通过行列复共轭的初等变换逐步化为低阶的对称矩阵.

证明方法二（配方法）：对厄米型的变量个数 n 用数归.

$\mathrm{n}=1$时，$f(x_1)=a_{11}|x_1|^{2}$，显然已是标准型.设$\mathrm{n}-1$元厄米型可通过非退化线性替换化为标准型，看 n 元情形：

（1）设某个$a_{jj}\neq 0$，则对x_j配方：

$f(x_1,x_2,\cdots,x_n)=$

$$a_{jj}\left(x_j+\frac{a_{j1}}{a_{jj}}x_1+\cdots+\frac{a_{jn}}{a_{jj}}x_n\right)^2+f'\left(x_1,\cdots,x_{j-1},x_{j+1},\cdots,x_n\right)$$

而$x_j\rightarrow x_j+\frac{a_{j1}}{a_{jj}}x_1+\cdots+\frac{a_{jn}}{a_{jj}}x_n$，其它$x_i\rightarrow x_i$，是可逆线性替换.

f'是$\mathrm{n}-1$元厄米型，故由归纳假设可证n元厄米型亦可化为标准形.

（2）若所有$a_{ii}=0$，则必有$a_{1j}\neq 0$，否则f中不含x_1，则令

$x_1=y_1+y_j$，$x_j=y_1+iy_j$，其它$x_i=y_i$，是可逆线性替换，故

$a_{1j}\overline{x_1}x_j+a_{j1}\overline{x_j}x_1=\left(a_{1j}+a_{j1}\right)|y_1|^2+i\left(a_{1j}-a_{j1}\right)|y_i|^2+\cdots$化为第(1)或第（2）种情况，证毕.

证明方法三（用幺正变换法）：上一章证明了，对厄米矩阵必可通过幺正变换相似对角化，即：存在$P^{+}P=1$，使$P^{-1}AP=diag$，而$P^{+}=P^{-1}$，故存在可逆的P，使$P^{+}AP=diag$.

注：以上定理证明了厄米型必定可化为标准型，但标准型不唯一！

既然二次型的标准型并不唯一，那么不同的标准型哪些性质一样呢？显然标准型中所含的平方项的个数一样，都等于厄米型的秩. 也就是合同对角化后，主对角上 0 的个数一样. 下一节将证明，不仅 0 的个数一样，而且正数和负数的个数也一样，即惯性定理.

厄米型的规范型

上一节给出了厄米型必可合同变换为标准型. 但指出：标准型不唯一.

设 $f\left(x_1, x_2, \cdots, x_n\right)=\sum_{i, j}a_{ij}\,x_i x_j=x^{+}AX(A^{+}=A)$

则可通过$X=CY(|C|\neq 0)$变成标准型：（r为二次型的秩）

$f\left(x_1, x_2, \cdots, x_n\right)=C_1|y_1|^2+C_2|y_2|^2+\cdots+C_r|y_r|^2$.

$C_i\left(i=1, 2, \cdots, n\right)$为实数，故可为正或负. 作变量的调换之后，可令

$\begin{cases} C_1, C_2, \cdots, C_p\text{为正} \\ C_{p+1}, C_{p+2}, \cdots, C_r\text{为负} \end{cases}$ 即：

$f\left(x_1, x_2, \cdots, x_n\right)$

$=d_1|y_1|^2+\cdots+d_p\left|y_p\right|^2-d_{p+1}\left|y_{p+1}\right|^2-\cdots-d_r|y_r|^2$

其中$d_i > 0(i = 1，2，\cdots，r)$，总可引入：

$y_1 = \frac{1}{\sqrt{d_1}} z_1$，$y_2 = \frac{1}{\sqrt{d_2}} z_2$，$\cdots$，$y_r = \frac{1}{\sqrt{d_r}} z_r$，$y_{r+1} = z_{r+1}$，$\cdots$，$y_n = z_n$，

则$f(x_1，x_2，\cdots，x_n) = |z_1|^2 + \cdots + |z_p|^2 - |z_{p+1}|^2 - \cdots - |z_r|^2$.

上述系数为± 1的模方和形式，称为厄米型的规范型.

定理8.4.4　C上任意一个n元厄米型都可通过非退化的线性替换化为如下形式的规范型（其中V为厄米型的秩）：

$$|t_1|^2 + |t_2|^2 + \cdots + |t_p|^2 - |t_{p+1}|^2 - \cdots - |t_r|^2$$

即：任意一个厄米型必定与一个规范型等价.

惯性定理，规范型的唯一性

规范型中，平方项的个数等于厄米型的秩，因此由厄米型唯一确定.

一个问题是，通过不同的方法化为规范型后，正负1的个数是不是也唯一由厄米型确定？若是，则规范型唯一.

定理8.4.5(惯性定理）一个秩为r的n元厄米型$f(x_1，x_2，\cdots，x_n)$，无论采用何种非退化的线性替代将其化为规范型,其中正项的个数p，负项的个数$\mathrm{r}-\mathrm{p}$都是唯一确定的！

注：由于上述定理，故p等于厄米型的矩阵的正本征值个数，$\mathrm{r}-\mathrm{p}$等于负本征值个数，r等于非零本征值的个数.

证明：（反证法）假设$X = C_1 Y$，$X = C_2 Z$分别化为：（C_1，C_1可逆）

$$\begin{cases} f(x_1，x_2，\cdots，x_n) = |y_1|^2 + |y_2|^2 + \cdots + |y_p|^2 - |y_{p+1}|^2 - \cdots - |y_r|^2. \\ f(x_1，x_2，\cdots，x_n) = |z_1|^2 + |z_2|^2 + \cdots + |z_q|^2 - |z_{q+1}|^2 - \cdots - |z_r|^2 \end{cases}$$

且$p \neq q$，不妨设$p > q$，由于$Z = {C_2}^{-1} C_1 Y$，

考虑关于$\mathrm{Y} = [y_1，y_2，\cdots，y_n]$的线性方程组：

$$\begin{cases} z_1 = 0 \\ z_2 = 0 \\ \quad\vdots \\ z_q = 0 \\ y_{p+1} = 0 \\ \quad\vdots \\ y_n = 0 \end{cases}$$

方程个数：$q+(n-p)=n+(q-p)<n$，未知元个数：n，故必有非零解，取其中之一为$\tilde{Y}$，可由$Z={C_2}^{-1}C_1Y$得到$\tilde{Z}={C_2}^{-1}C_1\tilde{Y}\neq 0$，将$\tilde{Y}$代入，可知二次型方程的值$f>0$，而将$\tilde{Z}$代入又有$f<0$，故矛盾！必有$p=q$，证毕.

定义：厄米型的规范型中正项个数 p 称该二次型的正惯性指数，负项个数$\mathrm{r}-\mathrm{p}$称该二次型的负惯性指数，$\mathrm{s}=\mathrm{p}-(r-p)=2p-r$ 称二次型的符号差，简称号差.

定理 8.2.5’（惯性定理） 任意一个 n 阶厄米矩阵 A 必定合同于一个如下的被称为规范型的矩阵：$\begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix}$，且规范型唯一.

注：其中$\mathrm{r}=\mathrm{r}(A)$，亦等于非零本征值的个数，而 p 等于正本征值的个数.

注：惯性定理亦可用标准型来表述：即无论用何种非退化的线性替换将厄米型化为标准型，其中正平方项的个数一定相同（负平方项个数亦相同）.因为如若不然的话，将这两种标准型再分别规范化，就与定理 8.2.4 矛盾.

推论 8.4.6 两个厄米型等价当且仅当它们有相同的秩和正惯性指数；两个厄米矩阵合同当且仅当它们有相同的秩和正惯性指数.

由上述定理，可将厄米型分类：同一类的厄米型等价：

推论 8.4.7　一切 n 元厄米型可分为$\frac{1}{2}(n+1)(n+2)$个类，属同一类的厄米型相互等价，不同类的厄米型不等价.

证：规范型$C_{r,\ p}=\begin{bmatrix} I_p & & \\ & -I_{r-p} & \\ & & 0 \end{bmatrix}$，r 可取$0$，$1$，$\cdots$，n，固定 r 后，p 可取$0$，$1$，$\cdots$，r，故共有$1+2+\cdots+(n+1)=\frac{1}{2}(n+1)(n+2)$中不同的规范型.

正定二次型的定义

一个 n 元厄米型 $f(x_1,\ x_2,\ \cdots,\ x_n)\ =\sum_{i,j} a_{ij}x_ix_j=x^+Ax$　可看成定义在 C 上的 n 个复变量的实值函数。

定义（正定厄米型，正定矩阵）　C 上的厄米型$f(x_1,\ x_2,\ \cdots,\ x_n)$，若对于任意一组不全为 0 的实数$C_1$，$C_2$，$\cdots$，$C_n$，恒有$f(C_1,\ C_2,\ \cdots,\ C_n)>0$，则称该厄米型为正定厄米型. 设 A 是一个厄米矩阵，若厄米型x^+Ax是正定的，则称 A 为正定矩阵.

注： 即：$\forall x\neq 0$，总有$x^+Ax>0$，类比内积$(\alpha,\ \alpha)>0$.

例 32　酉空间中的度规矩阵 A 是一个正定矩阵.

$(\ \forall\alpha\neq 0,\ (\alpha,\ \alpha)>0)$

正定厄米型的性质，及判定方法

定理 8.4.8　n 元厄米型$f(x_1,\ x_2,\ \cdots,\ x_n)$正定的充要条件是其正惯性指数$p=n$.

证明：$(\Longleftarrow)$　假设$p=n$，则由$X=CY$得：

$f(x)=g(x)=|y_1|^2+|y_2|^2+\cdots+|y_n|^2$，$\forall\tilde{X}\neq 0$，有$\widetilde{\mathrm{Y}}=C^{-1}\tilde{X}\neq 0$，代入有：

$f(\tilde{x})=g(\tilde{y})>0$，故正定.

$(\Longrightarrow)$ 设$f(x)$正定，则假设$p<n$，有：$X=CY$，

$f(x)=g(y)=|y_1|^2+\cdots+\left|y_p\right|^2-\left|y_{p+1}\right|^2-\cdots-|y_r|^2$ ，

取$\tilde{Y}=\left[0，0，\cdots，1\right]^T$，有：$\tilde{X}=C\tilde{Y}\neq 0$，代入得：$f(\tilde{x})=g(\tilde{y})\leq 0$，矛盾，故$p=n$. 证毕.

定理 8.4.8’ n 阶厄米矩阵正定的充要条件是其正惯性指数$p=n$.

推论 8.4.9 n 阶厄米矩阵正定的充要条件是其没有零本征值（满秩），且本征值全为正.

推论 8.4.10 n 阶厄米矩阵 A 正定的充要条件是其与I_nH 合同，或者说：存在可逆矩阵 D，使$A=D^+D$.

推论 8.4.11 若 A 为正定矩阵，则$|A|>0$.

证明：$|A|=\prod_{i=1}^{n}\lambda_i>0$ ，或$|A|=|D^+D|=\left||D|\right|^2>0$.

推论 8.4.11 不是充分条件，一个实用的判定方法是：

定理 8.4.12 厄米型$f(x)=x^+Ax$正定的充要条件是 A 的一切顺序主子式都大于零，即：

$A_1=a_{11}>0$，$A_2=\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}>0$，$A_3=\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}>0$，…，

$A_n=\begin{vmatrix}a_{11}&\cdots&a_{13}\\\vdots&\vdots&\vdots\\a_{31}&\cdots&a_{33}\end{vmatrix}=|A|>0$.

注： 主子式：取相同的行和列构成的子式.（子式：相同数目的行和列）顺序主子式：按顺序依次取相同的前个k行和列构成的子式.

证明：（必要性）设$f(x)=x^{+}Ax$正定，则可令

$X=\left[x_1，x_2，\cdots，x_k，0，0，\cdots，0\right]^T$代入$f(x)$得到一个$k$元的厄米型：$g\left(x_1，x_2，\cdots，x_k\right)=f(x)=x^{+}Ax$，由于$f$正定，故无论$X_k=\left[x_1，x_2，\cdots，x_k\right]^T$怎样选取$g(X_k)>0$，即$\mathrm{g}(X_k)$正定.

