线性代数期末复习

1 逆矩阵

1.1 定义

Definition 1.1

对于 n 阶矩阵 $\boldsymbol{A}$，如果有一个 n 阶矩阵 $\boldsymbol{B}$，使:

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{E} \tag{1}$$

其中 $\boldsymbol{E}$ 是单位矩阵。则称矩阵 $\boldsymbol{A}$ 是可逆的，并把矩阵 $\boldsymbol{B}$ 称为 $\boldsymbol{A}$ 的逆矩阵。

1.2 性质

1.2.1 性质 1

如果矩阵 $\boldsymbol{A}$ 是可逆的 (invertible)，那么 $\boldsymbol{A}$ 的逆矩阵是唯一的，而且矩阵 $\boldsymbol{A}$ 是非奇异矩阵，即 $|\boldsymbol{A}| \neq 0$。这是 $\boldsymbol{A}$ 是奇异矩阵的充分必要条件。奇异矩阵的行或列向量是线性相关的，因此奇异矩阵不是满秩 (full rank) 的（即 $\text{rank}(\boldsymbol{A}) < n$）而且奇异矩阵的行列式为 0。秩 (rank) 是矩阵中线性无关向量的数量。

Theorem 1.1

矩阵 $\boldsymbol{A}$ 是可逆的 $\Leftrightarrow$ 矩阵 $\boldsymbol{A}$ 是非奇异矩阵 $\Leftrightarrow$ 矩阵 $\boldsymbol{A}$ 是满秩的 $\Leftrightarrow$ 矩阵 $\boldsymbol{A}$ 的行列式不为 0

1.2.2 性质 2

若 $|\boldsymbol{A}| \neq 0$，则矩阵 $\boldsymbol{A}$ 可逆，且

$$\boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^* \tag{2}$$

其中 $\boldsymbol{A}^*$ 是 $\boldsymbol{A}$ 的伴随矩阵。在 n 阶行列式中，把 (i,j) 元 a_{ij} 所在的第 i 行和第 j 列划去后，留下来的 $n-1$ 阶行列式叫作 (i,j) 元 a_{ij} 的代数余子式，记作 M_{ij}。记:

$$A_{ij} = (-1)^{i+j}M_{ij} \tag{3}$$

其中 A_{ij} 叫作 a_{ij} 的代数余子式。

Definition 1.2

行列式 $|\boldsymbol{A}|$ 的各个元素的代数余子式 A_{ij} 所构成的矩阵的转置称作矩阵 $\boldsymbol{A}$ 的伴随矩阵。

1.2.3 性质 3

若 $\boldsymbol{A}, \boldsymbol{B}$ 为同阶矩阵且均可逆，则 $\boldsymbol{AB}$ 亦可逆，且：

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \tag{4}$$

1.3 求逆矩阵

对于一个矩阵：

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ -1 & -3 \end{bmatrix} \tag{5}$$

1.3.1 待定系数法求逆矩阵

假设所求的逆矩阵为：

$$\boldsymbol{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{6}$$

那么：

$$\boldsymbol{AA}^{-1} = \begin{bmatrix} 1 & 2 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+2c & b+2d \\ -a-3c & -b-3d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{7}$$

解得 $a = 3,\ b = 2,\ c = -1,\ d = -1$。所以：

$$\boldsymbol{A}^{-1} = \begin{bmatrix} 3 & 2 \\ -1 & -1 \end{bmatrix} \tag{8}$$

1.3.2 伴随矩阵求逆矩阵

矩阵 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|$ 计算得：

$$|\boldsymbol{A}| = 1 \times (-3) - (2 \times (-1)) = -1 \tag{9}$$

矩阵 $\boldsymbol{A}$ 的伴随矩阵可以计算为：

$$\boldsymbol{A}^{*} = \begin{bmatrix} A_{11} & A_{21} \\ A_{21} & A_{22} \end{bmatrix}^{T} = \begin{bmatrix} -3 & -(-1) \\ -(2) & 1 \end{bmatrix}^{T} = \begin{bmatrix} -3 & -2 \\ 1 & 1 \end{bmatrix} \tag{10}$$

所以:

$$A^{-1} = \frac{1}{|A|}A^* = \frac{1}{-1}\begin{bmatrix} -3 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -1 & -1 \end{bmatrix} \tag{11}$$

1.3.3 初等变换求逆矩阵(高斯-约当法)

写出增广矩阵 $\boldsymbol{A}|\boldsymbol{E}$，即矩阵 $\boldsymbol{A}$ 右侧放置一个同阶的单位矩阵，得到一个新矩阵。随后进行初等变换将这个矩阵的左侧化为一个单位矩阵，矩阵的右侧即为逆矩阵。

$$\begin{gathered}
\begin{bmatrix} 1 & 2 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \xrightarrow{R_2 \to R_2 + R_1} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{bmatrix} \\
\begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{bmatrix} \xrightarrow{R_1 \to R_1 + 2R_2} \begin{bmatrix} 1 & 0 & 3 & 2 \\ 0 & -1 & 1 & 1 \end{bmatrix} \\
\begin{bmatrix} 1 & 0 & 3 & 2 \\ 0 & -1 & 1 & 1 \end{bmatrix} \xrightarrow{R_2 \to -R_2} \begin{bmatrix} 1 & 0 & 3 & 2 \\ 0 & 1 & -1 & -1 \end{bmatrix}
\end{gathered} \tag{12}$$