由推论 8.4.11 可知 k 阶顺序主子式$A_k>0$.

(注：同样的方法可证明若 A 正定，则 A 的一切主子式均大于 0)

（充分性）：设 A 的一切顺序主子式均大于零. 对阶数 n 用数归.

$\mathrm{n}=1$时，$a_{11}>0$，显然$a_{11}\overline{x_1}x_1$正定，充分性成立.

假设 n-1 元厄米型充分性成立，则对 n 元次型$f(x)=x^{+}Ax$. 令

$\mathrm{A}=\begin{bmatrix}A_{n-1} & b\\ b & a_{nn}\end{bmatrix}$，则由归纳假设$A_{n-1}$正定，故存在

$Q_{n-1}^{+}A_{n-1}Q_{n-1}=I_{n-1}$, 取 Q$=\begin{bmatrix}Q_{n-1} & Y\\ 0 & 1\end{bmatrix}$，有

$$Q^{+}A\,Q=\begin{bmatrix}Q_{n-1}^{+}A_{n-1}Q_{n-1} & Q_{n-1}^{+}A_{n-1}Y+Q_{n-1}^{+}b\\ Y^{+}A_{n-1}Q_{n-1}+b^{+}Q_{n-1} & Y^{+}A_{n-1}Y+Y^{+}b+b^{+}Y+a_{nn}\end{bmatrix},$$

取$Y=-A_{n-1}^{-1}\ b$可知: $Q^{+}A\,Q=\begin{bmatrix}I_{n-1} & 0\\ 0 & a_{nn}-b^{+}A_{n-1}^{-1}b\end{bmatrix}$, 由于$|A|>0$，故$\left|Q^{+}A\,Q\right|>0$，即：$a_{nn}-b^{+}A_{n-1}^{-1}b>0$，故$Q^{+}A\,Q$正定，从而 A 正定，证毕.

负定、准正定、准负定、不定二次型

定义：C 上的厄米型$f\left(x_1，x_2，\cdots，x_n\right)=x^{+}Ax$，若对于任意一组不全为 0 的复数$C_1，C_2，\cdots，C_n$，恒有：

（1）$f\left(C_1，C_2，\cdots，C_n\right)<0$，则称该厄米型为负定厄米型，称 A 为负定矩阵.

（2）$f\left(C_1，C_2，\cdots，C_n\right) \geq 0$，则称该厄米型为准正定厄米型，称 A 为准正定矩阵.

（3）$f\left(C_1，C_2，\cdots，C_n\right) \leq 0$，则称该厄米型为准负定厄米型，称 A 为准负定矩阵. 若$f(x)$既不是准正定的，也不是准负定的，则称为不定厄米型.

显然一个厄米型$f(x) = x^{+}Ax$是负定的，当且仅当厄米型

$-f(x) = x^{+}(-A)x$是正定的. 由此根据正定厄米型的定理直接可得：

定理 8.4.13 设$f\left(x_1，x_2，\cdots，x_n\right) = x^{+}Ax$是一个 n 元厄米型，则下列论述等价：

（5）$f(x)$是负定的；

（6）$f(x)$的负惯性指数是 n;

（7）$f(x) = x^{+}Ax$中矩阵 A 的 n 个本征值全部小于零；

（8）$f(x) = x^{+}Ax$中矩阵 A 的顺序主子式A_k满足：

$A_1 = a_{11} < 0$，$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$，$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0$，$\cdots$，

即：$(-1)^k A_k > 0$.

对准正定二次型有如下类似的定理：

定理 8.4.14 设$f\left(x_1，x_2，\cdots，x_n\right) = x^{T}Ax$是一个 n 元厄米型，则下列论述等价：

（1）$f(x)$是准正定的；

（2）$f(x)$的正惯性指数等于秩；

（3）$f(x) = x^{+}Ax$中矩阵 A 的本征值全部大于或等于零；

（4）$f(x) = x^{+}Ax$中矩阵 A 的所有主子式都大于或等于零：

注： 尤其要注意第（4）条与正定的区别：要所有主子式，不仅是顺序主子式.

证明：$(1) \Leftrightarrow (2)$ 与定理 8.4.8 的证明类似，只须将证明中的 n 改为 r，并将>改为≥即可.

$(2) \Leftrightarrow (3)$ 显然等价，因为非零本征值个数等于秩，而正本征值个数等于正惯性指数.

$(1) \Leftrightarrow (4)$的证明：

$(1) \Rightarrow (4)$ 设$f(x_1, x_2, \cdots, x_n) = x^+Ax$准正定，则任取 k 个$x_i$非零，其余的$x_i$均为 0，代入得$f(x_1, x_2, \cdots, x_n) = g(x_{i1}, x_{i2}, \cdots, x_{ik})$可知：$k$元厄米型$g$必准正定（恒大于或等于 0）. 故由第（3）条的推论可知，g的矩阵行列式必大于或等于 0，而g的矩阵行列式正好是 A 的 k 阶主子式. 故 A 的任意k阶主子式均大于零$(k = 1,2,\cdots,n)$.

$(4) \Rightarrow (1)$ 设 A 的所有主子式均大于零，则构造$B = \lambda I + A(\lambda > 0)$

令A_k和B_k分别是 A 和 B 的 k 阶顺序主子式构成的矩阵，于是：

$$|B_k| = |\lambda I_k + A_k| = \begin{vmatrix} \lambda + a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & \lambda + a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k1} & a_{k2} & \cdots & \lambda + a_{kk} \end{vmatrix}$$

$$= \lambda^k + p_1\lambda^{k-1} + \cdots + p_{k-1}\lambda + p_k$$

其中p_i为A_k中一切i级主子式之和，由题设可知 A 的一切主子式均大于或等于 0，故$p_i \geq 0$，故$\lambda > 0$时$|B_k| > 0$，$B = \lambda I + A$为正定矩阵.

假设 A 不是准正定的，则存在$\tilde{x} \in R^n$使$\tilde{x}^+A\tilde{x} = -C < 0$，那么，取$\lambda = C/\tilde{x}^T\tilde{x} > 0$，则有：$\tilde{x}^+B\tilde{x} = \tilde{x}^+(\lambda I + A)\tilde{x} = \lambda\tilde{x}^+x - C = C - C = 0$，矛盾，故 A 必为准正定矩阵，证毕.

注意到厄米型$f(x)$为准负定的充要条件是$-f(x)$为准正定，故可得到与上述定理类似的结论.

例 33 若厄米矩阵 A 可逆，求证A^{-1}与AH合同.

证：$A = C^{+}FC$ 其中$C^{+}C = I$，故

$A^{-1} = C^{+}F^{-1}C = C^{+}F^{-1}FF^{-1}C = (F^{-1}C)^{+}F(F^{-1}C)$

例 34 求证：（1）数域 F 上的 n 阶反厄米矩阵必与如下形式矩阵 H 合同：

$H = i\begin{bmatrix} I_p & & \\ & -I_{r-p} & \\ & & 0 \end{bmatrix}$. 提示：A 反厄米，则$(iA)^{+} = iA$，厄米！

例 35 若A，B均为同阶正定矩阵，$k, l \in R^{+}$, 求证$kA + kB$正定.

证明：用正定的定义：$\forall x \in C^{n} \neq 0$, 有：

$x^{+}(kA + lB)x = k(x^{+}Ax) + l(x^{+}Bx) > 0$, 证毕.

例 36 设 A 为正定矩阵，求证A^{-1}，kA，A^{m}，A^{*}均正定（$k \in R^{+}$，$m \in Z^{+}$）.

证：由本征值来判定！注意$A^{*} = |A|A^{-1}$.

例 37 A 为厄米矩阵，求证：（1）当ε充分小时，$I + \varepsilon\mathrm{A}$正定；（2）当 t 充分大时，$\mathrm{t}I + \mathrm{A}$正定. 提示：$I + \varepsilon\mathrm{A}$和$\mathrm{t}I + \mathrm{A}$的本征值分别为$1 + \varepsilon\lambda$和$t + \lambda$.

例 38 设 A 为一正定厄米矩阵，求证存在另一个正定矩阵 S 使$A = S^{2}$.

证：注意到厄米矩阵必可相似对角化，故必有：

$\mathrm{A} = P^{-1}FP$，$F = diag[\lambda_1，\cdots，\lambda_n]$，A 正定，故$\lambda_i > 0$，故$F = \tilde{F}^{2}$，其中$\tilde{F} = drag[\sqrt{F_1}，\cdots，\sqrt{F_n}]$，即：

$A=P^{-1}\tilde{F}^2P=P^{-1}\tilde{F}PP^{-1}\tilde{F}P=\left(P^{-1}\tilde{F}P\right)^2$

例 39 设$f(x)=x^{+}Ax$，$g(x)=x^{+}Bx$是两个厄米型且g正定，求证存在一个非退化的线性替换$X=CY$分别将f和g化为

$\lambda_1|y_1|^2+\cdots+\lambda_n|y_n|^2$和$|y_1|^2+\cdots+|y_n|^2$（同时标准化）.

证：首先，存在C_1使${C_1}^{+}BC_1=I_n$，此时${C_1}^{+}AC_1=A'$，A'厄米，故存在正交矩阵C_2，使${C_2}^{+}A'C_2=F$，此时${C_2}^{+}I_nC_2={C_2}^{-1}I_nC_2=I_n$，综合上述，令$C=C_1C_2$，则$\begin{cases} C^{+}AC={C_2}^{+}{C_1}^{+}AC_1C_2={C_2}^{+}A'C_2=F \\ C^{+}BC={C_2}^{+}{C_1}^{+}BC_1C_2={C_2}^{-1}I_nC_2=I_n \end{cases}$.

例 40 设 A, B 为同阶的正定和非正定矩阵，且$B\neq 0$，求证

$|A+B|>|A|+|B|$.

证：类似上题，可证存在 C 可逆，使：

$C^{+}AC=I$，$C^{+}BC=F=diag[\lambda_1，\cdots，\lambda_n]$，$\lambda_i>0$不全为零，故

$|A+B|=\left|{C^{-1}}^{+}C^{+}\ (A+B)CC^{-1}\right|=\left||C^{-1}|\right|^2|I+F|$

$>\left||C^{-1}|\right|^2(|I|+|F|)$

$=\left|{C^{-1}}^{+}IC^{-1}\right|+\left|{C^{-1}}^{+}FC^{-1}\right|=|A|+|B|$.

例 41 设 A 是 n 阶正定厄米矩阵，求证$f\left(x_1，x_2，\cdots，x_n\right)=\begin{vmatrix} A & X \\ X^{+} & 0 \end{vmatrix}$是负定矩阵.

证：$f(\mathrm{x})=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & x_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n \\ \overline{x_1} & \overline{x_2} & \cdots & \overline{x_n} & 0 \end{vmatrix}$，令$M_{ij}$为 A 的$a_{ij}$的余子式，$A_{ij}$为 A 的$a_{ij}$的代数余子式，故$|x_i|^2$的系数为

$(-1)^{n+1+i}(-1)^{n+i}M_{ii}=-(-1)^{i+i}M_{ii}=A_{ii}$，$\overline{x_i}x_j$的系数为

$(-1)^{n+1+i}(-1)^{n+j}\left(M_{ij}\right)=-A_{ij}$，故$f(\mathrm{x})=-X^{+}A^{*}X$

（注：这一点亦可用另一方法来证：

$\begin{bmatrix} A & X \\ X^+ & 0 \end{bmatrix}\begin{bmatrix} I & -A^{-1}X \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A & 0 \\ X^+ & -X^+A^{-1}X \end{bmatrix}$），取行列式即：

$f(\mathrm{x}) = -|A|X^+A^{-1}X = -X^+A^*X$，由于 A 正定，故$A^*$正定，故$f(\mathrm{x})$负定，证毕.

例 42 厄米矩阵 A 正定，求证$|A| \le a_{nn}|A_{n-1}|$，其中$|A_{n-1}|$是 A 的 n-1 阶顺序主子式.

证：A 正定，故A_{n-1}亦正定，故$\begin{bmatrix} A_{n-1} & X \\ X^+ & 0 \end{bmatrix}$负定.

$$\text{而}A = \begin{vmatrix} a_{11} & \cdots & a_{1n-1} & a_{1n} \\ \vdots & & \vdots & \vdots \\ a_{n-11} & \cdots & a_{n-1n-1} & a_{n-1n} \\ \overline{a_{n1}} & \cdots & a_{nn-1} & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n-1} & 0 \\ \vdots & & \vdots & \vdots \\ a_{n-11} & \cdots & a_{n-1n-1} & 0 \\ \overline{a_{n1}} & \cdots & \overline{a_{nn-1}} & a_{nn} \end{vmatrix}$$

$$\le a_{nn}|A_{n-1}|.$$

例 43 若厄米矩阵 A 正定，求证$|A| \le a_{11}a_{22}\cdots a_{nn}$.

证：反复利用上一例题的结论.

例 44 设 A 为 n 阶厄米矩阵，求证：$|A|^2 \le \prod_{i=1}^{n}(|a_{1i}|^2 + |a_{2i}|^2 + \cdots + |a_{ni}|^2)$

证：$|A| = 0$时显然成立，$|A| \ne 0$时，A^+A为正定矩阵，利用上一题的结论有$|A|^2 = |A^+A| \le (A^+A)_{11}(A^+A)_{22}\cdots(A^+A)_{nn}$，证毕.

例 45（矩阵的 SVD 分解---奇异值分解）：设$\mathrm{A} \in C^{m\times n}$，且$r(A) = r$，求证：存在幺正矩阵$\mathrm{U} \in C^{m\times n}$和$\mathrm{V} \in C^{n\times n}$，使得：$\mathrm{A} = \mathrm{U}\begin{bmatrix} \Delta & 0 \\ 0 & 0 \end{bmatrix}V^+$，

其中$\Delta = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}$，$\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$称作 A 的奇异值.

证：首先，$r(A^+A) = r(A) = r$，其次，A^+A是厄米矩阵：$(A^+A)^+ = A^+A$

最后，A^+A非正定，因为$\forall$有$x^+A^+Ax=(Ax)^+(Ax)\geq 0$，故存在幺正

的 V,使$V^+A^+AV=\begin{bmatrix}\lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0\end{bmatrix}$，令$\sigma_i=\sqrt{\lambda_i}$

令$V=[\varepsilon_1，\varepsilon_2，\cdots，\varepsilon_r，\varepsilon_{r+1}，\cdots，\varepsilon_n]=[V_1，V_2]$,

则$(AV)^+(AV)=\begin{bmatrix}\Delta^2 & 0\\ 0 & 0\end{bmatrix}$，故$A\varepsilon_{r+1}=A\varepsilon_{r+2}=\cdots=A\varepsilon_n=0$，即：

$AV_2=0$. 又$V^+A^+AV=\begin{bmatrix}{V_1}^+\\ {V_2}^+\end{bmatrix}A^+A[V_1，V_2]=\begin{bmatrix}{V_1}^+A^+AV_1 & {V_1}^+A^+AV_2\\ {V_2}^+A^+AV_1 & {V_2}^+A^+AV_2\end{bmatrix}$

故${V_1}^+A^+AV_1=\Delta^2$，即：$\Delta^{-1}{V_1}^+A^+AV_1\Delta^{-1}=I_r$. 令$U_1=AV_1\Delta^{-1}$，故

${U_1}^+U_1=I_r$, U_1 是$m\times r$的矩阵，r 个列向量构成C^m中的标准正交组，

可将其扩充为C^m中的标准正交基，得 m 阶方阵$U=[U_1，U_2]$，故U_2的

列向量与U_1的正交，即${U_2}^+U_1=0$，此时有：

$$U^+AV=\begin{bmatrix}{U_1}^+\\ {U_2}^+\end{bmatrix}A[V_1，V_2]=\begin{bmatrix}{U_1}^+AV_1 & {U_1}^+AV_2\\ {U_2}^+AV_1 & {U_2}^+AV_2\end{bmatrix}=\begin{bmatrix}{U_1}^+U_1\Delta & 0\\ {U_2}^+U_1\Delta & 0\end{bmatrix}=$$

$\begin{bmatrix}\Delta & 0\\ 0 & 0\end{bmatrix}$，证毕.

矩阵做 SVD 分解后，对应的线性映射在选取标准正交基之后会变得相对简单---SVD 分解的用途.

§8.5 本征值问题的极值性质

本征值问题的极值性　极大-极小值原理　　本征值的极值性所得的一般结论　　本征向量组的完备性

本征值问题的极值性

本节讨论线性空间中厄米变换的本征值问题与厄米型的极值问题的等价性（所有的讨论限制在实空间上即对称变换的相应结论）.

有限维线性空间中，下文将要讨论的性质都是显然的，因此本节主要讨论无穷维空间.

定理 8.5.1（本征值的极值性） 设 V 是一酉空间（可以是无穷维）H 是 V 上的一个厄米变换（算符），设 H 的本征值有下限，并依次排列如下：$\lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots (H\alpha_i = \lambda_i\alpha_i)$，则$\left(\alpha，H\alpha\right)/\left(\alpha，\alpha\right)$的极小值是：（1）$\lambda_0$，若$\alpha$是 V 中任意非零向量；

（2）λ_1，若α是 V 中任意非零向量，且满足：$\left(\alpha_0，\alpha\right) = 0$

（3）λ_m，若α是 V 中任意非零向量，且满足：

$\left(\alpha_0，\alpha\right) = \left(\alpha_1，\alpha\right) = \left(\alpha_2，\alpha\right) = \cdots = \left(\alpha_{m-1}，\alpha\right) = 0$

注： 若 V 是有限维酉空间，则上述定理显然成立. 因为 n 维时，

$V \cong C^n$，而 H 对应于一个 n 阶厄米矩阵. 厄米矩阵的本征向量构成了C^n的一个基（n 个，线性无关），故 H 的全体本征向量可构成 V 的一个标准正交基，即：$\left\{\alpha_0，\alpha_2，\cdots，\alpha_{n-1}\right\}$构成 V 的标准正交基.

$\forall\alpha \in V$, 有：$\alpha = \sum_{i=0}^{n-1} x_i\alpha_i$

故$\frac{(\alpha，H\alpha)}{(\alpha，\alpha)} = \lambda_0\frac{|x_0|^2}{X^TX} + \lambda_1\frac{|x_1|^2}{X^TX} + \cdots + \lambda_{n-1}\frac{|x_{n-1}|^2}{X^TX}$

$= \lambda_0|y_0|^2 + \lambda_1|y_1|^2 + \cdots + \lambda_{n-1}|y_{n-1}|^2$ ①

其中$|y_0|^2+|y_1|^2+\cdots+|y_{n-1}|^2=1$ ②，显然结论（1）成立，

对（2），①和②两式的求和中无λ_0项，故亦成立；

对（3），①和②两式的求和均从 m 开始到 n-1，故亦成立.

对无穷维，H 的本征向量是否完备待证，故上述论述不成立.

证明：（1）$\frac{(\alpha,\ H\alpha)}{(\alpha,\ \alpha)}$是$\mathrm{V}\to\mathrm{R}$的一个映射，$\forall\alpha\in V$，有一实数与之对应，

这是一个泛函，令$\lambda=\frac{(\alpha,\ H\alpha)}{(\alpha,\ \alpha)}$，则$\lambda=\lambda[\alpha]$.假设$\lambda[\alpha]$在$\alpha=\tilde{\alpha}$处取极值，

则令$\alpha=\tilde{\alpha}+\varepsilon\beta$，$\varepsilon\in\mathrm{R}$，$\beta\in V$，$\lambda[\tilde{\alpha}+\varepsilon\beta]$对任意给定的$\beta$来说，是$\varepsilon$的一个函数，且应在$\varepsilon=0$处取极值，故$d\lambda/d\,\varepsilon\big|_{\varepsilon=0}=0$，而：

$$d\lambda\big|_{\varepsilon=0}=\lambda[\tilde{\alpha}+d\varepsilon\beta]-\lambda[\tilde{\alpha}]$$

$$=\frac{\left(\tilde{\alpha}+d\varepsilon\beta,\ H(\tilde{\alpha}+d\varepsilon\beta)\right)}{(\tilde{\alpha}+d\varepsilon\beta,\ \tilde{\alpha}+d\varepsilon\beta)}-\frac{(\tilde{\alpha},\ H\tilde{\alpha})}{(\tilde{\alpha},\ \tilde{\alpha})}$$

$$=\frac{(\tilde{\alpha},\ H\tilde{\alpha})+d\varepsilon(\beta,\ H\tilde{\alpha})+d\varepsilon(\tilde{\alpha},\ H\beta)}{(\tilde{\alpha},\ \tilde{\alpha})+d\varepsilon(\beta,\ \tilde{\alpha})+d\varepsilon(\tilde{\alpha},\ \beta)}-\frac{(\tilde{\alpha},\ H\tilde{\alpha})}{(\tilde{\alpha},\ \tilde{\alpha})}$$

$$=\frac{1}{(\tilde{\alpha},\ \tilde{\alpha})}\left[\left(\beta,\ (H-\tilde{\lambda})\tilde{\alpha}\right)+\left((H-\tilde{\lambda})\tilde{\alpha},\ \beta\right)\right]d\varepsilon$$

其中$\tilde{\lambda}=\frac{(\tilde{\alpha},\ H\tilde{\alpha})}{(\tilde{\alpha},\ \tilde{\alpha})}$，由于$\lambda[\alpha]$在$\alpha=\tilde{\alpha}$处取极值，故$\frac{d\lambda}{d\varepsilon}\big|_{\varepsilon=0}$，即对$\forall\beta\in V$，

均有：$\left(\beta,\ (H-\tilde{\lambda})\tilde{\alpha}\right)+\left((H-\tilde{\lambda})\tilde{\alpha},\ \beta\right)=0$，

取$\beta=(H-\tilde{\lambda})\tilde{\alpha}$代入即得：$(H-\tilde{\lambda})\tilde{\alpha}=0$，即：$H\tilde{\alpha}=\tilde{\lambda}\tilde{\alpha}$，即：$\lambda[\alpha]$在 H 的本征向量处取极值，而 H 的本征值为$\lambda_0\le\lambda_1\le\lambda_2\le\cdots$，故$\lambda[\alpha]$的极小值为$\lambda$.

由于$\frac{(\alpha，H\alpha)}{(\alpha，\alpha)}=\left(\frac{\alpha}{|\alpha|}，H\frac{\alpha}{|\alpha|}\right)$,故极小值可转换成条件极值,即:$\left(\alpha，H\alpha\right)$在$\left(\alpha，\alpha\right)=1$的约束之下的极值问题.从这个角度看,（1）要证的是:

$\forall\beta\in V$，$\lambda(\varepsilon)=\left(\tilde{\alpha}+\varepsilon\beta，H(\tilde{\alpha}+\varepsilon\beta)\right)$在约束条件

$\left(\tilde{\alpha}+\varepsilon\beta，(\tilde{\alpha}+\varepsilon\beta)\right)=1$之下，在$\varepsilon=0$处取极值.由拉格朗日乘子法，

即：$F(\varepsilon)=\lambda(\varepsilon)+k\left[\tilde{\alpha}+\varepsilon\beta，H(\tilde{\alpha}+\varepsilon\beta)-1\right]$的极值在$\varepsilon=0$，于是：

$\frac{dF}{d\varepsilon}\Big|_{\varepsilon=0}=\left(\beta，(H-k)\tilde{\alpha}\right)+\left(\tilde{\alpha}，(H-k)\beta\right)$，对于$\forall\beta\in V$均为零，

故$H\tilde{\alpha}=k\tilde{\alpha}$，即$\lambda[\alpha]$在$\alpha$为 H 的本征值处取极值，而 H 最小值为$\lambda$. 故得证.