2 LU 分解 (LU decomposition)

Definition 2.1

LU 分解是一种将一个矩阵分解为两个特殊矩阵的算法，通常表示为:

$$\boldsymbol{A} = \boldsymbol{LU} \tag{13}$$

其中: $\boldsymbol{A}$ 是一个 $n \times n$ 的矩阵。$\boldsymbol{L}$ 是一个下三角矩阵 *(lower triangular matrix)*。$\boldsymbol{U}$ 是一个上三角矩阵 *(upper triangular matrix)*。如果要求其中的 $\boldsymbol{L}$ 或者 $\boldsymbol{U}$ 为单位三角矩阵，那么分解是唯一的。

2.1 列方程进行 LU 分解

对矩阵 $\boldsymbol{A}$ 进行 LU 分解，设 $\boldsymbol{L}$ 为单位三角矩阵。

$$\boldsymbol{A} = \begin{bmatrix} 4 & 3 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix} \Rightarrow \begin{cases} 1 \times u_{11} + 0 \times 0 = 4 \\ l_{21} \times u_{11} + 1 \times 0 = 6 \\ 1 \times u_{12} + 0 \times u_{22} = 3 \\ l_{21} \times u_{12} + 1 \times u_{22} = 3 \end{cases} \Rightarrow \begin{cases} u_{11} = 4 \\ l_{21} \times u_{11} = 6 \\ u_{12} = 3 \\ l_{21} \times u_{12} + u_{22} = 3 \end{cases} \Rightarrow \begin{cases} u_{11} = 4 \\ l_{21} = 1.5 \\ u_{12} = 3 \\ u_{22} = -1.5 \end{cases} \tag{14}$$

即：

$$\boldsymbol{L}=\begin{bmatrix}1 & 0\\ -1.5 & 1\end{bmatrix},\ \boldsymbol{U}=\begin{bmatrix}4 & 3\\ 0 & -1.5\end{bmatrix} \tag{15}$$

2.2 用 Gauss 变换进行 LU 分解

2.2.1 前置知识

矩阵右乘列向量的计算方式为：

$$\begin{bmatrix}a & b & c\\ d & e & f\\ g & h & i\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=x\begin{bmatrix}a\\ d\\ g\end{bmatrix}+y\begin{bmatrix}b\\ e\\ h\end{bmatrix}+z\begin{bmatrix}c\\ f\\ i\end{bmatrix}=\begin{bmatrix}*\\ *\\ *\end{bmatrix} \tag{16}$$

矩阵左乘行向量的计算方式为：

$$\begin{bmatrix}x & y & z\end{bmatrix}\begin{bmatrix}a & b & c\\ d & e & f\\ g & h & i\end{bmatrix}=x\begin{bmatrix}a & b & c\end{bmatrix}+y\begin{bmatrix}d & e & f\end{bmatrix}+z\begin{bmatrix}g & h & i\end{bmatrix}c \tag{17}$$

矩阵运算的核心是对行或者列进行独立操作。对于两个矩阵：

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{a}_2\\ \vdots\\ \boldsymbol{a}_i\\ \vdots\\ \boldsymbol{a}_m\end{bmatrix},\ \boldsymbol{B}=\begin{bmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\\ \vdots\\ \boldsymbol{b}_i\\ \vdots\\ \boldsymbol{b}_n\end{bmatrix} \tag{18}$$

其中 $\boldsymbol{a}_1\ldots\boldsymbol{a}_m$ 和 $\boldsymbol{b}_1\ldots\boldsymbol{b}_n$ 是行向量，矩阵的乘法可以表示为：

$$\boldsymbol{AB}=\begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{a}_2\\ \vdots\\ \boldsymbol{a}_i\\ \vdots\\ \boldsymbol{a}_m\end{bmatrix}\boldsymbol{B}=\begin{bmatrix}\boldsymbol{a}_1\boldsymbol{B}\\ \boldsymbol{a}_2\boldsymbol{B}\\ \vdots\\ \boldsymbol{a}_i\boldsymbol{B}\\ \vdots\\ \boldsymbol{a}_m\boldsymbol{B}\end{bmatrix} \tag{19}$$

因为：

$$\boldsymbol{a}_i=\begin{bmatrix}a_{i1} & a_{i2} & \cdots & a_{in}\end{bmatrix} \tag{20}$$

所以:

$$\boldsymbol{a}_i\boldsymbol{B}=\boldsymbol{a}_i\begin{bmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\\ \vdots\\ \boldsymbol{b}_n\end{bmatrix}=\begin{bmatrix}a_{i1} & a_{i2} & \cdots & a_{in}\end{bmatrix}\begin{bmatrix}\boldsymbol{b}_1\\ \boldsymbol{b}_2\\ \vdots\\ \boldsymbol{b}_n\end{bmatrix}=a_{i1}\boldsymbol{b}_1+a_{i2}\boldsymbol{b}_2+\cdots+a_{in}\boldsymbol{b}_n \tag{21}$$

所以:

$$\boldsymbol{AB}=\begin{bmatrix}\boldsymbol{a}_1\boldsymbol{B}\\ \boldsymbol{a}_2\boldsymbol{B}\\ \vdots\\ \boldsymbol{a}_i\boldsymbol{B}\\ \vdots\\ \boldsymbol{a}_m\boldsymbol{B}\end{bmatrix}=\begin{bmatrix}a_{11}\boldsymbol{b}_1+a_{12}\boldsymbol{b}_2+\cdots+a_{1n}\boldsymbol{b}_n\\ a_{21}\boldsymbol{b}_1+a_{22}\boldsymbol{b}_2+\cdots+a_{2n}\boldsymbol{b}_n\\ \vdots\\ a_{i1}\boldsymbol{b}_1+a_{i2}\boldsymbol{b}_2+\cdots+a_{in}\boldsymbol{b}_n\\ \vdots\\ a_{m1}\boldsymbol{b}_1+a_{m2}\boldsymbol{b}_2+\cdots+a_{mn}\boldsymbol{b}_n\end{bmatrix} \tag{22}$$