（2）可转化为$\lambda[\alpha]=\left(\alpha，H\alpha\right)$在条件$\left(\alpha，\alpha\right)=1$，$\left(\alpha_0，\alpha\right)=0$之下的极值，$\forall\beta\in V$，总可选取 t 令$\left(\alpha_0，\beta-t\alpha\right)=0$，记$\beta-t\alpha_0=\tilde{\beta}$，

假设$\lambda[\alpha]$在$\alpha=\tilde{\alpha}$处取极值，$\left(\tilde{\alpha}，\tilde{\alpha}\right)=1$，$\left(\alpha_0，\tilde{\alpha}\right)=0$，

则$\lambda(\varepsilon)=\left(\tilde{\alpha}+\varepsilon\beta，H(\tilde{\alpha}+\varepsilon\beta)\right)$

对ε来说在约束$\left(\tilde{\alpha}+\varepsilon\beta，(\tilde{\alpha}+\varepsilon\beta)\right)=1$之下，

在$\varepsilon=0$处取极值.同（1）的分析，必有：$\left(\beta，(H-k)\tilde{\alpha}\right)+\left(\tilde{\alpha}，(H-k)\beta\right)=0$，而

$\left(\alpha_0，(H-k)\tilde{\alpha}\right)=\left((H-k)\alpha_0，\tilde{\alpha}\right)=\left((\lambda_0-k)\alpha_0，\tilde{\alpha}\right)=0$，故

$\left(\beta，(H-k)\tilde{\alpha}\right)+\left(\tilde{\alpha}，(H-k)\beta\right)=0$，仍对$\forall\beta\in V$成立，于是

$H\tilde{\alpha}=k\tilde{\alpha}$，而$\left(\alpha_0，\tilde{\alpha}\right)=0$，故（2）亦得证，（3）类似.

极大—极小值原理

定理 8.5.2（极大—极小值原理） 设 H 是酉空间 V 中的厄米算符，其本征值有下限，依次为$\lambda_0\le\lambda_1\le\lambda_2\le\cdots$ $(H\alpha_i=\lambda_i\alpha_i)$，则

（1）令$\beta \in V$为给定向量，定义$F(\beta)$是$\frac{(\alpha，H\alpha)}{(\alpha，\alpha)}$的极小值，其中$\alpha$是满足$(\beta，\alpha)=0$的任意向量.显然$F(\beta)$依赖$\beta$的选取，求证通过改变$\beta$的选取，$F(\beta)$的极大值是 H 的第二个本征值$\lambda_1$.

（2）令$\beta_1，\beta_2，\cdots，\beta_n \in V$是 n 个给定的线性无关向量，定义$F[\beta_1，\beta_2，\cdots，\beta_n]$是$\frac{(\alpha，H\alpha)}{(\alpha，\alpha)}$的极小值，其中$\alpha$是满足

$(\beta_1，\alpha)=\cdots=(\beta_n，\alpha)=0$的任意向量. 求证通过改变$\beta_i$的选取，$F[\beta_1，\beta_2，\cdots，\beta_n]$的极大值时 H 的第$\mathrm{n}+1$个本征值$\lambda_n$.

注：同定理 8.5.1，此处亦等于$(\alpha，\alpha)=1$之下$(\alpha，H\alpha)$的极值问题.

证明：显然（1）是（2）的特殊情况（$\mathrm{n}+1$），下面直接证明（2）.

若取$\beta_1=\alpha_0，\beta_2=\alpha_1，\cdots，\beta_n=\alpha_{n-1}$，则由定理 8.5.1，有：

$F[\beta_1，\beta_2，\cdots，\beta_n]=F[\alpha_0，\alpha_1，\cdots，\alpha_{n-1}]=\lambda_n$，故只需证明其它的$\beta_i$选取情况下，$F[\beta_1，\beta_2，\cdots，\beta_n]$永远不大于$\lambda_n$，亦即在其它的$\beta_i$选取情况下，总可找到一个满足条件：$(\alpha，\alpha)=1$，

$(\beta_1，\alpha)=(\beta_2，\alpha)=\cdots=(\beta_n，\alpha)=0$的$\alpha$（记作$\tilde{\alpha}$），使$(\tilde{\alpha}，H\tilde{\alpha})$的值不大于$\lambda_n$（由于$F[\beta_i]$是满足上述条件的$\alpha$之下$(\alpha，H\alpha)$的极小值，故必定不大于$(\tilde{\alpha}，H\tilde{\alpha})$，故不大于$\lambda_n$）为此，构造 $\tilde{\alpha}=\sum_{i=0}^{\mathrm{n}} k_i\alpha_i$

由$(\beta_1，\alpha)=(\beta_2，\alpha)=\cdots=(\beta_n，\alpha)=0$，给出$\mathrm{n}+1$个$k_i$满足的 n 个方程，必定有非零解. 归一化之后即有$|k_0|^2+|k_1|^2+\cdots+|k_n|^2=1$，此时$\tilde{\alpha}$满足条件，而$(\tilde{\alpha}，H\tilde{\alpha})=|k_0|^2\lambda_0+|k_1|^2\lambda_1+\cdots+|k_n|^2\lambda_n \le \lambda_n$，证毕.

本征值的极值性质所得的一般性结论：

根据极大—极小值原理 8.5.2，求本征值的问题可转化为求一系列的极大—极小值.而当任意选取一组向量β_i后，若我们加强对向量α可取范围的限制，则得到的极小值$F\left[\beta_1，\beta_2，\cdots，\beta_n\right]$必定不小于原来的情况.于是改变向量组$\beta_i$的选取，这些极小值中的极大值也必定不小于原来情况.由于这些极大—极小值对应于厄米算符的由小到大排列的所有本征值，因此可得到一般的结论如下：

若加强对向量可取范围的限制后，则厄米算符的本征值必定不小于原来情况下相应的本征值；反之，若放宽向量可取范围的限制，则厄米算符的本征值必定不大于原来情况下相应的本征值.

即：考虑厄米算符 H 的本征值问题，假设所有的向量均受到约束 C，相应地，H 的本征值和本征向量变为：$\lambda_i \to {\lambda_i}'$，$\alpha_i \to {\alpha_i}'$，则必有：

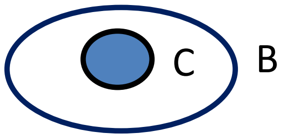
${\lambda_0}' \geq \lambda_0$，${\lambda_1}' \geq \lambda_1$，$\cdots$，${\lambda_i}' \geq \lambda_i$，$\cdots$

思考：为什么不能直接由极值性质 8.5.1 来得到上述结论？

第一个本征值的结论可直接由 8.5.1 得到，但考虑后面的本征值时，由于本征向量α_0在改变α的可取范围后还不相同了，故考虑λ_1的变化时，α的取值范围并不就是原始情况下α可取范围的子集。

例 46 考虑边界 B 固定的膜的振动，具本征频率ω_κ由方程：$-\nabla^2\phi = \omega_n^2\phi$来确定振幅$\phi$在边界 B 上等于 0（从而保证了运算符$-\nabla^2$是一个厄米算符$\int_\Omega d\omega = \int_{\alpha\Omega}\omega \to \int_{\alpha\Omega} d\vec{\iota} = \int_\Omega (d\vec{s} \times \vec{\nabla})$，故$\int_\Omega \phi_1 \nabla^2\phi_2 dx_1 dx_2 + \int_\Omega \nabla\phi_1 \nabla\phi_2 dx_1 dx_2 = \int_{\alpha\Omega} \phi_1\alpha_1\phi_2 dx_2 - \phi_1\alpha_2\phi_2 dx_1$）将这些频率依次排为$\omega_0 \leq \omega_1 \leq \omega_2 \leq \cdots$。若对ϕ加上约

束使ϕ在膜内一闭合曲线 C 中为零，则相应的ω_i变成ω_i'，对所有 i,有 $\omega_i' \geq \omega_i$.



本征向量组的完备性：

对有限维酉空间，厄米变换与一个有限维的厄米矩阵相对应。由于n阶厄米矩阵的本征向量组可构成C^n的一个标准正交基，因此是完备的（任何一个向量可用该向量组线性展开），从而有限维空间中厄米变换的本征向量组也是完备的。

对于无穷维酉空间，厄米变换的本征向量的正交性证明与有限维相同。值得研究的问题是，此时本征向量组是否仍然完备？

一个向量组$\{\alpha_i\}$ $(i = 0,1,2, \dots + \infty)$使得：

$\lim_{m\to\infty}(R_m, R_m) = 0$，其中：$R_m = \alpha - \sum_{i=0}^{m} c_i\alpha_i$（余项）

按照上定义，虽然有$m \to \infty$时，$\lim_{n\to\infty} R_m = 0$(零向量)，故$\forall\alpha \in V$均有：$\alpha = \sum_{i=0}^{+\infty} c_i\alpha_i$，即可用$\{\alpha_i\}$线性展开！

推论：$\lim_{m\to\infty} c_m \to 0$，若$\alpha_i$是单倍向量的话。(注意到$R_{m-1} - R_m = C_m\alpha_m$即得！)

定理 8.5.3（完备性） 设 H 是无穷维酉空间 V 的厄米算符，若 H 的本征值有下限无上限，即本征值按从小到大排列为：

$$\lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_i \leq \cdots$$

且$\lim_{n\to\infty} \lambda_n = +\infty$，则 H 的本征向量组$\{\alpha_i\}(i = 0,1,2, \dots, +\infty)$是完备的。

证：不失一般性，可设$\lambda_0 > 0$（若$\lambda_0 \le 0$，总可得$H' = H + cI$，选择足够大的 c，可使H'的本征值均大于零。代替H，我们考虑H'），对$\forall \alpha \in V$,选取$C_i = (\alpha_i, \alpha)$ ，则令：$R_m = \alpha - \sum_{i=0}^{m} c_i \alpha_i$

R_m满足：$(\alpha_0, R_m) = (\alpha_1, R_m) = \cdots = (\alpha_m R_m) = 0$，故根据定理 8.5.1 有：$\frac{(R_m, HR_m)}{(R_m, R_m)} \ge \lambda_{m+1} \ge \lambda_m$ 。故当 $m \to \infty$ 时，$\lambda_m \to +\infty$，故 $\frac{(R_m, HR_m)}{(R_m, R_m)} \to +\infty$。

另一方面，$\left(R_m, HR_m\right) = (\alpha - \sum_i c_i d_i, H\alpha - \sum_i c_i \lambda_i d_i)$

$=(\alpha, H\alpha) - \sum_i \bar{c}_i c_i \lambda_i - \sum_i \bar{c}_i c_i \lambda_i + \sum_i \bar{c}_i c_i \lambda_i$

$= (\alpha, H\alpha) - \sum_{i=0}^{m} |c_i|^2 \lambda_i \le (\alpha, H\alpha)$

故：$(R_m, R_m) \le \frac{1}{\lambda_m}(R_m, HR_m) \le \frac{1}{\lambda_m}(\alpha, H\alpha)$

$\forall \alpha \in V$，$(\alpha, H\alpha) \in R$为一与 m 无关的实数，故当$m \to \infty$时，$\lambda_m \to +\infty$ ，从而$(R_m, R_m) \to 0$

即：$\lim_{m\to\infty}(R_m, R_m) = 0$　　$\forall \alpha \in V$均成立，证毕。

例 47 $[-\pi, \pi]$上的周期性函数构成线性空间。$-\frac{d^2}{dx^2}$是一厄米算符，基本征值方程$-\frac{d^2u}{dx^2} = \lambda u$ 给出，

$\lambda = 0,1,2, \ldots k, \ldots + \infty u_\lambda = 1, e^{\pm ix}, e^{\pm 2ix}, \ldots e^{\pm ikx}, \ldots$

由傅立叶级数可知$\left\{e^{\pm iny}, n = 0,1,2, \ldots, +\infty\right\}$构成$[-\pi, \pi]$上的完备函数组，与定理 8.5.3 结论一致！

例 48　S-L 型本征值问题，可证 S-L 算符本征值无上限（希尔伯特书），从而具本征函数组构成一个完备组（《数理方法》（二）课程）

关于基于泛函数极值的本征问题的更多讨论，参见柯朗、希尔伯特《数学物理方法》卷一第 6 章。（例如本征值的无限增大，本征值的渐进分步等）

本节讨论了厄米型（一类泛函）的极值问题，另一类重要的泛函极值问题是由定积分定义的函数到数域的映射（函数的函数）这一类泛函的极值问题，这类问题的讨论即变分法的主要内容。

$$\mathrm{J[y(x)]} = \int_{x_0}^{x_1} F(x, y(x), y'(x))dx$$

变分的基本问题：求曲线的泛函的最大值或最小值，这些泛函是由一些定积分来表述的。（见第九章）

第九章、变分学

§9.1 泛函和变分的意义

函数的极值：f($x_{1,}$，$x_{2,}$，$\cdots x_n$)，必要条件 df=0 或f_{xi}=0(i=1,2⋯,n).

泛函（函数的函数）：f(x) →J 映射（J∈R）.例如：

垂直的 x-y 平面上，A(x_0，y_0)，B(x_0，y_0)为两给定点，

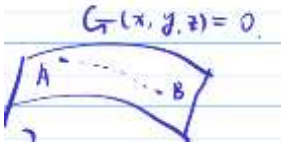
粒子沿过 A,B 的光滑曲线从 A 自由下落到 B 的时间 J,是曲线的函数。

$$J[y(x)] = \int dt = \int \frac{ds}{v} = \int_A^B \frac{\sqrt{Hy'^2}}{\sqrt{2gy}} dx = \sqrt{\frac{1}{2g}} \int_{x_0}^{x_1} \sqrt{\frac{1+y'^2}{y}} dx$$

要描述到 B 点的时间，不是知道 B 点的状态就足够，而是需要对从 A 到 B 的整个历史都积分。凡事具有历史性的现象，例如滞后效应、金属性质等物理现象，以及国家经济状况、人口状况等社会现象，都要对过去的历史进行积分才能知道它们现在的情况；也就是说，需要泛函来描述，这就是泛函的意义。

变分法的意义在于：研究泛函的极值问题

例 1 最速降线问题：在所有断点固定在 A、B 的光滑的微曲面中，沿哪一条下落时所用的时间最短？



例 2 曲面上的测地线问题：G(x,y,z)=0 为一给定的曲面， A(x_0，y_0)和 B(x_0，y_0)为曲面上两给定点，问曲面上连接 A、B 的最短曲线是哪条？

设曲线的方程为$\begin{cases} y = y(x) \\ z = z(x) \end{cases}$ ，由于端点必须是 A,B. 故$\begin{cases} y(x_0) = y_0 \\ z(x_0) = z_0 \end{cases}$

$\begin{cases} y(x_1) = y_1 \\ z(x_1) = z_1 \end{cases}$ 则：曲线长度

$J = J[y(x), z(x)] = \int_A^B ds = \int_A^B \sqrt{dx^2 + dy^2 + dz^2} =$
$\int_{x_0}^{x_1} \sqrt{1 + y'^2 + z'^2}\, dx$

要求是 J 在约束 $G(x, y(x), z(x))=0$ 之下的极值。

例 3 求一给定长度的闭曲线使它包围的面积尽量大。

解：设曲线为$\rho = \rho(\theta)$

曲线长度：$L[\rho(\theta)] = \int_0^{2\pi} \sqrt{d\rho^2 + (\rho d\theta)^2} = \int_0^{\pi} \sqrt{\rho^2 + (\dot{\rho})^2} d\theta = l$。

曲线所围面积：$S[\rho(\theta)] = \int_0^{2\pi} \frac{1}{2} \rho^2 d\theta$

故要求的是：在约束 $L[\rho(\theta)] = \int_0^{\pi} \sqrt{\rho^2 + (\dot{\rho})^2} d\theta = l$ 之下 $S[\rho(\theta)] = \int_0^{2\pi} \frac{1}{2} \rho^2 d\theta$的极大值.（注：以几何上考虑，显然是圆。）

§ 9.2 Euler 变分方程

9.2.1 变分学基本问题：

在所有定义在区间 上。固定边界（$y(x_0)=y_0$，$y(x_1)=y_1$）时，连续且有连续的一阶玉二阶导数的所有函数（记作类C^2）$y=y(x)$之中，寻找一个函数使得泛函：$\mathrm{J}[y(x)]=\int_{x_0}^{x_1}F(x,y,y')dx$ 具有最小（或最大的）可能值。

为寻找上述泛函 J 取极值的必要条件，任取一个C^2类函数

$$\eta(x_0)=\eta(x_1)=0$$

假设函数$y=y(x)$使泛函 J 取极值，构造一个函数$y(x)+\epsilon\eta(x)$:，则

$$J[\varepsilon]=J[y(x)+\epsilon\eta(x)]=\int_{x_0}^{x_1}F(x,y+\epsilon\eta,y'+\epsilon\eta')dx$$

$J[\varepsilon]$应在$\varepsilon=0$处取极值，即：

$$\frac{\mathrm{d}J[\varepsilon]}{\mathrm{d}\varepsilon}=0$$

而$\frac{\mathrm{d}J[\varepsilon]}{\mathrm{d}\varepsilon}$=$\int_{x_0}^{x_1}\big(F_y(x,y,y')\eta(x)+F_{y'}(x,y,y')\eta'(x)\big)dx$

$=\int_{x_0}^{x_1}\eta(x)\left(F_y(x,y,y')-\frac{d}{dx}F_{y'}(x,y,y')\eta'(x)\right)dx+[F_y,\eta(x)]_{x_0}^{x_1}$

$=\int_{x_0}^{x_1}\eta(x)\left(F_y(x,y,y')-\frac{d}{dx}F_{y'}(x,y,y')\eta'(x)\right)dx$

变分法基本引理：

基本引理 9. 1 设函数$\eta(x)$在区域$[x_0,x_1]$内是$C^{\prime}$类函数，且在端点处等于零，即：$\eta(x_0)=\eta(x_1)=0$. 那么，若对任意这样的函数$\eta(x)$，积分$\int_{x_0}^{x_1}f(x)\,\eta(x)dx$恒等于零（其中$f(x)$是$[x_0,x_1]$上的某个连续的函数），则$f(x)$在$[x_0,x_1]$内恒等于零，即$f(x)=0$，$x\in[x_0,x_1]$。

证明：反证法，设在$x=\xi\in[x_0,x_1]$处$f(x)$不等于零。例如$f(\xi)>0$.则$f(x)$是连续函数，一定存在ξ的某个邻域$[\xi_1,\xi_2]\subseteq[x_0,x_1]$,使$f(x)$在$[\xi_1,\xi_2]$内恒正。故构造$\eta(x)$如下：

$$\eta(x)=\begin{cases}0 & x\in[x_0,\xi_1]\\ (x-\xi_1)^2(x-\xi_2)^2 & x\in[\xi_1,\xi_2]\\ 0 & x\in[\xi_2,x_1]\end{cases}$$

$\eta(x)$满足引理的条件。此时积分

$\int_{x_0}^{x_1}f(x)\,\eta(x)dx=\int_{\xi_1}^{\xi_2}f(x)\,(x-\xi_1)^2(x-\xi_2)^2dx>0$矛盾

基本引理 9.2 设函数设函数$\eta(x,y)$在区域 x−y 平面上某区域 B 内的内是$C^{‘}$类函数，且在 B 的边界 l 上$\eta(x,y)$等于零，则，若对任意这样的函数$\eta(x,y)$，积分$\iint_B f(x,y)\eta(x,y)dxdy$恒等于零（其中$f(x,y)$为某个 B 上的连续函数），那么$f(x,y)$在区域 B 上必恒等于零，即

$$f(x,y)=0,(x,y)\in B$$

同样用反证法，假设$f(\xi,\eta)>0$,则存在(ξ,η)为圆心半径为 ρ 的圆，圆内$f(x,y)>0$. 构造：$\eta(\xi,\eta)=\begin{cases}0(\textit{圆外}) \\ [(x-\xi)^2+(y-\eta)^2-\rho^2]^2\ (\textit{圆内})\end{cases}$

剩余证明完全类似引理 9.1.

引理条件改成C^n类函数，结论不变，只需将证明过程构造的函数指数改成更高即可。对三重积分以及更一般的任何重积分，引理也很容易证明。回到基本问题上来。由 $\frac{\mathrm{d}J[\varepsilon]}{\mathrm{d}\varepsilon}=0$及引理一，立即得到：给泛函 J 以极值的曲线 必定满足如下的微分方程：

$$F_y(x,y,y')-\frac{d}{dx}F_{y'}(x,y,y')=0$$

此即 Euler 方程。

Euler 方程亦可写成：

$$-\mathrm{F}_y+\frac{\partial \mathrm{F}_{y'}}{\partial x}+\frac{\partial \mathrm{F}_{y'}}{\partial y}\frac{dy}{dx}+\frac{\partial \mathrm{F}_{y'}}{\partial y'}\frac{dy'}{dx}=0$$

即$\mathrm{F}_{y'y'}y''+\mathrm{F}_{y'y'}y'+\mathrm{F}_{xy'}-\mathrm{F}_y=0$

例 4 速降线问题：$\mathrm{J[y(x)]}=\int_{x_0}^{x_1}\sqrt{\frac{1+y\prime^2}{y}}dx$

$\mathrm{F(x,y,y')}=\sqrt{\frac{1+y\prime^2}{y}}$　　故：$F_y=\frac{\partial \mathrm{F}}{\partial y}=-\frac{\sqrt{1+y\prime^2}}{2y^{3/2}}$

$F_{y\prime}=\frac{\partial \mathrm{F}}{\partial y'}=-\frac{2y'}{2\sqrt{1+y'^2}\sqrt{y}}$

故 Euler 方程：$-\frac{\sqrt{1+y'^2}}{2y^{3/2}}-\frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}\sqrt{y}}\right)=0$

方程不显含 x,可分离变量，将 $\frac{d}{dx}=y'\frac{d}{dy}$代入有：

即:

$$\frac{\sqrt{1+y'^2}}{\sqrt{y^3}}+\ y'\frac{d}{dy}\left(\frac{y'}{\sqrt{y}\sqrt{1+y'^2}}\right)=0$$

整理后有：

$$\frac{y'dy'}{1+y'^2}+\frac{\mathrm{dy}}{\mathrm{2y}}=0$$

解之，得$x=\sqrt{2C_1y-y^2}+C_1\,cos^{-1}\frac{C_1-y}{C_1}+C_2$

引入　$\theta=cos^{-1}\frac{C_1-y}{C_1}$, 有：$\begin{cases}x=C_1\theta-C_1\sin\theta+C_2\\ y=C_1-C_1\cos\theta\end{cases}$

C_1,C_1由边界条件确定

此即旋轮线

9.2.2，Euler 表达式恒等于零的情形：

若泛函$J[y(x)]=\int_{x_0}^{x_1}F(x,y,y')dx$的欧拉表达式$F_{y'}-\frac{\partial F}{\partial y'}$恒等于零，则： $F_y-F_{xy'}-F_{yy'}y'-F_{y'y'}y''$对任意函数 *y=y(x)*恒为零

将其看成（x,y,y',y''）的一个函数，由于对任意给定的一个四元数组 ，总能找到一条曲线，使得

故$F_y-F_{xy'}-F_{yy'}y'-F_{y'y'}y''$对任意函数 *y=y(x)*恒为零即代表：其作为（x,y,y',y''）的一个函数，对任意的取值(C_1,C_2,C_3,C_4)均恒为零，而它又是y''的一次函数，故y''的函数$F_{y'y'}$必为零，即$F_{y'y'}=0$，故$F=ay'+b$(a,b 是 x,y 的函数）