2.2.2 细节过程

对于要分解的矩阵:

$$\boldsymbol{A}=\begin{bmatrix}1 & 4 & 7\\ 2 & 5 & 8\\ 3 & 6 & 10\end{bmatrix} \tag{23}$$

首先设第一个消元矩阵为:

$$\boldsymbol{L}_1=\begin{bmatrix}1 & 0 & 0\\ l_{21} & 1 & 0\\ l_{31} & 0 & 1\end{bmatrix} \tag{24}$$

根据定义，我们要让 $\boldsymbol{L}_1\boldsymbol{A}=\boldsymbol{U}_1$ 的 $(2,1)$ 元和 $(3,1)$ 元为 0。这给出:

$$l_{21}\times 1+1\times 2+0\times 3=0,\ l_{31}\times 1+0\times 2+1\times 3=0 \tag{25}$$

所以 $l_{21}=-2,\ l_{31}=-3$。将这个结果回带到 $\boldsymbol{L}_1$ 并继续计算 $\boldsymbol{U}_1$:

$$\boldsymbol{L}_1\boldsymbol{A}=\begin{bmatrix}1 & 0 & 0\\ -2 & 1 & 0\\ -3 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 4 & 7\\ 2 & 5 & 8\\ 3 & 6 & 10\end{bmatrix}=\begin{bmatrix}1 & 4 & 7\\ 0 & -3 & -6\\ 0 & -6 & -11\end{bmatrix} \tag{26}$$

接着设第二个消元矩阵为:

$$\boldsymbol{L}_2=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & l_{32} & 1\end{bmatrix} \tag{27}$$

根据定义，我们要让 $\boldsymbol{L}_2\boldsymbol{U}_1 = \boldsymbol{U}$ 的 $(3,2)$ 元为 0。这给出：

$$l_{32} \times -3 + 1 \times -6 = 0 \tag{28}$$

所以 $l_{32} = -2$。将这个结果回带到 $\boldsymbol{L}_2$ 并继续计算 $\boldsymbol{U}$：

$$\boldsymbol{L}_2\boldsymbol{U}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{bmatrix} = \boldsymbol{U} \tag{29}$$

所以：

$$\boldsymbol{L}_2(\boldsymbol{L}_1\boldsymbol{A}) = \boldsymbol{U} \Rightarrow (\boldsymbol{L}_2\boldsymbol{L}_1)\boldsymbol{A} = \boldsymbol{U} \Rightarrow \boldsymbol{A} = (\boldsymbol{L}_2\boldsymbol{L}_1)^{-1}\boldsymbol{U} = \boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\boldsymbol{U} \Rightarrow \boldsymbol{L} = \boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1} \tag{30}$$

所以：

$$\boldsymbol{L} = \boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} \tag{31}$$

3 矩阵的秩

3.1 定义

Definition 3.1

在 $m \times n$ 的矩阵 $\boldsymbol{A}$ 中，任取 k 行与 k 列 $(k \leq m, k \leq n)$，位于这些行列交叉处的 k^2 个元素，不改变它们在 $\boldsymbol{A}$ 中所处的位置次序而得到的 k 阶行列式，称为矩阵 $\boldsymbol{A}$ 的 k 阶<u>子式</u>。

Definition 3.2

设在矩阵 $\boldsymbol{A}$ 中有一个不等于 0 的 r 阶子式 D，且所有 $r+1$ 阶子式（如果存在的话）全等于 0，那么 D 称为矩阵 $\boldsymbol{A}$ 的<u>最高阶非零子式</u>，数 r 称为矩阵 $\boldsymbol{A}$ 的<u>秩</u>，记作 $R(\boldsymbol{A})$。$R(\boldsymbol{A})$ 也即为 $\boldsymbol{A}$ 的非零子式的最高阶数。

一个阶数为 n 的可逆矩阵 $\boldsymbol{B}$ 的最高阶非零子式一定是 $|\boldsymbol{B}|$，所以可逆矩阵的秩等于矩阵的阶数，不可逆矩阵的秩小于矩阵的阶数。因此，可逆矩阵又称<u>满秩矩阵</u>。

矩阵的秩等于它的列数的矩阵称为<u>列满秩矩阵</u>，类似地，若矩阵的秩等于它的行数，则称它为<u>行满秩矩阵</u>。当矩阵为方阵时，列满秩矩阵就称为满秩矩阵。

矩阵的秩的性质有：

1. $0 \leq R(\boldsymbol{A}_{m\times n}) \leq \min\{m, n\}$
2. $R(\boldsymbol{A}^T) = R(\boldsymbol{A})$

3. 若 $\boldsymbol{A} \sim \boldsymbol{B}$，则 $R(\boldsymbol{A}) \sim R(\boldsymbol{B})$

4. 若 $\boldsymbol{P}, \boldsymbol{Q}$ 可逆，则 $R(\boldsymbol{PAQ}) = R(\boldsymbol{A})$

5. $\max\{R(\boldsymbol{A}), R(\boldsymbol{B})\} \leq R(\boldsymbol{A}, \boldsymbol{B}) \leq R(\boldsymbol{A}) + R(\boldsymbol{B})$