将$F=ay'+b$带入 Euler 方程$F_{yy'}y'+F_{xy'}-F_y=0$ 有：

$$a_yy'+a_x-a_yy'-b_y=0$$

恒成立。

即a_x-b_y恒为零。在此条件下，泛函为：

$$J[y(x)]=\int(a(x,y)y'+b(x,y))dx=\int a(x,y)dy+b(x,y)dx$$

$$=\int dU(x,y)=U(x_1,y_1)-U(x_0,y_0)$$

即：此时泛函称为一个全微分的积分，只与端点取值有关，与路径（不同的函数选取）无关。

9.2.3 Euler 方程的形式不变性

当作参量的替换：$x=x(t),y(x)=y(x(t)),\ x\to t$

$$\mathrm{J}[y(x)]=\int_{x_0}^{x_1}F(x,y,y')dx=\int_{t_0}^{t_1}F(x(t),y(t),y'(t)\dot{x}(t))dt$$

$$=\int_{t_0}^{t_1}G(t,y(t),\dot{y}(t))dt=\mathrm{J}[y(t)]$$

故假设$y=y(x)$使泛函$\mathrm{J}[y(x)]$取极值的话，参量代换$x=x(t)$之后，$y=y(t)$也应使$\mathrm{J}[y(t)]$取极值，而

$y=y(x)$满足方程 $\frac{\partial F}{\partial y}-\frac{d}{dx}\frac{\partial F}{\partial y\prime}=0$

$y=y(t)$满足方程 $\frac{\partial G}{\partial y}-\frac{d}{dt}\frac{\partial G}{\partial y\prime}=0$

故

$$F_y-\frac{d}{dx}F_{y'}=0\Leftrightarrow G_y-\frac{d}{dt}G_{y'}=0$$

其中$G=G(t,y(t),\dot{y}(t))=F(x(t),y(t),y'(t))\dot{x}(t)$

可见，做参数的代换之后，Euler 方程的形式不变（可直接对 Euler 方程做变量替换来验证）。该性质使得某些情况下可选用合适的参数(广义坐标)来表述运动方程

9.2.4　形式标记——变分和变分导数

引入$y(x)\rightarrow y(x)+\epsilon\eta(x)$之后,

$$\mathrm{J}[y(x)]\rightarrow \mathrm{J}[\varepsilon]=\int F(x,y+\epsilon\eta,y'+\epsilon\eta')dx$$

故$\mathrm{J}[\varepsilon]$的一阶微分是（在$\epsilon=0$处）:

$$dJ(0)=\int[\frac{\partial F}{\partial y}\eta(x)d\epsilon+\frac{\partial F}{\partial y'}\eta'(x)d\epsilon]dx$$

$$=\int[F_y(x,y,y')-\frac{d}{dx}F_{y'}(x,y,y')]\,d\varepsilon\eta(x)dx+[d\varepsilon\eta(x)F_{y'}]_{x_0}^{x_1}$$

固定边界$y(x_0)=y_0,y(x_1)=y_1$，$\eta(x_0)=\eta(x_1)=0$,故:

$$dJ(0)=\int[F_y(x,y,y')-\frac{d}{dx}F_{y'}(x,y,y')]\,d\varepsilon\eta(x)dx$$

若引入变量标记$\delta y(x)=d\varepsilon\eta(x)$，$\delta \mathrm{J}[y(x)]=dJ(0)$

则：$\delta \mathrm{J}[y(x)]=\int[F_y-\frac{d}{dx}F_{y'}]\,\delta y(x)dx$ 称作泛函 J 的一阶变分，

定义$\frac{\delta \mathrm{J}}{\delta y(x)}=F_y-\frac{d}{dx}F_{y'}$，称作泛函 J 的一阶变分导数（泛函导数）

故: $\delta \mathrm{J}[y(x)]=\int\frac{\delta \mathrm{J}}{\delta y(x)}\delta y(x)dx$（变分导数的定义式）

不太严谨地（物理学家的通常做法），一阶变分$\delta \mathrm{J}$可这样求：

将$\delta y(x)$作为小量，并计算$\mathrm{J}[y(x)+\delta y(x)]-\mathrm{J}[y(x)]$，只保留到$\delta y(x)$的一阶。

例如： $\mathrm{J[y]}=\int_{x_0}^{x_1}F(x,y,y')\,dx$　　固定边界。

$$\mathrm{J[y}+\delta y]-\mathrm{J[y]}=\int_{x_0}^{x_1}[\,F(x,y+\delta y,y'+\delta y')-F(x,y,y')\,]dx$$

$$=\int_{x_0}^{x_1}(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y')\,dx+\cdots$$

$$=\int_{x_0}^{x_1}(F_y-\frac{d}{dx}F_{y'})\delta y\,dx+\left[\frac{\partial F}{\partial y'}\delta y'\right]_{x_0}^{x_1}+\cdots$$

故　$\delta \mathrm{J}[y(x)]=\int_{x_0}^{x_1}(F_y-\frac{d}{dx}F_{y'})\delta y\,dx$

故 $\frac{\delta \mathrm{J}}{\delta y(x)}=F_y-\frac{d}{dx}F_{y'}$

泛函取极值的必要条件是：对任意的$\delta y(x)$，一阶变分$\delta \mathrm{J}=0$。或者说$\mathrm{J[y]}$的一阶变分导数等于零：$\frac{\delta \mathrm{J}}{\delta y(x)}=0$

可推广上述讨论来定义高阶变分和高阶变分导数，不再赘述。

泛函和多元函数的类比：$\begin{cases}f=f(y_1,y_2,\cdots,y_n)\\ J=J[y(x)]\end{cases}$（无穷多个自变量的多元函数）

$$\begin{cases} df = \dfrac{\partial f}{\partial y_1}dy_1 + \dfrac{\partial f}{\partial y_2}dy_2 + \cdots + \dfrac{\partial f}{\partial y_n}dy_n \\ \delta J = \displaystyle\int \frac{\delta J}{\delta y(x)}\delta y(x)dx \end{cases}$$

计算的一些形式规则(自己课后验证)

（1）$y_1, y_2, \cdots, y_n$为独立变量，故$\frac{\partial y_i}{\partial y_i} = \delta_{ij}$，类比即有:

$$\frac{\delta y(x_1)}{\delta y(x_2)} = \delta(x_1 - x_2) \Leftarrow \delta y(x_1) = \int \delta(x_1 - x_2)\,\delta y(x_2)dx_2$$

（2）复合泛函：$\frac{\delta}{\delta y(x)}F[G(y(x))] = \int \frac{\delta F}{\delta G(y(x_1))}\frac{\delta G(y(x_1))}{\delta y(x)}dx_i$

（3）注意乘积的变分导数:

$$\frac{\delta}{\delta y(x)}\{F[y]G[y]\} = \frac{\delta F[y]}{\delta y(x)}G[y] + F[y]\frac{\delta G[y]}{\delta y(x)}$$

（3）多元泛函的一阶变分

$$\delta F[y_1(x), y_2(x)] = \int (\frac{\delta F[y]}{\delta y_1(x)}\delta y_1(x) + \frac{\delta F[y]}{\delta y_2(x)}\delta y_2(x))\mathrm{dx}$$

9.2.5、含有高阶导数的情景：

例 5　$J[y(x)] = \int_{x_0}^{x_1} F(x, y(x), y'(x), y''(x)dx$

在固定边界条件$\begin{cases} y(x_0) = c_0 \\ y'(x_0) = d_0 \end{cases}$ $\begin{cases} y(x_1) = c_1 \\ y'(x_1) = d_1 \end{cases}$之下的极值问题

仍然引入 y→ $y(x) + \varepsilon\eta(x)$ ，并假设y(x)使J取极值

则$J(\varepsilon) = J[y(x) + \varepsilon\eta(x)]$在$\varepsilon = 0$处取极值

而

$$\frac{dJ(0)}{d\varepsilon} = \int (\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta' + \frac{\partial F}{\partial y''}\eta'')dx$$

$\int \frac{\partial F}{\partial y'}\eta'\, dx = \int (-\frac{d}{dx}F_{y'})\,\eta(x)dx$　　$\int \frac{\partial F}{\partial y''}\eta'' dx = \cdots\cdots$

故$\frac{dJ(0)}{d\varepsilon} = \int (F_y - \frac{d}{dx}F_{y'} + \frac{d^2}{dx^2}F_{y''})\,\eta(x)dx = 0$

由基本引理有$F_y-\frac{d}{dx}F_{y\prime}+\frac{d^2}{dx^2}F_{y\prime\prime}=0$

含有更高阶导数的情形不难以此类推。

9.2.6 含有多个自变函数的情形

例 6 $J[y_1,y_2]=\int_{x_0}^{x_1}F(x,y_1,y_2,y_1',y_2')\,dx$

在固定边界条件$\begin{cases}y_1(x_0)=c_0\\y_1(x_1)=c_1\end{cases}$ $\begin{cases}y_2(x_0)=d_0\\y_2(x_1)=d_1\end{cases}$ 下的极值问题

此时假设 J 在$y_1=y_1(x),y_2=y_2(x)$处取极值，则引入：

$y_1\to y_1(x)+\varepsilon_1\eta(x)$ $\qquad y_2\to y_2(x)+\varepsilon_2\eta(x)$

$$\mathrm{J}(\varepsilon_1,\varepsilon_2)=\mathrm{J}[y_1+\varepsilon_1\eta,y_2+\varepsilon_2\eta]$$

故函数$\mathrm{J}(\varepsilon_1,\varepsilon_2)$应在$\varepsilon_1=\varepsilon_2=0$处取极值。

由$\frac{\partial \mathrm{J}(0,0)}{\partial\varepsilon_1}=0$及$\frac{\partial \mathrm{J}(0,0)}{\partial\varepsilon_2}=0$ 分别得到： $\begin{cases}\frac{\partial F}{\partial y_1}-\frac{d}{dx}F_{y_1\prime}=0\\\frac{\partial F}{\partial y_2}-\frac{d}{dx}F_{y_2\prime}=0\end{cases}$

更多自变函数以此类推。

9.2.7 含有多个自变量的情形：

例 7 设 D 是x_1-x_2平面内由曲线 T 围成的区域，

$$\mathrm{J}[\mathrm{u}(x_1,x_2)]=\iint_D F(x_1,x_2,u,u_1',u_2')$$

$(u_i'=\frac{\partial}{\partial x_i}u(x_1,x_2))$ 在固定边界条件：Γ上$\mathrm{u}(x_1,x_2)$取给定的值得情况下，问$\mathrm{u}(x_1,x_2)$取何函数形式，*J[u]*有极值？

仍假设$\mathrm{u}=\mathrm{u}(x_1,x_2)$使 J 取极值，构造$\mathrm{u}\to\mathrm{u}(x_1,x_2)+\varepsilon\eta((x_1,x_2))$（$(x_1,x_2)$在$\Gamma$上为零）

则$\mathrm{J}(\varepsilon)=\mathrm{J}[\mathrm{u}(x_1,x_2)+\varepsilon\eta((x_1,x_2)]$在$\varepsilon=0$处取极值。

而

$$\frac{\mathrm{dJ}(0)}{\mathrm{d\varepsilon}}=\iint(\frac{\partial F}{\partial u}\eta+\frac{\partial F}{\partial u_1'}\frac{\partial\eta}{\partial x_1}+\frac{\partial F}{\partial u_2'}\frac{\partial\eta}{\partial x_2})dx_1dx_2$$

$$=\iint(\frac{\partial F}{\partial u}\eta-\frac{\partial}{\partial x_1}\frac{\partial F}{\partial u_1'}-\frac{\partial}{\partial x_2}\frac{\partial F}{\partial u_2'})dx_1dx_2$$

$$+\iint[\frac{\partial}{\partial x_1}\left(\eta\frac{\partial F}{\partial u_1'}\right)+\frac{\partial}{\partial x_2}(\eta\frac{\partial F}{\partial u_2'})]dx_1dx_2$$