3.2 求矩阵的秩

求矩阵 $\boldsymbol{A}$ 的秩的方法有：

1. 如果矩阵 $\boldsymbol{A}$ 是可逆的（即 $|\boldsymbol{A}| \neq 0$，则矩阵是满秩的，矩阵的秩就等于矩阵的阶数。

2. 矩阵的秩等于矩阵非零子式的最高阶数，这也是矩阵化成阶梯型矩阵或最简形矩阵的非零行个数。

3. 最大线性无关向量组（简称最大无关组）所含向量的个数即为向量组 $\boldsymbol{A}$ 的秩。

4 线性方程组的解

4.1 线性方程组的解的数量与存在性

设有 n 个未知数 m 个方程的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{32}$$

该式可以写成以向量 $\boldsymbol{x}$ 为未知元的形式

$$\boldsymbol{Ax} = \boldsymbol{b} \tag{33}$$

利用系数矩阵 $\boldsymbol{A}$ 和增广矩阵 $\boldsymbol{B} = (\boldsymbol{A}, \boldsymbol{b})$ 的秩，可以方便地讨论线性方程组是否有解以及有解时解是否唯一等问题。其结论是：

对于 n 元线性方程组，即 $\boldsymbol{Ax} = \boldsymbol{b}$：

1. 无解的充分必要条件是 $R(\boldsymbol{A}) < R(\boldsymbol{A}, \boldsymbol{b})$

2. 有唯一解的充分必要条件是 $R(\boldsymbol{A}) = R(\boldsymbol{A}, \boldsymbol{b}) = n$

3. 有无限多解的充分必要条件是 $R(\boldsymbol{A}) = R(\boldsymbol{A}, \boldsymbol{b}) < n$

4.2 线性方程组的解的结构

4.2.1 齐次线性方程组

齐次线性方程组的解集 $S=\{\boldsymbol{x}\,|\,\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\}$ 是一个向量空间，只要找到向量空间的一个基，就能精确描述该空间。这个解空间又称为零空间 (null space)，该空间的一个基即称为该齐次线性方程组的基础解系，要求齐次线性方程组的通解，只需求出它的基础解系。

零空间的维数称为零空间维数 (nullity)，通常用 nullity($\boldsymbol{A}$) 表示。根据矩阵的秩-零定理 (rank-nullity theorem)：

$$\text{nullity}(\boldsymbol{A})=n-R(\boldsymbol{A}) \tag{34}$$

即零空间维数等于该齐次线性方程组的自由未知数的个数。

Example： 求齐次线性方程组

$$\begin{cases} x_1+x_2-x_3-x_4=0\\ 2x_1-5x_2+3x_3+2x_4=0\\ 7x_1-7x_2+3x_3+x_4=0 \end{cases} \tag{35}$$

的基础解系与通解。

Solution： 对系数矩阵 $\boldsymbol{A}$ 作初等行变换变为行最简形矩阵，有：

$$\boldsymbol{A}=\begin{bmatrix}1&1&-1&-1\\2&5&3&2\\7&-7&3&1\end{bmatrix}\stackrel{*}{\to}\begin{bmatrix}1&0&-\frac{2}{7}&-\frac{3}{7}\\0&1&-\frac{5}{7}&-\frac{4}{7}\\0&0&0&0\end{bmatrix} \tag{36}$$

这给出：

$$\begin{cases} x_1=\frac{2}{7}x_3+\frac{3}{7}x_4\\ x_2=\frac{5}{7}x_3+\frac{4}{7}x_4\\ x_3=x_3+0x_4\\ x_4=0x_3+x_4 \end{cases} \tag{37}$$

令 $\begin{bmatrix}x_3\\x_4\end{bmatrix}$ 取 $\begin{bmatrix}1\\0\end{bmatrix}$ 及 $\begin{bmatrix}0\\1\end{bmatrix}$，则对应的 $\begin{bmatrix}x_1\\x_2\end{bmatrix}$ 为 $\begin{bmatrix}\frac{2}{7}\\\frac{5}{7}\end{bmatrix}$ 及 $\begin{bmatrix}\frac{3}{7}\\\frac{4}{7}\end{bmatrix}$，即得基础解系：

$$\boldsymbol{\xi}_1=\begin{bmatrix}\frac{2}{7}\\\frac{5}{7}\\1\\0\end{bmatrix},\boldsymbol{\xi}_2=\begin{bmatrix}\frac{3}{7}\\\frac{4}{7}\\0\\1\end{bmatrix} \tag{38}$$

并由此写出通解:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = c_1 \begin{bmatrix} \frac{2}{7} \\ \frac{5}{7} \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} \frac{3}{7} \\ \frac{4}{7} \\ 0 \\ 1 \end{bmatrix} \ (c_1, c_2 \in \mathbb{R}) \tag{39}$$

4.2.2 非齐次线性方程组

如果求得方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的一个解 $\boldsymbol{\eta}^*$【称为特解 (particular solution)】, 那么该方程的通解为:

$$\boldsymbol{x} = k_1\boldsymbol{\xi}_1 + k_2\boldsymbol{\xi}_2 + \cdots + k_{n-r}\boldsymbol{\xi}_{n-r} + \boldsymbol{\eta}^* \ (k_1, k_2, \cdots, k_{n-r} \in \mathbb{R}) \tag{40}$$

其中 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_n$ 是方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的基础解系。

Example: 求解方程组

$$\begin{cases} x_1 - x_2 - x_3 + x_4 = 0 \\ x_1 - x_2 + x_3 - 3x_4 = 1 \\ x_1 - x_2 - 2x_3 + 3x_4 = -\frac{1}{2} \end{cases} \tag{41}$$

Solution: 对增广矩阵 $\boldsymbol{B}$ 施行初等行变换:

$$\boldsymbol{B} = \begin{bmatrix} 1 & -1 & -1 & 1 & 0 \\ 1 & -1 & 1 & -3 & 1 \\ 1 & -1 & -2 & 3 & -\frac{1}{2} \end{bmatrix} \xrightarrow{*} \begin{bmatrix} 1 & -1 & 0 & -1 & \frac{1}{2} \\ 0 & 0 & 1 & -2 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{42}$$

可见 $R(\boldsymbol{A}) = R(\boldsymbol{B}) = 2$, 故方程有解, 并有:

$$\begin{cases} x_1 = x_2 + x_4 + \frac{1}{2} \\ x_3 = 2x_4 + \frac{1}{2} \end{cases} \tag{43}$$

取 $x_2 = x_4 = 0$, 则 $x_1 = x_3 = \frac{1}{2}$, 即得方程组的特解:

$$\boldsymbol{\eta}^* = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{bmatrix} \tag{44}$$

在对应的齐次线性方程组 $\begin{cases} x_1 = x_2 + x_4 \\ x_3 = 2x_4 \end{cases}$ 中, 取 $\begin{bmatrix} x_2 \\ x_4 \end{bmatrix}$ 取 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 及 $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, 则对应的 $\begin{bmatrix} x_1 \\ x_3 \end{bmatrix}$ 为 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 及 $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, 即得对应的齐次线性方程组的基础解系:

$$\boldsymbol{\xi}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{\xi}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix} \tag{45}$$

于是所求通解为:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{bmatrix} \ (c_1, c_2 \in \mathbb{R}) \tag{46}$$

5 向量组的线性相关性

5.1 向量组及其线性组合

5.1.1 线性组合与线性表示

Definition 5.1

n 个有次序的数 $a_1, a_2, \cdot, a_n$ 所组成的数组称为 n 维向量，这 n 个数称为该向量的 n 个分量，第 i 个数 a_i 称为第 i 个分量。

Definition 5.2

3 维向量的全体所组成的集合

$$\mathbb{R}^3 = \left\{ \boldsymbol{r} | \boldsymbol{r} = (x, y, z)^T,\ x, y, z \in R \right\} \tag{47}$$

叫做 *3* 维向量空间。在点空间取定坐标系之后，空间中的点 $P(x, y, z)$ 与三维向量 $\boldsymbol{r} = (x, y, z)^T$ 之间有一一对应的关系。

类似地，n 维向量的全体所组成的集合记为:

$$\mathbb{R}^n = \left\{ \boldsymbol{x} | \boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T,\ x_1, x_2, \cdots, x_n \in R \right\} \tag{48}$$

Definition 5.3

若干个同维数的列向量（或行向量）所组成的集合叫做向量组。

例如一个 $m \times n$ 矩阵的全体列向量是一个含 n 个 m 维列向量的向量组，它的全体行向量是一个含 m 个 n 维向量的向量组。

Definition 5.4

给定向量组 $A : \boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$，对于任何一组实数 $k_1, k_2, \cdots, k_m$，表达式

$$k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + \cdots + k_m\boldsymbol{a}_m \tag{49}$$

称为线性组 A 的一个线性组合，$k_1, k_2, \cdots, k_m$ 称为这个线性组合的系数。

Definition 5.5

给定向量组 $A: \boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 和向量 $\boldsymbol{b}$，若存在一组数 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 使

$$\boldsymbol{b} = \lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \cdots + \lambda_m \boldsymbol{a}_m \tag{50}$$

则向量 $\boldsymbol{b}$ 是向量组 A 的线性组合，这时称向量 $\boldsymbol{b}$ 能由向量组 A 线性表示。

这也就是方程组

$$x_1 \boldsymbol{a}_1 + x_2 \boldsymbol{a}_2 + \cdots + x_m \boldsymbol{a}_m = \boldsymbol{b} \tag{51}$$

有解。根据现行方程组的解的条件，这要求矩阵 $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m \end{bmatrix}$ 的秩等于矩阵 $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m & \boldsymbol{b} \end{bmatrix}$ 的秩。这即是向量 $\boldsymbol{b}$ 能由向量组 A 线性表示的充分必要条件。

Example： 设

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ -1 \\ 4 \\ 0 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ 3 \\ 1 \end{bmatrix} \tag{52}$$

试证明 $\boldsymbol{b}$ 能由向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 线性表示，并求出表达式。

Solution： 根据刚才的结论，我们要证矩阵 $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix}$ 与 $\boldsymbol{B} = (\boldsymbol{A}, \boldsymbol{b})$ 的秩相等。为此，把 $\boldsymbol{B}$ 化成行最简形矩阵：

$$\boldsymbol{B} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 0 \\ 2 & 1 & 4 & 3 \\ 2 & 3 & 0 & 1 \end{bmatrix} \overset{*}{\to} \begin{bmatrix} 1 & 0 & 3 & 2 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{53}$$

可见 $R(\boldsymbol{A}) = R(\boldsymbol{B})$，因此，向量 $\boldsymbol{b}$ 能由向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 线性表示。由上述最简形矩阵，可得方程 $\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \boldsymbol{b}$ 的解为：

$$\begin{cases} x_1 + 3x_3 = 2 \\ x_2 - 2x_3 = -1 \end{cases} \tag{54}$$

设 $x_3 = c$ 可以得到通解：

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = c \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -3c+2 \\ 2c-1 \\ c \end{bmatrix} \tag{55}$$

例题之前讨论了向量 $\boldsymbol{b}$ 能由向量组 A 线性表示的充分必要条件。现在讨论向量组 B 能由向量组 A 线性表示的充分必要条件。

Definition 5.6

设有两个向量组 $A:\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 及 $B:\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_l$，若向量组 B 中的每个向量都能由向量组 A 线性表示，则称向量组 B 能由向量组 A 线性表示。

"向量组 $B:\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_l$ 能由向量组 $A:\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 线性表示" 等价于 "有矩阵 $\boldsymbol{K}$，使 $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{K}$", 进一步等价于 "方程 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$ 有解"。这三种叙述都可对应到充分必要条件: $R(\boldsymbol{A})=R(\boldsymbol{A},\boldsymbol{B})$。