由 Green 公式

$\frac{\mathrm{dJ}(0)}{\mathrm{d\varepsilon}}$的第二个积分项可化为 $\oint_\Gamma\eta\frac{\partial F}{\partial u_1'}dx_2-\eta\frac{\partial F}{\partial u_2'}dx_1$

故由$\frac{\mathrm{dJ}(0)}{\mathrm{d\varepsilon}}=0$及基本引理 2 有：极值处 u 满足方程：

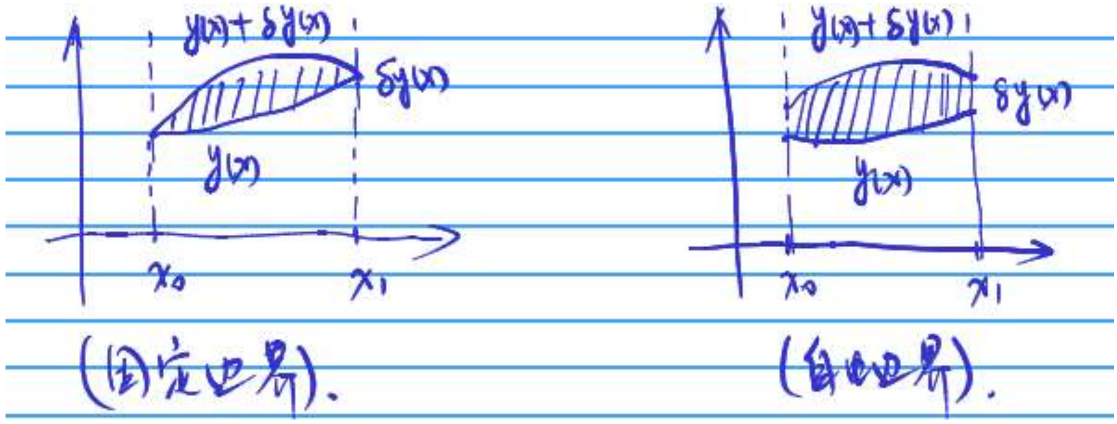
$$\frac{\partial F}{\partial u}-\frac{\partial}{\partial x_1}\frac{\partial F}{\partial u_{x_1}}-\frac{\partial}{\partial x_2}\frac{\partial F}{\partial u_{x_2}}=0$$

含有更多自变量的情形以此类推。

§ 9.3 非固定边界条件问题

上一节均假定在边界处的自变函数（有时还需自变函数的导数）取给定的值之下，讨论泛函的极值问题。本节看看其他边界条件。

9.3.1 自由边界条件



考虑泛函$\mathrm{J}[\mathrm{y}(\mathrm{x})]=\int_{x_0}^{x_1}F(x,y(x),y'(x))dx$，问$y(x)$在区间$\left[x_0,x_1\right]$上取何函数形式时，J[y(x)]取极值？

仍假设时$\mathrm{y}=y(x)$取极值。构造$\mathrm{y}\rightarrow y(x)+\varepsilon\eta(x)$与固定边界不同，η在$x_0,x_1$处不必等于零。此时$\mathrm{J}(\varepsilon)=\mathrm{J}[\mathrm{y}(\mathrm{x})+\varepsilon\eta(\mathrm{x})]$应在$\varepsilon=0$处取极值。而

$$\frac{dJ(0)}{d\varepsilon}=\int(\frac{\partial F}{\partial y}-\frac{d}{dx}\frac{\partial F}{\partial y'})\eta(x)dx+[\eta(x)\frac{\partial F}{\partial y'}]_{x_0}^{x_1}$$

此时边界项不能随意扔掉。上式对任意的$\eta(x)$必须恒为零。

可先考虑$\eta(x_0)=\eta(x_1)=1$的任意η。故由$\frac{dJ(0)}{d\varepsilon}$可得$F_y-\frac{d}{dx}F_{y'}=0$，代入上式有：$\eta(x_1)\frac{\partial F}{\partial y'}|_{x=x_1}-\eta(x_0)\frac{\partial F}{\partial y'}|_{x=x_0}=0$ ，对任意η成立。故η的端点值可随意选取，要使上式成立，只能是$F_{y'}|_{x=x_1}=$

$F_{y'}|_{x=x_0}=0$

综上，这种情况下变分取极值的必要条件是：$\begin{cases} F_y-\frac{d}{dx}F_{y'}=0 \\ F_{y'}|_{x=x_1}=F_{y'}|_{x=x_0}=0 \end{cases}$

9.3.2 横交条件（约束端点问题）：

为简单计，考虑其中一个端点固定。

设泛函$\mathrm{J}[\mathrm{y}(\mathrm{x})]=\int_{x_0}^{x_1}F(x,y(x),y'(x))dx$

其中自变函数$y=y(x)$的曲线一端固定在$\left(x_0,y_0\right)$处，另一端$\left(x_1,y_1\right)$在另一条曲线$\phi(\mathrm{x},\mathrm{y})=0$上滑动，问y取何函数形式时，泛函J[y]取极值？

仍假设$\mathrm{y}=\mathrm{y}(\mathrm{x})$时，J取极值，构造$\mathrm{y}\to y(x)+\delta y(x)$，相应地积分的上限$x_1$亦有改变：$x_1\to x_1+\delta x_1$，以使得 均在 上。

考虑一阶变分$\delta\mathrm{J}$: $(\mathrm{y}(x_0)=y_0)$

$$\delta\mathrm{J}\approx\int_{x_0}^{x_1+\delta x_1}F(x,y+\delta y,y'+\delta y')dx-\int_{x_0}^{x_1}F(x,y,y')dx$$

$$\approx\int_{x_0}^{x_1}(F(x,y+\delta y,y'+\delta y')-F(x,y,y'))dx+\ F(x_1,y_1,y'_1)$$

$$=\int_{x_0}^{x_1}(\frac{\partial F}{\partial y}-\frac{d}{dx}\frac{\partial F}{\partial y'})\delta y dx_1+\frac{\partial F}{\partial y'}\ \delta y|_{x=x_1}+F|_{x=x_1}\delta x_1=0$$

上式必须对任意$\delta y(x)$均成立，同样先考虑端点x_1处$\delta\mathrm{y}=0$的变分。

可知$\frac{\partial F}{\partial y}-\frac{d}{dx}\frac{\partial F}{\partial y'}=0$ ，代入上式有：$\frac{\partial F}{\partial y'}\ \delta y|_{x=x_1}+F|_{x=x_1}\delta x_1=0$

但是注意：$\delta y|_{x=x_1}\neq\delta y_1$

故：$\delta y|_{x=x_1}\approx\delta y_1-(y'(x)+\delta y'(x))|_{x=x_1}\cdot\delta x_1\approx\delta y_1-y'(x_1)\delta x_1$

而B,B'两点均在$\phi(x,y)=0$之上，故 $\frac{\delta\mathrm{y}_1}{\delta\mathrm{x}_1}=-\frac{\phi_x}{\phi_y}|_{x=x_1}$

故$(\mathrm{F}-y'F_{y'})\ \phi_y-F_{y'}\ \phi_x|_{x=x_1}=0$

另一种推导:引入曲线的参数形式以避免积分限的变动，见希尔伯特《数学物理方法》卷一第四章 Page163

§9.4 条件极值问题

9.4.1 函数的条件极值问题——Lagrange 乘子方法

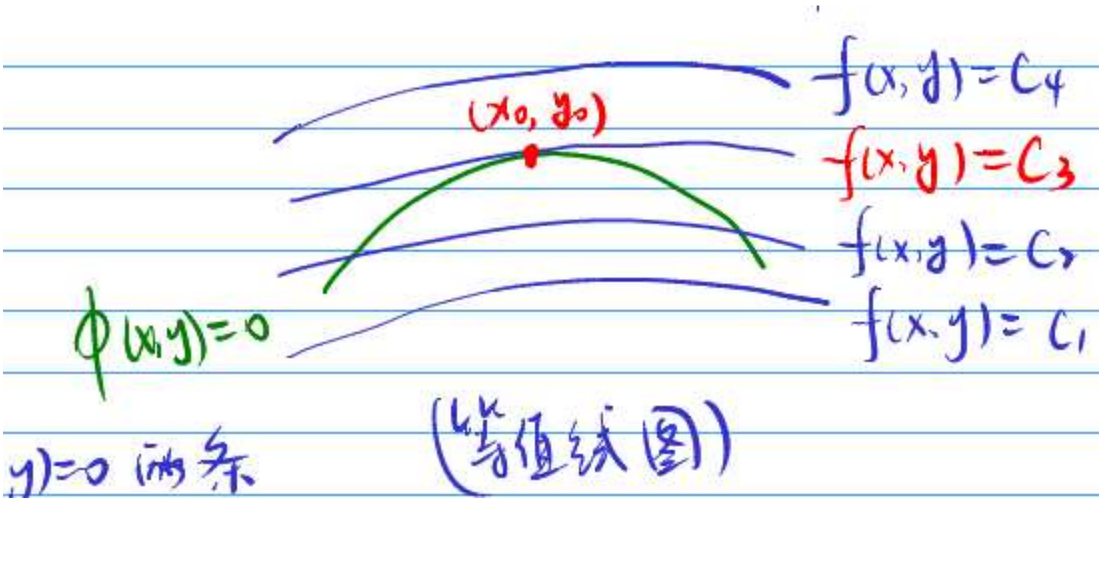
问题：设函数$f=f(x,y,z)\in C,\varphi(x,y,z)=0$是光滑曲面，则$f$在约束 φ 下的极值问题等价于 $F(x,y,z,\lambda)=f(x,y,z)+\lambda\,\varphi(x,y,z)$的无约束极值。

问题的本质在于求 f 的微分时， 不再是完全独立地变动，二十在满足约束的情况下变动，要求 df=0，故需用约束条件消除不独立的微元，将 df 用完全独立的微元表示出来，则此时由 df=0 即可得到微元前的系数必须为零。

证明方法一：（隐函数方法）由$\varphi(x,y,z)=0\Rightarrow z=z(x,y)$，故$f(x,y,z)=f(x,y,z(x,y))$而 x,y 是独立变量，故$\frac{\partial f}{\partial x}=\frac{\partial f}{\partial y}=0$

证明方法二：（几何图形）减少一维考虑：$f=f(x,y)$在约束$\varphi(x,y)=0$之下的极值？

从下侧图形可以看出：



$f(x,y)$在约束极值点(x_0,y_0)满足：$f(x,y)=C_3,\phi(x,y)=0$ 两条曲线相切。故 $f_x:f_y=\phi_x:\phi_y$．推广到$f(x,y,z)$在约束$\phi(x,y,z)=0$ 下的极值。则必须两曲面相切，故$f_x:f_y:f_z=\phi_x:\phi_y:\phi_z$，整理后即 Lagrange 乘子方法

法三：普遍方法。考虑$f(x_1,x_2,\cdots,x_n)$在$\begin{cases}\phi_1(x_1,x_2,\cdots,x_n)=0\\ \phi_2(x_1,x_2,\cdots,x_n)=0\\ \cdots\cdots\cdots\\ \phi_n(x_1,x_2,\cdots,x_n)=0\end{cases}$下的极值问题（m<n）

首先写出$df=\frac{\partial f}{\partial x_1}dx_1+\frac{\partial f}{\partial x_2}dx_2+\cdots+\frac{\partial f}{\partial x_n}dx_n$，由于有 m 个约束，故 n 个$dx_i$中必有 m 个是不独立的。而只有 n-m 个是独立的。不妨假设有 m 个约束可解出$\begin{cases}x_{n-m+1}=x_{n-m+1}(x_1,x_2,\cdots,x_{n-m})\\ x_{n-m+2}=x_{n-m+2}(x_1,x_2,\cdots,x_{n-m})\\ \cdots\cdots\cdots\\ x_{n-m+m}=x_n(x_1,x_2,\cdots,x_{n-m})\end{cases}$