同时这也要求 $R(\boldsymbol{B})\leq R(\boldsymbol{A})$, 因为 $\boldsymbol{A}$ 的列向量组成的向量空间的维度必须大于等于 $\boldsymbol{B}$ 的向量空间的维度以张成该空间。或者也可以通过相等不等关系进行推导: $R(\boldsymbol{B})<R(\boldsymbol{A},\boldsymbol{B})=R(\boldsymbol{A})$。注意 $R(\boldsymbol{B})\leq R(\boldsymbol{A})$ 是必要条件而非充分条件。

Theorem 5.1

向量组 $B:\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_l$ 能由向量组 $A:\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 线性表示的充分必要条件是 $R(\boldsymbol{A})=R(\boldsymbol{A},\boldsymbol{B})$，必要条件是 $R(\boldsymbol{B})\leq R(\boldsymbol{A})$。

Example: 已知向量组

$$A:\boldsymbol{a}_1=\begin{bmatrix}0\\1\\2\\3\end{bmatrix},\boldsymbol{a}_2=\begin{bmatrix}3\\0\\1\\2\end{bmatrix},\boldsymbol{a}_3=\begin{bmatrix}2\\3\\0\\1\end{bmatrix},\ B:\boldsymbol{b}_1=\begin{bmatrix}2\\1\\1\\2\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}0\\-2\\1\\1\end{bmatrix},\boldsymbol{b}_3=\begin{bmatrix}4\\4\\1\\3\end{bmatrix}\tag{56}$$

证明向量组 B 能由向量组 A 线性表示，但向量组 A 不能由向量组 B 线性表示。

Solution: 向量组 B 能由向量组 A 线性表示要求方程 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$ 有解，即 $R(\boldsymbol{A})=R(\boldsymbol{A},\boldsymbol{B})$。

首先求 $R(\boldsymbol{A})$:

$$\boldsymbol{A}=\begin{bmatrix}0&3&2\\1&0&3\\2&1&0\\3&2&1\end{bmatrix}\overset{*}{\to}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\\0&0&0\end{bmatrix}\tag{57}$$

这给出 $R(\boldsymbol{A})=3$。再求 $R(\boldsymbol{B})$:

$$\boldsymbol{B}=\begin{bmatrix}2&0&4\\1&-1&4\\1&1&1\\2&1&3\end{bmatrix}\overset{*}{\to}\begin{bmatrix}1&0&2\\0&1&-1\\0&0&0\\0&0&0\end{bmatrix}\tag{58}$$

这给出 $R(\boldsymbol{B})=2$。因为 $R(\boldsymbol{B})<R(\boldsymbol{A})$，所以向量组 A 一定不能由向量组 B 线性表示。为了判断向

量组 B 能不能由向量组 A 线性表示，还需要计算 $R(\boldsymbol{A}, \boldsymbol{B})$。

$$\boldsymbol{A}, \boldsymbol{B} = \begin{bmatrix} 0 & 3 & 2 & 2 & 0 & 4 \\ 1 & 0 & 3 & 1 & -1 & 4 \\ 2 & 1 & 0 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 & 1 & 3 \end{bmatrix} \overset{*}{\to} \begin{bmatrix} 1 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} & -\frac{3}{4} & \frac{5}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{59}$$

这给出 $R(\boldsymbol{A}, \boldsymbol{B}) = R(\boldsymbol{A}) = 3$，所以向量组 B 能由向量组 A 线性表示。

5.1.2 向量组等价

Definition 5.7

若向量组 A 与向量组 B 能相互线性表示，则称这两个向量组等价。

Theorem 5.2

向量组 $A: \boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 与向量组 $B: \boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_l$ 等价的充分必要条件是

$$R(\boldsymbol{A}) = R(\boldsymbol{B}) = R(\boldsymbol{A}, \boldsymbol{B}) \tag{60}$$

其中 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是向量组 A 和 B 所构成的矩阵。

5.2 向量组的线性相关性

Definition 5.8

给定向量组 $A: \boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$，若存在不全为零的数 $k_1, k_2, \cdots, k_m$，使

$$k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + \cdots + k_m\boldsymbol{a}_m = 0 \tag{61}$$

则称向量组 A 是线性相关的，否则称它线性无关。

对于只含一个向量 $\boldsymbol{a}$ 的向量组，当 $\boldsymbol{a} = \boldsymbol{0}$ 时是线性相关的，当 $\boldsymbol{a} \neq \boldsymbol{0}$ 时是线性无关的；对于含两个向量 $\boldsymbol{a}_1, \boldsymbol{a}_2$ 的向量组，它线性相关的充分必要条件是 $\boldsymbol{a}_1, \boldsymbol{a}_2$ 对应成比例，其几何意义是两向量共线；三个向量线性相关的几何意义是三向量共面。

Theorem 5.3

向量组 $A: \mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_m}$ $(m \geq 2)$ 线性相关的充分必要条件是向量组 A 中至少有一个向量能由其余 $m-1$ 个向量线性表示。

向量组 $A: \boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 构成矩阵 $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m \end{bmatrix}$，向量组 A 线性相关，就是齐次线性方程组

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_m\boldsymbol{a}_m = 0 \tag{62}$$

即 $\boldsymbol{Ax} = \boldsymbol{0}$ 有非零解。所以

Theorem 5.4

向量组 $A: \boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 线性相关的充分必要条件是它所构成的矩阵 $\mathbf{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m \end{bmatrix}$ 的秩小于向量个数 m；向量组 A 线性无关的充分必要条件是 $R(\boldsymbol{A}) = m$。

秩等于向量（未知数）个数时一定使其中一个解为 0，故线性无关。

Example： 已知

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} 2 \\ 4 \\ 7 \end{bmatrix} \tag{63}$$

试讨论向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 的线性相关性。

Solution： 对矩阵 $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix}$ 施行初等行变换变成行阶梯矩阵，即可同时看出矩阵 $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix}$ 及 $\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{bmatrix}$ 的秩，利用定理 5.4 即可得到结论。

$$\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 2 & 4 \\ 1 & 5 & 7 \end{bmatrix} \overset{*}{\to} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \tag{64}$$

可见 $R(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3) = 2$，故向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 线性相关；同时可见 $R(\boldsymbol{a}_1, \boldsymbol{a}_2) = 2$，故向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2$ 线性无关。

5.3 向量空间 (vector space)

Definition 5.9

设 V 为 n 维向量的集合，如果集合 V 非空，且集合 V 对于向量的加法及数乘两种运算封闭，那么就称集合 V 为向量空间。

所谓封闭，是指：若 $\boldsymbol{a} \in V, \boldsymbol{b} \in V$，则 $\boldsymbol{a} + \boldsymbol{b} \in V$；若 $\boldsymbol{a} \in V, \lambda \in \mathbb{R}$，则 $\lambda \boldsymbol{a} \in V$。

3 维向量的全体 $\mathbb{R}^3$ 是一个向量空间。因为任意两个 3 维向量之和仍然是 3 维向量，数 λ 乘 3 维向量也仍然是 3 维向量，它们都属于 $\mathbb{R}^3$。类似地，n 维向量的全体 $\mathbb{R}^n$ 也是一个向量空间。

Example： 设 $\boldsymbol{a}, \boldsymbol{b}$ 为两个已知的 n 维向量，集合

$$V = \{\lambda \boldsymbol{a} + \mu \boldsymbol{b} \mid \lambda, \mu \in \mathbb{R}\} \tag{65}$$

是一个向量空间。因为若 $\boldsymbol{x}_1 = \lambda_1 \boldsymbol{a} + \mu_1 \boldsymbol{b}$, $\boldsymbol{x}_2 = \lambda_2 \boldsymbol{a} + \mu_2 \boldsymbol{b}$, $k \in \mathbb{R}$，则有

$$\begin{aligned} \boldsymbol{x}_1 + \boldsymbol{x}_2 &= (\lambda_1 + \lambda_2)\boldsymbol{a} + (\mu_1 + \mu_2)\boldsymbol{b} \in V \\ k\boldsymbol{x}_1 &= (k\lambda_1)\boldsymbol{a} + (k\mu_1)\boldsymbol{b} \in V \end{aligned} \tag{66}$$

这个向量空间称为由向量 $\boldsymbol{a}, \boldsymbol{b}$ 所生成的向量空间，也可记为 $\mathrm{Span}\{\boldsymbol{a}, \boldsymbol{b}\}$。一般地，由向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 所生成的向量空间为

$$L = \{\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \cdots + \lambda_m \boldsymbol{a}_m \,|\lambda_1, \lambda_2, \cdots, \lambda_m \in \mathbb{R}\} \tag{67}$$

Definition 5.10

设有向量空间 V_1 及 V_2，若 $V_1 \subseteq V_2$，就称 V_1 是 V_2 的子空间。

例如任何由 n 维向量所组成的向量空间 V，总有 $V \subseteq \mathbb{R}^n$，这样的向量空间总是 $\mathbb{R}^n$ 的子空间。

Definition 5.11

设 V 为向量空间，如果 r 个向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r \in V$，且满足

1. $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, a_r$ 线性无关；
2. V 中任一向量都可由 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r$ 线性表示。

那么向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r$ 称为向量空间 V 的一个基，r 称为向量空间 V 的维数，记为 $\dim V = r$，并称 V 为 r 维向量空间。如果向量空间 V 没有基，那么 V 的维度为 0。0 维向量空间只含一个零向量 $\mathbf{0}$。

Theorem 5.5

设由向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 所生成的向量空间

$$L = \{\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \cdots + \lambda_m \boldsymbol{a}_m \,|\lambda_1, \lambda_2, \cdots, \lambda_m \in \mathbb{R}\} \tag{68}$$

显然向量空间 L 与向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 等价，所以向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 的最大无关组就是 L 的一个基，向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 的秩就是 L 的维数。若这个最大无关组是 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r$，则 V 可表示为

$$L = \{\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \cdots + \lambda_r \boldsymbol{a}_r \,|\lambda_1, \lambda_2, \cdots, \lambda_r \in \mathbb{R}\} \tag{69}$$

即 V 是基所生成的向量空间 $\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r\}$。

Important

若把向量空间 V 看作向量组，则由最大无关组的等价定义可知，V 的基就是向量组的最大无关组，V 的维数就是向量组的秩。

5.4 基

Theorem 5.6

若一组向量可生成 n 维空间且该组向量线性无关，那么这组向量的个数一定为 n。这即是说 n 维空间的基的维数一定是 n。因为当向量个数（也就是维数）小于 n 时，向量无法生成 n

维空间；当向量个数大于 n 时，向量组必定线性相关，因为向量组的维数（向量个数）$\geq$ 向量的维数 $r \geq$ 向量组的秩。

据此有以下推论：

1. 我们可以说：在 n 维空间中，任意 n 个线性无关的向量，必是该 n 维空间的一组基；在 n 维空间中，若 n 个向量可生成空间，那这 n 个向量必是该 n 维空间的一组基。这两句话中隐含有两个条件，即线性无关和生成空间。

2. 对于一个 n 维方阵 $\boldsymbol{A}$，若该方阵的列向量线性无关，则该方阵的列向量可以生成 n 维空间。这些列向量即是 n 维空间的一组基。由于"方阵 $\boldsymbol{A}$ 的列向量线性无关"这个命题等价于"方阵 $\boldsymbol{A}$ 可逆"，所以：