故从df中消除$dx_{n-m+1},\cdots$，dx_n问题就解决了。采用对称做法，由于$d\phi_i=0\Rightarrow\frac{\partial\phi_i}{\partial x_1}dx_1+\frac{\partial\phi_i}{\partial x_2}dx_2+\cdots+\frac{\partial\phi_i}{\partial x_n}dx_n=0$

给出了 n 个dx_i的线性约束，假设约束极值在 A 点取得，选取 m 个常数λ_i，使得在 A 点有：

$$\frac{\partial f}{\partial x_j}+\lambda_1\frac{\partial\phi_1}{\partial x_j}+\lambda_2\frac{\partial\phi_2}{\partial x_j}+\cdots+\lambda_m\frac{\partial\phi_m}{\partial x_j}|_{x=A}=0(j=n-m+1,\cdots,n)$$

在相应的 Jacobi 行列式非零的时候，这是可以做到的。因为这是关于 m 个λ_i的 m 个线性方程。选取如上的λ_i之后，由于 A 点是满足约束条件的极值点，故在满足约束的条件下做变动有：

$$\mathrm{d}f|_A=0,\mathrm{d}\phi_1|_A=0,\mathrm{d}\phi_2|_A=0,\cdots,\mathrm{d}\phi_m|_A=0$$

故 $\quad \mathrm{d}f+\lambda_1\mathrm{d}\phi_1+\lambda_2\mathrm{d}\phi_2+\cdots+\lambda_m\,\mathrm{d}\phi_m|_A=0$

由于λ_i的选取，使得上式中$dx_j(j=n-m+1,\cdots,n)$的系数正好为零。故

$$\sum_{i=1}^{n-m}(\frac{\partial f}{\partial x_i}+\lambda_1\frac{\partial \phi_1}{\partial x_i}+\lambda_2\frac{\partial \phi_2}{\partial x_i}+\cdots+\lambda_m\frac{\partial \phi_m}{\partial x_i}|_A)dx_i=0$$

而$dx_i(i=1,2,\cdots,n-m)$是完全独立的。故：

$$\frac{\partial f}{\partial x_i}+\lambda_1\frac{\partial \phi_1}{\partial x_i}+\lambda_2\frac{\partial \phi_2}{\partial x_i}+\cdots+\lambda_m\frac{\partial \phi_m}{\partial x_i}|_A=0$$

综上，有：

$$\begin{cases}\frac{\partial f}{\partial x_i}+\lambda_1\frac{\partial \phi_1}{\partial x_i}+\lambda_2\frac{\partial \phi_2}{\partial x_i}+\cdots+\lambda_m\frac{\partial \phi_m}{\partial x_i}|_A=0\\ \phi_1|_A=0,\phi_2|_A=0,\cdots,\phi_m|_A=0\end{cases}$$

上述结论正好可看成：

$$\mathrm{F}(x_1,\cdots,x_n,\lambda_1,\cdots,\lambda_m)=\mathrm{f}(x_1,\cdots,x_n)+\sum_{i=1}^{m}\lambda_i\phi_i(x_1,\cdots,x_m)$$

关于$x_1,\cdots,x_n,\lambda_1,\cdots,\lambda_m$这 n+m 个变量的无约束机制问题，此即 Lagrange 乘子法.

9.4.2 曲面上的测地线问题:泛函的 Lagrange 乘函法

考虑泛函： $J[y_1(x),y_2(x)]=\int_{x_0}^{x_1}F(x,y_1,y_2,{y_1}',{y_2}')dx$ 在约束$\mathrm{G}(x,y_1,y_2)=0$之下的固定边界极值问题。

$\mathrm{G}(x,y_1,y_2)=0$中 G 不含${y_1}'$等项称为整约束;含有导数的约束，例如$H(x,y_1,y_2,{y_1}',)=0$，称为非整约束！

方法一：隐函数方法，由$\mathrm{G}(x,y_1,y_2)=0$可解出$y_2=y_2(x,y_1(x))$，

将之带入 F 中有

$$\mathrm{F}=F(x,y_1(x),y_2(x,y_1(x)),{y_1}'(x),{y_2}'(x,y_1(x)))$$

注意$y_2{}' = \frac{\partial y_2}{\partial x} + \frac{d}{dx}\frac{\partial y_2}{\partial y_1}$+，最后求出关于$y_1$的欧拉方程，用 F 和 G 表述即可。

方法二：首先

$$\delta J = \int_{x_0}^{x_1}[(F_{y_1} - \frac{d}{dx}F_{y_1{}'})\delta y_1 + (F_{y_2} - \frac{d}{dx}F_{y_2{}'})\delta y_2]\,dx \quad (*)$$

问题的困难在于δy_1和δy_2不都是独立的，因为y_1和y_2的变动要满足$\mathrm{G}(x, y_1, y_2) = 0$.故$G_{y_1}\delta y_1 + G_{y_2}\delta y_2 = 0$。在构造一个函数$\lambda(x)$。有$y_1, y_2$满足约束时，$\int_{x_0}^{x_1} G(x, y_1, y_2)\lambda(x)dx = 0$

即在满足约束的变动下，

$$\int_{x_0}^{x_1}[G_{y_1}\lambda(x)\delta y_1 + G_{y_2}\lambda(x)\delta y_2]\,dx = 0 \qquad (\&)$$

选取$F_{y_2} - \frac{d}{dx}F_{y_2{}'} + G_{y_2}\lambda(x) = 0$

则将式(*)与(&)相加，并注意到δy_1是独立的变分，故：

$$F_{y_1} - \frac{d}{dx}F_{y_1{}'} + G_{y_1}\lambda(x) = 0$$

故整约束下，极值函数的方程为$\begin{cases} F_{y_1} - \frac{d}{dx}F_{y_1{}'} + G_{y_1}\lambda(x) = 0 \\ F_{y_2} - \frac{d}{dx}F_{y_2{}'} + G_{y_2}\lambda(x) = 0 \\ \mathrm{G}(x, y_1, y_2) = 0 \end{cases}$

上述方程组正好是泛函:

$$J'[y_1(x), y_2(x), \lambda(x)] = J[y_1(x), y_2(x)] + \int_{x_0}^{x_1} G(x, y_1, y_2)\lambda(x)dx$$

无约束的极值问题的欧拉方程。 $\lambda(x)$称为 Lagrange 乘函。上述结论即 Lagrange 乘函法.

对于有多个约束，以及约束是非整约束的情形，可仿照上述分析得到相应的结论。可参见 Smirnov《高等数学教程》第四卷，或希尔伯特《数学物理方法》第一卷

9.4.3、等周问题：泛函的 Lagrange 乘子法

考虑泛函：$J[y(x)] = \int_{x_0}^{x_1} F(x, y, y')dx$，在约束条件$\mathrm{L}[y(x)] = \int_{x_0}^{x_1} G(x, y(x), y(x)')dx = 0$之下的固定边界极值问题。这种情况下，隐函数方法失效，约束并不能隐式地解除。若直接写$\delta J[y] = \int_{x_0}^{x_1}[(F_y - \frac{d}{dx}F_{y'})\delta y]\,dx = 0$，则由于$\delta y$不是完全任意的变动，而是必须满足$\mathrm{L}[y(x)] = 0$之下的变动，故不能直接利用基本引理。实际上，$\delta y = \epsilon\eta$的话，$\epsilon$完全无法变动，因为约束给出$\mathrm{L}(\epsilon) = \mathrm{L}[y(x) + \epsilon\eta(x)] = 0$。因此，代替引入一个$\epsilon$，我们引入两个：

假设$J[y(x)]$在$y = y(x)$处取条件极值，则构造$y \to y(x) + \epsilon_1\eta_1(x) + \epsilon_2\eta_2(x)$（$\eta_1, \eta_2$在边界为 0）考虑满足约束的变动，则 满足约束：$\mathrm{L}(\epsilon_1, \epsilon_2) = L[y(x) + \epsilon_1\eta_1(x) + \epsilon_2\eta_2(x)] = 0$

故$\mathrm{J}(\epsilon_1, \epsilon_2) = J[y(x) + \epsilon_1\eta_1(x) + \epsilon_2\eta_2(x)]$应该在上述约束之下在$\epsilon_1 = \epsilon_2 = 0$处取极值。故

$$\begin{cases} \dfrac{\partial}{\partial\epsilon_1}(J(0,0) + \lambda L(0,0)) = 0 \\ \dfrac{\partial}{\partial\epsilon_2}(J(0,0) + \lambda L(0,0)) = 0 \end{cases}$$

其中λ是常数. 由$\eta_1(x)$和$\eta_2(x)$的任意性，上述两个方程均导致同一个结论：（$F_y + \lambda G_y) - \frac{d}{dx}(F_{y'} + \lambda G_{y'}) = 0$，连同约束：$\mathrm{L}[y(x)] = \int_{x_0}^{x_1} G(x, y(x), y(x)')dx = 0$一起，构成了最终的结论。可看成:

$$J'[y(x)](\lambda) = J[y(x)] + \lambda[y(x)] = \int_{x_0}^{x_1}[F(x, y, y') + \lambda G(x, y, y')]dx$$

关于$y(x)$和λ的无约束极值问题，注意到$J'[y(x)](\lambda)$是关于$y(x)$的泛函，亦是关于λ的函数，故J'无约束极值的必要条件是：

$$\frac{\delta J'}{\delta y(x)}=0,\frac{dJ'}{d\lambda}=0$$

此即泛函的 Lagrange 乘子法。

一个与前面泛函形式略微不同的泛函的条件极值问题：考虑$J[y(x)]=\iint y(x_1)k(x_1,x_2)\,y(x_2)dx_1dx_2$($k$为给定的函数)，在约束条件$\mathrm{G}(y(x))=\int(y(x))^2dx=1$下的极值问题。

虽然不能直接套用前面的欧拉公式，但基本思想是完全相同的，即引入$y(x)+\epsilon\eta(x)$（无约束下）或$y\to y(x)+\epsilon_1\eta_1(x)+\epsilon_2\eta_2(x)$（有约束下），将之转换为函数的（约束）极值问题。自己课后完成推导，可参见彭桓武《数学物理基础》9.5.2 节。实际上就是一个“二次型”！

结论：极值条件是$\begin{cases}\int k(x_1,x_2)y(x_2)dx_2=\lambda y(x_1)\\ \int(y(x))^2dx=1\end{cases}$

需推出的是：该条件极值问题与泛函

$$J'[y]=\frac{\iint y(x_1)k(x_1,x_2)\,y(x_2)dx_1dx_2}{\int(y(x))^2dx}$$

的无约束极值问题等价。

§9.5 物理学的变分原理

变分法在物理学中主要用于：

1、 近似求解（例如 Ritz 方法）

2、 归纳定律

差不多一切自然定律都能用变分原理的形式（代替微分方程的形式）来表述。对于不同领域的现象，选择一个含话的 Lagrange 函数，则由变分$\delta\int Ldt=0$得到的 Euler 方程就给出了该现象的规律。

例 8 牛顿力学：引入

$$\mathrm{S}[x_i(\mathrm{t})]=\int_A^B Ldt=\int_A^B[T(t,x_i,\dot{x}_i)-V(x_i)]dt$$

由$\delta S=0$即可给出运动方程。（S:作用量，L：拉格朗日量）。

用变分原理归纳定律有诸多好处：

（1）相比运动方程(组)，变分原理表述更紧凑简洁

（2）偏微分方程组的相容性是个问题，用方程组表述的规律可能会无解。但变分原理得到的微分方程组都是相容的。

（3）便于推广到新的物理领域。对新的相互作用，只需在 Lagrange 中加上描述新规律的项，由变分原理就可得到耦合的运动方程；

（4）作用量 S 是标量，便于对称性的处理；

（5）导致对称性和守恒律内在联系的发现；

（6）量子力学量子场论中……

（7） ……