   Theorem 5.7

   可逆矩阵的列向量是一组基。/ 行列式不为 0 的矩阵的列向量是一组基。

3. 2 维空间的一组基一定由 2 个线性无关（不共线）的向量组成。显然这样的基向量组有无数个。这个结论扩展到 n 维空间也成立：

   Theorem 5.8

   n 维空间的基向量组有无数个。

6 向量的内积与正交性

Definition 6.1

设有 n 维向量

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{70}$$

令

$$(\boldsymbol{x}, \boldsymbol{y}) = x_1y_1 + x_2y_2 + \cdots + x_ny_n \tag{71}$$

$(\boldsymbol{x}, \boldsymbol{y})$ 就称为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积。

内积是两个向量之间的一种运算，其结果是一个实数。当 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 都是列向量时，有

$$(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T\boldsymbol{y} \tag{72}$$

内积具有下列性质（其中 $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ 为 n 维向量，λ 为实数）：

1. $(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{y},\boldsymbol{x})$

2. $(\lambda\boldsymbol{x},\boldsymbol{y})=\lambda(\boldsymbol{x},\boldsymbol{y})$

3. $(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{z})+(\boldsymbol{y},\boldsymbol{z})$

4. 当 $\boldsymbol{x}=0$ 时，$(\boldsymbol{x},\boldsymbol{x})=0$；当 $\boldsymbol{x}\neq 0$ 时，$(\boldsymbol{x},\boldsymbol{x})>0$

Definition 6.2

令

$$\|\boldsymbol{x}\|=\sqrt{(\boldsymbol{x},\boldsymbol{x})}=\sqrt{{x_1}^2+{x_2}^2+\cdots+{x_n}^2} \tag{73}$$

$\|\boldsymbol{x}\|$ 就称为 n 维向量 $\boldsymbol{x}$ 的长度（或范数）。

当 $\|\boldsymbol{x}\|=1$ 时，称 $\boldsymbol{x}$ 为单位向量，若 $\boldsymbol{a}\neq 0$，取 $\boldsymbol{x}=\frac{1}{\|\boldsymbol{a}\|}\boldsymbol{a}$，则 $\boldsymbol{x}$ 是一个单位向量，由向量 $\boldsymbol{a}$ 得到 $\boldsymbol{x}$ 的过程称为把向量 $\boldsymbol{a}$ 单位化。

当 $\boldsymbol{x}\neq 0,\boldsymbol{y}\neq 0$ 时，

$$\theta=\arccos\frac{(\boldsymbol{x},\boldsymbol{y})}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} \tag{74}$$

称为 n 维向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的夹角。当 $(\boldsymbol{x},\boldsymbol{y})=0$ 时，称向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交。显然零向量与任何向量都正交。

Theorem 6.1

若 n 维向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 是一组两两正交的非零向量，则 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 线性无关。

Definition 6.3

设 n 维向量 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_r$ 是向量空间 $V(V\subseteq\mathbb{R}^n)$ 的一个基，若 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_r$ 两两正交，且都是单位向量，则称 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_r$ 是 V 的一个标准正交基。

设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 是空间 V 的一个基，要求 V 的一个标准正交基。这也就是要找一组两两正交的单位向量 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_r$，使 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_r$ 与 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 等价。这个过程称为施密特 (Schmidt) 正交化。

取

$$\begin{gathered}\boldsymbol{b}_1=\boldsymbol{a}_1,\\ \boldsymbol{b}_2=\boldsymbol{a}_2-\frac{(\boldsymbol{a}_2,\boldsymbol{b}_1)}{(\boldsymbol{b}_1,\boldsymbol{b}_1)}\boldsymbol{b}_1,\\ \cdots\\ \boldsymbol{b}_r=\boldsymbol{a}_r-\frac{(\boldsymbol{a}_r,\boldsymbol{b}_1)}{(\boldsymbol{b}_1,\boldsymbol{b}_1)}\boldsymbol{b}_1-\frac{(\boldsymbol{a}_r,\boldsymbol{b}_2)}{(\boldsymbol{b}_2,\boldsymbol{b}_2)}\boldsymbol{b}_2-\cdots-\frac{(\boldsymbol{a}_r,\boldsymbol{b}_{r-1})}{(\boldsymbol{b}_{r-1},\boldsymbol{b}_{r-1})}\boldsymbol{b}_{r-1}.\end{gathered} \tag{75}$$

容易验证 $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_r$ 两两正交，且 $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_r$ 与 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 等价。

然后把它们单位化，即取

$$\boldsymbol{\xi}_1=\frac{1}{\|\boldsymbol{b}_1\|}\boldsymbol{b}_1,\boldsymbol{\xi}_2=\frac{1}{\|\boldsymbol{b}_2\|}\boldsymbol{b}_2,\cdots,\boldsymbol{\xi}_r=\frac{1}{\|\boldsymbol{b}_r\|}\boldsymbol{b}_r \tag{76}$$

就是 V 的一个标准正交基。

$$\mathrm{d}\boldsymbol{J}=tr\left(\left(\nabla_{\boldsymbol{Z}^{[2]}}\boldsymbol{J}\right)^T\boldsymbol{W}^{[2]}\mathrm{d}\boldsymbol{A}^{[1]}\right)=tr\left(\boldsymbol{W}^{[2]}\left(\nabla_{\boldsymbol{Z}^{[2]}}\boldsymbol{J}\right)^T\mathrm{d}\boldsymbol{A}^{[1]}\right)=tr\left(\left(\nabla_{\boldsymbol{Z}^{[2]}}\boldsymbol{J}\boldsymbol{W}^{[2]T}\right)^T\mathrm{d}\boldsymbol{A}^{[1]}\right) \tag{77}$$